# Aerospace Technologies Advancements

# Aerospace Technologies Advancements

Edited by

Dr. Thawar T. Arif

*Intech*

Published by Intech

**Intech**
Olajnica 19/2, 32000 Vukovar, Croatia


Free online edition of this book you can find under www.sciyo.com
Additional copies can be obtained from:
publication@sciyo.com

First published January 2010
Printed in India

Technical Editor: Teodora Smiljanic
Cover designed by Dino Smrekar

Aerospace Technologies Advancements, Edited by Dr. Thawar T. Arif
p. cm.
ISBN 978-953-7619-96-1

# Preface

Space technology has become increasingly important after the great development and rapid progress in information and communication technology as well as the technology of space exploration. This book deals with the latest and most prominent research in space technology.

The first part of the book (first six chapters) deals with the algorithms and software used in information processing, communications and control of spacecrafts.

The second part (chapters 7 to 10) deals with the latest research on the space structures.

The third part (chapters 11 to 14) deals with some of the latest applications in space.

The fourth part (chapters 15 and 16) deals with small satellite technologies.

The fifth part (chapters 17 to 20) deals with some of the latest applications in the field of aircrafts.

The sixth part (chapters 21 to 25) outlines some recent research efforts in different subjects.

The following is a brief description of the subjects that are covered in each chapter:

**Chapter 1** reviews some examples of how machine learning is useful for Geoscience and remote sensing.

**Chapter 2** describes how reuse activities can fit into building the next generation of aerospace data processing systems by providing guidance on methods for improving reuse practices in order to realize the benefits of reuse.

**Chapter 3** introduces reconfigurable computing as an emerging technology with important implications for space systems.

**Chapter 4** presents the decentralized minimal controller synthesis as an effective algorithm for controlling the spacecraft attitude.

**Chapter 5** presents the benefits of commercially available FPGA development platforms from Xilinx for the development of NASA's future on-board processing capabilities.

**Chapter 6** describes the employed mitigation techniques for the A3P product family, to attain the radiation levels of the RT-product and presents the results issued from the Total Ionizing Dose and the Single Event Effects characterization of both of the A3P and the A3PL (the Low-Power version of ProASIC3).

**Chapter 7** defines a framework for Evolving Systems, develops theory and control solutions for fundamental characteristics of Evolving Systems, and provides illustrative examples of Evolving Systems and their control with adaptive key component controllers.

**Chapter 8** provides a thorough end-to-end description of the process for evaluation of three different data-driven algorithms for anomaly detection to select the best candidate for deployment as part of a suite of IVHM (Integrated Vehicle Health Management) technologies.

**Chapter 9** aims at proving the feasibility of low-cost satellites using COTS (Commercial Off The Shelf) devices.

**Chapter 10** presents a literature survey on robot mobility systems for planetary surface exploration.

**Chapter 11** outlines the Multi-platform Atmospheric Sounding Testbed (MAST) which is an end-to-end simulation environment at Jet Propulsion Laboratory.

**Chapter 12** describes recent U.S. technology investments in electric propulsion thrusters with emphasis on mission application and low-thrust mission design for interplanetary trajectories and geosynchronous transfer using primary electric propulsion.

**Chapter 13** presents Radio occultation, which is a remote sensing sounding technique in which a radio frequency signal emitted from a spacecraft passes through an intervening planetary atmosphere before arriving at the receiver.

**Chapter 14** reports the progress in the development of the real time failure detection and prognostics technique which is used for the solid rocket motors.

**Chapter 15** describes in details the design of Low-cost Telecommunications CubeSatclass Spacecraft by using one digital signal processor with multitasking operating system which integrates all the intelligences of the satellite.

**Chapter 16** explores the future of system engineering of microsatellites.

**Chapter 17** discusses an open question in air traffic management that is whether or not algorithms can be realistically shown to meat Federal Aviation Administration in USA.

**Chapter 18** presents the application of ray tracing techniques, which are based on geometric optics, to model an IEEE 802.11a wireless system propagation map.

**Chapter 19** proposes to measure the impacts of the use of new technology sensors in the tracking systems currently used for Air Traffic Control applications.

**Chapter 20** presents a regime recognition algorithm developed based on hidden Markov model for Helicopter usage monitoring.

**Chapter 21** outlines some progresses made in simulating dragonfly compound eye imaging system, establishing elementary mathematical model developing compound eye equipment and introducing the electronic image stabilization.

**Chapter 22** introduces a new technique that could be used with the blind or non blind adaptive algorithms to enhance their performance.

**Chapter 23** presents an improved cloud detection technique at South China Sea.

**Chapter 24** introduces a new tunable filter concept for potential application in multispectral and hyperspectral imaging systems.

**Chapter 25** presents how to estimate the value imbedded in the risk transfer from the contractor to the government in a Multi-Year government Procurement (MYP) contract using real options analysis.

Editor

**Dr. Thawar T. Arif**
*Applied Science University*
*Jordan*

# Contents

# Part I

# 1

# Artificial Intelligence in Aerospace

David John Lary
*Joint Center for Earth Systems Technology (JCET) UMBC, NASA/GSFC*
*United States*

## 1. Introduction

Machine learning has recently found many applications in aerospace and remote sensing. These applications range from bias correction to retrieval algorithms, from code acceleration to detection of disease in crops. As a broad subfield of artificial intelligence, machine learning is concerned with algorithms and techniques that allow computers to "learn". The major focus of machine learning is to extract information from data automatically by computational and statistical methods.

Over the last decade there has been considerable progress in developing a machine learning methodology for a variety of Earth Science applications involving trace gases, retrievals, aerosol products, land surface products, vegetation indices, and most recently, ocean products (*Yi and Prybutok*, 1996, *Atkinson and Tatnall*, 1997, *Carpenter et al.*, 1997, *Comrie*, 1997, *Chevallier et al.*, 1998, *Hyyppa et al.*, 1998, *Gardner and Dorling*, 1999, *Lary et al.*, 2004, *Lary et al.*, 2007, *Brown et al.*, 2008, *Lary and Aulov*, 2008, *Caselli et al.*, 2009, *Lary et al.*, 2009). Some of this work has even received special recognition as a NASA Aura Science highlight (*Lary et al.*, 2007) and commendation from the NASA MODIS instrument team (*Lary et al.*, 2009). The two types of machine learning algorithms typically used are neural networks and support vector machines. In this chapter, we will review some examples of how machine learning is useful for Geoscience and remote sensing, these examples come from the author's own research.

## 2. Typical applications

One of the features that make machine-learning algorithms so useful is that they are "universal approximators". They can learn the behaviour of a system if they are given a comprehensive set of examples in a training dataset. These examples should span as much of the parameter space as possible. Effective learning of the system's behaviour can be achieved even if it is multivariate and non-linear. An additional useful feature is that we do not need to know a priori the functional form of the system as required by traditional least-squares fitting, in other words they are non-parametric, non-linear and multivariate learning algorithms.

The uses of machine learning to date have fallen into three basic categories which are widely applicable across all of the Geosciences and remote sensing, the first two categories use machine learning for its regression capabilities, the third category uses machine learning for its classification capabilities. We can characterize the three application themes are as follows: First, where we have a theoretical description of the system in the form of a deterministic

model, but the model is computationally *expensive*. In this situation, a machine-learning "wrapper" can be applied to the deterministic model providing us with a "code accelerator". A good example of this is in the case of atmospheric photochemistry where we need to solve a large coupled system of ordinary differential equations (ODEs) at a large grid of locations. It was found that applying a neural network wrapper to the system was able to provide a speed up of between a factor of 2 and 200 depending on the conditions. Second, when we do not have a deterministic model but we have data available enabling us to empirically learn the behaviour of the system. Examples of this would include: Learning inter-instrument bias between sensors with a temporal overlap, and inferring physical parameters from remotely sensed proxies. Third, machine learning can be used for classification, for example, in providing land surface type classifications. Support Vector Machines perform particularly well for classification problems.
Now that we have an overview of the typical applications, the sections that follow will introduce two of the most powerful machine learning approaches, neural networks and support vector machines and then present a variety of examples.

## 3. Machine learning

### 3.1 Neural networks

Neural networks are multivariate, non-parametric, 'learning' algorithms (*Haykin*, 1994, *Bishop*, 1995, 1998, *Haykin*, 2001a, *Haykin*, 2001b, 2007) inspired by biological neural networks. Computational neural networks (NN) consist of an interconnected group of artificial neurons that processes information in parallel using a connectionist approach to computation. A NN is a non-linear statistical data-modelling tool that can be used to model complex relationships between inputs and outputs or to find patterns in data. The basic computational element of a NN is a model neuron or node. A node receives input from other nodes, or an external source (e.g. the input variables). A schematic of an example NN is shown in Figure 1. Each input has an associated weight, $w$, that can be modified to mimic synaptic learning. The unit computes some function, $f$, of the weighted sum of its inputs:

$$y_i = \sum_j f\left(w_{ij} y_j\right) \tag{1}$$

Its output, in turn, can serve as input to other units. $w_{ij}$ refers to the weight from unit $j$ to unit $i$. The function $f$ is the node's activation or transfer function. The transfer function of a node defines the output of that node given an input or set of inputs. In the simplest case, $f$ is the identity function, and the unit's output is $y_i$, this is called a linear node. However, non-linear sigmoid functions are often used, such as the hyperbolic tangent sigmoid transfer function and the log-sigmoid transfer function. Figure 1 shows an example feed-forward perceptron NN with five inputs, a single output, and twelve nodes in a hidden layer. A perceptron is a computer model devised to represent or simulate the ability of the brain to recognize and discriminate. In most cases, a NN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase.
When we perform neural network training, we want to ensure we can independently assess the quality of the machine learning 'fit'. To insure this objective assessment we usually

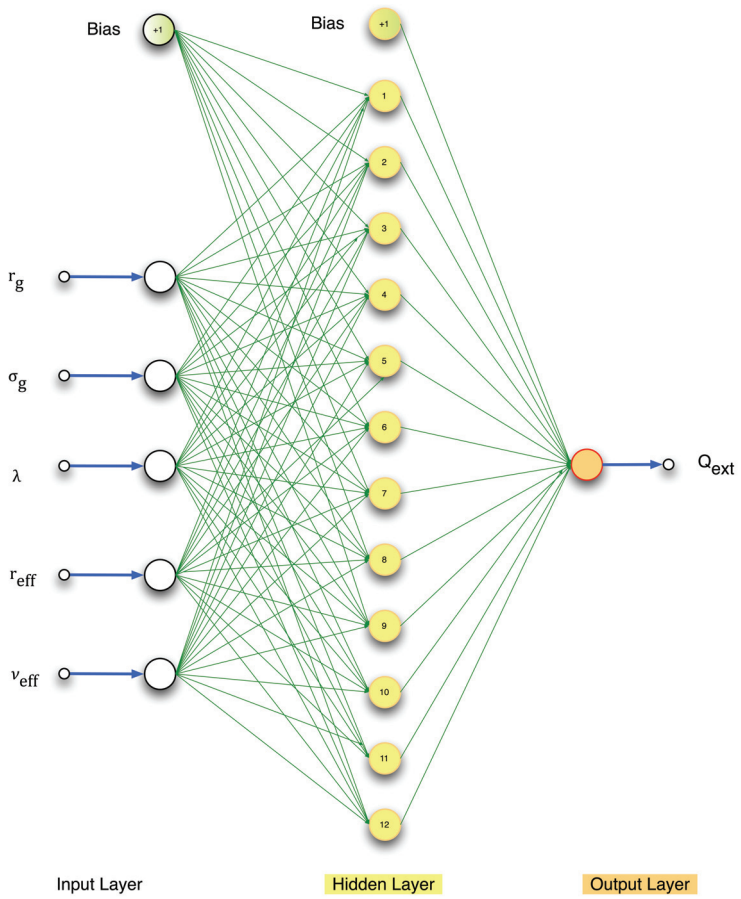



Fig. 1. Example neural network architecture showing a network with five inputs, one output, and twelve hidden nodes.

randomly split our training dataset into three portions, typically of 80%, 10% and 10%. The largest portion containing 80% of the dataset is used for training the neural network weights. This training is iterative, and on each training iteration we evaluate the current root mean square (RMS) error of the neural network output. The RMS error is calculated by using the second 10% portion of the data that was not used in the training. We use the RMS error and the way the RMS error changes with training iteration (epoch) to determine the convergence of our training. When the training is complete, we then use the final 10% portion of data as a totally independent validation dataset. This final 10% portion of the data is randomly chosen from the training dataset and is not used in either the training or RMS evaluation. We only use the neural network if the validation scatter diagram, which plots the actual data from validation portion against the neural network estimate, yields a straight-line graph with a slope very close to one and an intercept very close to zero. This is a stringent, independent and objective validation metric. The validation is global as the data

is randomly selected over all data points available. For our studies, we typically used feed-forward back-propagation neural networks with a Levenberg-Marquardt back-propagation training algorithm (*Levenberg*, 1944, *Marquardt*, 1963, *Moré*, 1977, *Marquardt*, 1979).

### 3.2 Support Vector Machines

Support Vector Machines (SVM) are based on the concept of decision planes that define decision boundaries and were first introduced by Vapnik (*Vapnik*, 1995, 1998, 2000) and has subsequently been extended by others (*Scholkopf et al.*, 2000, *Smola and Scholkopf*, 2004). A decision plane is one that separates between a set of objects having different class memberships. The simplest example is a linear classifier, i.e. a classifier that separates a set of objects into their respective groups with a line. However, most classification tasks are not that simple, and often more complex structures are needed in order to make an optimal separation, i.e., correctly classify new objects (test cases) on the basis of the examples that are available (training cases). Classification tasks based on drawing separating lines to distinguish between objects of different class memberships are known as hyperplane classifiers.

SVMs are a set of related supervised learning methods used for classification and regression. Viewing input data as two sets of vectors in an n-dimensional space, an SVM will construct a separating hyperplane in that space, one that maximizes the margin between the two data sets. To calculate the margin, two parallel hyperplanes are constructed, one on each side of the separating hyperplane, which are "pushed up against" the two data sets. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the neighboring data points of both classes, since in general the larger the margin the better the generalization error of the classifier. We typically used the SVMs provided by LIBSVM (*Fan et al.*, 2005, *Chen et al.*, 2006).

## 4. Applications

Let us now consider some applications.

### 4.1 Bias correction: atmospheric chlorine loading for ozone hole research

Critical in determining the speed at which the stratospheric ozone hole recovers is the total amount of atmospheric chlorine. Attributing changes in stratospheric ozone to changes in chlorine requires knowledge of the stratospheric chlorine abundance over time. Such attribution is central to international ozone assessments, such as those produced by the World Meteorological Organization (*Wmo*, 2006). However, we do not have continuous observations of all the key chlorine gases to provide such a continuous time series of stratospheric chlorine. To address this major limitation, we have devised a new technique that uses the long time series of available hydrochloric acid observations and neural networks to estimate the stratospheric chlorine ($Cl_y$) abundance (*Lary et al.*, 2007).

Knowledge of the distribution of inorganic chlorine $Cl_y$ in the stratosphere is needed to attribute changes in stratospheric ozone to changes in halogens, and to assess the realism of chemistry-climate models (*Eyring et al.*, 2006, *Eyring et al.*, 2007, *Waugh and Eyring*, 2008). However, simultaneous measurements of the major inorganic chlorine species are rare (*Zander et al.*, 1992, *Gunson et al.*, 1994, *Webster et al.*, 1994, *Michelsen et al.*, 1996, *Rinsland et al.*, 1996, *Zander et al.*, 1996, *Sen et al.*, 1999, *Bonne et al.*, 2000, *Voss et al.*, 2001, *Dufour et al.*, 2006,

*Nassar et al.*, 2006). In the upper stratosphere, the situation is a little easier as $Cl_y$ can be inferred from HCl alone (e.g., (*Anderson et al.*, 2000, *Froidevaux et al.*, 2006b, *Santee et al.*, 2008)). Our new estimates of stratospheric chlorine using machine learning (*Lary et al.*, 2007) work throughout the stratosphere and provide a much-needed critical test for current global models. This critical evaluation is necessary as there are significant differences in both the stratospheric chlorine and the timing of ozone recovery in the available model predictions.
Hydrochloric acid is the major reactive chlorine gas throughout much of the atmosphere, and throughout much of the year. However, the observations of HCl that we do have (from UARS HALOE, ATMOS, SCISAT-1 ACE and Aura MLS) have significant biases relative to each other. We found that machine learning can also address the inter-instrument bias (*Lary et al.*, 2007, *Lary and Aulov*, 2008). We compared measurements of HCl from the different instruments listed in Table 1. The Halogen Occultation Experiment (HALOE) provides the longest record of space based HCl observations. Figure 2 compares HALOE HCl with HCl observations from (a) the Atmospheric Trace Molecule Spectroscopy Experiment (ATMOS), (b) the Atmospheric Chemistry Experiment (ACE) and (c) the Microwave Limb Sounder (MLS).

| Instrument | Temporal Coverage | Species | References | Median Observation Uncertainty |
|---|---|---|---|---|
| ACE v2.2 | 2004-2006 | HCl, $ClONO_2$ | *Bernath et al.* [2005] | 8% (HCl), 30% ($ClONO_2$) |
| | | ClO, and HOCl | | >100% (ClO), >100% (HOCl) |
| ATMOS | 1991, 1993, 1994 | HCl, $ClONO_2$ | *Zander et al.* [1992] | 8% (HCl), 60% ($ClONO_2$) |
| Aura MLS v1 | 2004-2006 | HCl, ClO and HOCl | *Froidevaux et al.* [2006a] | 12% (HCl), 76% (ClO), >100% (HOCl) |
| CLAES v9 | 1991-1993 | $ClONO_2$ | *Roche et al.* [1993] | >100% |
| CRISTA | 1994, 1997 | $ClONO_2$ | *Offermann et al.* [1999] | 61% |
| HALOE v19 | 1991-2005 | HCl | *Russell et al.* [1993] | 4% |
| UARS MLS v5 | 1991-1999 | ClO | *Waters et al.* [1996] | >100% (ClO) |

Table 1. The instruments and constituents used in constructing the $Cl_y$ record from 1991-2006. The uncertainties given are the median values calculated for each level 2 measurement profile and its uncertainty (both in mixing ratio) for all the observations made. The uncertainties are larger than usually quoted for MLS ClO because they reflect the single profile precision, which is improved by temporal and/or spatial averaging. The HALOE uncertainties are only estimates of random error and do not include any indications of overall accuracy.

A consistent picture is seen in these plots: HALOE HCl measurements are lower than those from the other instruments. The slopes of the linear fits (relative scaling) are 1.05 for the HALOE-ATMOS comparison, 1.09 for the HALOE-MLS, and 1.18 for the HALOE-ACE. The offsets are apparent at the 525 K isentropic surface and above. Previous comparisons among HCl datasets reveal a similar bias for HALOE (*Russell et al.*, 1996, *Mchugh et al.*, 2005, *Froidevaux et al.*, 2006a, *Froidevaux et al.*, 2008). ACE and MLS HCl measurements are in much better agreement (Figure 2d). Note, the measurements agree within the stated observational uncertainties summarized in Table 1.
To combine the above HCl measurements to form a continuous time series of HCl (and then $Cl_y$) from 1991 to 2006 it is necessary to account for the biases between data sets. A neural network is used to learn the mapping from one set of measurements onto another as a function of equivalent latitude and potential temperature. We consider two cases. In one case ACE HCl is taken as the reference and the HALOE and Aura HCl observations are adjusted to agree with ACE HCl. In the other case HALOE HCl is taken as the reference and the Aura and ACE HCl observations are adjusted to agree with HALOE HCl. In both cases we use equivalent latitude and potential temperature to produce average profiles. The

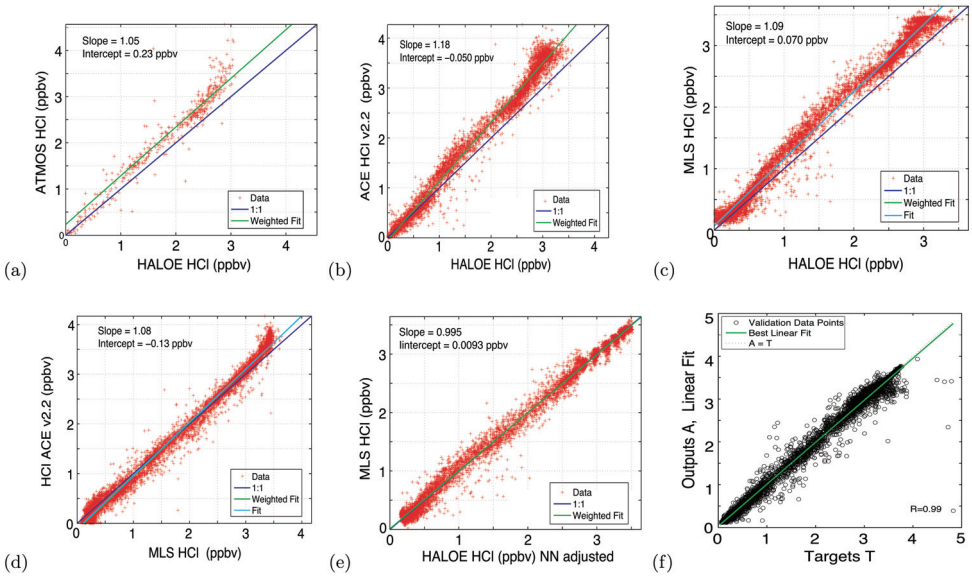



Fig. 2. Panels (a) to (d) show scatter plots of all contemporaneous observations of HCl made by HALOE, ATMOS, ACE and MLS Aura. In panels (a) to (c) HALOE is shown on the x-axis. Panel (e) correspond to panel (c) except that it uses the neural network 'adjusted' HALOE HCl values. Panel (f) shows the validation scatter diagram of the neural network estimate of $Cl_y \approx HCl + ClONO_2 + ClO + HOCl$ versus the actual $Cl_y$ for a totally independent data sample *not* used in training the neural network.

purpose of the NN mapping is simply to learn the bias as a function of location, not to imply which instrument is correct.

The precision of the correction using the neural network mapping is of the order of ±0.3 ppbv, as seen in Figure 2 (e) that shows the results when HALOE HCl measurements have been mapped into ACE measurements. The mapping has removed the bias between the measurements and has straightened out the 'wiggles' in 2 (c), i.e., the neural network has learned the equivalent PV latitude and potential temperature dependence of the bias between HALOE and MLS. The inter-instrument offsets are not constant in space or time, and are not a simple function of $Cl_y$.

So employing neural networks allows us to: Form a seamless record of HCl using observations from several space-borne instruments using neural networks. Provide an estimated of the associated inter-instrument bias. Infer $Cl_y$ from HCl, and thereby provide a seamless record of $Cl_y$, the parameter needed for examining the ozone hole recovery. A similar use of machine learning has been made for Aerosol Optical Depths, the subject of the next sub-section.

### 4.2 Bias correction: aerosol optical depth

As highlighted in the 2007 IPCC report on Climate Change, aerosol and cloud radiative effects remain the largest uncertainties in our understanding of climate change (*Solomon et al.*, 2007). Over the past decade observations and retrievals of aerosol characteristics have

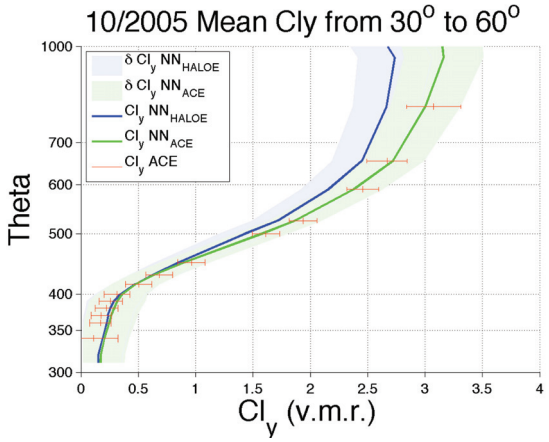



Fig. 3. $Cl_y$ average profiles between 30° and 60°N for October 2005, estimated by neural network calibrated to HALOE HCl (blue curve), estimated by neural network calibrated to ACE HCl (green), or from ACE observations of HCl, $ClONO_2$, ClO, and HOCl (red crosses). In each case, the shaded range represents the total uncertainty; it includes the observational uncertainty, the representativeness uncertainty (the variability over the analysis grid cell), the neural network uncertainty. The vertical extent of this plot was limited to below 1000 K (≈35 km), as there is no ACE v2.2 ClO data for the upper altitudes. In addition, above ≈750 K (≈25 km), ClO constitutes a larger fraction of $Cl_y$ (up to about 10%) and so the large uncertainties in ClO have greater effect.

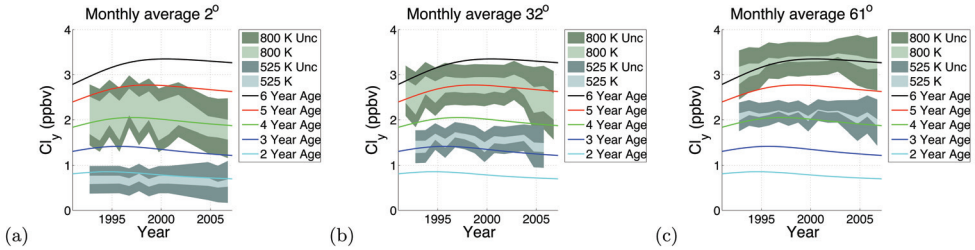



Fig. 4. Panels (a) to (c) show October $Cl_y$ time-series for the 525 K isentropic surface (≈20 km) and the 800 K isentropic surface (≈30 km). In each case the dark shaded range represents the total uncertainty in our estimate of $Cl_y$. This total uncertainty includes the observational uncertainty, the representativeness uncertainty (the variability over the analysis grid cell), the inter-instrument bias in HCl, the uncertainty associated with the neural network inter-instrument correction, and the uncertainty associated with the neural network inference of $Cl_y$ from HCl and $CH_4$. The inner light shading depicts the uncertainty on $Cl_y$ due to the inter-instrument bias in HCl alone. The upper limit of the light shaded range corresponds to the estimate of $Cl_y$ based on all the HCl observations calibrated by a neural network to agree with ACE v2.2 HCl. The lower limit of the light shaded range corresponds to the estimate of $Cl_y$ based on all the HCl observations calibrated to agree with HALOE v19 HCl. Overlaid are lines showing the $Cl_y$ based on age of air calculations (*Newman et al.*, 2006). To minimize variations due to differing data coverage months with less than 100 observations of HCl in the equivalent latitude bin were left out of the time-series.

been conducted from space-based sensors, from airborne instruments and from ground-based samplers and radiometers. Much effort has been directed at these data sets to collocate observations and retrievals, and to compare results. Ideally, when two instruments measure the same aerosol characteristic at the same time, the results should agree within well-understood measurement uncertainties. When inter-instrument biases exist, we would like to explain them theoretically from first principles. One example of this is the comparison between the aerosol optical depth (AOD) retrieved by the Moderate Resolution Imaging Spectroradiometer (MODIS) and the AOD measured by the Aerosol Robotics Network (AERONET). While progress has been made in understanding the biases between these two data sets, we still have an imperfect understanding of the root causes. (*Lary et al.*, 2009) examined the efficacy of empirical machine learning algorithms for aerosol bias correction.

Machine learning approaches (Neural Networks and Support Vector Machines) were used by (*Lary et al.*, 2009) to explore the reasons for a persistent bias between aerosol optical depth (AOD) retrieved from the MODerate resolution Imaging Spectroradiometer (MODIS) and the accurate ground-based Aerosol Robotics Network (AERONET). While this bias falls within the expected uncertainty of the MODIS algorithms, there is still room for algorithm improvement. The results of the machine learning approaches suggest a link between the MODIS AOD biases and surface type. From figure 5 we can see that machine learning algorithms were able to effectively adjust the AOD bias seen between the MODIS instruments and AERONET. Support vector machines performed the best improving the correlation coefficient between the AERONET AOD and the MODIS AOD from 0.86 to 0.99 for MODIS Aqua, and from 0.84 to 0.99 for MODIS Terra.

Key in allowing the machine learning algorithms to 'correct' the MODIS bias was provision of the surface type and other ancillary variables that explain the variance between MODIS and AERONET AOD. The provision of the ancillary variables that can explain the variance in the dataset is the key ingredient for the effective use of machine learning for bias correction. A similar use of machine learning has been made for vegetation indices, the subject of the next sub-section.

### 4.3 Bias correction: vegetation indices

Consistent, long term vegetation data records are critical for analysis of the impact of global change on terrestrial ecosystems. Continuous observations of terrestrial ecosystems through time are necessary to document changes in magnitude or variability in an ecosystem (*Tucker et al.*, 2001, *Eklundh and Olsson*, 2003, *Slayback et al.*, 2003). Satellite remote sensing has been the primary way that scientists have measured global trends in vegetation, as the measurements are both global and temporally frequent. In order to extend measurements through time, multiple sensors with different design and resolution must be used together in the same time series. This presents significant problems as sensor band placement, spectral response, processing, and atmospheric correction of the observations can vary significantly and impact the comparability of the measurements (*Brown et al.*, 2006). Even without differences in atmospheric correction, vegetation index values for the same target recorded under identical conditions will not be directly comparable because input reflectance values differ from sensor to sensor due to differences in sensor design (*Teillet et al.*, 1997, *Miura et al.*, 2006).

Several approaches have previously been taken to integrate data from multiple sensors. (*Steven et al.*, 2003), for example, simulated the spectral response from multiple instruments

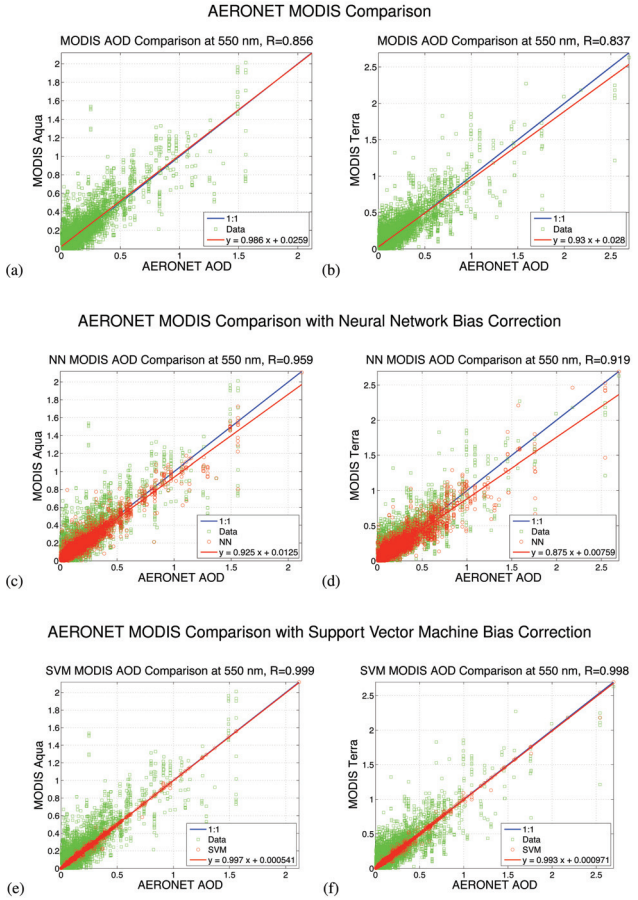



Fig. 5. Scatter diagram comparisons of Aerosol Optical Depth (AOD) from AERONET (x-axis) and MODIS (y-axis) as green circles overlaid with the ideal case of perfect agreement (blue line). The measurements shown in the comparison were made within half an hour of each other, with a great circle separation of less than 0.25° and with a solar zenith angle difference of less than 0.1°. The left hand column of plots is for MODIS Aqua and the right hand column of plots is for MODIS Terra. The first row shows the comparisons between AERONET and MODIS for the entire period of overlap between the MODIS and AERONET instruments from the launch of the MODIS instrument to the present. The second row shows the same comparison overlaid with the neural network correction as red circles. We note that the neural network bias correction makes a substantial improvement in the correlation coefficient with AERONET. An improvement from 0.86 to 0.96 for MODIS Aqua and an improvement from 0.84 to 0.92 for MODIS Terra. The third row shows the comparison overlaid with the support vector regression correction as red circles. We note that the support vector regression bias correction makes an even greater improvement in the correlation coefficient than the neural network correction. An improvement from 0.86 to 0.99 for MODIS Aqua and an improvement from 0.84 to 0.99 for MODIS Terra.

and with simple linear equations created conversion coefficients to transform NDVI data from one sensor to another. Their analysis is based on the observation that the vegetation index is critically dependent on the spectral response functions of the instrument used to calculate it. The conversion formulas the paper presents cannot be applied to maximum value NDVI datasets because the weighting coefficients are land cover and dataset dependent, reducing their efficacy in mixed pixel situations (*Steven et al.*, 2003). (*Trishchenko et al.*, 2002) created a series of quadratic functions to correct for differences in the reflectance and NDVI to NOAA-9 AVHRR-equivalents (*Trishchenko et al.*, 2002). Both the (*Steven et al.*, 2003) and the (*Trishchenko et al.*, 2002) approaches are land cover and dataset dependent and thus cannot be used on global datasets where multiple land covers are represented by one pixel. (*Miura et al.*, 2006) used hyper-spectral data to investigate the effect of different spectral response characteristics between MODIS and AVHRR instruments on both the reflectance and NDVI data, showing that the precise characteristics of the spectral response had a large effect on the resulting vegetation index. The complex patterns and dependencies on spectral band functions were both land cover dependent and strongly non-linear, thus we see that an exploration of a non-linear approach may be fruitful.

(*Brown et al.*, 2008) experimented with powerful, non-linear neural networks to identify and remove differences in sensor design and variable atmospheric contamination from the AVHRR NDVI record in order to match the range and variance of MODIS NDVI without removing the desired signal representing the underlying vegetation dynamics. Neural networks are 'data transformers' (*Atkinson and Tatnall*, 1997), where the objective is to associate the elements of one set of data to the elements in another. Relationships between the two datasets can be complex and the two datasets may have different statistical distributions. In addition, neural networks incorporate a priori knowledge and realistic physical constraints into the analysis, enabling a transformation from one dataset into another through a set of weighting functions (*Atkinson and Tatnall*, 1997). This transformation incorporates additional input data that may account for differences between the two datasets.

The objective of (*Brown et al.*, 2008) was to demonstrate the viability of neural networks as a tool to produce a long term dataset based on AVHRR NDVI that has the data range and statistical distribution of MODIS NDVI. Previous work has shown that the relationship between AVHRR and MODIS NDVI is complex and nonlinear (*Gallo et al.*, 2003, *Brown et al.*, 2006, *Miura et al.*, 2006), thus this problem is well suited to neural networks if appropriate inputs can be found. The influence of the variation of atmospheric contamination of the AVHRR data through time was explored by using observed atmospheric water vapor from the Total Ozone Mapping Spectrometer (TOMS) instrument during the overlap period 2000-2004 and back to 1985. Examination of the resulting MODIS fitted AVHRR dataset both during the overlap period and in the historical dataset will enable an evaluation of the efficacy of the neural net approach compared to other approaches to merge multiple-sensor NDVI datasets.

Remote sensing datasets are the result of a complex interaction between the design of a sensor, the spectral response function, stability in orbit, the processing of the raw data, compositing schemes, and post-processing corrections for various atmospheric effects including clouds and aerosols. The interaction between these various elements is often non-linear and non-additive, where some elements increase the vegetation signal to noise ratio (compositing, for example) and others reduce it (clouds and volcanic aerosols) (*Los*, 1998). Thus, although other authors have used simulated data to explore the relationship between

AVHRR and MODIS (*Trishchenko et al.*, 2002, *Van Leeuwen et al.*, 2006), these techniques are not directly useful in producing a sensor-independent vegetation dataset that can be used by data users in the near term.

There are substantial differences between the processed vegetation data from AVHRR and MODIS. (*Brown et al.*, 2008) showed that neural networks are an effective way to have a long data record that utilizes all available data back to 1981 by providing a practical way of incorporating the AVHRR data into a continuum of observations that include both MODIS and VIIRS. The results (*Brown et al.*, 2008) showed that the TOMS data record on clouds, ozone and aerosols can be used to identify and remove sensor-specific atmospheric contaminants that differentially affect the AVHRR over MODIS. Other sensor-related effects, particularly those of changing BRDF, viewing angle, illumination, and other effects that are not accounted for here, remain important sources of additional variability. Although this analysis has not produced a dataset with identical properties to MODIS, it has demonstrated that a neural net approach can remove most of the atmospheric-related aspects of the differences between the sensors, and match the mean, standard deviation and range of the two sensors. A similar technique can be used for the VIIRS sensor once the data is released.

Figure 6 shows a comparison of the NDVI from AVHR (panel a), MODIS (panel p), and then a reconstruction of MODIS using AVHRR and machine learning (panel c). Figure 7 (a) shows a time-series from 2000 to 2003 of the zonal mean difference between the AVHRR and MODIS NDVIs, this highlights that significant differences exist between the two data products. Panel (b) shows a time series over the same period after the machine learning has been used to "cross-calibrate" AVHRR as MODIS, illustrating that the machine learning has effectively learnt how to cross-calibrate the instruments.

So far, we have seen three examples of using machine learning for bias correction (constituent biases, aerosol optical depth biases and vegetation index biases), and one example of using machine learning to infer a useful proxy from remotely sensed data ($Cl_y$ from HCl). Let us look at one more example of inferring proxies from existing remotely sensed data before moving onto consider using machine learning for code acceleration.

### 4.4 Inferring proxies: tracer correlations

The spatial distributions of atmospheric trace constituents are in general dependent on both chemistry and transport. Compact correlations between long-lived species are well-observed features in the middle atmosphere. The correlations exist for all long-lived tracers - not just those that are chemically related - due to their transport by the general circulation of the atmosphere. The tight relationships between different constituents have led to many analyses using measurements of one tracer to infer the abundance of another tracer. Using these correlations is also as a diagnostic of mixing and can distinguish between air-parcels of different origins. Of special interest are the so-called 'long-lived' tracers: constituents such as nitrous oxide ($N_2O$), methane ($CH_4$), and the chlorofluorocarbons (CFCs) that have long lifetimes (many years) in the troposphere and lower stratosphere, but are destroyed rapidly in the middle and upper stratosphere.

The correlations are spatially and temporally dependent. For example, there is a 'compact-relation' regime in the lower part of the stratosphere and an 'altitude-dependent' regime above this. In the compact-relation region, the abundance of one tracer is uniquely determined by the value of the other tracer, without regard to other variables such as

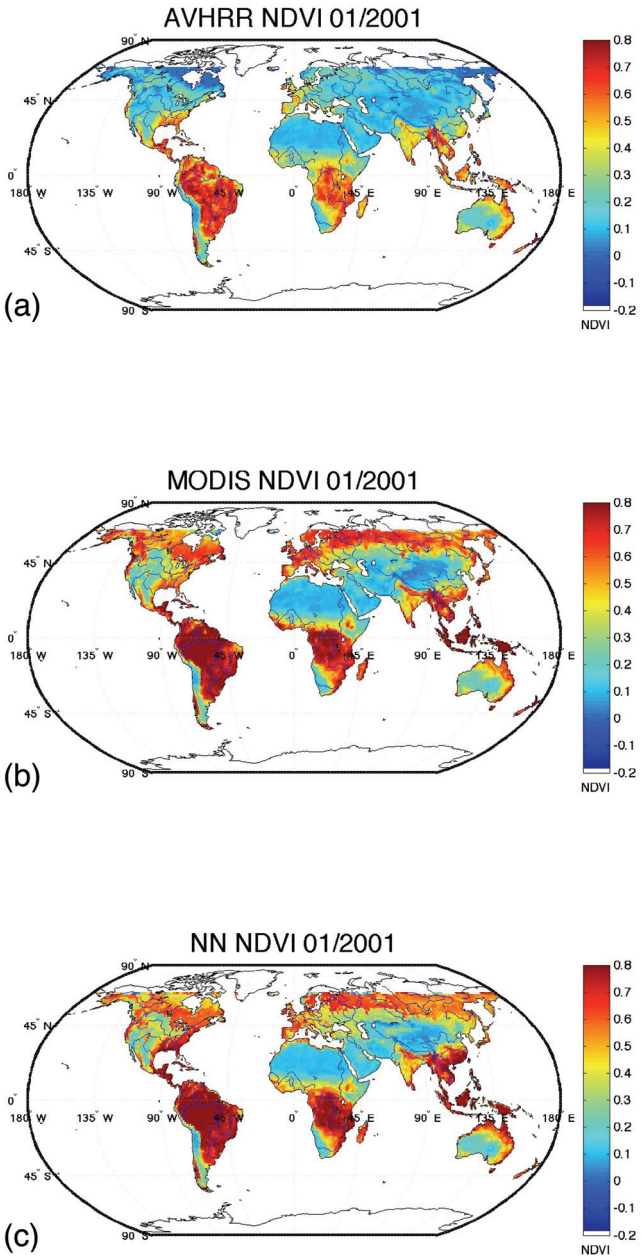



Fig. 6. A comparison of the NDVI from AVHR (panel a), MODIS (panel p), and then a reconstruction of MODIS using AVHRR and machine learning (panel c). We note that the machine learning can successfully account for the large differences that are found between AVHRR and MODIS.

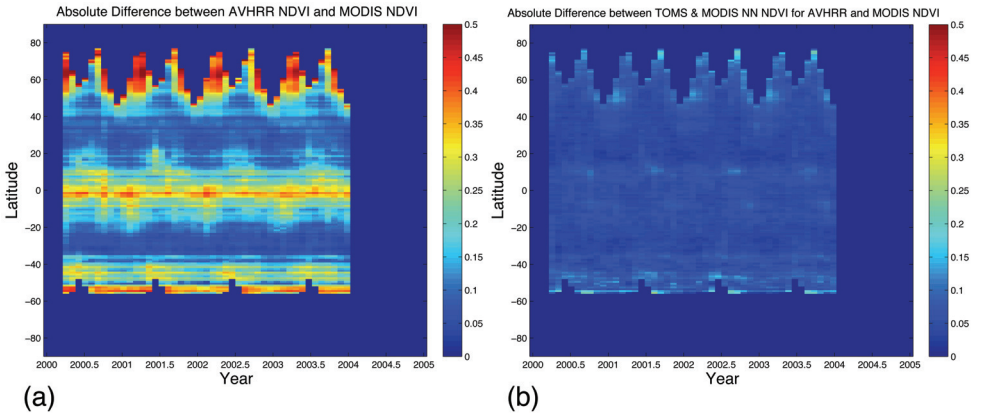



Fig. 7. Panel (a) shows a time-series from 2000 to 2003 of the zonal mean (averaged per latitude) difference between the AVHRR and MODIS NDVIs, this highlights that significant differences exist between the two data products. Panel (b) shows a time series over the same period after the machine learning has been used to "cross-calibrate" AVHRR as MODIS, showing that the machine learning has effectively learnt how to cross-calibrate the instruments.

latitude or altitude. In the altitude-dependent regime, the correlation generally shows significant variation with altitude.

A family of correlations usually achieves the description of such spatially and temporally dependent correlations. However, a single neural network is a natural and effective alternative. The motivation for this case study was preparation for a long-term chemical assimilation of Upper Atmosphere Research Satellite (UARS) data starting in 1991 and coming up to the present. For this period, we have continuous version 19 data from the Halogen Occultation Experiment (HALOE) but not observations of $N_2O$ as both ISAMS and CLAES failed. In addition, we would like to constrain the total amount of reactive nitrogen, chlorine, and bromine in a self-consistent way (i.e. the correlations between the long-lived tracers is preserved). Tracer correlations provide a means to do this by using HALOE $CH_4$ observations.

Machine learning is ideally suited to describe the spatial and temporal dependence of tracer-tracer correlations. The neural network performs well even in regions where the correlations are less compact and normally a family of correlation curves would be required. For example, the methane $CH_4$-$N_2O$ correlation can be well described using a neural network (*Lary et al.*, 2004) trained with the latitude, pressure, time of year, and $CH_4$ volume mixing ratio (v.m.r.). *Lary et al.* (2004) used a neural network to reproduce the $CH_4$-$N_2O$ correlation with a correlation coefficient between simulated and training values of 0.9995. Such an accurate representation of tracer-tracer correlations allows more use to be made of long-term datasets to constrain chemical models. For example, the Halogen Occultation Experiment (HALOE) that continuously observed $CH_4$ (but not $N_2O$) from 1991 until 2005.

Figure 8 (a) shows the global $N_2O$-$CH_4$ correlation for an entire year, after evaluating the efficacy of 3,000 different functional forms for parametric fits, we overlaid the best, an order 20 Chebyshev Polynomial. However, this still does not account for the multi-variate nature

of the problem exhibited by the 'cloud' of points rather than a compact 'curve' or 'line'. However, in Figure 8 (b) we can see that a neural network is able to account for the non-linear and multi-variate aspects, the training dataset exhibited a 'cloud' of points, the neural network fit reproduces a 'cloud' of points. The most important factor in producing a 'spread' in the correlations is the strong altitude dependence of the $N_2O$-$CH_4$ correlation.

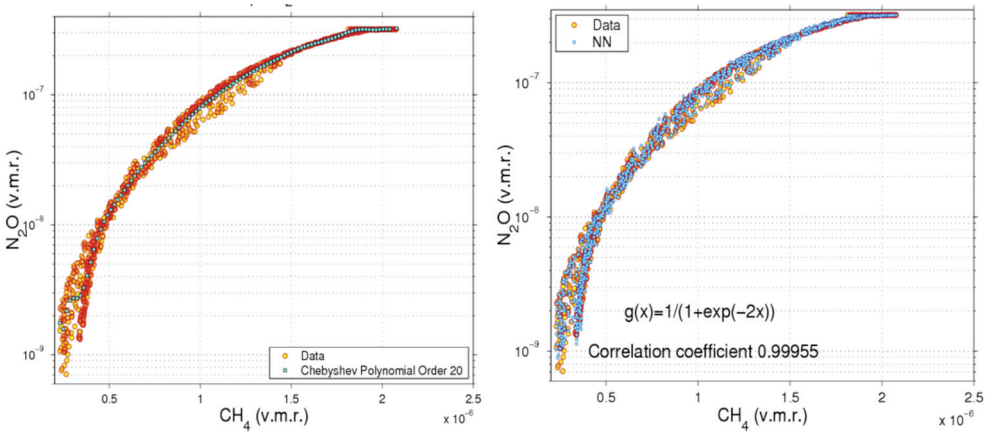



Fig. 8. Panel (a) shows the global $N_2O$-$CH_4$ correlation for an entire year, after evaluating the efficacy of 3,000 different functional forms for parametric fits, we overlaid the best, an order 20 Chebyshev Polynomial. However, this still does not account for the multi-variate nature of the problem exhibited by the 'cloud' of points rather than a compact 'curve' or 'line'. However, in panel (b) we can see that a neural network is able to account for the non-linear and multi-variate aspects, the training dataset exhibited a 'cloud' of points, the neural network fit reproduces a 'cloud' of points. The most important factor in producing a 'spread' in the correlations is the strong altitude dependence of the $N_2O$-$CH_4$ correlation.

### 4.5 Code acceleration: example from ordinary differential equation solvers

There are many applications in the Geosciences and remote sensing which are computationally expensive. Machine learning can be very effective in accelerating components of these calculations. We can readily create training datasets for these applications using the very models we would like to accelerate.

The first example for which we found this effective was solving ordinary differential equations. An adequate photochemical mechanism to describe the evolution of ozone in the upper troposphere and lower stratosphere (UT/LS) in a computational model involves a comprehensive treatment of reactive nitrogen, hydrogen, halogens, hydrocarbons, and interactions with aerosols. Describing this complex interaction is computationally expensive, and applications are limited by the computational burden. Simulations are often made tractable by using a coarser horizontal resolution than would be desired or by reducing the interactions accounted for in the photochemical mechanism. These compromises also limit the scientific applications. Machine learning algorithms offer a means to obtain a fast and accurate solution to the stiff ordinary differential equations that comprise the photochemical calculations, thus making high-resolution simulations including the complete photochemical mechanism much more tractable.

For the sake of an example, a 3D model of atmospheric chemistry and transport, the GMI-COMBO model, can use 55 vertical levels and a 4° latitude x 5° longitude grid and 125 species. With 15-minute time steps the chemical ODE solver is called 119,750,400 times in simulating just one week. If the simulation is for a year then the ODE solver needs to be called 6,227,020,800 (or $6x10^9$) times. If the spatial and temporal resolution is doubled then the chemical ODE solver needs to be called a staggering $2.5x10^{10}$ times to simulate a year. This represents a major computational cost in simulating a constituent's spatial and temporal evolution. The ODEs solved at adjacent grid cells and time steps are very similar. Therefore, if the simulations from one grid cell and time step could be used to speed up the simulation for adjacent grid cells and subsequent time steps, we would have a strategy to dramatically decrease the computational cost of our simulations.

Figure 9 shows the strategy that we used for applying a neural wrapper to accelerate the ODE solver. Figure 10 shows some example results for ozone after using a neural wrapper around an atmospheric chemistry ODE solver. The x-axis shows the actual ozone abundance as a volume mixing ratio (vmr) using the regular ODE solver without neural networks. The y-axis shows the ozone vmr inferred using the neural network solution. It can be seen that we have excellent agreement between the two solutions with a correlation coefficient of 1. The neural network has learned the behaviour of the ozone ODE very well. Without the adaptive error control the acceleration could be up to 200 times, with the full adaptive error control the acceleration was less, but usually at least a factor of two. Similarly, in Figure 11 the two panels below show the results for formaldehyde (HCHO) in the GMI model. The left panel shows the solution with SMVGear for level 1 at 01:00 UT and the right panel shows the corresponding solution using the neural network. As one would hope, the two results are almost indistinguishable.

**Neurological Acceleration of stiff ODE solver with error monitoring and dynamic training**

| Phase | Standard Stiff ODE Solver Action | Neural Network Action | Description |
|---|---|---|---|
| **A** | Always running | None | Collecting training data. |
| **B** | None | Initial training | Train *n* networks, one per constituent ODE. |
| **C** | Always running | Testing | Compare standard stiff ODE solver with neural network solution to train error networks, one per constituent ODE. |
| **D** | Always running | Update training | Update training of *n* ODE networks. Train *n* error networks giving neural network solution uncertainty. |
| **E** | Sometimes running | Operational with periodic training updates | **Operational.** When error networks indicate an error from the neural network solution that is less than the required threshold for each ODE use the neural network solution, otherwise use the regular stiff ODE solver. However, even when error network indicates an acceptable neural network solution periodically use the full stiff ODE solver solution to gather further training data to periodically update the training of all the networks. |

Fig. 9. Strategy for applying a neural wrapper to accelerate the ODE solver.

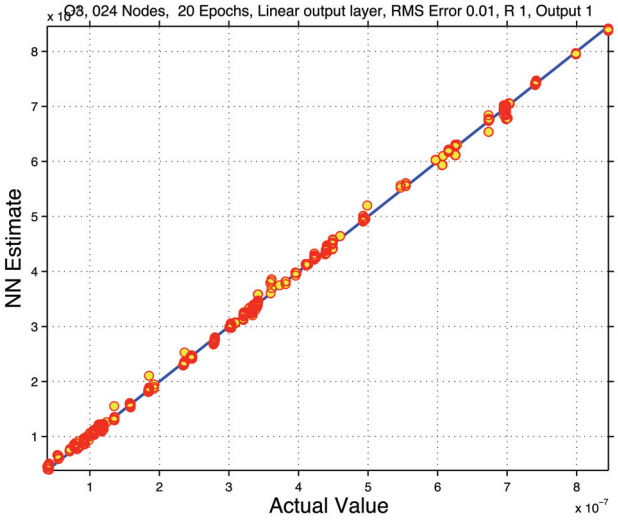



Fig. 10. Example results for using a neural wrapper around an atmospheric chemistry ODE solver. The x-axis shows the actual ozone v.m.r. using the regular ODE solver without neural networks. The y-axis shows the ozone v.m.r. inferred using the neural network solution. It can be seen that we have excellent agreement between the two solutions with a correlation coefficient of 1. The neural network has learned the behaviour of the ozone ODE very well.

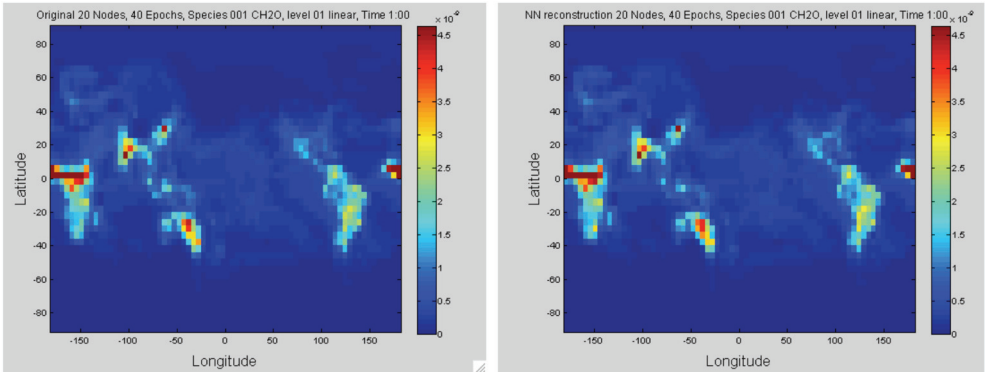



Fig. 11. The two panels below show the results for formaldehyde (HCHO) in the GMI model. The left panel shows the solution with SMVGear for level 1 at 01:00 UT and the right panel shows the corresponding solution using the neural network. As one would hope, the two results are almost indistinguishable.

### 4.6 Classification: example from detecting drought stress and infection in cacao

The source of chocolate, theobroma cacao (cacao), is an understory tropical tree (*Wood*, 2001). Cacao is intolerant to drought (*Belsky & Siebert*, 2003), and yields and production patterns are severely affected by periodic droughts and seasonal rainfall patterns. (*Bae et al.*,

2008) studied the molecular response of cacao to drought and have identified several genes responsive to drought stress (*Bailey et al.*, 2006). They have also been studying the response of cacao to colonization by an endophytic isolates of Trichoderma including Trichoderma hamatum, DIS 219b (*Bailey et al.*, 2006). One of the benefits to colonization Trichoderma hamatum isolate DIS 219b is tolerance to drought as mediated through plant growth promotion, specifically enhanced root growth (*Bae et al.*, 2008).

In characterizing the drought response of cacao considerable variation was observed in the response of individual seedlings depending upon the degree of drought stress applied (*Bae et al.*, 2008). In addition, although colonization by DIS 219b delayed the drought response, direct effects of DIS 219b on cacao gene expression in the absence of drought were difficult to identify (*Bae et al.*, 2008). The complexity of the DIS 219b/cacao plant microbe interaction overlaid on cacao's response to drought makes the system of looking at individual genes as a marker for either drought or endophyte inefficient.

There would be considerable utility in reliably predicting drought and endophyte stress from complex gene expression patterns, particularly as the endophyte lives within the plant without causing apparent phenotypic changes in the plant. Machine - learning models offer the possibility of highly accurate, automated predictions of plant stress from a variety of causes that may otherwise go undetected or be obscured by the complexity of plant responses to multiple environmental factors, to be considered status quo for plants in nature. We examined the ability of five different machine - learning approaches to predict drought stress and endophyte colonization in cacao: a naive Bayes classifier, decision trees (DTs), neural networks (NN), neuro-fuzzy inference (NFI), and support vector machine (SVM) classification. The results provided some support for the accuracy of machine-learning models in discerning endophyte colonization and drought stress. The best performance was by the neuro-fuzzy inference system and the support vector classifier that correctly identified 100% of the drought and endophyte stress samples. Of the two, the approaches the support vector classifier is likely to have the best generalization (wider applicability to data not previously seen in the training process).

Why did the SVM model outperform the four other machine learning approaches? We noted earlier that SVMs construct separating hyperplanes that maximize the margins between the different clusters in the training data set (the vectors that constrain the width of the margin are the support vectors). A good separation is achieved by those hyperplanes providing the largest distance between neighbouring classes, and in general, the larger the margin the better the generalization of the classifier.

When the points in neighbouring classes are separated by a nonlinear dividing line, rather than fitting nonlinear curves to the data, SVMs use a kernel function to map the data into a different space where a hyperplane can once more be used to do the separation. The kernel function may transform the data into a higher dimensional space to make it possible to perform the separation. The concept of a kernel mapping function is very powerful. It allows SVM models to perform separations even with very complex boundaries. Hence, we infer that, in the present application, the SVM model algorithmic process utilizes higher dimensional space to achieve superior predictive power.

For classification, the SVM algorithmic process offers an important advantage compared with neural network approaches. Specifically, neural networks can suffer from multiple local minima; in contrast, the solution to a support vector machine is global and unique. This characteristic may be partially attributed to the development process of these algorithms;

SVMs were developed in the reverse order to the development of neural networks. SVMs evolved from the theory to implementation and experiments; neural networks followed a more heuristic path, from applications and extensive experimentation to theory.
In handling this data using traditional methods where individual gene responses are characterized as treatment effects, it was especially difficult to sort out direct effects of endophyte on gene expression over time or at specific time points. The differences between the responses of non-stressed plants with or without the endophyte were small and, after the zero time point, were highly variable. The general conclusion from this study was that colonization of cacao seedlings by the endophyte enhanced root growth resulting in increased drought tolerance but the direct effects of endophyte on cacao gene expression at the time points studied were minimal. Yet the neuro-fuzzy inference and support vector classification methods of analysis were able identify samples receiving these treatments correctly.
In this system, each gene in the plants genome is a potential sensor for the applied stress or treatment. It is not necessary that the genes response be significant in itself in determining the outcome of the plants response or that it be consistent in time or level of response. Since multiple genes are used in characterizing the response it is always the relative response in terms of the many other changes that are occurring at the same time as influenced by uncontrolled changes in the system that is important. With this study the treatments were controlled but variation in the genetic make up of each seedling (they were from segregating open pollinated seed) and minute differences in air currents within the chamber, soil composition, colonization levels, microbial populations within each pot and seedling, and even exact watering levels at each time point, all likely contributed to creating uncontrolled variation in the plants response to what is already a complex reaction to multiple factors (drought and endophyte). This type of variation makes accessing treatment responses using single gene approaches difficult and the prediction of cause due to effect in open systems almost impossible in complex systems.

## 5. Future directions

We have seen the utility of machine learning for a suite of very diverse applications. These applications often help us make better use of existing data in a variety of ways. In parallel to the success of machine learning we also have the rapid development of publically available web services. So it is timely to combine both approached by providing online services that use machine learning for intelligent data fusion as part of a workflow that allows us to cross-calibrate multiple datasets. This obviously requires care to ensure the appropriate of datasets. However, if done carefully, this could greatly facilitate the production of seamless multi-year global records for a host of Earth science applications.
When it comes to dealing with inter-instrument biases in a consistent manner there is currently a gap in many space agencies' Earth science information systems. This could be addressed by providing an extensible and reusable open source infrastructure that gap that could be reused for multiple projects. A clear need for such an infrastructure would be for NASA's future Decadal Survey missions.

## 6. Summary

Machine learning has recently found many applications in the geosciences aerospace and remote sensing. These applications range from bias correction to retrieval algorithms, from

code acceleration to detection of disease in crops. Machine-learning algorithms can act as "universal approximators", they can learn the behaviour of a system if they are given a comprehensive set of examples in a training dataset. Effective learning of the system's behaviour can be achieved even if it is multivariate and non-linear. An additional useful feature is that we do not need to know a priori the functional form of the system as required by traditional least-squares fitting, in other words they are non-parametric, non-linear and multivariate learning algorithms.

The uses of machine learning to date have fallen into three basic categories which are widely applicable across all of the Geosciences and remote sensing, the first two categories use machine learning for its regression capabilities, the third category uses machine learning for its classification capabilities. We can characterize the three application themes are as follows: First, where we have a theoretical description of the system in the form of a deterministic model, but the model is computationally *expensive*. In this situation, a machine-learning "wrapper" can be applied to the deterministic model providing us with a "code accelerator". Second, when we do not have a deterministic model but we have data available enabling us to empirically learn the behaviour of the system. Third, machine learning can be used for classification.

## 7. References


Anderson, J., Russell, J. M., Solomon, S. & Deaver, L. E. (2000) Halogen occultation experiment confirmation of stratospheric chlorine decreases in accordance with the montreal protocol. *Journal of Geophysical Research-Atmospheres,* 105, 4483-4490.

Atkinson, P. M. & Tatnall, A. R. L. (1997) Introduction: Neural networks in remote sensing. *International Journal of Remote Sensing,* 18, 699 - 709.

Bae, H., Kim, S. H., Kim, M. S., Sicher, R. C., Lary, D., Strem, M. D., Natarajan, S. & Bailey, B. A. (2008) The drought response of theobroma cacao (cacao) and the regulation of genes involved in polyamine biosynthesis by drought and other stresses. *Plant Physiology and Biochemistry,* 46, 174-188.

Bailey, B. A., Bae, H., Strem, M. D., Roberts, D. P., Thomas, S. E., Crozier, J., Samuels, G. J., Choi, I. Y. & Holmes, K. A. (2006) Fungal and plant gene expression during the colonization of cacao seedlings by endophytic isolates of four trichoderma species. *Planta,* 224, 1449-1464.

Belsky, J. M. & Siebert, S. F. (2003) Cultivating cacao: Implications of sun-grown cacao on local food security and environmental sustainability. *Agriculture and Human Values,* 20, 277-285.

Bishop, C. M. (1995) *Neural networks for pattern recognition,* Oxford, Oxford University Press.

Bishop, C. M. (1998) *Neural networks and machine learning,* Berlin; New York, Springer.

Bonne, G. P., Stimpfle, R. M., Cohen, R. C., Voss, P. B., Perkins, K. K., Anderson, J. G., Salawitch, R. J., Elkins, J. W., Dutton, G. S., Jucks, K. W. & Toon, G. C. (2000) An examination of the inorganic chlorine budget in the lower stratosphere. *Journal of Geophysical Research-Atmospheres,* 105, 1957-1971.

Brown, M. E., Lary, D. J., Vrieling, A., Stathakis, D. & Mussa, H. (2008) Neural networks as a tool for constructing continuous NDVI time series from AVHRR and MODIS. *International Journal of Remote Sensing,* 29, 7141-7158.

Brown, M. E., Pinzon, J. E., Didan, K., Morisette, J. T. & Tucker, C. J. (2006) Evaluation of the consistency of long-term NDVI time series derived from AVHRR, spot-vegetation,

seawifs, MODIS and landsat etm+. *IEEE Transactions Geoscience and Remote Sensing,* 44, 1787-1793.

Carpenter, G. A., Gjaja, M. N., Gopal, S. & Woodcock, C. E. (1997) Art neural networks for remote sensing: Vegetation classification from landsat tm and terrain data. *IEEE Transactions on Geoscience and Remote Sensing,* 35, 308-325.

Caselli, M., Trizio, L., De Gennaro, G. & Ielpo, P. (2009) A simple feedforward neural network for the pm10 forecasting: Comparison with a radial basis function network and a multivariate linear regression model. *Water Air and Soil Pollution,* 201, 365-377.

Chen, P. H., Fan, R. E. & Lin, C. J. (2006) A study on smo-type decomposition methods for support vector machines. *Ieee Transactions on Neural Networks,* 17, 893-908.

Chevallier, F., Cheruy, F., Scott, N. A. & Chedin, A. (1998) A neural network approach for a fast and accurate computation of a longwave radiative budget. *Journal of Applied Meteorology,* 37, 1385-1397.

Comrie, A. C. (1997) Comparing neural networks and regression models for ozone forecasting. *Journal of the Air & Waste Management Association,* 47, 653-663.

Dufour, G., Nassar, R., Boone, C. D., Skelton, R., Walker, K. A., Bernath, P. F., Rinsland, C. P., Semeniuk, K., Jin, J. J., Mcconnell, J. C. & Manney, G. L. (2006) Partitioning between the inorganic chlorine reservoirs HCl and ClONO2 during the arctic winter 2005 from the ace-fts. *Atmospheric Chemistry and Physics,* 6, 2355-2366.

Eklundh, L. & Olsson, L. (2003) Vegetation index trends for the african sahel 1982-1999. *Geophysical Research Letters,* 30.

Eyring, V., Butchart, N., Waugh, D. W., Akiyoshi, H., Austin, J., Bekki, S., Bodeker, G. E., Boville, B. A., Bruhl, C., Chipperfield, M. P., Cordero, E., Dameris, M., Deushi, M., Fioletov, V. E., Frith, S. M., Garcia, R. R., Gettelman, A., Giorgetta, M. A., Grewe, V., Jourdain, L., Kinnison, D. E., Mancini, E., Manzini, E., Marchand, M., Marsh, D. R., Nagashima, T., Newman, P. A., Nielsen, J. E., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Schraner, M., Shepherd, T. G., Shibata, K., Stolarski, R. S., Struthers, H., Tian, W. & Yoshiki, M. (2006) Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past. *Journal of Geophysical Research-Atmospheres,* 111.

Eyring, V., Waugh, D. W., Bodeker, G. E., Cordero, E., Akiyoshi, H., Austin, J., Beagley, S. R., Boville, B. A., Braesicke, P., Bruhl, C., Butchart, N., Chipperfield, M. P., Dameris, M., Deckert, R., Deushi, M., Frith, S. M., Garcia, R. R., Gettelman, A., Giorgetta, M. A., Kinnison, D. E., Mancini, E., Manzini, E., Marsh, D. R., Matthes, S., Nagashima, T., Newman, P. A., Nielsen, J. E., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Schraner, M., Scinocca, J. F., Semeniuk, K., Shepherd, T. G., Shibata, K., Steil, B., Stolarski, R. S., Tian, W. & Yoshiki, M. (2007) Multimodel projections of stratospheric ozone in the 21st century. *Journal of Geophysical Research-Atmospheres,* 112.

Fan, R. E., Chen, P. H. & Lin, C. J. (2005) Working set selection using second order information for training support vector machines. *Journal of Machine Learning Research,* 6, 1889-1918.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Fuller, R. A., Marcy, T. P., Popp, P. J., Gao, R. S., Fahey, D. W., Jucks, K. W., Stachnik, R. A., Toon, G. C., Christensen, L. E., Webster, C. R., Bernath, P. F., Boone, C. D., Walker,

K. A., Pumphrey, H. C., Harwood, R. S., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P. & Wagner, P. A. (2008) Validation of Aura microwave limb sounder HCl measurements. *Journal of Geophysical Research-Atmospheres,* 113.

Froidevaux, L., Livesey, N. J., Read, W. G., Jiang, Y. B. B., Jimenez, C., Filipiak, M. J., Schwartz, M. J., Santee, M. L., Pumphrey, H. C., Jiang, J. H., Wu, D. L., Manney, G. L., Drouin, B. J., Waters, J. W., Fetzer, E. J., Bernath, P. F., Boone, C. D., Walker, K. A., Jucks, K. W., Toon, G. C., Margitan, J. J., Sen, B., Webster, C. R., Christensen, L. E., Elkins, J. W., Atlas, E., Lueb, R. A. & Hendershot, R. (2006a) Early validation analyses of atmospheric profiles from eos MLS on the Aura satellite. *IEEE Transactions on Geoscience and Remote Sensing,* 44, 1106-1121.

Froidevaux, L., Livesey, N. J., Read, W. G., Salawitch, R. J., Waters, J. W., Drouin, B., Mackenzie, I. A., Pumphrey, H. C., Bernath, P., Boone, C., Nassar, R., Montzka, S., Elkins, J., Cunnold, D. & Waugh, D. (2006b) Temporal decrease in upper atmospheric chlorine. *Geophysical Research Letters,* 33.

Gallo, K. P., Ji, L., Reed, B. C., Dwyer, J. & Eidenshink, J. C. (2003) Comparison of MODIS and AVHRR 16-day normalized difference vegetation index composite data. *Geophysical Research Letters,* 31, L07502-5.

Gardner, M. W. & Dorling, S. R. (1999) Neural network modelling and prediction of hourly NOx and NO2 concentrations in urban air in london. *Atmospheric Environment,* 33, 709-719.

Gunson, M. R., Abrams, M. C., Lowes, L. L., Mahieu, E., Zander, R., Rinsland, C. P., Ko, M. K. W., Sze, N. D. & Weisenstein, D. K. (1994) Increase in levels of stratospheric chlorine and fluorine loading between 1985 and 1992. *Geophysical Research Letters,* 21, 2223-2226.

Haykin, S. (2001a) *Kalman filtering and neural networks*, Wiley-Interscience.

Haykin, S. S. (1994) *Neural networks : A comprehensive foundation,* New York, Toronto, Macmillan.

Haykin, S. S. (2001b) *Kalman filtering and neural networks,* New York, Wiley.

Haykin, S. S. (2007) *New directions in statistical signal processing : From systems to brain,* Cambridge, Mass., MIT Press.

Hyyppa, J., Hyyppa, H., Inkinen, M., Engdahl, M., Linko, S. & Zhu, Y. H. (1998) Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management.* Lake Buena Vista, Florida.

Lary, D. J. & Aulov, O. (2008) Space-based measurements of HCl: Intercomparison and historical context. *Journal of Geophysical Research-Atmospheres,* 113.

Lary, D. J., Muller, M. D. & Mussa, H. Y. (2004) Using neural networks to describe tracer correlations. *Atmospheric Chemistry and Physics,* 4, 143-146.

Lary, D. J., Remer, L., Paradise, S., Macneill, D. & Roscoe, B. (2009) Machine learning and bias correction of MODIS aerosol optical depth. *IEEE Trans. on Geoscience and Remote Sensing*

Lary, D. J., Waugh, D. W., Douglass, A. R., Stolarski, R. S., Newman, P. A. & Mussa, H. (2007) Variations in stratospheric inorganic chlorine between 1991 and 2006. *Geophysical Research Letters,* 34.

Levenberg, K. (1944) A method for the solution of certain problems in least squares. *Quart. Appl. Math.,* 2, 164-168.

Los, S. O. (1998) Estimation of the ratio of sensor degradation between noaa AVHRR channels 1 and 2 from monthly NDVI composites. *IEEE Transactions on Geoscience and Remote Sensing,* 36, 206-213.

Marquardt, D. W. (1963) An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics,* 11, 431-441.

Marquardt, D. W. (1979) Citation classic - algorithm for least-squares estimation of non-linear parameters. *Current Contents/Engineering Technology & Applied Sciences*, 14-14.

Mchugh, M., Magill, B., Walker, K. A., Boone, C. D., Bernath, P. F. & Russell, J. M. (2005) Comparison of atmospheric retrievals from ace and HALOE. *Geophysical Research Letters,* 32.

Michelsen, H. A., Salawitch, R. J., Gunson, M. R., Aellig, C., Kampfer, N., Abbas, M. M., Abrams, M. C., Brown, T. L., Chang, A. Y., Goldman, A., Irion, F. W., Newchurch, M. J., Rinsland, C. P., Stiller, G. P. & Zander, R. (1996) Stratospheric chlorine partitioning: Constraints from shuttle-borne measurements of [HCl], [clno3], and [ClO]. *Geophysical Research Letters,* 23, 2361-2364.

Miura, T., Huete, A. & Yoshioka, H. (2006) An empirical investigation of cross-sensor relationships of NDVI and red/near-infrared reflectance using eo-1 hyperion data. *Remote Sensing of Environment,* 100, 223-236.

Moré, J. J. (1977) The levenberg-marquardt algorithm: Implementation and theory. IN Watson, G. A. (Ed.) *Numerical analysis*. Springer Verlag.

Nassar, R., Bernath, P. F., Boone, C. D., Clerbaux, C., Coheur, P. F., Dufour, G., Froidevaux, L., Mahieu, E., Mcconnell, J. C., Mcleod, S. D., Murtagh, D. P., Rinsland, C. P., Semeniuk, K., Skelton, R., Walker, K. A. & Zander, R. (2006) A global inventory of stratospheric chlorine in 2004. *Journal of Geophysical Research-Atmospheres,* 111.

Newman, P. A., Nash, E. R., Kawa, S. R., Montzka, S. A. & Schauffler, S. M. (2006) When will the antarctic ozone hole recover? *Geophysical Research Letters,* 33.

Rinsland, C. P., Gunson, M. R., Salawitch, R. J., Michelsen, H. A., Zander, R., Newchurch, M. J., Abbas, M. M., Abrams, M. C., Manney, G. L., Chang, A. Y., Irion, F. W., Goldman, A. & Mahieu, E. (1996) ATMOS/atlas-3 measurements of stratospheric chlorine and reactive nitrogen partitioning inside and outside the november 1994 antarctic vortex. *Geophysical Research Letters,* 23, 2365-2368.

Russell, J. M., Deaver, L. E., Luo, M. Z., Park, J. H., Gordley, L. L., Tuck, A. F., Toon, G. C., Gunson, M. R., Traub, W. A., Johnson, D. G., Jucks, K. W., Murcray, D. G., Zander, R., Nolt, I. G. & Webster, C. R. (1996) Validation of hydrogen chloride measurements made by the halogen occultation experiment from the UARS platform. *Journal of Geophysical Research-Atmospheres,* 101, 10151-10162.

Santee, M. L., Mackenzie, I. A., Manney, G. L., Chipperfield, M. P., Bernath, P. F., Walker, K. A., Boone, C. D., Froidevaux, L., Livesey, N. J. & Waters, J. W. (2008) A study of stratospheric chlorine partitioning based on new satellite measurements and modeling. *Journal of Geophysical Research-Atmospheres,* 113.

Scholkopf, B., Smola, A. J., Williamson, R. C. & Bartlett, P. L. (2000) New support vector algorithms. *Neural Computation,* 12, 1207-1245.

Sen, B., Osterman, G. B., Salawitch, R. J., Toon, G. C., Margitan, J. J., Blavier, J. F., Chang, A. Y., May, R. D., Webster, C. R., Stimpfle, R. M., Bonne, G. P., Voss, P. B., Perkins, K.

K., Anderson, J. G., Cohen, R. C., Elkins, J. W., Dutton, G. S., Hurst, D. F., Romashkin, P. A., Atlas, E. L., Schauffler, S. M. & Loewenstein, M. (1999) The budget and partitioning of stratospheric chlorine during the 1997 arctic summer. *Journal of Geophysical Research-Atmospheres,* 104, 26653-26665.

Slayback, D. A., Pinzon, J. E., Los, S. O. & Tucker, C. J. (2003) Northern hemisphere photosynthetic trends 1982-99. *Global Change Biology,* 9, 1-15.

Smola, A. J. & Scholkopf, B. (2004) A tutorial on support vector regression. *Statistics and Computing,* 14, 199-222.

Solomon, S., Intergovernmental Panel on Climate Change. & Intergovernmental Panel on Climate Change. Working Group I. (2007) *Climate change 2007 : The physical science basis : Contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change,* Cambridge ; New York, Cambridge University Press.

Steven, M. D., Malthus, T. J., Baret, F., Xu, H. & Chopping, M. J. (2003) Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment,* 88, 412-422.

Teillet, M., Staenz, K. & Williams, D. J. (1997) Effects of spectral, spatial and radiometric characteristics on remote sensing vegetation indices of forested regions. *Remote Sensing of Environment,* 61, 139-149.

Trishchenko, A. P., Cihlar, J. & Li, Z. (2002) Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. *Remote Sensing of Environment,* 81, 1-18.

Tucker, C. J., Slayback, D. A., Pinzon, J. E., Los, S. O., Myneni, R. B. & Taylor, M. G. (2001) Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999. *International Journal of Biometeorology,* 45, 184-190.

Van Leeuwen, W., Orr, B. J., Marsh, S. E. & Herrmann, S. M. (2006) Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications. *Remote Sensing of Environment,* 100, 67-81.

Vapnik, V. N. (1995) *The nature of statistical learning theory,* New York, Springer.

Vapnik, V. N. (1998) *Statistical learning theory,* New York, Wiley.

Vapnik, V. N. (2000) *The nature of statistical learning theory,* New York, Springer.

Voss, P. B., Stimpfle, R. M., Cohen, R. C., Hanisco, T. F., Bonne, G. P., Perkins, K. K., Lanzendorf, E. J., Anderson, J. G., Salawitch, R. J., Webster, C. R., Scott, D. C., May, R. D., Wennberg, P. O., Newman, P. A., Lait, L. R., Elkins, J. W. & Bui, T. P. (2001) Inorganic chlorine partitioning in the summer lower stratosphere: Modeled and measured [ClONO2]/[HCl] during polaris. *Journal of Geophysical Research-Atmospheres,* 106, 1713-1732.

Waugh, D. W. & Eyring, V. (2008) Quantitative performance metrics for stratospheric-resolving chemistry-climate models. *Atmospheric Chemistry and Physics,* 8, 5699-5713.

Webster, C. R., May, R. D., Jaegle, L., Hu, H., Sander, S. P., Gunson, M. R., Toon, G. C., Russell, J. M., Stimpfle, R. M., Koplow, J. P., Salawitch, R. J. & Michelsen, H. A. (1994) Hydrochloric-acid and the chlorine budget of the lower stratosphere. *Geophysical Research Letters,* 21, 2575-2578.

Wmo (2006) Scientific assessment of ozone depletion: 2006. WMO Global Ozone Res. and Monitor. Proj., Geneva.

Wood, G. A. R. (2001) *CACAO,* 620.

Yi, J. S. & Prybutok, V. R. (1996) A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environmental Pollution,* 92, 349-357.

Zander, R., Gunson, M. R., Farmer, C. B., Rinsland, C. P., Irion, F. W. & Mahieu, E. (1992) The 1985 chlorine and fluorine inventories in the stratosphere based on ATMOS observations at 30-degrees north latitude. *Journal of Atmospheric Chemistry,* 15, 171-186.

Zander, R., Mahieu, E., Gunson, M. R., Abrams, M. C., Chang, A. Y., Abbas, M., Aellig, C., Engel, A., Goldman, A., Irion, F. W., Kampfer, N., Michelsen, H. A., Newchurch, M. J., Rinsland, C. P., Salawitch, R. J., Stiller, G. P. & Toon, G. C. (1996) The 1994 northern midlatitude budget of stratospheric chlorine derived from ATMOS/atlas-3 observations. *Geophysical Research Letters,* 23, 2357-2360.

# 2

# Building the Next Generation of Aerospace Data Processing Systems by Reusing Existing Software Components

James J. Marshall[1], Robert R. Downs[2], and Shahin Samadi[1]
[1]*Innovim / NASA Goddard Space Flight Center,*
[2]*CIESIN, Columbia University*
*USA*

## 1. Introduction

Software is a key ingredient when developing any aerospace system. It is used in embedded electronics, in flight dynamics, in ground and space data processing, and in the current generation of data products. For example, the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center's Innovative Partnerships Program Office offers licensing opportunities for software and technologies from a variety of areas relevant to the hardware and software requirements of Earth and space science missions and projects: aerospace/aeronautics, computer software, data processing/analysis, electromechanical devices, electronics, manufacturing equipment, mechanical technologies, nanotechnology, optics and photonics, sensor and detector technologies, subassemblies and components, telecommunications and internet, and test and measurement (IPP Office, 2009). Reuse of existing experience and artifacts eliminates having to "reinvent the wheel" and is a key element to achieving progress in many areas of complex aerospace system development.

Originally, in the absence of vendor-provided solutions and commercial off-the-shelf software components, many data and information systems were designed and built as custom applications. However, as the practice of systems and applications development has matured, facilitating reuse of software and reusing previously developed software have been recognized as beneficial for meeting the challenges of developing and maintaining complex systems. Some of the challenges commonly faced by system developers can include dealing with very large quantities of data (e.g., terabytes per day), working with a distributed knowledge base, the expense and complexity of required technology infrastructure, and the need for domain-specific knowledge in software development (Samadi et al., 2007). In software development, reuse can assist today's development teams in various aspects of the system development life cycle, especially when they share common goals (Samadi et al., 2007).

The development of new systems can benefit from the efforts that contributed to the development of current and previous generations of systems. Considering the costs of building new systems, and the learning curve that contributes to such costs, leveraging the results of previous system development activities has the potential to reduce system

development costs and improve capabilities for new development. Previously developed system components, such as plans, design requirements, system documentation, algorithms, and procedures, offer the results of previous design, testing activities, scientific algorithms, and learning experiences that were part of the initial systems development effort. Such legacy resources contain the evolved expertise of previous generations and can have continuing value if reused when building the next generation of systems or enhancing existing systems.

While the potential gains from software reuse appear quite promising, these gains also can be accompanied by costs and risks. Effectively preparing software components and other artifacts for potential reuse requires efforts to ensure that such artifacts can be reused in a manner that offers adopters efficiency gains that can be realized through reuse. Without such preparation efforts, adopters might not achieve the efficiency benefits anticipated for their reuse activities. However, the costs of preparing software for reuse could be small in comparison to the potential gains to be attained from reuse. For example, in cases where a planned or "forward-looking" approach to software reuse has been employed in the aerospace industry, improvements have been observed in terms of increasing the quality and reducing the costs of software reuse (Finnigan & Blanchette, 2008).

Similarly, risks can be incurred through software reuse. The adoption of previously-developed software can pose risks for those who contribute software as well as for those who adopt software. For example, assumptions about the origination, validation, and verification of software can increase risks to software projects, especially if the software affects mission-critical operations or safety-critical systems (Orrego & Mundy, 2007). While costs can be controlled and risks can be mitigated, they need to be recognized and assessed to determine their potential effects on software reuse. Recognizing the existence of the possible costs and risks also enables the involved parties to engage in software reuse with grounded expectations. Taking a deliberate, planned, and systematic approach to software reuse will help producers and adopters of reusable software to attain the benefits of reuse while limiting its costs and managing its risks.

Reuse has been shown to be effective in developing large and complex data systems to support space-based missions. One example is the National Polar-orbiting Operational Environmental Satellite System Preparatory Project where high levels of reuse have enabled system developers in the Science Data Segment, which provides ground systems that support data product assessment and quality verification, to reduce development effort as described by Samadi et al. (2007). Software reuse also contributes to flight software, such as the instrument manager software for the Mars Science Laboratory rover (Murray et al., 2009). This chapter describes how reuse activities can fit into building the next generation of aerospace data processing systems by providing guidance on methods for improving reuse practices in order to realize the benefits of reuse.

## 2. Background on reuse

The reuse of existing software and related artifacts is commonly referred to as software reuse. More specifically, the reuse of a software artifact is its integration into another context (Leyton, 2008). The reuse of software has been practiced informally and formally by software developers, since the practice offers potential benefits for developers and development organizations to attain improvements in productivity and software quality (Ajila & Wu, 2007; Mohagheghi & Conradi, 2007; Nazareth & Rothenberger, 2004).

Decreases in software product development time and software development costs also have been reported (Frakes & Kang, 2005). In addition, systematic software reuse could provide additional benefits if the organization establishes incentives to motivate developers (Sherif et al., 2006).

The results of surveying members of the Earth science community on their reuse practices and experiences showed that the reuse of existing technologies is most often performed in order to save time, save money, and ensure the reliability of the product (Marshall et al., 2006). When measuring potential gains in productivity, quality, and time-to-market from software reuse, metrics must be analyzed from the entire software development life cycle as well as from software maintenance efforts (Mascena et al., 2005). If such gains can be realized, then efforts to reuse software code and other artifacts should be pursued by software developers and their employing organizations. Likewise, such efforts should be documented to identify and measure actual benefits attained. Furthermore, measuring the benefits of software reuse can assist in benchmarking the activities that led to the measured gains and enable the identification of additional gains that can be compared to such benchmarks.

In addition to the potential benefits previously described, reliability is another objective of software reuse (Frakes & Kang, 2005). NASA recognizes the value of reuse in providing "a lower likelihood that the item will fail tests and generate rework" (NASA, 2007). Anticipating potential uses of software artifacts should help to reduce failures of reusing such artifacts within new domains and operating environments (Suri & Garg, 2008). In addition, efforts to ensure the reliability of software for reuse within new environments should begin during the early stages of the software development process (Immonen & Niemela, 2008).

Various software artifacts, in addition to source code, can be reused to the benefit of software developers and their organizations. Some reusable software components are listed in Table 1 (derived from Samadi et al., 2007, with additions).

| **Reusable Software Components** |
|---|
| Operational source code |
| Analysis and design specifications |
| Plans (project management) |
| Data (testing) |
| Synthetic data generators and analysis tools |
| Expertise/experience (life cycle model, quality assurance) |
| Information used to create software and documentation |
| Testing procedures |
| Documentation |
| System and user guides |
| Use cases and user scenarios |

Table 1. Reuseable Software Components

In order to make use of reusable components, information about "how" they work must be included along with information about "what" they do (Neighbors, 1992). Additional work

products, such as comments and other artifacts, also can contribute to the reusability of source code and are often needed to migrate software components between platforms, contexts, and projects. In many cases, documentation, such as user guides, installation guides, and other specifications, also is essential to facilitate the reusability of software components.

Establishing reusability as a requirement for the development of software components can contribute to their potential for reuse. However, even in cases where reusability was not planned or originally intended during initial software development, an artifact may still be reused in domains and contexts similar to the original. In such cases, domain analysis may be required to facilitate reuse, even when the reuse is within an application or problem domain that is similar to the original. Domain analysis includes identifying, capturing, and organizing information that was used in developing the original system to facilitate its potential reuse within a new system (Preito-Diaz, 1990), enabling subsequent evolution of the original information infrastructure.

## 3. Adoption process for reusable components

Reusing previously developed components within a new system requires a directed effort. In some cases, a top-down and single-project life cycle model such as the "waterfall" model might not be appropriate for software reuse (Simos, 1988). Software reuse is not inherently top-down, and it needs a perspective which may be beyond the development activities of a single project or program. The reuse process needs to follow a structured approach such as the approach recommended by Basili et al. (1989):

1. Specifying the object to be created
2. Searching the project, domain, and general databases for reuse candidates
3. Evaluating the candidates to determine which (if any) should be used
4. Modifying, if necessary, to fit specific needs
5. Integrating the reusable components
6. Validating
7. Feeding back the knowledge regarding the payoff of reuse

Initially, potentially reusable components need to be identified through domain analysis, and the candidates for reuse need to be assessed in terms of how they might be reused. For example, a categorized list of sub-routines might inform current development activities even if the subroutines are not considered to be reusable. The extent of reuse could depend on common aspects between the original functionality of the reusable components and the planned purposes for their potential reuse. Also, developers need to determine whether the candidate reuse would be cost-effective when assessing how much preparation is required for potential reuse. The more mature the software is, in a reusability sense, the more cost-effective its reuse is likely to be. Less mature software is likely to take more time, cost, effort, etc. than building the software from scratch, and is less likely to provide cost-effective solutions.

### 3.1 Locating reusable components

In addition to identifying candidates for reuse from current and legacy systems, other systems, either within the same organization or available externally, also should be considered when seeking reusable components. Developers from various fields are

contributing to open source collections of software that are available for reuse by other adopters. SourceForge (2009) is one example of such a collection, and the site covers essentially all domains. Other collections may be focused on one or more particular domains, such as the Netlib for mathematics (2009) and the Space Telescope Science Institute's list of software and hardware products for astronomy (2009).

There are also collections for particular programming languages, tools/formats, and applications; some examples of these are the Comprehensive Perl Archive Network (2009), the software available at the HDF-EOS Tools and Information Center (2009), and the Comprehensive TeX Archive Network (2009). While adopters are most likely to find reusable components for their particular domain, programming language, tool/format, or application in these more specific collections, they should check the more general collections as well, since some components may be suitable candidates for reuse even if they were originally developed for a different domain.

Often, these open source resources are offered with non-restrictive licenses that allow and encourage others to reuse the components. Open source communities also encourage developers to share improvements to open source resources so that the resources and the adopters will benefit from enhancements that have been developed and offered by others.

### 3.2 Assessing reusable components

Various tools are available for systems developers when assessing the potential for reusing legacy resources in the development of a new system. One example is reuse readiness levels (Marshall & Downs, 2008), which can be used to assess the potential for existing software being considered for reuse. Reuse Readiness Levels (RRLs) also can be used to assess the value of software that is being developed and to determine areas where improvements could contribute to the reusability of the software, particularly through the use of the topic area levels from which the overall RRLs were developed. A brief introduction to this method of assessing the reusability maturity of software components follows.

As described in Marshall & Downs (2008), the RRLs were developed by members of the Earth science data systems software development community through an iterative process. Initially, the needs of both software contributors and software adopters were considered in order to identify topic areas that could be important for measuring the reuse maturity of software components. Nine such areas were identified; alphabetically, they are documentation, extensibility, intellectual property issues, modularity, packaging, portability, standards compliance, support, and verification/testing. Nine teams, each consisting of at least two people who have been involved with software development, worked together to write a set of levels for these nine topic areas; descriptions of the topic level areas can be found in Marshall et al. (2008). Next, members of these teams looked across all topic areas at each level to draft summaries of the overall RRLs, combining key information from all topics at the same level. Based on feedback received from other members of the community, the overall RRLs and topic area levels were revised and improved. This process of revising the levels and obtaining additional feedback to use in further revisions is repeated as the community continues to improve the RRLs. A summary of the RRLs is presented in Table 2.

Additional details on the RRLs and the topic area levels from which they were developed can be found in materials produced by the NASA Earth Science Data Systems Software Reuse Working Group, many of which are publicly available on the group's web site (2009).

| RRL | Summary |
|---|---|
| 1 | Limited reusability; the software is not recommended for reuse. |
| 2 | Initial reusability; software reuse is not practical. |
| 3 | Basic reusability; the software might be reusable by skilled users at substantial effort, cost, and risk. |
| 4 | Reuse is possible; the software might be reused by most users with some effort, cost, and risk. |
| 5 | Reuse is practical; the software could be reused by most users with reasonable cost and risk. |
| 6 | Software is reusable; the software can be reused by most users although there may be some cost and risk. |
| 7 | Software is highly reusable; the software can be reused by most users with minimum cost and risk. |
| 8 | Demonstrated local reusability; the software has been reused by multiple users. |
| 9 | Demonstrated extensive reusability; the software is being reused by many classes of users over a wide range of systems. |

Table 2. Reuse Readiness Levels

The Earth science data systems development community also has identified a number of potential uses for the RRLs. As metadata for reusable software components stored in catalogs and repositories, RRLs provide a guide to software adopters. They can help adopters quickly asses the maturity of candidate components for their reuse efforts, narrowing down the number of possible solutions they must consider in detail. The RRLs, and the RRL topic area levels in particular, can serve as an indicator of areas to focus on when creating reusable components, as a guide to contributors. Each topic area was selected because it was identified as an important factor that contributes to the reusability of software. By assessing their software components in each of the topic areas, contributors can identify the strengths and weaknesses of their components and work to improve the reusability of the components by using the levels as a guide. The RRLs are being evaluated and work is in progress to develop specific use cases for the RRLs, for both software contributors and software adopters.

It also has been recognized that RRLs could eventually become part of requests for proposals or contracts, which require a reuse approach or explanation of how components are being made reusable. Projects could undertake reusability improvement efforts, indicating that software that begins at one RRL will be developed to and released at a higher RRL. By maturing the reusability of the software, the chances of it being reused would be higher, enabling more projects to benefit from previous efforts. Projects that involve new software development could propose to make their resulting work available for reuse and indicate the planned RRL that would be targeted for release of the software. Projects that will be reusing software could indicate the RRLs of the component(s) being considered for reuse and demonstrate how this reuse provides benefits to the proposed work (e.g., by reducing development time and costs).

Tools such as the RRLs can be very helpful in both creating and adopting reusable software components. Developers are advised to take advantage of such tools when possible.

## 4. Building components for reuse

Creating reusable software and other components also offers benefits for system developers and their employing organizations. In addition to the benefits of reusing software, system developers also need to recognize the potential value that their current system development activities could have for future system development efforts. Besides contributing to the new system that is being developed, the results of current system development projects also can contribute to the development of future systems that are created by the same organization or by other organizations. Considering the potential value of reusable components for future systems, building software components so that they can be reused could be economically beneficial for organizations involved in software development. There is evidence that reuse offers cost reduction, cost avoidance, and increased profit even though there can be additional costs involved in developing software for reuse (Lim, 1998). Furthermore, building software components to be reusable should be considered good business practice (Stephens, 2006), enabling organizations to reduce redundancy, avoid increasing maintenance costs, and meet evolving requirements.

### 4.1 Motivations for building reusable components

Other factors also motivate individual developers to create reusable software components, either for reuse within the same organization that originally created the software or for reuse by others. Open source software developers can be motivated both intrinsically, for altruistic reasons, and extrinsically, for potential gains (Hars & Ou, 2002; Wu et al., 2007). Such motivations can be complementary for individual software developers (Roberts et al., 2006) and also could evolve (Shah, 2006). However, extrinsic factors, such as improving reputation and self-development, could be more motivating than intrinsic factors, such as altruism (Oreg & Nov, 2008). Factors motivating an individual developer's involvement in open source Internet communications software development projects include personal software needs, expectations of skill development, reputation enhancement, and enjoyment (Xu et al., 2009). Community factors such as leadership effectiveness and interpersonal relationships also were found to be motivators for involvement in these software development projects (Xu et al., 2009).

While systems development technology and available resources and tools can be expected to change, the purposes for developing future systems could have many aspects in common with the systems that are being developed today. In addition, as techniques and capabilities for developing reusable components improve, future systems could gain even more from current software development efforts. Software development organizations also could contribute to the motivations of their employees to engage in software reuse activities by providing the resources to create an infrastructure that facilitates software reuse (Sherif et al., 2006). Commitment to reuse by the top management of the organization also could be a contributing factor (Morisio et al., 2002).

### 4.2 Planning and developing reusable components

Planning is necessary when developing systems components for potential reuse. In some cases, initial efforts to ensure the reusability of system components can offer future benefits without adding to development costs. If additional development costs will be required to produce reusable components, these must be weighed against the potential benefits and cost

savings of reusing the components in other projects in the future. If a small additional expenditure on the current project can save a larger amount of money in future projects, it would be justified. However, if the organization developing the reusable components is not expected to be the one to reap the benefits of reusing the components in the future, it may be more difficult to justify the additional cost of preparing components for future reuse. Engaging in a planned, systematic approach to software reuse, including the identification of its benefits, can reduce the potential risks of software reuse (Frakes & Isoda, 1994).

Certainly, many of the activities that foster reusability, such as effectively documenting the work, should be conducted routinely during system development even if reuse is not planned to ensure that enhancements can be completed efficiently. While seemingly trivial and sometimes overlooked by software developers, adding comments to source code can assist in its subsequent analysis and reuse. Furthermore, the inclusion of comments was found to be a success factor for module design that contributed to reuse in large-scale systems at NASA (Selby, 2005).

Developing to standards also can facilitate reuse by others. By developing to established standards for software (see, for example, Baldo et al., 1997 and IEEE, 2004) and data (Baxter et al., 2006), reuse of both software and data can be fostered.

Furthermore, establishing standards for system components enables others to develop to such standards and potentially foster reuse of components that meet the standards. For example, NASA has proposed an open architecture standard for software-defined radio systems that could facilitate the reuse of design expertise as well as software code, offering potential gains in efficiency for future development and modification efforts (Reinhart et al., 2008).

Independent of whether components have been developed to meet established standards, it is necessary to provide sufficient information about resources available for reuse so that they can be identified and assessed for adoption by potential adopters. In addition, enabling reuse also requires developers to grant appropriate intellectual property rights to eliminate legal barriers to reuse, as described in the following section on contributing components for reuse.

## 5. Contributing components for reuse

Contributing reusable system components to open source software collections and catalogs can enable others to reuse such components if they are licensed as shareable resources that foster reuse by others. Enabling others to reuse system components can contribute to the pool of resources available for all members of the open source community and result in further improvements to these resources by others. If there are barriers to adoption, it will be more difficult for potential adopters to reuse software and other available components. Such barriers can appear at the individual level or the organizational level, and the interaction of barriers at the organizational level can create barriers at the individual level (Sherif & Vinze, 2003). It is important for software developers to consider the potential barriers that exist at all levels and to take the steps that are necessary to address such barriers in order for their software components and related artifacts to be reused more easily by potential adopters.

### 5.1 Authorizing reuse

Potential adopters of reusable components need to be authorized to reuse components that they identify for potential adoption. Without the legal rights to adopt available components,

it would be imprudent for potential adopters to incur potential risks by adopting components that they have not been authorized to use. Furthermore, considering the effort involved in adopting reusable components and integrating such components into systems, such investments would not be recommended without first ensuring that the rights for reuse have been obtained. Given these considerations, it is necessary for contributors of components to grant the intellectual property rights necessary for others to adopt components that have been developed for reuse. In addition, it also is necessary to adequately communicate information about the rights for reuse so that potential adopters can easily determine whether they have the right to adopt a particular component and determine any conditions for reuse that might be associated with the adoption.
Several licensing options are available for system and software developers to authorize others to utilize components that have been developed for reuse. One such option is open source, for which there are a number of available licenses; see, for example, Open Source Initiative (2009) for a list. However, prior to considering the available choices for granting rights for reuse, developers should identify any policies or rules within their own organization that might govern their choices. Some organizations have policies that prevent granting rights to others. In such cases, a waiver might have to be obtained from the appropriate authority within the organization so that authorization for reuse can be granted. Organizations also might be concerned about liability and could require review to ensure that granting rights for reuse does not expose the organization to such liability. In some cases, organizations that routinely create reusable components might have stock licensing statements available for developers to offer components for reuse. Such licensing and liability statements should be reviewed by an attorney.

### 5.2 Propagating reuse

Adopters of reusable software should consider sharing any enhancements that they have made to reusable software so that others also may benefit from their improvements. The changes that one organization finds useful might also be useful for other organizations. In addition, other software developers could evaluate the contributed changes and improve them even further. Likewise, such improvements might be considered valuable by the original adopters who had shared their enhancement. In such cases, the original adopters might decide to adopt the new version that includes the improvements to their original enhancements. These kinds of scenarios could be compounded as the community that adopts and contributes enhancements, for a particular reusable resource, continues to grow. Such cycles of individual contributions can improve libraries of software components, offering benefits to the software development community that participates in the software reuse process.

## 6. Summary

The reuse of software components and other artifacts offers potential benefits for software developers of aerospace systems and other systems. However, a planned and systematic approach to software reuse is recommended for software developers and software development organizations to realize the benefits of software reuse while minimizing potential costs and risks of reuse. The tools and techniques for software reuse that have been described offer options for developing the capabilities to engage in software reuse as

contributors and adopters of reusable software and related artifacts. In addition, suggestions for improving software reuse practices have been offered for software contributors and for software adopters.

## 7. Acknowledgements

The authors appreciate the contributions of the current and previous members of the National Aeronautics and Space Administration (NASA) Earth Science Data Systems Software Reuse Working Group to some of the work presented here, and very much appreciate the support received from the NASA for the work reported in this paper, including the support for Robert Downs under Contract NAS5-03117.

## 8. References

Ajila, S.A. & Wu, D. (2007) Empirical study of the effects of open source adoption on software development economics. *Journal of Systems and Software,* 80, 9, (September 2007) 1517–1529, ISSN 0164-1212

Baldo, J.; Moore, J. & Rine, D. (1997). Software reuse standards. *StandardView,* 5, 2, (June 1997) 50–57, ISSN 1067-9936

Basili, V.R.; Rombach, H.D.; Bailey, J.; Delis, A. & Farhat, F. (1989). Ada Reuse Metrics, In: *Guidelines Document for Ada Reuse and Metrics*, P. A. Lesslie, R. O. Chester, and M. F. Theofaanos (Eds.), 11–29, Martin Marietta Energy Systems, Inc., Oak Ridge, Tennessee, under contract to U.S. Army, AIRMICS

Baxter, S.M.; Day, S.W.; Fetrow, J.S. & Reisinger, S.J. (2006). Scientific Software Development Is Not an Oxymoron. *PLoS Computational Biology,* 2, 9, e87, ISSN 1553-734X (Print) 1553-7358 (Online)

Comprehensive Perl Archive Network (2009). http://www.cpan.org/

Comprehensive TeX Archive Network (2009). TeX Users Group, http://www.ctan.org/

Earth Science Data Systems Software Reuse Working Group (2009). http://www.esdswg.com/ softwarereuse/

Finnigan, J.V. & Blanchette, J. (2008). A Forward-Looking Software Reuse Strategy. *Proceedings of the 2008 IEEE Aerospace Conference,* pp. 1–9, ISBN 978-1-4244-1487-1, Big Sky, Montana, March 2008, IEEE Press, New York

Frakes, W.B. & Isoda, S. (1994). Success factors of systematic reuse. *IEEE Software*, 11, 5, (September 1994) 14–19, ISSN 0740-7459

Frakes, W.B. & Kang, K. (2005). Software Reuse Research: Status and Future. *IEEE Transactions on Software Engineering*, 31, 7, (July 2005) 529–536, ISSN 0098-5589

Hars, A. & Ou, S. (2002). Working for Free? Motivations for Participating in Open-Source Projects, *International Journal of Electronic Commerce,* 6, 3, (Spring 2002), 25–39, ISSN 1086-4415

HDF-EOF Tools and Information Center (2009). The HDF Group, http://hdfeos.org/ software.php

Innovative Partnerships Program (IPP) Office (2009). Featured Technologies, http://ipp.gsfc.nasa.gov/ technologies.html

Institute of Electrical and Electronics Engineers, Inc. (2004). *IEEE Standard for Information Technology – Software Life Cycle Processes – Reuse Processe*s, IEEE Press, ISBN 0-7381-1735-8, New York

Immonen, A. & Niemela, E. (2008). Survey of reliability and availability prediction methods from the viewpoint of software architecture. *Software and Systems Modeling*, 7, 1 (February 2008) 49–65, ISSN 1619-1366 (Print) 1619-1374 (Online)

Leyton, M. (2008). Defining Reuse, presentation available online at: http://www.esdswg.com/softwarereuse/Resources/library/7th-esds-wg-meeting/ESDSWG7DefinitionOfReuse.pdf

Lim, W.C. (1998). *Managing Software Reuse: A Comprehensive Guide to Strategically Reengineering the Organization for Reusable Components,* Prentice Hall PTR, ISBN 0-13-552373-7, Upper Saddle River, NJ

Marshall, J.J. & Downs, R.R (2008). Reuse Readiness Levels as a Measure of Software Reusability, *Proceedings of 2008 Geoscience and Remote Sensing Symposium, Volume 3,* pp. III-1414–III-1417, ISBN 978-1-4244-2807-6, Boston, Massachusetts, July 2008, IEEE Press, New York

Marshall, J.J.; Downs, R.R.; Samadi, S.; Gerard, N.S.; Wolfe, R.E. (2008). Software Reuse to Support Earth Science. *Journal of Frontiers of Computer Science and Technology,* 2, 3 (June 2008), 296–310, ISSN 1673-9418

Marshall, J.J.; Olding, S.W.; Wolfe, R.E. & Delnore, V.E. (2006). Software Reuse Within the Earth Science Community, *Proceedings of 2006 IEEE International Conference on Geoscience and Remote Sensing Symposium,* pp. 2880–2883, ISBN 0-7803-9510-7, Denver, Colorado, July–August 2006, IEEE Press, New York

Mascena, J.C.C.P.; de Almeida, E.S. & de Lemos Meira, S.R. (2005). A Comparative Study on Software Reuse Metrics and Economic Models from a Traceability Perspective, *Proceedings of IEEE International Conference on Information Reuse and Integration*, pp. 72–77, ISBN 0-7803-9093-8, Las Vegas, August 2005, IEEE Press, New York

Mohagheghi, P. & Conradi, R. (2007). Quality, productivity and economic benefits of software reuse: a review of industrial studies. *Empirical Software Engineering*, 12, 471–516, ISSN 1382-3256 (Print) 1573-7616 (Online)

Morisio, M.; Ezran, M. & Tully, C. (2002). Success and Failure Factors in Software Reuse. *IEEE Transactions on Software Engineering*, 28, 4, (April 2002) 340–357, ISSN 0098-5589

Murray, A.; Schoppers, M. & Scandore, S. (2009). A Reusable Architectural Pattern for Auto-Generated Payload Management Flight Software, *Proceedings of the 2009 IEEE Aerospace Conference,* pp. 1–11, ISBN 978-1-4244-2621-8, Big Sky, Montana, March 2009, IEEE Press, New York

National Aeronautics and Space Administration (2007). *Systems Engineering Handbook,* SP-2007-6105 Rev1, Washington, D.C.

Nazareth, D.L. & Rothenberger, M.A. (2004). Assessing the cost-effectiveness of software reuse: a model for planned reuse. *Journal of Systems and Software,* 73, 2, (October 2004) 245–255, ISSN 0164-1212

Neighbors, J.M. (1992). The Evolution from Software Components to Domain Analysis. *International Journal of Software Engineering and Knowledge Engineering*, 2, 3, (September 1992) 325–354, ISSN 0218-1940

Netlib (2009). Netlib Repository at UTK and ORNL. http://www.netlib.org/

Open Source Initiative (2009). Open Source Licenses, http://www.opensource.org/licenses

Oreg, S. & Nov, O. (2007). Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values, *Computers in Human Behavior,* 24, 5, (September 2008) 2055–2073, ISSN 0747-5632

Orrego, A.S. & Mundy, G.E. (2007). A Study of Software Reuse in NASA Legacy Systems. *Innovations in Systems and Software Engineering,* 3, 3, (September 2007) 167–180, ISSN 1614-5046 (Print) 1614-5054 (Online)

Prieto-Diaz, R. (1990). Domain Analysis: An Introduction. *ACM Software Engineering Notes,* 15, 2, (April 1990) 6–16, ISSN 0163-5948

Reinhart, R.C.; Johnson, S.K.; Kacpura, T.J.; Hall, C.S.; Smith, C.R.; & Liebetreu, J. (2008). Open Architecture Standard for NASA's Software-Defined Space Telecommunications Radio Systems, NASA/TP—2008-214941, pp. 1–8, May 2008, National Aeronautics and Space Administration, Glenn Research Center, Cleveland, Ohio. http://gltrs.grc.nasa.gov/reports/2008/TP-2008-214941.pdf

Roberts, J.A.; Hann, I. & Slaughter, S.A. (2006). Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects, *Management Science,* 52, 7, (July 2006), 984–999, ISSN 0025-1909 (Print) 1526-5501 (Online)

Samadi, S.; Gerard, R.; Hunter, M.; Marshall, J.J.; Schweiss, R.J.; Wolfe, R.E. & Masuoka, E.J. (2007). Reusing Software to Build Data Processing Systems, *Proceedings of 2007 IEEE Aerospace Conference*, pp. 1–12, ISBN 1-4244-0525-4, Big Sky, Montana, March 2007, IEEE Press, New York

Selby, R.W. (2005). Enabling reuse-based software development of large-scale systems, *IEEE Transactions on Software Engineering,* 31, 6, (June 2005) 495–510, ISSN 0098-5589

Shah, S. K. (2006). Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development, *Management Science,* 52, 7, (July 2006), 1000–1014, ISSN 0025-1909 (Print) 1526-5501 (Online)

Sherif, K.; Appan, R. & Lin, Z. (2006). Resources and Incentives for the Adoption of Systematic Software Reuse. *International Journal of Information Management*, 26, 1, (February 2006) 70–80, ISSN 0268-4012

Sherif, K. & Vinze, A. (2003). Barriers to adoption of software reuse: A qualitative study, *Information & Management,* 41, 2, (December 2003) 159–175, ISSN 0378-7206

Simos, M.A. (1988). The Domain-Oriented Software Life Cycle: Towards an Extended Process Model for Reusability, In *Software Reuse: Emerging Technology,* Will Tracz (Ed.), 354–363, IEEE Computer Society Press, ISBN 0-8186-0846-3, Los Alamitos, California

SourceForge, Inc. (2009). SourceForge.net: Find and Develop Open-Source Software. http://sourceforge.net/

Space Telescope Science Institute (2009). Software and Hardware Products from STScI. Baltimore, Maryland, http://www.stsci.edu/resources/software_hardware

Stephens, K. (2006) Web Services and Asset Reuse: Implications for Developers. International Business Machines Corporation, Armonk, New York, http://www.ibm.com/developerworks/webservices/library/ws-assetreuse/

Suri, P. K. & Garg, N. (2008). Simulator for evaluating Reliability of Reusable Components in a Domain Interconnection Network. *International Journal of Computer Science and Network Security*, 8, 3, (March 2008) 251–259, ISSN 1738-7906

Wu, C.; Gerlach, J.H. & Young, C.E. (2007). An empirical analysis of open source software developers' motivations and continuance intentions, *Information & Management,* 44, 3, (April 2007) 253–262, ISSN 0378-7206

Xu, B.; Jones, D.R. & Shao, B. (2009). Volunteers' involvement in online community based software development, *Information & Management*, 46, 3, (April 2009) 151–158, ISSN 0378-7206

# 3

# Reconfigurable Computing for Space

Donohoe, Gregory W.[1] and Lyke, James C.[2]
[1]*University of Idaho*
[2]*Air Force Research Laboratory*
*USA*

## 1. Introduction

Reconfigurable computing is an emerging technology with important implications for space systems. Spacecraft serve as sensor platforms to gather, format, possibly screen or interpret, and then downlink sensor data. Once launched, a spacecraft may operate unattended for decades. The trend in spacecraft design is toward more complex and voluminous data, including multispectral and hyperspectral images and synthetic aperture radar, which require extensive on-board data processing. Other on-board computing tasks are navigation, orientation, and communication. As spacecraft grow more sophisticated, the need for higher-performance computing grows also.

Reconfigurable computing allows computing hardware to be configured or "wired" by software to optimize the architecture for the problem at hand. Some early discussions of reconfigurable computing are found in (DeHon, 1996), and Villasenor (Villasenor, 1997). Reconfigurable computers may be implemented on specially-designed hardware, or on Field Programmable Gate Arrays (FPGAs). Hauck and DeHon cover FPGA-based reconfigurable computing very thoroughly their book (Hauck, 2008).

The space environment presents special challenges to the system designer. Chief among these are:

1. Power efficiency. Electrical power is a scarce and expensive commodity for orbiters and deep space probes. In addition, removing heat generated by electronics is very challenging in space, where air cooling is not an option.
2. Minimum size and weight. Increased size and weight increase launch costs, and require more fuel for on-orbit maneuvering.
3. High reliability for long mission life. On-orbit repair is rarely feasible, and the time and monetary cost of replacing a spacecraft are high.
4. Rapid deployment or redeployment of spacecraft to meet new or changing mission requirements.
5. Radiation tolerance. The specific requirements depend on the mission profile, but all spacecraft systems must be designed to cope with cosmic radiation, chiefly heavy ions.

In some space missions, rapid deployment of a new spacecraft, or redeployment of an existing craft, are of high priority. Many of potential benefits of reconfigurability for space are laid out in (Lanza, 2004). Reconfigurability can help designers meet challenges (1) and (2) above by enabling hardware to be reconfigured for different phases of a computational task, resulting in smaller size, lower weight, and reduced power consumption than would

be the case if each task required dedicated hardware. For example, in an instrument calibration task, three configurations of a processor were used to capture and format calibration data, and a fourth configuration was used in operational mode, to calibrate data in real time (Sabde, 2003 & 2004).

### 1.1 Radiation effects and their mitigation

It is the need for radiation tolerance that sets spacecraft systems apart. There are three basic approaches to radiation protection: shielding, system-level redundancy, and chip-level mitigation. Most spacecraft employ some kind of radiation shielding, but this adds size and weight, and provides only partial protection. System level redundancy, such as Triple Module Redundancy (TMR), protects against radiation-induced soft errors, but increases complexity, weight, and power consumption up to five fold. There is much to be gained by designing microelectronic circuits and systems to be intrinsically resistant to the effects of radiation (Holmes, 2002).

The primary radiation effects of concern in Complementary Metal Oxide (CMOS) integrated circuits are divided into Total Ionizing Dose (TID) and Single Event Effects (SEE), which include Single Event Latchup (SEL), Single Event Upset (SEU), and Single Event Transients (SET).

**Total Ionizing Dose (TID).** This is caused by the accumulation of ions in the integrated circuit materials, particularly low-grade "field" oxides. The electric charges carried by the trapped ions produce an electric field that alters the behavior of the circuits, most commonly causing leaking between adjacent components. The significance of TID depends on the space mission, as some regions of space carry higher radiation concentrations than others. TID effects are mitigated by choosing an appropriate CMOS process in which materials are pure, and less likely to trap charged particles, and feature dimensions are small enough to enable particles to anneal out as fast as they accumulate. Most modern processes are adequately resistant to TID for space applications.

**Single Event Latchup (SEL).** This is a pathological condition in which a particle strike induces a current loop in parasitic transistors, which can disable or destroy the device. SEL is mitigated in bulk CMOS processes by adding guard bands of heavily doped silicon which drain away these charges, and weaken the parasitic circuits. Silicon-on-Insulator (SOI) processes are inherently immune to SEL.

**Single Event Upset (SEU**). This is a soft memory error, induced when the charge deposited by an ion sets or resets a latch, disrupting the state of the digital logic. Latches can be designed at the circuit and layout levels to resist SEU by incorporating redundancy.

**Single Event Transients (SET).** When an ion strike changes the state of combinational logic, which then propagates into the downstream logic and memory circuits, causing logical malfunction, this is called a Single Event Transient. It is mitigated at the circuit and layout levels by adding redundancy and temporal separation between signals.

Radiation Hardening by Design (RHBD) is a practical strategy for achieving radiation tolerance in digital integrated circuits fabricated with commercial processes. The first step is to select a process that is adequately immune to TID, and to design a standard logic cell library that incorporates SEE and SEL mitigation. A processor can then be synthesized from a hardware description language (HDL) using this cell library. To achieve a practical level of radiation tolerance, the final layout must be completed carefully. For example, the radiation sensitivity of critical circuit elements such as clock trees can be enhanced by increasing the size and drive strength of these elements, enabling them to harmlessly absorb the extra

charge. Radiation Hardening by Design (RHBD) can often achieve acceptable levels of radiation tolerance, but exacts a performance cost, as the resulting processor is slower, larger, and consumes more power than an un-hardened equivalent.

## 2. Data processor options

Embedded computing systems are commonly built around one or more core processor chips. Microprocessors, microcontrollers and digital signal processors (DSPs) are essentially sequential processors, executing one pre-defined instruction at a time. Some processor architectures feature multiple execution units and instruction pipelining to increase throughput, but the execution model is still basically sequential. The design is completed through system and application software. Sequential processors are unquestionably the most flexible processor choice, but they face certain performance limitations. The primary way to increase data throughput in a sequential processor is to reduce the instruction cycle time. Radiation hardening adds delays, which increases instruction cycle time, often very significantly.

High throughput can often be achieved with application-specific integrated circuits (ASICs), semi-custom chips often synthesized from a hardware description language (HDL). ASICs fabricated from commercial processes can be made sufficiently radiation-tolerant for space applications by using a radiation-hard-by-design (RHBD) techniques. As the name implies, however, these chips are inflexible, and cannot usually be retargeted to meet different system requirements. Design schedules are long – months to years – and the non-recurring engineering (NRE) and fabrication costs are high.

Another choice is high-performance field programmable gate arrays (FPGAs), which can be configured to implement computing functions, and are often the platform of choice for reconfigurable computing. By departing from the sequential computing model, data path parallelism, and hence throughput, can be increased. With today's high-performance FPGAs, however, it is difficult to meet power and radiation requirements for space applications. Originally intended to replace logic gates, they are based on a fine-grained architecture, often requiring millions of configuration registers and interconnect switches. The required circuitry consumes silicon area are and leads to high power consumption. FPGA systems can be hardened against single-event radiation effects or soft errors, at the system level with triple-module redundancy (TMR), but this complicates the hardware significantly and greatly increases the power consumption.

The chapter will focus on processors that are designed specifically for reconfigurable computing, using the Field Programmable Processor Array as a case study (Donohoe, 2007). The goal is to achieve data processing rates comparable to ASICs with flexibility approaching that of software. Energy efficiency is achieved by selecting a data granularity appropriate to numerical processing, which reduces interconnect and configuration complexity by up to three orders of magnitude compared to FPGAs. Radiation tolerance can be built in at the gate level using RHBD techniques, thus avoiding the area and power costs of triple-module redundancy.

### 2.1 Architectures

Spacecraft on-board processing fits in the category of embedded computing, as distinct from general-purpose computing associated with laptop or desktop computers. An embedded

computer is designed into a product, is often not detectable as a computer by the users, and is targeted to a comparatively limited range of processing tasks. Thus, some flexibility may often be sacrificed to maximize throughput and energy efficiency.
Spacecraft data processing applications that demand high throughput are often streaming processes such as signal and image processing, instrument calibration, and feedback control loops. These are well-suited to Single Instruction-Multiple Data (SIMD) computational models, something they share with multimedia processing. Most reconfigurable computing implementations perform best on these kinds of problems, which fit naturally into a dataflow computing model. A. Dehon provides a survey of computing models for reconfigurable computing in Chapter 5 of (Hauck, 2008), with emphasis on dataflow models.
Some representative reconfigurable computing architectures are listed below.
**PipeRench.** A product of research at Carnegie-Mellon University, the PipeRench (Goldstein, 2000; Tessier, 2 001) is a very innovative architecture that interleaves processing elements and data registers to form a virtualized data path, portions of which map onto hardware at any time. This virtualization model enables implementation of computational problems too large to fit directly on the hardware. When implemented in a 0.18 micron process, it was competitive with commercial DSP chips in throughput, consuming less than one watt for a 20-tap FIR filter.
**RAW Processor.** Developed at the Massachusetts Institute of Technology, the RAW processor (Taylor, 2002) is essentially an on-chip network of RISC processor cores with flexible interconnect to enable it to set up a multicore computing fabric. It represents the high end of complexity and flexibility. It is not intended for low-power applications.
**Stream Processor.** The Stream Processor (Khailany, 2008) is an example of an SIMD architecture designed for signal processing and multimedia applications. Implemented as a System on Chip (SOC) in a 0.13 micron process, it is reported to deliver performance of 256 16-bit Giga-operations per second (GOPS) with excellent power efficiency. The architecture is quite comprehensive, optimized for processing data streams.
For more examples, see "Reconfigurable Computing for Digital Signal Processing: A Survey" (Tessier, 2001).

## 3. Field programmable processor array

The Field Programmable Processor Array a reconfigurable processor architecture and chip developed for high-throughput, low-power, on-board data processing for space (Donohoe, 2007). The design philosophy is to minimize hardware features to incorporate only those necessary to implement a class of challenge problems, which include examples of instrument calibration, Fourier transform hyperspectral image processing, wavelet-based operations, image registration, and pattern recognition. One processor chip contains 16 processing elements (PEs), each capable of multiplication and addition, an on-board execution unit with eight nested hardware loops, five 16-bit I/O ports, a special dynamically switchable output port, and interface circuits. FPPA chips can be tiled to form a larger computing fabric. Internal data paths are 16 and 32 bits. Multiple precision is supported.

### 3.1 FPPA architecture

The Field Programmable Processor Array implements a dynamic streaming dataflow computing model, as described in DeHon's survey (Hauck, 2008, Ch. 5). Each processing

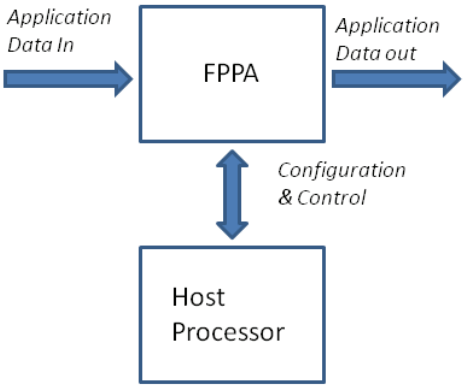



Fig. 1. Arrangement of FPPA and host

element serves as node in the data flow network. A host processor is required to initialize and configure the FPPA, as shown in Figure 1. Figure 2 shows a data path pipeline constructed from four processing elements, an I/O port (IOP0) configured for input, and an I/O port (IOP1) configured for output, all of which form nodes in the dataflow network. In the synchronous dataflow computing model the sequence of operations is encoded in the topology of the network. Processing occurs at a node when valid data appear at all the inputs to a node, and the node "fires", that is, it latches input data and begins processing. In the FPPA, the "Fire" signals (Fire IOP0, etc.) initiate processing at a node; data appear at the output when computation at that node is complete. This is sometimes called a "blocking read, non-blocking write" structure.

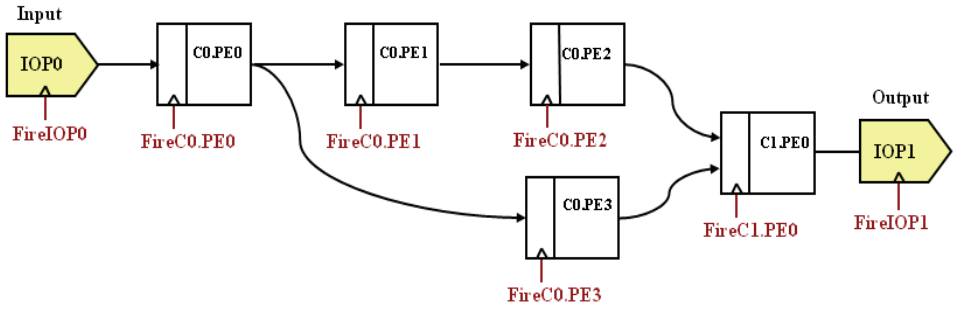



Fig. 2. Synchronous dataflow processing pipeline

A degree of flexibility is provided by the ability to conditionally select one or another input to a node based on a computational result. In DeHon's taxonomy, the FPPA implements a dynamic streaming dataflow compute model (Hauck 2008).

Operation occurs in two phases: (1) configuration, in which processing elements are configured by a host processor to perform data processing functions, and (2) execution, where the processor operates autonomously on a stream of data, without intervention by the host. In addition to the I/O and output ports, the processor implements sixteen one-bit

output control signals and sixteen one-bit input control signals. The output controls signal internal state to external devices. The input controls implement a "wait on event" function, halting FPPA processing until an external condition is satisfied. In this way, the FPPA can synchronize itself with external devices.
The goals of high throughput with deterministic sample rate, low power consumption, and radiation tolerance are addressed below.
High throughput is achieved through through data path pipelining. Multiple chips can be connected to increase the size of the computational fabric in increments of 32 PEs. Interconnect paths are established during a configuration, and persists until the next configuration phase.
Low power consumption is achieved by concentrating chip resources in the data path. The control mechanism is very simple and distributed, minimizing the complexity of control circuitry and communication.
The chip is implemented using Radiation Hard by Design (RHBD) techniques built into a standard cell library designed for a commercial bulk CMOS process. The chip was synthesized using a commercial software tool (Synopsys). The library mitigates single even latchup (SEL) with guard bands, and single event upset (SEU) with a 14-transistor SEL-resistant latch circuit.

### 3.2 Interconnect structure

The interconnect structure provides dedicated data paths within a configuration cycle. There are no shared busses, although conditional data path selection enables data streams to be merged and split. As a result, there is no bus contention, and execution time is deterministic.
Up to 90% of the area of a high-performance FPGA might be dedicated to programmable interconnect, including routing channels, switches, and configuration registers. This is due to the fine-grained, bit-oriented architecture, and the high flexibility of the configurable logic elements. The result is large chip area and relatively high power consumption. The logic elements of the FPPA are more coarse-grained, directly implementing a16-bit arithmetic and logic functions. The required interconnect are and power consumption are proportionally less.
The datapth Interconnect requirements are:

1. Read in stream(s) of data
2. Break data streams out into separate substreams for processing
3. Reassemble substreams into output stream(s) by combining with arithmetic/logic operations, or interleaving
4. Write output streams
5. Synchronize input and output data streams with external devices

The interconnect implementation is subject to these constraints:

1. No throughput penalty for adding processing nodes
2. Acceptable interconnect area
3. Minimal interconnect stages, hence delays
4. Low interconnect power dissipation
5. Small number of reconfiguration bits to enable fast and energy-efficient configuration

The interconnect structure designed to meet these requirements is shown in Figure 3. Data stream input and output requirements #1 and #4 are met with clocked data I/O ports.

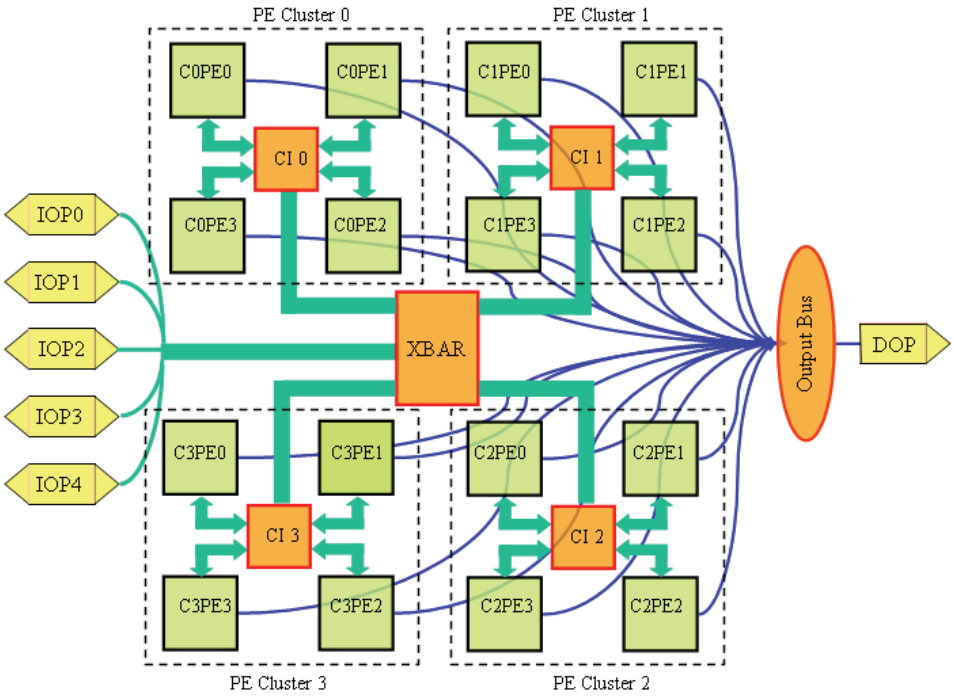



Fig. 3. FPPA datapath

Requirement #2 is implemented with input registers on each of the PEs, under the control of a program memory; this enables each PE to pick the required data items out of an input stream to form the head of a new chain of processing nodes. Recombining data streams, requirement #3, is accomplished in one of these ways:

a. Streams can be combined in a PE with arithmetic operations (add, subtract, multiply) or logical operations (AND, OR, etc.)
b. Conditional data path selection enables selection of one of two paths based on a condition code resulting from an operation. For example, selection might be based on magnitude (greater than, less than, equal to) or other arithmetic or logical relations.
c. The Dynamic Output Port (DOP) allows the output data word to be chosen, under program control, from any of the PEs. This can be used, for example, to reshuffle the data in a Discrete Fourier Transform calculation to produce the desired order.

Constraint #1 precludes shared buses; the chosen architecture uses dedicated point-to-point paths curing a configuration phase. To meet constraints #2 and #3, we assume that applications exhibit spatial locality – that is, it is possible to concentrate high-speed activities requiring rich interconnect between processing elements into a small spatial region, while sparser communication can be tolerated between spatially separated nodes. The resulting architecture is hierarchical, with fourteen interconnect channels within a PE cluster. This supports data-intensive local operations such as multiprecision accumulation. Six channels are provided between each cluster and the central routing switch. Finally, six ports are

provided for each FPPA chip. Constraint #5 is met by the sixteen bit granularity and the pre-built arithmetic and logic circuits, which require many fewer configuration bits than a bit-oriented FPGA.

### 3.2 FPPA software

A novel computing paradigm requires a new approach to software. Programming languages designed for sequential, random-access machines are not easily adapted to highly concurrent architectures. The challenge for dataflow architectures (and similar computational structures implemented in FPGAs), is to create "circuitizeable" code. Where flow control in a sequential computer is determined by a sequence of instructions, in a dataflow architecture the flow of data is controlled through registers and data path selectors, e.g. multiplexers.

Systems that target FPGAs for reconfigurable computing often use hardware description languages such as VHDL in order to achieve detailed low-level control of the fine-grained architectures. While well-suited to describing hardware circuits, VHDL is too low-level to be effective for implementing general algorithms.

Software developed to support the FPPA includes a functional simulator, assembly-level software translators to produce configuration and run-time files, and design-entry tools, including a graphical design entry program and an adaptation of the Single-Assignment C (SA-C) language compiler (Bohm, 2002).

SA-C is a C-like language originally developed to target FPGAs then modified for the FFPA. To ensure that the code is circuitizable, the Single-Assignment C (SA-C) language imposes the constraint that a variable can only be written (assigned) at one point in the program.

Figure 4 shows a flow diagram for application software development for the FPPA. The source can be written in SA-C or the FPPAflo graphical design entry program. Both compile to a dataflow graph, an intermediate representation that is suitable for automatic optimization. The key optimization parameters are throughput and computing resources, primarily processing elements and routing. Resources may be distributed over multiple FPPA chips.

The source code must be translated into two files: a configuration file, and a runtime file. The configuration file contains configuration instructions to set up the input and output ports, the processing elements, and the internal routing paths. The runtime file contains the runtime program specification. These are translated into bit streams by the configuration assembler and runtime assembler, respectively. The final output is a file of packets to be read by the host processor to configure the FPPA, in the FPPA-specific host-computer interface (HCI) format.

### 3.3 Memory

The FPPA is a shallow-memory device; the only data memory is the input registers to the processing elements and the I/O ports, and an 8-word programmable delay line in each PE. The latter is provided to aid in aligning parallel paths in the dataflow pipeline. Many applications, such as one-dimensional linear and nonlinear filters and wavelets, can be implemented without additional memory.

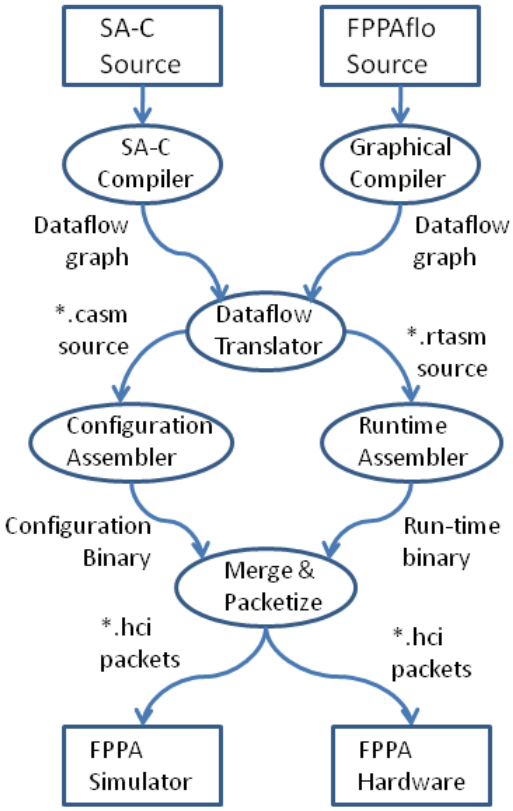



Fig. 4. FPPA software design flow

Many applications require larger-scale data memory, however. For example, two-dimensional image filters typically require "line buffers" to hold past image values. Sequential processors are typically coupled with random access memory to store instructions and data, accessible over shared busses. For streaming applications, data memory is often organized in regular data structures with well-defined access modes: first-in, first-out (FIFO) queues, last-in, first-out (LIFO) stacks, sequential access, circular buffers, or two-dimensional matrices representing images or subimages. High-performance FPGAs are typically equipped with user-definable memory. These large-scale on-chip platforms are not yet available in radiation-hard form, however.

In the FPPA, program memory is small (128 x 128 bits); each bit controls one module. The program memory is distributed over the chip into 1x128 bit latch arrays, each co-located with the module it controls. Data memory is not provided on chip because of the challenges of implementing large-scale memory in a logic process, and the problem of deciding how

much memory to incorporate and how to organize it, as this is highly application dependent.

The FPPA can work with random access memory, but this is very costly in terms of I/O ports and board-level interconnects, as shared busses are not supported, and dedicated data busses are required for addresses and data. For this reason, a Reconfigurable Memory Module (RMM) was conceived to aid in coupling the FPPA with commercial memory chips. This consists of an interface module configured to generate addresses for a particular application, interfaced with a commercial memory chip. In one instantiation of the RMM, the following access modes were implemented:

**RMM Data Access Modes**

1. Sequential access
2. Last In, First Out (LIFO) stack
3. First In, Last Out (FIFO) buffer
4. Circular (ring) buffer
5. Lookup Table (LUT)
6. Region of Interest (ROI)
7. Histogram

The circular buffer defines a beginning and end of a buffer in the memory. Access is sequential until the end of the buffer is reached; then it wraps around to the beginning. Lookup table mode (mode 5) is just random access, using one of the data ports for an address and the other for data. Region of interest mode (mode 6) models the random access memory as a two-dimensional array (often used to store images), and defines a subarray, or subimage; sequential accesses sweep horizontally across a row of data elements; when the end of the row is reached, the address counter moves to the beginning of the next row.

Histogram mode (mode 7) supports histogram calculations, which are sometimes used in image enhancement and in signal and image analysis. For example, estimating signal entropy requires calculating

$$H(X) = \sum_{x=0}^{N-1} p(x_n) \log_2 p(x_n) \tag{1}$$

Where $X$ is a set of $N$ values $x_0$ to $x_{N-1}$ and $p(x_n)$ is the probability of occurrence of value of element $x_n$. This probability can be estimated by counting occurrences of $x_n$ for each $n$, i.e., forming a histogram, then dividing by the total number of samples. This histogram is initialized by setting each element to zero using sequential access mode. To calculate the histogram, a data sample is read from a source, and this is used as an index into the histogram in a read-modify-write operation: the histogram values is read, incremented, then written back. Finally, the histogram can be read out in sequential access mode. Thus, three configurations of the RMM are required.

Figure 5 shows a typical arrangement for an FPPA systems consisting of one FPPA, a host processor, and the RMM. The host configures and initializes both the FPPA and the RMM. One FPPA data port is used for writing from the FPPA to the RMM, and one for reading from the RMM into the FPPA. The host communicates with the RMM via an 8-bit data bus and two asynchronous handshake signals, RTX and CTX. Once the FPPA and RMM are set up, and the FPPA begins to execute, the host is not longer required. Two control bits, command_0 and command_1, are sufficient for the FPPA to control the RMM.

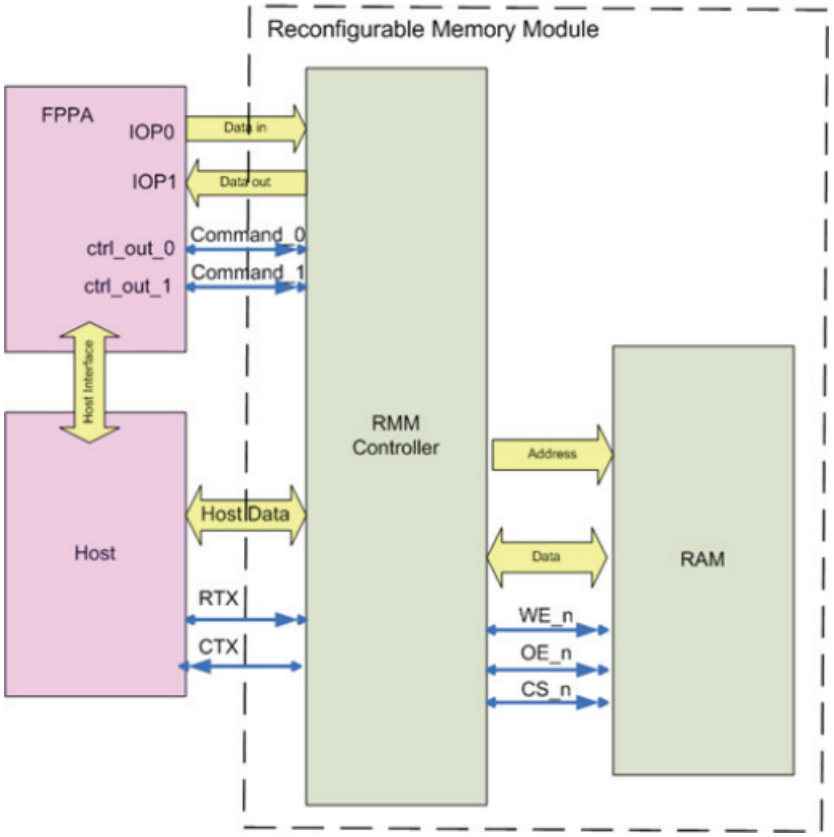



Fig. 5. Example system consisting of FPPA, host, and RMM

### 3.4 Prototype chip

A prototype of the FPPA was synthesized from VHDL using radiation-hard-by-design standard cell library implemented in a bulk 0.25 micron commercial CMOS process[1]. The design then fabricated and tested. The die measures 7 x 7 mm, contains 241,306 gates, and uses 200 pins of a 208 pin flat-pack package.

The chip was tested at 22.5 MHz on two challenge problems: a cloud masking algorithm for hyperspectral image data (Zhou, 2003), and wavelet-based dimensionality reduction for hyperspectral images (Sinthop, 2003). Both applications were completely parallelizable, so the throughput was one output sample per input sample. Both applications were implemented so that they fit in one FPPA chip, without external data memory. To facilitate comparison with other processors, a power-performance metric was defined, mega-

[1] The prototype chip is not fully radiation tolerant, as single-event-upset resistant latches were not used.

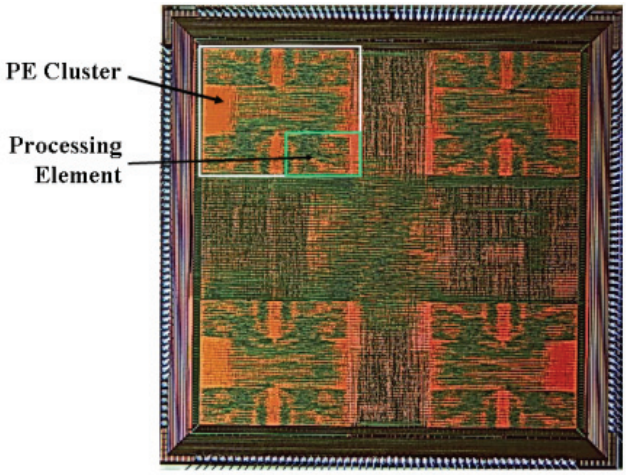



Fig. 6. Prototype FPPA chip

operations/second-watt (MOPS/watt). Instructions sets vary between processors, so that one processor may require more machine level instructions to implement a computation than another. To the system designer, the relevant instructions are those that transform data as defined in a high-level description of the problem. Here we define an "instruction" as an arithmetic operator (add, multiply), or a logic operator (shift, AND, etc.).

For the cloud masking algorithm, the FPPA ran at 742 MOPS consuming 0.23 watts, or 3,167 Mops/sec-watt. The FPPA executed the wavelet application with 1080 MOPS consuming 0.48 watts, or 1,849 MOPS/watt.

## 4. Conclusion

Spacecraft represent one of the application domains that have the most to gain from reconfigurability, particularly due to the high cost and long delays associated with designing, building, qualifying and launching a new craft (Lanza 2004). Areas in which spacecraft can benefit include:

- Tele-alteration: the ability to change configuration and function of a system remotely, increasing mission agility and extending the useful life of the system.
- Resilience and robustness: reconfigurability makes it possible to map out failed components and restore functionality.
- Functionality on demand : when a new function is required, if the configurable resources are in place, remotely configure the required functionality

Adapting commercial technologies to reconfigurable space systems is complicated by the requirements of small size and weight, low power consumption, and radiation hardness. This chapter presented an example of a reconfigurable processor developed for space using radiation-hard-by-design techniques. The goals of the design were to achieve an appropriate level of functionality to solve a wide range of space computing problems, extensible to larger-scale problems than can be implemented in one chip, minimal power consumption, and sufficient radiation hardness for space applications. The architecture implements a

synchronous dataflow computing model, and is extensible by tiling multiple chips to increase the size of the dataflow network, and adding reconfigurable memory modules for applications requiring external data memory.

Reconfigurable computing is becoming a reality, but we can envision extending the concept of reconfigurability beyond logic to other domains. Software radio is a familiar example, in which key components of a radio communication system can be altered to adapt to different frequencies and communication protocols. Reconfigurable analog systems could permit in situ configuration of sensing and control functions. Reconfigurable pathways (wiring, optical, fluidic, thermal) could enable software-defined physical systems. A reconfigurable power infrastructure, and even reconfigurable matter can be conceived to accommodate as-yet unimagined demands for spacecraft systems.

## 5. References


Bohm, W.; Hammes, J.; Draper, B; Chawath, M.; Ross, C.; Rinker, R. & Majjar, W. (2002). Mapping a Single Assignment Programming Language to Reconfigurable Systems, *Supercomputing* 21:117-130, 2002.

DeHon, A. (1996). *Reconfigurable Architectures for General-Purpose Computing*, MIT AI Lab Report No. 1586, Massachusetts Institute of Technology, Cambridge, MA, USA, 1996.

Donohoe, G; Buehler, D; Hass, K; Walker, W. & Yeh, P.-S. (2007). Field Programmable Processor Array: Reconfigurable Computing for Space, *IEEE Aerospace Conference*, Big Sky, MT, March 3-10, 2007.

Goldstein, S; Schmit,H., Budiu, M., Cadambi, S., Moe,M. & Taylor, R. (2000) "Piperench: a Reconfigurable Architecture and Compiler", IEEE Computer, Vol. 33, Issue 4, April 2000, pp. 70-77.

Hauck, S. & Dehon, Andre (2008), Reconfigurable Computing: The Theory and Practice of FPGA-Based Computation, Elsevier/Morgan Kaufmann, Amsterdam, ISBN 978-0-12-370552-8.

Holmes, A. & Adams, L. (2002), Handbook of Radiation Effects, Oxford University Press, Oxford, ISBN-13 978-0198507338.

Lanza, D; Lyke, J.; Zetocha, P; Fronterhouse, D; & Melanson, D. (2004). Responsive Space through Adaptive Avionics, Space 2004 Conference and Exhibit, American Institute of Avionics and Aeronautics, AIAA 2004-6116, San Diego, CA, 2004.

Sabde, J; Buehler, D. & Donohoe, G. (2003) Focal Plane Array Sensor Readout Correction on a Reconfigurable Processor, Proc. 11th NASA Symposium on VLSI Design, Coeur d'Alene, ID, USA, May 2003.

Sabde, J. (2004). Sensor Data Processing on a Reconfigurable Processor, Master of Science Thesis, University of Idaho, USA.

Sinthop,K.; Le Moigne, J. & El-Ghazawi, T. (2003). "Automatic Reduction of Hyperspectral Imagery using Wavelet Spectral Analysis," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 41, No. 4, April 2003.

Tessier, R. & Burleson, W. (2001). Reconfigurable Computing for Digital Signal Processing: A Survey, *Journal of VLSI Signal Processing* 28, 7-27, 2001.

Villasenor, J. & and W. Mangione-Smith, W. (1997). "Configurable Computing", *Scientific American*, June 1997.

Zhou, Y.; P.-S. Yeh,, P.-S.; Wiscombe, W. & S.-C.Tsay, S.-C. (2003) "Cloud context-based onboard data compression," *Proc. IGARS 2003*, Toulouse, July 21-25, 2003.

# 4

# Spacecraft Attitude Control

Thawar Arif
*Applied Science University*
*Jordan*

## 1. Introduction

The motion of a spacecraft is specified by its position, velocity, attitude, and attitude motion. The first two quantities describe the translational motion of the center of mass of the spacecraft and are the subject of what is called Orbit Determination. The latter two quantities describe the rotational motion of the body of the spacecraft about the center of mass and are the subject of what is called Attitude Determination and Control.
The attitude of a spacecraft is its orientation in space with respect to a defined frame of reference. This chapter[1] discusses the aspects of spacecraft attitude control. It is an engineering discipline aiming at keeping the spacecraft pointing in the right direction.
In this work, the attitude control of flexible spacecraft is studied. The flexible satellite is considered as a large scale system since it comprises several coupled subsystems. The control of large-scale systems, which are composed of interconnected subsystems, usually goes hand in hand with poor knowledge of the subsystem parameters. As a result, the use of adaptive schemes is particularly appropriate in such a situation. Even assuming perfect parameter knowledge, the design and implementation of a single centralized controller for a large scale system turns out to be a formidable task from the point of view of design complexity as well as associated expenditure (Datta, 1993). Consequently, decentralized adaptive control schemes, whereby each subsystem is controlled independently on the basis of its own local performance criterion and locally available information, have been proposed in the literature (Lyou, 1995; Spooner & Passino, 1996). The advantages of the decentralized schemes are (Benchoubane & Stoten, 1992):

1. The controller equations are structurally simpler than the centralized equivalent.
2. No communication is necessary between the individual controllers.
3. Parallel implementations are possible. One of the main advantages to the practical control engineer would be that as the system is expanded, new controller loops could be implemented with no changes to those already in existence. One of the powerful decentralized adaptive control schemes is that developed by (Benchoubane & Stoten, 1992) called the Decentralized Minimal Controller Synthesis which is an extension of the Minimal Controller Synthesis scheme developed earlier by (Benchoubane & Stoten, 1990a). The minimal controller synthesis strategy is based on a model reference

[1] This chapter received support towards its publication from the Deanship of Research and Graduate Studies at Applied Science University, Amman, Jordan.

adaptive control scheme using positivity and hyperstability concepts in its design procedure to ensure asymptotic stability. The minimal controller synthesis algorithm requires a minimal amount of computation. Various theoretical and experimental studies have shown that it possesses the stability and robustness features essential to any successful adaptive control scheme.

To date, most of the minimal controller synthesis implementation studies have been made on controlling robotic systems (Stoten & Hodgson, 1992), chaos (Di Bernardo & Stoten, 2006), X-38 crew return vehicle (Campbell & Lieven, 2002), or substructuring of dynamical systems (Wagg & Stoten, 2001).

The Decentralized Minimal Controller Synthesis is adopted in this work for controlling the attitude of flexible spacecraft. Equations of motion are written with respect to a coordinate system fixed in the spacecraft and oriented along its principal axis. The control is by means of three reaction wheels which are also oriented along the principal axes of the spacecraft. It is assumed, for simplicity, that there is no wheel damping and that wheel torque can be controlled precisely.

Many spacecraft attitude control systems, which use Euler angles or direction cosine matrix for parameterization of the attitude kinematics, are based on a sequence of rotations about each of the three principal axes separately (Pande & Ventachalam, 1982). However, the time needed to realize such a reorientation increases by a factor of two or three, compared with one single three axes slew, which is obtained when the quaternion is used for parameterization (Luo et al, 2005). The quaternion is adopted in this work.

## 2. Minimal controller synthesis

The minimal controller synthesis algorithm (Benchoubane & Stoten, 1990b) is a significant extension to model reference adaptive control (Landau, 1979). In a similar manner to model reference adaptive control, the aim of minimal controller synthesis is to achieve excellent closed-loop control despite the presence of plant parameters variations, external disturbances, dynamic coupling within the plant and plant non-linearities. However, unlike model reference adaptive control, minimal controller synthesis requires no plant model identification (apart from the general structure of a state space equation) or linear controller synthesis.

Considering a single-input single-output plant described by the following state-space equation:

$$\dot{\mathbf{x}}_p(t) = \mathbf{A}\mathbf{x}_p(t) + \mathbf{b}u_p(t) + \mathbf{d}(t) \tag{1}$$

where $\mathbf{x}_p$ is an n-vector, $u_p$ is a control signal, $\mathbf{A}$ is an nxn plant coefficient matrix, and $\mathbf{b}$ is an nx1 control coefficient vector. The term $\mathbf{d}(t)$ represents an nx1 vector aggregate of unknown external disturbances applied to the plant, plant non-linearities, and any unmodelled terms.

In general, $\mathbf{d}(t) \neq \mathbf{0}_{n,1}$, and if $\mathbf{x}_p(t) \neq \mathbf{0}_{n,1}$, then $\mathbf{d}(t)$ can be represented as (Benchoubane & Stoten, 1990a):

$$\mathbf{d}(t) = \delta\mathbf{A}_1(t)\mathbf{x}_p(t) \tag{2}$$

The term $\delta\mathbf{A}_1$ can be considered as an unknown variation in the $\mathbf{A}$ matrix, structured according to any admissible variations in $\mathbf{A}$. Also some other admissible variations in

matrices {**A**, **b**} can occur, owing to system parameter and/or environmental changes. Let these changes be denoted by $\delta\mathbf{A}_2(t)$ and $\delta\mathbf{b}(t)$, respectively; also let:

$$\delta\mathbf{A}(t) = \delta\mathbf{A}_1(t) + \delta\mathbf{A}_2(t) \tag{3}$$

Then the state equation (1) can be rewritten as:

$$\dot{\mathbf{x}}_p(t) = (\mathbf{A} + \delta\mathbf{A}(t))\mathbf{x}_p(t) + (\mathbf{b} + \delta\mathbf{b}(t))u_p(t) \tag{3}$$

In common with model reference adaptive control, the objective of the minimal controller synthesis is to ensure that the system state $\mathbf{x}_p(t)$ faithfully follows the state of a reference model despite the effects of the unknown variations $\delta\mathbf{A}(t)$ and $\delta\mathbf{b}(t)$. The reference model is known exactly as (Benchoubane & Stoten, 1992):

$$\dot{\mathbf{x}}_m(t) = \mathbf{A}_m\mathbf{x}_m(t) + \mathbf{b}_m u_m(t) \tag{4}$$

Where $\mathbf{x}_m(t)$ is an nx1 model reference state vector, $u_m(t)$ is a reference signal, $\mathbf{A}_m$ is an nxn model reference coefficient matrix with constant elements, and $\mathbf{b}_m$ is an nx1 reference signal coefficient vector with constant elements. The control law of the model reference adaptive control is given by (Wertz, 1980):

$$u_p(t) = (-\mathbf{k}_p + \delta\mathbf{k}_p(t))\mathbf{x}_p(t) + (k_u + \delta k_u(t))u_m(t) \tag{5}$$

where $\mathbf{k}_p$ is a 1xn constant feedback gain vector and $k_u$ is a constant feedforward gain. The $\delta\mathbf{k}_p$ and $\delta k_u$ terms are adaptive changes to these gains that usually result from the effects of $\mathbf{d}(t)$ on the state trajectory, $\mathbf{x}_p(t)$. Whilst the control law of the minimal controller synthesis is given by setting $\mathbf{k}_p = \mathbf{0}_{n,1}$, $k_u = 0$, so that (Benchoubane & Stoten, 1990b):

$$u_p(t) = \delta\mathbf{k}_p(t)\mathbf{x}_p(t) + \delta k_u(t)u_m(t) \tag{6}$$

In equation (5), the linear model reference controller gains $\mathbf{k}_p$ and $k_u$ can be found in closed form, assuming that Erzberger's conditions are satisfied (Isermann, 1992). The satisfaction of Erzberger's conditions tends to restrict the choice of reference model. In particular, equation (6) contravenes the conditions whilst retaining robustness.
Substituting equation (6) into (3) gives:

$$\dot{\mathbf{x}}_p(t) = \mathbf{A}_p(t)\mathbf{x}_p(t) + \mathbf{b}_p(t)(\delta\mathbf{k}_p(t)\mathbf{x}_p(t) + \delta k_u(t)u_m(t)) \tag{7}$$

where $\mathbf{A}_p(t) = \mathbf{A} + \delta\mathbf{A}(t)$ and $\mathbf{b}_p(t) = \mathbf{b} + \delta\mathbf{b}(t)$.
Therefore, the closed-loop plant dynamics becomes:

$$\dot{\mathbf{x}}_p(t) = (\mathbf{A}_p(t) + \mathbf{b}_p(t)\delta\mathbf{k}_p(t))\mathbf{x}_p(t) + \mathbf{b}_p(t)\delta k_u(t)u_m(t) \tag{8}$$

From equations (4) and (8), the error dynamics of the closed loop system are given by:

$$\dot{\mathbf{x}}_e(t) = \mathbf{A}_m\mathbf{x}_e(t) + (\mathbf{A}_o(t) - \mathbf{b}_p(t)\delta\mathbf{k}_p(t))\mathbf{x}_p(t) + (\mathbf{b}_m - \mathbf{b}_p(t)\delta k_u(t))u_m(t) \tag{9}$$

where $\mathbf{x}_e(t)$ is an nx1 error state vector which is given by $\mathbf{x}_e(t) = \mathbf{x}_m(t) - \mathbf{x}_p(t)$. While $\mathbf{A}_o(t) = \mathbf{A}_m - \mathbf{A}_p(t)$.
From equation (9), let:

$$\mathbf{v}(t) = (\mathbf{A}_o(t) - \mathbf{b}_p(t)\delta\mathbf{k}_p(t))\mathbf{x}_p(t) + (\mathbf{b}_m - \mathbf{b}_p(t)\delta k_u(t))u_m(t) \tag{10}$$

so that:

$$\dot{\mathbf{x}}_e(t) = \mathbf{A}_m\mathbf{x}_e(t) + \mathbf{I}_n\mathbf{v}(t) \tag{11}$$

where $\mathbf{I}_n$ is an nxn identity matrix.

The absolute stability of equation (11) is investigated by the application of hyperstability theory and Popov's criterion to the equivalent non-linear closed-loop system (Landau, 1979). In this system, shown in figure (1), let: $\mathbf{v}(t) = -\mathbf{v}_e(t)$ and $\mathbf{v}_e(t)$ is generated by a necessarily non-linear function of the output error vector $\mathbf{y}_e(t)$ (this constitutes the adaptive block); where: $\mathbf{y}_e(t) = \mathbf{P}\mathbf{x}_e(t)$. $\mathbf{P}$ is an nxn positive definite symmetric matrix which is the solution of the Lyapunov matrix equation:

$$\mathbf{P}\mathbf{A}_m + \mathbf{A}_m^T\mathbf{P} = -\mathbf{Q} \tag{12}$$

where $\mathbf{Q}$ is an nxn positive definite matrix.

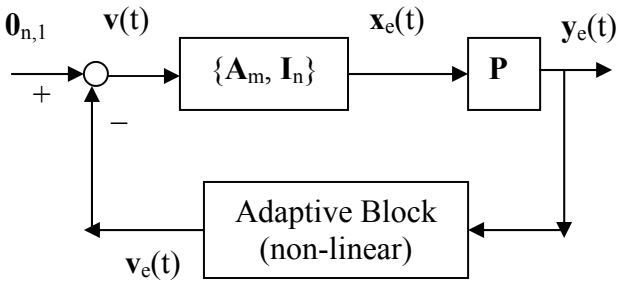



Fig. 1. Closed-Loop System Equivalent to Equations (10) & (11)

The system (11) is hyperstable if the block $\{\mathbf{A}_m, \mathbf{I}_n, \mathbf{P}\}$ is a hyperstable block, i.e. satisfies Lyapunov matrix equation (12), and the following Popov's inequality is satisfied (Landau, 1979):

$$\int_{to}^{t1} \mathbf{y}_e^T(t)\,\mathbf{v}_e(t)\,dt \geq -\mu^2, \qquad \mu > 0; \text{for all } t_1 \geq t_o \tag{13}$$

For a given reference model and arbitrary positive definite matrix $\mathbf{Q}$, the Lyapunov matrix equation (12) can be solved to yield the positive definite symmetric matrix $\mathbf{P}$ (Landau & Courtiol, 1974). It remains to satisfy equation (13), which can be rewritten, using equation (10), as:

$$\int_{to}^{t1} \mathbf{y}_e^T(t)\left(\mathbf{b}_p(t)\delta\mathbf{k}_p(t) - \mathbf{A}_o(t)\right)\mathbf{x}_p(t)\,dt \geq -\mu_1^2 \tag{14-a}$$

And

$$\int_{to}^{t1} \mathbf{y}_e^T(t)\left(\mathbf{b}_p(t)\delta k_u(t) - \mathbf{b}_m\right)u_m(t)\,dt \geq -\mu_2^2 \tag{14-b}$$

Where

$$\mu_1^2 + \mu_2^2 = \mu^2$$

The satisfaction of equations (14) is explained in (Arif, 2008).

## 3. Decentralized minimal controller synthesis

It is assumed that the multivariable system to be controlled can be modeled as an interconnection of (m) single-input single-output subsystems, whose individual dynamics are described by:

$$\dot{\mathbf{x}}_{pi} = \mathbf{A}_{pi}\mathbf{x}_{pi} + \mathbf{b}_{pi}u_{pi} + \mathbf{d}_i + \sum_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{f}_{ij}(t, \mathbf{x}_{pj}) + \sum_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{b}_{pij}u_{pj} \; ; (i = 1, 2, \ldots, m) \tag{15}$$

where for this $i_{th}$ subsystem:
$\mathbf{x}_{pi}$ is the state vector of dimension $n_i$ defined as:

$$\mathbf{x}_{pi} = [x_{pi1} \;\; x_{pi2} \;\; \ldots \; x_{pini}]$$

$\mathbf{d}_i$ is the bounded vector of dimension $n_i$ containing the subsystem nonlinearities and external disturbances, $u_{pi}$ is the control variable, $\mathbf{f}_{ij}(t, \mathbf{x}_{pj})$ and $\mathbf{b}_{pij}u_{pj}$ are vectors of dimension $n_i$ representing the bounded interactions with the other subsystem states and control.
Further, the matrix $\mathbf{A}_{pi}$ and the vector $\mathbf{d}_i$ have unknown parameters, but with the assumed structures:

$$\mathbf{A}_{pi} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{pi1} & -a_{pi2} & -a_{pi3} & \cdots & -a_{pini} \end{bmatrix} \tag{16-a}$$

$$\mathbf{b}_{pi} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b_{pini} \end{bmatrix} \tag{16-b}$$

$$\mathbf{b}_{pij} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b_{pijni} \end{bmatrix} \tag{16-c}$$

$$\mathbf{d}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ d_{ini} \end{bmatrix} \tag{16-d}$$

The state interaction terms $\mathbf{f}_{ij}(t, \mathbf{x}_{pj})$ are of the form:

$$\mathbf{f}_{ij}(t, \mathbf{x}_{pj}) = [0 \ldots 0 \ \mathbf{f}_{ij}]^T \tag{17}$$

which satisfy the inequality:

$$||\mathbf{f}_{ij}(t, x_{pj})|| \leq c_{ij} ||\mathbf{x}_{pj}|| \tag{18}$$

where $c_{ij}$ are finite positive, unknown, coefficients.
The system dynamics in a full multivariable guise can be written as:

$$\dot{\mathbf{x}}_p(t) = \mathbf{A}_p\mathbf{x}_p(t) + \mathbf{b}_p u_p(t) + \mathbf{d}(t) + \mathbf{f}(t, \mathbf{x}_p) \tag{19}$$

where

$$\mathbf{x}_p(t) = [\ \mathbf{x}_{p1}^T\ \mathbf{x}_{p2}^T\ \ldots\ \mathbf{x}_{pm}^T\ ]^T = \text{the complete state vector.}$$

$$u_p(t) = [u_{p1}\ \ u_{p2}\ \ \ldots\ \ u_{pm}]^T = \text{the complete control vector.}$$

$$\mathbf{d}(t) = [\ \mathbf{d}_1^T\ \mathbf{d}_2^T\ \ldots\ d_m^T\ ]^T = \text{the complete disturbance/nonlinearities vector.}$$

$$\mathbf{f}(t, \mathbf{x}_p) = [\mathbf{f}_{ij}(t, \mathbf{x}_{pj})] \text{ if } i \neq j$$

$$\mathbf{f}(t, \mathbf{x}_p) = [\mathbf{0}] \text{ if } i = j$$

and

$$\mathbf{A}_p = \text{diag}[\mathbf{A}_{pi}]$$

$$\mathbf{b}_p = [\mathbf{b}_{pij}]; \text{ where } \mathbf{b}_{pii} = \mathbf{b}_{pi}$$

the objective of decentralized minimal controller synthesis is to drive the control signal $u_{pi}$ for each subsystem given by equation (15), using local information, so that the corresponding states track those of a local reference model, described by:

$$\dot{\mathbf{x}}_{mi}(t) = \mathbf{A}_{mi}\mathbf{x}_{mi} + \mathbf{b}_{mi}u_{mi} \tag{20}$$

where $\mathbf{x}_{mi}$ is the $i_{th}$ reference model state vector of dimension $n_i$ and $u_{mi}$ is the bounded reference input. Furthermore, the matrix $\mathbf{A}_{mi}$ and the vector $\mathbf{b}_{mi}$ are defined as:

$$\mathbf{A}_{mi} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{mi1} & -a_{mi2} & -a_{mi3} & \cdots & -a_{mini} \end{bmatrix} \tag{21-a}$$

$$\mathbf{b}_{mi} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b_{min\,i} \end{bmatrix} \tag{21-b}$$

Therefore, the local information available to the $i_{th}$ subsystem is the set of variables $\{\mathbf{x}_{mi}, \mathbf{x}_{pi}, u_{mi}\}$. The error vector $\mathbf{x}_{ei}$ corresponding to the $i_{th}$ subsystem is:

$$\mathbf{x}_{ei} = \mathbf{x}_{mi} - \mathbf{x}_{pi} \tag{22}$$

and by using equations (15) and (20) we get:

$$\dot{\mathbf{x}}_{ei} = \mathbf{A}_{mi}\mathbf{x}_{ei} - \mathbf{b}_{pi}u_{pi} - (\mathbf{A}_{pi} - \mathbf{A}_{mi})\mathbf{x}_{pi} + \mathbf{b}_{mi}u_{mi} - \mathbf{d}_i - \sum_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{f}_{ij}(t, \mathbf{x}_{pj}) - \sum_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{b}_{pij}u_{pj} \tag{23}$$

Or

$$\dot{\mathbf{x}}_{ei}(t) = \mathbf{A}_{mi}\mathbf{x}_{ei} - \boldsymbol{\lambda}_{i1} - \boldsymbol{\lambda}_{i2} \tag{24}$$

where

$$\boldsymbol{\lambda}_{i1} = \mathbf{b}_{pi}u_{pi} + (\mathbf{A}_{pi} - \mathbf{A}_{mi})\mathbf{x}_{pi} - \mathbf{b}_{mi}u_{mi}$$

$$\boldsymbol{\lambda}_{i2} = \mathbf{d}_i + \sum_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{f}_{ij}(t, \mathbf{x}_{pj}) + \sum_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{b}_{pij}u_{pj}$$

Following the form of the minimal controller synthesis; the control law (6) is proposed for each subsystem, as follows:

$$u_{pi} = \delta\mathbf{k}_{pi}\mathbf{x}_{pi} + \delta k_{ui}u_{mi} \tag{25}$$

where

$$\delta\mathbf{k}_{pi} = \int_0^t \alpha_i\, y_{ni}\mathbf{x}_{pi}^T d\tau + \beta_i y_{ni}\mathbf{x}_{pi}^T \tag{26}$$

$$\delta k_{ui} = \int_0^t \alpha_i y_{ni}u_{mi}\, d\tau + \beta_i y_{ni}u_{mi} \tag{27}$$

$$y_{ni} = \boldsymbol{\Gamma}_i^T \mathbf{P}_i\mathbf{x}_{ei} \tag{28}$$

$$\boldsymbol{\Gamma}_i = [0 \ldots 0\ \ 1]^T = (n_i \times 1) \text{ vector} \tag{29}$$

$\mathbf{P}_i$ is the symmetric positive definite solution of the following Lyapunov equation:

$$\mathbf{P}_i\mathbf{A}_{mi} + \mathbf{A}_{mi}^T \mathbf{P}_i = -\mathbf{Q}_i \tag{30}$$

where $\mathbf{Q}_i = \text{diag}(q_{11} \; q_{22})$ is a positive definite matrix. The elements of $\mathbf{Q}$ are to be selected by the designer. $\alpha_i$ and $\beta_i$ are constant gains.

For a given reference model and arbitrary positive definite matrix $\mathbf{Q}$, the Lyapunov matrix equation (30) can be solved to yield the positive definite symmetric matrix $\mathbf{P}$ (Landau & Courtiol, 1974).

With the aid of equations (25) to (29) the error dynamics given by equation (23) can be rewritten as follows:

$$\dot{\mathbf{x}}_{ei} = \mathbf{A}_{mi}\mathbf{x}_{ei} - \mathbf{\Gamma}_i \mathbf{\Phi}_i^T \mathbf{Z}_i - \mathbf{\Gamma}_i \mathbf{\Psi}_i^T \mathbf{Z}_i - \boldsymbol{\lambda}_{i2} \tag{31}$$

where

$$\mathbf{\Phi}_i^T = \left[b_{pini}\int_0^t \alpha_i y_{ni} x_{pi1} d\tau - (a_{mi1} - a_{pi1}), \ldots, b_{pini}\int_0^t \alpha_i y_{ni} x_{pini} d\tau - (a_{mini} - a_{pini}), b_{pini}\int_0^t \alpha_i y_{ni} u_{mi} d\tau - b_{mini}\right] \tag{32}$$

$$\mathbf{Z}_i = [\mathbf{x}_{pi}^T \; u_{mi}]^T \tag{33}$$

$$\mathbf{\Psi}_i = b_{pini}\beta_i y_{ni}\mathbf{Z}_i \tag{34}$$

### 3.1 Stability and robustness of decentralized minimal controller synthesis algorithm

Equations (31) to (34) define the closed-loop dynamics of the system described by equation (15), under the decentralized minimal controller synthesis control strategy described by equations (25) to (30). These closed-loop equations can be guaranteed hyperstable if the parameters in $\boldsymbol{\lambda}_{i2}$ (equations (24) and (31)) vary slowly, i.e. compared with the speed of the individual adaptive control loops.

The procedure now follows the approach taken in (Benchoubane & Stoten, 1992), whereby the approximately constant parameters are incorporated into the corresponding entries of $\mathbf{\Phi}_i$. Thus, rewrite $\boldsymbol{\lambda}_{i2}$ as:

$$\boldsymbol{\lambda}_{i2} = \delta\mathbf{A}_{pi}u_{mi} \tag{35}$$

where $\delta\mathbf{A}_{pi}$ is an unknown ($n_i$ x 1) vector defined as:

$$\delta\mathbf{A}_{pi} = [0 \ldots 0 \; \delta a_{pi}]^T \tag{36}$$

Therefore, the speed of variation of $\delta\mathbf{A}_{pi}$ is determined by both the speeds of variation of $u_{mi}$ and $\boldsymbol{\lambda}_{i2}$. However, in many practical situations, the reference inputs are relatively slowly varying, and therefore the speed of variation of $\delta\mathbf{A}_{pi}$ is only dependent upon $\boldsymbol{\lambda}_{i2}$. Thus, if the terms $\boldsymbol{\lambda}_{i2}$ is slowly varying, the terms $\delta\mathbf{A}_{pi}$ can be considered as approximately constant and incorporated into the last entry of each $\mathbf{\Phi}_i$:

$$\Phi_i^T = [b_{pini}\int_0^t \alpha_i y_{ni} x_{pi1} d\tau - (a_{mi1} - a_{pi1}), \ldots, b_{pini}\int_0^t \alpha_i y_{ni} x_{pini} d\tau - (a_{mini} - a_{pini}), b_{pini}\int_0^t \alpha_i y_{ni} u_{mi} d\tau + \delta a_{pi} - b_{mini}] \quad (37)$$

Now equation (31) may be rewritten as:

$$\dot{\mathbf{x}}_{ei} = \mathbf{A}_{mi}\mathbf{x}_{ei} - \mathbf{\Gamma}_i \mathbf{\Phi}_i^T \mathbf{Z}_i - \mathbf{\Gamma}_i \mathbf{\Psi}_i^T \mathbf{Z}_i \quad (38)$$

It then follows that equation (38) defines a hyperstable system if (Isermann, 1992):

$$\alpha_i b_{pini} > 0 \quad (39)$$

$$\beta_i b_{pini} \geq 0 \quad (40)$$

and if the system parameter variations are slow compared with the speed of the individual adaptive control loops.
In summary, if the disturbances $\mathbf{d}_i$ and the coupling terms are slowly varying, then the errors are globally asymptotically stable.

## 4. Attitude control of flexible spacecraft

Many spacecrafts have large solar panel arrays with significant structural flexibility. The structural mode interaction with attitude control system has been one of the primary concerns for the design of 3-axis stabilized spacecraft (Wie & Plescia, 1984).
The main objective of the control system is to control the attitude of the spacecraft, which includes the contributions of both the rigid-body modes and the elastic modes. Considerable volume of literature (Metzger, 1979; Breakwell, 1981; Skaar et al., 1986) on flexible spacecraft focused on controlling only elastic modes, with the premise that control of rigid modes is straightforward and can be dealt with separately. This approach suffers from the fact that rigid modes are coupled with the elastic modes through control inputs and sensor observations, and cannot be separated from the elastic modes for controller design (Joshi et al., 1995).
The very simple flexible spacecraft configuration is shown in figure (2), where
x, y, z = principal axes of inertia of the undeformed spacecraft,
X, Y, Z = axes through the spacecraft center of mass, fixed in space,
$\theta_x$, $\theta_y$, $\theta_z$ = Euler angles: pitch, roll, and yaw.

### 4.1 Modeling of the flexible spacecraft

Flexible arrays are modeled as a mass-spring-dashpot system mounted on a rigid massless rod attached to the main body of the spacecraft via a coil spring to simulate torsional effects. The spacecraft and its flexible solar arrays are modeled as a rigid central body with each solar array represented as a point mass with two degrees of freedom, displacements in the roll-yaw plane, and by a disc having torsional rotation only about the pitch array axis (x-axis). The disc is circular spring whose mass is negligible compared with the spacecraft's body mass, as shown in figure (3). $K_x$, $K_y$, and $K_z$ are the spring constants and $C_x$, $C_y$, and $C_z$ are the damping factors. The coordinates $\alpha_y$, $\alpha_z$, $\beta_y$, and $\beta_z$ (deflection) and $\alpha_x$ and $\beta_x$ (rotation), which describe the position of movable parts with respect to the main body (Van Woerkom, 1985).

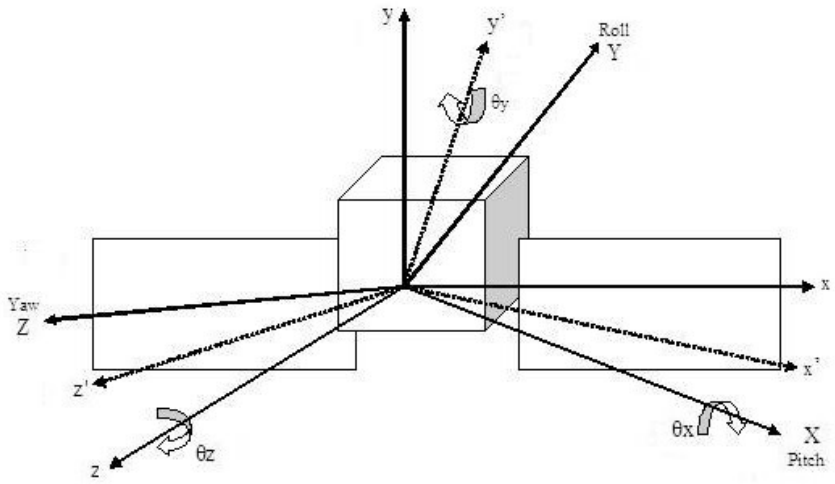



Fig. 2. Configuration of Flexible Spacecraft

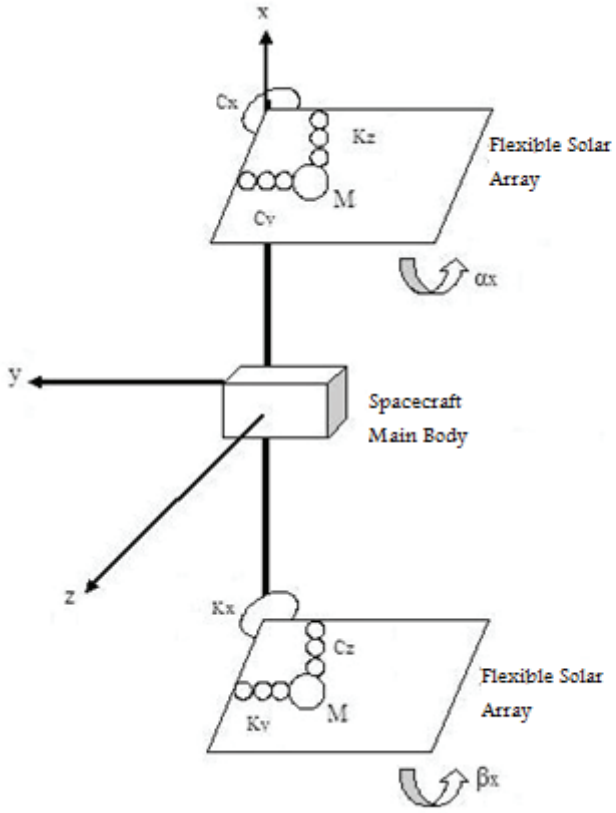



Fig. 3. Dynamic Model of Flexible Spacecraft

By having:
M = value of the point mass,
L = distance of the point mass from the spacecraft center of mass,
$I_f$ = moment of inertia of the movable material frame with respect to the x-axis,
the moments of inertia of the undeformed spacecraft are:

$$I_{xx} = I_x + 2I_f \tag{41}$$

$$I_{yy} = I_y + 2ML^2 \tag{42-a}$$

$$I_{zz} = I_z + 2ML^2 \tag{42-b}$$

where $I_x$, $I_y$, and $I_z$ are the principal moments of inertia of the rigid body. Assuming small deflections/rotations of the appendages (i.e. $\alpha_y$, …, $\beta_x$ small). Also, assuming that the spacecraft main body angular velocity components referred to the principal axes of inertia ($\Omega_x$, $\Omega_y$, and $\Omega_z$) are very small, therefore neglecting their products.
Then the dynamics of the flexible spacecraft become (Lorenzo, 1975):

$$I_{xx}\dot{\Omega}_x(t) + I_f(\ddot{\alpha}_x(t) + \ddot{\beta}_x(t)) = T_x(t) \tag{43-a}$$

$$I_{yy}\dot{\Omega}_y(t) + ML(\ddot{\beta}_z(t) - \ddot{\alpha}_z(t)) = T_y(t) \tag{43-b}$$

$$I_{zz}\dot{\Omega}_z(t) + ML(\ddot{\alpha}_y(t) - \ddot{\beta}_y(t)) = T_z(t) \tag{43-c}$$

$$\ddot{\alpha}_x(t) + \frac{C_x}{I_f}\dot{\alpha}_x(t) + \frac{K_x}{I_f}\alpha_x(t) = -\dot{\Omega}_x(t) \tag{43-d}$$

$$\ddot{\beta}_x(t) + \frac{C_x}{I_f}\dot{\beta}_x(t) + \frac{K_x}{I_f}\beta_x(t) = -\dot{\Omega}_x(t) \tag{43-e}$$

$$\ddot{\alpha}_y(t) + \frac{C_y}{M}\dot{\alpha}_y(t) + \frac{K_y}{M}\alpha_y(t) = -\dot{\Omega}_z(t)L \tag{43-f}$$

$$\ddot{\beta}_y(t) + \frac{C_y}{M}\dot{\beta}_y(t) + \frac{K_y}{M}\beta_y(t) = \dot{\Omega}_z(t)L \tag{43-g}$$

$$\ddot{\alpha}_z(t) + \frac{C_z}{M}\dot{\alpha}_z(t) + \frac{K_z}{M}\alpha_z(t) = \dot{\Omega}_y(t)L \tag{43-h}$$

$$\ddot{\beta}_z(t) + \frac{C_z}{M}\dot{\beta}_z(t) + \frac{K_z}{M}\beta_z(t) = -\dot{\Omega}_y(t)L \tag{43-i}$$

where $T_x(t)$, $T_y(t)$. and $T_z(t)$ are the torques applied to the spacecraft main body. Moreover, if:

$$f_x = \alpha_x + \beta_x$$

$$f_y = \alpha_y - \beta_y$$

$$f_z = -\alpha_z + \beta_z$$

the dynamics of the flexible spacecraft are described by the following decoupled set of equations (Lorenzo, 1975):

$$I_{xx}\dot{\Omega}_x(t) + I_f\ddot{f}_x(t) = T_x(t) \tag{44-a}$$

$$I_f\ddot{f}_x(t) + C_x\dot{f}_x(t) + K_x f_x(t) = -2I_f\dot{\Omega}_x(t) \quad (44\text{-b})$$

$$I_{yy}\dot{\Omega}_y(t) + ML\ddot{f}_z(t) = T_y(t) \quad (44\text{-c})$$

$$M\ddot{f}_z(t) + C_z\dot{f}_z(t) + K_z f_z(t) = -2ML\dot{\Omega}_y(t) \quad (44\text{-d})$$

$$I_{zz}\dot{\Omega}_z(t) + ML\ddot{f}_y(t) = T_z(t) \quad (44\text{-e})$$

$$M\ddot{f}_y(t) + C_y\dot{f}_y(t) + K_y f_y(t) = -2ML\dot{\Omega}_z(t) \quad (44\text{-f})$$

### 4.2 Kinematics

The attitude motion of a spacecraft is basically represented by a set of two equations: (i) Euler's dynamic equation, which describes the time evolution of the angular velocity vector, and (ii) the kinematic equation, which relates the time derivatives of the orientation angles to the angular velocity vector. Several kinematic parameterizations exist to represent the orientation angles, including singular, three-parameter representations (e.g., the Euler angles, Gibbs vector) (Costic et al, 2000). Four-parameter attitude representations such as quaternions avoid the problem of singular points and have better numerical properties than more conventional three-parameter representations (Kristiansen et al, 2009).

The use of a quaternion $(q_o, q_x, q_y, q_z)^T$ in describing the orientation of a rigid body lend themselves well to calculation with aid of an onboard computer since only products and no goniometric relations (which arises in using the Euler angles) exist in the formula. The quaternion equation yields (Wen & Kruetz, 1991):

$$\begin{bmatrix} \dot{q}_o \\ \dot{q}_x \\ \dot{q}_y \\ \dot{q}_z \end{bmatrix} = 0.5 \begin{bmatrix} 0 & -\Omega_x & -\Omega_y & -\Omega_z \\ \Omega_x & 0 & \Omega_z & -\Omega_y \\ \Omega_y & -\Omega_z & 0 & \Omega_x \\ \Omega_z & \Omega_y & -\Omega_x & 0 \end{bmatrix} \begin{bmatrix} q_o \\ q_x \\ q_y \\ q_z \end{bmatrix} \quad (45)$$

With

$$q_o^2 + q_x^2 + q_y^2 + q_z^2 = 1 \quad (46)$$

Four variables $q_o$, $q_x$, $q_y$ and $q_z$, coupled by means of Equation (46), uniquely describe the orientation of the spacecraft in space.

### 4.3 Reference model

The angular velocity $\boldsymbol{\Omega}(t)$ of the spacecraft is measured with gyros. Its orientation $\mathbf{q}(t)$ is calculated via the quaternion (45). Consequently, the state of the spacecraft ($\boldsymbol{\Omega}$ and $\mathbf{q}$) is accurately known. In this work, a model reference adaptive controller is proposed to realize a fast, three axes slew about the Euler axis, even in the presence of parameter variations inside the spacecraft and internal (friction of the reaction wheel) and external disturbances.

The reference model is used for realizing one single three axes slew. It calculates the trajectory in space from the present to the desired orientation. The spacecraft is forced to follow this trajectory so that it will perform a three axes slew as well and it will reach its desired attitude in space. The proposed reference model is selected to be a linearized, decoupled model of the spacecraft without disturbance and gyroscopic coupling; i.e. the reference model will exhibit an ideal trajectory.

The proposed control signal of the reference model, $u_m(t)$, is derived from the corresponding quaternion $q_m(t)$ and the angular velocity $\Omega_m(t)$ for each axis separately, namely:

$$u_m(t) = -k_{m1}q_m(t) - k_{m2}\Omega_m(t) \tag{47}$$

where $k_{m1}$ and $k_{m2}$ are constant feedback gains. This control strategy guarantees zero steady-state error in the model response. The reference model will always reach the desired attitude in space by one single three axes slew. Before a new slew can be made, a new reference frame has to be calculated, such that the required orientation will be the origin of this new reference frame. So, at the end of each slew $q_{mx} = q_{my} = q_{mz} = 0$ and $q_{mo} = 1$. In general, $q_{mo}$ will be large compared with $q_{mx}$, $q_{my}$, and $q_{mz}$. Therefore, the quaternion (45) can be linearized and decoupled. Taking the x-axis, for example, then:

$$\dot{q}_{mx}(t) = 0.5\Omega_{mx}(t) \tag{48}$$

This approximation, by which the model of one axis becomes linear and decoupled, allows the gains $k_{m1x}$ and $k_{m2x}$ to be calculated analytically. The authors in (Van Den Bosch et al., 1986) have chosen the undamped natural frequency $w_n$ and the relative damping ratio $\xi$ as design parameters, such that the feedback gains are calculated as follows:

$$k_{m1x}(t) = \frac{w_n^2 I_{mx}}{g_{mx}} \tag{49}$$

and

$$k_{m2x}(t) = \frac{2\xi w_n I_{mx}}{g_{mx}} \tag{50}$$

where $I_{mx}$ is the moment of inertia of the model x-axis and $g_{mx}$ is the gain of the model reaction wheels. This idea is adopted in this work.

### 4.4 Control with reaction wheels

Reaction wheels are momentum exchange devices which provide reaction torque to a spacecraft and store angular momentum. When reaction wheels are used to provide momentum H with respect to the body axes, the torques exerted on the spacecraft are (Wie et al. 1985):

$$T_x(t) = -\dot{H}_x(t) \tag{51-a}$$

$$T_y(t) = -\dot{H}_y(t) - \Omega_z(t)H_o \tag{51-b}$$

$$T_z(t) = -\dot{H}_z(t) + \Omega_y(t)H_o \tag{51-c}$$

Where $H_x(t)$, $H_y(t)$, and $H_z(t)$ are the components of H along the body axes. $H_o$ is the nominal value of $H_x(t)$, the largest component of H. If equations (51) are substituted into Equations (44), the set of equations describing the dynamics of the flexible spacecraft may be written as (Lorenzo, 1975; Wie et al. 1985):

$$I_{xx}\dot{\Omega}_x(t) + I_f\ddot{\eta}_x(t) = -\dot{H}_x(t) \tag{52-a}$$

$$I_f\ddot{f}_x(t) + C_x\dot{f}_x(t) + K_x f_x(t) + 2I_f\dot{\Omega}_x(t) = 0 \quad (52\text{-b})$$

$$I_{yy}\dot{\Omega}_y(t) + H_o\Omega_z(t) + ML\ddot{f}_z(t) = -\dot{H}_y(t) \quad (53\text{-a})$$

$$M\ddot{f}_z(t) + C_z\dot{f}_z(t) + K_z f_z(t) + 2ML\dot{\Omega}_y(t) = 0 \quad (53\text{-b})$$

$$I_{zz}\dot{\Omega}_z(t) - H_o\Omega_y(t) + ML\ddot{f}_y(t) = -\dot{H}_z(t) \quad (53\text{-c})$$

$$M\ddot{f}_y(t) + C_y\dot{f}_y(t) + K_y f_y(t) + 2ML\dot{\Omega}_z(t) = 0 \quad (53\text{-d})$$

The system is broken into two linear systems. The first (Equations (52)), describes the behavior around the x axis only. The second (Equations (53)), describes the behavior of the coupled y and z axes.

For simplicity, the usual practice is to consider decoupled actuating torques (Lorenzo, 1975; Wie et al. 1985). This means that there are no interferences among the spacecraft axes due to the rotation of each reaction wheel, so that Equations (52) and (53) which represent the dynamics of the flexible spacecraft remain linear. But, practically, this is not true because each reaction wheel, rotating about one axis, will cause gyroscopic coupling among the other two axes.

In this work, the dynamics of flexible spacecraft is taken coupled and nonlinear. Although this will complicate the control algorithm but at the same time it will further enhance the performance of the attitude control system.

The complete set of equations used here to describe the flexible spacecraft is:

$$\dot{q}_o(t) = 0.5(-\Omega_x(t)q_x(t) - \Omega_y(t)q_y(t) - \Omega_z(t)q_z(t)) \quad (54)$$

$$\dot{q}_x(t) = 0.5(+\Omega_x(t)q_o(t) + \Omega_z(t)q_y(t) - \Omega_y(t)q_z(t)) \quad (55)$$

$$\dot{q}_y(t) = 0.5(+\Omega_y(t)q_o(t) - \Omega_z(t)q_x(t) + \Omega_x(t)q_z(t)) \quad (56)$$

$$\dot{q}_z(t) = 0.5(+\Omega_z(t)q_o(t) + \Omega_y(t)q_x(t) - \Omega_x(t)q_y(t)) \quad (57)$$

$$\dot{\Omega}_x(t) = \frac{C_x}{I_x}\dot{f}_x(t) + \frac{K_x}{I_x}f_x(t) + \frac{g_x}{I_x}u_{px}(t) - \frac{I_{zz}\Omega_z(t) + J_z w_z(t)}{I_x}\Omega_y(t) + \frac{I_{yy}\Omega_y(t) + J_y w_y(t)}{I_x}\Omega_z(t) \quad (58)$$

$$\dot{\Omega}_y(t) = \frac{C_z L}{I_y}\dot{f}_z(t) + \frac{K_z L}{I_y}f_z(t) - \frac{H_o}{I_y}\Omega_z(t) + \frac{g_y}{I_y}u_{py}(t) + \frac{I_{zz}\Omega_z(t) + J_z w_z(t)}{I_y}\Omega_x(t) - \frac{I_{xx}\Omega_x(t) + J_x w_x(t)}{I_y}\Omega_z(t) \quad (59)$$

$$\dot{\Omega}_z(t) = \frac{C_y L}{I_z}\dot{f}_y(t) + \frac{K_y L}{I_z}f_y(t) + \frac{H_o}{I_z}\Omega_y(t) + \frac{g_z}{I_z}u_{pz}(t) - \frac{I_{yy}\Omega_y(t) + J_y w_y(t)}{I_z}\Omega_x(t) + \frac{I_{xx}\Omega_x(t) + J_x w_x(t)}{I_z}\Omega_y(t) \quad (60)$$

$$\ddot{f}_x(t) = -\frac{C_x I_{xx}}{I_f I_x}\dot{f}_x(t) - \frac{K_x I_{xx}}{I_f I_x}f_x(t) - 2\frac{g_x}{I_x}u_{px}(t) + 2\frac{I_{zz}\Omega_z(t) + J_z w_z(t)}{I_x}\Omega_y(t) - 2\frac{I_{yy}\Omega_y(t) + J_y w_y(t)}{I_x}\Omega_z(t) \quad (61)$$

$$\ddot{f}_y(t) = -\frac{K_y I_{zz}}{MI_z}\dot{f}_y(t) - \frac{K_y I_{zz}}{MI_z}f_y(t) - 2L\frac{H_o}{MI_z}\Omega_y(t) - 2L\frac{g_z}{I_z}u_{pz}(t) + 2L\frac{I_{yy}\dot{\Omega}_y(t) + J_y\dot{w}_y(t)}{I_z}\Omega_x(t) - 2L\frac{I_{xx}\dot{\Omega}_x(t) + J_x\dot{w}_x(t)}{I_z}\Omega_y(t) \quad (62)$$

$$\ddot{f}_z(t) = -\frac{K_z I_{yy}}{MI_y}\dot{f}_z(t) - \frac{K_z I_{yy}}{MI_y}f_z(t) + 2L\frac{H_o}{MI_y}\Omega_z(t) - 2L\frac{g_y}{I_y}u_{py}(t) - 2L\frac{I_{zz}\dot{\Omega}_z(t) + J_z\dot{w}_z(t)}{I_y}\Omega_x(t) + 2L\frac{I_{xx}\dot{\Omega}_x(t) + J_x\dot{w}_x(t)}{I_y}\Omega_z(t) \quad (63)$$

where J is the moment of inertia of each reaction wheel; $w_x$, $w_y$, and $w_z$ are the angular velocities of the reaction wheels; g is the gain of each reaction wheel; and $u_{px}$, $u_{py}$ and $u_{pz}$ are the control signals for each axis.

Equations (54) to (57) represent the kinematics of the spacecraft. Equations (58) to (60) represent the dynamics of flexible spacecraft main body. While equations (61) to (63) represent the dynamics of the solar arrays.

The proposed model of the three axes of the reference model is given by:

$$\dot{q}_m(t) = 0.5\Omega_m(t) \quad (64)$$

$$\dot{\Omega}_m(t) = -\frac{K_{m1}g_m}{I_m}q_m(t) - \frac{K_{m2}g_m}{I_m}\Omega_m(t) \quad (65)$$

for each axis individually, i.e. linearized and decoupled.

In this work, a decentralized adaptive control strategy which uses a threefold single axis reference model is proposed for controlling the attitude of flexible spacecraft.

## 5. Simulation results

To test the decentralized minimal controller synthesis algorithm, the following data of a flexible satellite are simulated for a large-angle slew of (0°, 60°, 60°) (Lorenzo, 1975):

$I_x$ = 250 kgm²; $I_y$ = 1500 kgm²; $I_z$ = 1800 kgm²; $I_f$ = 25 kgm²; $C_x$ = 1 kgm²rad/s; $C_y$ = 1.75 kg/s; $C_z$ = 1.75 kg/s; $K_x$ = 1 kgm²rad/s²; $K_y$ = 2.8 kg/s²; $K_z$ = 2.8 kg/s²; L = 10 m; M = 70 kg; $H_o$ = 100 Nms; J=$J_m$= 0.072 kgm²; g=$g_m$= 0.08 Nm/A; $\boldsymbol{\Omega}(0)=[0, 0, 0]^T$ (rad/s); $\mathbf{w}(0) = [-70, 10, -30]^T$ (rad/s); $\xi$ = 0.95; $w_n$ = 0.035 rad/s.

Even with the high nonlinearity of the presented decentralized adaptive controller, the simulation results show that this controller is extremely efficient when it's implemented for maneuvering a flexible satellite. The angular displacement response is shown in figure (4). The problem of existence of oscillations in the transient response of the flexible satellite, which causes the excitation of the solar arrays, is usually encountered in the literature (Marchal, 1983; Van Den Bosch, 1986; Singh, 1988). But, in this work, due to the robustness of the presented adaptive algorithm this problem is avoided and the solar arrays exhibit only few oscillations during the maneuvering of the flexible satellite as shown in figure (5).

## 6. Conclusion

A model reference adaptive control methodology is presented by using hyperstability theory. The main contribution of this work effort lies in the use of decentralized minimal controller synthesis algorithm to guarantee the overall asymptotic stability and to highly enhance the robustness of the attitude control system of flexible spacecraft against disturbances.

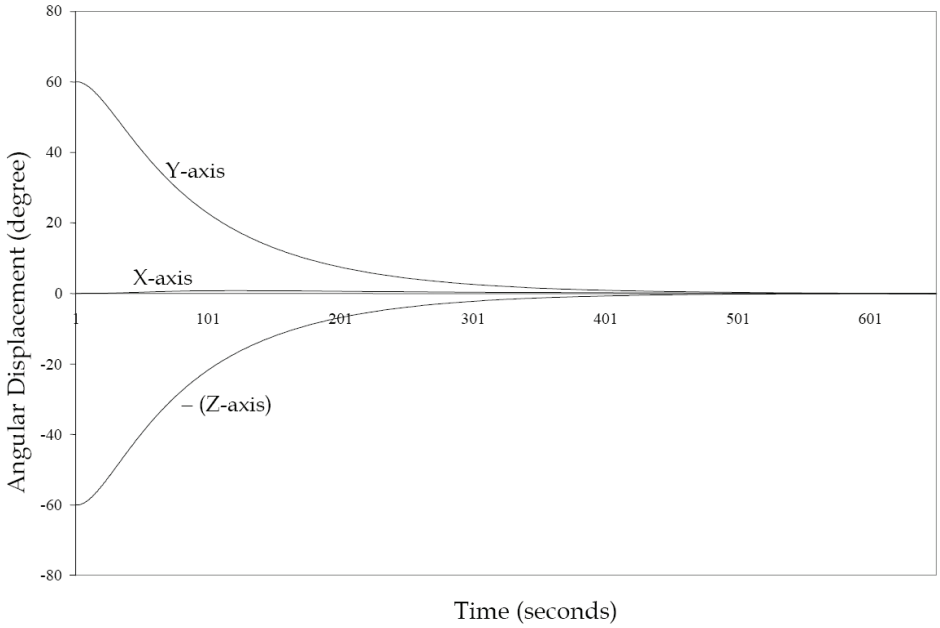



Fig. 4. Angular Displacement Response of Main Body

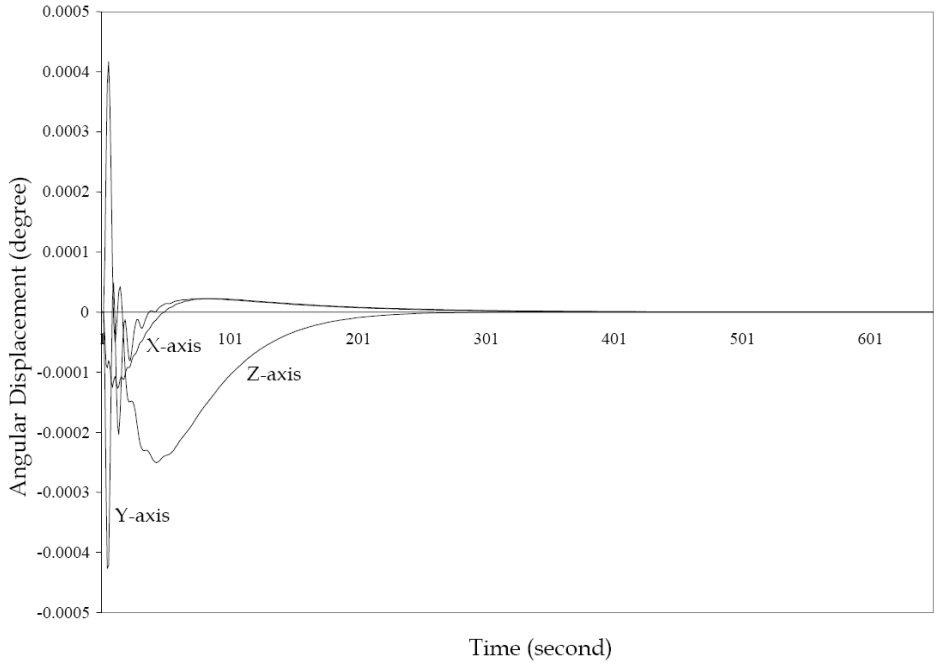



Fig. 5. Angular Displacement Response of Solar Array

It is concluded that for a large-scale control system whose dynamics are coupled and nonlinear, the recommendation is to use the decentralized adaptive control schemes for controlling such system since they give better performance than that of centralized adaptive control schemes.
The dynamics of the main body of the flexible spacecraft are usually considered, in some literature, separately from that of the flexible arrays. In other literature, the torques exerted by the reaction wheels is taken decoupled for simplicity. In this work, the dynamics of the main body and the solar arrays are considered as coupled. Moreover, the reaction wheels' torques is taken coupled, too. This leads to the complexity of the attitude control system algorithm because the flexible spacecraft dynamics become highly coupled and nonlinear. Nevertheless, the performance of the attitude control system is highly enhanced.
The application of the decentralized minimal controller synthesis algorithm for the attitude control of flexible spacecraft gives excellent results concerning the behavior of the solar arrays during the maneuvering of the spacecraft.

## 7. References


Arif, T. (2008). Adaptive Control of Rigid Body Satellite, *International Journal of Automation and Computing*, Vol. 5, No. 3, pp. 296-306, ISSN: 1476-8186.

Benchoubane, H. & Stoten, D. (1992). The Decentralized Minimal Controller Synthesis Algorithm, *International Journal of Control*, Vol. 56, No. 4, pp. 867-983, ISSN: 0020-7179.

Breakwell, J. (1981). Optimal Control Slewing of Flexible Spacecraft, *Journal of Guidance and Control*, Vol. 4, No. 4, pp. 472-479, ISSN 0731-5090.

Campbell, B. & Lieven, N. (2002). Fault Tolerant Control Using Minimal Controller Synthesis for the X-38 Crew Return Vehicle. *Proceedings of the 5th ESA International Conference on Spacecraft Guidance, Navigation and Control Systems*, pp. 185-190, ISBN: 92-9092-826-3, Italy, October 2002, Frascati.

Costic, B.; Dawson, D.; de Queiroz & Kapila, V. (2000). A Quaternion-Based Adaptive Attitude Tracking Controller without Velocity Measurements, *Proceedings of the 39th IEEE Conference on Decision and Control*, Vol. 3, pp. 2424-2429, ISBN 0-7803-6638-7, Australia, December 2000, IEEE Press, Sydney.

Datta, A. (1993). Performance Improvement in Decentralized Adaptive Control: A Modified Reference Scheme, *IEEE Transactions on Automatic Control*, Vol. 38, No. 11, pp. 1717-1722, ISSN: 0018-9286.

Di Bernardo, M. & Stoten, D. (2006). Minimal Controller Synthesis Adaptive Control of Nonlinear Systems: Utilizing the Properties of Chaos, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 364, No. 1846, pp. 2397-2415, ISSN: 1472-2962.

Isermann, R. (1992). *Adaptive Control Systems*, Prentice Hall, ISBN: 0131374560, New York, USA.

Joshi, S.; Maghami P. & Kelkar, A. (1995). Design of Dynamic Dissipative Compensators for Flexible Space Structures, *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 31, No. 4, pp. 1314-1323, ISSN: 0018-9251.

Kristiansen, R.; Nicklasson, J. & Gravdahl, J. (2009). Satellite Attitude Control by Quaternion-Based Backstepping, *IEEE Transactions on Control Systems Technology*, Vol. 17, No. 1, pp. 227-232, ISSN: 1036-6536.

Landau, I. & Courtiol, B. (1974). Survey of Model Reference Adaptive Techniques – Theory and Applications, *Automatica*, Vol. 10, No. 4, pp. 353-379, ISSN: 0005-1098.

Landau, I. (1979). *Adaptive Control: the Model Reference Approach*, Marcel Dekker, ISBN: 0824765484, New York, USA.

Lorenzo, R.; Santinelli, A. & Sciacovelli, D. (1975). Some Attitude Control Synthesis for Satellites with Flexible Appendages, *Automatica*, Vol. 11, pp. 161-170, ISSN: 0005-1098.

Luo, W.; Chu, Y. & Ling, K. (2005). Inverse Optimal Adaptive Control for Attitude Tracking of Spacecraft, *IEEE Transactions on Automatic Control*, Vol. 50, No. 11, pp. 1639-1654, ISSN: 0018-9286.

Marchal, C. (1983). Minimum Horizontal Rotations, *Optimal Control Applications & Methods*, Vol. 4, pp. 357-363, ISSN: 0143-2087.

Metzger, R. (1979). A Simple Criterium for Satellite with Flexible Appendages, *Proceedings of the 8th IFAC Symposium on Automatic Control in Space*, pp. 63-69, ISBN: 0080244491, England, July 1979, Oxford.

Pande, K. & Ventachalam, R. (1982). Optimal Solar Pressure Attitude Control of Spacecraft, *Acta Astronautica*, Vol. 9, No. 9, pp.533-545, ISSN: 0094-4765.

Singh, S. (1988). Flexible Spacecraft Maneuver: Inverse Attitude Control and Modal Stabilization, *Acta Astronautica*, Vol. 17, No. 1, pp. 1-9, ISSN: 0094-4765.

Skaar, S.; Tang L. & Yalda, I. (1986). On-Off Attitude Control of Flexible Satellites, *Journal of Guidance*, Vol. 9, No. 4, pp. 507-510, ISSN: 0731-5090.

Spooner, T. & Passino, K. (1996). Adaptive Control of a Class of Decentralized Nonlinear Systems, *IEEE Transactions on Automatic Control*, Vol. 41, No. 2, pp. 280-284, ISSN: 0018-9286.

Stoten, D. & Benchoubane, H. (1990a). Empirical Studies of an MRAC Algorithm with Minimal Controller Synthesis, *International Journal of Control*, Vol. 51, No. 4, pp. 823-849, ISSN: 0020-7179.

Stoten, D. & Benchoubane, H. (1990b). Robustness of a Minimal Controller Synthesis Algorithm, *International Journal of Control*, Vol. 51, No. 4, pp. 851-861, ISSN: 0020-7179.

Stoten, D. & Hodgson, S. (1992). The Application of the Minimal Control Synthesis Algorithm to the Hybrid Control of Class 1 Manipulator, *International Journal of Control*, Vol. 56, No. 3, pp. 499-513, ISSN: 0020-7179.

Van Den Bosch, P.; Jongkind, W. & Swieten, A. (1986). Adaptive Attitude Control for Large-Angle Slew Maneuvers, *Automatica*, Vol. 22, No. 2, pp. 209-215, ISSN: 0005-1098.

Van Woerkom, P. (1985). Mathematical Models of Flexible Spacecraft Dynamics: A Survey of Order Reduction Approaches, *Proceedings of the 10th IFAC Symposium on Automatic Control in Space*, pp. 127-135, France, June 1985, Toulouse.

Wagg, D. & Stoten, D. (2001). Substructuring of Dynamical Systems via the Adaptive Minimal Controller Synthesis Algorithm, *Earthquake Engineering & Structural Dynamics*, Vol. 30, No. 6, pp. 865-877, ISSN: 0098-8847.

Wen, J. & Kreutz-Delgado, K. (1991). The Attitude Control Problem, *IEEE Transactions on Automatic Control*, Vol. 36, No. 10, pp. 1148-1161, ISSN: 0020-7179.

Wertz, J. (1980). *Spacecraft Attitude Determination and Control*, D. Reidel, ISBN: 9027712042, Boston, MA, USA.

Wie, B. & Plescia, C. (1984). Attitude Stabilization of Flexible Spacecraft during Stationkeeping Maneuvers, *Journal of Guidance*, Vol. 7, No. 4, pp. 430-436, ISSN: 0731-5090.

Wie, B.; Lehner, J. & Plescia, C. (1985). Roll/Yaw Control of a Flexible Spacecraft Using Skewed Bias Momentum Wheels, *Journal of Guidance*, Vol. 8, No. 4, pp. 447-453, ISSN: 0731-5090.

# 5

# Advancing NASA's On-Board Processing Capabilities with Reconfigurable FPGA Technologies

Paula J. Pingree
*Jet Propulsion Laboratory, California Institute of Technology*
*United States of America*

## 1. Introduction

Future NASA missions will require measurements from high data rate active and passive instruments. Recent internal studies at NASA's Jet Propulstion Laboratory (JPL) estimate approximately 1–5 Terabytes per day of raw data (uncompressed) are expected, for example, from spectroscopy instruments. Implementations of on–board processing algorithms to perform lossless data reduction are required to drastically reduce data volumes to within the downlink capabilities of the spacecraft and existing ground stations. Reconfigurable Field Programmable Gate Arrays (FPGAs) such as the Xilinx™ Virtex–4 and Virtex–5 series devices can include dual core PowerPC processors thereby providing a flexible hardware and software co–design architecture to meet the on–board processing challenges of these missions while reducing the essential resources of mass and volume of earlier generation flight–qualified computing platforms such as the BAE Rad750 single board computer (SBC). Reconfigurable FPGAs also offer unique advantages over one–time programmable (OTP) FPGAs with flexible prototype development platforms that provide an important "path–to–flight" for spaceborne instruments. Reconfigurable FPGA technologies also provide in–flight flexibility with the ability to update processing algorithms as needed post–launch.

This chapter will discuss these comparative technologies and present the benefits of commercially available FPGA development platforms from Xilinx for the development of NASA's future on-board processing capabilities. Additionally, commercially available tools such as Impulse C™ have been used to adapt legacy C–code into Verilog or VHDL for implementation in FPGA fabric to achieve hardware acceleration. Key features and requirements of future NASA missions proposed within the National Research Council's Decadal Survey will be described with ideas on how these reconfigurable FPGA technologies and development tools can combine to achieve breakthrough on–board processing performance to meet their science objectives. To provide specific demonstrations of these ideas, three unique and recent design implementations on the Xilinx V4FX60 and Virtex–5 FPGAs targeted to enable future NASA missions will be presented. They include on–board processing algorithms for a) Support Vector Machine (SVM) Classifiers similar to those in operation on the Earth Observing 1 (EO-1) Hyperion instrument, b) a Fourier transform infrared (FTIR) spectrometer, and c) a new Multiangle Spectropolarimetric

Imager (MSPI). Finally, the important issue of radiation effects from single event upsets (SEUs) and necessary mitigation schemes for SRAM–based reconfigurable FPGA devices will be addressed particularly as it depends on science data return requirements for both Earth and planetary missions.

## 2. Earth science objectives & the need for advanced on-board processing

Many of the missions proposed within the National Research Council's Decadal Survey (NRC, 2007) will require measurements from high data rate active and passive instruments. Internal studies at JPL estimate approximately 1-5 Terabytes of expected raw (uncompressed) data per day from spectroscopy instruments. When lossless compression algorithms are applied, this volume can be reduced to 100s of Gigabytes per day, but many of the Decadal Survey missions actually carry multiple instruments. Implementations of on-board processing algorithms to perform science data reductions can dramatically reduce the data rates, but there are additional advantages. An on-board processing platform of sufficient capability can simplify the instrument design itself leading to reductions in mass, power, and volume for missions where multiple instrument components necessitate optimization of these resources. In this way, strategic application of information technology advances can lead to measurable improvements in both instrument data reduction and design. This chapter will highlight two Decadal Survey Missions for which reconfigurable FPGA technologies are targeted to meet the on-board processing requirements of the key instrument payloads. Both the Geostationary Coastal and Air Pollution Events (GEO-CAPE) and Aerosol-Cloud-Ecosystem (ACE) missions are recommended for launch in the 2013-2016 timeframe.

The GEO-CAPE mission will analyze atmospheric gas columns for air quality forecasts and analyze ocean color for coastal ecosystem health and climate emissions (Sander et al., 2008). The key instruments for GEO-CAPE are a high-spatial resolution hyperspectral spectrometer, a low-spatial resolution imaging spectrometer and an IR correlation radiometer. The Panchromatic Fourier Transform Spectrometer (PanFTS), a new project within the NASA Instrument Incubator Program, will measure all of the trace species called out in the Decadal Survey for GEO-CAPE. With continuous sensitivity from 0.26 to 15 micron and high spectral resolution, PanFTS combines the functionality of separate UV, visible and IR instruments in a single package. These capabilities also permit PanFTS to meet the requirements for high spatial resolution hyperspectral imaging of the coastal zone. Two key enabling technologies under development for PanFTS are high-speed, high-dynamic range CMOS hybrid focal plane arrays (FPAs), and parallel, co-aligned optical trains for the ultraviolet-Visible-Near-infrared (UV-Vis-NIR) and mid-IR bands. Section 4 of this chapter describes the role of reconfigurable FPGAs in the development of the PanFTS instrument to meet the on-board processing requirements for the GEO-CAPE mission.

The ACE mission will analyze aerosol and cloud profiles for climate and water cycles and analyze ocean color for open ocean biogeochemistry. The Decadal Survey identifies a multi-angle, multispectral, high-accuracy polarization imager as a key requirement for the ACE mission. JPL has been developing a Multiangle SpectroPolarimetric Imager (MSPI) as a candidate to fill this need. Using passive remote sensing to determine aerosol and cloud microphysical parameters necessitates a variety of constraints to overcome retrieval non-uniqueness, and places stringent design requirements on a cohesive, next-generation passive satellite imager. Fusion of ultraviolet (UV) to SWIR multispectral, multiangular, and polarimetric observations is important because each brings sensitivity to different aspects of

particle spatial distributions and microphysics (Yu et al., 2006; Kahn et al., 2004). UV wavelengths are sensitive to aerosol absorption and height, visible/near-infrared (VNIR) wavelengths to fine mode aerosol size distributions, and SWIR wavelengths to coarse mode aerosol and cloud particle sizes. Multiangular radiances distinguish particle shapes, and improve sensitivity to optical depth, notably over bright surfaces. Polarimetry is the only means of determining particle real refractive index, providing compositional information. Moderately high spatial resolution resolves intercloud scales and aerosol gradients in urban settings. A broad swath is important for providing environmental context to aerosol and cloud spatial relationships, distinguishing airmass types, validating transport models, and overflying field campaigns and suborbital instruments. The technological design considerations of integrating all of these attributes into a single instrument are compounded by the need to acquire accurate multispectral intensity imagery (3% uncertainty) simultaneously with accurate degree of linear polarization (DOLP) imagery (0.5% uncertainty). The MSPI technology development plan has been systematically reducing risk in critical areas, with consequent direct benefit to the ACE mission. An essential part of this plan is on-board processing for digital signal handling. The role of reconfigurable FPGAs for MSPI on-board processing is discussed further in Section 4 of this chapter.

## 3. Comparative technologies

On-board computation has become a bottleneck for advanced science instrument and engineering capabilities. Currently available spacecraft processors have high power consumption, are expensive, require additional interface boards, and are limited in their computational capabilities. Recently developed hybrid FPGAs, such as the Xilinx Virtex-4FX60 and Virtex-5, offer the versatility of running diverse software applications on embedded processors while at the same time taking advantage of reconfigurable hardware resources, all on the same chip package. These tightly coupled hardware/software co-designed systems are lower power and lower cost than general-purpose SBCs, and promise breakthrough performance over radiation-hardened SBCs, leading to a new architecture for future spaceborne instrument development.

| Computational Platform | Performance (DMIPS) |
|---|---|
| RAD750 SBC | 240 |
| Xilinx Virtex-II Pro | 450 |
| Xilinx Virtex-4 | 680 |
| Xilinx Virtex-5 | > 1000 |

Table 1. Performance: SBC vs. Embedded FPGAs (http://www.xilinx.com)

The Rad750 SBC has flight-proven heritage serving as the primary spacecraft computer for JPL's 2005 Deep Impact mission (http://deepimpact.jpl.nasa.gov). For future science instruments, however, baselining a Rad750 SBC is likely to exceed the instrument's allocations for mass and volume, especially when there are multiple science instruments to be accommodated on a single spacecraft platform. In some cases a single Rad750 SBC may not even meet the science instruments requirements for on-board data processing as was the case for an instrument called MATMOS that was proposed for the 2007 Mars Scout mission

opportunity (See Section 4.2). MATMOS is a precursor instrument that led to the development of PanFTS targeted for the GEO-Cape mission. The on-board processing FPGA development started on MATMOS has been extended toward a baseline architecture for PanFTS using the Xilinx Virtex-5 FPGA with embedded PowerPC440 processors.
The advantage reconfigurable FPGAs offer in flexibility over OTP FPGAs in the instrument development and operations cycles can be illustrated by the following timeline. NASA's upcoming Juno mission to Jupiter is scheduled to launch in August 2011 and arrive at Jupiter in 2016 for a nominal one year mission orbiting the planet. A key instrument on the Juno spacecraft is the Microwave Radiometer (MWR), under development at JPL (Pingree et al., 2008-a). The MWR Electronics Unit contains a single OTP FPGA for the command and telemetry interface to the spacecraft as well as to control the instrument's six (6) receivers. To meet the MWR instrument development schedule, the final FPGA design was programmed and installed on the flight electronics board in June 2009. The implication is that 2+ years before the spacecraft is launched the MWR FPGA design cannot be modified, improved or corrected without re-working the flight board at a significant impact to the project's cost and schedule. Furthermore, there is an additional 5 years of cruise to Jupiter that will include MWR instrument calibration activities, and still the FPGA design remains fixed. Irregardless of potential in-flight findings where updating the FPGA may be desireable to improve the instrument's capabilities to better meet the science objectives of the mission, doing so is impossible with a OTP FPGA. This would not be the case with an SRAM-based reprogrammable FPGA such as the Xilinx Virtex-5. Especially with the future trend for increasing complexity in on-board processing requirements, it is particularly beneficial to have the flexibility to improve and optimize algorithms throughout the instrument development and operations cycles.
Other JPL missions have used Xilinx reprogrammable FPGAs. Four Xilinx XQR4062XL FPGA devices controlled the Mars Exploration Rovers (MER) lander pyrotechnics that were crucial to the Entry, Descent and Landing phase of the mission. The MER Spirit and Opportunity vehicles, still in operation after over 5 years on the surface of Mars, contain the Xilinx XQVR1000 FPGA in their motor controller boards. The Mars Reconnaissance Orbiter (MRO) Electra software radio uses the Xilinx VirtexII-8000 FPGA. Each of these devices, however are not hybrid FPGAs; they do not contain embedded PowerPC processors. The Xilinx hybrid FPGAs such as the V4FX60 and Virtex-5 showcased in the next Section's design cases do not have space-flight heritage and therefore represent the cutting edge of computing platforms required to meet the requirements of the Decadal Survey missions.

## 4. Recent on-board processing FPGA design cases

To provide specific demonstrations of the ideas and advantages of reprogrammable FPGAs that have been discussed, three unique and recent design implementations on the Xilinx V4FX60 and Virtex-5 FPGAs targeted to enable future NASA missions are presented. They include on-board processing algorithms for a) Support Vector Machine (SVM) Classifiers similar to those in operation on the EO-1 Hyperion instrument, b) a Fourier transform infrared (FTIR) spectrometer, and c) a new Multiangle Spectropolarimetric Imager (MSPI).

### 4.1 SVMs for hyperspectral classification

Fast and accurate on-board classification of image data is a critical part of modern satellite image processing. For Earth sciences and other applications, space-based payloads make use

of intelligent, machine-learning algorithms and instrument autonomy to detect and identify natural phenomena such as flooding, volcanic eruptions, and sea ice break-up. JPL has developed support vector machine (SVM) classification algorithms used on board spacecraft to identify high-priority image data for downlinking to Earth. These algorithms also provide onboard data analysis to enable rapid reaction to dynamic events. These onboard classifiers help reduce the amount of data downloaded to Earth, greatly increasing the science return of the instrument.

SVM classification algorithms are flying today using computational platforms such as the RAD6000 and Mongoose V processors; these SBCs are of an even older generation than the Rad750 discussed in Section 3 . These legacy processors have only limited computing power, extremely limited active storage capabilities, and are no longer considered state-of-the-art. For this reason, onboard classification has been limited to only the simplest functions running on only a subset of the full instrument data: for example, only 11 of 242 image bands are classified in the case of the Hyperion instrument on the EO-1 satellite. FPGA coprocessors are an ideal candidate for future implementations of these algorithms to provide significant improvement in onboard classification capability and accuracy when compared to the legacy processing platforms now flying.

SVMs are also well-suited to onboard autonomy applications. The property that makes SVMs particularly applicable is the asymmetry of computational effort in the training and testing stages of the algorithm. Classifying new data points requires orders of magnitude less computation than training because the process of training a SVM requires solving a quadratic optimization problem. SVM training requires $O(n^3)$ operations, where $n$ is the number of training examples. In contrast, testing a new vector with a trained SVM requires only O(n) operations. Faster algorithms that exploit the specific structure of the SVM optimization problem exist (Platt, 1998), but the training remains the primary computational bottleneck.

After a SVM is trained, many of the weights, $w_i$, will be equal to zero. This means that these terms can be ignored in the classification formula. Those input vectors that have a corresponding non-zero weight are called *support vectors.* Even more computational savings can be realized in the case of using a linear kernel function. The weighted sum over the kernel function is associative, so all the support vectors can be collapsed into a single vector with a single weight. Reducing the number of support vectors is key to successfully deploying a SVM classifier onboard a spacecraft where there are severe constraints on the available computational resources. Previously deployed classifiers (Castano et als., 2006) have used such reduced-set methods, but were still constrained to operate on only a subset of the available classification features. Removing such bottlenecks is critical to realizing the full potential of SVMs as an onboard autonomy tool.

In 2007, a small JPL team evaluated the effectiveness of FPGAs for SVM algorithms for a legacy snow-water-ice-land (SWIL) classifier, originally developed for the Hyperion instrument, on the Xilinx ML410 development board that includes the Virtex-4 FX60 FPGA (Pingree et al., 2008-b). For rapid development, the Impulse C™ autocode toolset, provided by Impulse Accelerated Technologies, was used to adapt legacy C-code into HDL[1] for implementation in the FPGA fabric. The Impulse C compiler generates synthesizable HDL files ready to use with the Xilinx EDK tools. In addition to generating HDL files, the Impulse

[1] High-level Design Language (HDL)

compiler also generates additional files required by the EDK tools, including PLB (Processor Local Bus) and APU (Auxiliary Processor Unit) bus interfaces. The Impulse C compiler performs a variety of low-level optimizations, including C statement scheduling and loop pipelining, saving application developers a great deal of time that would otherwise be spent performing tedious, low-level hardware optimization. The Impulse C compiler performs these optimizations and generates hardware in the form of either VHDL or Verilog. This hardware can then be synthesized using FPGA tools such as Xilinx ISE™ software and Platform Studio.
On the processor side, the compiler generates run-time libraries ready for use on the embedded PowerPC processor. While is is arguable that auto-code generation tools may not produce Verilog or VHDL that is as efficient as a programmer may develop directly by hand, but for already-existing legacy C-code algorithms that are intended for acceleration by FPGA implementation, the Impulse C tools can provide an efficient path toward rapid prototype development. This approach to implementing the Hyperion linear SVM on the Virtex-4 FX60 FPGA, as well as additional experiments that were performed using an increased number of data bands and a more sophisticated SVM kernel, provided rapid results to show the potential for more efficient, higher performance onboard classification using FPGA-embedded algorithms.
When using FPGAs with embedded processors, efficient partitioning of algorithms between software and hardware is important to achieve high performance. The image file input and classification file output are managed within the FPGA's embedded PowerPC processor. Figure 1 shows an example architecture for this FPGA-based implementation.

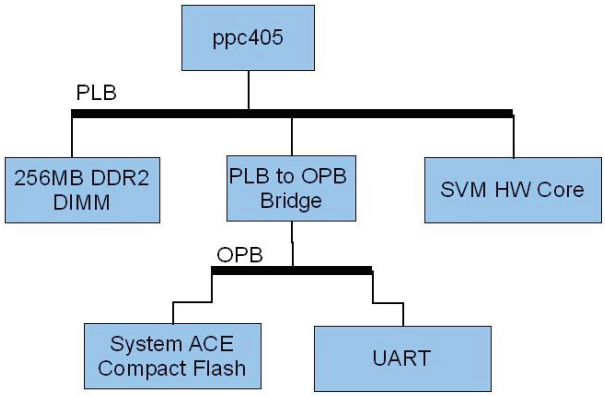



Fig. 1. SVM Classifier FPGA Architecture (Pingree et al., 2008-b).

In support of partitioned software/hardware applications such as this, the Impulse tools include a library of C-compatible functions that implement a number of process-to-process communication methods. These methods include streaming, shared memory, and message passing. For this application, the Impulse C streaming programming model was the obvious choice. In Impulse C streaming applications, hardware and software processes communicate primarily through buffered data streams implemented directly in hardware. This buffering of data makes it possible to write parallel applications at a relatively high level of abstraction without the cycle-by-cycle synchronization that would otherwise be required. Impulse C functions are used to open and close data streams, read or write data on the streams, and, if desired, send status messages or poll for results. In the case of the

Virtex-4FX, stream reads and writes can be specified as operations that take advantage of either the PLB or the auxiliary peripheral unit (APU) interface.

The software side of the SVM application is coded in C and compiled to the embedded PowerPC 405 processor using the Xilinx EDK tools. The embedded software application reads an input image file consisting of 857,856 pixels (based on actual Hyperion image size). The software-side application streams the image data to the SVM, which is also written in C but has been compiled (using the Impulse C-to-FPGA compiler) to FPGA hardware. The SVM hardware process performs the required SVM operation on the image and streams the results back to the PowerPC 405 processor. The processor then writes the pixel classifications (e.g., snow, water, ice, land, cloud, or unclassified) to an output file.

The Impulse C-converted HDL code is synthesized for the Virtex-4FX FPGA using the Xilinx ISE & EDK development environment. Synthesis of this SVM application results in the following resource utilization for the V4FX60 device on the ML410 development platform (Table 2). As shown, the FPGA can easily accommodate the algorithm that was severely constrained within the EO-1 Hyperion software implementation on the Mongoose processor.

| FPGA RESOURCES | V4FX60 (on ML410) |
|---|---|
| Number of Slices: | 1151 out of 25280 **(4%)** |
| Number of Slice Flip Flops: | 1290 out of 50,560 **(2%)** |
| Number of 4 input LUTs: | 1838 out of 50560 **(3%)** |
| Number of FIFO/RAMB16s | 2 out of 232 **(1%)** |
| Number of DSP48s: | 4 out of 128 **(3%)** |

Table 2. Impulse C Resource Report for SVM Application.

Post-processed image visualizations of the Virtex-4FX60 FPGA imlementation of the SVM classification algorithm show excellent agreement with the results from EO-1 (Figures 2 & 3). Having successfully implemented the legacy SVM designed for Hyperion with ample FPGA resources available, extensions to the C-language algorithm were applied to increase overall system performance for on-board classification. Recall that in the Hyperion instrument algorithm currently in use on the EO-1 spacecraft, a linear SVM kernel operates on 11 of 242 available bands of pixel data. Future implementations for SVM classification with FPGAs can consider operating on a larger number of bands for a more complete classification set and/or consider using a nonlinear kernel for improved classification accuracy. Using the C-to-hardware compiler enables quick experimention with alternative implementations such as those shown in Figure 4.

The Hyperion instrument, as part of the NASA New Millennium Program Earth Observing – 1 Mission, was the first imaging spectrometer to routinely acquire science grade data from Earth orbit (Pearlman, 2003). EO-1 was launched on a Delta 7320 rocket from Vandenberg Air Force Base on November 21, 2000. The Hyperion instrument was built by TRW, Inc. (now Northrop Grumman Space Technology) and was designed to provide high quality calibrated data for hyperspectral application evaluations (Pearlman et al., 2001). Hyperion had a single telescope and two spectrometers: one visible/near infrared (VNIR) spectrometer and one short-wave infrared (SWIR) spectrometer producing 242 bands in the range of 400-2500 nm with 10 nm bandwidths. With Hyperion, each pushbroom image frame captured the spectra from an area 30 m along-track by 7.7 km cross-track that was

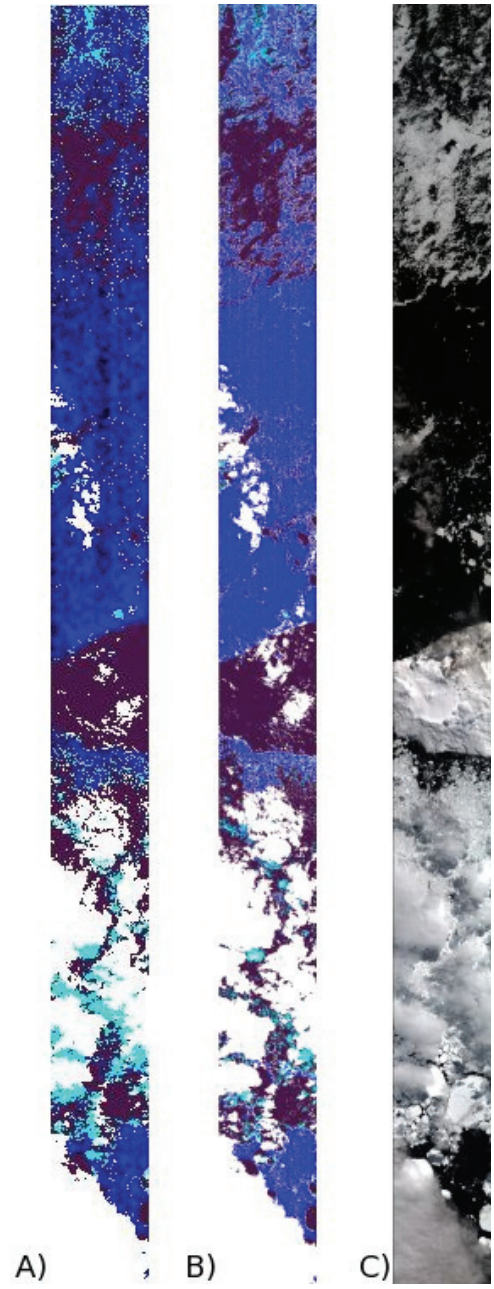



Fig. 2. A comparison of the results from A) the FPGA SVM implementation, B) the legacy C-code SVM, and C) the original hyperspectral image. The color key is blue = water, cyan = ice, dark purple = snow, lavender = unclassified.

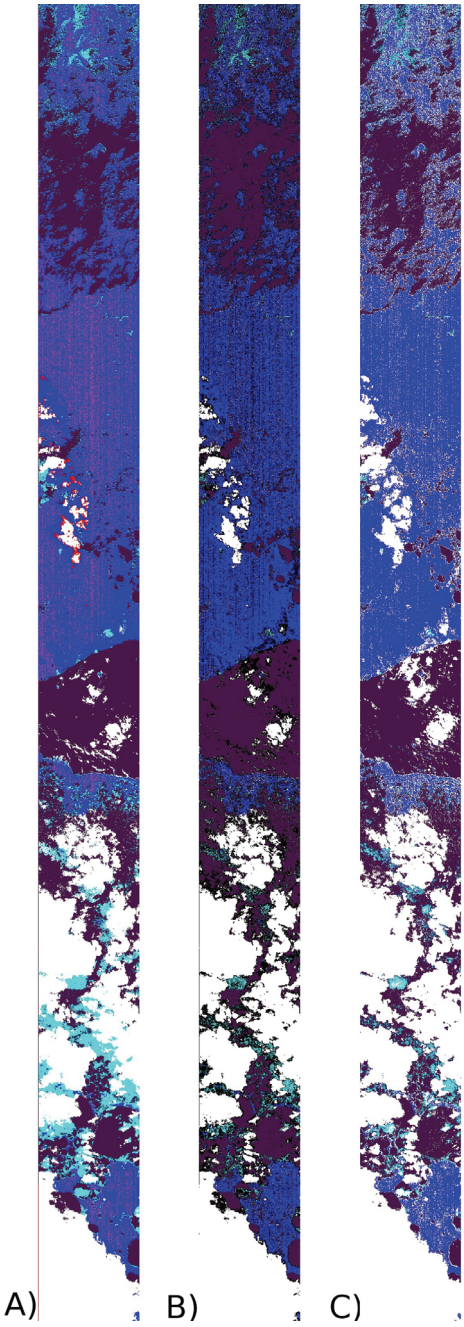



Fig. 3. The black pixels in image (B) indicate the indices where the classifications of pixels (A) and (C) were not identical.

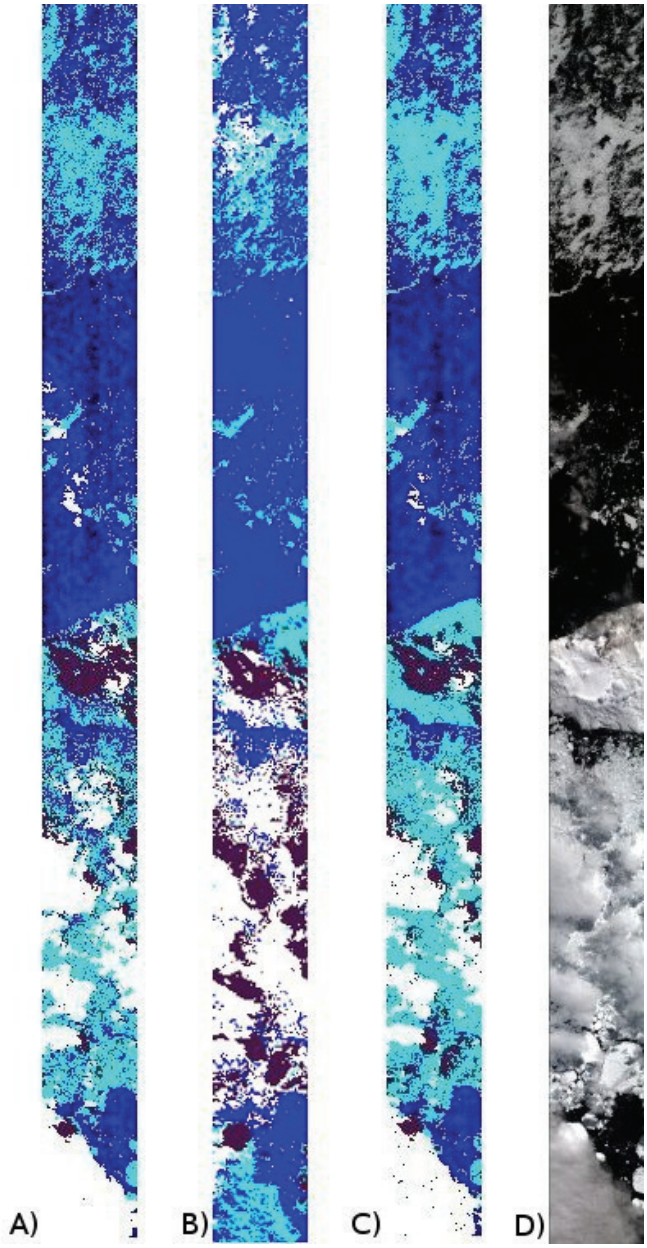



Fig. 4. A comparison of the results from A) the 11-band linear SVM hardware implementation, B) a 30-band linear SVM hardware implementation, C) a non-linear (polynomial) SVM hardware implementation, to D) the original EO-1 hyperspectral image. Color differences between image B) and the other 2 images (A & C) are due to the different bands (qty. 11 vs. 30) selected for each classification.

sufficient to address most land cover issues. The forward motion of the satellite created a sequence of frames that were combined into a two-dimensional spatial image with a third dimension of spectral information (called a "3-d data cube").

In January 2004, as part of the New Millennium Program ST-6 Autonomous Sciencecraft Experiment (ASE), autonomy software has been executing on the EO-1 Mongoose V processor using Hyperion image data for on-board science analysis and autonomous planning and execution of instrument observations (http://ase.jpl.nasa.gov). ASE employs four unique classifiers on Hyperion image pixel data. One is the SWIL classifier described previously in this Section. Following pixel classification, ASE autonomy software triggers new observations based on the analysis of the current observation. This is known as event detection. Specific events of interest on EO-1 include sea ice break-up and formation, lake freeze and thaw, and active volcanism. The limitations of executing the software SWIL classifier on the Mongoose V processor are clear. Of the 242 available bands of Hyperion image data only 11 bands can be processed.

Accommodating reprogrammable FPGAs in the architecture and design of future NASA instruments can provide a highly capable and flexible computing platform for on-board science data processing. The Decadal Survey's Hyperspectral Infrared Imager (HyspIRI) Mission is poised to benefit from reprogrammable FPGAs. The HyspIRI mission has heritage in the EO-1 Hyperion imaging spectrometer as well as in JPL's Moon Mineralogy Mapper ($M^3$) instrument on the Indian Moon-orbiting mission, Chandrayaan-1. Using imaging spectroscopy, the HyspIRI mission would obtain global coverage from LEO with a repeat frequency of 30 days at 45-m spatial resolution. A pointing capability is required for frequent and high-resolution imaging of critical events, such as volcanos, wildfires and droughts. While there is no present plan to include on-board classifiers such as the SWIL classifier on HyspIRI, a cloud detection algorithm implemented in a reconfigurable FPGA is under evaluation for the mission (LeMoigne, 2007).

### 4.2 FTIR Spectrometer

The Panchromatic Fourier Transform Spectrometer (PanFTS) is a NASA Instrument Incubator Program (IIP) funded development to build and demonstrate a single instrument capable of meeting or exceeding the requirements of the GEO-CAPE mission. An FTS is a Michelson interferometer in which the optical path difference of light rays is continuously varied with moving mirrors. Using photovoltaic detectors, this modulated light is converted to an electric signal known as an interferogram. For broadband sources, the interferogram exhibits a large peak at zero path difference where all wavelengths interfere constructively.

The PanFTS development effort has strong heritage in the MATMOS FTIR[2] spectrometer instrument concept designed to measure the Mars atmospheric composition using solar occultation from orbit as proposed (but not selected) for the 2007 Mars Scout mission. The MATMOS instrument would measures the 850-4300 *cm-1* region of the IR spectrum of sunlight as it shines through the Martian atmosphere. MATMOS would record roughly 26 spectra per occultation, with each containing 172,500 spectral elements. There are two occultations per orbit, one for sunrise and one for sunset. The duration of an occultation is

[2] Fourier Transform Infra Red (FTIR)

between 78 and 169 seconds, thus requiring that each spectrum be collected in 3.0 to 6.5 seconds (Pingree et al., 2007). The MATMOS FTS utilizes three separate detectors in the process of collecting occultation spectra. An *HgCdTe* detector is used to collect longer wavelengths (5 μ*m* to 2 μ*m*) and an *InSb* detector collects shorter wavelengths (5 μ*m* to 2 μ*m*). A *Ge* detector is used to collect the reference laser interferogram (used to measure the path difference – internal to the FTS). For each orbit, the three detectors produce 659 Mbytes of raw data that must be processed and compressed prior to downlink.

The data processing consists of five steps: interferogram resampling, phase correction, FFT (Fast Fourier Transform), spectra averaging, and lossless compression. Re-sampling converts the time-domain signal to the path difference domain, removing frequency modulation in the process and reducing the number of points for each solar detector. Phase correction makes the interferogram symmetrical about the zero path difference (ZPD), a point where the two moving mirrors inside the interferometer are at equal distance to the beam splitter. This allows the two symmetrical halves of the interferogram to be averaged together. The spectrum is then computed with an FFT, reducing the dynamic range of the interferogram thus allowing it to be represented with fewer data bits. Averaging scans taken above the atmosphere and then performing lossless compression further reduces the volume of data to be transmitted to Earth.

In a 2007 technology demonstration, the Xilinx V4FX60 FPGA was evaluated for its FTIR spectrometer data processing capability targeting the MATMOS instrument development (Bekker et al., 2008). By optimizing floating-point calculations necessary for processing and compression of MATMOS data prior to downlink, a more than 8x reduction in execution time was achieved on the FPGA however, these results still lagged behind the Rad750's processing capabilities. In 2008, the FTIR spectrometry algorithm was targeted to the most recently available Virtex-5FXT FPGA (Bekker et al., 2009). The V5FXT FPGA contains the more powerful PPC440 processor, more cache, and improved memory interfaces over those of the V4FX, as well as an improved auxiliary processor unit (APU) controller and floating-point unit (FPU). Preliminary results for the MATMOS FTIR on-board processing algorithm on the V5FXT show a nearly 5-x improvement over the V4FX implementation and execution times that now surpass the Rad750. The FTIR V5FXT system is shown in Figure 5.

The PanFTS design for the GeoCAPE mission combines atmospheric measurement capabilities in the IR and UV-VIS[3] with the ability to measure ocean color by using imaging FTS to provide full spatial coverage. For the atmospheric composition, the instrument includes up to four Focal Plane Arrays (FPA) of 128 x 128 pixels that are read at a frame rate of 16 kHz. JPL has developed an interface that records pixel data from commercially available IR FPAs that are capable of the required frame rate at a lower spatial coverage. This interface uses high speed ADCs[4] and the Xilinx Virtex-5FXT FPGA. FTIR on-board processing development efforts on the V5FXT FPGA continue at JPL for the PanFTS instrument.

[3] Ultra Violet Visible (UV-VIS)

[4] Analog to Digital Converter (ADC)

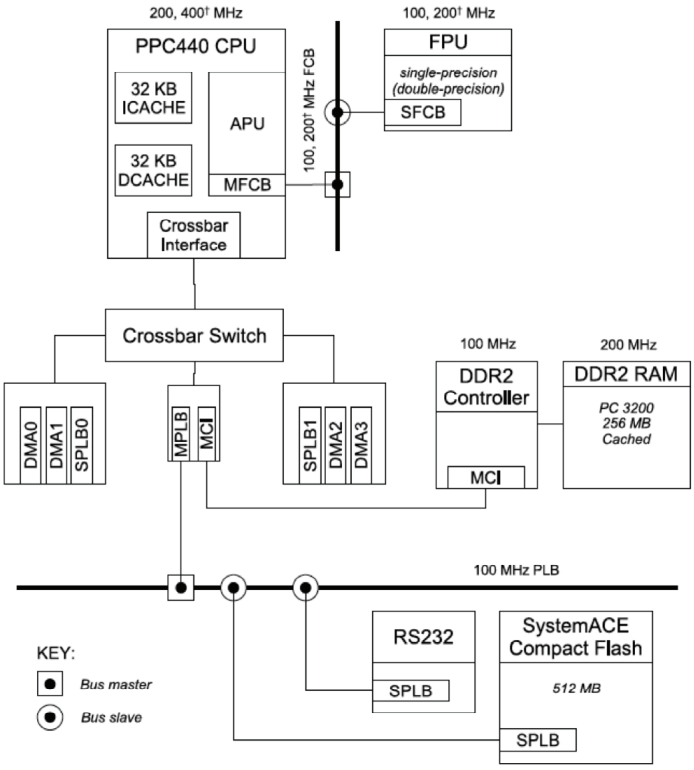



Fig. 5. V5FXT FTIR system with FPU co-processor, shown in multiple configurations (Bekker et al., 2009). † indicates an alternate configuration.

### 4.3 The Multi-angle Spectro-Polarimetric Imager (MSPI)

The Multi-angle Spectro-Polarimetric Imager (MSPI) is an advanced instrument concept in development at JPL to produce a highly accurate multi-angle, multi-wavelength polarimeter to measure cloud and aerosol properties as called for by the Aerosol-Cloud-Ecosystem (ACE) mission concept in the Earth Sciences Decadal Survey. The MSPI instrument will use a set of 9 cameras (8-fixed and 1-gimballed)[5], each associated with a given along-track view angle in the $0^o$-$70^o$ range (see Figure 6). Each camera must eventually process a raw video signal rate around 95 Mbytes/sec over 16 channels.

The greatest challenges of the MSPI instrument are the stringent demand on degree of linear polarization (DOLP) tolerance over a wide swath, and the need to acquire polarimetric and multispectral intensity imaging simultaneously from the UV-SWIR[6]. In an attempt to achieve necessary accuracy of the DOLP of better than 0.5%, the light in the optical system is subjected to a complex modulation designed to make the overall system robust against

[5] Number of cameras is not finalized; may be 7-9.

[6] Ultra-Violet Short Wave Infra Red (UV-SWIR)

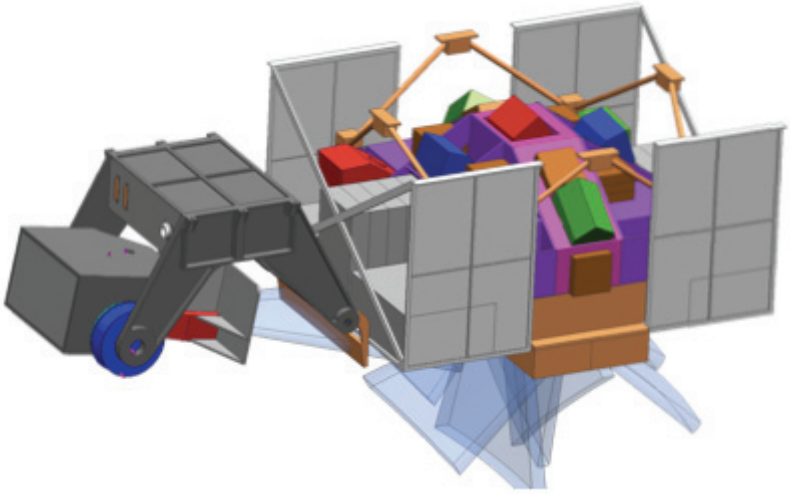


Fig. 6. A conceptual layout of the MSPI instrument. A set of fixed cameras view different angles and a gimbaled camera provides high angular resolution for selected Earth targets and camera-to-camera calibration (Diner et al., 2007).

many instrumental artifacts that have plagued such measurements in the past. This scheme involves two photoelastic modulators that are beating in a carefully selected pattern against each other (Diner et al., 2007). In order to properly sample this modulation pattern, each of the proposed nine cameras in the system needs to read out its imager array about 1000 times per second, resulting in two orders of magnitude more data than can typically be downlinked from the satellite.

A key technology development needed for MSPI is on-board signal processing to calculate polarimetry data as imaged by each of the 9 cameras forming the instrument. With funding from NASA's Advanced Information Systems Technology (AIST) Program, JPL is solving the real-time data processing requirements to demonstrate, for the first time, how signal data at 95 Mbytes/sec over 16-channels for each of the 9 multi-angle cameras in the spaceborne instrument can be reduced on-board to 0.45 Mbytes/sec. This will produce the intensity and polarization data needed to characterize aerosol and cloud microphysical properties. The onboard processing required to compress this data involves least-squares fits of Bessel functions to data from every pixel, effectively in real-time, thus requiring an on-board computing system with advanced data processing capabilities in excess of those commonly available for space flight. A Xilinx Virtex-5 FPGA-based computing platform is currently under development at JPL to meet MSPI's on-board processing (OBP) requirements.

As a brief polarimetry imaging overview, DOLP is calculated by equation (1), where $I$ is the total intensity, and $Q$ and $U$ describe linear polarization.

$$\text{DOLP} = \sqrt{(Q/I)^2 + (U/I)^2} = \sqrt{q^2 + u^2} \quad (1)$$

To achieve the high degree of accuracy in DOLP, two photo-elastic modulators (PEMs) are included in the MSPI optical path to modulate the Q and U polarization components of the Stokes vector. One full cycle of the modulated polarization signal occurs in the time of one 40-msec frame, set by the beat frequency of the two PEMS. Each cycle of the modulation must be "oversampled" to create a hi-fidelity digital representation of the polarization components. The baseline is to sample the modulation 32 times per frame – thereby creating 32 sub-frames per frame. Compared to MISR[7] cameras, each with 4 spectral channels, the raw video data rate that must be handled by MSPI is increased by a factor of 256 (32x due to oversampling; 4x due to expansion of the number of channels, and 2x due to correlated double sampling to suppress read noise in the Si-CMOS readout). A single 16-channel MSPI camera (one of nine) must process 95 Mbytes/sec of raw video data. A computationally intensive linear least-squares algorithm must also be applied to perform data reduction for video processing of the signal output from the photo-detector array. These data reductions can be performed (without sacrificing the information content of the camera product for science) based on how the calculations for digital signal processing are implemented in the reconfigurable FPGA.

The MSPI on-board processor collects data as it streams out from the focal plane of the camera, calculates the basis function values, and computes the least-squares fit of the data using the basis functions. The result of the on-board processing is the reduction of dozens of samples acquired during a 40-msec frame to five parameters. In 2008, the Xilinx Virtex-4FX60 FPGA, including PowerPC405 processors, was used to implement a least-squares fit Bessel function fitting algorithm to generate a pixel data stream (Norton et al., 2009). The algorithm extracts intensity and polarimetric parameters in real-time thereby substantially reducing the image data volume for spacecraft downlink without loss of science information. The accuracy results of the FPGA design indicate that the OBP contribution to the MSPI degree of linear polarization (DOLP) accuracy requirement of less than 0.5% error is on the order of only 0.001%. The Virtex-4FX60 FPGA-based design for MSPI OBP is shown in Figure 7. It provides a successful prototype for the 3-channel ground-based instrument and indicates a path-to-flight for the full 16-channel space-flight instrument proposed for the ACE mission.

Current efforts to advance the MSPI OBP design target the Xilinx Virtex-5 FPGA for its advanced radiation hardness and increased performance capabilities. This 3-year task funded by the NASA ROSES AIST Program will meet the following objectives: (1) complete the design of the 16-channel polarimetric processing algorithm with migration and testing on the Xilinx Virtex-5 FPGA and development board; (2) integrate the on-board processor into the camera brassboard system; (3) perform FPGA design trades to optimize performance and explore how DSP features can be incorporated into the design; and (4) perform laboratory and airborne validation of the OBP system with real-time retrieval of polarimetry data.

[7] Multi-angle Imaging SpectroRadiometer (MISR) is an in-flight instrument on the Terra spacecraft for the Earth Observing System (EOS) Mission, launched in 1999.

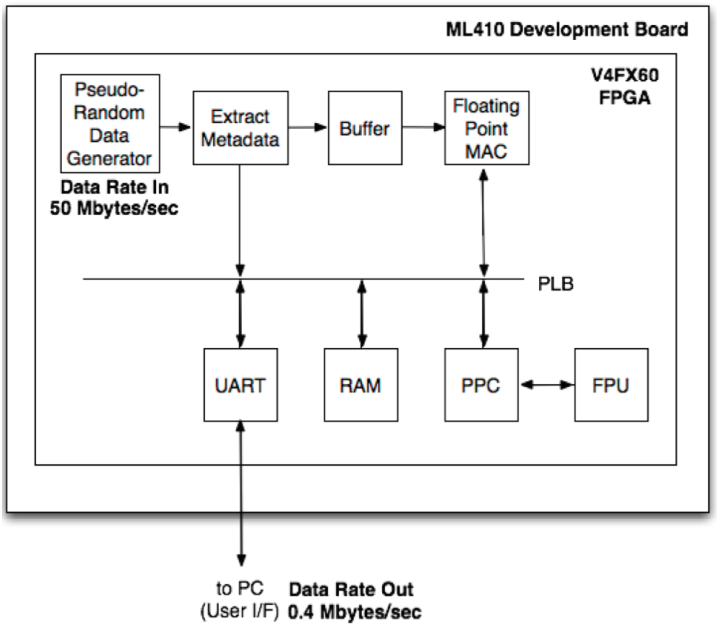



Fig. 7. Top Level Block Diagram of MSPI On-Board Processing Co-Design on the Xilinx Virtex-4 FPGA (Norton et al., 2009).

## 5. Radiation effects and Single Event Upset (SEU) mitigation

From NASA's Preferred Reliability Practice No. PD-ED-1258, Space Radiation Effects on Electronic Components in Low-Earth Orbit, April 1996:

> "Radiation in space is generated by particles emitted from a variety of sources both within and beyond our solar system. Radiation effects from these particles can not only cause degradation, but can also cause failure of the electronic and electrical systems in space vehicles or satellites. Even high altitude commercial airliners flying polar routes have shown documented cases of avionics malfunctions due to radiation events."
>
> "Experience with many spacecraft since Explorer I shows that higher electron concentrations are observed between 45 degrees and 85 degrees latitude in both the northern and southern hemispheres, indicating that the belts descend to a lower altitude in these regions. For low inclination orbits, less than 30 degrees, the electron concentrations are relatively low. Due to the earth's asymmetric magnetic field, a region in the Atlantic near Argentina and Brazil, known as South Atlantic Anomaly (SAA), has relatively high concentrations of electrons. The SAA is known to cause problems such as: single event upsets (SEU)."

SRAM-based reconfigurable FPGA devices are susceptible to SEUs. A necessary feature of any space-flight qualified Xilinx SRAM-based FPGA design, such as those described for

development on the Virtex-4 and Virtex-5 FPGAs, is to mitigate the effects of radiation SEUs. For Virtex-4 designs, SEU mitigation techniques such as Triple Modular Redundancy (TMR) to triplicate logic in the FPGA, presuming there are sufficient remaining resources in the device, as well as running the dual-core processors in lock-step may be employed. The Virtex-5 FPGA is advertised by Xilinx to be Rad-Hard By Design (RBDH), potentially eliminating the need for SEU mitigation techniques to be added into the design.

For future low Earth-orbiting science instruments such as PanFTS (for the GEO-CAPE mission) and MSPI (for the ACE mission), the tolerance to occasional SEUs may be acceptable. The simplest approach for these instruments may be to include only SEU detection in the design and when detection occurs re-load the FPGA configuration file, a key advantage to these reconfigurable computing platforms. This is a viable strategy for non-critical applications that can withstand occasional interruption for re-configuration as may be the case for global mapping science instruments.

## 6. Conclusions

Hybrid or system-on-a-chip (SOC) FPGAs with embedded processors are demonstrating levels of performance and efficiency that were previously impossible using traditional processors for spaceborne computational platforms. Hardware acceleration of science instrument algorithms promises to dramatically improve onboard data processing in future NASA science missions as required by the Decadal Survey. Software-to-hardware autocode design tools can play an important role in the fast prototyping and development of legacy algorithms into hardware accelerated FPGA implementations. The Xilinx FPGA development platforms provided an excellent and cost-effective prototyping environment and a path-to-flight for future instrument on-board processing technology development.

The Xilinx Virtex-4 and Virtex-5 FPGA-based developments and capabilities presented in the design cases of Section 4 respond directly to the future needs of the Decadal Survey missions for instrument science data on-board processing. The results to date demonstrate the benefits of FPGA-based processing for spectrocopy and image processing. The new RHBD architecture of the Virtex-5 FPGA promises to resolve the SRAM-based FPGA limitation of SEU susceptibility.

(Part of) This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

## 7. References

Bekker, D.; M. Lukowiak, M. Shaaban, J-F. Blavier, & P. Pingree (2008). A Hybrid-FPGA System for On-Board Data Processing Targeting the MATMOS FTIR Instrument, *Proceedings of IEEE Aerospace Conference,* Big Sky, MT, March 2008.

Bekker, D.; J-F. Blavier, G. Toon, & C. Servais (2009). An FPGA-Based Data Acquisition and Processing System for the MATMOS FTIR Instrument, *Proceedings of IEEE Aerospace Conference*, Big Sky, MT, March 2009.

Castano, R.; N. Tang, T. Dogget, S. Chien, D. Mazzoni, R. Greely, B. Cichy, & A. Davis (2006). Onboard classifiers for science event detection on a remote sensing spacecraft, *Proceedings of the 12th ACM SIGKDD International conference of Knowledge Discovery and Data Mining*, ACM Press, (2006), 845 – 851.

Diner, D.; A. Davis, B. Hancock, G. Gutt, R. Chipman, & B. Cairns (2007). Dual-photoelastic-modulator-based polarimetric imaging concept for aerosol remote sensing, *Applied Optics, Vol. 46 Issue 35*, pp.8428-8445 (2007).

Kahn, R.A., J.A. Ogren, T.P. Ackerman, J. Bösenberg, R.J. Charlson, D.J. Diner, B.N. Holben, R.T. Menzies, M.A. Miller, & J.H. Seinfeld (2004). Aerosol data sources and their roles within PARAGON. *Bull. Amer. Meteorol. Soc.* 85, 1511-1522.

LeMoigne, Jacqueline (2008). "A Reconfigurable Computing Environment for On-Board Data Reduction and Cloud Detection", *2.7 Buckner Hyspiri Techonology Investments Presentation_Final.pdf.*

National Research Council (NRC), Committee on Earth Science and Applications from Space (2007). *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond.* The National Academies Press, Washington, DC, 437 pp.

Norton, C.; T. Werne, P. Pingree, & S. Geier (2009). An Evaluation of the Xilinx Virtex-4 FPGA for On-Board Processing in and Advanced Imaging System, *Proceedings of IEEE Aerospace Conference,* Big Sky, MT, March 2009.

Pearlman, J.; S. Carman, C. Segal, P. Jarecke, P. Barry, & W. Browne (2001). Overview of the Hyperion imaging spectrometer for the NASA EO-1 mission, *IEEE International Geoscience and Remote Sensing Symposium*, 6, pp 3504-3506, 2001.

Pingree, P.; J-F. Blavier, G. Toon, & D. Bekker (2007). An FPGA/SoC Approach to On-Board Data Processing Enabling New Mars Science with Smart Payloads, *Proceedings of IEEE Aerospace Conference,* Big Sky, MT, March 2007.

Pingree, P.; M. Janssen, J. Oswald, S. Brown, J. Chen, K. Hurst, A. Kitiyakara, F. Maiwald, & S. Smith (2008-a). Microwave Radiometers from 0.6 to 22 GHz for Juno, A Polar Orbiter around Jupiter, *Proceedings of IEEE Aerospace Conference*, Big Sky, MT, March 2008.

Pingree, P.; L. Scharenbroich, T. Werne, & C. Hartzell (2008-b). Implementing Legacy-C Algorithms in FPGA Co-Processors for Performance Accelerated Smart Payloads, *Proceedings of IEEE Aerospace Conference,* Big Sky, MT, March 2008.

Platt, J. (1998). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines, *Microsoft Research Technical Report MSR-TR-98-14*, (1998).

Sander, S. P.; R. Beer, J-F. Blavier, K. Bowman, A. Eldering, D. Rider, G. Toon, W. Traub, & J. Worden (2008). Panchromatic Fourier Transform Spectrometer (PanFTS) for the Geostationary Coastal and Air Pollution Events (GEO-CAPE) Mission, American Geophysical Union, Fall Meeting, San Fransciso, CA, December 2008.

Yu, H., Y.J. Kaufman, M. Chin, G. Feingold, L.A. Remer, T.L. Anderson, Y. Balkanski, N. Bellouin, O. Boucher, S. Christopher, P. DeCola, R. Kahn, D. Koch, N. Loeb, M. S. Reddy, M. Schulz, T. Takemura, & M. Zhou (2006). A review of measurement-based assessments of the aerosol direct radiative effect and forcing. *Atmos. Chem. Phys.* 6, 613-666.

# 6

# New Reprogrammable and Non-Volatile Radiation-Tolerant FPGA: RT ProASIC®3

Sana Rezgui
*Actel Corporation*
*USA*

## 1. Introduction

Non-Volatile and Reconfigurable Field Programmable Gate Arrays (FPGAs) present an attractive solution for high-level system integration in various aerospace and military applications. Commercially available Low-Power Flash-based FPGAs, 0.13-µm ProASIC3/L (A3PL) and its extended family product (A3PEL) are non-volatile while providing remote in-system reprogramming to support future design iterations and field upgrades. Flash-based technology provides them the advantage of being a secure, low-power, single-chip solution [Morris, 2006]. Unlike SRAM based-FPGAs, the configuration memories are not volatile and hence don't require additional non-volatile memory to reload the device configuration data at system-power-up or due to radiation effects [Swift et al., 2004] in addition to Triple Module Redundancy (TMR) of its entire set of configuration bits [Carmichael, 2001]. This reduces cost, power, and initialization time and improves system reliability. However, despite the SEE immunity of their configuration memory, their Floating Gate (FG) switches and CMOS logic gates are susceptible to both effects of the Total Ioninzing Dose (TID) and the Single Event Effects (SEE).
For TID effects, the primary issue is the radiation-induced charge loss in the floating gate [Snyder et al., 1989, Cellere et al., 2004, Wang et al., 2004, Guertin et al., 2006], resulting in the change of the FPGA electrical performances (maximum speed, current, etc.). While for SEE, the primary concern resides in the upset of its registers (state of the flip-flop) due to a particle hit, resulting in the disruption of the normal operation of the FPGA-design [Rezgui et al., 2007a & 2007b]. The new Radiation-Tolerant ProASIC3 (RT ProASIC3 or RT3P), sharing the same silicon of the Low-Power A3PL FPGAs is hardened for TID and SEE by software means in a transparent manner to the user [Rezgui et al., 2008a]. The Single Event Transients (SET) tolerance is hardened by single or duplication filtering [Shuler et al., 2005 & 2006, Balasubramanian et al., 2005, Baze et al., 2006, Mavis & Eaton, 2007, Rezgui et al., 2007a] and Single Event Upsets (SEU) are hardened by TMR or Error Detection and Correction (EDAC) to soft error rates less than $10^{-10}$ upsets/bit-day and $LET_{th}$ larger than 40 MeV•cm$^2$/mg for clock frequency up to 100 MHz. While their TID limit is improved by simple reprogrammimg of the FPGA resulting in the restoration of the charge loss from their configuration FG swicthes.
This chapter describes the employed mitigation techniques for the A3P product family, to attain the radiation levels of the RT-product and presents the results issued from the TID and the SEE characterization of both of the A3P and the A3PL (the Low-Power version of

ProASIC3). The SET characterization or mitigation will not be addressed in this chapter, but detailed analyses and measurements of SET cross-sections are provided in [Rezgui et al., 2007a, 2008b & 2009]. This chapter includes a brief description of the RT ProASIC3 FPGA from architectural and device perspectives as well as detailed analyses of the radiation test results issued from 1) the TID characterization, 2) the SEE characterization and 3) the TID Effects on the SEE Sensitivities.

## 2. New radiation-tolerant 0.13-µm flash-FPGAs

Based on its low-power capabilities and its increased IO features in the Extended (E) family product, the 0.13-µm ProASIC3EL (A3PEL) part is selected as the silicon foundation of the new Radiation-Tolerant Flash-based FPGA (RT ProASIC3). Additionally, RT ProASIC3 FPGAs are assembled in hermetically-sealed ceramic packages, which are available as either Column Grid Array (CG, with Six Sigma solder columns attached) or Land Grid Array (LG, no solder columns attached). Qualification, inspection, assembly, and testing are performed in accordance with MIL-STD-883 Class B [MIL-STD-883G]. In the following, a brief description of these products at the architectural and the device levels as well as of the differences between the A3P and A3PL product families are given.

### 2.1 The ProASIC3 internal architecture

The A3PEL product family has up to 3 million system gates, 504 kbits of true dual-port SRAM, 620 single-ended I/Os, and 300 differential I/O pairs. They also include 1 kbits of on-chip, programmable, non-volatile Flash-ROM (FROM) memory storage as well as up to 6 integrated phase locked loops (PLL). The FPGA core consists of logic tiles, called "VersaTiles", and routing structures. Each logic tile is a combination of CMOS logic and flash switches and can be configured as a three-input logic function or as a D-flip-flop with an optional enable, or as a latch by programming the appropriate flash switch interconnections. The logic tiles are connected with each other through routing structures and FG switches. These flash switches are distributed throughout the device to provide reconfigurable programming to connect signal lines to the appropriate logic-tile inputs and outputs [ProAISC3 Handbook], as shown in Fig. 1. The Flash-FPGAs are reprogrammable through the JTAG port and contain programming control circuits composed of charge pumps, sense amplifiers, Digital to Analog Converters (DAC), CMOS logic, High-Voltage (HV) NMOS transistors and FG cells to store the factory parameters.

### 2.2 Floating gate device

As shown in Fig. 1 and detailed in [Wang et al., 2004a, 2006a & 2006b], the FPGA switch circuit is a set of two NMOS transistors: 1) Sense Transistor to program the floating gate and sense the current during the threshold voltage measurement and 2) Switch Transistor to turn ON or OFF a data-path in the FPGA. The two transistors share the same control gate and floating gate. The threshold voltage is determined by the stored charge in the FG. Fowler-Nordheim tunneling through the thin gate oxide (100 Å) is the mechanism that modulates the stored charge during program and erase of the FG. The FG switch is programmed to a low threshold voltage state to turn the switch ON and erased to a high threshold voltage state to turn it OFF. Fig. 2 shows the structure of the FG transistor: an NMOS transistor with a stacked gate. Between the silicon substrate and the floating gate is

the tunnel oxide and between the FG and the control gate the inter-poly oxide-nitride-oxide (ONO) composite dielectric.

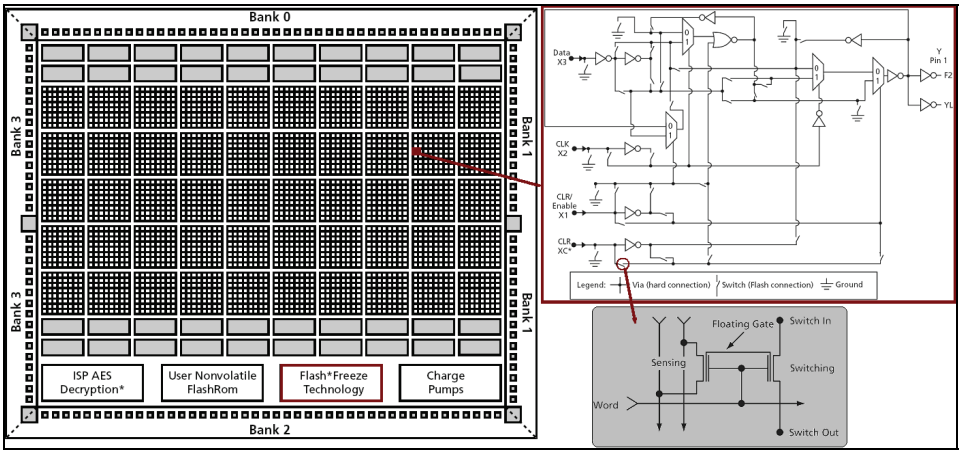



Fig. 1. ProASIC3 FPGA Core, VersaTile (Logic Tile) and Flash-Based Switch. Each logic tile is a combination of CMOS logic and flash switches.

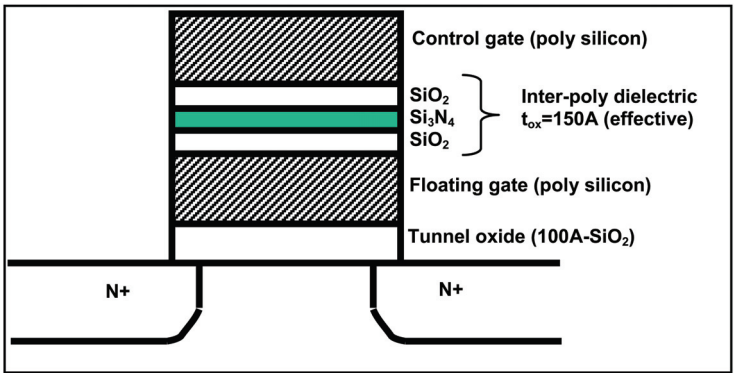



Fig. 2. Floating Gate Transistor in the Flash-Based FPGA is a set of two NMOS transistors: 1) Sense Transistor to program the floating gate and sense the current during the threshold voltage measurement and 2) Switch Transistor to turn ON or OFF a data-path in the FPGA.

### 2.3 ProASIC3E and low-power ProASIC3EL

The Low-Power A3PEL parts although different at the process level from the A3PE are identical at the design and architectural levels and are pin to pin compatible with the A3PE, except for a new added feature called "flash-freeze". This feature provides a low-power static mode that retains all SRAM and register information with rapid recovery to "active" (operating) mode, by simply asserting a single input. The device then enters a low-power mode in 1µs, in which case, clocks are frozen, I/Os are tri-stated, and core registers and memories maintain state. In this mode, external signals driving the FPGA I/Os and clocks can still be toggled without impact on the device power consumption. For instance, in the

flash-freeze mode, the power consumption of the low-power FPGAs ranges in the tens of microwatts [ProASIC3 Handbook].
Furthermore and because of their basic process differences, resulting mainly in the increase of their threshold voltages, the A3PEL products have much lower power consumption than the A3PE part. For instance, the A3PEL operates at 40 percent lower dynamic power and 90 percent lower static power than the ProASIC3E FPGAs, and orders of magnitude lower power than the SRAM-based FPGAs, with up to 350 MHz operation. These process differences between the two product families (A3PE and A3PLE) are only induced in the CMOS transistors used to build the FPGA logic blocks but not in the FG transistors. Since the TID effects are much lower on the CMOS transistors than on the FG transistors, the same TID performance should be expected for both of the A3PE and A3PEL parts when both are operated at a 1.5 V core voltage. In addition and at the opposite of the A3PE FPGA which could operate only at 1.5V, the A3PEL can operate at all core voltages between 1.2 and 1.5V which allows more reduction in their power consumption when operated at 1.2V.
In the following sections, the test results issued from TID and SEE test experiments of the ProASIC3EL are reported and discussed along with additional suggestions on mitigation methodologies suitable for the target device. For these experiments, a few devices from the ProASIC3EL product family were selected for the TID characterization in x-rays and gamma-rays (the A3P250 and the A3PL600) and the SEE characterization in heavy-ions (HI) and protons beams (the A3P250 and the A3P1000). Since the A3P/E and the A3PL/E share the same FPGA core, the radiation test results are expected to be very similar.

## 3. TID characterization

This section covers the TID performance at the product level of the A3P and the Low-Power RT PRoASIC3 (A3PL) product families. Radiation tests for the selected products were performed in x-rays at ARACOR facility, in Sunnyvale, CA and in gamma-rays at the Defense MicroElectronics Activity (DMEA), in Sacramento, CA. The x-rays irradiation tests are performed by an ARACOR 4100 x-rays Irradiator. The TID test results are reported and discussed, along with additional suggestions on ways to extend the TID lifetime of the Flash-FPGAs.
The purpose of this characterization is to study the TID effects on 1) the FPGA core (CMOS logic and FG devices) and 2) the programming control circuit (FG devices, charge pumps, analog circuits and HV NMOS devices). TID irradiation tests for the selected features were performed in x-rays and in gamma-rays. Most of the results presented in this chapter are obtained in x-rays beams whose effects are estimated to be approximately 2.9 times less effective than those measured in Gamma rays [Wang et al., 2004]. This calibration factor between the x-rays and the Gamma-ray data was calculated experimentally using the same methodology previously applied in [Palkuti & LePage, 1982]. Additionally, all the x-rays irradiation tests were performed on the A3P parts (A3P250-PQ208) while Gamma test experiments at DMEA, were performed on the A3PL part (A3P600L-FG484), both when operated at 1.5V core voltage. During all x-rays and Gamma dose irradiations, except for the power pins, all the Device Under-Test (DUT) inputs are grounded; the ambient is at room temperature.

### 3.1 TID effects on floating gate transistors

Three radiation-induced mechanisms detailed in [Wang et al., 2004, Brown & Brewer, 2002] can affect the threshold voltage of the FG devices: 1) holes injected into the FG, 2) holes

trapped into the oxides and 3) electrons emitted over the polysilicon/oxide barriers. Electron-hole pairs initiated from radiation test results in the injection of holes in the FG and the trapping of holes in the oxides. Hole injection and trapping have a similar effect since they both reduce the threshold voltage in the FG device. The third radiation phenomenon: electron-emission occurs mainly when radiation-induced photons possess an energy exceeding the potential barrier. The emitted electrons are then swept to the substrate or control gate by the electric field, which reduces the FG threshold voltage. Fig. 3 shows an example of threshold voltage ($V_t$) shift in both the program and erase distributions of the FG devices, when irradiated in x-rays.

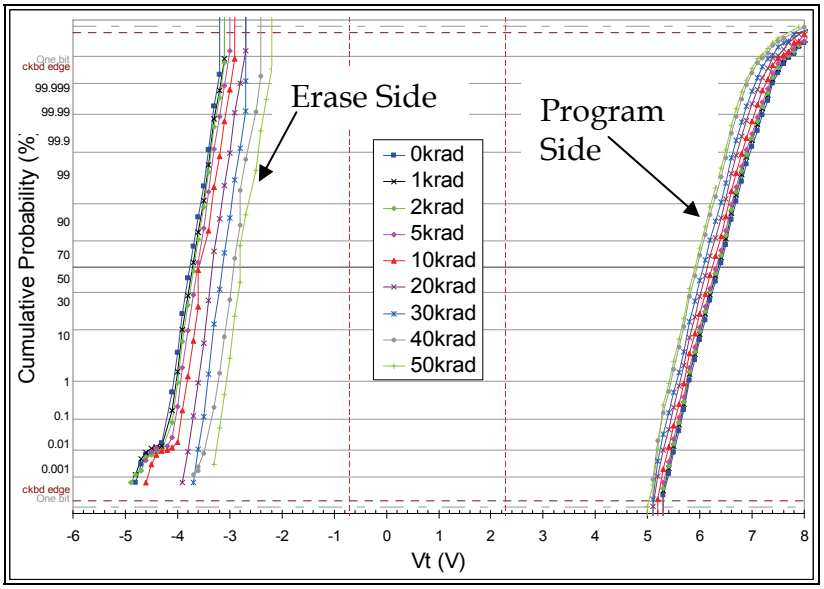



Fig. 3. Threshold Voltage Degradation vs. TID of x-rays Irradiation for an A3P250-PQ208 DUT. The charge loss effects on the FG were investigated by x-rays irradiation in [4-7].

The fundamental, consistent physical process of charge generation, separation and trapping in the dielectrics surrounding the floating gate will modulate the threshold voltage ($V_t$) of the floating gate (FG) device and subsequently the function of the FPGA. The major key TID-indicating electrical parameters on a given FPGA design are 1) the propagation delay, which is best measured on an inverter-string design and 2) the maximum allowed frequency of the circuit registers. In the following, TID-induced effects on a given design will be discussed for both of the DUTs mentioned above (A3P and A3PL).

### 3.2 TID performance of the FPGA core

#### 3.2.1 Test design and test procedure

To measure the TID effects on the FPGA core, three A3P FPGAs were configured with three sub-designs: D1) an inverter-string with 1000 stages, D2) a shift register with 1000 D-flip-flops (DFF) running at 350 MHz and D3) a shift register with 310 DFFs combined with combinational logic (12 inverters) between each consecutive flip-flops running at 135 MHz.

Before x-rays irradiation, at 0 Krad, both of the rising and falling edges of the D1 output signal are measured and on average are approximately 530 ns. The maximum attained frequency of the D2 design was 350 MHz, as stated in the [ProASIC3 Handbook], while the maximum frequency for D3 is about 135 MHz. For both of the D2 and the D3 test designs, the input-data are toggling at half of the clock frequency and at the positive edge of the clock-input, while their output-data are switching at the clock negative edge. After exposure to a certain dose, the rising and falling edges for the D1 output signal, and the maximum attained frequency for the D2 and the D3 sub-designs, were measured. A block diagram of the DUT design is given in Fig. 4.

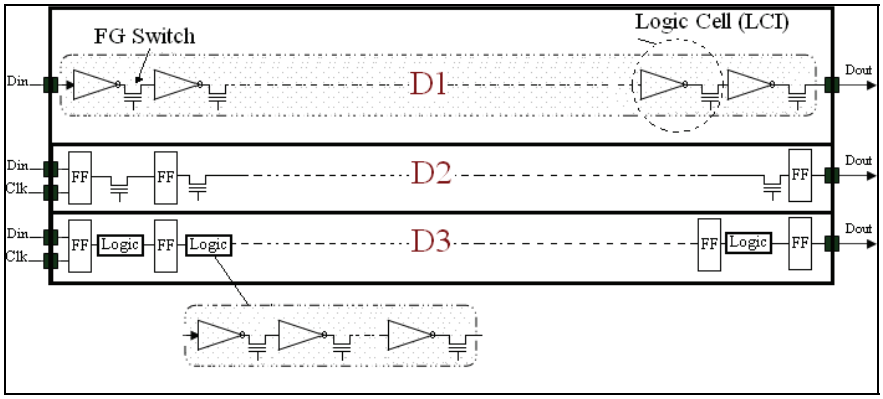



Fig. 4. Block Diagram of the DUT Design. This design is shared in three sub-designs: D1) an inverter-string, D2) a shift register and D3) a shift register combined with combinational logic between each consecutive flip-flops.

The input signals for each sub-design are supplied from an off-chip pulse generator while the electrical parameters of the three output signals were observed and recorded on the scope off-beam after two minutes from each DUT irradiation. The same tests applied to the A3P part, combining combinational and sequential logic, have been repeated in gamma-rays for the A3PL600-FG484 FPGA at DMEA and the issued results are reported. The dose rate during these tests was varied between 4 and 25 Krad/min (67 and 461 rad/s), which is higher than the dose rate required by the TM1019.7 (50 rad/s) [MIL-STD-883G].

### 3.2.2 X-Rays test results

The test circuits were exposed continuously to TID until one of the three sub-design's output state became unstable off beam and required annealing to recover normal operation. This instability in the output signals was always accompanied with an increase of the current in the FPGA core (from 1 to 33 mA in the worst observed case) and was mainly observed starting from an x-rays total dose of 175 Krad (60 Krad in Gamma Rays). The obtained results for the A3P FPGA, displayed in Fig. 5, show that for the A3P parts (DUT 3), the 10% degradation in the propagation delay was obtained at 66 Krad (22 Krad in gamma-rays).

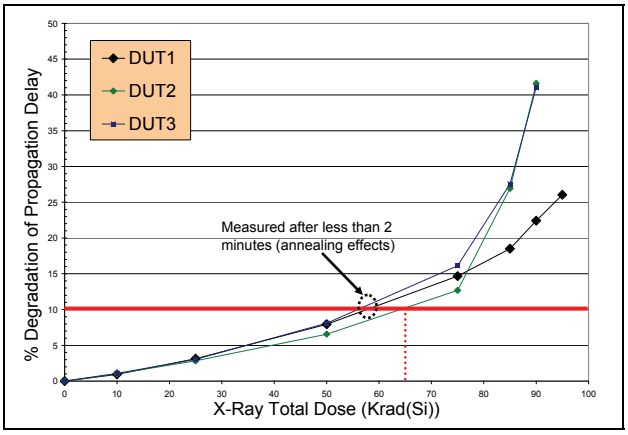



Fig. 5. % Propagation-Delay Degradation vs. TID of x-rays Irradiation for three A3P250-PQ208 DUTs. The 10% degradation in the propagation delay was obtained at 66 Krad.

Furthermore, as shown in Fig. 6, until a TID of 78 Krad, no differences in the maximum allowed frequency for the D2 was noticed, which means that all the DFFs can still operate at 350 MHz. This means that all the timing requirements (setup time, etc.) needed for the DFF were still valid. However, when combining both of the combinational and sequential logic in one single design (D3), the TID limit to observe a variation in the maximum frequency was reduced to 70 Krad as shown in Fig. 7. Indeed, the true maximum frequency of a DFF is about 2 GHz but because of the IOs and the internal FPGA's routing, the maximum frequency is reduced to 350 MHz. Therefore, although an actual reduction in the maximum speed of a DFF has occurred during TID irradiation, it does not show much until a high TID, which means a high reduction in the maximum frequency.

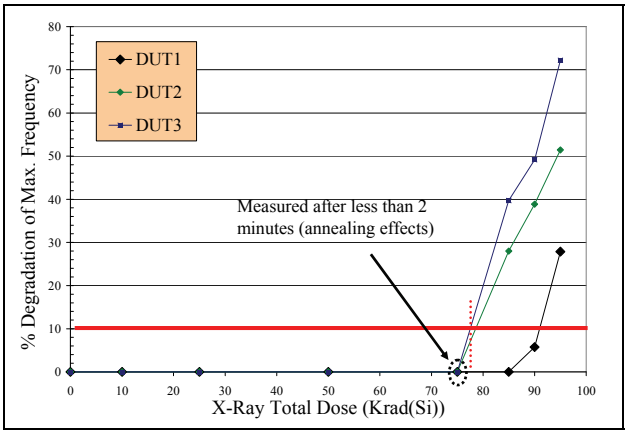



Fig. 6. % D2-Frequency Degradation vs. TID of x-rays Irradiations for three A3P250-PQ208 DUTs. Degradation in the D2 maximum frequency was observed only at 75 Krad.

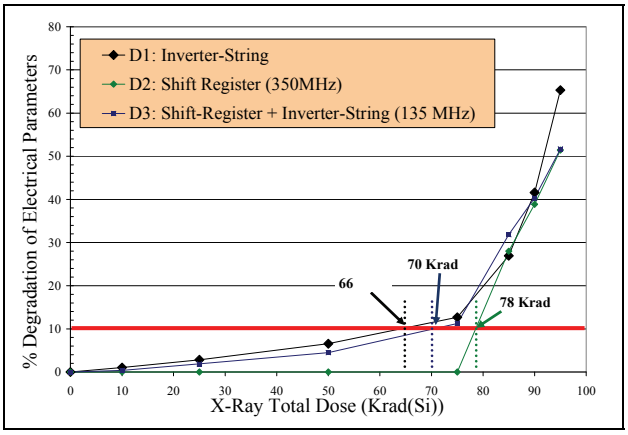



Fig. 7. % D1, D2 and D3 Electrical Parameters Degradation vs. TID of x-rays Irradiations for three A3P250-PQ208 DUTs. 10% Degradation for the D1 was observed at 66 Krad, D2 (78 Krad) and D3 (70 Krad).

In the following, the same test data will be compared to gamma-rays to calculate the circuits' TID performances as well as to verify the 2.9 factor between the x-rays and the gamma radiations.

### 3.2.3 Gamma-rays test results

The A3PL600-FG484 was exposed to Gamma irradiation at the core voltage of 1.5V. Fig. 8 shows the measured degradation in the propagation delay in the inverter-string along with the extrapolated data from the previously obtained data in x-rays irradiation (Fig. 7) based on the 2.9 factor. The obtained results show that the measured and predicted TID limit correlate quite well confirming the correctness of the 2.9 factor between the Gamma and x-rays dose irradiations. However, as the only purpose of this comparison is the correlation

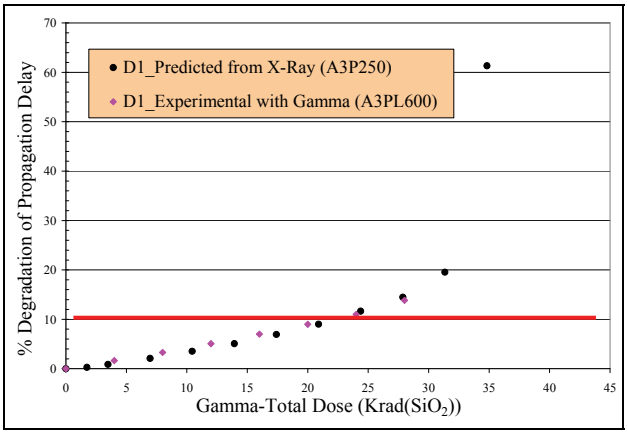



Fig. 8. % Propagation-Delay Degradation vs. TID of Gamma-ray Irradiation for A3P600-FG484 DUT with the correlation factor (2.9).

between both of the x-rays and the Gamma-ray radiation data, it is certainly not the objective to show which one has the higher TID effects. On the other hand, this data confirm that the TID limit of the A3PL part is around the 22 Krad relative to Gamma Rays. Additionally, and as shown in Fig. 9, the obtained data for the D2 show no degradation in the flip-flops maximum frequency till a TID of 28 Krad (the last tested value). This confirms the same x-rays test results, proving again that a degradation in the speed performances of a logic tile configured as a Flip-Flop is less observable than on a logic tile configured as an inverter.

However, as for the x-rays TID testing, the TID performance of Design 3, although slightly better (28 Krad), follows the trend of the TID performance of the Design 1 (the inverter-string). This is expected since Design 3 combines both sequential and combinational logic.

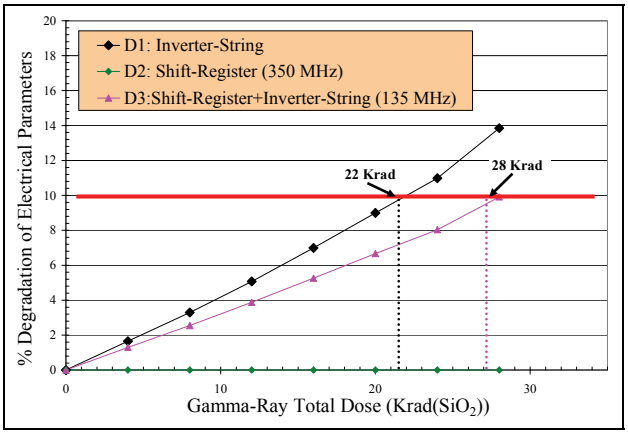



Fig. 9. % D1, D2 and D3 Electrical Parameters Degradation vs. TID of x-rays Irradiation for three A3P250-PQ208 DUTs.

### 3.3 TID performance of the programming control circuit

The main function of this circuit is to erase, program and measure the threshold voltages ($V_t$) of each sense FG device. As a consequence, the test flow consists of reprogramming the part, which invokes erasing, reprogramming and verifying the correctness of the configured design by measuring the $V_t$ of all the sense devices. For clarity purposes, the entire procedure will be called reprogramming or refreshing of the part. The test flow, applied on the A3P parts, consisted of reprogramming the part off-beam after its irradiation to a certain dose (10 Krad in x-rays in this case) until failure to reprogram was observed.

The test results showed that the maximum TID at which the programming procedure passed was 40 Krad, since it failed at 50 Krad, which suggests that the TID limit of this sub-circuit is between 40 and 50 Krad in x-rays. Note that all the three tested parts that were exposed to 50 Krad recovered the reprogramming capability at room temperature after few days. This means that this part is subject to annealing effects. The following section will show some of these effects. Also and as mentioned above, the TID limit in x-rays irradiation for the FPGA core was about 66 Krad, while for the programming control circuit, it is about 40 Krad. This difference in the TID limits could be due to the FG devices located in the programming control circuit, the thick-oxide HV devices, possibly the analog circuits or the

charge pumps. Since the TID tests were done at the product level, it is not possible to conclude on the first failing part to TID in the programming control circuits.

### 3.4 FG refreshing & annealing effects on the product's TID limit

#### 3.4.1 Test procedure

As explained in [Wang et al., 2004 & 2006], the percentage of the degradation in the propagation delay is mainly due to the charge loss in the FG devices (whether in the erase or the program state). Therefore, a first TID mitigation solution would be to attempt to restore that charge to these FG cells. This refresh could simply be done by erasing and reprogramming the Flash-FPGA. However, since the previous results showed that the programming circuit is limited to 40 Krad in x-rays irradiation unless annealing effects are taken in account, the test flow consisted in reprogramming the part off-beam after having been irradiated to 10, 20, 30 and 40 Krad (x-rays). On the other hand, when starting from a much higher TID (85 Krad in x-rays), the measurements of the electrical parameters of the D1, D2 and D3 became variable with time, requiring longer time than 2 minutes to get a stable value of the output states. These electrical parameters improved with annealing time and were then recorded after 2, 15 and 30 minutes, starting from a TID of 85 Krad. Indeed, as shown in Fig. 10, three data points are displayed at 85, 95 and 105 Krad. An improvement of 10% was observed between each measurement taken at 2, 15 and 30 minutes at these three TID values, clearly showing the annealing impact on the FG devices.

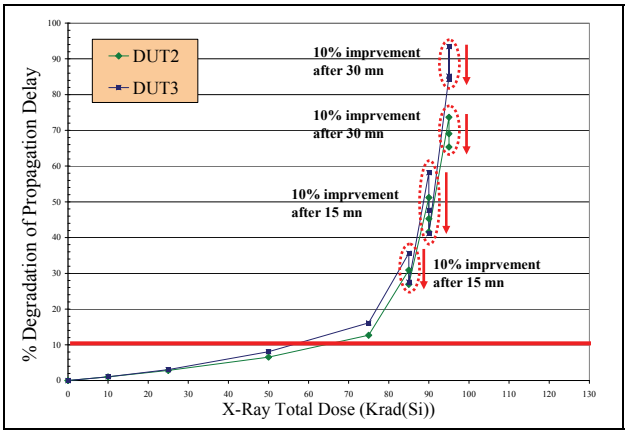



Fig. 10. Annealing Effects on the A3P250 DUTs. These effects are clearly observed for TID higher than 85 Krad.

#### 3.4.2 Test results of the refreshing effects

The obtained results, shown in Fig. 11, demonstrate clearly the efficacy of the employed refresh technique in restoring the lost charge from the FG devices. They also show that at each refresh, the three sub-designs restore completely the original operational parameters (rising and falling times as well the maximum frequencies). Indeed, the maximum TID limit (based on 10% degradation in the propagation delay) was increased by 18 Krad, improving it from 22 to 40 Krad. This suggests that if the programming circuitry was more robust to TID effects, the overall TID lifetime of the FPGA core could be extended to higher than 40

Krad. Note, that the predicted data shown in Fig. 11 was extracted from the TID measurements during the DUT exposition to x-rays. Both of the x-rays and Gamma induced-radiation correlate again quite well and confirm the 2.9 factor. Furthermore, after each refresh cycle (10 Krad irradiation in x-rays), the threshold voltages were measured. The obtained $V_t$ distributions, similarly to what has been shown in Fig. 3, prove that all the FG devices have regained their lost charge because of TID and shifted back to their original $V_t$ whether on the program or the erase side.

Note that when employing the refresh techniques and except for the device de-rating aspects of it, the three sub-designs remained functional proving that no switching of the FG transistors from ON to OFF and vice versa has occurred, until a TID of 275 Krad in x-rays which should be equivalent to 95 Krad when exposed to gamma-rays. Furthermore, since the three sub-designs use 99% of the FPGA logic tiles, and remained fully-functional although with much lower timing performance, it is then clear that there are no stuck bits because of x-rays or gamma irradiations.

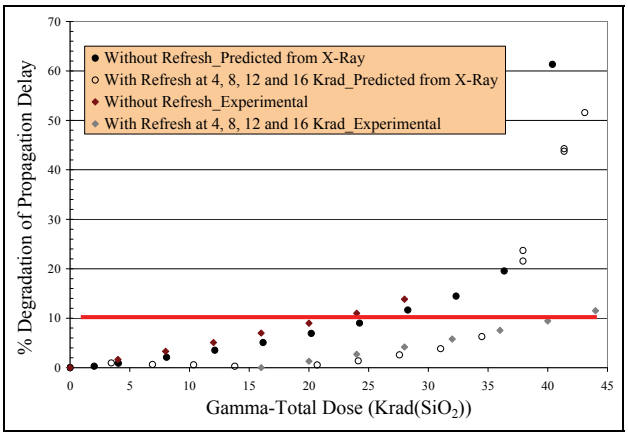



Fig. 11. Refresh Effects on the A3P250 DUTs. The reprogramming of the A3P part in Gamma and x-rays restore the lost charge from the FG devices and increase the product's TID limit.

In summary, the obtained results showed TID sensitivity in the FPGA core and the programming control circuit of the FPGA. A degradation of 10% in the propagation delays was attained at 22 Krad and the part could not be reprogrammed after 16 Krad when exposed to gamma-rays. However, two phenomena to mitigate the TID effects on the FG devices have been observed: 1) the considerable annealing effects and 2) the impact of the FPGA refreshing to restore the FG-lost charge. Indeed, after each refresh of the FPGA core, the latter recovers the original electrical parameters, as if it has not been irradiated. Nevertheless and because of the low TID performance of the programming control circuit, the TID limit of the FPGA core could not be improved to higher than 40 Krad in gamma-rays. In the next section, the SEE characterization and mitigation of the 0.13-µm ProASIC3 FPGAs will be heavily addressed [Rezgui et al., 2007a, 2008b & 2009].

## 4. SEE characterization

The SEE characterization of the ProASIC3 FPGA was performed in HI and proton beam experiments. HI beam experiments were performed at the facility of Texas A&M University

(TAMU) and at the Lawrence Berkeley National Laboratories (LBNL) while proton radiation experiments were conducted at the Crocker Nuclear Laboratory of California in Davis (CNL). HI beam experiments were performed with a wide ion-cocktail (Neon, Argon, Copper, Krypton and Xenon) at normal incidences and two additional tilt angles (30° and 45°). No testing with rolling angles was performed nor is differentiation in the data between the data collected at normal incidence or tilt angles is provided in this chapter.
Radiation tests targeted primarily the five programmable architectures in the ProASIC3: 1) FPGA Core, 2) Clock Network and PLL, 3) Flash-ROM (non-volatile memory) and 4) SRAM. The schemes of the DUT designs for the testing of these programmable blocks as well as the derived beam test results showing some SEE sensitivity in most of the programmable architectural features of the FPGA except in the FROM, are described and discussed in the following.

### 4.1 Devices under-test & experimental test setup

For the beam test experiments, two devices from the ProASIC3 product family were selected: the A3P250 and the A3P1000. Each selected part is mounted in a PQ208 package. Table 1 shows the features of the two selected parts. The test primarily targets the circuitry used for the DUT erase and programming depicted in the bottom of Fig. 1 as the block for "Charge Pumps" as well as the 5 configurable architectures in the A3P FPGA, as shown also in Fig. 1: 1) the FPGA Core, 2) the Clock Network and the PLL, 3) the FROM and 4) the SRAM.

| Part | A3P250 | A3P1000 |
|---|---|---|
| System Gates | 250K | 1M |
| D-Flip-Flops | 6,144 | 24,576 |
| RAM Kbits | 36 | 144 |
| Flash-ROM | 1K | 1K |
| Secure (AES) ISP | Yes | Yes |
| Integrated PLL | 1 | 1 |
| Global Signals | 18 | 18 |
| I/O Banks | 4 | 4 |
| Single-Ended I/O | 151 | 154 |
| Differential I/O Pairs | 34 | 35 |

Table 1. Features of the Selected DUTs: the A3P250 and the A3P1000. Both are mounted on a PQ208 package.

A new test setup was built for the A3P radiation testing. As shown in Fig. 12, it includes two boards: 1) a "master" board for the monitoring and control of the DUT operation in-beam and 2) a "slave" board for the communication between the host PC and the master board through two USB ports. The "master" board includes an A3P1000-FG484, called "master" FPGA, and a DUT (A3P-PQ208). IO "channels" of an input (SE or LVDS) routed immediately to a nearby output are also added between the "master" FPGA and the DUT. There are 38 SE and 13 LVDS I/O channels on both FPGAs. This board architecture allows the implementation of several separate designs on the same DUT to be tested simultaneously. The slave board includes an A3P1000-PQ208; it allows the data acquisition and data transfer to the host PC.

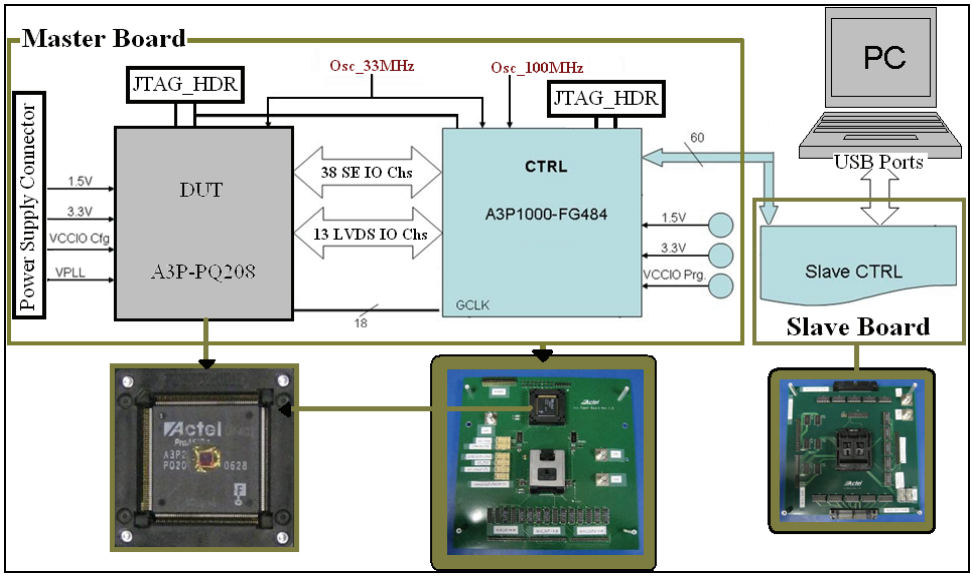



Fig. 12. Block diagram of the A3P Test Setup. It includes two boards: a "master" board for the monitoring of the DUT operation in-beam and a "slave" board for the communication between the host PC and the master board.

For communication with the host PC, a generic user interface was designed to communicate with the slave board. The communication protocol between the slave board and the host PC remains always the same for easy and fast implementation of any new SEE test experiment. Indeed, there are always a maximum of 64 display counters available to the designer, which names are adjustable according to the running experiments. These counters are usually used for display of number of SEE events among other indicators of the operation of the DUT design. In addition, this user interface allows the self-monitoring of the test system itself, by testing each board and FPGA individually as shown in the "Mode" knob on the top left of Fig. 13. Among other features, it also allows the pattern selection to be accomplished by the "pattern" knob (all zeroes, all ones, checkerboard or inversion of checkerboard) exercised on the DUT inputs and the frequency at which the DUT design is running by using the "Frequency" knob.

### 4.3 Test designs and experimental results

#### 4.3.1 FPGA core SEE characterization (flip-flops)

The purpose of this testing is to determine the SEE cross-section of an A3P logic tile configured as a DFF. This should lead to the highest possible upset cross-section of a logic tile. The basic test design is a shift register (SR) using 86 logic tiles with each one of them configured as a DFF and one global clock signal but no reset signal. Note that if the SR design was using a reset line, this signal would be a global and using a global IO pad in the same way as any other global clock signal, whose cross-section will be given below.

On the other hand, since this is a 0.13-μm technology, the part might be sensitive to Multiple Bit Upsets (MBU) [Quinn et al., 2005], which in some cases cannot be mitigated effectively by TMR. For instance, if the MBU affects two TMR paths out of three, the output TMR result

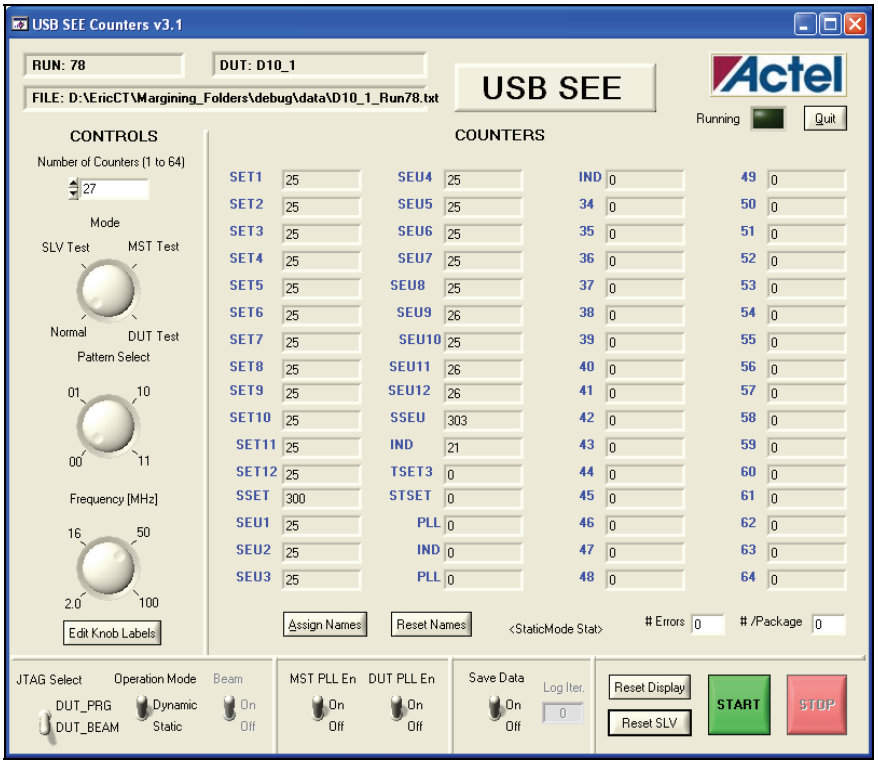



Fig. 13. SEE Software User Interface with a maximum of 64 display counters.

will be wrong. Therefore using TMR as a test methodology constitutes a good approach to detect some of the MBU or SEE on the FPGA's global signals. Note that the design should be using at least 99% of the FPGA resources and the three paths of a TMR circuit should be as close as possible to simulate the worst case of a TMR implementation. Hence in addition to the version (D1) having SR without mitigation, two versions of the TMR'd design have been implemented on the same DUT: 1) D2: TMR'd SR using one single global clock, where voters and IOs are also tripled and 2) D3: TMR'd SR where every I/O signal is tripled, including the global clock signal. All three flip-flops of a TMR'd DFF are always placed directly next to each other.

*4.3.1.1 Test Design*

Among the 37 Single-Ended (SE) channels, the non-mitigated test design D1 uses 28 SE channels of the DUT. Between each input/output of these 28 channels, a shift register (86 DFF) is inserted. In total, the D1 design uses 28 Input/Output and 2408 (86×28) DFF. D2 uses three copies of a TMR'd SR with no triplication of the clock signal, i.e. nine SE channels and one global clock, while D3 uses 4 copies of the TMR'd SR, i.e. 12 LVDS IO channels and 3 global clocks. D1 and D2 use 2 SE IO banks and D3 uses two LVDS IO Banks. The three versions of the design occupy 98% of the A3P250-PQ208. A detailed block diagram of these 3 design implementations, D1, D2 and D3, is given in Fig. 14. The testing was performed at the clock frequency of 2, 16 and 50 MHz.

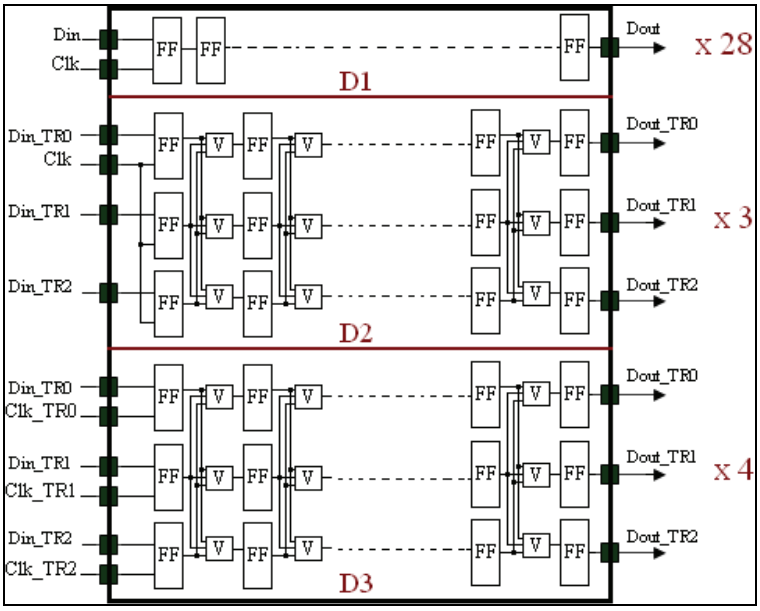



Fig. 14. Block Diagram of D1, D2 and D3 Test Designs. D1 uses 2048 FFs. D2 uses three copies of a TMR'd SR with no triplication of the clock signal and D3 uses four copies of a TMR'd shift-register.

Note that implementing the same design D1, D2 or D3 on several channels will help check the repeatability and the consistency of the tests for its non-dependency of different tested channels. Moreover, it allows checking for SEE on common global signals other than the user global clock and reset signals. For example, an SEE in global signals that link an IO bank can cause a simultaneous soft error in every channel using the same IO bank [Rezgui et al., 2007a]. Indeed, a transient event was observed on all the IO channels belonging to a single IO bank with a cross-section of $2.37\times10^{-6}$ cm² per IO-bank. The threshold LET of this event is around 7 MeV•mg/cm². This suggests that if a design is using all the tripled IOs in the same bank, its cross-section will be no less than $2.37\times10^{-6}$ cm² per IO-bank.

*4.3.1.2 HI and Protons Beam Test Results*

For the Design D1, the obtained HI results showed three types of errors: 1) single error on one channel, 2) multiple errors on one single or few channels, and 3) single or multiple errors on all the IO channels associated to a common IO bank. All errors were transient and did not require any reconfiguration or power cycle of the FPGA. Type 1 was most likely due to an SEU in the DFF or to an SET in the clock signal associated to this DFF. Type 2 could be due to the clock signal or to another global signal besides the IOs since we didn't see all the IO channels disrupted at the same time. Type 3 was most likely due to the aforementioned event for the IO testing and observed in a single IO bank. Fig. 15 shows the single DFF cross-sections at three different frequencies obtained from D1-test data. There was no dependency of cross sections on the frequency; this was expected for soft errors in the flip-flops when the static SEU rate dominates. Note that for better visibility, WEIBULL curves in Fig. 15 (also in Fig. 16 and 17) have been drawn only for the 50MHz data.

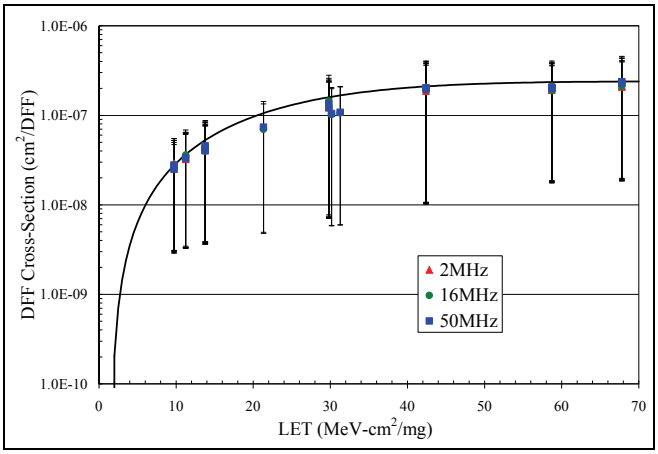



Fig. 15. A3P250-PQ208 DFF Cross-Section at three different frequencies (2, 16 & 50 MHz).

Although not visible in Fig. 15, these data include global error cross-sections due to the IO bank or clock global signals; this subject will be discussed in detail in the following section. The global-error cross-sections are dependent on the clock frequency because they are due to the SET in the IO bank or clock global signals. It is well known that SET induced errors have a strong dependence on the clock frequency [Berg et al., 2006]. For the design D2, only errors type 2 and 3 have been observed, while for D3 only errors type 3 have been observed, which means that each SEE observed on the TMR'd design (D3) always affected an entire IO bank. To compare the SEE response of the three test designs and to validate the efficacy of the increase of mitigation level, TMR of the DFF and the triplication of the global clock signal, the SEE cross-sections were averaged on three channels for each design, since D2 was using only three channels. These cross-sections are given in Fig. 16. It is clear that increasing the frequency increases the SEE cross-sections of D2 and D3.

Fig. 16 shows a clear reduction in the SEE cross-sections from D1 to D2 and finally to D3 with the increase of the level of mitigation. In addition, the results show that each observed error on the design D3, where all the resources have been TMR'd, always originates from an SET which affects an entire IO bank. The cross-section of the TMR'd design ($4\times10^{-6}$ cm$^2$ per design) in D3 is very close to twice the IO-bank SET cross-section deduced from SET errors in designs D1 and D2. This is expected because D3 uses the banks 1 and 3 for the differential IOs while D1 or D2 only uses the bank 2 for single-ended IOs. The IO-bank-SET is suspected to be due to SET occurring on the enable signal of a single IO bank. To accomplish complete SEE immunity, all the tripled IOs have to be separated on three different IO banks; this had been fully demonstrated in [Rezgui et al., 2007a].

Furthermore, if we increase the number of usage of the FPGA core of D2 and D3, the SEE cross-sections should not increase because they are are dominated by SET on the global signals, i.e. Clock or IO bank enable signals. These cross-sections depend on the number of used global clock signals (18 maximum), the used IO banks (4 maximum for the A3P and 8 for the A3PE) or the operation frequency. On the other hands, if the usage of resources of D1 should increase, its cross-section should increase linearly. Note that for the design D1, the events where all the disrupted IO channels are not counted for this comparison. Fig. 17 shows the clock global cross-section; it is acquired simply by measuring the difference between the designs D2 and D3.

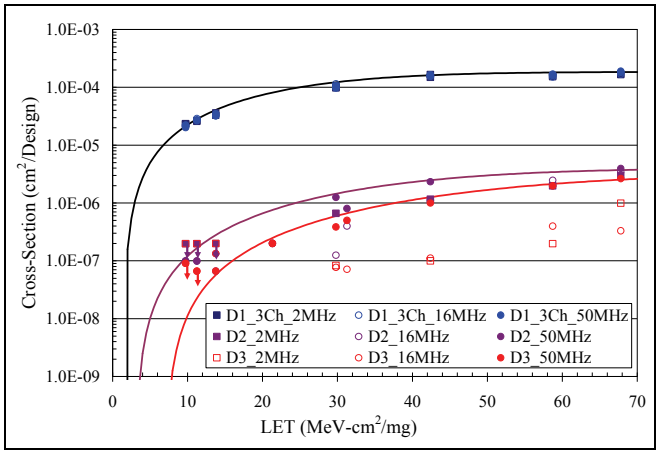



Fig. 16. D1, D2 and D3 SEE Cross-Sections at 2, 16 and 50 MHz.

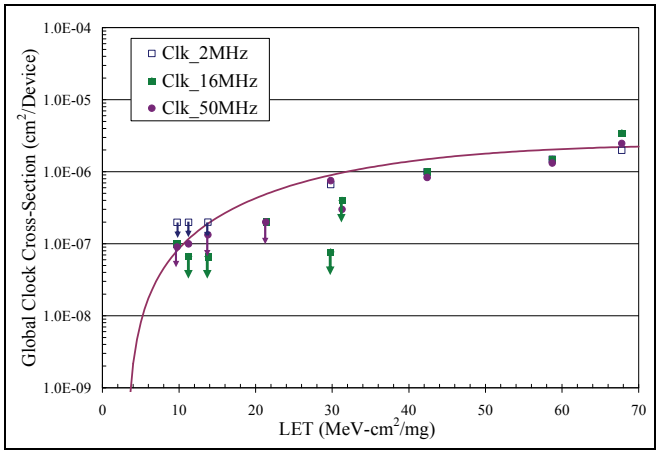



Fig. 17. A3P250-PQ208 Global Clock Cross-Section. This SET cross-section is very similar to an IO Bank cross-section, proving that most SETs inducing errors on the clock network are due to SET on the IO bank.

On the other side, proton-beam test experiments showed very little SEE sensitivity at proton energy of 63.5 MeV and when running the design at 50 MHz. Indeed, the DFF SEU cross-section was measured at $5.18 \times 10^{-14}$ cm²/DFF. Note also that at this energy and for a fluence of $6.49 \times 10^{12}$ of proton particles, no SET in the configuration logic tiles, on the enable signal of the IO banks, on the IOs themselves or the global clock signal was observed. Because of such low SEU cross-section, the DFF design was not tested at lower energies, although it is advised to measure the threshold energy for the A3P DFF in future experiments. No errors were observed on the TMR'd channels, proving the efficacy of the TMR technique in fully mitigating SEUs. Automated software SEU mitigation, a user-selected TMR for the design's registers, is offered for the RT3P FPGAs.

### 4.3.2 PLL SEE characterization

A PLL macro uses the CLKA input to drive its reference clock. It uses the GLA and optionally the GLB and GLC global outputs to drive the global networks (Fig. 18). A PLL macro can also drive the YB and YC regular core outputs, but if the GLB (or GLC) global output is used, the YB (or YC) output [ProASIC3 Handbook] cannot be reused. The purpose of this test design is the identification of all the PLL error modes due to beam irradiation.

*4.3.2.1 Test Design*

The test design uses a PLL whose output (GLA) clocks a triple DFF. Its input signal CLKA is using the 33MHz oscillator output and its GLA signal is running at 50 MHz. The three DFFs have three different inputs and three different outputs. The only common point between the three of them is the PLL output clock signal (DUTCLK). On the master FPGA, the three outputs of the DUT DFF are voted and their output is compared continuously with the DFF input provided from the master FPGA, which is clocked at 16 MHz. Any mismatch between the DFF voted value and the expected value (the input value), is counted as an error.
The test design allows also the monitoring of the PLL LOCK signal. This signal should always be high indicating that the PLL is working properly; if it goes low then the PLL is unlocked and this will also be counted as an error. The objective of these radiation tests is the classification of the detected error types and the test of the efficiency of self-correction through the PLL POWERDOWN signals (Fig. 18) without having to power-cycle the entire FPGA.

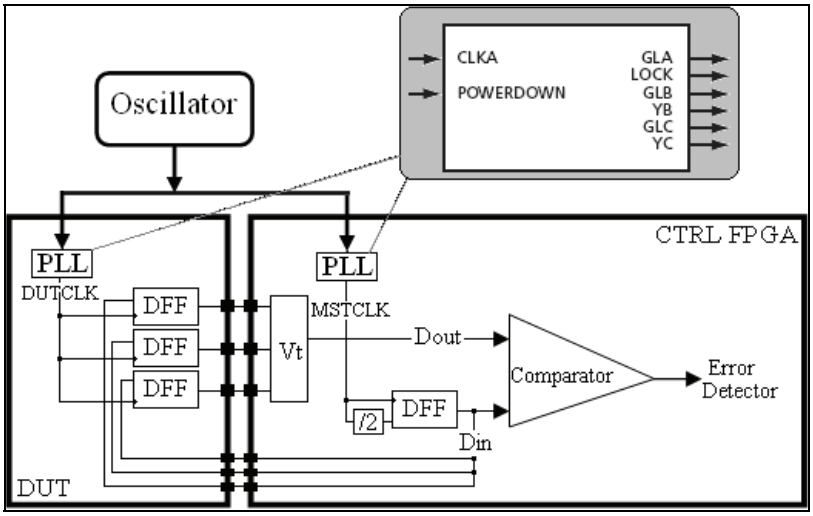



Fig. 18. Block Diagram of the PLL Test Design

The test design is implemented so six types of errors, called error-type 1 to error-type 6 summarized in Table 2, can be detected during the beam test experiments. In the case of a mismatch between the Din and Dout signals of Fig. 18, the error would be counted as an error-type 1, which is similar to an SET event on the PLL clock signal if the error does not persist. However, if the error persists for longer than two clock cycles but less than 100 cycles, it will be counted instead as error-type 2. If the same error persists for longer than 100 clock cycles, it will be considered as error-type 3 and the master FPGA will then power cycle the PLL through the POWERDOWN signal and restart normal operation.

| Error Type | Error Description |
|---|---|
| 1 | An SET has occurred on the DUTCLK signal. |
| 2 | A mismatch between Din and Dout that lasts less than 100 clock cycles. |
| 3 | A mismatch between Din and Dout that lasts longer than 100 clock cycles. |
| 4 | An SET has occurred on the LOCK signal. |
| 5 | The LOCK signal remains at '0' for less than 100 cycles and the PLL recovers by itself. |
| 6 | The LOCK signal remains at '0' for more than 100 cycles and the PLL can not recover by itself. |

Table 2. PLL Error Modes in Beam.

Simultaneously, the master FPGA is continuously checking for the status of the PLL LOCK signal. If this signal goes low, the master FPGA counts it as an SET on the LOCK signal (error type 4) and waits for 2 clock cycles. If the LOCK signal remains at '0' logic for less than 100 clock cycles and the PLL recovers by itself then the error is counted as a PLL lock case and considered instead an error-type 5. In the case where an error-type 5 would last longer than 100 cycles, it will be considered as an error-type 6 and the master FPGA would then power cycle the DUT PLL through the POWERDOWN signal. The block diagram of this test design is given in Fig. 18. Note that the actually implemented test design runs the DUT design at 50 MHz while the error checking on the master side is at 16 MHz.

*4.3.2.2 HI and Protons Beam Test Results*

The MSTCLK was exercised at two frequencies (2 and 16 MHz). In both cases, among the six expected types of errors, only two have been observed: errors from type 2 and 6. The latter was always combined with a difference between the Din and Dout signals lasting for more than 100 clock cycles. Only toggling the PLL POWERDOWN signal could restart the operation of the PLL in that case. As shown in Fig. 19, the test results indicate little variation between the cross-sections of error-type 6 obtained at both test frequencies (2 and 16 MHz). Error type 2 has been observed only at 16 MHz (frequency of the master FPGA). The $LET_{th}$

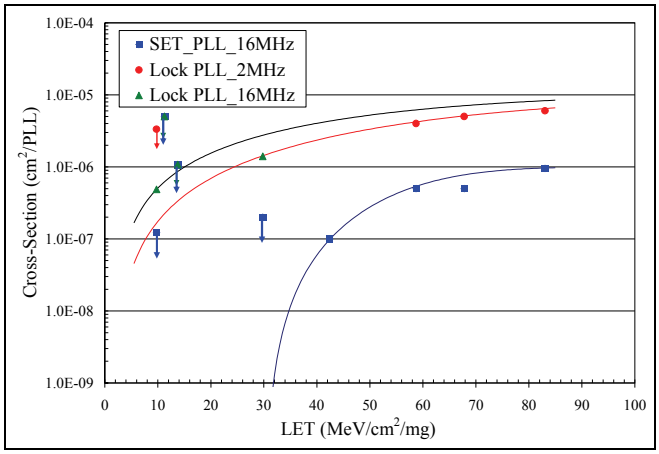



Fig. 19. A3P250-PQ208 PLL SEE Cross-Section

for this type of errors is shown in Fig. 19 to be around 32 MeV-cm²/mg. This value might seem high if the SET on the clock signal generated from the PLL occurred on the FG switches that links this signal to the tripled DFF. However, it might be expected if it is related to the internal PLL circuit. Only collecting more data could clarify this point. The saturation cross-section of the PLL in LOCK mode is $10^{-5}$ cm².
Finally, no SEE was observed on the PLL during beam irradiation tests for a fluence of $9x10^{10}$ of proton particles having energy of 63.5 MEV, which was expected considering the low sensitivity of the FPGA core itself.

### 4.3.3 Flash-ROM (FROM) memory SEE characterization

*4.3.3.1 Test Design*

ProASIC3 devices have 1 kbits of on-chip nonvolatile Flash memory that can be read from the FPGA core fabric. The Flash ROM is arranged in 8 banks of 128 bits during programming. The 128 bits in each bank are addressable as 16 bytes during the read back of the FROM from the FPGA core. The FROM will be configured initially with a pattern that reflects the byte address and the master FPGA will be simply checking its content. The frequency of the FROM read was varied between 2 and 16 MHz to check the speed effects and quantify the number of SETs that had occurred during the beam testing. The FROM was read during and after irradiation. In beam, each FROM address was read 3 times successively to avoid counting SEE on the peripheral gates (7 DFF automatically connected to FROM address bus, 8 DFF connected at the data outputs, routing switches and active regions of the IO pads).

*4.3.3.2 HI and Protons Beam Test Results*

The data showed no observable SEE sensitivity on the FROM during beam irradiation tests for LET < 83 MeV•cm²/mg (Fig. 20) and for a fluence of $9x10^{10}$ of proton particles having an energy of 63.5 MeV. This demonstrates the SEE hardness of the embedded FROM and opens its possibilities for space applications; for example it can be used as a boot memory for the embedded processors in the A3P FPGA.

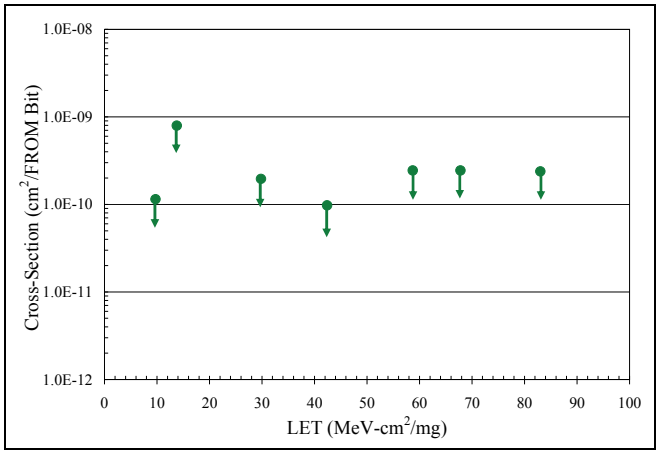



Fig. 20. FROM Bit SEU Cross-Section. The data shows non-sensitivity to SEE in HI beams.

### 4.3.4 SRAM memory SEE characterization

The selected ProASIC3 devices (A3P250 and A3P1000) have embedded SRAM blocks along the north and south sides of the devices. To meet the needs of high-performance designs, the memory blocks operate strictly in synchronous mode for both read and write operations. The read and write clocks are completely independent and each may operate at any desired frequency up to 350 MHz. To have better statistics, an A3P1000 was used as the DUT, which has 144 Kbits of SRAM bits, four times more than that in an A3P250.

During beam-test experiments, the "master" FPGA initially writes a checkerboard pattern into the embedded SRAM and continuously checks its contents. When an upset is detected in the SRAM bits, the upset counter is incremented and the memory content is flipped back. Note that for ease of implementation, only one organization of SRAM was used: "RAM512x9". In the DUT design, all the logic used to interface with the SRAM, such as IOs, address decoder, read and write signals of the 32 SRAM blocks, are TMR'd and therefore mitigated to SEE. This means also that only SEE on the SRAM will be counted. This should avoid the overestimation of the SRAM SEE cross-section due to the SEE sensitivity of other programmable circuits used in the DUT test design. In this test, the maximum SRAM frequency is 16 MHz. The block diagram of the test design is given in Fig. 21.

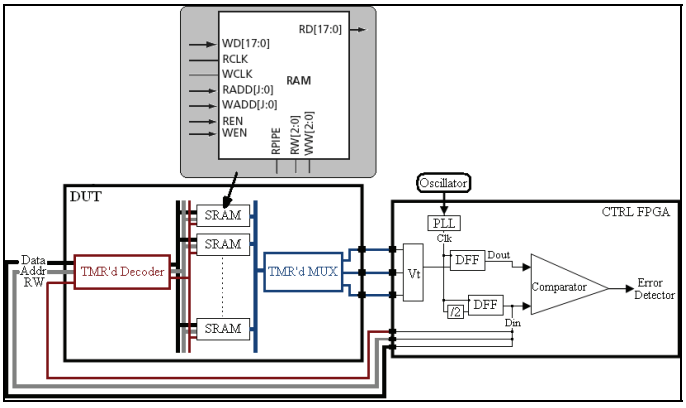



Fig. 21. Block Diagram of the SRAM Test Design

The test results show no SRAM SEE cross-section dependence on the frequency, indicating that most of the SET effects on the peripheral combinational logic are filtered out and only SEU on the SRAM blocks are counted. Also, no MBU were observed in the SRAM bits. Measured SEU cross-sections are given in Fig. 22. The saturation cross-section is approximately $4.22 \times 10^{-8}$ cm$^2$/SRAM-bit.

The LET threshold is around 0.5 MeV-cm$^2$/mg, which is considered very low, but correlate well with the published SEU cross-sections of the Virtex-II configuration bits from Xilinx, since they are also based 0.13-µm SRAM bits [Rezgui et al., 2004]. It should be mentioned also that additional testing should be done to find out about MBU in the SRAM blocks. Static tests should be used where the SRAM is read at the end of each run preferably irradiated at low fluxes to avoid hiding some of the bit-errors because of multiple hits. SEE mitigation solutions for the SRAM, based mainly on EDAC approach such as the one employed for the SRAM of the RTAX-S FPGAs [Wang et al., 04b] have been implemented and are ready to use for the RT3P products.

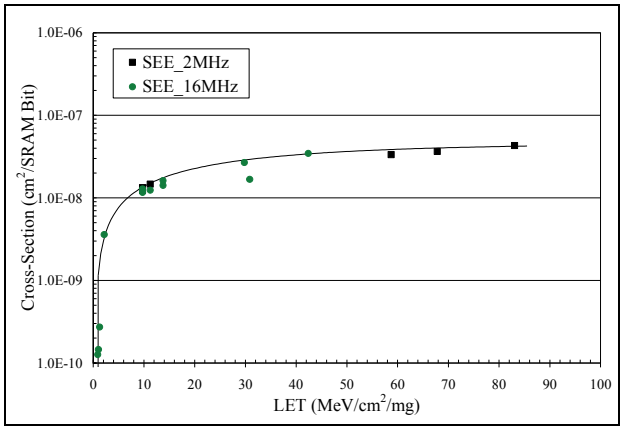



Fig. 22. HI SRAM Bit SEU Cross-Section

Finally, in comparison with the other FPGA resources, the embedded SRAM blocks when operated at 16 MHz, showed an SEU cross-section in protons beams, for a cocktail of energies of 63.5, 30, 19.5 and 16.5 MEV. The obtained results are shown in Fig. 23. Additional tests shall be performed to establish the threshold proton energy to induce upsets in the SRAM bits.

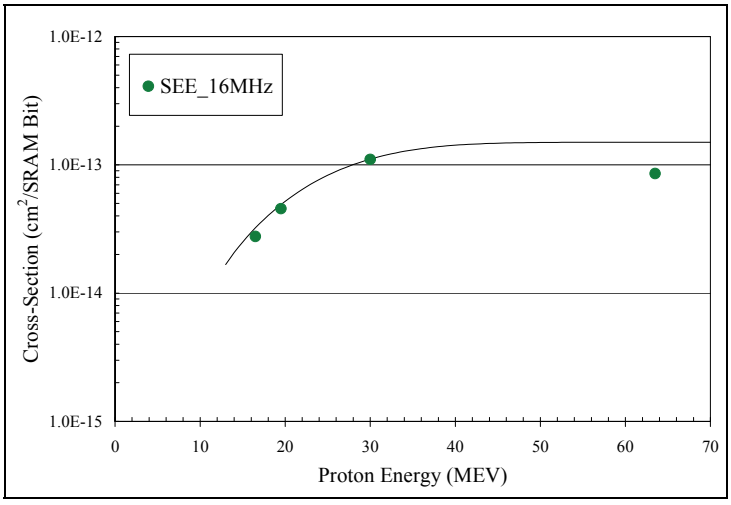



Fig. 23. Proton SRAM Bit SEU Cross-Section

## 5. Preliminary studies of TID effects on SEE sensitivities

### 5.1 Proton characterization of the programming and erase circuitry

One major advantage of the Flash-based FPGAs compared to the previous generation of ACTEL FPGAs, based on the Antifuse technology, is the re-programmability feature. However, during erase and reprogramming of the part, high voltages are applied (±17.5 V)

and one might think that there is a risk of permanent damage on the FG cells or other overhead circuitry if an ion hit during that mode. Therefore, radiation test experiments during the erase and the programming of this part are required to measure the SEE sensitivity of this specific part of the FPGA (charge pumps) and the overall consequences from an ion hit.

Ten A3P250-PQ208 circuits have been exercised in proton beams during the erase, reprogramming and verification of the programmed FG cells. The shift register design using 98% of the FPGA logic tiles (A3P250-PQ208) was used as a reference design. For each beam run, consecutive erase, reprogramming and verify cycles are launched and the functionality of the design is always checked at the end of each run. At least four full cycles of erase, program and verify cycle are executed during each beam run; each cycle requires 41 seconds. Each run exposes a new DUT to a dose of 13.4 Krad due to proton beam exposition and uses a fluence of $10^{11}$ of proton particles. Table 3 summarizes the obtained results.

| Behavior Type | Error Description | # DUTs |
|---|---|---|
| 1 | All 4 programming and erase cycles have passed successfully | 9 |
| 2 | One erase/program cycle among 4 failed and the next one passed | 1 |
| 3 | Failure of the 5th cycle of erase / programming because of total exposure to TID (13.4 Krad) requiring annealing | 2 |

Table 3. Programming and Erase Error Modes in Proton Beams

Three types of behavior have been observed during the proton irradiation testing, as summarized in Table 3. Type 1 is showing the case where four erase, programming and verifying cycles have been performed without any failure including the design's operation. Type 2 shows the one case where only one verifying failure has been observed (second cycle), which could be due to the programming of false information in the FG cells (ON state instead of OFF state and vice versa). This type of errors was easily mitigated by running a second cycle of erase, reprogramming and verifying of the FG cells allowing the DUT to recover normal operation and has a cross-section of $10^{-12}$ $cm^2$/FPGA. Type 3 is the one where a fifth cycle was started and did fail because we reached a dose of 13.4 Krad, which is considered high for the normal operation of the charge pump circuit, according to TID tests in gamma-rays at DMEA, shown above in the Section 3 and also considering the high dose rate exercised in this case (58 rad/s).

During all these runs, there was no permanent damage on the circuit and all errors that have been observed during these test cycles disappeared after annealing. Indeed, the two parts that have failed programming on the 5th time recovered functionality after annealing of the DUT at room temperatures for many days. Although these preliminary results are encouraging and since the annealing effects on the floating gates are still under study, it is well-advised to avoid erasing and reprogramming the DUT in or off-beam after its exposure to a dose higher than 16 Krad. This statement is valid only if the applied dose rate from heavy ions, protons or gamma is around 50 rad/s as required by the JEDEC test standards [MIL-STD-883G, TM1019.7]. In the case of the actual protons testing, the dose rate was around 58 rad/s, which might explain the observation of some failures on the 5th cycle of erase and programming at 13.4 Krad. Also the cross-section of writing wrong information

($10^{-12}$ $cm^2$/FPGA) could be fundamentally due to the very little SEE sensitivity to protons of the A3P FPGA. Heavy ion data is hence required to confirm that no catastrophic failures could result from programming and erasing in beam since the FPGA's SEE sensitivities under HI irradiation are much higher relative to the proton sensitivity.

### 5.2 Testing beyond the TID limit

Most of the collected data for the measurements of the SEE cross-sections in this chapter has been obtained for TID less than 25 Krad in gamma-rays. Data provided in the Section 3, showed the TID performance of this device to be 16 Krad for the programming and erase circuitry and 22 Krad for the FPGA core itself (the FG cells). For the latter, the TID performance was mainly obtained when a degradation of 10% in the propagation delay of the logic tiles configured as a chain of buffers is attained, but no permanent damage on the FPGA was noted.

The purpose of this new specific test is to check the designs' functionality and their SEE performance for TID higher than 25 Krad as well as the maximum TID to which the design is still functional. The SRAM test design was selected for this study, since it uses various resources of the FPGA: 8.24 % of the FPGA logic tiles (configured as combinational or sequential logic), 100 % of the embedded SRAM memories, the embedded PLL and FROM and 44 % of the IOs. This design was also selected because of the SRAM high SEE sensitivity compared to the other FPGA resources, which could help monitoring the functionality and the SEE cross-sections if they do increase.

The DUT was exposed to beam for 5 consecutive runs, each at a fluence of $4x10^{10}$ of 16.5 MeV proton particles. This corresponds approximately to a TID of 15 Krad per run, and to a total of 75 Krad for the five runs. During all these runs, the DUT design was functional and the error cross-section per run was consistent without any noticeable increase in the SEE sensitivities as shown in Table 4. It should also be noted that for all of the five runs, the detection of errors stops with the end of the beam time. This confirms that the FG cells are still functional upon a TID of 75 Krad. However, upon the start of the $6^{th}$ run, the design stopped functioning, which could be due to a high charge loss in the FG cells. After four months of annealing in room temperature, the design did recover functionality but not the reprogramming capability. Time is needed to check if more annealing time will allow the recovering of the full operation of the charge pumps needed for the FPGA re-programming.

| Run | Accumulated TID [Krad] | SRAM Bit SEE Cross-Section [MeV-$cm^2$/mg] | Fluence [16.5 MEV Proton-Particles] |
|---|---|---|---|
| 1 | 15 | $2.48x\ 10^{-14}$ | $4x10^{10}$ |
| 2 | 30 | $2.29x\ 10^{-14}$ | $4x10^{10}$ |
| 3 | 45 | $2.51x\ 10^{-14}$ | $4x10^{10}$ |
| 4 | 60 | $2.80x\ 10^{-14}$ | $4x10^{10}$ |
| 5 | 75 | $2.71x\ 10^{-14}$ | $4x10^{10}$ |
| 6 | 90 | Design lost functionality right in the beginning of the run but recovered after annealing in room temperature | $4x10^{10}$ |

Table 4. TID Effects from Proton Irradiation (Energy = 16.5 MEV) on the SEE Cross-Sections of an SRAM-Bit

It should also be stated that an accurate estimation of the TID effects on the SEE cross-sections requires a better measurement of the accumulated dose. Indeed, until today, only gamma rays could provide an accurate measurement of the exposed dose and therefore it would be advised to expose the part to a certain dose in gamma rays and then measure the SEE cross-sections, within 2 hours or few days if transported in dry ice to avoid annealing effects.
In addition, it should be mentioned also that among the 60 parts, tested in all the HI experiments, 59 of them have recovered the DUT programming and erasing capabilities after many months of annealing in room temperature and did never loose functionalities in or off-beam. The TID for the 59 parts varied between 5 and 40 Krad. The only DUT that did not recover yet the programming capability was exposed to a TID of 41.5 Krad. Knowing that after annealing, we could erase this part led us to assume that we might need more time to be able to reprogram it again. On the other side, all of the 24 parts that have been tested in protons could be erased but seven of them could not be reprogrammed. Time is needed to make sure that the seven remaining parts will recover this feature.
The main conclusion from these test experiments is that most of the tested parts did recover the programming and erase features after annealing in room temperature for many months. None of them lost functionality for dose that approximate 40 Krad even at the highest LET (83 MeV-cm$^2$/mg) or 63.5 MeV in protons. It is clear though that the recovering of the erase functionality is much quicker than the recovering of the programming capability. This is certainly not a quantitative study but rather qualitative to make sure that there is no permanent damage from HI or protons on the part due to TID. Additional testing is hence mandatory to calculate accurately the annealing effects on the FG cells and the circuitry used for the erase and the reprogramming of the FPGA. More work has been done since to show and explain the annealing effects on the Flash-memories [Bagatin et al., 09].

## 6. Conclusion

This chapter detailed the extensive radiation tests of the new Radiation-Tolerant Flash based-FPGAs (RT ProASIC3) to determine its sensitivities to TID and SEE as well as some suitable methodologies for its mitigation to these effects. Based on the measurements of the degradation in the propagation delay of an inverter-string, the TID performance of the RT3P was characterized to be 22 Krad. However, if programming in space is allowed then the TID limit of this part can be improved to 40 Krad. Note that safe reprogramming of the RT3P FPGAs is allowed only till 16 Krad because of the TID effects on the programming control circuits.
Furthermore, the obtained results from the SEE characterization showed some radiation sensitivity in most of the programmable architectural features of the FPGA; the exception is the embedded FROM, which is very radiation hard. If mitigation solutions of TMR and SET filtering are adopted for the logic and clock in A3P FPGA, the only remaining cross-section would be due to the transient event on the IO banks used for SE or LVDS IOs observable mostly at high frequencies. On the other hand, if a complete SEE immunity is required at high frequencies (50 MHz and above), triplication of IOs is mandatory in addition to their separation on three different IO banks. Finally, as expected for a non-volatile FPGA, no observed error-event required a reconfiguration of the Flash-based FPGA nor were there any destructive SEE events even during the erase, the programming and the verifying of the

FPGA. SEU mitigation by software user-selective-TMR and software Intellectual Property (IP) to implement EDAC for the embedded SRAMs are available to the user of the Radiation-Tolerant Flash-based FPGAs, guaranteeing its full-immunity to SEUs.

## 7. Acknowledgements

I am greatly indebted to JJ Wang, Brian Cronquist, John McCollum, Minal Sawant, and Ken O'Neill from Actel Corporation for many fruitful discussions and suggestions. I would like to express a great thanks to Natalie Charest, Eric Chan Tung, Yinming Sun, and Durwyn D'Silva from the University Of Toronto, Canada, who worked as intern-students in Actel Corporation each for a year period.

## 8. References

Bagatin, M.; Gerardin, S.; Cellere, G.; Paccagnella, A.; Visconti, A.; Bonanomi, M.; Beltrami, S. (2009) "Error Instability in Floating Gate Flash Memories Exposed to TID", *NSREC 2009*, to be published at *IEEE TNS, Vol. 56, NO. 6*, Dec. 2009, Quebec City, Quebec, Canada.

Baze, M. P.; Wert, J.; Clement, J. W.; Hubert, M.G.; Witulski, A.; Amusan, O.A.; Massengill, L. & McMorrow, D. (2006). "Propagating SET Characterization Technique for Digital CMOS Libraries", *IEEE TNS, Vol. 53, NO. 6*, Dec. 2006, pp 3472-3478.

Berg, M.; Wang, J.J.; Ladbury, R.; Buchner, S.; Kim, H.; Howard, J.; Label, K.; Phan, A.; Irwin, T. & Friendlich, M. (2006). "An Analysis of Single Event Upset Dependencies on High Frequency and Architectural Implementations within Actel RTAX-S Family Field Programmable Gate Arrays", *IEEE TNS, Vol. 53, NO. 6*, Dec. 2006, pp 3569-3574.

Brown, W.D. & Brewer, J. (2002). "Nonvolatile Semiconductor Memory Technology: A Comprehensive Guide to Understanding and Using NVSM Devices", IEEE Press Series on Microelectronic Systems.

Carmichael, C. (2001) "Triple Module Redundancy Design Techniques for Virtex FPGAs, "Xilinx Application Note XAPP197, Nov. 2001, available at www.xilinx.com/bvdocs/appnotes/xapp197.pdf

Balasubramanian, A.; Bhuva, B.L.; Black, J.D.; Massengill, L.W. (2005). "RHBD Techniques for Mitigating Effects of Single-Event Hits Using Guard-Gates", IEEE TNS, Vol. 52, NO 6, Dec. 2005, pp 2531 – 2535.

Cellere, G.; Paccagnella, A.; Visconti, A.; Bonanomi, M.; Caprara, P. & Lora, S. (2004). "A Model for TID Effects on Floating gate Memory Cells", *IEEE TNS, Vol. 51, NO. 6*, Dec. 2004, pp 3753-3758.

Guertin, S.; Nguyen, D. & Patterson, J. (2006). "Microdose Induced Dose Data Loss on floating Gate Memories", *IEEE TNS, Vol. 53, NO. 6*, Dec. 2006, pp 3518-3524.

MIL-STD-883G, TM1019.7, http://www.aspentechnologies.com/files/rur89tn69a.pdf

Mitra, S.; Zhang, M.; Seifert, N.; Gill, B.; Waqas, S.; Kim, K.S. (2006). "Combinational Logic Soft Error Correction", IEEE ITC, Nov. 2006.

Mavis, D. & Eaton, P. (2007). "SEU and SEU Modeling and Mitigation in Deep-Submicron Technologies", IRPS 2007, pp 293-305, Albuquerque, USA.

Morris, K. (2006). *FPGA Journal*, available at: http://www.fpgajournal.com/articles_2006/pdf/20060829_igloo.pdf

Palkuti, L.J.; & LePage, J.J. (1982). "X-rays Wafer Probe for Total Dose Testing", *IEEE TNS, Vol. 29, NO. 6*, Dec. 1982, pp 1832-1837.

ProASIC3 FPGA Handbook, available at: http://www.actel.com/documents/PA3_HB.pdf

Quinn, H.; Graham, P.; Krone, J.; Caffrey, M. & Rezgui, S. (2005). "Radiation-Induced Multi-Bit Upsets in SRAM-Based FPGAs", IEEE TNS, Vol. 53, NO. 6, Dec. 2005, pp. 2455-2461.

Rezgui, S.; Swift, G. & Xilinx SEE Consortium (2004). "Xilinx Single Event Effects First Consortium Report Virtex-II Static SEU Characterization", available at: http://parts.jpl.nasa.gov/docs/swift/virtex2_0104.pdf

Rezgui, S.; Wang, J.J.; Chan Tung, E.; McCollum, J. & Cronquist, B. (2007). "New Methodologies for SET Characterization and Mitigation in Flash-Based FPGAs", *IEEE TNS, Vol. 54, NO. 6*, Dec. 2007, pp 2512-2524.

Rezgui, S.; Wang, J.J.; Sun, Y.; Cronquist, B. & McCollum, J. (2008a). "New Reprogrammable and Non-Volatile Radiation Tolerant FPGA: RTA3P", *IEEE Aerospace 2008*, Big Sky, MT.

Rezgui, S.; Wang, J.J.; Sun, Y.; Cronquist, B. & McCollum, J. (2008b). "Configuration and Routing Effects on the SET Propagation in Flash-Based FPGAs", *IEEE TNS, Vol. 55, NO. 6*, Dec. 2008, pp 3328-3335.

Rezgui, S.; Wang, J.J.; Won, R. & McCollum, J. (2009) " Design and Layout Effects on SET Propagation in ASIC and FPGA 90-nm Test Structures", *NSREC 2009*, to be published at *IEEE TNS, Vol. 56, NO. 6*, Dec. 2009, Quebec City, Quebec, Canada.

Snyder, E.S.; McWhorter, P.J.; Dellin, T.A. & Sweetman, J.D. (1989). "Radiation Response of Floating Gate EEPROM Memory Cells", *IEEE TNS, Vol. 36, NO. 6*, Dec. 1989, pp 2131-2139.

Shuler, R.L.; Kouba, C. & O'Neill, P.M. (2005). "SEU Performance of TAG Based Flip-Flops", IEEE TNS, Vol. 52, NO 6, Dec. 2005, pp 2550 – 2553.

Shuler, R.L.; Balasubramanian, A.; Narasimham, B.; Bhuva, B.L.; O'Neil, P.M. & C. Kouba (2006). "The effectiveness of TAG or Guard-Gates in SET Suppression Using Delay and Dual-Rail Configurations at 0.35 um", *IEEE TNS, Vol. 53, NO. 6*, Dec. 2006, pp 3428 -3431.

Swift, G.; Rezgui, S.; George, J.; Carmichael, C.; Napier, M.; Maksimowictz, J.; Moore, J.; Lesea, A.; Koga, R. & Wrobel, T.F. (2004) "Dynamic Testing of Xilinx Virtex-II Field Programmable Gate Array (FPGA) Input/Output Blocks (IOBs)", *IEEE TNS, Vol. 51, NO. 6,* Dec. 2004, pp 3469-3479.

Wang, J.J.; Samiee, S.; Chen, H.S.; Huang, C.K.; Cheung, M.; Borillo, J.; Sun, S.N.; Cronquist, B. & McCollum, J. (2004a). "Total Ionizing Dose Effects on Flash-based Field Programmable Gate Array", *IEEE TNS, Vol. 51, NO. 6*, Dec. 2004, pp 3759-3766.

Wang, J.J. (2004b) "RTAX EDAC-RAM Single Event Upset Test Report", available at: http://www.actel.com/documents/RTAX-S%20SEE%20EDAC%20RAM.pdf

Wang, J.J.; Kuganesan, G.; Charest, N. & Cronquist, B. (2006a). "Biased-Irradiation Characteristics of the Floating Gate Switch in FPGA", *NSREC 2006*, Ponte Vedra, FL.

Wang, J.J.; Charest, N.; Kuganesan, G.; Huang, C.K.; Yip, M.; Chen, H.S.; Borillo, J.; Samiee, S.; Dhaoui F.; Sun, J.; Rezgui, S.; McCollum, J. & Cronquist, B. (2006b). "Investigating and Modeling Total Ionizing Dose and Heavy Ion effects in Flash-Based Field Programmable Gate Array", *RADECS 2006*, Athens, Greece.

# Part II

# 7

# Evolving Systems and Adaptive Key Component Control

Susan A. Frost[1] and Mark J. Balas[2]
[1]*NASA Ames Research Center*
[2]*University of Wyoming*
*U.S.A.*

## 1. Introduction

We propose a new framework called Evolving Systems to describe the self-assembly, or autonomous assembly, of actively controlled dynamical subsystems into an Evolved System with a higher purpose. An introduction to Evolving Systems and exploration of the essential topics of the control and stability properties of Evolving Systems is provided. This chapter defines a framework for Evolving Systems, develops theory and control solutions for fundamental characteristics of Evolving Systems, and provides illustrative examples of Evolving Systems and their control with adaptive key component controllers.

Evolving Systems provide a framework that facilitates the design and analysis of self-assembling systems. The components of an Evolving System self-assemble, or mate, to form new components or the Evolved System. The mating of the subsystem components can be self-directed or agent controlled. The Evolving Systems framework provides a scalable, modular architecture to model and analyze the subsystem components, their connections to other components, and the Evolved System. Ultimately, once all the components of an Evolving System have joined together to form the fully Evolved System, it will have a new, higher purpose that could not have been achieved by the individual components collectively.

Autonomous assembly of large, complex structures in space, or on-orbit assembly, is an excellent application area for Evolving Systems. For example, the Solar Power Satellite (SPS) is a conceptual space structure that collects solar energy, which is then transmitted to Earth as microwaves (NASA, 1995). The solar array of the SPS, as envisioned in fig. 1, is a complex structure that could be assembled from many actively controlled components to form a new system with a higher purpose.

System stability is a trait that could be exhibited by an Evolving System or their components. We say that a subsystem trait is *inherited* by an Evolving System when the system retains the properties of the trait after assembly. The inheritance of subsystem traits, or genetics, such as controllability, observability, stability, and robustness, in Evolving Systems is an important research topic.

A critical element of successful on-orbit assembly of flexible space structures is the autonomous control of a structure during and after the connection of two or more subsystem components. The inheritance of stability in Evolving Systems is crucial in space applications due to potential damage and catastrophic losses that can result from unstable

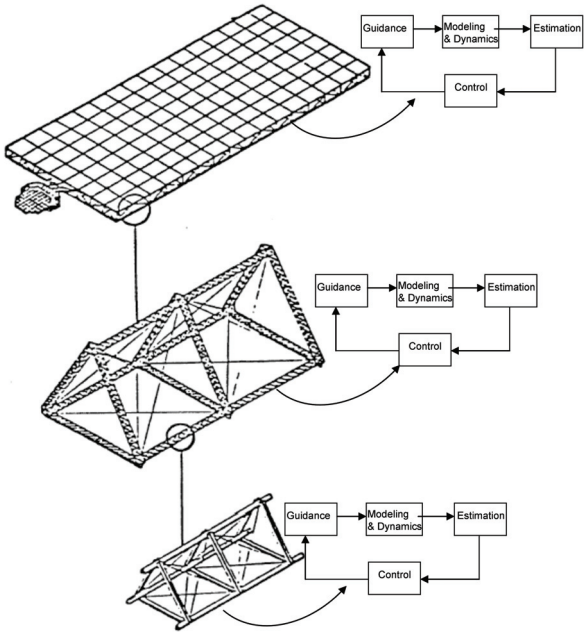



Fig. 1. Solar array component of a Solar Powered Satellite, image credit NASA.

space systems. The subsystem components of an Evolving System are designed to be stable as free-fliers, or unconnected components, but the Evolving System might fail to inherit stability at any step of the assembly, resulting in an unstable Evolved System. The fundamental topic of stability in Evolving Systems has been a primary focus of our Evolving Systems research (Balas & Frost, 2007; Frost & Balas, 2007a;b; 2008b;a; Balas & Frost, 2008; Frost, 2008). In this chapter, we develop an adaptive key component control method to ensure that stability is inherited in flexible structure Evolving Systems.

### 1.1 Description of evolving systems

Evolving Systems are dynamical systems that are self-assembled from actively controlled subsystem components. Central to the concept of Evolving Systems is the idea that an Evolved System has a higher functioning purpose than that of its subsystem components. For instance, the subsystem components might include a truss system, optical equipment, control systems, and communications equipment. If these components are assembled to form a space-based telescope, this would have a higher purpose than that of the individual components. Subsystems could consist of deployed components and self-assembled components. One could imagine that a space-based telescope, such as the Hubble Space Telescope, could be built as an Evolving System. The higher functioning purpose of the Evolving System would most likely come about not directly from the assembly of the subsystem components into a new system, but as a result of a new controller or agent taking over operation of the Evolving System after the subsystem components are fully assembled.
It is assumed that the components of an Evolving System would self-assemble, either through their own knowledge, or through the knowledge of an external agent. Note that the

agent would not be a human, but an autonomous agent with knowledge of the assembly requirements of the Evolving System. In the Evolving Systems framework presented here, it is assumed that the positioning of the subsystem components in space and time would be handled by the agent or the components themselves. Once the components are positioned, they would be self-directed or agent-directed to assemble with the appropriate components.
The actual connections made between subsystem components in an Evolving System are envisioned as compliant connections, so no degrees of freedom would be lost as a consequence of two components joining together in a rigid manner. A key concept in Evolving Systems is an evolutionary connection parameter, $\epsilon$, that enables the compliant connection to smoothly go from not existing at all ($\epsilon = 0$), to the full compliance of the connection ($\epsilon = 1$). The evolution of the connection parameter would occur independent of time. In Evolving Systems of flexible structures, the compliant connection might be modeled by a spring joining two components. Formation flying of imaging satellites to create synthetic apertures could be modeled as Evolving Systems with virtual forces representing the distance maintained between members of the satellite constellation.
In the formulation of Evolving Systems presented here, the evolution of the connection between components occurs independent of time. We are ignoring time in our formulation because it is assumed that the mating of the components is not time critical. We are interested in studying the joining of subsystem components to form an Evolved System, which is controlled by the evolution of the connection parameter going from zero to one. We say an Evolving System is *fully* evolved when all of the connection parameters joining the subsystem components equal one. An Evolving System is said to be *partially* evolved when at least one of its connection parameters never attains the value of 1 due to some event. In the case of a partially Evolved System, some of the components have failed to completely join together to form the prescribed configuration of the Evolving System.
Evolving Systems could be used for the design and analysis of self-assembling systems at all scales. Self-assembly occurs in nature and technology starting at the molecular or nanoscale (formation of crystals and nanostructures) to the macro-scale (formation of netted computer systems). See Whitesides & Grzybowski (2002) for an excellent survey of present and future applications of self-assembly.
The Evolving Systems framework is ideal for systems that are modular and can be scaled for complexity. If a system can be decomposed into modules, the detailed design process for each module needs to be performed only once. Parameter variations affecting the module can often be accommodated by the original design with significantly less effort than a new design would require. Once the design and validation of the module is complete, scaling the system to include more modules would be cost effective within the Evolving Systems framework.

### 1.2 Motivation for evolving systems

Future space missions will require on-orbit assembly of large aperture (greater than 10 meters) space systems, possibly at distant locations that prohibit astronaut intervention (Flinn, 2009). Historically, deployable techniques, sometimes in combination with astronaut assistance, have been used for fielding space systems. As the aperture size of the fielded space structure increases, deployable fielding techniques can become overly complex and unreliable. The increasing complexity of space structures, including such missions as the International Space Station (ISS) and the Hubble Space Telescope, often results in the need for extraordinary astronaut and ground crew assistance for assembly, servicing, and

upgrades. Evolving Systems research could facilitate self-assembly and autonomous servicing of complex space systems (Saleh et al., 2002). Additionally, future space missions might entail systems where the scale, complexity, and distance preclude astronaut assistance due to the inherent risks and costs associated with direct human involvement in these missions. These considerations suggest the need for an Evolving Systems framework and methodologies to enable on-orbit autonomous assembly and servicing of space systems with little or no direct human involvement.

Once an autonomous assembly problem has been solved with the Evolving Systems approach, the same solution can be used repeatedly or scaled to solve a similar problem. For example, the assembly of a large truss structure can be broken down into the assembly of smaller components. These components might consist of a small number of beams that are assembled into certain configurations. The designer only needs to develop the methods to assemble a certain type of component once, then this solution can be repeated to create any number of similar components. One can envision the development of a repository of designs that could be reused in different platforms with small modifications or parameterization changes for new dimensions, configurations, or other characteristics of the components. Evolving Systems enables the scaling and reuse of subsystem components, allowing new platforms to leverage existing technologies or reuse demonstrated solutions.

Flexible structure Evolving Systems are actively controlled, self-assembling flexible structures. The autonomous assembly of space structures provides an efficient means to build very large space structures with the elimination of space walk missions. Removing the astronaut from the assembly of space structures removes the dependency on transportation of the astronaut to the structure, eliminates the risk to human life, and eliminates the high costs associated with transporting humans to space. On-orbit assembly also gives the capability to build and service space systems at distant locations in space that are inaccessible to humans. A key benefit of Evolving Systems is its ability to enable on-orbit servicing and upgrades to existing space systems, thereby leveraging our existing space assets to their fullest capability.

The Evolving Systems framework is ideal for exploiting the inherent modularity and scalability of flexible structure space systems to potentially deliver more reliable systems at lower costs. Space systems that are self-assembled from components can lead to greater launch packing efficiency than can be achieved in traditionally deployed systems. The component aspect of Evolving Systems aids in the mitigation of vibration damage associated with the launch environment by allowing subsystem components to be individually enclosed in energy absorbing packaging. The modular framework of Evolving Systems allows designers to easily add redundancy to systems, thereby mitigating risks. Evolving Systems has the potential to solve difficult autonomous assembly and on-orbit servicing missions of flexible structure space systems, hence, the framework and the control problems investigated here are tailored to the application of flexible structure Evolving Systems.

### 1.3 Previous research

Decentralized control theory and analysis has been applied to the control of large interconnected systems; see the excellent survey paper by Nils Sandell (Sandell, Jr. et al., 1978) on this topic. Generally, decentralized control has been used to decrease the complexity of the control issues affecting large interconnected systems. Several researchers have proposed methods to decompose large interconnected systems into subsystems which can then be analyzed for stability properties and for the use of decentralized control

methodologies (Michel, 1983; Willems, 1986; Corfmat & Morse, 1976b). These ideas are related, but not equivalent to the Evolving Systems viewpoint.
Formations or constellations of satellites, nano-satellites, or micro-spacecraft could be included in the Evolving Systems framework. These formations of multiple, low cost spacecraft enable missions to accomplish complex objectives with the benefit of greater redundancy, improved performance, and reduced cost. An especially challenging control problem for constellations having large numbers of satellites is the task of coordinating and controlling the relative distances and phases between members of the fleet (Mueller et al., 2001; Kapilal, 1999). The solutions proposed in this work are specific to the application of constellations of satellites, and so are not as general as the Evolving Systems framework we are presenting here.
On the experimental side, a research group at the Information Sciences Institute at the University of Southern California (USC) has been conducting research in self-reconfigurable, autonomous robots and systems. They have conducted experimental work to study the feasibility of techniques for assembling large space structures as part of their FIMER (Free-flying Intelligent MatchmakER robots) project (Suri et al., 2006; Shen et al., 2003). This group uses a distributed control method with simple proportional derivative control laws for the selfassembly of components.

## 2. Theoretical formulation of evolving systems

This section provides the general theoretical formulation of Evolving Systems, expanding on work first presented in (Balas et al., 2006). In the previous section, we introduced the reader to the variety of dynamical systems that can be modeled by Evolving Systems and some of the benefits applications can obtain by using the Evolving Systems approach. Flexible structures are relatively simple, generally well understood mechanical dynamic systems, so they will be used to illustrate many ideas presented here. The state space representation developed in this section will be for general linear time-invariant (LTI) Evolving Systems, although the framework can be easily extended to account for nonlinear time-invariant and time varying Evolving Systems.

### 2.1 General formulation of evolving systems

In this section we give the general mathematical formulation of Evolving Systems. Consider a system of $L$ individually actively controlled components, where the components are given by

$$\begin{cases} \dot{\mathbf{x}}_i &= f_i(\mathbf{x}_i, \mathbf{u}_i); \quad \mathbf{x}_i(0) = \mathbf{x}_{i0} \\ \mathbf{y}_i &= g_i(\mathbf{x}_i, \mathbf{u}_i) \end{cases} \tag{1}$$

where $i = 1,2,\ldots,L$, $\mathbf{x}_i \equiv \left[x_1^i x_2^i \cdots x_{n_i}^i\right]^T$ is the component state vector, $\dot{\mathbf{x}}_i \equiv \left[\dot{x}_1^i \dot{x}_2^i \cdots \dot{x}_{n_i}^i\right]^T$, $\mathbf{u}_i \equiv \left[u_1^i u_2^i \cdots u_{m_i}^i\right]^T$ is the control input vector, $\mathbf{y}_i \equiv \left[y_1^i y_2^i \cdots y_{p_i}^i\right]^T$ is the vector of sensed outputs, and $\mathbf{x}_{i_0}$ is the vector of initial conditions. Note that $n_i$ is the dimension of the state vector $\mathbf{x}_i$, $m_i$ is the dimension of the control vector $\mathbf{u}_i$, and $p_i$ is the dimension of the output vector $\mathbf{y}_i$. Each component has an objective to be satisfied by the perfomance cost function $J_i$. Local control that depends only on local state or local output information will be used to keep the components stable and to meet the component performance requirements, $J_i$. In general, the local controller for a Evolving System component would have the form given by

$$\begin{cases} \mathbf{u}_i &= h_i(\mathbf{z}_i) \\ \dot{\mathbf{z}}_i &= l_i(\mathbf{z}_i, \mathbf{y}_i) \end{cases} \tag{2}$$

where $h_i$ and $l_i$ are control operators and $\dot{\mathbf{z}}$ represents the dynamical part of the control law.
The components are the building blocks of the Evolving System. When these individual components join to form an Evolving System, the interconnections between components $I$ and $j$ are represented by the function $k_{ij}(\mathbf{x},\mathbf{u})$. The connection parameter, $\epsilon_{ij}$, multiplies the interconnections between components $i$ and $j$.
The subsystem components of the Evolving System with the interconnections included is given by

$$\begin{cases} \dot{\mathbf{x}}_i &= f_i(\mathbf{x}_i, \mathbf{u}_i) + \sum_{j=1}^{L} \epsilon_{ij} k_{ij}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y}_i &= g_i(\mathbf{x}_i, \mathbf{u}_i) \end{cases} \tag{3}$$

where $\mathbf{x} = [\mathbf{x}_1 \mathbf{x}_2 \ldots \mathbf{x}_L]^T$, $\mathbf{u} = [\mathbf{u}_1 \mathbf{u}_2 \ldots \mathbf{u}_L]^T$, and $0 \le \epsilon_{ij} \le 1$.
The connection parameter, $\epsilon_{ij}$, is a mathematical construct representing the evolutionary joining of components in an Evolving System. The connection parameter evolves continuously from zero to one as the components assemble. The connection parameter is zero when the components are *unconnected*, or free-fliers. In the free-flier configuration, the components are completely independent of each other. The concept of *partial* evolution versus *full* evolution is an important distinction in Evolving Systems. *Full* evolution of two components occurs when the evolution parameter controlling the connection of the components evolves completely, resulting in the connection reaching its full magnitude and the components being joined together. *Partial* evolution is the case where, for some reason, the connection parameter $\epsilon_{ij}$ joining two components fails to attain the value of 1, resulting in the failure of the two components to join together. An important characteristic of the Evolving Systems framework is that the *evolution process* of a system comprises the homotopies $0 \le \epsilon_{ij} \le 1$, not just the endpoints where $\epsilon_{ij} = 0$ or $\epsilon_{ij} = 1$. In Evolving Systems, the mating of components is independent of the evolution of time in the system. The time parameter and the connection parameter are uncoupled in Evolving Systems because the connection parameter completely defines the evolutionary joining of components.
When the subsystem components join to form an Evolved System, the new entity becomes

$$\begin{cases} \dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= g(\mathbf{x}, \mathbf{u}) \end{cases} \tag{4}$$

### 2.2 Finite element method formulation of evolving systems of linear flexible structures

A flexible structure Evolving System is a mechanical dynamical system consisting of actively controlled flexible structure components that are joined together by compliant forces, e.g., springs. A practical and well accepted representation of flexible structures is based on the finite element method (FEM), see Balas (1982); Meirovitch (2001). The fundamental law governing mechanical systems is Newton's second law, which we use to write the dynamical equations describing a flexible structure. The FEM of the lumped model in physical coordinates, $\mathbf{q}_i$, for an arbitrary actively controlled flexible structure component, $i$, with $n$ elements, $m$ control inputs, and $p$ outputs is given in matrix form as

$$\begin{cases} \mathbf{M}_i\ddot{\mathbf{q}}_i(t) + \mathbf{D}_i\dot{\mathbf{q}}_i(t) + \mathbf{K}_i\mathbf{q}_i(t) = \mathbf{B}_i\mathbf{u}_i(t); \mathbf{q}_i(0) = \mathbf{q}_{i0} \\ \mathbf{y}_i(t) = \mathbf{C}_i\mathbf{q}_i(t) + \mathbf{E}_i\dot{\mathbf{q}}_i(t) \end{cases} \tag{5}$$

Where $\mathbf{M}_i \equiv \text{diag}(m_1, m_2, \ldots, m_n)$ is the $n \times n$ element mass matrix, $\mathbf{q}_i(t) \equiv [q_1(t) q_2(t) \ldots q_n(t)]^T$ is the displacement vector, $\dot{\mathbf{q}}_i(t) \equiv \left[\dot{q}_1(t)\dot{q}_2(t)\cdots\dot{q}_n(t)\right]^T$ is the velocity vector, $\ddot{\mathbf{q}}_i(t) \equiv \left[\ddot{q}_1(t)\ddot{q}_2(t)\cdots\ddot{q}_n(t)\right]^T$ is the acceleration vector, $\mathbf{D}_i$ is the $n \times n$ element damping matrix, $\mathbf{K}_i$ is the $n \times n$ element stiffness matrix, $\mathbf{B}_i$ is the $n \times m$ matrix of control input constants, $\mathbf{u}_i(t) \equiv [u_1(t) u_2(t) \ldots u_m(t)]^T$ are the control inputs, $\mathbf{y}_i(t) \equiv [y_1(t) y_2(t) \ldots y_p(t)]^T$ is the vector of sensed outputs of the component, and $\mathbf{C}_i$ and $\mathbf{E}_i$ are the $p \times n$ matrices of output constants.

The damping in space structures in orbit above the atmosphere is expected to be quite small and can be well modeled by Rayleigh damping (Balas, 1982) as given by

$$\mathbf{D}_i = \alpha_1\mathbf{M}_i + \alpha_2\mathbf{K}_i \tag{6}$$

Because the damping is quite small, it is customary to use the undamped generalized eigenproblem for eq. 5 given by

$$\left(\mathbf{K}_i - \omega_k^2\mathbf{M}_i\right)\phi_k = 0 \tag{7}$$

where $k$ =1, 2, . . . ,$H$,$\mathbf{M}_i$ is symmetric, positive definite, $\mathbf{K}_i$ is symmetric, positive semidefinite, and $H$ is equal to the number of degrees of freedom (DOF) in the physical model. The mode shapes $\phi_k$ and the mode frequencies $\omega_k$ are calculated from the generalized eigenproblem. Modal coordinates, $\mathbf{z}$, are obtained from the transformation

$$\mathbf{q} = \Phi\mathbf{z} \tag{8}$$

where $\Phi = [\phi_1\ \phi_2 \ldots\ \phi_H]$. Generally, the number of modes computed for design and analysis is much smaller than the number of DOF included in the physical model (Bansenauer & Balas, January-February 1995).

The active control of each flexible structure component is local in the sense that the controller only uses the input and output ports located on its component. In the examples presented here, the active component control is in the form of Proportional Derivative (PD) control or Proportional Integral Derivative (PID) control.

The flexible structure components are the building blocks of the Evolving System. Any number of components can join together in an arbitrary, but predetermined, configuration to form an Evolved System. The components of an Evolving System are joined by connection forces operating on the displacements of physical coordinates within the components. The connection forces joining the components are modeled by linear springs connecting two elements, one from each component. Note that the connections could also be made through the velocities of the physical coordinates, with dampers connecting the components.

For the flexible structure Evolving System being described here, each connection force, or spring, joining physical coordinates from two components will be multiplied by a connection parameter. The symbol $\epsilon_{ij}$ will denote the *connection parameter* that multiplies the forces joining the $i^{th}$ and the $j^{th}$ components. For simplicity, the formulation of Evolving System presented here will only allow one connection parameter to multiply the forces joining two components. However, it would be possible to construct more complex flexible

structure Evolving Systems which have multiple, distinct connection parameters corresponding to the forces joining different physical coordinates of two subsystem components.

The connection forces between components of an Evolving System are represented in the connection matrix, $\mathbf{K}_{\mathrm{E}}(\epsilon_{ij})$, which multiplies the displacements of the component elements and has the form

$$\mathbf{K}_{\mathbf{E}}(\epsilon_{ij}) = \left[\mathbf{K}_{ij}(\epsilon_{ij})\right] \tag{9}$$

where $i$, $j$ = 1, 2, . . . $L$. The connection parameter, $\epsilon_{ij}$, multiplies the elements of the connection matrix corresponding to the connection forces joining physical coordinates of component $i$ to coordinates of component $j$, where $i \neq j$, i.e., components $i$ and $j$ are separate components. There is only one connection parameter connecting component $i$ to component $j$, so $\epsilon_{ij} \equiv \epsilon_{ji}$.

If there are no connections between any elements of components $i$ and $j$, then $\epsilon_{ij} \equiv 0$ and $\mathbf{K}_{ij}(\epsilon_{ij}) \equiv \mathbf{K}_{ji}(\epsilon_{ij}) \equiv 0$. The connection matrix has zero entries for the elements of components that have nothing connected to them. There are no cyclic connections within components represented in the connection matrix, so $\epsilon_{ii} \equiv 0$. Since $\mathbf{K}_{\mathrm{E}}(\epsilon_{ij})$ is a matrix of connection forces that are symmetric, $\mathbf{K}_{ji}(\epsilon_{ij}) \equiv \mathbf{K}_{ij}(\epsilon_{ij})^T$.

The off-diagonal elements of the connection matrix have the form

$$\mathbf{K}_{ij}(\epsilon_{ij}) = -\epsilon_{ij} \cdot \begin{bmatrix} k(i^1, j^1) & \cdots & k(i^1, j^{j_n}) \\ \vdots & \ddots & \vdots \\ k(i^{i_n}, j^1) & \cdots & k(i^{i_n}, j^{j_n}) \end{bmatrix} \tag{10}$$

where $i^n$ represents the $n^{th}$ element of the $i^{th}$ component, $j^m$ represents the $m^{th}$ element of the $j^{th}$ component, $k(i^n, j^m)$ is the connection force exerted by the $n^{th}$ element of component $i$ on the $m^{th}$ element of component $j$, and the values $i_n$ and $j_n$ represent the number of elements in the $i^{th}$ and the $j^{th}$ component FEM representations, respectively.

The block diagonal elements, $\mathbf{K}_{ii}(\epsilon_{ij})$, of the connection matrix are more complex, since they represent the connection forces of all of the components in the Evolving System which connect to component $i$. If more than one component connects to a given component $i$, then $\mathbf{K}_{ii}(\epsilon_{ij})$ will include multiple connection parameters. A general form of the block diagonal elements of the connection matrix is given by

$$\mathbf{K}_{ii}(\epsilon_{ij}) = \operatorname{diag}\left( \sum_{j=1}^{L} \epsilon_{ij} \sum_{m=1}^{j_n} k(i^1, j^m), \sum_{j=1}^{L} \epsilon_{ij} \sum_{m=1}^{j_n} k(i^2, j^m), \ldots, \sum_{j=1}^{L} \epsilon_{ij} \sum_{m=1}^{j_n} k(i^{i_n}, j^m) \right) \tag{11}$$

We can write an individual component of the flexible structure Evolving System as

$$\begin{cases} \mathbf{M}_i \ddot{\mathbf{q}}_i(t) + \mathbf{D}_i \dot{\mathbf{q}}_i(t) + \mathbf{K}_i \mathbf{q}_i(t) + \sum_{j=1}^{L} \mathbf{K}_{ij}(\epsilon_{ij}) \mathbf{q}_j(t) = \mathbf{B}_i \mathbf{u}_i(t);\ \mathbf{q}_i(0) = \mathbf{q}_{i0} \\ \mathbf{y}_i(t) = \mathbf{C}_i \mathbf{q}_i(t) + \mathbf{E}_i \dot{\mathbf{q}}_i(t) \end{cases} \tag{12}$$

The Evolving System consisting of $L$ interconnected components can now be written in matrix form as

$$\begin{cases} \mathbf{M_0}\ddot{\mathbf{q}}(t) + \mathbf{D_0}\dot{\mathbf{q}}(t) + \mathbf{K_0}\mathbf{q}(t) + \mathbf{K_E}(\epsilon_{ij})\mathbf{q}(t) = \mathbf{B_0}\mathbf{u}(t); \quad \mathbf{q}(0) = \mathbf{q}_0 \\ \mathbf{y}(t) = \mathbf{C_0}\mathbf{q}(t) + \mathbf{E_0}\dot{\mathbf{q}}(t) \end{cases} \tag{13}$$

where $\mathbf{M_0} \equiv \text{diag}(\mathbf{M_1},\mathbf{M_2}, \ldots ,\mathbf{M_L})$, $\mathbf{B_0} = \text{diag}(\mathbf{B_1},\mathbf{B_2}, \ldots ,\mathbf{B_L})$, $\mathbf{q}(t) \equiv [\mathbf{q_1}(t)\mathbf{q_2}(t) \ldots \mathbf{q_L}(t)]^T$, $\dot{\mathbf{q}}(t) \equiv [\dot{\mathbf{q}}_1(t)\ \dot{\mathbf{q}}_2(t) \ldots \dot{\mathbf{q}}_L(t)]^T$, $\ddot{\mathbf{q}}(t) \equiv [\ddot{\mathbf{q}}_1(t)\ \ddot{\mathbf{q}}_2(t) \ldots \ddot{\mathbf{q}}_L(t)]^T$, $\mathbf{u}(t) \equiv [\mathbf{u_1}(t)\mathbf{u_2}(t) \ldots \mathbf{u_L}(t)]^T$, $\mathbf{K_0} = \text{diag}(\mathbf{K_1},\mathbf{K_2}, \ldots ,\mathbf{K_L})$, $\mathbf{K_E}(\epsilon_{ij})$ is the connection matrix as described above, $\mathbf{y}(t) = [\mathbf{y_1}(t)\mathbf{y_2}(t) \ldots \mathbf{y_L}(t)]^T$, $\mathbf{C_0} \equiv \text{diag}(\mathbf{C_1},\mathbf{C_2}, \ldots ,\mathbf{C_L})$, and $\mathbf{E_0} \equiv \text{diag}(\mathbf{E_1},\mathbf{E_2}, \ldots ,\mathbf{E_L})$. For simplicity of notation, the time parameter, $t$, will be omitted henceforth.

The evolution of the different connection parameters, $\epsilon_{ij}$, that enable the mating of components is controlled by the self-assembling components or the agent. The various connection parameters can evolve at different times during the assembly of the Evolving System, thereby orchestrating the assembly order of the components. It is also possible to envision the evolution of the connection parameters proceeding at different rates, effectively causing the assembly of some components to be accomplished more rapidly than others. However, we will not address the evolution rate of the connection parameter at this time. The Evolving Systems framework allows the evolution of two components to be undone, i.e., connected components can *devolve*, where their connection parameter goes from one down to zero.

The flexible structure Evolving Systems described above are some of the simplest examples of dynamic behavior where the nature of Evolving Systems can be investigated. The approaches developed here can be generalized rather easily and scaled for much more complex Evolving Systems.

The matrix notation which we have developed for flexible structure Evolving Systems is useful for understanding the formulation of Evolving Systems, however, a model based on a state space representation of Evolving Systems will prove essential as we analyze more complex Evolving Systems.

### 2.3 State space description of flexible structure evolving systems

A state space representation of linear time-invariant Evolving Systems is developed here. Suppose we have a flexible structure Evolving System consisting of $L$ individual components as described in section 2.2 and given by the FEM

$$\begin{cases} \mathbf{M}_i\ddot{\mathbf{q}}_i + \mathbf{D}_i\dot{\mathbf{q}} + \mathbf{K}_i\mathbf{q}_i = \mathbf{B}_i^0\mathbf{u}_i; \quad \mathbf{q}_i(0) = \mathbf{q}_{i0} \\ \mathbf{y}_i = \mathbf{C}_i^0\mathbf{q}_i + \mathbf{E}_i\dot{\mathbf{q}}_i \end{cases} \tag{14}$$

We can represent the individual flexible structure components given by eq. 14 by the state-space description

$$\begin{cases} \dot{\mathbf{x}}_i &= \mathbf{A}_i\mathbf{x}_i + \mathbf{B}_i\mathbf{u}_i;\ \mathbf{x}_i(0) \equiv \mathbf{x}_{i_0} \\ \mathbf{y}_i &= \mathbf{C}_i\mathbf{x}_i \end{cases} \tag{15}$$

where $i = 1,2,\ldots,L$, $\mathbf{x}_i \equiv \left[x_1^i x_2^i \cdots x_{n_i}^i\right]^T$ is the component state vector, $\dot{\mathbf{x}}_i \equiv \left[\dot{x}_1^i \dot{x}_2^i \cdots \dot{x}_{n_i}^i\right]^T$, $\mathbf{u}_i \equiv \left[u_1^i u_2^i \cdots u_{m_i}^i\right]^T$ is the control input vector, $\mathbf{y}_i \equiv \left[y_1^i y_2^i \cdots y_{p_i}^i\right]^T$ is the vector of sensed outputs, $\mathbf{x}_{i_0}$ is the vector of initial conditions, and $\mathbf{A}_i$, $\mathbf{B}_i$, and $\mathbf{C}_i$ are constant matrices of dimension $n_i \times n_i$, $n_i \times m_i$, and $p_i \times n_i$, respectively. Since the state space description comes from the dynamical equations given by eq. 14, we have that $\mathbf{A}_i = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}_i^{-1}\mathbf{K}_i & -\mathbf{M}_i^{-1}\mathbf{D}_i \end{bmatrix}$, $\mathbf{B}_i = \begin{bmatrix} 0 \\ \mathbf{M}_i^{-1}\mathbf{B}_i^0 \end{bmatrix}$ and

$\mathbf{C}_i = \begin{bmatrix}\mathbf{C}_i^0 & \mathbf{E}_i^0\end{bmatrix}$. Note that $n_i$ is the dimension of the state vector $\mathbf{x}_i$, $m_i$ is the dimension of the control vector $\mathbf{u}_i$, and $p_i$ is the dimension of the output vector $\mathbf{y}_i$. The local controller on component $i$ is given by

$$\mathbf{u}_i = l_i \mathbf{y}_i \tag{16}$$

where $l_i$ is a linear control operator.

The subsystem components are the building blocks of the Evolving System. The connection forces between two components, $i$ and $j$, of an Evolving System are represented in the connection matrix, $\mathbf{A}_{ij}$, which is multiplied by the connection parameter, $\epsilon_{ij}$, where $i, j = 1, 2, \ldots, L$, $i \neq j$, and $0 \leq \epsilon_{ij} \leq 1$. Even though connections may exist between the states of different components of the Evolving System, the component inputs and outputs are still local, i.e., there is no sharing of component inputs or outputs between components.

The FEM representation of a flexible structure Evolving System component is given by

$$\begin{cases} \mathbf{M}_i\ddot{\mathbf{q}}_i + \mathbf{D}_i\dot{\mathbf{q}}_i + \mathbf{K}_i\mathbf{q} + \sum_{j=1}^{L} \mathbf{K}_{ij}(\epsilon_{ij})\mathbf{q}_j = \mathbf{B}_i^0\mathbf{u}_i;\ \mathbf{q}_i(0) = \mathbf{q}_{i0} \\ \mathbf{y}_i = \mathbf{C}_i^0\mathbf{q}_i + \mathbf{E}_i\dot{\mathbf{q}}_i \end{cases} \tag{17}$$

The state space equations for an individual component including connections to other components in an Evolving System are given by

$$\begin{cases} \dot{\mathbf{x}}_i &= \mathbf{A}_i\mathbf{x}_i + \mathbf{B}_i\mathbf{u}_i + \sum_{j=1}^{L} \epsilon_{ij}\mathbf{A}_{ij}\mathbf{x};\ \mathbf{x}_i(0) \equiv \mathbf{x}_{i_0} \\ \mathbf{y}_i &= \mathbf{C}_i\mathbf{x}_i \end{cases} \tag{18}$$

where $\mathbf{x} \equiv [\mathbf{x_1 x_2} \ldots \mathbf{x_L}]^T$ is the concatenated state vectors of the entire Evolving System, $\mathbf{A}_{ij}$ is the connection matrix, and $0 \leq \epsilon_{ij} \leq 1$ is the connection parameter. The connection matrix, $\mathbf{A}_{ij}$, has dimension $n_i$ by dim($\mathbf{x}$), where $n_i$ is the dimension of the state vector $\mathbf{x}_i$ corresponding to component $i$ and $\dim(\mathbf{x}) \equiv \sum_{k=1}^{L} n_k$. In eq. 17, the matrix $\mathbf{K}_{ij}(\epsilon_{ij})$ multiplies the vector $\mathbf{q}_j$. The elements of the matrix $\mathbf{A}_{ij}$ are related to the elements of $\mathbf{K}_{ij}(\epsilon_{ij})$, except that they are mass normalized by $\mathbf{M}_i^{-1}$and rearranged so that they multiply the elements of $\mathbf{x}$ corresponding to $\mathbf{q}_j$. The other elements of $\mathbf{A}_{ij}$ are set to zero.

The connection parameter, $\epsilon_{ij}$, multiplies the forces connecting the physical coordinates of component $i$ to physical coordinates of component $j$. The connection parameter is the same in the state space representation as in the FEM flexible structure model. If there is no connection between any states of components $i$ and $j$, then $\epsilon_{ij} \equiv 0$ and the connection matrix $\mathbf{A}_{ij} \equiv 0$.

When a system of $L$ individual components, as described by eq. 18, mate to form an Evolving System, the new entity becomes

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{A}(\epsilon_{ij})\mathbf{x} + \mathbf{B}\mathbf{u};\ \mathbf{x}(0) \equiv \mathbf{x_0} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \end{cases} \tag{19}$$

where $\mathbf{x} \equiv [\mathbf{x_1 x_2} \ldots \mathbf{x_L}]^T$, $\dot{\mathbf{x}} \equiv [\dot{\mathbf{x}}_1 \dot{\mathbf{x}}_2 \ldots \dot{\mathbf{x}}_L]^T$, $\mathbf{u} \equiv [\mathbf{u_1 u_2} \ldots \mathbf{u_L}]^T$, $\mathbf{y} \equiv [\mathbf{y_1 y_2} \ldots \mathbf{y_L}]^T$, $\mathbf{B} \equiv \mathrm{diag}(\mathbf{B_1}, \mathbf{B_2}, \ldots, \mathbf{B_L})$, $\mathbf{C} \equiv \mathrm{diag}(\mathbf{C_1}, \mathbf{C_2}, \ldots, \mathbf{C_L})$, $\mathbf{A}(\epsilon_{ij}) \equiv \mathrm{diag}(\mathbf{A}_i) + \left[ \sum_{j=1}^{L} \epsilon_{1j} \mathbf{A}_{1j} \sum_{j=1}^{L} \epsilon_{2j} \mathbf{A}_{2j} \cdots \sum_{j=1}^{L} \epsilon_{Lj} \mathbf{A}_{Lj} \right]^T$ and $0 \le \epsilon_{ij} \le 1$. The system given by eq. 19 will also be represented by the standard state space notation of $(A,B,C)$.

We can form the closed-loop Evolving System by taking the Evolving System given by eq. 19 and connecting each of the local component controllers, $\mathbf{u}_i$, to their corresponding input and output ports, i.e., close the loops in each of the components. The closed-loop Evolving System can be written as

$$\dot{\mathbf{z}} = \bar{\mathbf{A}}(\epsilon_{ij})\mathbf{z}; \ \mathbf{z}(0) \equiv \left[\mathbf{x_0 0}\right]^T \tag{20}$$

where $\mathbf{z} = [\mathbf{x} \ \eta]^T$ is the augmented state vector, $\bar{\mathbf{A}}(\epsilon_{ij})$ is the closed-loop system, and $\epsilon_{ij}$ is the connection parameter. The closed-loop Evolving System given by eq. 20 will be used for stability analysis.

Flexible structure Evolving Systems can be written in a form that is mass normalized and the component state vectors can be rearranged to appear as one flexible structure, instead of multiple component state vectors concatenated together. The state space description of flexible structure Evolving Systems can be easily extended to describe other applications of Evolving Systems.

### 2.4 Impedance-admittance formulation of contact dynamics in evolving systems

In this section, we formulate the contact dynamics in Evolving Systems in terms of mechanical impedance and admittance, as first described in (Frost & Balas, 2007a). For many dynamical systems, the impedance-admittance form is a useful tool for modeling the contact dynamics of components, see Harris & Crede (1976).

**Definition 2.1** *The impedance of a mechanical system is determined by the equation $f = Z(v)$, where $f$ is the force exerted by the system, $v$ is the velocity of the system, and $Z$ is the impedance of the system.*

**Definition 2.2** *The admittance of a mechanical system, $Y$, is the inverse of the impedance of the system, e.g., $Y \equiv Z^{-1}$ and $v = Y(f)$.*

Impedance and admittance can be seen as nonlinear operators describing the relationship between the output of a mechanical system, or the force it exerts at a contact point, with the input of the system, or the velocity at the contact point. When two components join at a point of contact, their velocities are equal and the forces exerted are equal and opposite. If the contact points of the two components are represented as ( $f_1$,$v_1$) and ( $f_2$,$v_2$) with displacements $q_1$ and $q_2$, then we can write

$$\begin{cases} f_1 = -f_2 \\ v_1 \equiv \dot{q}_1 = v_2 \equiv \dot{q}_2 \end{cases} \tag{21}$$

This formulation can also be seen as the feedback connection of two components in an Evolving System, where the admittance of component 1 is connected in feedback with the impedance of component 2, as shown in fig. 2. We introduce two nonlinear operators $Y_1$ and $Z_2$ that provide the admittance and impedance formulation of the contact dynamics of nonlinear Evolving Systems components. These operators relate the force and velocity at the contact point of two mating components as given by the equations $\mathbf{v}_1 = Y_1(\mathbf{f}_1)$ and $\mathbf{f}_2 = Z_2(\mathbf{v}_2)$.

In linear time-invariant systems, these operators can be easily calculated using Laplace transforms. For nonlinear components, the admittance and impedance operators cannot be easily found. However, this does not invalidate the analysis provided in this chapter, which will provide a foundation for adaptive key component control.

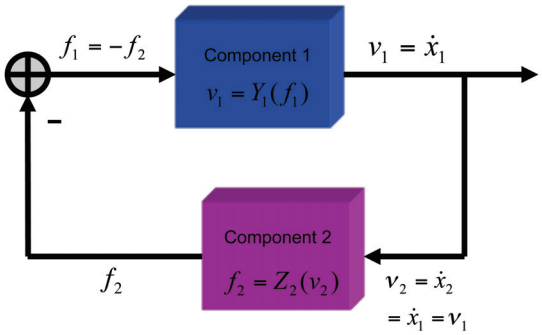



Fig. 2. Admittance-impedance feedback connection of two components.

## 3. Stability inheritance in evolving systems

The application of Evolving Systems to self-assembly of structures in space imposes the need for the inheritance of stability. Many textbooks (Vidyasagar, 1993; Brogan, 1991; Ogata, 2002; Slotine & Li, 1991) give excellent discussions of linear and nonlinear systems stability analysis. For linear time-invariant Evolving Systems, we will examine the closed-loop poles of the system to evaluate the stability of the system as it evolves. In particular, we will examine the eigenvalues of the matrix $\bar{\mathbf{A}}(\epsilon_{ij})$ from the state space equation of the closed-loop Evolving System given by eq. 20 as $\epsilon_{ij}$ goes from 0→1. The system is unstable if any of the closed-loop poles cross the $j\omega$-axis.

Example 1 is a two component Evolving System where each of the components is actively controlled and stable, but the Evolving System fails to inherit the stability traits of the components. This particular system becomes unstable during the evolution process and remains unstable when the system is fully evolved. Consider the fully actuated, fully sensed three mass Evolving System shown in fig. 3. Component 1 contains only one mass with local control. The dynamical equations for component 1 are

$$\begin{cases} m_1\ddot{q}_1 &= u_1 \\ \mathbf{y_1} &= \begin{bmatrix} q_1 & \dot{q}_1 \end{bmatrix}^T \end{cases} \tag{22}$$

where $m_1 = 30$ is the mass of mass 1, $q_1$ is the displacement of mass 1, and $u_1 = -(0.9s + 0.1)q_1$ is the local controller for component 1 with the Laplace variable $s$. The dynamical equations for component 2 are

$$\begin{cases} m_2\ddot{q}_2 &= u_2 - k_{23}(q_2 - q_3) \\ m_3\ddot{q}_3 &= u_2 - k_{23}(q_3 - q_2) \\ \mathbf{y_2} &= \begin{bmatrix} q_2 & \dot{q}_2 \end{bmatrix}^T \\ \mathbf{y_3} &= \begin{bmatrix} q_3 & \dot{q}_3 \end{bmatrix}^T \end{cases} \tag{23}$$

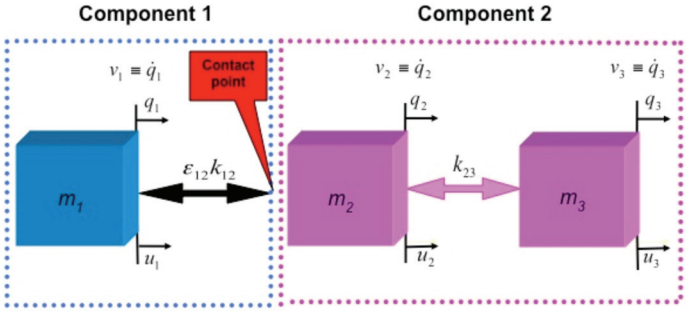



Fig. 3. Ex. 1: A two component flexible structure Evolving System.

where $m_2 = 1.0$ is the mass of mass 2, $m_3 = 1.0$ is the mass of mass 3, $q_2$ is the displacement of mass 2, $q_3$ is the displacement of mass 3, and $k_{23} = 1.0$. The controllers on component 2 are

$$\begin{cases} u_2 &= -\left(\dfrac{0.1}{s} + 0.2s + 0.5\right) q_2 \\ u_3 &= -(0.6s + 1)\, q_3 \end{cases} \tag{24}$$

The controllers for components 1 and 2 have been designed to produce stable behavior when the components are unconnected. The two components are joined by a spring connecting mass 1 with mass 2. The Evolving System comprised of these two components can be written in the matrix form of eq. 13 as

$$\begin{cases} \mathbf{M_0}\ddot{\mathbf{q}} &= \mathbf{B_0 u} - \mathbf{K_0 q} - \mathbf{K_E}(\epsilon_{ij})\mathbf{q} \\ \mathbf{y} &= \mathbf{C_0}\begin{bmatrix}\mathbf{q} & \dot{\mathbf{q}}\end{bmatrix}^T \end{cases} \tag{25}$$

where $\mathbf{M_0} = \text{diag}(m_1,m_2,m_3)$, $\ddot{\mathbf{q}} = [\ddot{q}_1\ddot{q}_2\ddot{q}_3]^T$, $\mathbf{q} = [q_1q_2q_3]^T$ $\mathbf{u} = [u_1u_2u_3]^T$, $\mathbf{y} = [y_1y_2y_3]^T$, $\mathbf{B_0} = \mathbf{I_3}$, $\mathbf{C_0} = \mathbf{I_6}$, $\mathbf{K_0} = \begin{bmatrix} \epsilon_{12}k_{12} & -\epsilon_{12}k_{12} & 0 \\ -\epsilon_{12}k_{12} & \epsilon_{12}k_{12} & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{K_E}(\epsilon_{ij}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & k_{23} & -k_{23} \\ 0 & -k_{23} & k_{23} \end{bmatrix}$, $k_{12} = 1.0$, and $0 \le \epsilon_{12} \le 1$.

Matlab and Simulink models of this system were created. To determine the stability of the Evolving System, we connect the local component controllers to their inputs and outputs, and we examine the closed-loop poles, or the eigenvalues, of the resulting composite system. Figure 4 shows the closed-loop poles of the Evolving System given by equation 25 as the system evolves, i.e., as $\epsilon_{12}$ goes from 0 to 1. Note that two of the closed-loop poles of ex. 1 cross the $j\omega$-axis for some $\epsilon_{12} > 0$, demonstrating that the Evolving System loses stability during evolution. When the system is fully evolved, i.e., $\epsilon_{12} = 1$, the Evolved System is unstable, i.e., it fails to inherit the stability of its components, as seen in fig. 4.

In the next section, we explore a method to restore stability to Evolving Systems that would otherwise fail to inherit the stability traits of their components.

## 4. Key component controllers

In this section we introduce the idea of controllers that stabilize flexible structure Evolving Systems during evolution. Often the design requirements for an Evolving System dictate that the individual components remain unchanged as much as possible. For situations

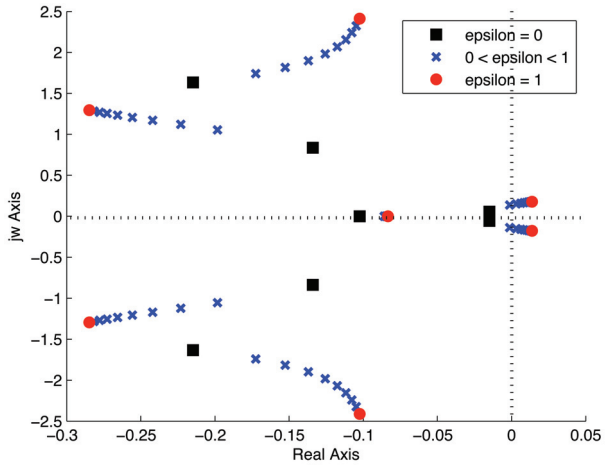



Fig. 4. Closed-loop poles of Ex. 1.

where stability is not inherited during evolution, in many cases it would be advantageous to augment the controller on only one component to restore stability to the entire Evolving System, thereby leaving the other components and their controllers unmodified. Furthermore, it is desirable for the augmented controller to only use the input-output ports on the component on which it resides. In this section, we introduce the idea of key component controllers that restore stability to an Evolving System by augmenting the controller on a single subsystem component, using only the input-output ports on that component. The key component controller using fixed gains was first proposed in (Frost & Balas, 2007b).

In the key component controller design approach, one key component is chosen from the Evolving System to have additional local control added to it with the objective of maintaining system stability during the entire evolution of the system. The control and sensing of the other subsystem components will be unaltered and remain local. The key component controller operates solely through a single set of input and output ports on the key component, see fig. 5. For components that lose stability when assembling, the individual components could mate with the key component one at a time. The key component would compensate for any component which caused instability, thereby restoring stability to the system.

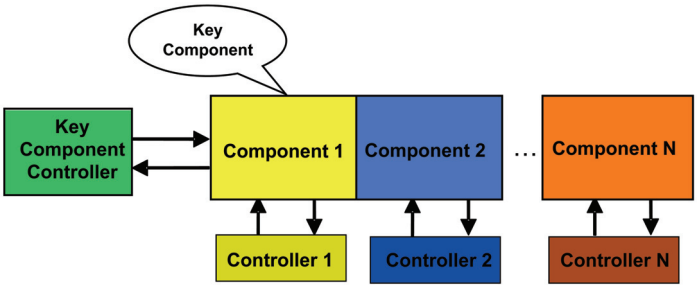



Fig. 5. Block diagram of key component controller and Evolving System.

A clear advantage of the key component design approach is that components can be reused in many different configurations of Evolving Systems without needing redesign from a stability point of view. Redesign of existing components is unnecessary because the key component will be responsible for maintaining overall system stability. The reuse of components that are space-qualified, or at least previously designed, built, and verified and validated, could reduce overall system development and validation time and could result in higher quality systems with potentially significant cost savings and risk mitigation.
The key component controller design requires the controllability and observability of the states of the Evolving System from a set of input and output ports on the key component. In the case of LTI Evolving Systems, we can use a method of applying local output feedback through specified input ports to obtain controllability and observability from a single set of input-output ports. Details of the method are given in (Corfmat & Morse, 1976a). Applying local output feedback on a component is seen as a minor modification that still preserves the idea of leaving the nonkey components mostly unmodified.

### 4.1 Adaptive key component controllers for restoring stability in evolving systems

We present a key component controller that uses a direct adaptive control law to restore stability to an Evolving System. In many aerospace environments and applications, the parameters of a system are poorly known and difficult and costly to obtain. Control laws that use direct adaptation are a good design choice for systems where access to precisely known parametric values is limited, since these control laws adapt their gains to the system output. We propose the use of adaptive control laws in a key component controller to provide a practical solution to the problems described above. This approach was first proposed in (Frost & Balas, 2008b).
The adaptive key component controller adapts its gains based on the system outputs to ensure that the Evolving System remains stable during component assembly. The adaptive key component design has the same advantages as the fixed gain key component controller without the need to schedule the gains based on the value of $\epsilon$.
We consider an Evolving System consisting of two components given by:

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{A}(\epsilon)\mathbf{x} + \mathbf{B}\mathbf{u};\ \mathbf{x}(0) \equiv \mathbf{x_0} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \end{cases} \tag{26}$$

where $\mathbf{x} \equiv [\mathbf{x_1}\mathbf{x_2}]^T$, $\dot{\mathbf{x}}\_ [\dot{\mathbf{x}}_1\dot{\mathbf{x}}_2]^T$, $\mathbf{u} \equiv [\mathbf{u_1}\mathbf{u_2}]^T$, $\mathbf{y} \equiv [\mathbf{y_1}\mathbf{y_2}]^T$, $\mathbf{B} \equiv \text{diag}(\mathbf{B_1},\mathbf{B_2})$, $\mathbf{C} \equiv \text{diag}(\mathbf{C_1},\mathbf{C_2})$, $\mathbf{A}(\epsilon) \equiv \text{diag}(\mathbf{A_1},\mathbf{A_2}) + \left[\sum_{j=1}^{2} \epsilon\mathbf{A}_{1j} \sum_{j=1}^{2} \epsilon\mathbf{A}_{2j}\right]^T$ and $0 \le \epsilon \le 1$.
Now we give the equations for an Evolving System with a key component controller. Without loss of generality, we can let component 1 be the key component since the system can be rewritten to switch component 1 with component 2. Also, we may think of component 2 as being the rest of the Evolving System to which the key component and its adaptive controller will be connected. The adaptive key component controller on component 1 is given by

$$\begin{cases} \mathbf{u}_1^A &= \mathbf{G}\mathbf{y}_1^A \\ \dot{\mathbf{G}} &= -\mathbf{y}_1^A(\mathbf{y}_1^A)^T\mathbf{H};\ \mathbf{H} > 0 \end{cases} \tag{27}$$

The adaptive key component controller only uses the input and output ports located on component 1. Component 1, which is the key component of the system, can be written as

$$\begin{cases} \dot{\mathbf{x}}_1 &= \mathbf{A}_1\mathbf{x}_1 + \mathbf{B}_1\mathbf{u}_1 + \mathbf{B}_1^{\mathbf{A}}\mathbf{u}_1^{\mathbf{A}} + \sum_{j=1}^{2}\epsilon\mathbf{A}_{1j}(\mathbf{x}); \\ \mathbf{y}_1 &= \mathbf{C}_1\mathbf{x}_1 \\ \mathbf{y}_1^{\mathbf{A}} &= \mathbf{C}_1^{\mathbf{A}}\mathbf{x}_1 \end{cases} \tag{28}$$

and component 2 can be written as

$$\begin{cases} \dot{\mathbf{x}}_2 &= \mathbf{A}_2\mathbf{x}_2 + \mathbf{B}_2\mathbf{u}_2 + \mathbf{B}_2^{\mathbf{A}}\mathbf{u}_2^{\mathbf{A}} + \sum_{j=1}^{2}\epsilon\mathbf{A}_{2j}(\mathbf{x}) \\ \mathbf{y}_2 &= \mathbf{C}_2\mathbf{x}_2 \\ \mathbf{y}_2^{\mathbf{A}} &= \mathbf{C}_2^{\mathbf{A}}\mathbf{x}_2 \end{cases} \tag{29}$$

where the augmented control $\mathbf{u}_2^{\mathbf{A}}$ would only be present if additional output feedback control were needed to satisfy sufficient condition for the adaptive controller. Next, we give some useful definitions.

**Definition 4.1** *Consider a linear system* (**A**,**B**,**C**) *with closed-loop transfer function,* $T_c(s) \equiv \mathbf{C}(s\mathbf{I} - \mathbf{A})\mathbf{B}$. *We say the system* (**A**,**B**,**C**) *is strict positive real (SPR) when for all* $\omega$ *real and for some* $\sigma > 0$

$$Re[T_c(-\sigma + j\omega)] \geq 0 \tag{30}$$

**Definition 4.2** *We say a linear system* (**A**,**B**,**C**) *is almost strict positive real (ASPR) when it can be made strict positive real by adding output feedback.*

**Remark:** A linear system (**A**,**B**,**C**) is ASPR if it has no nonminimum phase zeros and $\mathbf{CB} > \mathbf{0}$.

The following result from (Fuentes & Balas, 2000) gives the sufficient condition for a linear time-invariant system with an adaptive control law as described above, to be guaranteed to have bounded gains and asymptotic output tracking.

**Theorem 4.3** *Assume the linear time-invariant system given by*

$$\begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t);\ \mathbf{x}(\mathbf{0}) = \mathbf{x_0} \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{cases} \tag{31}$$

*is ASPR. Then the direct adaptive control law*

$$\begin{cases} \mathbf{u}(t) &= \mathbf{G}\mathbf{y}(t); \\ \dot{\mathbf{G}} &= -\mathbf{y}(t)\mathbf{y}(t)^T\mathbf{H};\ \mathbf{H} > 0 \end{cases} \tag{32}$$

*produces bounded adaptive gains,* **G**, *and* $\mathbf{y} \to 0$ *as* $t \to \infty$

This result suggests that the sufficient condition for an Evolving System with an adaptive key component controller to have guaranteed bounded gains and asymptotic tracking is that the system be ASPR. This idea will be developed further in a subsequent section. Note that the theory developed in (Fuentes & Balas, 2000) could also be applied to design the key component adaptive controller to track a desired reference model and reject disturbances.

### 4.2 Results of restoring stability to Ex. 1 with adaptive key component controllers

A Simulink model was created to implement the adaptive key component controller for ex. 1. Simulations were run with the connection parameter, $\epsilon_{12}$, ranging from 0 to 1, allowing

the system to go from unconnected components to a fully Evolved System. The key component controller was able to maintain system stability during the entire evolution process when it used the input-output ports on mass 1 of component 1, see fig. 6. When component 1 was the key component, the Evolving System is ASPR.

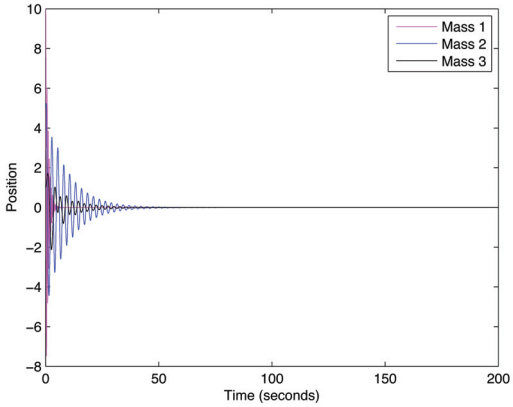



Fig. 6. Nondimensional position displacements of Ex. 1 with adaptive key component controller on component 1.

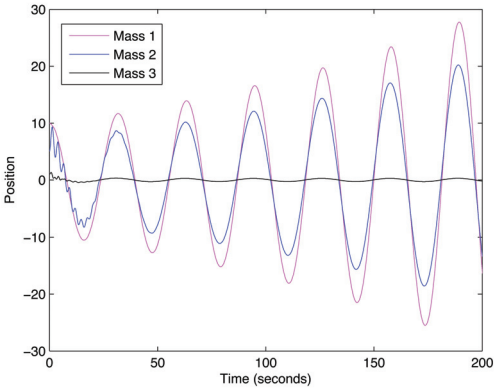



Fig. 7. Nondimensional position displacements of Ex. 1 with adaptive key component controller on mass 3 of component 2.

When the key component controller was located on component 2 and used the inputoutput ports on mass 3, stability was not maintained, see fig. 7. The adaptive key component controller was not able to restore stability on mass 3 because that system had nonminimum phase zeros at $0.00515 \pm 0.2009i$, i.e., the system was not ASPR.

## 5. Inheritance of passivity properties in evolving systems

In this section we explore the inheritance of different types of passivity in Evolving Systems. First we give some theorems on the inheritance of these traits in systems connected in

feedback. Then we use the admittance-impedance formulation of Evolving Systems developed in Section 2.4 to determine the condition under which passivity traits are inherited in Evolving Systems. We use these results to determine the sufficient condition for LTI Evolving Systems with an adaptive key component controller to be guaranteed to have bounded gains and asymptotic state tracking.

Intuitively, a system is passive if the energy stored by the system is less than or equal to the energy supplied. Physical systems satisfy energy conservation equations of the form

$$\frac{d}{dt}[\text{Stored Energy}] = [\text{External Power Input}] + [\text{Internal Power Generation}] \tag{33}$$

**Definition 5.1** *We say that a nonlinear system of the form*

$$\begin{cases} \dot{\mathbf{x}} &= f(\mathbf{x},\mathbf{u});\ \mathbf{x}(0) = \mathbf{x_0} \\ \mathbf{y} &= \mathbf{h}(\mathbf{x},\mathbf{u}) \end{cases} \tag{34}$$

*is passive if it has an positive definite energy storage function, $V(\mathbf{x})$, that satisfies*

$$\dot{V}(\mathbf{x}) = \langle \mathbf{y},\mathbf{u} \rangle - S(\mathbf{x}) = \mathbf{y}^T\mathbf{u} - S(\mathbf{x}) \tag{35}$$

*where $S(\mathbf{x})$ is a positive semi-definite function, i.e., $S(\mathbf{x}) \geq 0$.*

The term $\dot{V}(\mathbf{x})$ in eq. 35 represents the energy storage rate of the system. The external power input term in eq. 33 is represented by the inner product of the input and the output of the system, i.e., $\mathbf{y}^T\mathbf{u}$. Note that $V(\mathbf{x})$ can also be seen as a Lyapunov candidate function. Excellent references exist on passivity in linear and nonlinear systems, see Vidyasagar (1993); Wen (1988); Slotine & Li (1991); Isidori (1995).

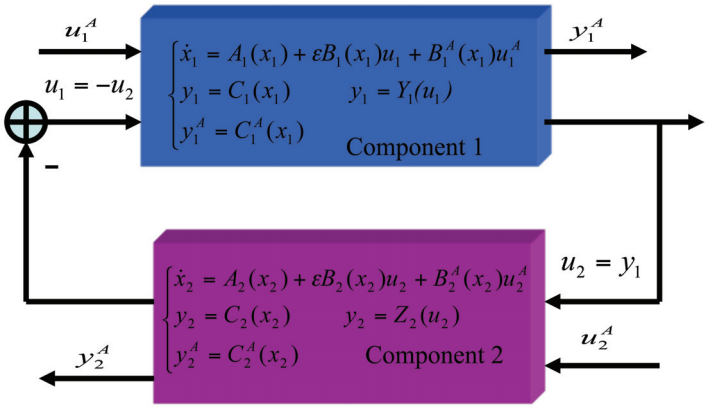



Fig. 8. Admittance-Impedance feedback connection of two nonlinear subsystems.

We can use the nonlinear impedance and admittance operators introduced in section 2.4 to find the state space representation of the impedance and admittance of Evolving System components. The nonlinear state space representation of the admittance of one component connected to the impedance of a second component is shown in fig. 8. We use the following representation for components that are nonlinear in state

$$\left\{\begin{array}{lcl} \dot{\mathbf{x}}_i &=& \mathbf{A}_i(\mathbf{x}_i) + \epsilon \mathbf{B}_i(\mathbf{x}_i)\mathbf{u}_i + \mathbf{B}_i^A(\mathbf{x}_i)\mathbf{u}_i^A \\ \mathbf{y}_i &=& \mathbf{C}_i(\mathbf{x}_i) \\ \mathbf{y}_i^A &=& \mathbf{C}_i^A(\mathbf{x}_i) \end{array}\right. \tag{36}$$

**Definition 5.2** *Consider a system that is nonlinear in state and is given by*

$$\left\{\begin{array}{lcl} \dot{\mathbf{x}} &=& \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u} \\ \mathbf{y} &=& \mathbf{C}(\mathbf{x}) \end{array}\right. \tag{37}$$

*We say that this system is Strictly Passive when* $\exists V(\mathbf{x}) > 0\ \forall \mathbf{x} \neq 0$ *such that*

$$\dot{V}(\mathbf{x}) = \langle \mathbf{u}, \mathbf{y} \rangle - S(\mathbf{x}) \tag{38}$$

*with* $S(\mathbf{x}) > 0\ \forall \mathbf{x} \neq 0$.

**Definition 5.3** *Consider a nonlinear system of the form given by*

$$\left\{\begin{array}{lcl} \dot{\mathbf{x}} &=& \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})u \\ \mathbf{y} &=& \mathbf{C}(\mathbf{x}) \end{array}\right. \tag{39}$$

*We say that this system is Almost Strictly Passive (ASP) when there is some output feedback,* $\mathbf{u} = \mathbf{G}\mathbf{y} + \mathbf{u}_r$, *that makes it strictly passive.*

We can state the following result about the inheritance of strict passivity in systems connected in feedback.

**Theorem 5.4** *Suppose we have a pair of subsystems of the form*

$$\left\{\begin{array}{lcl} \dot{\mathbf{x}}_i &=& \mathbf{A}_i(\mathbf{x}_i) + \mathbf{B}_i(\mathbf{x}_i)\mathbf{u}_i + \mathbf{B}_i^A(\mathbf{x}_i)\mathbf{u}_i^A \\ \mathbf{y}_i &=& \mathbf{C}_i(\mathbf{x}_i) \\ \mathbf{y}_i^A &=& \mathbf{C}_i^A(\mathbf{x}_i) \end{array}\right. \tag{40}$$

*where* $i = 1,2$ *and both subsystems* $\left(\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_1^A\end{bmatrix}, \begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_1^A\end{bmatrix}\right)$ *and* $\left(\begin{bmatrix}\mathbf{u}_2\\ \mathbf{u}_2^A\end{bmatrix}, \begin{bmatrix}\mathbf{y}_2\\ \mathbf{y}_2^A\end{bmatrix}\right)$ *are strictly passive with energy storage functions* $V_1(\mathbf{x}_1)$ *and* $V_2(\mathbf{x}_2)$. *Then the feedback connection of the two subsystems, where* $(\mathbf{y}_1 = \mathbf{u}_2)$ *and* $(\mathbf{u}_1 = -\mathbf{y}_2)$, *will leave the resulting composite system* $\left(\mathbf{u}_A \equiv \begin{bmatrix}\mathbf{u}_1^A\\ \mathbf{u}_2^A\end{bmatrix}, \mathbf{y}_A \equiv \begin{bmatrix}\mathbf{y}_1^A\\ \mathbf{y}_2^A\end{bmatrix}\right)$ *strictly passive.*

**Proof:** Form the composite system by connecting $(\mathbf{u}_1,\mathbf{y}_1)$ in feedback with $(\mathbf{u}_2,\mathbf{y}_2)$, which gives us $\mathbf{y}_1 = \mathbf{u}_2$ and $\mathbf{u}_1 = -\mathbf{y}_2$. Let the energy storage function for the composite system be $V(\mathbf{x}) = V_1(\mathbf{x}_1) + V_2(\mathbf{x}_2)$. Using the fact that both components are strictly passive, and making the substitutions for the feedback connection, we have

$$\begin{array}{rcl} \dot{V}(\mathbf{x}) &=& \left\langle \begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_1^A\end{bmatrix}, \begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_1^A\end{bmatrix} \right\rangle + \left\langle \begin{bmatrix}\mathbf{u}_2\\ \mathbf{u}_2^A\end{bmatrix}, \begin{bmatrix}\mathbf{y}_2\\ \mathbf{y}_2^A\end{bmatrix} \right\rangle - (S_1(\mathbf{x}_1) + S_2(\mathbf{x}_2)) \\ &=& \left(\langle \mathbf{u}_1, \mathbf{y}_1 \rangle + \langle \mathbf{u}_1^A, \mathbf{y}_1^A \rangle\right) + \left(\langle \mathbf{u}_2, \mathbf{y}_2 \rangle + \langle \mathbf{u}_2^A, \mathbf{y}_2^A \rangle\right) - S(\mathbf{x}) \\ &=& \left(\langle -\mathbf{y}_2, \mathbf{y}_1 \rangle + \langle \mathbf{u}_1^A, \mathbf{y}_1^A \rangle\right) + \left(\langle \mathbf{y}_1, \mathbf{y}_2 \rangle + \langle \mathbf{u}_2^A, \mathbf{y}_2^A \rangle\right) - S(\mathbf{x}) \\ &=& \left\langle \begin{bmatrix}\mathbf{u}_1^A\\ \mathbf{u}_2^A\end{bmatrix}, \begin{bmatrix}\mathbf{y}_1^A\\ \mathbf{y}_2^A\end{bmatrix} \right\rangle - S(\mathbf{x}) \end{array} \tag{41}$$

with $S(\mathbf{x}) \equiv S_1(\mathbf{x}_1) + S_2(\mathbf{x}_2) > 0 \ \forall \mathbf{x} \equiv \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \neq 0$. Therefore, by the definition of strict passivity, the composite system, given by $\left( \mathbf{u}_A \equiv \begin{bmatrix} \mathbf{u}_1^A \\ \mathbf{u}_2^A \end{bmatrix}, \mathbf{y}_A \equiv \begin{bmatrix} \mathbf{y}_1^A \\ \mathbf{y}_2^A \end{bmatrix} \right)$ remains strictly passive. □

We have shown that the feedback connection of two strictly passive systems results in a composite system that is strictly passive. Hence, strict passivity is inherited by systems connected in feedback. We now give a result on the inheritance of almost strict passivity.

**Theorem 5.5** *Suppose we have a pair of subsystems of the form*

$$\begin{cases} \dot{\mathbf{x}}_i &= \mathbf{A}_i(\mathbf{x}_i) + \epsilon \mathbf{B}_i(\mathbf{x}_i)\mathbf{u}_i + \mathbf{B}_i^A(\mathbf{x}_i)\mathbf{u}_i^A \\ \mathbf{y}_i &= \mathbf{C}_i(\mathbf{x}_i) \\ \mathbf{y}_i^A &= \mathbf{C}_i^A(\mathbf{x}_i) \end{cases} \tag{42}$$

*where* $i$ = 1,2 *and both subsystems* $\left( \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_1^A \end{bmatrix}, \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_1^A \end{bmatrix} \right)$ *and* $\left( \begin{bmatrix} \mathbf{u}_2 \\ \mathbf{u}_2^A \end{bmatrix}, \begin{bmatrix} \mathbf{y}_2 \\ \mathbf{y}_2^A \end{bmatrix} \right)$ *are almost strictly passive with energy storage function* $V_1(\mathbf{x}_1)$ *and* $V_2(\mathbf{x}_2)$. *Then the feedback connection of the two subsystems, where* ($\mathbf{y}_1 = \mathbf{u}_2$) and ($\mathbf{u}_1 = -\mathbf{y}_2$), *will leave the resulting composite system* $\left( \mathbf{u}_A \equiv \begin{bmatrix} \mathbf{u}_1^A \\ \mathbf{u}_2^A \end{bmatrix}, \mathbf{y}_A \equiv \begin{bmatrix} \mathbf{y}_1^A \\ \mathbf{y}_2^A \end{bmatrix} \right)$ *almost strictly passive*.

**Proof:** By the definition of almost strict passivity, there exists output feedback control that makes each of the subsystems strictly passive. Output feedback of the form $\mathbf{u} = \mathbf{G}\mathbf{y} + \mathbf{u}_r$ can be added to a system ($\mathbf{A}(\mathbf{x})$,$\mathbf{B}(\mathbf{x})$,$\mathbf{C}(\mathbf{x})$) to obtain

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{G}\mathbf{C}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u}_r \\ \mathbf{y} &= \mathbf{C}(\mathbf{x}) \end{cases} \tag{43}$$

Let $\mathbf{u}_i^A = \mathbf{G}_i^A \mathbf{y}_i^A + \mathbf{u}_i^r$, where $i$ = 1, 2, be the output feedback that makes the subsystems given by eq. 42 strictly passive. The subsystems with $\mathbf{u}_i^A$ defined as above are now both strictly passive. We can connect the two subsystems in feedback, with ($\mathbf{y}_1 = \mathbf{u}_2$) and ($\mathbf{u}_1 = -\mathbf{y}_2$). By Theo. 5.4, the composite system resulting from the feedback connection of two strictly passive systems is strictly passive. Thus, the composite system ($\mathbf{u}_A$,$\mathbf{y}_A$) is strictly passive. Now let $\begin{bmatrix} \mathbf{u}_1^A \\ \mathbf{u}_2^A \end{bmatrix} \equiv \begin{bmatrix} \mathbf{G}_1^A & 0 \\ 0 & \mathbf{G}_2^A \end{bmatrix} \begin{bmatrix} \mathbf{y}_1^A \\ \mathbf{y}_2^A \end{bmatrix} + \begin{bmatrix} \mathbf{u}_1^r \\ \mathbf{u}_2^r \end{bmatrix}$. Adding this output feedback to the composite system, ($\mathbf{u}_A$,$\mathbf{y}_A$) formed from the original subsystem components without the output feedback, is equivalent to adding the output feedback to the components and then connecting them in feedback. Since the two methods of adding output feedback are equivalent, we can add output feedback to the composite system, resulting in a strictly passive system. Hence by the definition of almost strict passivity, the composite system is almost strictly passive. Thus, almost strict passivity is inherited, and the result is true. □

**Theorem 5.6** *A LTI system given by* ($\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$) *is strict positive real iff it is strictly passive.*

**Proof:** First we show that if ($\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$) is SPR, then it is strictly passive. Since ($\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$) is SPR, the Kalman-Yacubovic Lemma (Vidyasagar, 1993) implies that $\exists\ \epsilon > 0$ such that

$$\begin{cases} (\mathbf{A} + \epsilon\mathbf{I})^T\mathbf{P} + \mathbf{P}(\mathbf{A} + \epsilon\mathbf{I}) = -\mathbf{Q} \\ \mathbf{P}\mathbf{B} = \mathbf{C}^T \end{cases} \tag{44}$$

with $\mathbf{Q} \geq 0$ and $\mathbf{P} > 0$. We can rearrange eq. 44 to obtain

$$\begin{cases} \mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} = -(\mathbf{Q} + 2\epsilon\mathbf{P}) \\ \mathbf{P}\mathbf{B} = \mathbf{C}^T \end{cases} \tag{45}$$

Since $\mathbf{P} > 0$ and $\mathbf{Q} \geq 0$, then $\mathbf{W}(\epsilon) \equiv \mathbf{Q} + 2\epsilon\mathbf{P} > 0$. Choose $V(\mathbf{x}) \equiv \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x}$ with $\mathbf{P}$ chosen as in eq. 44 and $\mathbf{P} > 0$. The time derivative along any state trajectory of $V(\mathbf{x})$ is given by

$$\begin{aligned} \dot{V}(\mathbf{x}) &= \tfrac{1}{2}\left(\mathbf{x}^T\mathbf{P}\mathbf{A}\dot{\mathbf{x}} + \dot{\mathbf{x}}^T\mathbf{P}\mathbf{B}\mathbf{u}\mathbf{x}\right) \\ &= \tfrac{1}{2}\mathbf{x}^T(\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{x} + \mathbf{x}^T\mathbf{P}\mathbf{B}\mathbf{u} \\ &= -\tfrac{1}{2}\mathbf{x}^T\mathbf{W}(\epsilon)\mathbf{x} + (\mathbf{C}\mathbf{x})^T\mathbf{u} \\ &= -S(\mathbf{x}) + \mathbf{y}^T\mathbf{u} \end{aligned} \tag{46}$$

Therefore (**A**,**B**,**C**) is strictly passive.

We now show that if (**A**,**B**,**C**) is strictly passive, then it is SPR. Since (**A**,**B**,**C**) is strictly passive, we have $V(\mathbf{x}) = -S(\mathbf{x}) + \mathbf{y}^T\mathbf{u}$ with $S(\mathbf{x}) > 0$. Choose $S(\mathbf{x}) \equiv \mathbf{W}(\epsilon)$, with $\mathbf{P}$ and $\mathbf{Q}$ as in eq. 44. Then all the previous arguments can be reversed, giving the desired result. □

In Section 2.4 we showed that the physical connection of two Evolving System components is equivalent to the feedback connection of the admittance of one component and the impedance of the other component. Consequently, if the subsystem components of an Evolving System are in admittance-impedance form, then by Theo. 5.4 and Theo. 5.5 we see that strict passivity and almost strict passivity are traits that are always inherited in nonlinear Evolving Systems. Therefore, if the impedance of one component and the admittance of the other component are both strictly passive, then their feedback connection will be strictly passive. The same is true for almost strict passivity.

The following result gives the sufficient condition for an LTI Evolving System with an adaptive key component controller to be guaranteed to have bounded gains and asymptotic output tracking.

**Theorem 5.7** *Consider a two component linear time-invariant Evolving System given by*

$$\begin{cases} \dot{\mathbf{x}}_i &= \mathbf{A}_i\mathbf{x}_i + \epsilon\mathbf{B}_1\mathbf{u}_1 + \mathbf{B}_1^A\mathbf{u}_1^A \\ \mathbf{y}_i &= \mathbf{C}_i\mathbf{x}_i \end{cases} \tag{47}$$

*where* $i = 1,2$. *Let component 1 have an adaptive key component controller with the following direct adaptive control law*

$$\begin{cases} \mathbf{u}_1^A &= \mathbf{G}\mathbf{y}_1^A \\ \dot{\mathbf{G}} &= -\mathbf{y}_1^A(\mathbf{y}_1^A)^T\mathbf{H};\ \mathbf{H} > 0 \end{cases} \tag{48}$$

*If both components of the Evolving System are almost strictly passive from an admittance-impedance point of view, then the adaptive gains,* $\mathbf{G}$, *are bounded and* $\mathbf{y} \to \mathbf{0}$ as $t \to \infty$.

**Proof:** By Theo. 5.5, since both components are almost strictly passive, then the composite system resulting from the feedback connection of the components is almost strictly passive. In Theo. 5.6, we showed that for linear time-invariant systems, strict passivity is equivalent to the strict positive real property. A system that is almost strict positive real (ASPR) is one that can be made strict positive real with output feedback. Hence, for LTI systems, almost strict positive real is equivalent to almost strict passivity. Since the Evolving System given by eq. 47 is an almost strictly passive LTI system, it is an almost strict positive real system.

Theorem 4.3 states that the sufficient condition for a LTI system with an adaptive control law given by eq. 48 to be guaranteed to have bounded gains and asymptotic output tracking is that the system be almost strict positive real. Therefore, by Theo. 4.3, the adaptive gains, **G**, are bounded and $\mathbf{y}\rightarrow\mathbf{0}$ as $t\rightarrow\infty$.
We think of the ports $\mathbf{u}_1$ and $\mathbf{u}_2$ as being the admittance-impedance ports through which the components make contact. For an Evolving System that has an adaptive key component controller, one of the ports $\mathbf{u}_1^A$ or $\mathbf{u}_2^A$ would be used for the key component controller to augment the system to restore stability if necessary. The ports $\mathbf{u}_1^A$ or $\mathbf{u}_2^A$ could also be used to add output feedback to make the Evolving System strictly passive.

## 6. Inheritance of dissipativity properties in evolving systems

In this section we briefly present several results that were presented in (Frost & Balas, 2010).
**Definition 6.1** *Consider a nonlinear system of the form given by*

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u} \\ \mathbf{y} &= \mathbf{C}(\mathbf{x}) \end{cases} \tag{49}$$

*We say that this system is Strictly Dissipative when* $\exists V(\mathbf{x}) > 0\ \forall \mathbf{x} \neq 0$ *such that* $\forall \mathbf{x}$

$$\begin{cases} \nabla V \mathbf{A}(\mathbf{x}) \leq -S(\mathbf{x}) \\ \nabla V \mathbf{B}(\mathbf{x}) = \mathbf{C}^T(\mathbf{x}) \end{cases} \tag{50}$$

where $\nabla V \equiv$ gradient $V$ and $S(\mathbf{x}) > 0\ \forall \mathbf{x} \neq 0$.
The function $V(\mathbf{x})$ is the Lyapunov candidate function for eq. 49. The function, $V(\mathbf{x})$, is related to $\nabla V$ by the following

$$\dot{V}(\mathbf{x}) \equiv \nabla V[\mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u}] \tag{51}$$

The above says that the storage rate is always less than the external power. This can be seen by using eq. 50 to obtain

$$\begin{aligned} \dot{V}(\mathbf{x}) &\equiv \nabla V[\mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u}] \\ &\leq -S(\mathbf{x}) + \mathbf{C}^T(\mathbf{x})\mathbf{u} \\ &= -S(\mathbf{x}) + \langle \mathbf{y}, \mathbf{u} \rangle \end{aligned} \tag{52}$$

Taking $\mathbf{u} \equiv \mathbf{0}$, it is easy to see that eq. 52 implies eq. 50(a), but not necessarily eq. 50(b). So eq. 50 implies eq. 52 but not conversely. The two are only equivalent if eq. 50(a) is an equality. If the inequalities in eq. 50 and eq. 52 are equalities, then the property is called Strict Passivity, which was defined in section 5.
**Definition 6.2** *Consider a nonlinear system of the form given by*

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u} \\ \mathbf{y} &= \mathbf{C}(\mathbf{x}) \end{cases} \tag{53}$$

*We say that this system is Almost Strictly Dissipative (ASD) when there is some output feedback,* $\mathbf{u} = \mathbf{G}\mathbf{y} + \mathbf{u}_r$, *that makes it strictly dissipative.*
**Theorem 6.3** *If a nonlinear system given by* $(\mathbf{A}(\mathbf{x}),\mathbf{B}(\mathbf{x}),\mathbf{C}(\mathbf{x}))$ *is strictly passive, then it is strictly dissipative.*

**Theorem 6.4** *A LTI system given by* (**A**,**B**,**C**) *is strictly dissipative iff it is strictly passive.*
**Theorem 6.5** *Suppose we have a pair of subsystems of the form*

$$\left\{ \begin{array}{rcl} \dot{\mathbf{x}}_i & = & \mathbf{A}_i(\mathbf{x}_i) + \epsilon \mathbf{B}_i(\mathbf{x})\mathbf{u}_i + \mathbf{B}_i^A(\mathbf{x})\mathbf{u}_i^A \\ \mathbf{y}_i & = & \mathbf{C}_i(\mathbf{x}_i) \\ \mathbf{y}_i^A & = & \mathbf{C}_i^A(\mathbf{x}_i) \end{array} \right. \tag{54}$$

*where* $i$ = 1,2 *and both subsystems* $\left( \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_1^A \end{bmatrix}, \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_1^A \end{bmatrix} \right)$ *and* $\left( \begin{bmatrix} \mathbf{u}_2 \\ \mathbf{u}_2^A \end{bmatrix}, \begin{bmatrix} \mathbf{y}_2 \\ \mathbf{y}_2^A \end{bmatrix} \right)$ *are almost strictly dissipative with energy storage function* $V_1(\mathbf{x}_1)$ *and* $V_2(\mathbf{x}_2)$, *and*

$$\nabla V_i \epsilon \mathbf{B}_i(\mathbf{x}_i) = \epsilon \mathbf{C}_i^T(\mathbf{x}_i) \tag{55}$$

*Then the feedback connection of the two subsystems, where* ($\mathbf{y}_1 = \mathbf{u}_2$) *and* ($\mathbf{u}_1 = -\mathbf{y}_2$), *will leave the resulting composite system* $\left( \mathbf{u}_A \equiv \begin{bmatrix} \mathbf{u}_1^A \\ \mathbf{u}_2^A \end{bmatrix}, \mathbf{y}_A \equiv \begin{bmatrix} \mathbf{y}_1^A \\ \mathbf{y}_2^A \end{bmatrix} \right)$ *almost strictly dissipative.*

A corollary of Theo. 6.5 is that strict dissipativity is inherited by systems connected in feedback.
**Corollary 6.6** *Suppose we have a pair of subsystems of the form*

$$\left\{ \begin{array}{rcl} \dot{\mathbf{x}}_i & = & \mathbf{A}_i(\mathbf{x}_i) + \epsilon \mathbf{B}_i(\mathbf{x})\mathbf{u}_i + \mathbf{B}_i^A(\mathbf{x})\mathbf{u}_i^A \\ \mathbf{y}_i & = & \mathbf{C}_i(\mathbf{x}_i) \\ \mathbf{y}_i^A & = & \mathbf{C}_i^A(\mathbf{x}_i) \end{array} \right. \tag{56}$$

*where* $i$ = 1,2 *and both subsystems* $\left( \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_1^A \end{bmatrix}, \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_1^A \end{bmatrix} \right)$ *and* $\left( \begin{bmatrix} \mathbf{u}_2 \\ \mathbf{u}_2^A \end{bmatrix}, \begin{bmatrix} \mathbf{y}_2 \\ \mathbf{y}_2^A \end{bmatrix} \right)$ *are strictly dissipative with energy storage function* $V_1(\mathbf{x}_1)$ *and* $V_2(\mathbf{x}_2)$, *and*

$$\nabla V_i \epsilon \mathbf{B}_i(\mathbf{x}_i) = \epsilon \mathbf{C}_i^T(\mathbf{x}_i) \tag{57}$$

*Then the feedback connection of the two subsystems, where* ($\mathbf{y}_1 = \mathbf{u}_2$) *and* ($\mathbf{u}_1 = -\mathbf{y}_2$), *will leave the resulting composite system* $\left( \mathbf{u}_A \equiv \begin{bmatrix} \mathbf{u}_1^A \\ \mathbf{u}_2^A \end{bmatrix}, \mathbf{y}_A \equiv \begin{bmatrix} \mathbf{y}_1^A \\ \mathbf{y}_2^A \end{bmatrix} \right)$ *strictly dissipative.*

Theorem 6.5 and Cor. 6.6 can both be used to show that two component nonlinear Evolving Systems with components that are either both almost strictly dissipative or strictly dissipative from an admittance-impedance point of view inherit the properties of their subsystem components. Thus strict dissipativity and almost strict dissipativity are traits that are always inherited in nonlinear Evolving Systems.
**Theorem 6.7** *Consider a two component nonlinear time-invariant Evolving System given by*

$$\left\{ \begin{array}{rcl} \dot{\mathbf{x}}_i & = & \mathbf{A}_i(\mathbf{x}_i) + \epsilon \mathbf{B}_i(\mathbf{x}_i)\mathbf{u}_i + \mathbf{B}_i^A(\mathbf{x}_i)\mathbf{u}_1^A \\ \mathbf{y}_i & = & \mathbf{C}_i(\mathbf{x}_i) \end{array} \right. \tag{58}$$

*where* $i$ = 1,2 *with energy storage functions* $V_1(\mathbf{x}_1)$ *and* $V_2(\mathbf{x}_2)$. *Let component 1 have an adaptive key component controller with the following direct adaptive control law*

$$\left\{ \begin{array}{rcl} \mathbf{u}_1^A & = & \mathbf{G}\mathbf{y}_1^A \\ \dot{\mathbf{G}} & = & -\mathbf{y}_1^A(\mathbf{y}_1^A)^T\mathbf{H};\ \mathbf{H} > 0 \end{array} \right. \tag{59}$$

*Assume that $V_1$ and $V_2$ are positive $\forall \mathbf{x} \neq 0$ and radially unbounded, and $(\mathbf{A}(\mathbf{x}),\mathbf{B}(\mathbf{x}),\mathbf{C}(\mathbf{x}))$ are continuous functions of $\mathbf{x}$ and $S(\mathbf{x})$ is positive $\forall \mathbf{x} \neq 0$ and has continuous partial derivatives in $\mathbf{x}$. Furthermore, assume:*

1. *Component 2, given by $(\mathbf{u}_2,\mathbf{y}_2)$, is strictly dissipative and in impedance form;*
2. *Component 1, given by $(\mathbf{u}_1^A,\mathbf{y}_1^A)$, is almost strictly dissipative;*
3. *Component 1, given by $(\mathbf{u}_1,\mathbf{y}_1)$, is in admittance form.*

*Then the adaptive key component controller given by eq. 59 produces global asymptotic state stability, i.e., $\mathbf{x} \equiv [\mathbf{x}_1\mathbf{x}_2]^T \rightarrow 0$ as $t \rightarrow \infty$ with bounded adaptive gains when component 1 is joined with component 2 into an Evolved System and the outputs $\mathbf{y}_i = \mathbf{C}_i(\mathbf{x}_i) \rightarrow 0$ as $t \rightarrow \infty$.*

The above results assume that the Lyapunov function, $V(\mathbf{x})$, is defined on the entire domain, $\mathbb{R}^n$, of the system. Thus all the stability and dissipativity results are global results. The same is true for the other results given in this chapter. For instance, Theo. 6.7 says that a nonlinear Evolving System with an adaptive key component controller as given by eq. 48 will have bounded gains and globally asymptotic state tracking. However, the Lyapunov function, $V(\mathbf{x})$, might only be defined on a neighborhood $N_i(0,r_i) \equiv \left\{ \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|} < r_i \right\}$ of the origin, in which case the results could only be local at the best.

Using Lemma 1 from (Balas et al., 2008), $\forall \delta > 0$ such that if the initial conditions of the system are close enough to the origin, i.e., within $N_\delta = (0, \delta)$, then the trajectories are guaranteed to stay in the neighborhood of the origin for which the Lyapunov function is defined. In such a case, then the results would be local. For instance, if the Lyapunov function $V(\mathbf{x})$ in Theo. 6.7 only has the assumed properties on a neighborhood $N_i(0,r_i) \equiv \left\{ \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|} < r_i \right\}$ of the origin and the trajectories all remain inside the neighborhood, then the stability is locally asymptotic to the origin. In that case, Theo. 6.7 gives the result that a nonlinear Evolving System with an adaptive key component controller as given by eq. 59 will have bounded gains and locally asymptotic state tracking.

## 7. Conclusions

In this chapter, we presented the motivation and the framework for Evolving Systems, a new area of aerospace research. We developed the adaptive key component controller approach to maintain stability in Evolving Systems that would otherwise fail to inherit the stability traits of their components. We showed that strict passivity, almost strict passivity, strict dissipativity, and almost strict dissipativity are inherited by systems connected in feedback. Using the impedance-admittance formulation of contact dynamics between components of Evolving Systems, we showed that these traits are also always inherited in nonlinear Evolving Systems. Finally, we gave sufficient conditions for the use of the adaptive key component controller with linear and nonlinear Evolving Systems.

## 8. References

Balas, M. J. (1982). Trends in large space structure control theory: Fondest hopes, wildest dreams, *IEEE Transactions on Automatic Control* AC-27: 522–535.

Balas, M. J. & Frost, S. A. (2007). Evolving systems: Inheriting stability with evolving controllers, *Proceedings 47th Israel Annual Conference on Aerospace Sciences*, Tel-Aviv, Israel.

Balas, M. J. & Frost, S. A. (2008). Adaptive key component control for nonlinear evolving systems, *Proceedings of the ASME Conference on Smart Materials, Adaptive Structures and Intelligent Systems*, Ellicott City, MD.

Balas, M. J., Frost, S. A. & Hadaegh, F. Y. (2006). Evolving systems: A theoretical foundation, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Keystone, CO.

Balas, M. J., Harvey, S. & Mehiel, E. (2008). Nonlinear adaptive control with persistent disturbances, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Honolulu, HI.

Bansenauer, B. T. & Balas, M. J. (January-February 1995). Reduced-order model based control of the flexible, articulated-truss space crane, *Journal of Guidance, Control, and Dynamics* 18(1): 135–142.

Brogan,W. L. (1991). *Modern Control Theory*, 3rd edn, Prentice-Hall, Englewood Cliffs, NJ. Corfmat, J. P. & Morse, A. S. (1976a). Control of linear systems through specified input channels, *SIAM Journal of Control and Optimization* 14(1): 163–175.

Corfmat, J. P. & Morse, A. S. (1976b). Decentralized control of linear multivariable systems, *Automatica* 12: 479–495.

Flinn, E. D. (2009). Delicate dance for mango and tango, *Aerospace America*.

Frost, S. A. (2008). *Evolving Systems: Control and Stability Inheritance in Self-Assembling Structures*, University ofWyoming, Electrical Engineering Doctoral Thesis, Laramie, WY.

Frost, S. A. & Balas, M. J. (2007a). Stability inheritance and contact dynamics of flexible structure evolving systems, *Proceedings 17th IFAC Symposium on Automatic Control in Aerospace*, Toulouse, France.

Frost, S. A. & Balas, M. J. (2007b). Stabilizing controllers for evolving systems with application to flexible space structures, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Hilton Head, SC.

Frost, S. A. & Balas, M. J. (2008a). Adaptive key component control and inheritance of almost strict passivity in evolving systems, *Proceedings AAS Landis Markley Symposium*, Cambridge, MD.

Frost, S. A. & Balas, M. J. (2008b). Adaptive key component controllers for evolving systems, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Honolulu, HI.

Frost, S. A. & Balas, M. J. (2010). Evolving systems: Adaptive key component control and inheritance of passivity and dissipativity, *preprint, Journal of Applied Mathematics and Computation, Special Issue in Honor of George Leitman* .

Fuentes, R. J. & Balas, M. J. (2000). Direct adaptive rejection of persistent disturbances, *Journal of Mathematical Analysis and Applications* 251(1): 28–39.

Harris, C. M. & Crede, C. E. (1976). *Shock and Vibration Handbook*, 2nd edn, McGraw-Hill, New York, NY.

Isidori, A. (1995). *Nonlinear Control Systems*, 3rd edn, Springer-Verlag London Limited, London, Great Britain.

Kapilal, V. (1999). Spacecraft formation flying: dynamics and control, *Proceedings of the 1999 American Control Conference*, San Diego, CA.

Meirovitch, L. (2001). *Fundamentals of Vibrations*, McGraw-Hill, International ed. New York, NY.

Michel, A. N. (1983). On the status of stability of interconnected systems, *IEEE Transactions on Automatic Control* AC-28(6): 639–653.

Mueller, J., Surka, D. M. & Udrea, B. (2001). Agent-based control of multiple satellite formation flying, *Proceedings. The 6th international symposium on artificial intelligence, robotics and automation in space: a new space odyssey,* Montreal, Canada.

NASA (1995). Space solar power, an advanced concepts study project, *Advanced Concepts Office,* NASA HQ, and *Advanced Space Analysis Office*, *NASA LaRC (SSP Technical Interchange Meeting,* Washington, D.C.) 1-2.

Ogata, K. (2002). *Modern Control Engineering*, 4th edn, Prentice-Hall, Upper Saddle River, NJ.

Saleh, J. H., Lamassoure, E. & Hastings, D. E. (2002). Space systems flexibility provided by on-orbit servicing: Part 1, *Journal of Spacecraft and Rockets* 39(4).

Sandell, Jr., N. R., Athans, M., Varaiya, P. & Safonov, M. G. (1978). Survey of decentralized control methods for large scale systems, *IEEE Transactions on Automatic Control* AC-23(2): 108–28.

Shen, W.-M., Will, P. & Khoshnevis, B. (2003). Self-assembly in space via self-reconfigurable robots, *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on* 2: 2516–2521 vol.2.

Slotine, J.-J. E. & Li,W. (1991). *Nonlinear Control Theory*, Prentice-Hall, Englewood Cliffs, NJ.

Suri, H., Will, P. & Shen, W.-M. (2006). System design of robots for application to in-space assembly, *Robotics and Automation, 2006. Proceedings. ICRA '06. IEEE International Conference on*.

Vidyasagar, M. (1993). *Nonlinear Systems Analysis*, 2nd edn, Prentice-Hall, Englewood Cliffs, NJ.

Wen, J. T. (1988). Time domain and frequency domain conditions for strict positive realness, *IEEE Transactions on Automatic Control* AC-33(10): 988–992.

Whitesides, G. M. & Grzybowski, B. (2002). Self-assembly at all scales, *Science* 295(5564): 2418– 2421.

Willems, J. L. (1986). The stabilizability of structurally fixed modes and nonstructurally fixed modes by time-varying feedback, *Proceedings of the 25th Conference on Decision and Control*, Athens, Greece.

# 8

# Evaluation of Anomaly Detection Capability for Ground-Based Pre-Launch Shuttle Operations

Rodney A. Martin, Ph.D.
*NASA Ames Research Center*
*Mail Stop 269-1*
*Moffett Field, CA 94035*
*U.S.A.*

## 1. Introduction

This chapter will provide a thorough end-to-end description of the process for evaluation of three different data-driven algorithms for anomaly detection to select the best candidate for deployment as part of a suite of IVHM (Integrated Vehicle Health Management) technologies. These algorithms will be evaluated based upon their capability for robustly detecting incipient faults or failures in the ground-based phase of pre-launch space shuttle operations, rather than based on system certifiability as performed in previous studies (Schwabacher & Waterman, 2008). Robust detection will allow for the achievement of pre-specified minimum false alarm and/or missed detection rates in the selection of alert thresholds. All algorithms will be optimized with respect to an aggregation of these same criteria. The final results will also include a formal cross-validation procedure, which will be used to perform optimization and alert threshold selection.

The data-driven algorithms to be evaluated in this study were deemed to be sufficiently mature for consideration as viable candidates for demonstration during the launch of Ares I-X. This launch represents the first test flight of Ares I, which will be the successor to the Space Shuttle for NASA's Constellation program. Our study relies upon the use of Shuttle data to act as a proxy for and in preparation for application to Ares I-X data, which uses a very similar hardware platform for the subsystems that are being targeted (TVC – Thrust Vector Control subsystem for the SRB (Solid Rocket Booster)).

Data-driven algorithms are just one of three different types being deployed, the details of which were presented in previous work (Iverson *et al.*, 2009); (Schwabacher & Waterman, 2008). The other two types of algorithms being deployed include a "rule-based" expert system, and a "model-based" system. Within these two categories, the deployable candidates have already been selected based upon non-quantitative factors such as flight heritage and system certifiability. For the rule-based system, SHINE (Spacecraft Health Inference Engine) has been selected for deployment, which is one of many components of BEAM (Beacon-based Exception Analysis for Multimissions) as described in its debut article (Mackey *et al.*, 2001). Other components of BEAM include various data-driven algorithms. BEAM is a patented technology developed at NASA's JPL (Jet Propulsion Laboratory) and

SHINE serves to aid in the management and identification of operational modes. For the "model-based" system, a commercially available package developed by QSI (Qualtech Systems, Inc.), TEAMS (Testability Engineering and Maintenance System) was highlighted in work subsequent to its debut (Cavanaugh, 2001), and has been selected for deployment to aid in diagnosis. In the context of this particular deployment, distinctions among the use of the terms "data-driven," "rule-based," and "model-based," can be found in the previously cited paper by Schwabacher and Waterman.

Although there are three different categories of algorithms that have been selected for deployment, our main focus in this chapter will be on the evaluation of three candidates for *data-driven* anomaly detection. The main thrust of the chapter is to provide instructive coverage on the topics of algorithmic optimization and alert threshold selection for the candidate data-driven algorithms. These algorithms will be optimized and compared using the AUC (Area under the ROC (Receiver Operating Characteristic) curve), which represents overall classification discriminability. The resulting optimized algorithmic parameters can then be used to perform threshold or alert selection based upon the corresponding ROC curve. This allows for the demonstration of a robust anomaly detection capability when performing alert threshold selection, which is based upon specified minimum false alarm and/or missed detection rates. As a practical measure we will also present a performance comparison among the candidate data-driven algorithms by evaluation of their computational complexity as well as the AUC. The final results will also include a formal cross-validation procedure, which will be used to provide uncertainty bounds for the optimization and alert threshold selection processes.

## 2. Motivation & background

Some insightful previous studies (Fragola, 1996); (Paté-Cornell and Dillon, 2001) discuss statistical analyses concerning general Space Shuttle risk management for the pre-Columbia era, and detail both lapses and advancements in the use of probabilistic risk analyses (PRA). We note here that probabilistic risk analysis covers the quantification of failure probabilities in the absence of large volumes of data for supporting analyses, but is not directly related to the post-design real-time monitoring phase for the detection of anomalous events. The primary author of the previously cited study (Paté-Cornell & Dillon, 2001) has long been considered a respected expert in the PRA field, and is often sought by NASA for opinions, judgment, and analyses on such issues. In a study by this same author (Paté-Cornell and Fischbeck, 1994) on the risk analyses of tiles on the space shuttle which predated the Columbia tragedy by almost a decade, it was found that the use of PRA admitted a conclusion that almost foretold the events of the eventual demise of the Space Shuttle Columbia.

Noted for the advocacy on the use of Bayesian statistics as opposed to classical frequentist statistics within PRA, it is often shown in Paté-Cornell's studies that a quantitative model-based formulation requires data gathered from a variety of sources. In the case of frequentist statistics, this requires a sufficient volume of data in order to generate a reasonably accurate estimate of risk and confidence. The use of Bayesian analysis is more apropos to the situation of developing risk estimates for novel engineered systems such as the Space Shuttle, and even more so for Ares I-X where there is an inherent scarcity of data. As such, there is the need to develop a well-informed prior distribution.

Furthermore, Paté-Cornell strongly cautions engineers against the use of a mixed approach involving both Bayesian and frequentist statistics in order to provide for consistency, and

possibly due to previous lessons learned from experience with the application of conservative risk estimates in tandem with probabilities. Although a pure Bayesian approach is not used in our study, nor we do not concern ourselves with the computation and quantification of risks and potential failures for PRA, we are motivated by another of Paté-Cornell's cautions that all too often qualitative measures are used for decision making rather than quantitative ones.

We do not present these ideas as a preamble for the investigation of the development of such PRA's in this chapter, but rather for the monitoring of systems ostensibly that PRA served to inform at the design and post-design certification stages. However, areas where these previous studies may relate to ours are in the establishment of reasonably accurate failure probabilities to inform evaluation of monitoring algorithms, and the use of a model-based approach in tandem with a data-driven approach for monitoring. A good overview of the need to establish a formal relationship between design and monitoring for space applications has been discussed thoroughly in previous work (Tumer, 2005).

We note that PRA relates to design and risk assessment, whereas the use of data driven statistical analyses performed in this study assumes that appropriate measures have already been taken in advance in order to compute and quantify probabilities of failure. As such, complementary maximum allowable probabilities of false alarm and missed detection are assumed to have already been established and available. These two metrics are the essence of the analysis we present in this chapter.

We also note as described previously that a model-based approach (TEAMS) is used in the same architecture as a data-driven approach. This is still not applicable nor indicative of an attempt to develop a PRA. The framework presented here relates to the monitoring of the designed systems rather than the assessment of risk. There is evidence that the use of a Bayesian approach has been embedded within this model-based tool (Tu *et al*., 2006), although these aspects of the system are not currently being used for monitoring in this particular demonstration or deployment. However, the availability of such functionality is a general step in the right direction nonetheless. Paté-Cornell alludes to the use of a Bayesian approach as being of paramount importance for inclusion in PRA, but it arguably has similar merit for monitoring. Even though the full Bayesian updating aspects of TEAMS are not currently being deployed for monitoring purposes, the option to include MTTF (Mean-Time-To-Failure) estimates based upon the results of PRA exists in TEAMS-RT, the real-time counterpart to our chosen model-based approach. These estimates can be used to calculate the probability of each suspect failure mode.

Our primary concern in this chapter, however, is with an optimization, evaluation, and performance assessment of candidate data-driven monitoring algorithms for the relevant platform. As such, we note that there have been other related studies *e.g.* (Hart, 1990) on the statistical analyses of data related to the exact subsystem that we are considering here. However, we have found very little indication that the results of these studies were adopted and implemented in any fashion. The mathematical sophistication of these studies was most likely well beyond the capability of practical implementation at the time, as they were performed nearly two decades ago. However, advancements in the use of ROC analysis that have recently gained traction or application in the aerospace IVHM community have motivated its current use for space-based propulsion systems of the exact type analyzed in previous work of the current author (Martin, 2007); (Martin *et al*., 2007). Furthermore, ROC analysis is a much more straightforward and broadly applicable tool used for statistical analysis than in the previously cited work (Hart, 1990), and more useful for the design of alarm and monitoring systems.

## 3. Approach

Ostensibly, operating procedures for the thrust vector control (TVC) subsystem on Ares I-X used to support gimbaling of the solid rocket booster are also patterned after operations of the Space Shuttle. As such, for the purposes of our analysis here we consider two primary phases of operation of the TVC subsystem based upon Space Shuttle operations. Testing of the TVC in the VAB (Vehicle Assembly Building) often occurs anywhere from several weeks to a few months prior to the point at which the vehicle is fully assembled and ready to be rolled out on the MLP (Mobile Launch Platform) to the pad for launch. Launch countdown begins while the assembled vehicle is already rolled out to the launch platform; however, occasionally certain tests that normally occur in the VAB are also performed at the pad.
To simplify the distinction between the two main phases of operation, we will only consider tests of the TVC subsystem when the vehicle is in the VAB, and when the fully assembled vehicle is at the pad after launch countdown has commenced. Data from previous Space Shuttle launches has been gathered from both phases, and as such both will be considered viable candidates for analysis using the methods described previously. The availability of the data during each phase and its ability to meet certain requirements will determine the choice of a particular phase of operation as the primary test phase for evaluation of data-driven algorithms. Specifically, the requirement to be met is the availability of a sufficient amount of data to undergo a formal cross-validation procedure that can be processed in a reasonable amount of time. *A priori* knowledge on computational complexity of each data-driven algorithm will aid in determining if they can be processed in a reasonable amount of time, given a fixed volume of data that sufficiently characterizes nominal behavior spanning multiple launches.
Finally, all data must be pre-processed and transformed into a standard format fit for utilization of all data-driven algorithms. This pre-processing can be decomposed into two basic steps. The first step requires resampling of data that is recorded only when parameter values change or that are derived from different data sources. This resampling technique is performed by holding the last known value, resulting in a new data file where each row corresponds to a single time step, and each column corresponds to a distinct parameter, regardless of its origin. Thus, the data is merged into one aggregate "matrix" representation. The second step necessitates normalization of all continuous-valued parameters so that any potential algorithmic biases resulting from parameters that have drastically different ranges are effectively mitigated. If any discrete-valued parameters or continuous parameters that have fixed values exist in the dataset, they will bypass this normalization step in order to prevent potential singularities in algorithmic processing. For the purposes of this chapter, unless otherwise stated it will be assumed henceforth that all data has been processed and transformed into this standard format.

### 3.1 IMS (Inductive Monitoring System)

The Inductive Monitoring System (IMS) is a distance-based anomaly detection tool that uses a data driven technique called clustering to extract models of normal system operation from archived data. IMS works with vectors of data values as described in the previous section. During the learning process, IMS analyzes data collected during periods of normal system operation to build a system model. It characterizes how the parameters relate to one another during normal operation by finding areas in the vector space where nominal data tends to fall. These areas are called nominal operating regions and correspond to clusters of nearby,

similar points found by the IMS clustering algorithm. IMS represents these nominal operating regions as hyper-boxes in the vector space, providing a minimum and maximum value limit for each parameter of a vector contained in a particular hyper-box. These hyper-box cluster specifications are stored in a knowledge base that IMS uses for real-time telemetry monitoring or archived data analysis. Figure 1a shows an overview of the IMS method.

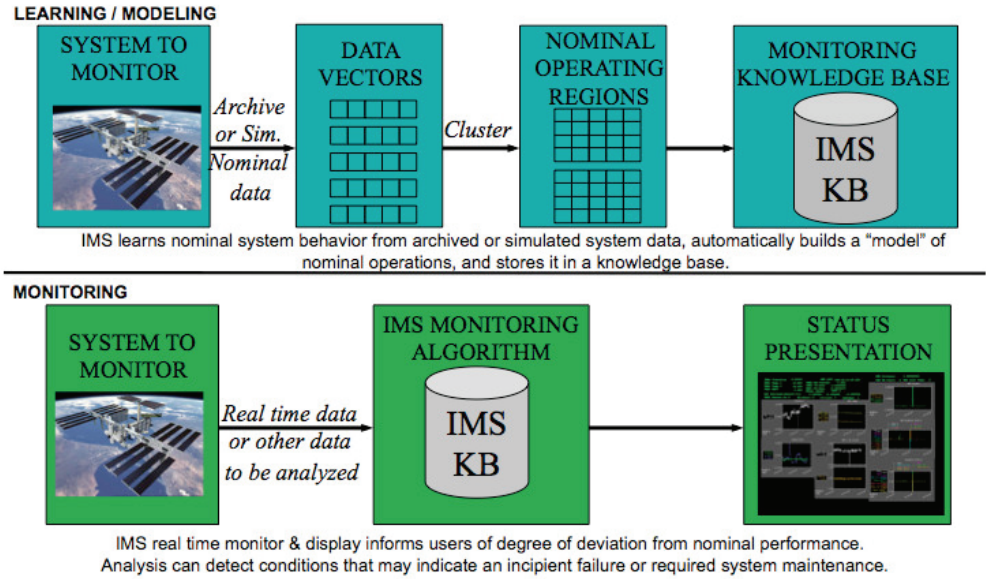



Fig. 1a. Inductive Monitoring System (IMS) Overview

### 3.1.1 IMS learning process

In general, the number and extent of nominal operating regions created during the IMS learning process is determined by three learning parameters: the "maximum cluster radius" can be used to adjust the size and number of clusters derived from a fixed number of training data points, the "initial cluster size" is used to adjust the tolerance of newly created nominal operating regions, and the "cluster growth percent" is used to adjust the percent increase in size of a nominal operating region when incorporating new training data vectors. More specifically, the learning algorithm builds a knowledge base of clusters from successively processed vectors of training data. As such, the clustering approach is incremental in nature, which distinguishes it from well-known methods such as k-means clustering where the resulting clusters are independent of the ordering of the vectors. With the processing of each new training data vector, the distance from this new vector to the centroid of the nearest cluster in the knowledge base is computed. If this distance is below a pre-specified value, the "maximum cluster radius," the new vector is summarily incorporated into that cluster. The upper or lower limits for each affected dimension of the cluster are expanded respectively according to the "cluster growth percent" parameter to reflect the inclusion of the new vector. This incremental, inductive process gives IMS an advantage over other clustering methods such as k-means, since it tends to group temporally related points during the learning process. The grouping of temporally related

points may also aid in discovering distinct system operations, which makes IMS more amenable to the specific goal of monitoring time series data for system operations.
The "cluster growth percent" parameter is used to adjust the learning rate. It establishes a fixed "growth" percentage difference for expansion of each dimension when updating previously formed clusters. This "cluster growth percent" learning parameter is therefore clearly proportional to the learning rate, due to the increased number of training data points that will be assigned to each new cluster per iteration for higher values of the "cluster growth percent" parameter. Naturally, the number of clusters in the knowledge base for a given training data set will increase as the "maximum cluster radius" and "cluster growth percent" values are decreased. Therefore, an inverse relationship between the maximum cluster radius and the number of clusters in the knowledge base exists. This dependence can be exploited to regulate the final size of the knowledge base in order to accommodate resource limitations in the computers running IMS, and to optimize selected metrics.
If the distance between a newly processed vector and the centroid of the nearest cluster in the knowledge base is above the pre-specified "maximum cluster radius" value, a new cluster is created. The formation of a new cluster is accomplished by creating a hyper-box whose dimensions are based upon forming a window around each element of the new training data vector. The window is defined by introducing the "initial cluster size" parameter which is used to adjust the learning tolerance. This "initial cluster size" learning parameter represents a fixed percentage of the value for each dimension of the new training vector and as such relates directly to the size of newly established clusters. The "initial cluster size" and "cluster growth percent" learning parameters also act as buffers which enable a provisional allowance for manufacturing sensor tolerances and for sensors that may have suffered from deterioration due to wear. Furthermore, these learning parameters provide increased coverage to compensate for training data that may not fully characterize the nominal performance envelope.

### 3.1.2 IMS monitoring process

During the monitoring operation, IMS reads and normalizes real-time or archived data values, formats them into the predefined vector structure, and searches the knowledge base of nominal operating regions to see how well the new data vector fits the nominal system characterization. After each search, IMS returns the distance from the new vector to the nearest nominal operating region, called the composite distance. Data that matches the normal training data well will have a composite distance of zero. If one or more of the data parameters is slightly outside of expected values, a small non-zero result is returned. As incoming data deviates further from the normal system data, indicating a possible malfunction, IMS will return a higher composite distance value to alert users to the anomaly. IMS also calculates the contribution of each individual parameter to the composite deviation, and the index of the nearest cluster to the monitored point, which can help identify and isolate the cause of the anomaly.

### 3.1.3 IMS score computations

This section describes the computation of the IMS composite score and individual contributing scores. The IMS scores are computed on a monitored data stream, and are functions of distances to the nearest clusters which have been computed from training data. Table 1 lists the notation used in the score calculation equations:

| | |
|---|---|
| $y$ | Raw input data in physical measurement units |
| $s$ | Sample index |
| $p$ | Parameter index |
| $c$ | Cluster index |
| $n_p$ | Number of parameters |
| $S$ | Scale factor for IMS internal computations; as of this writing S = 10,000 |
| $k_z$ | Standard deviation multiplier for normalization constant |
| $\mathcal{O}$ | Offset to make IMS internal numbers positive |
| $U_{p,c}$ | Upper cluster bound for cluster $c$ in dimension $p$ |
| $L_{p,c}$ | Lower cluster bound for cluster $c$ in dimension $p$ |

Table 1. Notation

1. During training, data is converted to a multiple $k_z$ of the Z-score over the training data. The mean and standard deviation of each parameter in the training data is computed. The normalized data is then multiplied by a scale factor S and an offset $\mathcal{O}$ is added to make all of the numbers positive.

   The modified Z-score calculation is shown in Eqn. 1.

$$z(y_{p,s}) = \frac{y_{p,s} - \overline{y_p}}{k_z \hat{\sigma}(y_p)} \tag{1}$$

   where $\hat{\sigma}$ is the sample standard deviation. The coordinates of each point, in IMS internal units, is calculated as shown in Eqn. 2.

$$r_{p,s} = z(y_{p,s})S + \mathcal{O} \tag{2}$$

2. The distance component between a monitor point $r_{p,s}$ for parameter $p$, sample $s$ to cluster $c$ is as shown in Eqn. 3.

$$\Delta_{p,s,c} = \text{dist}(r_{p,s}, c) = \begin{cases} r_{p,s} - U_{p,c}, & r_{p,s} > U_{p,c} \\ r_{p,s} - L_{p,c}, & r_{p,s} < L_{p,c} \\ 0, & L_{p,c} < r_{p,s} < U_{p,c} \end{cases} \tag{3}$$

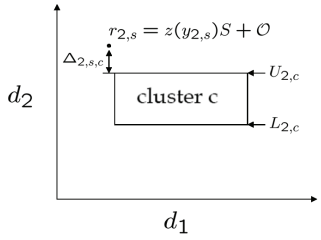



Fig. 1b. Distance from point $r_{2,s}$ to cluster c in dimension p=2

3. The squared Euclidean distance $\beta_{s,c}$ from each sample to each cluster is as shown in Eqn. 4.

$$\beta_{s,c} = \sum_p \Delta^2_{p,s,c} \tag{4}$$

4. The Euclidean distance $\gamma_s$ to the nearest cluster for each point is as shown in Eqn. 5.

$$\gamma_s = \sqrt{\min_c(\beta_{s,c})} \tag{5}$$

Eqn. 6 provides a formula for the nearest cluster, denoted as $\tilde{c}_s$.

$$\tilde{c}_s = \arg\min_c(\beta_{s,c}) \tag{6}$$

5. The composite score shown in Eqn. 7 is computed in a way which is intended to have a rough correspondance with a percentage of the size of the vector space. This correspondance is closer when the data was normalized over the range of the training data in each dimension. In the IMS implementation where the normalizing constant in each dimension is some multiple of the standard deviation for that dimension, the meaning is less clear.

$$\theta_s = (\frac{\gamma_s}{\sqrt{n_p S^2}})100 \tag{7}$$

The factor of $\sqrt{n_p}$ in the denominator of Eqn. 7 normalizes the IMS score across data sets with differing numbers of parameters.

6. The contributing channel score for an individual channel is computed as follows. Let $\tilde{d}(r_{p,s}, \tilde{c}_s) = \text{dist}(r_{p,s}, \tilde{c}_s)$. The contributing score for $\tilde{D}_{s,p}$ is shown in Eqn. 8.

$$\tilde{D}_{s,p} = (\frac{\tilde{d}_{p,s}}{S})100 \tag{8}$$

#### 3.1.4 IMS complexity

Figs. 2 and 3 illustrate empirically generated timing tests for IMS runs that demonstrate its computational complexity for datasets of increasing size. Fig. 2 illustrates only the time that it takes to train IMS, while Fig. 3 shows the time that it takes to normalize, train, test, and parse the results for IMS. In both figures, results are shown for data sets that have different numbers of parameters, which aids in gaining a sense of scalability as well as complexity. In Fig. 2, the size of the dataset containing 21 parameters was increased from a minimal size of 338 kB to its full size of 32 MB (blue dashed line). This data set was also noted to have been based upon resampled data, as distinct from "change only" data. This distinction is notable due to the fact that the resampled counterpart to a "change-only" data set is always of a much larger size, and a change only dataset is sufficiently representative for use as a training data set, both for obvious reasons. The resampled version is only necessary when performing monitoring and subsequent analyses.

In Fig. 2, both resampled and change only versions of a dataset containing 164 parameters are also shown. They are varied in size from a minimal ~ 6 MB to 218 MB (just over one-third the full size of the resampled dataset) and from 2.64 MB to 250 MB (the full size of the change only dataset), which are shown with solid blue dots and the solid blue line, respectively. The length of time spanned by the actual dataset (real-time) is shown in green,

and appears nearly as a vertical line due to the scale used for the graph, even though the slope is finite. As such, it is evident that for all timing tests shown in both Figs. 2 and 3 (in blue), IMS runs nearly in linear time and far faster than real-time, even when accounting for the time that it takes to normalize, train, test, and parse the results for IMS (In Fig. 3). It is also evident that the number of parameters in the dataset processed by IMS scales in a multiplicative fashion (as well as the size of the dataset, implicitly, due to aforementioned linear time characterization). However, for the purposes of performing a thorough analysis that involves optimization and an n-fold cross validation, even a linear time complexity needs to be considered carefully.

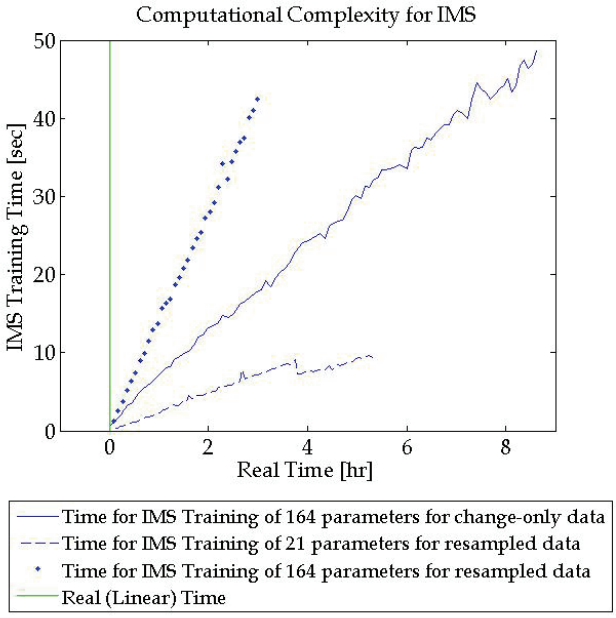



Fig. 2. IMS Training Complexity

To see how quickly time can add up, take as an example a 10-fold cross validation and optimization performed on 5-hr training and validation datasets with 21 parameters. If we base the total processing time for each point within a 150 point optimization grid on Fig. 3, it becomes clear that it would exceed ~ 30 sec x 10 cases X 150 optimization points (>12 hrs), in addition to the time for monitoring auxiliary fault data. If the validation dataset used to monitor auxiliary fault data consists of 2 nominal and 2 fault scenarios, and each scenario spans 5 hrs, then the time for monitoring fault data for 2 nominal and 2 fault scenarios exceeds ~ 4 sec X 10 cases X150 optimization points X 4 scenarios, based upon Fig. 2 (using ½ training time as a rough estimate for monitoring time). The total estimated time amounts to almost 20 hrs. of computational effort. And even so, these rough computations do not account for the fact that for file sizes above the max 250 MB used to create these plots, the number of clusters found by IMS may increase substantially, also adding significantly to the computational cost, possibly running in near real-time or even above real-time for very large files.

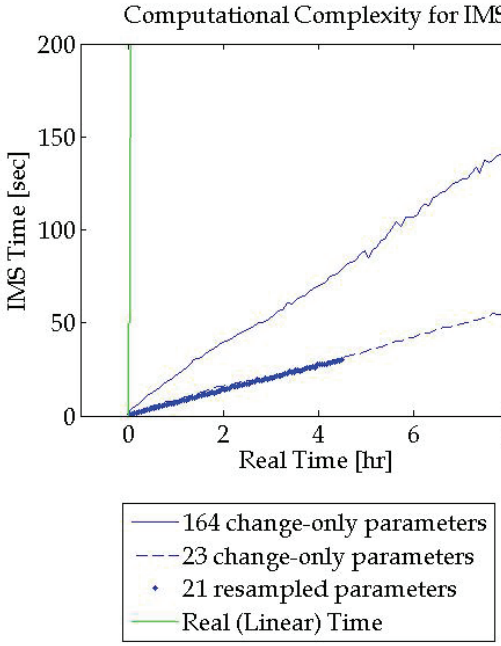



Fig. 3. IMS Normalization, Training, Testing, and Results Parsing Complexity

Evidence of this variation with number of clusters is shown in Fig. 4. Here, training /validation data file sizes are fixed for the entire experiment. The experiment implicitly varies the number of clusters with each subsequent training/validation run via adjustment of IMS's "maximum cluster radius" parameter. The x axis represents the number of sequential IMS training/validation set completions, the y axis on the left represents the date of completion, and the y axis on the right represents the number of clusters. As in Fig. 3, this involves IMS training, testing, and the overhead involved in wrapper code written for generation of statistics for optimization, which includes parsing of the results. The reason sequentially processing training/validation runs appears sublinear is due to the fact that the "maximum cluster radius" value is being adjusted from a very small value to 1. Recall that there is always an inverse relationship between the maximum cluster radius and the number of clusters. As such, we implicitly decrease the number of clusters per IMS knowledge base for each subsequent run from 976 to 18, causing the time for completion to be shortened decrementally.

For files of an increasing size which seeded the experimental results shown in Figs. 2 and 3, the effect seen in Fig. 4 may bias the attempt to control the experiment by holding the maximum cluster radius parameter fixed. When the file size is increased for the experiments yielding the results shown in Figs. 2 and 3, the number of clusters naturally increases due to the greater chance of visiting areas in the hyperspace that contain vectors previously unseen by the current IMS cluster population. As such, a similar sublinear effect as seen in Fig. 4 may result for an experiment in which the file size is decreased but the maximum cluster radius parameter is held fixed. Conversely, a limiting effect is apparent when the number of

clusters increases as a side effect of increasing the file size, implying that IMS will run in near linear time as file size increases (as is evident in Figs. 2 and 3). However, this limiting linear time effect does not imply that IMS will run in *real* time for an arbitrarily large number of clusters. In fact, as mentioned previously, it is possible for IMS to run above real-time for very large files, since we know that IMS scales in a multiplicative fashion with the size of the dataset. As such, an increase in the number of clusters resulting from an increase in file size will yield a substantially greater computational burden. In the previous example, rough estimates were provided that did not account for these effects.

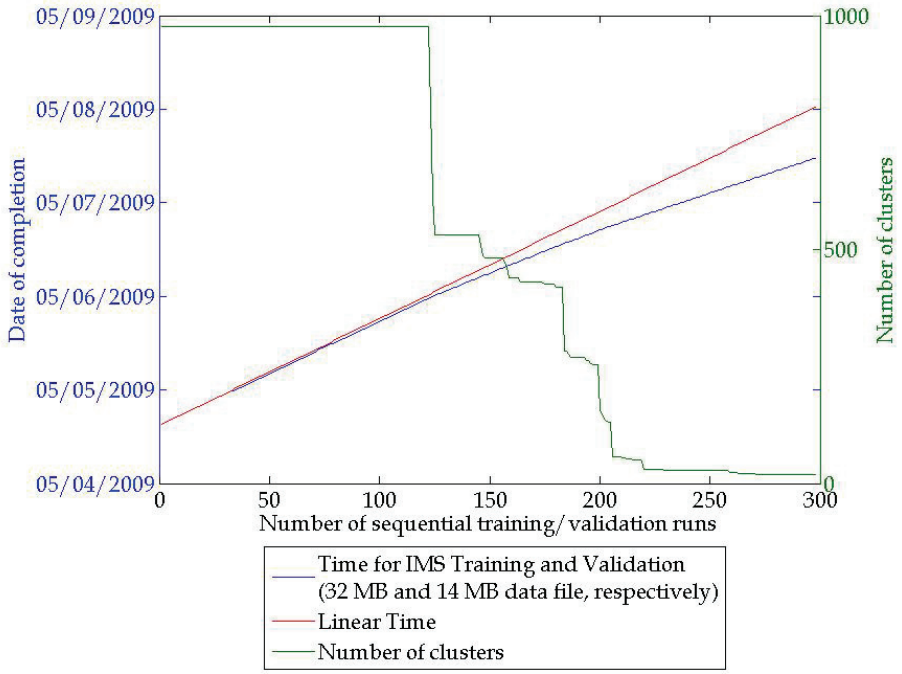



Fig. 4. IMS Complexity and corresponding number of clusters

### 3.2 Orca

Orca is a software tool that uses a nearest neighbor based approach to outlier detection which is based upon the Euclidean distance metric. It uses a modified pruning rule that allows for increased computational efficiency, running in near linear time as the number of "top score" outlier points selected for evaluation decreases. More information on this algorithm and some of its applications can be found in previous work (Bay and Schwabacher, 2003); (Schwabacher, 2005). This algorithm outputs a total score which represents the average distance to the nearest $k$ neighbors in the multi-dimensional feature space containing all of the variables. It also outputs the contribution of each variable to this score in order to show which variables cause each outlier to be classified as such.

Unlike IMS, Orca requires no training due the inherent comparative nature of the algorithm. Ostensibly, this should reduce the total computational burden. However, the greatest

computational advantages of using Orca are reaped only when applying a modified pruning rule for a reduced number of outlier points based upon having the highest score magnitude selected for evaluation. An implicit requirement to use all available points exists, to aid in the construction of an unbiased ROC curve representative of all score magnitudes and with full resolution. Therefore, Orca loses is computational advantage and defaults to a standard *k*-nn (nearest neighbor) algorithm that scales quadratically. A two-dimensional example is provided in Fig. 5 in order to illustrate the basic idea behind the *k*-nn algorithm. The red circle represents the point being monitored, and the average Euclidean distance to a fixed number of nearest neighbors, *k*, is used to compute a composite anomaly score. Isolation of the outlier point becomes possible by using this technique as the monitored point is indexed across the entire population, as is evident in the diagram.

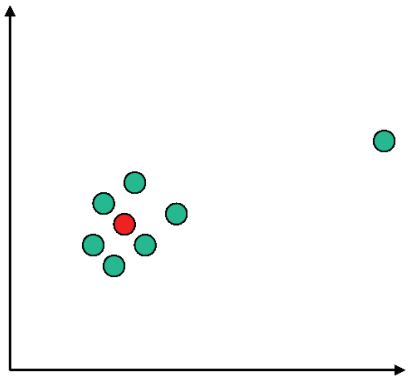


Fig. 5. Nearest Neighbor Approach to Anomaly Detection

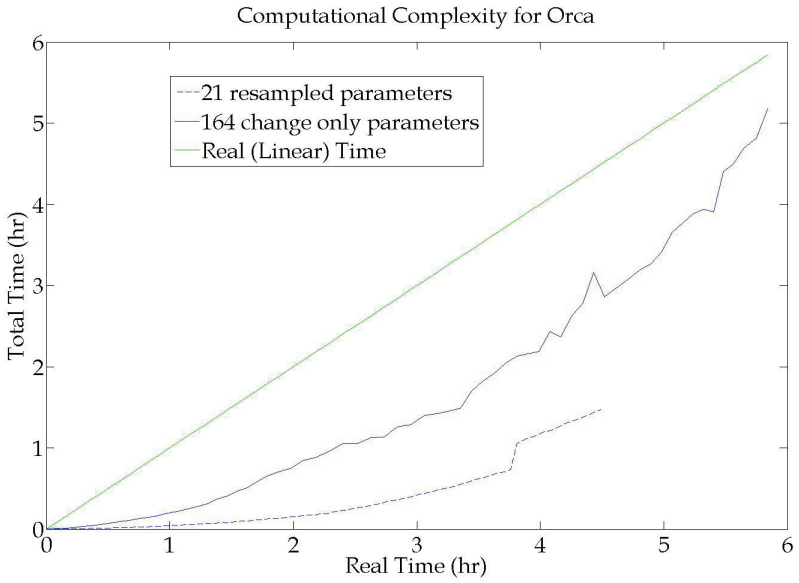



Fig. 6. Orca Normalization, Testing, and Results Parsing Complexity

Empirical evidence of quadratic complexity is shown in Fig. 6, which also illustrates the time required for normalization, application of Orca, and results parsing. This graph merits direct comparison to Fig. 3 showing complexity results for IMS, due to use of the same datasets and ranges in size of the datasets. Clearly, the computational complexity of Orca is far greater than that of using IMS. Furthermore, for datasets spanning a large period of time, it is readily apparent that the algorithm will eventually no longer be able to run faster than real time.

### 3.3 One-Class Support Vector Machine (SVM)

The one-class support vector machine is a very specific instance of a support vector machine which is geared for anomaly detection. The generic support vector machine (SVM) can be used to classify data in multiple dimensions by finding an appropriate decision boundary. Like neural networks, support vector machines perform classification using nonlinear boundaries, however they go one step beyond neural networks by finding the boundaries that provide the maximum margin between different classes of data. Additionally, using the support vector machine one can map data from a lower dimensional space that is not linearly separable to a higher (even infinite-dimensional) space where the data are linearly separable by a hyperplane. This is performed by using what is commonly known in machine learning as the "kernel trick," when using SVM's. A kernel function is chosen to map the data from the lower-dimensional space to the higher-dimensional space. It can be chosen arbitrarily so as to best suit the data and at the same time reduce the computational burden involved with generating the mapped values by direct evaluation. However, for our purposes, we choose the Gaussian radial basis function or kernel, given by Eqn. 9, which is the most widely used in the application of one-class SVM's.

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{2\sigma^2}} \tag{9}$$

$K$ is the kernel function, and the training data vectors are given by $\mathbf{x}_i$ and $\mathbf{x}_j$. The kernel width is given by $\sigma$, and is used to control the distribution of the kernel function around the training data point. The magnitude of this value also often has a direct effect on the algorithm's speed and complexity. "Support vectors" correspond to those data points that lie along the margin or closest to it. The maximum margin between classes is found by solving a quadratic optimization problem.

The one-class SVM differs from the generic version of the SVM in that the resulting quadratic optimization problem includes an allowance for a certain small predefined percentage of outliers, $\nu$, making it suitable for anomaly detection. As shown in Fig. 7, these outliers lie between the origin and the optimal separating hyperplane. All the remaining data fall on the opposite side of the optimal separating hyperplane, belonging to a single, nominal class, hence the terminology "one-class" SVM. The SVM outputs a score that represents the distance from the data point being tested to the optimal hyperplane. Positive values for the one-class SVM output represent normal behavior (with higher values representing greater normality) and negative values represent abnormal behavior (with lower values representing greater abnormality). For subsequent analyses and in the results section, the *negative* of this score will be used in order to establish a frame of reference commensurate with ROC analysis and for comparison to other algorithms. More technical details on the one-class SVM are available in previously published studies (Das *et al*., 2007); (Cohen *et al*., 2004).

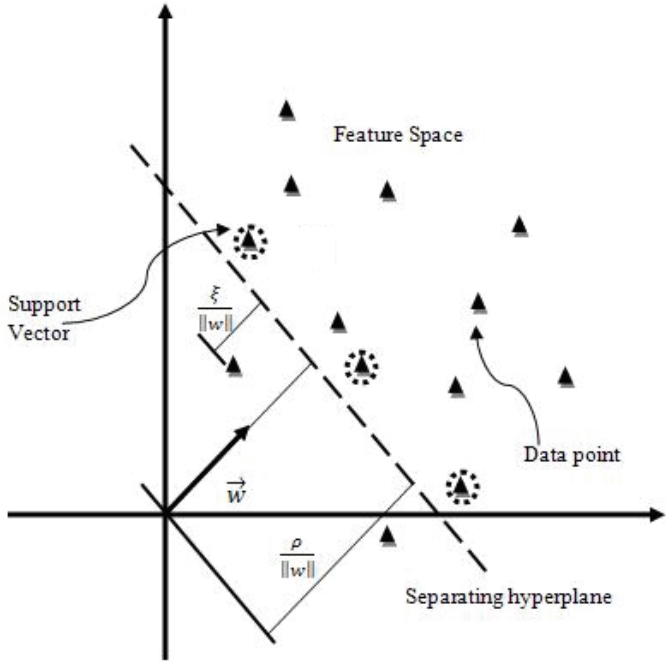



Fig. 7. Illustration of One-Class Support Vector Machine Framework

The one-class SVM differs from the other methods discussed in this paper because it determines whether or not a point is an outlier based on the distance of the point to a separating hyperplane in a feature space induced by a kernel operator, whereas most of the other methods rely on an analysis of the data in the original data space. For the one-class SVM, a single hyperplane separates the nominal data from the origin. Thus, for a system which undergoes nominal mode changes during its operation, all such changes will be characterized as nominal with a single hyperplane. Orca and IMS, on the other hand, characterize the anomalousness of a point based on local characteristics within the data space. This quality can make those algorithms more robust to significant mode changes compared with the one-class SVM.

The small predefined percentage of outliers, ν, is typically selected *a priori* and used to inform optimal selection of a value for the kernel width, σ, as is shown in Fig. 8. This is performed by finding the correct training classification rate as a function of the kernel width, and finding the value of σ that corresponds to 1-ν. It is often easy to find this value within a reasonable level of tolerance. In these cases, the kernel width, σ, corresponding to the first exceedance of 1-ν is used as the optimal value. However, when using a very small value of ν, it may be infeasible to find a training misclassification rate that will achieve this value (1-ν). This puts a lower bound on the predefined percentage of outliers that can be selected. As such, various candidates can be tested for feasibility in order to determine the lower bound on ν. For each of these candidates, a corresponding relevant metric can be computed, in addition to the corresponding optimal kernel width. In this case, the relevant test metric of interest is the AUC value. A value of ν that has been determined to be feasible,

and corresponds to the highest AUC value will thus be used for selection of the optimal kernel width.

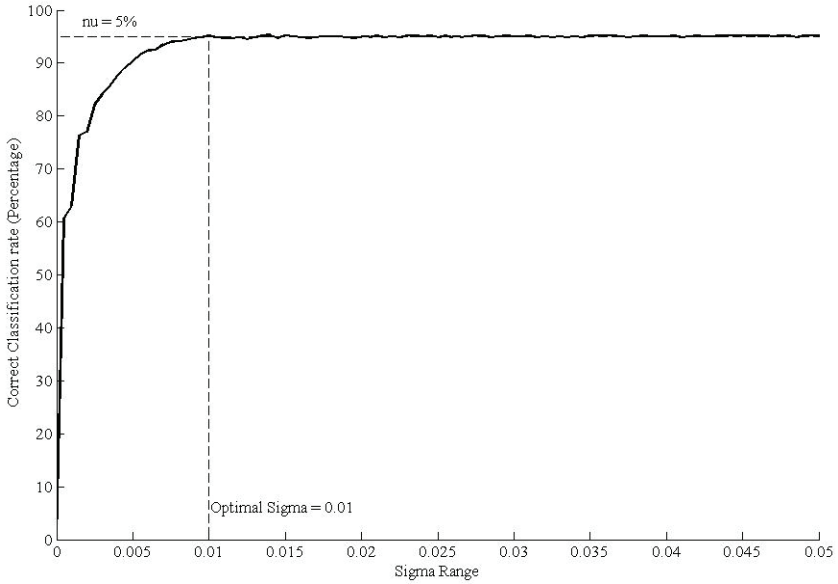



Fig. 8. Illustration of Selection of Optimal Kernel Width

The complexity of the one-class SVM algorithm is also of interest. As such, we provide an empirical complexity analysis shown in Fig. 9 to complement Figs. 2-4 and 6, which illustrate the complexity for IMS and Orca. Similar to Orca in Fig. 6, it is evident that the one-class support vector machine algorithm yields a superlinear complexity shown in Fig. 9. Furthermore, for the smaller dataset containing 21 resampled parameters, we compare the time for training only (shown as a dotted line), to the time required for normalization, training and validation (shown as a dash-dot line), which adds negligible time. It is also interesting to note that there is no noticeable multiplicative increase in complexity with the processing of additional parameters as opposed to Orca and IMS (cf. 21 resampled parameters to 164 change-only parameters), which is ostensibly due to the use of the kernel trick. However, regardless of this slight advantage, again the computational complexity of the one-class SVM algorithm is far greater than that of IMS, and on par with that of Orca. As seen in Fig. 9 , for datasets spanning a period of time greater than 7 hrs., it is readily apparent that the OCSVM algorithm is no longer able to run faster than real time.

## 4. Validation and evaluation criteria

In order to motivate a detailed discussion of the evaluation criteria, we will first need to introduce the means by which we will validate the algorithms discussed thus far. Specifically, we need to outline the type and volume of data required for validation. Recall

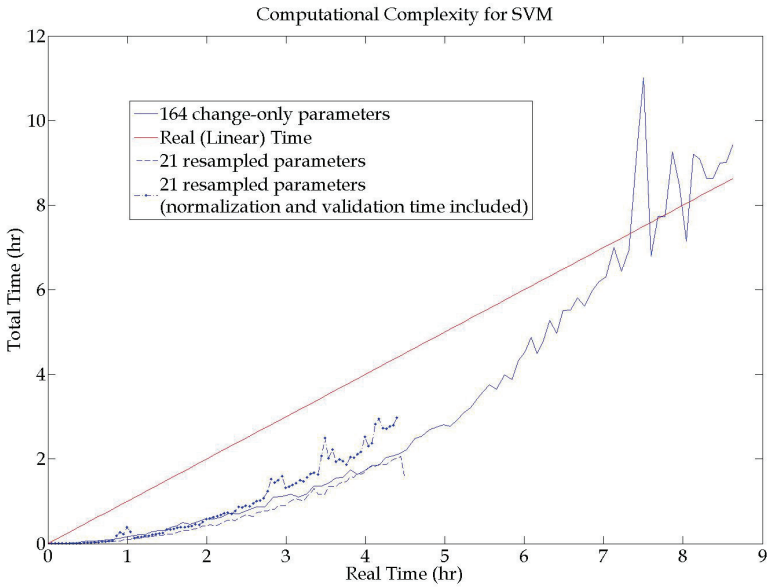



Fig. 9. One-Class Support Vector Machine Computational Complexity

the requirement to process a sufficient amount of data to undergo a formal cross-validation procedure that can be processed in a reasonable amount of time. The computational complexity of each data-driven algorithm is now well established, which facilitates the selection of various candidate cross-validation experiments. Again, for the purposes of our analysis here we will only consider two primary phases of operation of the TVC subsystem based upon meeting these specified data requirements. The two phases of interest involve both testing of the TVC in the VAB, and when the fully assembled vehicle is at the pad after launch countdown has commenced. For each of these phases, there are specific periods of time in which respective tests of interest or important elements of the operational sequence are conducted. Certain pre-specified failure modes from a FMECA (Failure Modes, Effects and Criticality Analysis) have been identified as potential adverse events that could occur during these critical periods of time.

Table 2 provides a summary of each failure mode under consideration for each phase of operation. The periods of time spanned by tests corresponding to each of these failure modes are also provided, which aid important considerations of computational complexity.

For the last three row entries in Table 2, the period of time spanned by the tests are largely dependent upon the specific flight, and are also driven by acquisition of a sufficient amount of data to accurately characterize nominal behavior in order to prevent false alarms. In light of this fact, the use of a 25 Hz sampling rate, and the example of how quickly time complexity for even the least burdensome algorithm adds up, we will use the first two row entries in Table 2 as the basis of experiments and results to be subsequently presented. Another advantage of using the two failure scenarios at the pad is due to the fact that they occur during the same time period within T – 1 min prior to launch, thus simplifying the experimental construction and allowing for fewer scenarios to investigate independently.

| Operational Phase | Failure Mode | Period of Time Spanned |
|---|---|---|
| Pad | FSM (Fuel Supply Module) Pressure Drop due to $N_2H_4$ (Hydrazine) Leak | Within T – 1 min prior to launch |
| Pad | Hydraulic Fluid Reservoir Level Drop due to hydraulic fluid leak | Within T – 1 min prior to launch |
| VAB | Actuator Stuck during Actuator positioning test | 2.5 min test in VAB |
| VAB | FSM (Fuel Supply Module) Pressure Drop due to $N_2H_4$ (Hydrazine) Leak | Within ≥34 minute period after calibration test in VAB |
| VAB | Hydraulic Fluid Reservoir Level Drop due to hydraulic fluid leak | Within ≥10 min period during TVC actuator tests in VAB |
| VAB | Hydraulic pumping unit overtemperature failure | Within ≥25 min period during tests in VAB |

Table 2. Failure Mode Summary

Due to the absence of real data to use for validation and testing for the two candidate failure modes under consideration in Table 2, they must be simulated using existing means. Increasing the range of fidelity of available failure simulations is an ongoing process. In previous work, (Schwabacher *et al.*, 2009) a high fidelity physics-based simulation of a leak in a liquid propulsion-based J-2X engine was used to validate data-driven algorithms. However, such a simulator does not currently exist for relevant components of the TVC subsystem. In future work, we plan to develop high fidelity physics-based simulations of a leak in the fuel supply module (FSM) of the TVC subsystem, which is a spherical tank containing liquid hydrazine ($N_2H_4$) pressurized by $GN_2$ (gaseous nitrogen). However, for our purposes here we have chosen to use a rather low fidelity simulation for the two scenarios. The simple use of linearly decreasing ramps from nominal operating conditions to off-nominal values for parameters specific to each failure mode will be used, given predefined rates of degradation and preselected off-nominal values.

The time of failure injection within the T- 1 min period prior to launch is selected to leave a significant nominal fraction of the period at from T -1 min until injection of the failure, which occurs at T-28.12 sec for the hydrazine leak at the pad, levelling off at its off-nominal value at T-15 sec. For the hydraulic fluid reservoir leak at the pad, the failure is injected at T-13 sec and levels off at its nominal value at T-12 sec. However, it is often useful to supplement our validation data with additional cases of purely nominal behavior as well. In our case, we will use a concatenation of nominal data from T-1 min to launch, and two additional failure datasets representing the scenarios described above. Now that we have established the basis for constructing the validation data sets, it is of interest to describe how the algorithms will be optimized and evaluated with the use of ROC curve analysis.

The ROC curve essentially plots the true positive rate against the false alarm rate for all possible threshold values, as shown in Fig. 10. It therefore can be used as a design tool in order to select an alert threshold according to pre-established requirements for minimum

missed detection and/or false alarm rates. However, in order to compute true positive and false alarm rates, it is necessary to obtain a "ground truth" representation for each monitored example. In this case, an "example" can represent an individual validation flight, or a single time point. Here we will compute false alarm and missed detection rates based upon the classification of each individual time point, rather than on a flight-by-flight basis, as was performed in previous studies (Martin *et al.*, 2007).

In order to compute true positive and false alarm rates with a reasonable level of accuracy, there is a need to obtain a statistically significant number of labeled examples, both nominal and anomalous. The availability of nominally classified examples often far exceeds the availability of anomalously classified examples. The deficit of the latter drove our efforts to simulate faults, and hence to use classified time points as monitored examples in lieu of individual validation flights for construction of the ROC curve. For our purposes here, the classified time points represent the "ground truth," and will be based upon the times of failure injection specified previously, for each failure scenario.

One advantage of using ROC curve analysis is in its inherent robustness against the use of skewed distributions between the populations of nominally and anomalously categorized examples. Thus, as long as the population of failure examples is statistically significant, it does not need to be on par with the population of nominal examples. Furthermore, unlike other alert threshold selection techniques, the ROC curve may also be used for optimized selection of algorithmic parameters. This can be performed by using the area under the ROC curve, which represents overall classification discriminability. Therefore, the AUC can be maximized with respect to free design parameters used for algorithmic tuning to optimize and control performance during the validation stage. As such, we can ensure that the resulting model used to perform threshold or alert selection has the best anomaly detection capability possible. Furthermore, the AUC analysis can be used to compare the performance of all candidate algorithms.

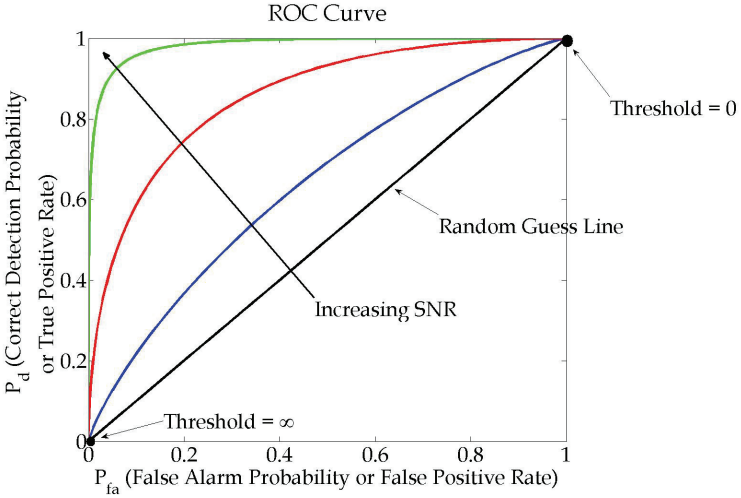



Fig. 10. Sample ROC Curve

As seen in Fig. 10, the area under the ROC curve has a classic increase in relation to the SNR (Signal-to-Noise ratio). This relationship is well established, and is derived from the origins

of the ROC curve for use in radar applications dating back to WWII. Unlike the SNR, it is often the case that algorithmic design parameters do not have a similar straightforward relationship to the AUC. However, this relationship is implicit in Fig. 10 due to the fact that optimization of algorithmic design parameters is performed by using the area under the ROC curve. In a sense, the ROC curve is "tuned" by attempting to maximize the AUC (area under the ROC curve) and choosing the appropriate algorithmic design parameter(s) to allow for maximum predictive capability. This can be thought of as choosing the ROC curve with the highest SNR based upon such parameters. In this case theoretically the "signal" might loosely be thought of as the *variability* of anomalous behavior, and the "noise" might be thought of as the *variability* of nominal operations.

## 5. Results and discussion

As previously discussed, a formal cross-validation procedure will be used to perform optimization and alert threshold selection. Doing so will help to prevent overfitting, as is often the case when using a single training and validation dataset. A *k*-fold cross-validation will be used, in which *k*-1 training examples are used, and the remaining example is used for validation. Each of the *k* examples is rotated and used one time for use as a validation set, so that there are *k* sets of statistics rather than a single one. The candidate training/validation dataset partitions will span 7 flights (STS-117, STS-107, STS-112, STS-113, STS-114, STS-120, and STS-122) containing 30 continuous-valued parameters, and a 7-fold cross-validation procedure will be used. Resulting AUC and ROC curves will be averaged over the results of the 7-fold cross-validation. The optimization parameters used and the respective optimized values derived from the average AUC across all 7 folds for each of the candidate data-driven algorithms are summarized in Table 3.

| Algorithm | Optimization Parameter | Optimized Value | Max. of averaged AUC Value Achieved | "Hyperparameter" Values |
|---|---|---|---|---|
| IMS | Maximum cluster radius (MAX_INTERP) | .01894 (403 clusters) | .9807 | Initial cluster size (INIT_TOLERANCE) = 0.01 Cluster growth percent (EXTRAP_PERCENT) = 0.05 |
| Orca | k (number of nearest neighbors) | 2 | 0.91077 | N/A |
| OCSVM | σ (kernel width) | 10.8081 | 0.92956 | ν = 0.2289 |
| OCSVM | σ (kernel width) | 2.121 | 0.93072 | ν = 0.2 |
| OCSVM | σ (kernel width) | 10.8081 | 0.90416 | ν = 0.1 |
| OCSVM | σ (kernel width) | 2.121 | 0.9051 | ν = 0.08 |
| OCSVM | σ (kernel width) | 0.74531 | 0.94666 | ν = 0.075 |
| OCSVM | σ (kernel width) | 2.121 | 0.90664 | ν = 0.07 |
| OCSVM | σ (kernel width) | 5.9636 | 0.90529 | ν = 0.05 |
| OCSVM | σ (kernel width) | 5.9636 | 0.90525 | ν = 0.01 |
| OCSVM | σ (kernel width) | 3.5565 | 0.89681 | ν = 0.001 |
| OCSVM | σ (kernel width) | infeasible | infeasible | ν = 0.0001 |

Table 3. Optimization Parameters

Additionally, the optimized maximum AUC values achieved by averaging across all 7 folds, and values of any supplementary "hyperparameters" that were used for each algorithm are provided in Table 3. Fig. 11 illustrates the average AUC values across all 7 folds for each of the candidate data-driven algorithms as a function of each of their design parameters. In addition, quasi-confidence bounds representing the maximum and minimum AUC values across all 7 folds are provided. In the lower right panel of Fig. 11, the maximum of the averaged AUC value achieved that lies in a feasible range is demarcated, even though it is clearly not the maximum over the entire range of kernel width values shown. The range of feasible kernel width values for the row in Table 3 highlighted in red represents any value greater than the optimized value shown, and corresponds to the lower right panel in Fig. 11. Both the highlighted row and the lower right panel of Table 3 and Fig. 11 respectively represent the maximum achievable AUC averaged across all 7 folds for the OCSVM algorithm. Thus, the optimized kernel width lies on the very border of the feasible range. Note that the last row of Table 3 indicates that a lower bound on the predefined percentage of outliers, $\nu_{min}$, has been identified for any kernel width.

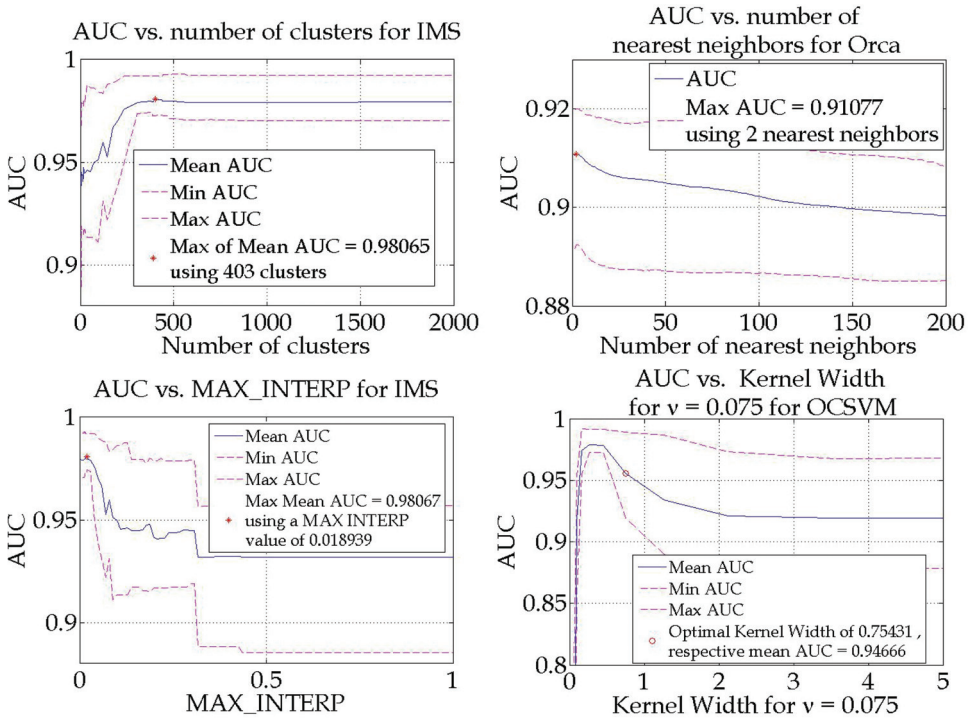



Fig. 11. AUC optimization with confidence intervals

Conflating the values presented in Table 3 with the graphs shown in Fig. 11, it becomes clear that IMS outperforms the other two algorithms, even when the hyperparameters for IMS are held fixed and not implicitly used to inform the optimization problem, as was performed with the one-class SVM. However, what may not be as clear from Fig. 11 is that some overlap between the confidence intervals among the algorithms exists. Specifically, for IMS,

the averaged optimized AUC value achieved falls within the range (0.9728, 0.9921), for Orca the range for optimized AUC values is (0.8924, 0.9201), and for OCSVM the range for optimized AUC values is (0.919, 0.9891). Therefore, there is some overlap between the optimum ranges for IMS and OCSVM, albeit a very small overlap of 0.0163 in which SVM might outperform IMS. Orca has no range overlap with IMS, but has an extremely small overlap with SVM of 0.0011.

Our next and final step is to perform threshold or alert selection based upon the ROC curves corresponding to the optimized parameters found from the previous step. Recall that this allows for alert threshold selection based upon specified minimum false alarm and/or missed detection rates with a robust anomaly detection capability and the best classification discriminability. Fig. 12 shows examples of the optimized ROC curves averaged over all 7 folds and their minimum/maximum confidence interval counterparts. Additionally, corresponding alert thresholds have been selected according to an established maximum allowable false positive rate of 0.01, and the results are shown in the legend.

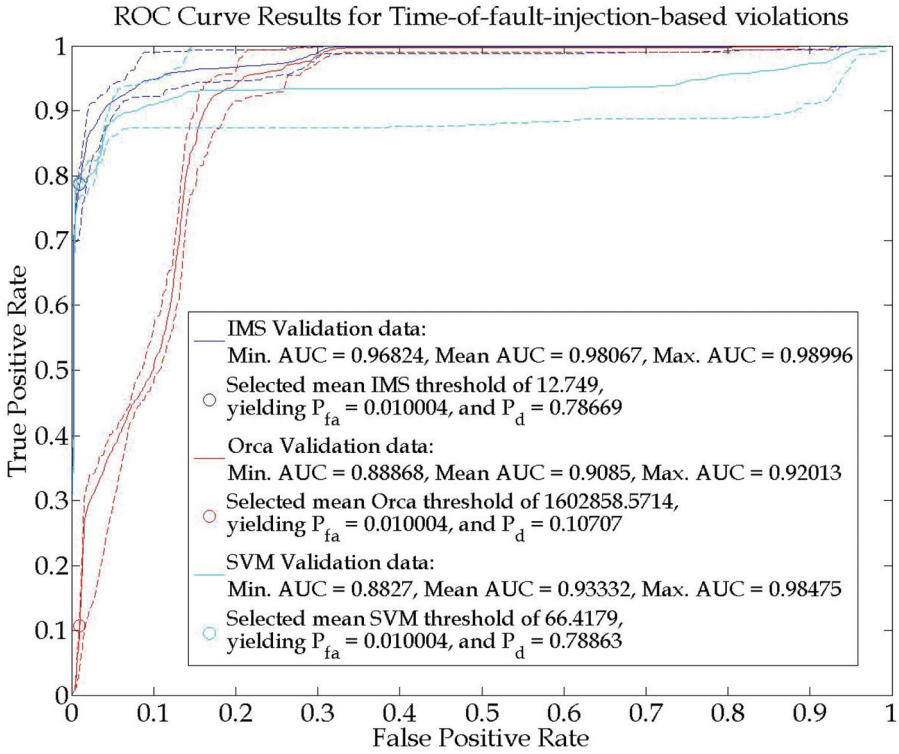



Fig. 12. ROC Curve Comparison and Alert Threshold Selection

The slight variation in ranges shown in the legend differ from the optimized AUC ranges previously provided due to the revised manner in which they have been computed. The minimum and maximum ROC curves were constructed in a modified manner in order to prevent any excursions of the mean ROC curve outside of the min/max envelope for visualization purposes. As such, the mean false positive rate is used consistently to construct

ROC curves representing the min/max bounds, while the min/max true positive rate is used respectively, resulting in different min/max AUC values than previously reported, however these variations are negligible. Note that for a fixed false alarm rate ceiling ($P_{fa}$) of 0.01, the resulting OCSVM threshold yields a true positive rate ($P_d$) on par with that of IMS (mean 0.79), and the resulting Orca threshold yields a much lower a true positive rate ($P_d$) having a mean of 0.11. In due fairness to Orca, it is possible that the AUC value may have increased beyond the range of the 200 nearest neighbors shown on the upper right panel of Fig. 11. However, due to scalability and complexity issues, the experiment was terminated as shown. Indeed, the complexity of running Orca with an increased number of nearest neighbors scales quadratically when not using any pruning rules.

Finally, Fig. 13 illustrates realizations of validation flight STS-117 to which optimized parameters and thresholds are applied.

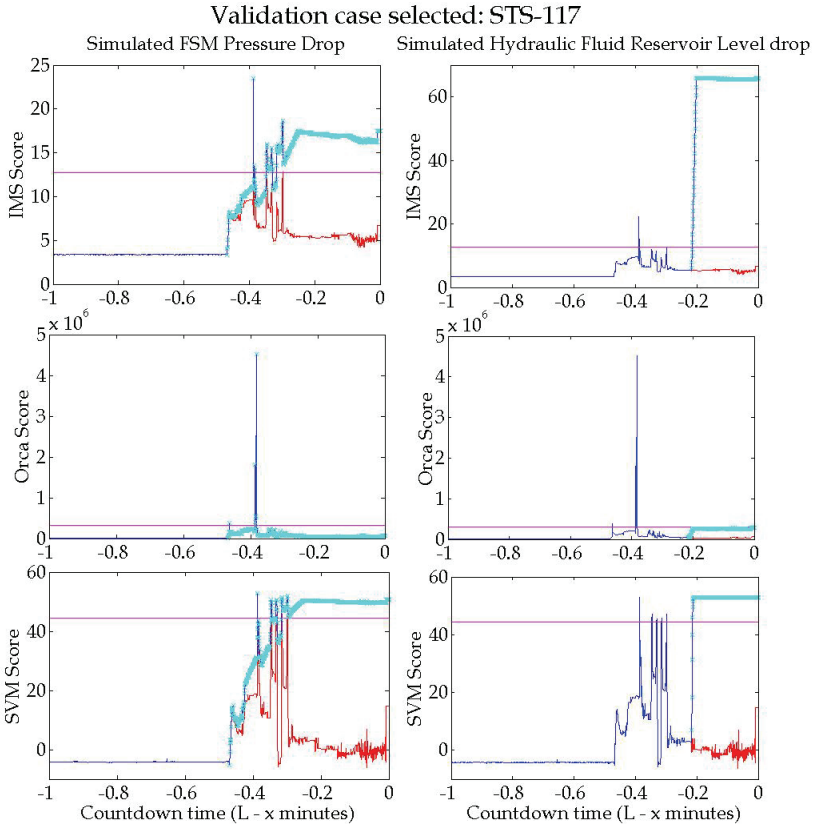



Fig. 13. Realizations of Optimized Parameters and Thresholds Applied to Validation Flight

Fig. 13 illustrates both the nominal (red lines) and fault injected scores (blue lines) from T – 1 min to launch at the pad for all algorithms. In all cases, both failure scenarios involving an FSM pressure drop (left panels) and a hydraulic fluid reservoir leak (right panels) are shown, using the optimized thresholds (fixed magenta colored lines) found from the ROC analysis. The actual threshold values may vary slightly from the values specified in Fig. 12

due to realizations being based on independent experiments, and as such should still fall within the confidence bands formed in both Figs. 11 and 12. The light blue highlights superimposed over the failure injected scores in dark blue represent the "ground truth" time of failure injection and duration, so as to give a feel for the false alarm and correct detection rates. Evidently, there is a clear bifurcation between nominal and anomalous scores for both IMS and SVM, and for Orca the same is true although it is less apparent. As can be discerned from Figs. 11 – 13, we have identified the fact that both in complexity and accuracy, IMS seems to be the best choice among all of the algorithms investigated. However, there is some overlap in the confidence intervals for IMS and SVM AUC values, and the alert thresholds applied for both corresponding ROC curves yield almost identical true positive rates.

## 6. Conclusion and next steps

We have provided a thorough end-to-end description of the process for evaluation of three different data-driven algorithms for anomaly detection. Through optimization of algorithmic parameters using the AUC, we were able to choose parameters yielding the best detection capability. The respective ROC curves corresponding to these parameters were then used to inform alert threshold selection by enforcement of a maximum allowable false alarm rate. It was found that IMS was the best performing algorithm when considering both computational complexity and accuracy. However, when evaluating the results based upon accuracy alone, the OCVSM approach is competitive with IMS due to overlapping confidence intervals present in the accuracy results.

In subsequent research studies, we will provide results of unseen hold out test cases to which optimized parameters and thresholds will be applied, in order to provide additional evidence demonstrating the superiority of a particular algorithmic technique. Furthermore, we will employ a variant of the AUC that only considers performance evaluation for algorithmic comparison restricted to low false positive rates. A slightly modified definition of false alarms and missed detections that accounts for pre-defined latencies and prediction horizons will also be investigated.

## 7. Acknowledgements

The author would like to acknowledge the support of the Ares I-X Ground Diagnostics Prototype activity, which was funded by the NASA Constellation Program, by the NASA Exploration Technology Development Program, and by the NASA Kennedy Space Center Ground Operations Project. Furthermore, the author graciously acknowledges the reviews of Dr. Mark Schwabacher, Bryan Matthews, and Dr. Ann Patterson-Hine. The author also extends appreciation to John Wallerius for his contribution of the subsection on IMS score computations, and his ideas pertaining to next steps on consideration of performance evaluation for algorithmic comparison restricted to low false positive rates. Finally, the author acknowledges the permission to use of Figs. 7 and 8 from Dr. Santanu Das, and Fig. 1a with supporting text from David Iverson.

## 8. References

Bay, S.D. & Schwabacher, M. (2003). Mining distance-based outliers in near linear time with randomization and a simple pruning rule, *Proceedings of The Ninth ACM SIGKDD*

*International Conference on Knowledge Discovery and Data Mining*, pp. 29–38, New York, NY, 2003.

Cavanaugh, K. (2001). An integrated diagnostics virtual test bench for life cycle support, *Proceedings of the IEEE Aerospace Conference*, pp. 7–3235–7–3246, ISBN: 0-7803-6599-2, Big Sky, Montana, March 2001.

Cohen, G.; Hilario, M. & Pellegrini, C. (2004). One-class support vector machines with a conformal kernel: a case study in handling class imbalance, In: *Structural, Syntactic, and Statistical Pattern Recognition*, A. Fred *et al.* (Eds.), pp. 850–858, Springer-Verlag, Berlin, Heidelberg, Germany.

Das, S.; Srivastava, A. & Chattopadhyah, A. (2007). Classification of Damage Signatures in Composite Plates using One-Class SVM's, *Proceedings of the IEEE Aerospace Conference*, Big Sky, MO, March 2007.

Fragola, J.R. (1996). Space shuttle program risk management, *Proceedings of the International Symposium on Product Quality and Integrity: Reliability and Maintainability Symposium*, ISBN: 0-7803-3112-5, pp. 133–142, Jan 1996.

Hart, G.F. (1990). Launch commit criteria performance trending analysis, *Annual Proceedings of the Reliability and Maintainability Symposium*, pp. 36–41, Jan 1990.

Iverson, D.L.; Martin, R.; Schwabacher, M.; Spirkovska, L.; Taylor, W.; Mackey, R. & Castle. J.P. (2009). General purpose data-driven system monitoring for space operations, *Proceedings of the AIAA Infotech@Aerospace Conference*, Seattle, Washington, April 2009.

Mackey, R.; James, M.; Park, H. & Zak. M. (2001). BEAM: Technology for autonomous self-analysis, *Proceedings of the IEEE Aerospace Conference*, Big Sky, MT, 2001.

Martin, R. (2007). Unsupervised anomaly detection and diagnosis for liquid rocket engine propulsion, *Proceedings of the IEEE Aerospace Conference*, Big Sky, MT, March 2007.

Martin, R.; Schwabacher, M.; Oza, N. & Srivastava, A. (2007). Comparison of unsupervised anomaly detection methods for systems health management using Space Shuttle main engine data, *Proceedings of the 54th Joint Army-Navy-NASA-Air Force Propulsion Meeting*, Denver, CO, May 2007.

Paté-Cornell, E. & Dillon, R. (2001). Probabilistic risk analysis for the NASA space shuttle: a brief history and current work. *Reliability Engineering & System Safety*, Vol. 74, No. 3, (Dec. 2001) pp. 345 – 352.

Paté-Cornell, E. & Fischbeck, P.S. (1994). Risk Management for the Tiles of the Space Shuttle. *Interfaces*, Vol. 24, No. 1, (Jan. – Feb. 1994) pp. 64–86.

Schwabacher, M. (2005). Machine learning for rocket propulsion health monitoring, *Proceedings of the SAE World Aerospace Congress*, pp. 1192–1197, Dallas, Texas, 2005.

Schwabacher, M. & Waterman, R. (2008). Pre-launch diagnostics for launch vehicles, *Proceedings of the IEEE Aerospace Conference*, Big Sky, MT, March 2008.

Schwabacher, M; Aguilar, R & Figueroa, F. Using Decision Trees to Detect and Isolate Simulated Leaks in the J-2X Rocket Engine, *Proceedings of the IEEE Aerospace Conference*, Big Sky, MT, 2009.

Tu, H.; Allanach J.; Singh S.; Pattipati K.R. & Willett, P. (2006). Information integration via hierarchical and hybrid Bayesian networks. *IEEE Transactions on Systems, Man and Cybernetics, Part A*, Vol. 36, No. 1 (Jan. 2006) pp.19–33.

Tumer, I. (2005). Design methods and practices for fault prevention and management in spacecraft, *Technical report, NASA Ames Research Center*, 2005.

# 9

# Design Solutions for Modular Satellite Architectures

Leonardo M. Reyneri, Claudio Sansoè, Claudio Passerone,
Stefano Speretta, Maurizio Tranchero, Marco Borri, and Dante Del Corso
*Politecnico di Torino*
*ITALY*

## 1. Introduction

The cost-effective access to space envisaged by ESA would open a wide range of new opportunities and markets, but is still many years ahead. There is still a lack of devices, circuits, systems which make possible to develop satellites, ground stations and related services at costs compatible with the budget of academic institutions and small and medium enterprises (SMEs). As soon as the development time and cost of small satellites will fall below a certain threshold (e.g. 100,000 to 500,000 €), appropriate business models will likely develop to ensure a cost-effective and pervasive access to space, and related infrastructures and services.

These considerations spurred the activity described in this paper, which is aimed at:

1. proving the feasibility of low-cost satellites using COTS (Commercial Off The Shelf) devices. This is a new trend in the space industry, which is not yet fully exploited due to the belief that COTS devices are not reliable enough for this kind of applications;
2. developing a flight model of a flexible and reliable nano-satellite with less than 25,000€;
3. training students in the field of avionics space systems: the design here described is developed by a team including undergraduate students working towards their graduation work. The educational aspects include the development of specific new university courses;
4. developing expertise in the field of low-cost avionic systems, both internally (university staff) and externally (graduated students will bring their expertise in their future work activity);
5. gather and cluster expertise and resources available inside the university around a common high-tech project;
6. creating a working group composed of both University and SMEs devoted to the application of commercially available technology to space environment.

The first step in this direction was the development of a small low cost nano-satellite, started in the year 2004: the name of this project was PiCPoT (Piccolo Cubo del Politecnico di Torino, Small Cube of Politecnico di Torino). The project was carried out by some departments of the Politecnico, in particular Electronics and Aerospace. The main goal of the project was to evaluate the feasibility of using COTS components in a space project in order to greatly reduce costs; the design exploited internal subsystems modularity to allow reuse and further cost reduction for future missions.

Starting from the PiCPoT experience, in 2006 we began a new project called ARaMiS (Speretta et al., 2007) which is the Italian acronym for Modular Architecture for Satellites. This work describes how the architecture of the ARaMiS satellite has been obtained from the lesson learned from our former experience. Moreover we describe satellite operations, giving some details of the major subsystems. This work is composed of two parts. The first one describes the design methodology, solutions and techniques that we used to develop the PiCPoT satellite; it gives an overview of its operations, with some details of the major subsystems. Details on the specifications can also be found in (Del Corso et al., 2007; Passerone et al, 2008). The second part, indeed exploits the experience achieved during the PiCPoT development and describes a proposal for a low-cost modular architecture for satellites.

## 2. The PiCPoT satellite

The PiCPoT design activity carried out at Dept. of Electronics, in tight cooperation with the Dept. of Aerospace Engineering and other departments of Politecnico, was aimed at developing and manufacturing a low-cost prototype of a fully operational nano-satellite. The design activity started in early 2004 and gathered about 10 people among professors and Ph.D. students, plus about 20 undergraduate students (the former for the whole Ph.D. program duration, the latter for shorter period, between 6 and 12 months each). The total effort of the project can be estimated as about 12 man-years (staff + student) for design, manufacturing and testing; a flight model and two engineering models of the PiCPoT satellite, shown in Figure 1, have been built.

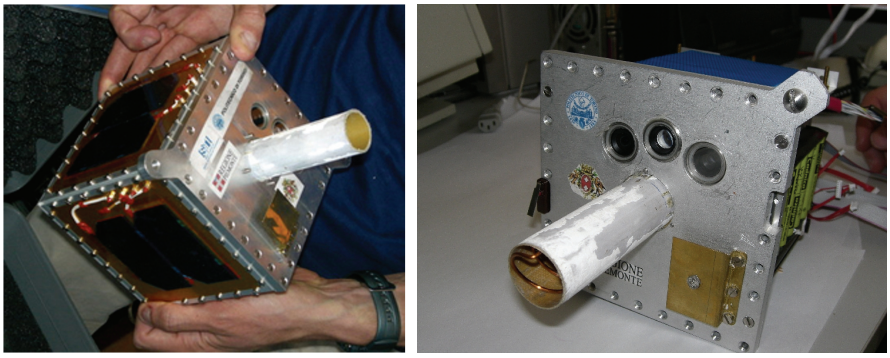


Fig. 1. PiCPoT engineering model.

The satellite has been completely designed using COTS devices, with the exception of solar panels. The basic architecture consist of five solar panels; six battery packs; three cameras with different focal lengths; five processors in full redundancy; two RX-TX communication modules with antennas operating at 437 MHz and 2.4 GHz, respectively. The on board electronics uses six PCBs hosted in a cubic aluminum case (developed by Dept. of Aerospace Engineering), 13 cm in side and 2.5 kg total mass. The main mission was to send telemetry data (temperatures, voltages and currents) to ground, and to take, store and transmit pictures of the Earth at different spatial resolutions.
The satellite was launched on July 26th 2006, from Baykonour. Unfortunately a failure of the launcher forced its destruction before being released in the planned orbit.

## 3. Design constraints

An airborne satellite must comply with hard constraints related to the severe space environment and the inability to repair the system in case of failure. Therefore, the design and the assembly of the device must abide by tighter rules than usual *good and safe design* criteria applied for any electronic system. This is particularly true when using COTS components and technology, which require the adoption of design techniques which guarantee system operation even in the presence of limited faults at the device level.
Other specific characteristics of a space application, although not directly related to failures of the system, further constrain the possible design solutions that can be adopted. These include the need to autonomously produce power, the limited visibility of the satellite from a ground station and the distance from it, the length of the mission, and so on.
In the following, the constraints and their implications that were considered in the design of PiCPoT, along with some solutions and ideas, are outlined.

### 3.1 Radiation

The planned orbit is close to the Van Allen belts, where a limited amount of radiation is present. This radiation might be in the form of high energy particles (protons, neutrons, alpha and beta particles) or ionizing electromagnetic rays from ultraviolet to X-rays. Due to the low orbit (polar, at 600km of altitude), and to the short lifetime assumed for the mission (3 months), total dose effects have not been considered. However, single-event effects (SEE) such as latch-up occurring in CMOS devices, and state upsets in memories and/or registers of digital circuits, might indeed induce wrong behaviors or even permanent faults. Thus the satellite circuits have been protected at the logical and system level against these events. Techniques that have been used include latch-up protection circuits, watchdog timers and redundancy at various levels. More details can be found in Sec. 4.

### 3.2 Electro-magnetic interference and signal integrity

Noise at various frequencies may come from both internal and external sources. However, the satellite outer structure is completely metallic, and all inner circuits are therefore well shielded against electro-magnetic interference (EMI) from the outside. Internal interference between different boards or within a single board is addressed by properly designing ground planes and the printed circuit board (PCB) layout of RF and digital units.

### 3.3 Temperature ranges

While it cycles through its orbit, the satellite alternates from broad daylight to deep Earth shadow. In these conditions, temperature may vary considerably. However, the orbital period is fast enough not to allow too much heat to build up or be released into space, preventing burning or frosting of the satellite. Thermal simulations allowed us to predict the actual temperature ranges for the outside and the inside faces of the aluminium plates that constitute the external structure of the satellite, and for the internal electronic boards. We considered the cases when the electronic boards are inactive, as well as when they are active and dissipating power (Caldera et al, 2005). The predicted outside temperature range with active electronics is [5, +50] °C; the parts subject to this range are external ones, such as solar panels and antennas. The temperature range inside the satellite is [+20, +70] °C, as shown in Figure 2, where the different curves represent the temperature of each board; all electronic circuits must comply with this range, which is compatible with standard commercial devices.

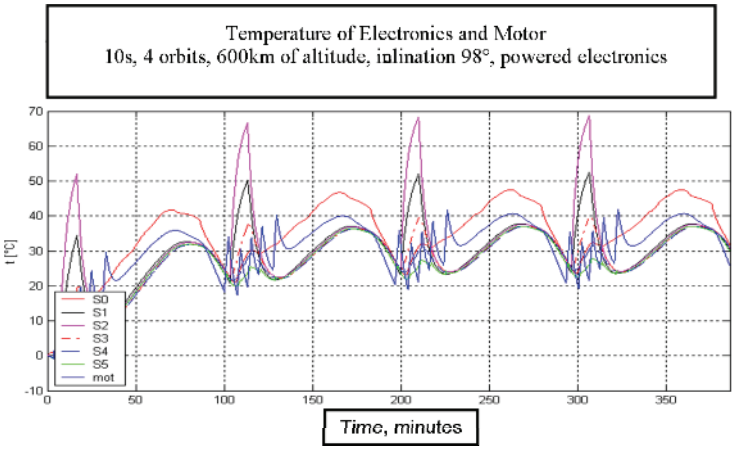



Fig. 2. Thermal analysis for powered electronic boards in the satellite

### 3.4 Vacuum

Vacuum is not a problem for sealed electronic components, but reduces the power dissipation capability due to missing convection, leaving only conduction and radiation to the outside. This problem is related to the temperature ranges outlined above. The board that dissipates more heat is the one responsible of data transmission, as it hosts the power amplifiers; we successfully tested it in a thermal vacuum chamber, with a temperature range of [-20, +50] °C and a pressure of 10Pa. While the expected pressure at the orbit altitude is some order of magnitude lower, we considered the level that we could achieve with in-house equipment sufficient to assess the board reliability. Other boards were simulated using their nominal characteristics, taking into account de-rating because of the absence of convection.

### 3.5 Vibrations

Forces and vibrations applied to the satellite during the launch are very high, and might cause physical damages, as well as disconnection of electronic devices and disengage of electrical connectors. A careful choice of packages (i.e., no BGA devices, more sensitive to vibrations), mounting technologies and overall structure is therefore mandatory.

PCBs (see Figure 3 for an example) have small size (about 12 × 8 cm$^2$), and are mechanically blocked at the four edges, therefore vibrations are kept within acceptable limits. More bulky components are secured to PCB, but connectors represent a critical point. Direct board-to-board connectors are kept in place by the mechanical fixture of boards. Other connections use flexible PCBs or small flat cables; in these cases silicon glue is used to keep in place the movable part.

Specifications and requirements with respect to static loads and vibrations were established by the launcher company (Kosmotras and Yuzhnoye Design Office, http://www.yuzhnoye.com), and verified by simulations and ground tests. Mechanical tests for the maximum longitudinal $g$-load of 10.0$g$ were conducted at Thales Alenia Space facilities in Torino, including random and sinusoidal vibrations. Shock and acoustic loads tests have been carried out by Yuzhnoye in Ukraine.

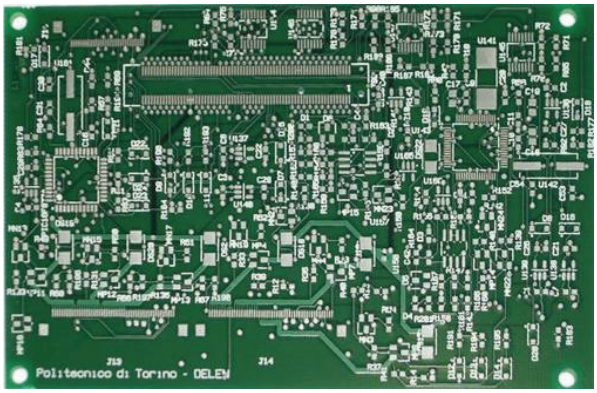



Fig. 3. An example of the PCB developed and used in PiCPoT

### 3.6 Orbit

The predicted polar orbit is at a height of around 600 km (370 miles) and takes roughly 90 minutes to complete one revolution. In optimal conditions (i.e., when the satellite passes through the zenith), the line-of-sight visibility of the satellite from any given point on the Earth lasts about 10 minutes. If we take into account the distance (which varies depending on the altitude of the satellite over the horizon) and absorption due to the atmosphere, an electromagnetic signal would on average be attenuated 160 dB. Given the available power at the transmitter on the satellite, the transmitting and receiving antenna gains, and the receiver characteristics, the maximum transfer rate, assuming a certain bit error rate, can be computed.

### 3.7 Power

The satellite has to generate its own power to function properly. The Sun is the only power source, and solar panels are used to transform light into electricity. At the Earth-to-Sun distance, the total power per square centimeter potentially available is 0.135 W. 5 out of 6 faces of the satellite are covered with solar panels, and only 3 of them are facing the Sun, with varying form factors (i.e., the angle between the solar panel and the incoming light ray). From these information, combined with orbit data the efficiency of the transformation process, the total available power can be computed. Since the satellite spends most of the time in a semi-idle state, power can be accumulated in batteries, to make it available at a later time. Our calculations show that solar panels provide an average of 1.68W of power, that we use to charge six battery packs, and gives an average power available for all electronic systems of 820mW, when worst case efficiencies of both the battery charger and the batteries themselves are taken into account. Total charge time is 63.4 hours (roughly 2.5 days), and the maximum available energy is 202kJ. Peak power consumption of the electronic subsystems can obviously exceed 820mW, provided that they are not used continuously.

### 3.8 Size and weight

Launch costs make a considerable fraction of the total costs of a small satellite, and are directly related to the size and the weight of the satellite itself. The shape and size of the

external enclosure should comply with requirements imposed by the launch vector (Kosmotras DNEPR LV, in our case), and in particular with the technique used to hold the satellite in place during launch and the way it is released when proper orbit is reached. Weight is the most important variable in computing the launch costs, since the amount of fuel needed to bring the satellite in orbit is directly proportional to it. The weight and size costs are grouped in "classes" (upper limit for weight and size); hence, the design constraint was to fit within the selected class limit, not true weight and size minimization. normal good design practice were applied in selecting components and sub-systems.

## 4. Design solutions

Most of the design efforts for using COTS components in a satellite are aimed to protect the system from fatal events. Techniques to achieve this goal can be classified as either physical or logical. The former includes shielding the sensitive parts and choosing devices that are less prone to errors due to radiation at a comparable price tag. The latter, while allowing events to take place, mitigates or completely eliminates their effects by acting at the system level. Examples of such techniques include error correction (i.e., in memories), redundancy at several abstraction levels, and watchdog timers to reset misbehaving devices or boards. We applied several such techniques in the design of the satellite, as described in the following.

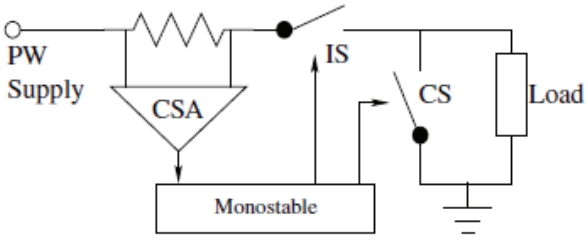



Fig. 4. latch-up protection circuit

### 4.1 Single Event Latch-up (SEL)

Latch-up (LU) occurs when a parasitic SCR made by the couple of complementary MOS devices is turned on by high input voltages (LU in ICs, caused by input over-voltages) or by high energy particles which induce a small current (this is the case for a space device) (Gray et al., 2001). The effect is a high, self-sustaining current flow, which can bring a high power dissipation and, in turn, device disruption.

LU-free circuits can be designed by avoiding CMOS all-together, or by using radiation hardened devices. Since one of the goals of PiCPoT is to explore the use of COTS components for space applications, we decided to keep only some critical parts LU-free by proper device selection, and to use standard CMOS devices in other circuits, made LU-safe with specific protection circuits.

The basic idea behind protection is to constantly measure current and to immediately turn the power off as soon as anomalous current consumption is detected. Once the transient event is over, normal operation can be restored. This technique is analogous to a watchdog timer, except that it actively monitors the circuit to be preserved, rather than waiting for the

expiration of a deadline. Each supply path should have its own protection circuit, which should itself be LU-free, e.g. by using only bipolar technology.
The block diagram of the protection circuit of a single supply path is shown in Fig. 4, and includes:

- a current sense differential amplifier (CSA),
- a mono-stable circuit with threshold input,
- isolating and current-steering switches (IS and CS).

When the current crosses the limit set for anti-latch-up intervention (usually 2× the maximum regular current), the mono-stable is triggered and isolates the load from the power sources for about 100 ms. To fully extinguish the LU, the shunt switch (CS) steers residual current away from the load.
The main problem in the design of LU protection is to balance the LU current threshold with current limit of the power supply. Namely, if the regulator current limit is activated before the LU, the current is limited but not brought to 0, and LU continues for indefinite time.

### 4.2 Single Event Upset (SEU)

PiCPoT contains several digital circuits, including 5 processors, different kind of memories and programmable logic devices. When a high-energy particle hits a circuit, it may cause a transient change in voltage levels. While this is usually not considered a problem with analog circuits, it might adversely affect digital circuits which typically involve high speed signals with steep edges, and especially memories that rely on tiny voltages to carry their information. If the final effect results only in a glitch (*Single Event Transient*, *SET*), then it can safely be ignored; however, if the event is latched, or directly upsets a bit (or multiple bits) in a memory or a register, it will probably lead to incorrect behaviors (*soft* errors). In extreme cases, such as when a configuration bit of a programmable logic device turns an input into an output, it can even cause severe damages.
In the less dramatic case of a soft event, we distinguished between three different kind of errors:

1. errors on dynamic data and/or in code segments resident in volatile memory;
2. errors on data stored in non-volatile memory;
3. errors on program code stored in non-volatile memory.

The outcome of such events may be wrong data, wrong behavior (if the event affects some data dependent control, for instance) or even a crash (i.e., if the upset results in a non-existent op-code for a processor).
The available solutions to address the problem are very diverse, each with its own advantages and shortcomings. Some cope with all three kind of errors, others address only some of them. We applied different techniques in various parts of the satellite, depending on the kind of protection we wanted to provide. The selection was driven by the need to keep the design simple and power consumption and total budget low. We did not employed radiation-hardened devices (which are too expensive and against the whole philosophy of the project to use COTS components), and memories with error corrections code (ECC, which are only useful for dynamic data and do not protect against multiple bits upsets).
The susceptibility of COTS components to radiation can be very different. Careful selection of the best devices for the application allows us to strongly reduce the probability of single event upsets. We examined several kind of memories in search for the best ones, and in particular we considered:

- *Dynamic RAM (DRAM)*: it is the most dense memory, used when large amount of memory is required. Being based on charge held on a capacitor, it is rather sensitive to radiations. Those parts of the satellite that depend on this kind of memory must be protected in some other way.
- *Static RAM (SRAM)*: the information is stored into a two-state device (flip-flop); it has been shown that these are more sensitive to radiation than dynamic RAMs (Ziegler et al., 1996), but have the advantage of consuming less power. Processor registers also use the very same technology.
- *Flash*: even more energy than conventional static RAM is needed to change the state of a bit. For this reason, flash devices are more tolerant to radiation and are a good candidate for important data and code. They are also non-volatile and cheap, but cannot be used for normal processor operations, since writing performance are extremely poor.
- *Ferro-electric RAM (FeRAM)*: this is a kind of memory (Nguyen & Scheick, 2001) based on ferro-electric phenomenon. A ferro-electric material (usually an alloy of zirconium or titanium) can be polarized by applying an external electric field. The polarization hysteresis allows to store information. Writing operations on an FeRAM require lower voltages (3.3 V, for instance) and are 2 to 3 order of magnitude faster than in flash memories. This allows energy saving and at the same time maintains the good tolerance to radiation of flash devices. Since few information about the behavior of FeRAM in space is available in the literature, its use on PiCPoT was limited to a single board.

We used a mix of all the above memories because strengths and weaknesses were often complementary. Dynamic and static memories were used for program execution, while Flash and FeRAM were used for permanent data and program storage. Being highly experimental, FeRAM was only used to hold non-vital data, such as the telemetry stream acquired from sensors.

Another technique to handle the problem of SEU is to use redundancy. In general, at least three replicated units are necessary to implement a voting mechanism, where the majority wins and allows correction of a fault. The replicated unit can be a complete board (processor, memories and peripherals), a physical device on a board (three instances of the same component) or an abstract unit within a device (three memory segments in the same chip, holding identical information). This method potentially allows active identification of an SEU even in RAMs during the execution of a program, and to promptly act to correct it. However, the space available inside the satellite did not allow us to replicate identical boards, or even devices within a board. Nonetheless, in some of the processor boards the program stored in flash memory is maintained in multiple copies and a procedure to search for SEUs can be explicitly activated. Data, such as pictures or telemetry, on the other hand, is not protected and if an SEU occurs, the information downloaded to ground will simply be incorrect.

Since RAMs, both static and dynamic, including registers inside the processors, are the most sensitive devices to SEU, and they are not replicated, other techniques must be used to ensure proper behavior. Our solution is to periodically turn off processor boards and start a complete boot procedure. Given that the program is stored in flash memory (possibly with some duplication) and that RAMs go through a power cycle and reset, the soft error will be completely eliminated. Obviously, whatever command was being executed at the instant the

SEU occurred will potentially result in wrong data or a crash. This however does not preclude the system to work correctly at the successive re-boot. The periodicity that was selected is 60 s: it allows all but the longest command to be executed with a good margin; the notable exception is the download of a picture to ground, which might need to be split into multiple commands acting on different portions of the image, if it is too large to be transmitted in 60 s at the available bit rates. This technique is similar to a watchdog, but the chosen periodicity is a hard deadline and cannot be extended by the controlled processor boards.

Communication between boards may also be affected by SEU, as well as by other noise sources. Long data streams (tens or even hundreds of kbytes), such as when transmitting a picture from one board to another for successive download to ground, are more subject to problems than very short (a few bytes) commands. For this reason, long communications are protected by a protocol that involves CRC computation and retransmission. Among the various alternatives, the *X-modem* protocol has been selected for its simple implementation and because it is often a standard feature of terminal emulation programs on PCs, which allowed easy testing of the boards before they were connected and assembled together.

### 4.3 Cumulative effect of radiations

Although in Section 3.1 we stated that total dose effects have not been considered, in fact we do provide protection against possible permanent failures, as opposed to the single event effects described in previous sections, in the satellite electronic boards. This is mainly achieved through three orthogonal techniques:

1. replication of functional chains;
2. differentiation of the replicated units with respect to the algorithms, topology, architecture and technology;
3. graceful performance degradation.

The former provides multiple alternative units to perform the same functionality. Any unit can be used, but only one should be selected. Unlike replication used to address single event effects, where all units work at the same time and on the same data, this technique does not provide the ability to correct a failure. Simply, if one chain fails for any reason, one or more backups exist to take over. In some cases, multiple units can be used to reach a particular goal, but failure of any of them does not preclude the overall system to work, although functionality and/or performance might be affected.

In order to prevent similar problems from affecting all the replicas, different implementation solutions are used in the various chains. We considered using different technologies (CMOS versus Bipolar, Flash versus FeRAM, NiCd versus LiPo), architectures (two different processors and instruction set, different memory hierarchies) and algorithms (chains were developed independently by different groups, so that bugs in the software, for instance, did not show up identical in replicas).

Examples of replication with differentiation are the power supply, which can survive several failures, although with performance degradation (less available power), the on-board computers, the timing unit and the communication unit. The latter provides two communication channels using separate antennas, at frequency of 437 MHz and 2.4 GHz respectively. More details about the implementation can be found in Section 5.6. The only non replicated unit is the camera control board (payload).

### 4.4 Shielding

In a satellite two kind of EMI must be handled: radio-frequency interferences and radiation. We developed special solutions to reduce problems related to RF phenomena. The outer structure is based on six aluminum alloy faces, electrically connected together, using screws which are less than $\lambda/4$ apart for the highest used frequency (2.4 GHz). The wires connecting solar panels (external) and switching converters (inner part) go through special feed-through filters.

Internally, only one board deals with RF and it is structured to limit interactions with other subsystems. The board is isolated from the others using multiple ground planes and placing the RF components on the face opposite the other modules.

There is not enough space to use thick shields to protect from high energy particles, so we used internal placement of boards, batteries and panels to reduce its influence. PCBs are lodged in the inner part surrounded by a "sandwich" made of solar panels, aluminum panels (external structure), battery packs and aluminum panels (internal structure), which reduces radiation effects. Other techniques, such as the one described in previous sections, further mitigate radiation induced problems, like latch-up and single event upsets.

### 4.5 Power consumption and dissipation

Being a battery-based system, the whole PiCPoT project was made on low-power concept. In order to reduce power consumption every component has been chosen in commercial low-power domain. When low-power commercial components were not readily available, such as in the case of the high performance image processing sub-system, our solution was to keep them either in idle state or completely switched off when not in use, or with reduced performance if allowed by the application.

Typical power consumption of on-board systems is summarized in Table 1, where both peak and average power are indicated in column 3 and 4, respectively. Column 2 shows, in percentage, the fraction of time each subsystem is expected to be turned on. Power Management is always on, while on-board processors, payload and communication are used only when necessary. Total average power is around 0.5W. Since the average power generated by solar panels is about 820mW, we have an average margin of about 300mW. The extra power is dissipated on shunts (zener diodes) inserted on the power subsystem to avoid over-voltages on the power bus.

RF transmission is the only part which needs a lot of power for a medium-long period, since the power level is related to the satellite distance from the Earth. On the RF board we have two different devices, each of them dealing with a different band (437MHz and 2.44GHz). Power amplifiers are the most power-hungry elements, as they have to generate an output power of about 2W each. The most critical is the 437MHz one whose efficiency is only 25%, while for 2.44GHz it raises to 40%. For these reasons we had to dissipate about 6W in the worst case. This has been met using different solutions:

- The PCB contains 3 ground-planes that extend their own area to all the space available, in this way heat generated by the PAs is distributed to the entire board.
- The PCB surface is covered with high-thermal-conductivity coating.
- Chip body is connected to metal face through a thermal conductive mat. The panel is aluminum black-anodized in order to allow maximum radiation.

Thermal analysis had shown that our satellite, in its orbit can reach at most 80°C. Boards have been tested in thermo-vacuum environment, showing good performances also in corner cases.

| Device | Duty Cycle | Peak Power | Avg. Power |
|---|---|---|---|
| Power Mgmt. | 100% | 20mW | 20mW |
| Proc A&B | 6% | 200mW | 12mW |
| Payload | 0.5% | 3.84W | 21mW |
| TxRx | 2.6% | 17.2W | 443mW |
| Total | | | 496mW |

Table 1. PiCPoT power budget.

### 4.6 Interconnection solutions

When the amount of space available is small the problem of interconnecting a complex system like a satellite can be hard. In our case we had to share multiple connections among the boards in order to allow:

- digital communications (for actuators, house-keeping, …);
- analog signals acquisition (mainly for sensors);
- power connections;
- RF communications links.

When using connectors for these links, care should be taken to avoid detachments caused by strong vibrations during launch.

This issue was solved using a series of stackable connectors, as shown in Fig. 5, which represent a CAD model of the mounted boards. This solution leads to a pseudo-shared bus, which ensures communication among modules and reduces problems related to vibrations. The connector is tightly connected to the board, ensuring the electrical link. On this connector were channelled all the communication signals among tiles and many of the analog signals (used to acquire sensors values). In this way we obtain an efficient vibration-proof connection for digital and analog signals. For this goal we use a main stack of indirect narrow-pitch 140 pin connectors, and a second 60 pin connector for selected signals.

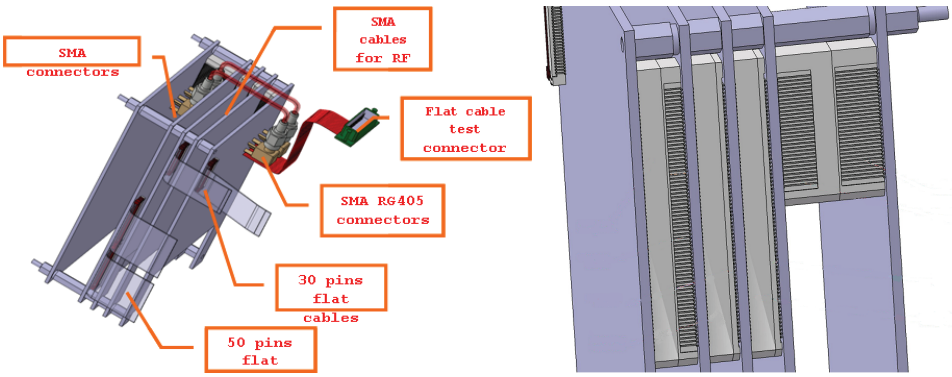



Fig. 5. CAD model of wirings: RF and batteries connections (a) and stackable connections among boards (b).

On the remaining signals (especially for power lines, and RF connections), instead, we have to use special media:

- SMA and coaxial cables for RF, in order to guarantee a controlled impedance and low losses between boards dealing with radio-frequency signals and antennas;
- multiple cables for connecting solar panels, batteries and power suppliers board, for achieving redundancy on these critical connections;
- flat cables to connect analog and digital signals to a board that was not stacked with the others.

Figure 5 shows the organization of the signal cables; it also includes a test connector which is available on one of the external plates of the satellite, in order to allow verification of the satellite electronics while it is closed inside its enclosure. Figure 6 illustrates the wiring of power cables when all the electronic boards are mounted in the satellite structure, and shows the test connector and cable, as well. Two sets of power cables are necessary: one to link solar panels to the batteries, and another to bring power from the batteries to the electronic boards.

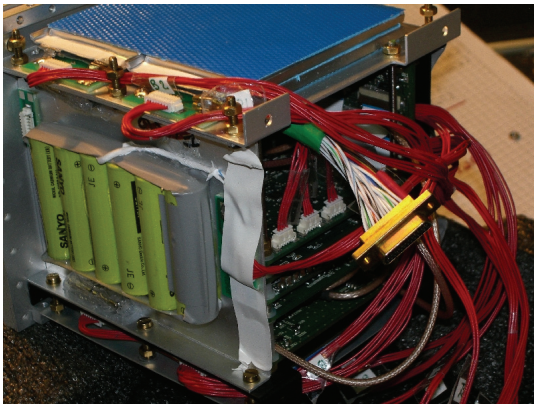



Fig. 6. Power cable wiring in the mounted satellite

## 5. Architecture and functional units

### 5.1 Satellite architecture

The complete architecture of PiCPoT is shown in Fig. 7. The core of PiCPoT satellite is a redundant central power management and timing unit (PowerSwitch) which drives two processing chains (A/B). Every minute the timing unit selects the most charged battery and turns chain A on.

The processor waits for a command from ground, which is decoded and executed. If no command is received within 5s, telemetry is sent to ground anyway and the chain power is turned off. If a latch-up occurs, power consumption rises quickly, and power is turned off to extinguish the latch-up.

A similar sequence of actions takes place at time shift of 30s on chain B, which implements the same functionalities as chain A, but with different components and using the other radio link.

### 5.2 Power supply

The main power sources are 5 triple junction GaAs solar panels. Each of them has a dedicated Maximum Power Point Tracker (MPPT) made with a switching power converter, using only bipolar IC, not sensitive to latch-up. The five converters allow the system to survive, even if four of them got damaged.

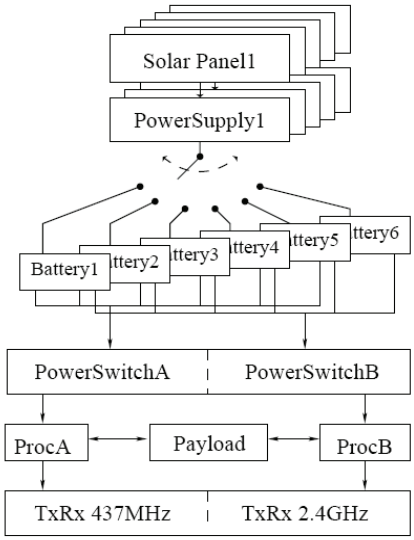



Fig. 7. Architecture of PiCPoT satellite

The satellite uses 6 battery packs (2×7.2V900mAh Ni-Cd, 4×7.2V1500mAh Li-Po), which feed two independent power busses.

### 5.3 Power switch

This board is composed of two (A/B) independent subsystems responsible for:

- Battery selection
- Voltage regulation
- Schedule the power up
- Latch-up events count

For design diversity, the A chain uses a Microchip PIC microcontroller, while chain B uses a Texas Instrument MSP430. The two power on cycles are shifted 30". Latch-up events are counted and transmitted to the ground station.

### 5.4 On-board processors A and B

We used two different commercial low-power processors: a Chipcon CC1010 (ProcA), and a TI MSP430 (ProcB). They have similar tasks but different design solutions to increase system reliability and to prevent a single bug to crash both chains.

The functions performed include: data acquisition, battery management, interpreting and executing commands received from ground.

### 5.5 Camera handling

The main payload is a set of three color cameras (commercial units with a standard PAL video output), with different focal lengths. The analog video is converted into a standard ITU-R BT.656-4 digital stream, then the interlaced raw image is converted into a compressed JPEG picture, which is divided into 9 zones and individually sent to ground. An Analog

Devices Blackfin DSP manages the board and implements the compression algorithm and permanent storage of the pictures.

### 5.6 RF transceivers

The satellite operates on two different frequencies: UHF at 437.480MHz and S-Band at 2440MHz (radio amateur satellite bands), connected respectively to the A + B chains. The UHF downlink is compatible with the amateur PK96 packet radio.
The S-Band link data are organized in a similar way but uses a modulation scheme not directly compatible with amateur stations. Link budget is summarized in Table 2.
The UHF link is based on the transceiver included in the ProcA OBC, Chipcon CC1010. The S-band link is based on Chipcon CC2400 transceivers. Two separate devices are used for TX and RX. The UHF system is equipped with a folded double helical antenna (Orefice & Dassano, 2007), while S-band uses a Planar Inverted-F Antenna (PIFA), as depicted in Figure 1; the same figure also shows the three on-board cameras.

| Link | 437 MHz Uplink | 437 MHz Downlink | 2.4 GHz Uplink | 2.4 GHz Downlink |
|---|---|---|---|---|
| TX Antenna Gain | 24 dBi | 1.5 dBi | 25 dBi | 4 dBi |
| Output Power | 47 dBm | 30 dBm | 40 dBm | 33 dBm |
| Attenuation | 154 dB | 154 dB | 169 dB | 169 dB |
| RX Antenna Gain | 1.5 dBi | 24 dBi | 4 dBi | 25 dBi |
| RX Sensitivity | -103 | -115 dBm | -118 | -130 dBm |
| Noise Temp. | 450 K | 450 K | 400 K | 400 K |

Table 2. PiCPoT link budget

### 5.7 Attitude measurement and control

No orbit control is provided in PiCPoT, as there is no room for an orbit-correction propulsion system. Anyway, the short predicted life-time (3 months) does not require adjusting the altitude of the satellite. Moreover, past university satellites with no orbit control showed a long period of activity with no correction at all.
On the other hand, attitude control is necessary for two reasons:

- Aiming the antennas to ground for communication.
- Aiming the cameras toward the earth for taking pictures.

The large field of view of even the highest resolution on-board camera allows a low pointing accuracy. Also, antennas are studied such that the transmission lobe spans a wide area. We therefore looked for low cost and easy technological solutions, which could exploit the favorable orientation of the Earth magnetic field in the geographical area of Europe. We provide two ways of controlling attitude:

- A passive mechanism based on permanent magnets to align the satellite with the Earth magnetic field, with hysteresis plates as dampers to minimize oscillations.
- An active reaction wheel driven by a brushless Maxon EC 32 motor, controlled through commands from Earth.

The permanent magnets (four Neodymium 35) can only control the satellite attitude around two axes perpendicular to the Earth magnetic field. The magnets are positioned and oriented such that the bottom panel, where the antennas and camera lenses are mounted, will face the Earth over the northern hemisphere, and space over the southern hemisphere (Fig. 8). The reaction wheel is mounted such that its axis is parallel to the magnetic field: by acting on it we can reduce the spin-axis rotation of the satellite. However, given the symmetry along the vertical axis of both pictures and radio communications, this last control was not strictly necessary.

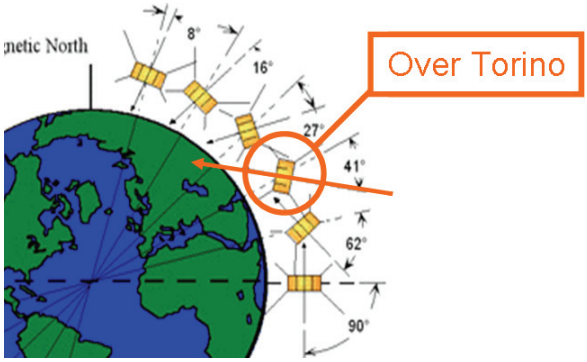



Fig. 8. Satellite orientation over our base-station.

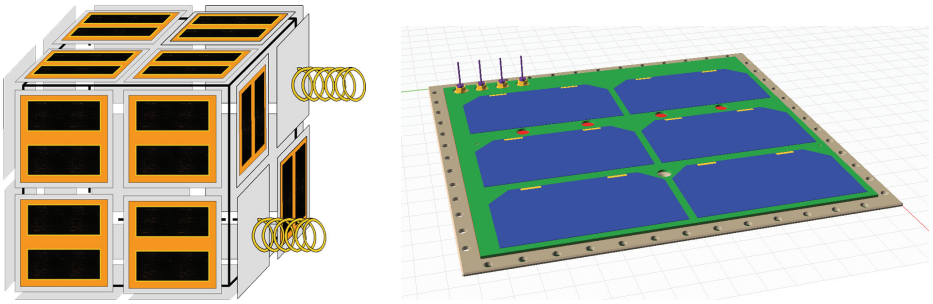
## 6. The ARaMiS modular architecture

The PiCPoT project was carried out as a dedicated design, aimed to the development of a single small satellite, with specific characteristics. For the design of other satellites, with different size and payloads, the design must be repeated from scratch. A more effective way to reduce the cost of a nano- or micro-satellite mission is to reduce design and non-recurrent fabrication costs, which usually account for more than 90% of the overall budget, by sharing the design among a large number of missions. Design reuse and modularity are the rationale behind the ARaMiS project, that is, to have a modular architecture based on a small number of flexible and powerful modules which can be reused as much as possible in different missions. Using the same module(s) more times obviously allows to share design, qualification and testing costs and to reduce the time-to-launch.

### 6.1 General description

The first step in the ARaMiS project has been to identify the subsystems which are used on every satellite and provide critical functions. We have then concentrated our efforts on these subsystems, i) mechanical subsystem; ii) power management subsystem; iii) telecommunication subsystem; iv) on-board processing subsystem; v) payload support; vi) ground segment.

The architecture of ARaMiS is based on modular smart tiles. They are placed on the outer surface of the satellite and have also mechanical purposes (Fig. 9). The inner part of the satellite is mostly left empty, to be filled by the user-defined payload. This last is the only part to be designed and manufactured ad-hoc for each mission; thanks to modularity and reuse, each tile is designed only once, but manufactured and tested in relatively large quantities.



(a) (b)

Fig. 9. Example of modular composition of developed tiles (a) and Power-Management tile, external view. (b)

Communication and power distribution among tiles and other units are also "standardized", and can exploit the same interfaces. Reuse also allows to put an increased design effort to compensate for the lower reliability of COTS devices, therefore achieving a reasonable system reliability at a reduced cost.
The outer tiles are of two types:

- Power management (see Sec. 7 and 8), which are composed of: i) a solar panel; ii) a rechargeable battery to store energy; iii) a battery charger; iv) a microcontroller-based housekeeping module to keep track of voltages, currents and temperatures inside the tile; v) an active magnetic and/or inertial asset control. A number of such tiles (depending on power requirements of the mission) are placed around a cubic or prismatic shape (or displaced after satellite release) and represent a pre-designed and pre-assembled modular power management subsystem.
- Telecommunication (see Sec. 10), which are composed of: i) a microcontroller-based programmable transceiver; ii) a 437MHz or 2.4 GHz modem; iii) a power amplifier (for transmission) and low-noise amplifiers (for reception); iv) an antenna system. At least one such tile is placed as one satellite face, preferably pointing to ground, and takes care of reliable data and command exchange to/from ground. The user, who needs not take care of any telecommunication detail, sends/receives data, via the internal data busses, to/from this tile and, consequently, to/from ground, in a completely transparent manner.

The inner modules can be of many types. At present we are developing only one of them: the on-board processor and payload support (see Sec. 9 and 11), which takes care of all data handling for housekeeping, and interfaces to the user-defined payload. This tile has some CPU power and memory available for the user, such that simple payloads may use it instead of having their own processor.

### 6.2 Mechanical subsystem

To make the mechanical structure of ARaMiS as simple as possible, the tiles have also structural properties, that is, they are a significant part of the mechanical subsystem. This allows to reduce the overall weight and cost and simplifies assembling and testing of the satellite. This approach applies to small satellites (that is, with side not longer than 50 cm), while larger one should have a stiff frame to carry the many modular units.
We have achieved this goal by:

1. standardizing the size of the tiles: 15×15 cm$^2$. We have decided not to use the CubeSat (Heidt et al., 2001) standard, as it was felt to be too much constraining for larger satellites, while not offering significant advantages for the intended applications.
2. using a self-bearing 1.5mm thick anodized Al, which is also the mechanical support for solar panels, batteries, antennas, processors, PCBs, etc. (depending on the tile);
3. having screw holes at 15mm apart, which ensures high mechanical strength together with an excellent electromagnetic sealing also at 2.4 GHz;
4. reducing the rest of the mechanical subsystem to one 6 × 6mm$^2$ Al square rod per edge, which is screwed to the outer tiles, fixing them together;
5. having a very simple internal mechanical frame, which is fixed to the external mechanical frame, to hold the predefined on-board processor and the user-defined payload hardware and circuits (if any).

This choice allows a certain degree of freedom in the shape and size of the satellite. A few possibilities are listed below:

- smallest cubic shape (15 × 15 × 15 cm$^3$): 5 power management tiles plus one telecommunication tile; one internal processing tile and the user-defined payload. About 1,000 cm$^3$ are available for the payload.
- larger cubic (or prismatic) shape: each surface is made of n×m or m×q or n×q tiles, either power management or telecommunication or payload-specific. Sides with at least 2×2 tiles require a cross-shaped rod; one or more internal processing tiles and the user-defined payload (see Fig. 9-a).
- small hexagonal/octagonal satellite: six/eight power management or telecommunication tiles with two hexagonal/octagonal Al sides (possibly telecommunication or payload-specific).
- larger satellites (larger than 50 cm in side) will require a stiffer mechanical frame to survive vibrations during launch.

### 6.3 Attitude Control System

The Attitude Control System (ACS) is used to stabilize the spacecraft and to orient the satellite in the desired direction. First of all, it is necessary to define the modes of operations of this system: De-tumbling, Stabilization, Spin compensation.

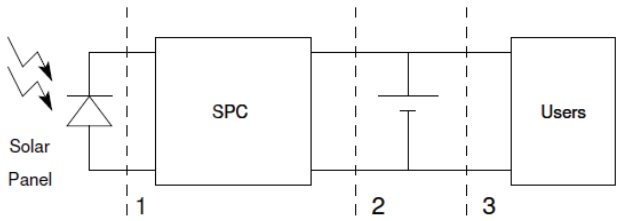



Fig. 10. Power generation and distribution.

In de-tumbling mode, the controller is responsible of dumping the initial angular velocity which the launcher gives the satellite upon release: only after this preliminary phase the real stabilization phase could begin. In stabilization mode the actuators are used to control the orientation of the satellite to align antennas toward the Earth while in spin compensation mode the controller is responsible of compensating the satellite spin-axis rotation. Sensors

are dual-axis magnetometers and photodiodes, both housed on Power Management tiles: data are collected by the OBC which also executes the attitude stabilization algorithm. Corrections to the attitude are sent to Power Management tiles, which control the actuators. For orienting the satellite we used a combination of magnetic torquers and small reaction wheels: each tile operates only on one axis but, using all the tiles together we get the 3 axis stabilization.

## 7. Power management

This module is responsible for gathering power from solar panels and control the energy storage process in order to grant an adequate power level to all subsystems during the orbit. Furthermore it is responsible for scheduling the power-on time for all the other subsystems, to reduce the overall power consumption, and it also houses the attitude control sensors and actuators.

### 7.1 Power generation and storage

As for most satellites, the main power sources in ARaMiS are solar panels mounted all around the satellite (see Figure 9-b). Their output power is highly variable: we therefore need a separate solar panel controller for each of them. Since our goal is to use COTS components to build a highly reliable system we need to implement different strategies to increase fault tolerance.

The basic Power Management tile is made of a solar cell, its control electronics, a rechargeable battery and the Al-alloy panel which holds everything together and encloses the satellite (as shown in Fig.10). All these basic units are connected with a power bus to share the generated power, and with a communication channel to exchange system status information (see Sec. 8.1 and 8.2 respectively). The power generation system is composed by this tile, replicated as many times as needed to get the desired power. All these tiles work in parallel and this redundant solution helps also from the fault tolerance point of view, making the system able to tolerate faults with graceful performance degradation.

The solar panel controllers (SPC), a Maximum Power Point Tracker (MPPT), is responsible for tracking the maximum power point of the solar panel according to instantaneous environmental conditions. The system is a switch-mode power supply which allows to change the load seen by the solar cells to extract maximum power. We adopted an hysteretic boost converter because it causes a reduced current stress to solar panels, requires fewer components and can be implemented without using CMOS integrated circuits, prone to latch-up in space environment. Temperature based quasi-MPPT controllers are quite widespread because their control scheme is simple: they infer the maximum power point from the solar panel temperature and track it. This requires the precise knowledge of the temperature/ voltage plot at the maximum power point: with this strategy solar cells should be precisely characterized and changing the solar cell means changing some parameters in the MPPT. Instantaneous power tracking solar panel controller is based instead on the measurement of instantaneous power: this type of controller has to sample solar panel voltage and current, and then multiply them in order to get the power. In ARaMiS we adopted both these control strategies because they allow an accurate maximum power tracking and reduces the probability of common mode faults. We in fact developed an analog temperature MPPT, built with bipolar devices to avoid latch-up events, and a digital maximum power MPPT implemented using TI MSP430 microcontroller: both systems act as

a single controller increasing fault tolerance. The analog controller is always available but it suffers from errors in the characterization of solar cells which leads to small losses in tracking the maximum power point. The digital controller is instead not always available (due to possible latch-up events) but it is able to correct the maximum power point estimated by the other controller to reduce losses.

### 7.2 Latch-up protection

Uses the same "anti-latch-up" circuit developed for PiCPoT, which become a library component block in the new modular design, in order to maintain the circuit safe even if based on COTS components. The circuit is described in Sec. 4.1.

## 8. Signal and power busses

On ARaMiS power distribution and system communications are shared infrastructures, and it is necessary to avoid a single faulty user to block the others or to damage the infrastructure.

### 8.1 Power bus

A power bus has to allow all the energy producers and all the energy consumers to respectively distribute and receive all the energy available to provide the power path from the solar panels to the users. This structure is replicated in each external tile, with redundant connections to increase fault tolerance. Joining together every power-path in the sections shown by dashed lines we can move energy from one point to another by-passing any faulty modules. These solutions can lead to advantages and drawbacks:

1. if we allow a solar panel to be shared by multiple SPCs we gain in redundancy, but we have to introduce extra-hardware to manage multiple panels and we dramatically increase the number of connections among modules and software complexity for SPC;
2. if we allow each battery to be re-charged by many SPC, again, we gain in redundancy, but, as in the previous case, we increase global complexity in hardware, software and interconnects;
3. joining the power-path after batteries is quite simple since it can be done using a simple diode and leaving the output voltage unregulated, with small extra-hardware overhead.

For the above reasons we keep separate power paths, merged towards the load with active switch circuits, to avoid power loss caused by diode voltage drops.

### 8.2 Communication bus

Each ARaMiS module needs to communicate with the others and with the radio-frequency transceiver. We need high reliability, some degree of fault tolerance, multi-master protocols, and a bit-rate of about 200 kbps. We use an isolated redundant bus as shared medium. Each tile is coupled to the others by means of isolation transformers, and data are encoded to remove DC component. We use differential signaling, so even if a fault shorts one of these lines, the other still remains available and continues transmitting data.

Since the bus is multi-master we use Carrier Sense and Collision Detection to deal with data collisions. Every data packet contains as the first part the univocal master ID. If a collision occurs, all the masters involved wait for a random period before starting a new bus access.

## 9. On-Board Computing

In ARaMiS the On-Board Computing (OBC) unit is mainly responsible of managing the system, in particular of: creating and transmitting (by Transceiver board) Beacon packets, decoding and executing commands, executing attitude control algorithm, storing housekeeping data, controlling Payload boards.
At turn on, the OBC unit verifies if a command had been received by the Transceiver board: it issues the corresponding commands to the other sub-systems of the satellite, gather replies from them and send data to ground. If no command was received the OBC gathers housekeeping data from all the boards and creates the beacon packet which will be sent again with the Transceiver board. In this way any people listening to the satellite could know the status of the satellite.
The attitude control algorithm is the most time consuming task the OBC should perform: it takes data from the magnetic field ad sun sensors on the Power Management tile and it controls the actuators. The main purpose of this system is to first de-tumble the satellite after the detachment from the launch vector and to keep antennas aiming the earth.
Data storage is accomplished with commercial Secure Digital Flash card To ensure also radiation hardness we developed a system that is similar in principle to RAID-1 storage units in commercial PC: data are stored at the same time in multiple Flash cards and they are then read and compared to each other to correct errors. Since the capacity of these devices is high (it easier to find 1GB devices than for example 16MB ones) it is also possible to store multiple copies on the same card also because housekeeping data to store per single orbit are only approximately 10kB.
The OBC is also responsible of controlling Payload boards, with the possibility to perform some mission specific tasks. The unit is made by TI MSP430 devices connected in a triple modular redundant structure, used to correct Single Event Upset errors originated from high energy particles.

## 10. Transceiver and communications

The communications subsystem of ARaMiS follows the modularity philosophy of the satellite: a basic communication tile is provided as standard, while dedicated tiles are foreseen in case of special applications. The communication system is connected to the cross-bus architecture. It is used to receive command and control packets from the Ground Station and to send back telemetry and status information. The bandwidth needed to exchange this kind of information is usually low, so the RF link was designed for low speed and low power.
In order to achieve fault tolerance, two different channels are used, in the bands allocated by ITU for satellite communications. The first channel in the UHF 437MHz band, the second in the SHF 2.4GHz band. The data contents of the two links are equivalent, thus providing two interchangeable possibilities to communicate with the satellite. Both channels use an half-duplex protocol, sharing the same frequency for downlink and uplink.
The UHF downlink was designed to be compatible with the amateur PK96 packet radio. It uses the UI frame defined in the AX.25 standard, following the subset for APRS. Enabling the reception of ARaMiS telemetry by radio-amateurs around the World makes possible to collect data from orbits unreachable from our Ground Station. The telemetry data packet format for each ARaMiS satellite will be publicly available from the web-site.
The S-Band link data are organized in a similar way but, to avoid the computational overhead of some of the operations required by AX.25 (scrambling and bit-stuffing), the

transceiver uses a modulation scheme not directly compatible with amateur stations. External users are not allowed to issue commands to the satellite, and the format of the uplink data packet will not be published. To reduce the probability of false command detection, data are strongly encoded.
The UHF link is based on a TI/Chipcon transceiver, model CC1020. This chip implements a complete digital UHF radio, with one output and one input channel, with good input sensitivity and output power in excess of 1mW. To complete the UHF channel it is therefore necessary to connect the transceiver to a microprocessor on the base band side, to a Power Amplifier (PA) on the transmit side and to a Low Noise Amplifier (LNA) on the receive side. An electromechanical switch is used to connect the single ARaMiS UHF antenna to the output of the PA or the input of the LNA. This device is more robust than an active switch (no radiation-related problems), and more power efficient than many diode based circuits. The power amplifier selected for the first implementation of the board is the RFMD model RF2175, which provides an output level of +34dBm. This device is very small, provides enough output power with good efficiency and is latch-up free. Depending on the amount of power available for communications, a power module can be cascaded to the PA to attain an output power in excess of 8W.
The processor supervising the UHF link is a TI MSP430. Its functions are essentially:

- to exchange base-band command/status/telemetry packets with OBC and Payload processors;
- to encode/decode packets, by performing scrambling/ descrambling, bit-stuffing, insertion and removal of prologue and header information, so that OBC and Payload to not need to cope with communications details;
- to supervise operation of the transceiver and RF subsystem (power sequencing, antenna switching).

The S-band link is based on the transceiver TI/Chipcon CC2510. This device incorporates a complete radio modem and a processor kernel of the 8051 family, thus it is not necessary to use an external processor as in the UHF subsystem. The CC2510 is not optimized for high power operation: it uses the same pin as the output of the internal PA (output power: 0dBm) and the input of the internal LNA. It is possible to build an external circuit to drive an external PA and get RF signal from an external LNA, but we decided to use two devices, one as receiver, the second as transmitter. The PCB layout is very simple, removing the risk of unwanted coupling between transmit and receive paths.
The amplifier on transmit side includes two devices, a NEC UPC2762 followed by a RFMD RF5189, with an output power about +32dBm. The power amplifier is pushed to its linearity limits, but this is not a problem because the GFSK modulation adopted in the link is not sensitive to nonlinearity. As for the UHF subsystem, the antenna switch is an electromechanical relay. On the receive side we inserted a low noise amplifier, Maxim MAX2644, and obtained a sensitivity of about -120dBm. An optional power module is available to increase output power to about 7W at the expenses of a much higher power consumption.
Antennas are strongly dependent on the satellite dimension and on the desired bandwidth. Many solutions are possible, but are not taken into consideration here.
As previously stated, the bandwidth of the standard RF links is quite low. The UHF connection data rate is 9.6kbps, while the SHF rate can be set from 10kbps to 1Mbps, but analyzing the link budget it is possible to show that only using the lower speeds a reliable communication is attained. Some of the foreseen applications (for example high resolution or real-time imaging) require a large bandwidth downlink. In this case the performances of

the standard links are not sufficient, even increasing the output power. To address this case we are designing a large bandwidth downlink tile, based on a large FPGA for base-band processing, a high speed dual channel DAC converter with internal digital filtering and over-sampling, a PLL, an I-Q modulator, followed by a PA. The architecture is much more complex than the standard one, but allows for Software Defined Radio implementation, to match the characteristics of the transmitter to those of the application. The obvious disadvantage of this solution is increased power drain from the satellite's resources and increased complexity, so its use will be limited to very demanding applications.

## 11. Payload

The payload is a mission dependent module that cannot be fully pre-designed. However, a number of applications often share a similar structure, where a central processor is used to control one or more instruments, translate commands from ground into actions, provide on-board storage capacity and communication facilities. Although some very small applications may leverage the presence of the OBCs described in Sec. 9, these tasks typically require a moderate to fairly powerful processor which exceeds to computational power supplied by the OBCs themselves.

Therefore, we decided to provide a reference design of a processor board which seamlessly interconnects with the rest of the satellite, as a starting point for a more detailed design which takes into account all the requirements of the applications. Nonetheless, the reference design can be used as-is in those cases where no additional features are necessary and design time should be short.

While designing the board, we tried to follow ARaMiS philosophy as much as possible, resulting in a set of characteristics that the board should satisfy:

- Modularity here intends as the ability to customize and expand in several ways the reference platform, both in terms of hardware and software.
- Standardization of the programming environment, based on the C language and on common operating systems, such as Windows CE or Linux with real time extensions, with hardware abstractions using device drivers.
- Fault tolerance, based on techniques such as hardware and software redundancy, watchdog timers, and error detection and correction; these methods should operate independently of the user program that controls the application.
- Low power consumption, through static or dynamic voltage and frequency scaling, as well as selective instantiation and/or activation of peripherals devices.
- Low cost, by using COTS components and commercial technology.

After considering several possible alternatives, we eventually chose a soft processor to be programmed on an FPGA. This allows us to achieve the maximum flexibility in the processor configuration, while providing spare logic for application specific enhancements.

The selected processor is a Leon 2, running on a Xilinx Virtex2 Pro FPGA. The only operating system which has currently been tested is Linux 2.6 with a full development suite. This processor is also available in a fault tolerant version against Single Event Upset (SEU), and can be implemented on a radiation hardened Atmel FPGA. Plans to upgrade to the Leon 3 processor family are under way.

One of the most important features in the design is the flexibility of communication of the processor with the rest of the system. We considered four different kind of communication media:

1. A peripheral implementing the telemetry bus standard defined in Sec. 8.2 is available. Messages from or to the on-board computers and communication boards can be transmitted on this bus.
2. Direct communication using standard peripherals: high priority or low latency messages may require a dedicated path other than telemetry bus. This is, for instance, the case when the payload processor should have a direct high speed connection with the communication board that drives the antennas. Most of these peripherals are available as IP cores.
3. Digital and Analog general purpose input/outputs can be used to interface with some payload instruments. The number and type of I/Os is configurable at system generation time. External AD and DA converters with analog multiplexers are used for analog signals processing. Digital inputs can be configured to generate interrupts, and a set of encoders to count events can be implemented on the FPGA. A simple device driver to access these I/Os is provided in Linux.
4. Debugging of the payload board can be performed using a set of dedicated communication links, such as an RS232 console, a JTAG connector for step-by-step execution, and optionally an Ethernet device. If computational power is sufficient, a built-in web server running on the processor can be used to monitor the system during earth bound test sessions.

Additionally, we provide support for video decoders if standard video sources need to be connected to the system. Any peripheral or component that is not strictly necessary can be omitted, thus saving precious power.

Fault tolerance can be both at the hardware and at the software level. The processor can be replicated using Triple Modular Redundancy (TRM) (Mitra et al. 2000) and a voter is used to protect against Single Event Upsets, which can occur both in the internal processor memory and in the FPGA configuration memory. Watchdog timers, on the other hand, allow to stop a faulty execution and restart the system. Upon restart, the FPGA configuration memory is also checked against possible upsets, and reloaded with correct code if necessary.

Software can also be replicated, especially the executable code. Error detection techniques, such as CRC, can be used to identify wrong copies and correct them. Data protection is more application dependent, as some errors maybe neglected, while others might have catastrophic effects, and thus should be taken into consideration.

## 12. Testing

The idea of using modular and flexible intelligent tiles also aims at a cheap testing strategy. Each manufactured tile shall be tested mechanically, functionally, and under radiation. Being tiles small, simple and standardized allows reducing the cost of testing significantly by: i) reusing the equipment which has to be developed *ad-hoc* for testing (both mechanical and functional and electrical); ii) using smaller equipment (especially for mechanical and radiation tests).

As standard testing strategies are often redundant for nano- and micro-satellites, we are also studying and developing appropriate testing strategies aimed at cheap satellites, as part of the ARaMiS project.

On the other hand, since each module contains a micro-controller, automatic test functions (e.g., BIST) can be embedded without any extra hardware, in order to simplify verification of correct satellite operation.

## 13. Conclusions

The paper presents the design issues of two small satellites, developed with different approaches.
The first one (PiCPoT) was a "custom" design: a unique set of specifications identifies size and functions of the satellite, and the design was developed following these "hard" constraints.
The second one (ARaMiS) rather than a single satellite represents an architecture: a fully modular approach brings to the development of fully reusable units (the "tiles"), which allows to build a variety of small satellites with different size, weight, and functionality.
This work is supported by "Regione Piemonte" within the MAESS-2006 project.
Both projects share some common issues, such as

- emphasis on education: the design teams include students at various levels (master and Ph.D.), and specific units or subsystems are developed within final projects;
- use of commercial devices (COTS), with proper selection and design choices to achieve the required reliability.

## 14. References

Caldera, M.; Corpino, S.; Masoero, M.; Viola, N. (2005). Analisi termica di un nano-satellite universitario, *Atti del XVII congresso nazionale AIDAA*, September 2005, Volterra, PI.

Del Corso, D.; Passerone, C.; Reyneri, L.M.; Sansoè, C.; Borri, M.; Speretta, S.; Tranchero, M. (2007). Architecture of a Small Low-Cost Satellite, *Proceedings of 10th Euromicro Conference on Digital System Design*, pp. 428−431, August 2007, Lubek, Germany.

Gray, P.R.; Hurst, P.J.; Lewis, S.H.; Meyer, R.G. (2001). *Analysis and design of analog integrated circuits,* John Wiley & Sons, Inc., ISBN 9971-51-354-4, New York.

Heidt, H.; Puig-Suari, J.; Moore, S.; Nakasuka, S.; Twiggs, R. J. (2001). CubeSat: A New Generation of Picosatellite, *Proceedings of the 14th Annual AIAA/USU Conference on Small Satellites,* August 2001, Logan, UT, USA.

Mitra, S.; Saxena, N.; Mc Cluskey, E. (2000). Common-Mode Failures in Redundant VLSI Systems: a Survey, *IEEE Transactions on Reliability*, vol. 49, no. 3, September 2000.

Miyahira T.; G. Swift, G.; The use of advanced flash memories in satellite applications: Radiation responses and mitigation choices, *Proceedings of the Military and Aerospace Applications of Programmable Devices and Technologies Conference*, September 1998.

Nguyen, D.N.; Scheick, L.Z. (2001). TID testing of ferroelectric nonvolatile RAM, *Proceedings of Radiation Effects Data Workshop*, pp. 56−61, July 2001.

Orefice, M.; Dassano, G. (2007). The PICPOT Satellite Antenna Systems, *Proceedings of the IEEE AP-S International Symposium*, pp.3029-3032, June 10-15 2007, Digest, Honolulu.

Passerone, C.; Tranchero, M.; Speretta, S.; Reyneri, L.M.; Sansoè, C.; Del Corso, D. (2008). Design Solutions for a University Nano-satellite, *Proceedings of the Aerospace Conference*, pp. 1-13, March 2008, Big Sky, MT.

Pradhan, D.K. (1996). *Fault-Tolerant Computer System Design.* Prentice Hall, 1996, Englewood Cliffs, NJ, USA.

Speretta, S.; Reyneri, L.M.; Sansoè, C.; Tranchero, M.; Passerone, C.; Del Corso, D. (2007). Modular Architecture for Satellites, *Proceedings of the 58th International Astronautical Congress*, September 2007, Hyderabad, India.

Ziegler, J.F.; *et al.* IBM experiments in soft fails in computer electronics, *IBM Journal of Research and Development*, vol. 40, no. 1, January 1996.

# 10

# Robot Mobility Systems for Planetary Surface Exploration – State-of-the-Art and Future Outlook: A Literature Survey

Aravind Seeni, Bernd Schäfer and Gerd Hirzinger
*Institute of Robotics and Mechatronics*
*German Aerospace Center (DLR)*
*Germany*

## 1. Introduction

Mobile robots have been an essential element of space exploration to perform science on distant lunar and planetary objects. Past orbiter missions to Moon have shown evidence on the presence of ice inside the permanently shadowed polar areas. Similarly, Mars holds clues of life in the distant past. While data sent by orbiters provides a wealth of information and gives rise to new speculations and theories, in-situ science data from mobile rovers and landers is essential to validate or confirm them. "Mobility" is a vital element for a space missions due to valuable science return potential from different sites as opposed to static landers. Over the years, technology development has given rise to numerous mobile systems. Some of the systems are spin-offs from terrestrial applications like automobiles that use wheels, military tanks that use tracks and aerial balloons. Others have been developed purely for space application like hoppers and hybrid systems. Since 1970s, twelve surface missions have reported using mobile robots. Most of them used wheels as their mobility element for locomotion. Since surface space-exploration of planetary objects across the solar system has gained increasing importance in recent years, it is important to understand the state-of-the-art in mobile robotics and the technology development of the mobility system in particular. This need will be addressed in this chapter.

## 2. Aim of study

With advancement in research and technology, many mobile systems have been developed with different geometries, sizes, and configurations. These systems share different performance qualities under certain operational condition. There are limitations with every system in terms of certain mobility performance such as slope climb, obstacle traverse, speed, power consumption rate etc. Each system can be chosen depending on the type of mission, operation, expected science return, reliability and so on.
So far, many robotic vehicles have been designed and launched to different lunar and planetary bodies. They have been designed to operate remotely either through independent communication links or human-operated (Table 1). Over the years, technology advancements have been made that resulted in development of new mobility systems, due

to ever-growing science requirements. Several such robot mobility systems have been reported in literature (Fig. 1). Here, we try to systematically classify all these systems depending on the type of mobility, as follows:

1. Aerial systems
2. Sliding systems
3. Rolling systems
4. Wheel-enabled systems
5. Leg-enabled systems
6. Track-enabled systems
7. Hoppers
8. Hybrid systems

| Robot Name | Mission | Launch Year | Body | Country | Mobility |
|---|---|---|---|---|---|
| Lunokhod 1 | Luna 17 | 10 Nov 1970 | Moon | Soviet Union | Wheels |
| Prop-M | Mars 2 | 19 May 1971 | Mars | Soviet Union | Skids |
| Prop-M | Mars 3 | 28 May 1971 | Mars | Soviet Union | Skids |
| Lunar Roving Vehicle | Apollo 15 | 26 Jul 1971 | Moon | USA | Wheels |
| Lunar Roving Vehicle | Apollo 16 | 16 Apr 1972 | Moon | USA | Wheels |
| Lunar Roving Vehicle | Apollo 17 | 07 Dec 1972 | Moon | USA | Wheels |
| Lunokhod 2 | Luna 21 | 08 Jan 1973 | Moon | Soviet Union | Wheels |
| -NA- | Phobos 2 | 12 Jul 1988 | Phobos | Soviet Union | Hopper |
| Sojourner | Mars Pathfinder | 04 Dec 1996 | Mars | USA | Wheels |
| MINERVA | Hayabusa | 09 May 2003 | Asteroid Itokawa | Japan | Hopper |
| Spirit | MER-A | 10 Jun 2003 | Mars | USA | Wheels |
| Opportunity | MER-B | 07 Jul 2003 | Mars | USA | Wheels |

Table 1. All reported space missions and mobile robots launched from the year 1970 to present

Aerial systems are particularly useful for "global" exploration cases. It can be useful in operations where there is presence of atmosphere. Other systems, that are listed above, while operated remotely in one, two, or few numbers would enable "local" exploration of the landing site. The maximum exploration range is limited to few tens of kilometres on the surface. But when deployed as a group or swarm, a "regional" exploration scale of a few square kilometers within reasonable mission duration is also possible. The geographic coverage can still be increased when deployed across different sites on the surface through a second auxiliary system (e.g. aerial).

In this Chapter, the result of a literature survey of mobility systems development in space agencies, research institutes and universities from the following space-faring nations - USA, Japan, Europe, Canada and Russia is presented. The Chapter aims to give an overview of already accomplished and currently ongoing research in mobile robot development

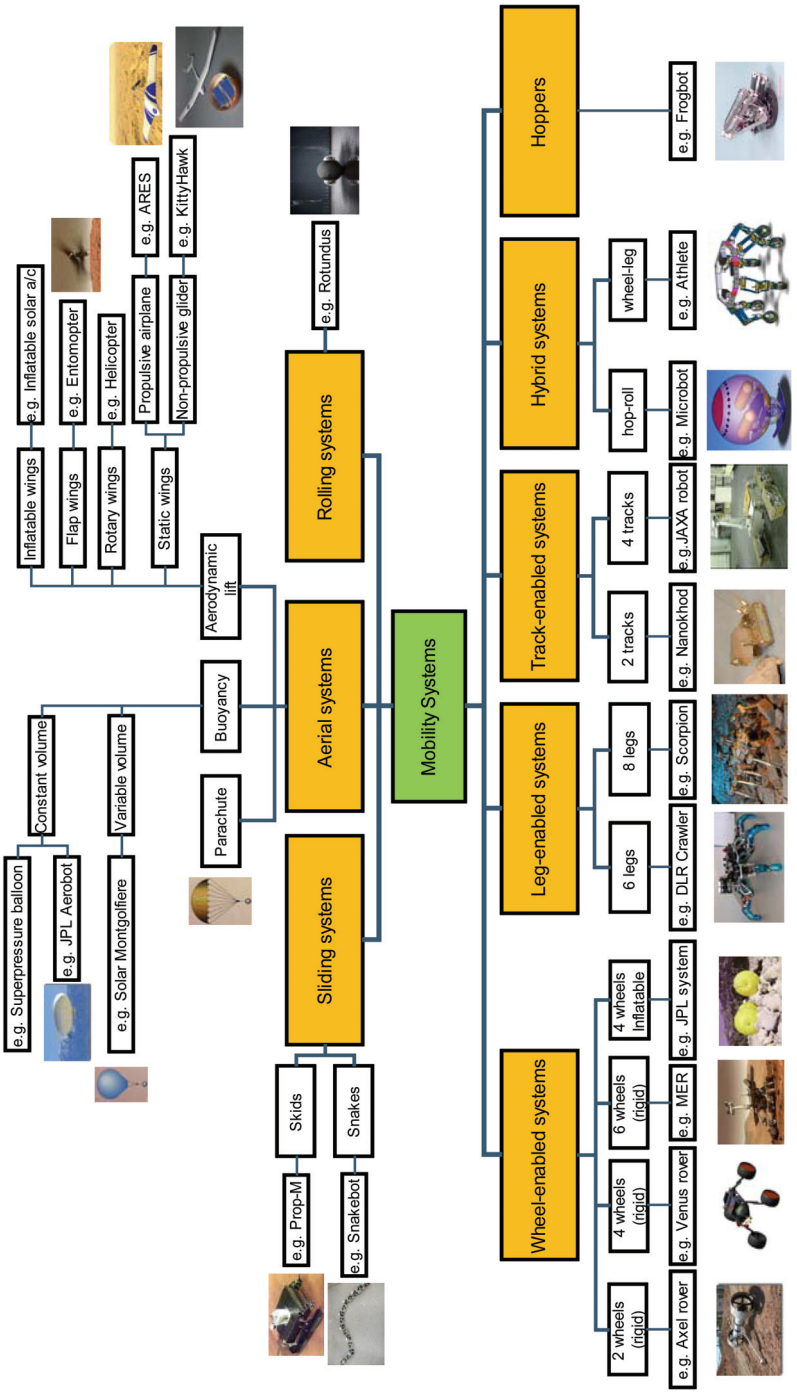



Fig. 1. Summary of past and present mobility systems technology development categorized based on mobility type

categorically and provide a brief description of the technical aspects of some of the mobility systems. In short, the study aims to answer the following questions:

1. What kind of robot mobility systems are developed or presently ongoing development for surface exploration?
2. What kind of mission-oriented applications would they fit in?
3. What are their pros and cons in terms of performance and reliability? What are their present readiness levels?

There are numerous ways to achieve mobility on an extra-terrestrial surface. This has been illustrated in Fig. 1. The chapter cannot provide a description of all the systems available, since such a scope is very exhaustive. Instead six systems are chosen based on their importance as well as literature coverage and then described subsequently with examples. The six systems chosen are wheel-enabled, leg-enabled, track-enabled, hoppers, wheel-leg hybrid and hop-roll hybrid systems. In the end, a comparative assessment of the six systems is also performed for different metrics qualitatively.

## 3. Wheel enabled systems

Wheels are commonly used for years to enable motion in terrestrial applications as in many on- and off-road vehicles. Wheel enabled systems or rovers can be categorized based on the number of wheels. Here two systems, eight and six-wheel rovers are discussed with examples.

### 3.1 Eight-wheel system

The well-demonstrated technology was first employed for space operations during the Lunokhod-1 and 2 missions in 1970 and 1973 respectively, to the lunar surface by former Soviet Union. Lunokhod-1 (Fig. 2) rover's total mass was 750 kg and its "undercarriage" or suspension system alone weighed 105 kg (Kermurjian, 1990). The suspension consisted of eight active wheels that could enable the rover move longitudinally at speeds between 0.8 and 2 km/hr. The wheels were not designed to swivel and rover turning was enabled by varying the wheels' rotational velocity at the left and right suspension. The suspension enabled Lunokhod-1 and 2 cover a distance of 10.54 and 37 km respectively on the surface during their operational lives.

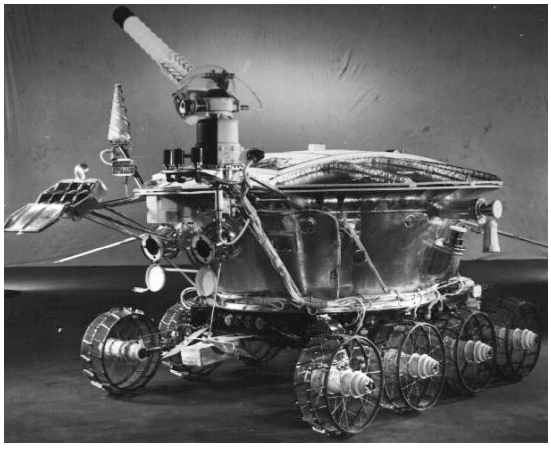


Fig. 2. Lunokhod 1 rover (Image Courtesy: Kermurjian, 1990)

The wheels of the rover were rigid with perforated, cleated rims. The rover also had a ninth free-rolling wheel that was used to monitor rover odometry and wheel slippage. Slip was estimated by comparing the number of revolutions of free-rolling and main drive wheels.

### 3.2 Six-wheel system

Eight-wheel enabled mobility systems as in Lunokhod are very heavy and are no longer being developed. In recent years, newer suspension concepts for rovers have emerged and are continuously being developed that promise high mobility performance without considerably impacting mass and power consumption constraints. The system typically consists of chassis, wheels, actuators, sensors, electronics, and steering mechanism. The notable one is the six wheel enabled rocker-bogie suspension system (Fig. 3) developed at NASA Jet Propulsion Laboratory and California Institute of Technology (Lindemann et al., 2005).

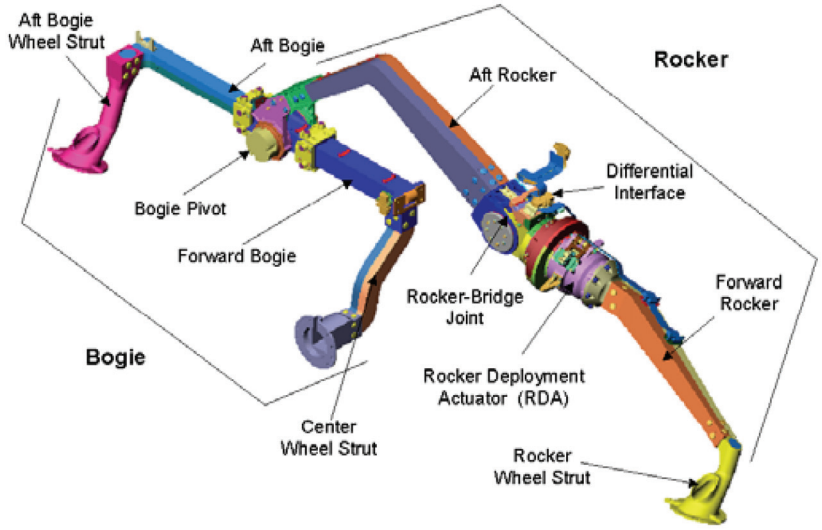



Fig. 3. Rocker-bogie suspension system (Image Courtesy: Lindemann et al., 2005)

The rocker-bogie suspension system uses a 6x6x4 wheel formula[1]. All the six wheels are independently actuated by DC motors and the front and rear wheels are steered by additional, identical DC motors. An internal differential mechanism in the rover's body connects the left and right rocker-bogie assemblies. The suspension system enables the vehicle to passively keep all wheels in contact with the ground even while traveling on severely uneven terrain. Typically, a vehicle with a rocker-bogie suspension system is capable of traversing obstacles with a height of at least the rover's wheel diameter. Tests have confirmed that the suspension system enables a rover to climb obstacles up to 1.5 times

[1] Wheel formula specifies the type of locomotion configuration associated with a rover. The formula is usually written in the form as follows:
*Wheel formula = Total no. of wheels × No. of actuated wheels × No. of steerable actuated wheels*

its diameter. The system also helps the rover in climbing loose-soil surfaces, during which the wheel's average pressure will be equilibrated. This is important for motion on soft terrain as excessive ground pressure results in wheel sinkage.

A limitation to the system as observed on Mars is excessive wheel slippage which results in total rover immobility. For example, Spirit and Opportunity rovers have remained trapped in loose soil for weeks. At the time of writing, Spirit was stuck in soft soil and unable to be recovered since. Such states also arise due to locomotion design constraints imposed by several parameters while designing the suspension system. One uncontrollable parameter is unknown local soil properties. Other constraints include lander's low stowage volume and lack of high-friction cleats on wheels due to possibility of entanglement with landing airbag during rover egress (Lindemann et al., 2006).

The rocker-bogie system is geometrically scaled to develop rovers of different mass class. Until now, JPL has used it in the Rocky7 testbed, Mars Pathfinder mission rover Sojourner, Mars Exploration Rovers (MER) Spirit (Fig. 4) and Opportunity. It is also being used to develop the 800 kg Curiosity (MSL) rover (Fig. 5) slated to be launched in 2016 (NASA/JPL/Caltech website, 2005).

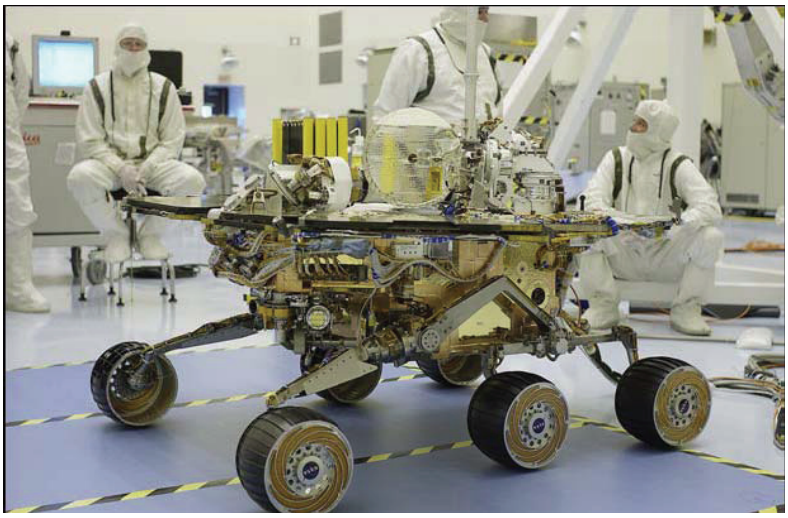


Fig. 4. Flight rover Spirit of Mars Exploration Rover missions (Image Courtesy: NASA/JPL/Caltech)

Enabling motion with wheels are not only limited to JPL rover applications. Studies from other space agencies such as ESA, JAXA also propose wheels (Roe et al., 2008; Kubota et al., 2005). This was evident during the early project phases of ExoMars and SELENE-II missions. However, unlike the rocker-bogie, these rovers differ in their system configuration. ExoMars employs longitudinal and traverse bogies (Roe et al., 2008). ExoMars is now planned to be launched in 2016. Presently, it is ongoing extensive modeling and simulation to analyse the terramechanics of the rover's mobility system at the authors' institute [Gibbesch et al., 2009]. SELENE-II uses the so-called Pentad Grade Assist Suspension or PEGASUS suspension system for their Micro5 testbed rover (Kubota et al., 2005).

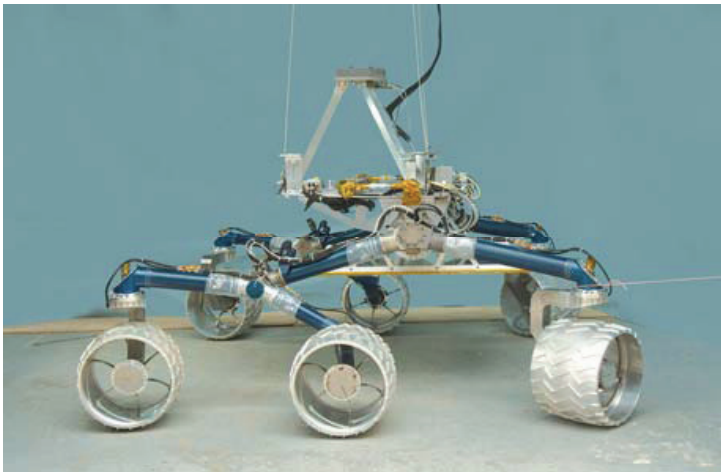


Fig. 5. Mobility model of Curiosity (MSL) rover (Image Courtesy: NASA/JPL/Caltech)

## 4. Track-enabled systems

Track enabled robots use crawl units or tracks that are commonly used for terrestrial mobile applications like military tanks and automobiles. These tracks are especially suited to motion on difficult terrain. Currently, track enabled systems are being considered for extra-terrestrial surface exploration also. Two types of such systems – Nanokhod and JAXA Track robot - are discussed below.

### 4.1 Twin-track system

The Nanokhod is a miniaturized track enabled robot (Fig. 6) that was developed based on Russian technology. Initially, it was foreseen to be launched with Beagle-1 Lander during ESA's 2003 Mars Express mission and was cancelled. The track system was then developed for BepiColombo mission to Mercury, which unfortunately was cancelled again. Since then, it is being immensely studied for lunar and other planetary missions (Klinker et al., 2005).

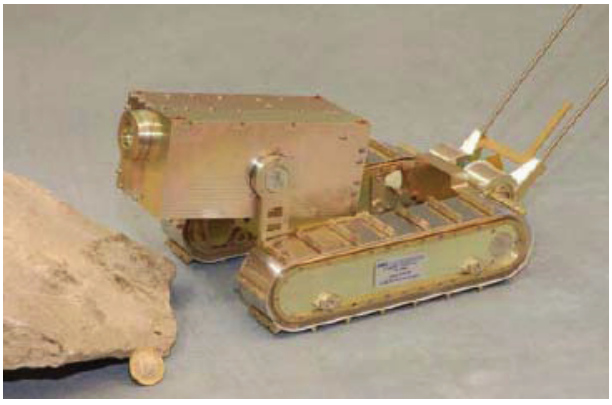


Fig. 6. Nanokhod dual-track system (Image Courtesy: Klinker, 2007)

The tracker consists of two "caterpillar" track units, a tether unit, and a payload cabin (Fig. 7). The caterpillar tracks are driven by four internal drive units. The drive units consist of a stepper motor attached to a 64:1 planetary gear in front of a clown and pinion stage. The output stage is a miniaturized harmonic drive whose input is coupled directly to the crown gear. The output is obtained through a flex spline and circular spline for track and arm drives respectively. The tracks are surrounded by protective walls.

Fig. 7. Nanokhod's miniaturized components (Image Courtesy: Klinker, 2007)

Nanokhod can move at a low speed of 5 m/hr (~0.14 cm/s). It is capable of climbing obstacles at least 10 cm high and trenched 10 cm wide sized (Klinker et al., 2007). The rover is small-sized. Electrical power is designed to be fed by a stationary platform (e.g. Lander) through tethers. Autonomous localization and vision-based navigation is performed based on model-based techniques that make use of Lander's fixed stereo camera images [Steinmetz et al., 2001].

### 4.2 Four-track system

The Advanced Space Technology research group in JAXA proposes a tracker for SELENE-II lunar mission (Wakabayashi et al., 2006). The mobility system for this tracker consists of four caterpillar crawl units with two on both sides (Fig. 8). Mesh structures are used in the crawl units. The units' links are meshed and equipped with L-shaped small lugs to increase traction. Metal mesh belts are used in each crawl unit.

Fig. 8. JAXA's four track system (Image Courtesy: Wakabayashi, 2006)

The nominal contact area of one unit is 100 $cm^2$. The crawl units (Fig. 9) take advantage of the compaction of regolith for mobility. The strategy of using mesh structures instead of track shoes enables reduction of cost, number of parts, and malfunctions.

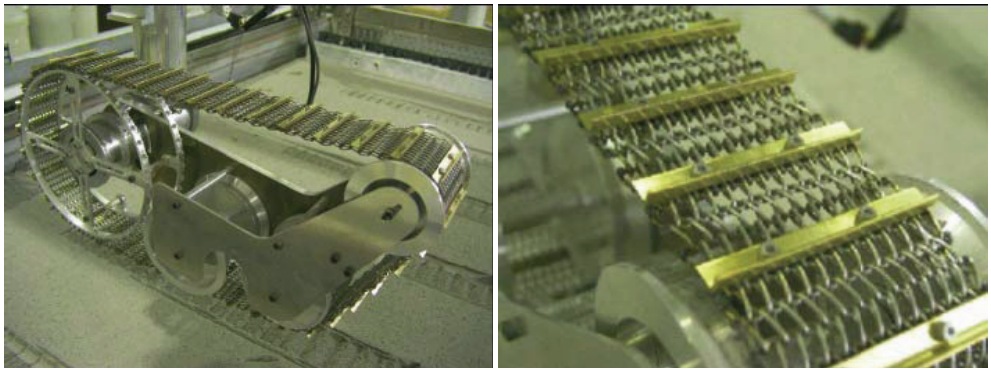


Fig. 9. Close look of a track (left); track mesh (right) (Image Courtesy: Wakabayashi, 2006)

## 5. Leg-enabled systems

Biological designs and neurobiological controls inspire robot development and have given rise to several robots. Biology offers working examples of robust and sustainable motion behavior. A few of the biology inspired designs that are leg enabled are discussed below.

### 5.1 Eight-leg system

SCORPION is an octapod (eight-legged), outdoor walking robot developed by German Research Center for Artificial Intelligence - Bremen, Defense Advanced Research Projects Agency and NASA (Fig. 10). It is designed to walk in dangerous, highly unstructured, rough terrain where mobility is crucial. The walker is 65 cm long and has a minimum height of 52 cm. In an M-shaped walking configuration, it is 40 cm wide. In a stretched-leg configuration, it is 35 cm high. The legs provide a ground clearance of 28 cm to the body. Each leg has three degrees of freedom (DOF). The legs consist of a thoracic joint for protraction and retraction, a basal joint for elevation and depression, and a distal joint for extension and flexion of the leg. The joints are actuated using 24 V, 6 W DC motors with high gear transmission ratio. The leg also features a spring element in the distal segment to reduce the mechanical stress (damping) and for measuring the ground contact force by an integrated linear potentiometer (Spenneberg & Kirchner, 2002).

The robot is powered by 3.0 Ah batteries. The robot is integrated with proprioceptive sensors namely motor encoders, Hall Effect motor current sensors (for each drive motor), and power management sensors. Motor encoders measure the relative joint angle whereas power management sensors measure current battery voltage and power drain.

### 5.2 Six-leg system

The DLR Walker is a Hexapod (six-legged) robot developed at the Institute of Robotics and Mechatronics of German Aerospace Center (Fig. 11). The legs of the robot are the fingers developed for DLR Hand-II (Görner, 2007; Borst et al., 2002). These legs/fingers consist of

Fig. 10. 8-legged system *SCORPION* (Image Courtesy: University of Bremen, DARPA, NASA)

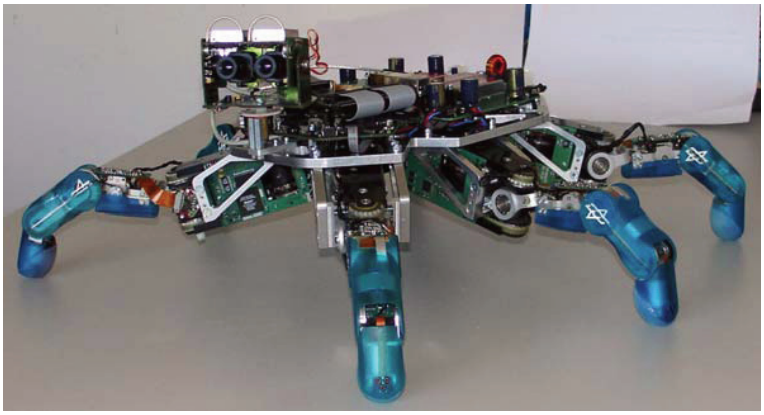


Fig. 11. DLR Walker robot (Image Courtesy: Görner, 2007)

three independent actuated joints that enable omni-directional walking. The base-joint actuating system, in each leg, essentially consists of brushless DC motors, tooth belts, harmonic drive gears and bevel gears. The base-joint is connected to the main body. It is of differential bevel gear type and movable in two DOF. The actuating system in the medial joint has relatively less power than the base joint and is designed to meet the conditions in the base joint when the leg is in stretched configuration. The third joint near the tip is passive and coupled to medial joint actuating system.

Each leg has three joint-position, three joint-torque, one forcetorque, three motor-torque, and six temperature sensors. The leg tip has a force-torque sensor that measures forces in six dimensions. Due to the force/torque sensing at the gear output it is most favorable to apply advanced so-called soft control algorithms. The most promising concept is to apply virtual

springs/damping forces to each joint in order to allow any kind of softly forced motion. Different walking gaits have been realized to overcome any complexly structured terrain.
Presently, this DLR walker is primarily used to study advanced navigation algorithms that autonomously map the 3D environment, localize itself and determine the next safe trajectory in this mapped environment. This is made possible by using the in-house developed Semi-Global Matching Method algorithm [Hirschmüller, 2008].

## 6. Hoppers

Hopping robots or hoppers were proposed as a cost-effective solution for planetary surface mobility to other systems. Unlike for other systems, where every small rock or hole on the surface is an obstacle to motion, hoppers are designed to hop over these obstacles. The importance given to design of these systems is simplicity and not high mobility performance. Hopping is performed at a certain velocity on a chosen trajectory without the need for accuracy of the destination. When used in swarms or groups of large numbers, these systems are capable of exploring a wide area.
An example of a hopping robot being designed is the Micro-hopper being developed by Canadian Space Agency (Dupuis et al., 2005). Hopping is enabled by a Shape Memory Alloy (SMA) actuator that utilizes heat from the Sun to store energy for the hop. SMA actuator elongates during the variation in day-night temperature. As a result, the current design reported in (Dupuis et al., 2006), allows one jump per day on Mars. A cylindrical structure with scissors is used to transfer energy for the hop. The hop is performed at a fixed angle of attack. The maximum horizontal distance covered in one hop is 3 m. The geometry of the whole robot is designed to be a regular tetrahedron as shown in Fig. 12. This allows landing in any direction and recovery after the hop.

Fig. 12. Illustration of CSA's Micro-hopper (Image Courtesy: Dupuis et al., 2005)

## 7. Hybrid systems

The wheeled, tracked, and legged locomotion systems discussed earlier have their own advantages and disadvantages. For example, while wheels are capable of higher speeds than trackers and walkers on a flat terrain, it is relatively less capable of traversing obstacles as the other two. Hybrid robots consist of a combination of two mobility concepts that shares the advantages of both. Two kinds of such systems are discussed in the following sub-chapters.

### 7.1 Wheel-leg hybrid

Wheeled-leg bybrids have the advantage of higher mobility of walkers combined with the energy efficiency of rovers. The system as a whole can be designed to be highly modular, reusable, redundant, reconfigurable and with adequate margins. One such example is the DLR's hybrid concept that is proposed two have two mobility concepts in one: A six-wheeled rover system that carries a six-legged walking system [DLR Status Report, 2009].
The system can be favorably used to carry a small legged system on a flat terrain to a nonfriendly terrain where wheels cannot roll safely. In such circumstances, the legged system can be either carry the whole robot as shown in Fig. 13 (left) or dislodge from the top of the body and still carry the payload as in Fig. 13 (right). An example of the former scenario is using the main robot to the rim of a crater and then deploying the smaller robot to move down to the depths.

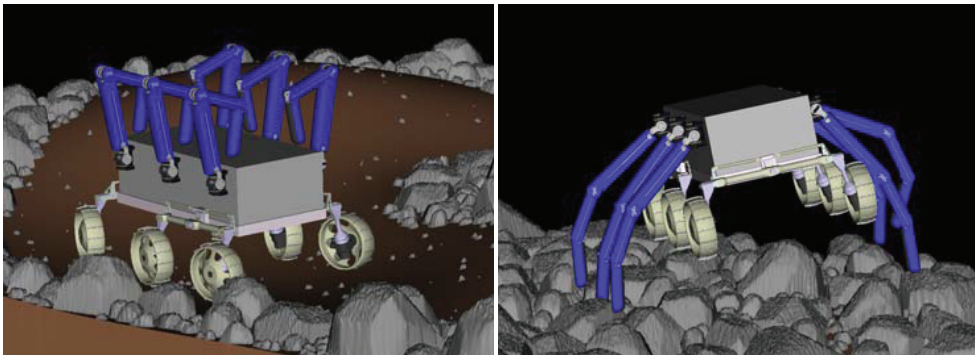


Fig. 13. Simulated DLR hybrid mobility concept in "wheel rolling" mode on flat terrain (left); "leg walking" mode on rugged terrain (right)

Wheels are used to roll over smooth terrain and legs for extreme terrain. The hybrid concept can also be used with unprecedented mobility capabilities. The concept is being realized in a hybrid lunar lander development called ATHLETE (Wilcox, 2007) by NASA (Fig. 14) primarily being developed to construct lunar outposts and assist astronauts (Morring, 2007).

### 7.2 Hop-roll hybrid

Hop-roll hybrid systems share the mobility concepts of both hopping and roll systems. The system can hop and the body has a ball-shaped design to aid rolling on inclined terrain. One such design called "Microbot" (Fig. 15) is proposed by Massachussetts Institute of Technology (Dubowsky et al., 2005). The system is designed to work with a swarm of other microbots to enable vast scale exploration and inter-communication possibilities between them (Fig. 16).

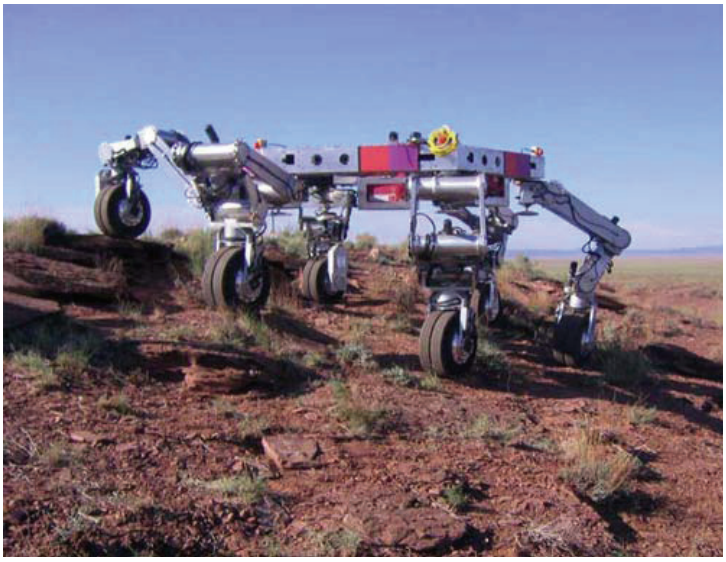


Fig. 14. Wheel-leg Hybrid system ATHLETE (Image Courtesy: NASA/JPL/Caltech)

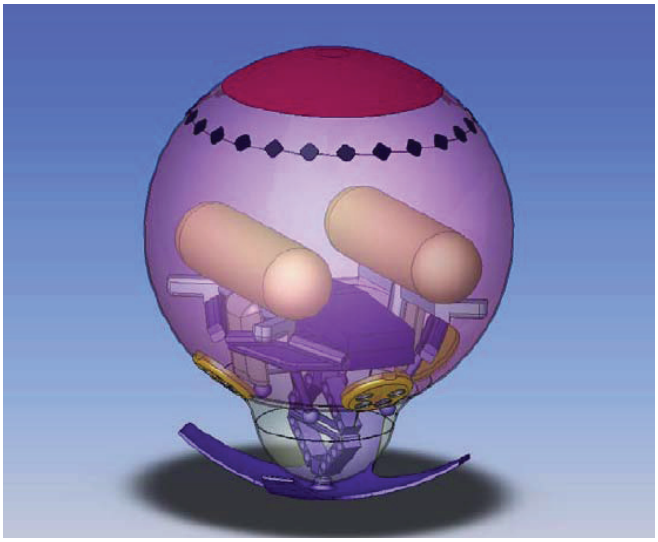


Fig. 15. Illustration of MIT proposed Microbot (Image Courtesy: Dubowsky et al., 2004)

Each microbot is proposed to be of low-mass (150 g), with a body diameter of 10 cm and has a horizontal distance hop of 1.5 m. The hop mobility is enabled by a bistable mechanism that is activated by dielectric elastomer actuators also called as Electro-active Polymer Muscle Actuators. Energy is transferred to the actuators over a few seconds or minutes, which are then continuously transferred by a short, high power stroke for a hop. Power is supplied by fuel cells. The design of the body, allows non-powered bouncing and rolling motion in slopes or in any inclined terrain.

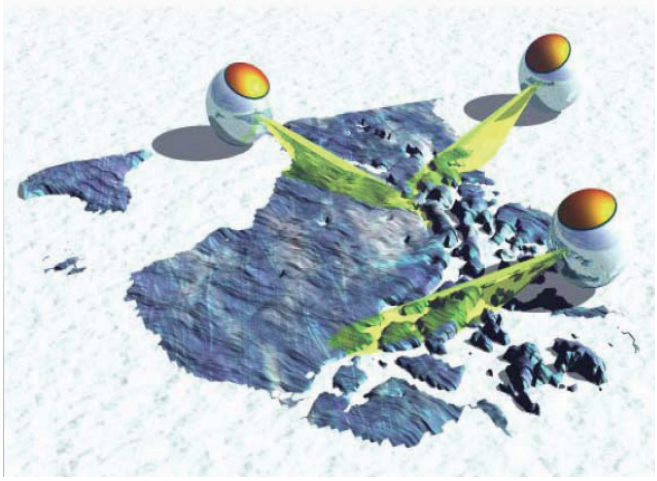


Fig. 16. Illustration of a swarm exploration scenario with microbots (Image Courtesy: Dubowsky et al., 2004)

## 8. Comparison of locomotion systems

A comparison of advantages and disadvantages of the different locomotion concepts will be treated hereafter. The comparison will focus on the aspects of mobility performance, reliability, energy consumption, obstacle negotiation and technical readiness levels. To compare locomotion systems comprehensively, experimental or simulation data based on operations in planetary terrain-like environment (planetary test-bed) is required. These experiments should be performed on similar test conditions. Such comparative data and subsequent comparative analysis of different locomotion concepts (wheel, leg, track enabled) are scarce in literature as it is influenced by many discrete parameters such as external dimensions, mass, number of wheels/legs, diameter of wheels etc. During this study, a step towards comparison of different locomotion concepts has been attempted qualitatively by inferring data from current state-of-the-art locomotion systems available in recent literature.

### 8.1 Metrics based comparison

The metrics are initially benchmarked with capabilities of a fictional locomotion system. These capabilities are then compared with the following 6 locomotion concepts discussed earlier – Wheel enabled system, track enabled system, leg enabled walker, hoppers, wheeled-leg hybrid, and hop-roll hybrid. The major strengths and weaknesses of each system were then marked qualitatively against different metrics. The selected metrics reflect the general performance of a robot's mobility based on state-of-the-art in present technology.

The metrics that are used and their description are as follows:

**Maximum speed capability** - Capability to move fast on a flat surface. This metric provides a relative measure to compare the distance reachable per day on a lunar or planetary surface.

| Metrics | Assumptions |
|---|---|
| Maximum speed capability | Capability to move on flat terrain at a speed of 30 cm/s (on Earth) |
| Obstacle traverse capability | Step climbing ability equivalent to mobile element size (wheel diameter, length of legs). |
| Slope climb capability | Slope climb ability of 30° inclination |
| Soil sinkage | Medium sinkage of mobile element during motion on soil |
| Mechanical simplicity | Moderate level of mechanical complexity |
| Mobile element redundancy | Ability to perform and continue with motion after malfunction of two wheel assembly, assuming the vehicle having six components |
| Energy consumption rates | Capability to move with less battery power consumed per unit distance |
| Payload Mass Fraction capacity | In the range of 8 -12% |
| Soil interaction | Body contact with soil during motion process is moderate |
| Technology readiness level | Breadboard validated in space environment, i.e. TRL 5 |

Table 2. Benchmark vehicle assumptions

**Obstacle traverse capability** - Capability to move over obstacles/boulders relative to their mobile element size (wheels, legs etc.)
**Slope climb capability** - Ability to climb slopes covered with soft soil smoothly without excessive loads on one or more particular mobile elements (rear wheels, legs etc.)
**Soil sinkage** - Ability to move on soft soil over flat surface by having the least contact pressure without large slip, minimal sinkage and mobility resistance (depends on vehicle's mass)
**Mechanical simplicity** - Less complexity of the locomotion subsystem with regard to number of parts, linkages etc.; less moving parts
**Mobile element redundancy** - Capacity of the vehicle to continue on the mission objective in case of failure of primary mobile elements (wheels, tracks, or legs)
**Energy efficiency** - Ability to move with low power requirement/unit. In the case of walkers, the power consumed while raising and lowering of legs is added along with forward motion energy rates
**Payload mass fraction capacity –** Mass of payload to total robot mass ratio. Ability to carry significant science payload mass relative to total mass
**Soil interaction** - Ability to move by adequate ground clearance to corrosive planetary soil without unwanted interaction. This influences the capacity to withstanding hostile environment and long term effects (tolerance to corrosiveness).
**Technology readiness level** - Technology maturity and demonstration

### 8.2 Criteria

The criteria set for comparing the mobility systems are set in five different grade scales in Table 3. The assumptions of the benchmark vehicle (BMV) described in Table 2 are comparable with the third quality grade *** and given as BMV.

| **Metrics** | * | ** | *** | **** | ***** |
|---|---|---|---|---|---|
| Maximum speed capability | lower than 50% | up to 50% low | BMV | up to 50% high | higher than 50% |
| Obstacle traverse capability | < 20 cm | up to 20 cm | BMV | up to 30 cm | > 30 cm |
| Slope climb capability | < 15° | up to 25° | BMV | up to 35° | >35° |
| Soil sinkage | Very high | High | BMV | Low | Very low |
| Mechanical simplicity | Very low | Low | BMV | High | Very high |
| Mobile element redundancy | No redundancy | Low | BMV | High | Very high |
| Energy consumption rates | Very high | High | BMV | Low | Very less |
| Payload mass fraction capacity | < 5% | 5 – 8% | BMV | 12 – 15% | > 15% |
| Soil interaction | Very high | High | BMV | Low | Very low |
| Technology readiness level | TRL 1 or 2 | TRL 3 or 4 | BMV | TRL 6 or 7 | TRL 8 or 9 |

Table 3. Criteria for evaluating metrics

### 8.3 Result and discussion

Based on information reported in literature, the present state-of-the-art is understood and the metrics are graded accordingly. The result is given in Table 4.

Some advantages and disadvantages of the compared systems are summarized in Table 5 in general. In addition, the following points are worthy to note: A wheel enabled rover can roll at better speed on flat terrain and has moderate sinkage characteristics on soft terrain. A present-technology track enabled system has better slope climbing and obstacle traversing capabilities, but at the expense of high energy consumption. Generally, current systems for space missions are quite mechanically complex. Wheel enabled systems show considerable difficulties while climbing slopes over 15°, since the slip ratio drops suddenly. Track systems have better slope climbing and obstacle traversing capabilities than rovers.

Not many leg enabled systems developed for space applications are existent today, although a few vehicles are currently being developed in the US and Europe. These systems are highly stable while moving over obstacles, possess better mobility during downhill motion due to their Center of Mass position. However, it has the disadvantage of high power consumption, since power is needed for both lifting and forward motion of legs. Conversely, a wheeled rover is energy efficient.

Wheel enabled systems also possess good reliability and redundancy. Assume three six-wheel, six-track, and six-leg systems powered by motors individually at their mobility joints. Three malfunction wheels can still offer sufficient locomotion capacity in a wheel-enabled system as seen with MER. A track system has low reliability due to possibility of jamming of tracks and track units malfunction. Failure of one of the tracks may

| Metrics | Wheel-enabled system | Track-enabled system | Leg-enabled system | Hoppers | Wheel-leg hybrid | Hop-roll hybrid |
|---|---|---|---|---|---|---|
| Maximum speed capability | *** | *** | *** | ***** | ***** | ***** |
| Obstacle traverse capability | *** | **** | ***** | ***** | ***** | ***** |
| Slope climb capability | ** | *** | *** | ***** | ***** | ***** |
| Soil sinkage | *** | **** | ** | ***** | *** | ***** |
| Mechanical simplicity | *** | ** | ** | **** | * | **** |
| Mobile element redundancy | **** | * | **** | * | ***** | * |
| Energy consumption rates | **** | ** | * | **** | * | ***** |
| Payload mass fraction capacity | *** | *** | ** | **** | ***** | **** |
| Soil interaction | *** | ** | **** | ** | *** | * |
| Technology readiness level | ***** | *** | ** | ** | *** | ** |

Table 4. Qualitative mobility systems comparison

lead to partial or total immobility. A leg enabled system can still move provided the failure no longer affects the motors and stow the malfunctioned legs over the body to avert contact with surface. The jamming of tracks in track systems can be averted by low-speed operation. Such compromises are less required for a wheel system. In any case, this argument validity depends on which motors powering the wheels, tracks and legs are malfunctioned. Failure of all three motor units on one particular side, will render any system go totally immobile.

Hoppers apply an easy, simple method of achieving mobility on low-gravity surfaces. However, the impact during every landing is a source of concern, since this may result in loss or damage to the entire system. Such is the case with hop-roll hybrids as well. An additional mobility advantage is that hop-roll systems can use their round body to bounce or roll easily in declined slopes without any power consumption.

The operational speed of a robotic vehicle is usually limited in the range of 10 cm/s due to safety concerns. It is limited by the type of gears used in the motor and power availability.

| System | Advantages | Disadvantages |
|---|---|---|
| Wheels | • Better speed in even terrain<br>• Simple and m ature technology<br>• Adequate redundancy (mobility)<br>• Payload weight-to- mechanism weight ratio high<br>• Relatively low power consumption rates and energy efficient | • Relatively low slope climb capacity due to wheel slippage<br>• Obstacle traverse capability relatively less compared to other concepts |
| Tracks | • Good terrain capability<br>• Technology well understood in terrestrial applications<br>• Better traction capability on loose soil • Handles large hinders, small holes, ditches better<br>• Good payload capacity | • Inefficient due to friction of tracks<br>• Low speed operation<br>• Slip turning and friction<br>• Low redundancy, jamming of parts and prone to failure |
| Legs | • Highly adapted to uneven terrain and hence better obstacle and slope traverse capability | • Mechanically complex<br>• Control of walking is complex<br>• Slow mobility<br>• Impact after each step<br>• Poor payload weight-to-mechanism weight ratio |
| Hoppers | • Better obstacle traverse capabilities<br>• If power availability is flexible, can enable large scale exploration due to better speed | • Impact during landing after hopping had large risk of failure |
| Hybrids | • Shares the advantages of two locomotion concepts<br>• Miniaturized hybrid hop-roll systems operating as swarm, enables exploration of larger area in a short time | • More complexity<br>• Low technology maturity |

Table 5. Advantages and disadvantages of mobility systems

The size of the vehicle is also a crucial factor while considering the mission scenario. It can be said that a mini-rover in the mass range of 30-100 kg is capable of accomplishing many science tasks by accommodating more integrated instruments and payloads. Also minirovers are capable of generating enough power for surviving the entire mission. Microrovers (5-30 kg) and nano-rovers (<5 kg) can be better suited for accomplishing specific mission objectives within limited range and power availability. These rovers are not capable of accomplishing a wide range of objectives like mini-rovers. However, in the case of nanorovers, there are various flexibilities possible in choosing a mission. Nano-rovers may not be independent in operation and deployed from Lander through tethers.

## 9. Conclusion

Trends in mobility technology are expected to have a major impact on surface exploration with robots in the future. Until the near future, the current trend of sending large payload mass using a single, large vehicle would be followed as can be seen in the case of Curiosity rover development that is bound for Mars. However, like in the case of earth-observation exploration, the increasing development of small satellites has made a major impact. A similar trend is expected in the surface robot development also. The current low to medium TRL of some mobility technology suggests that the present approach would change course to "lighter, cheaper and faster" missions. Some papers report exploration of Mars would follow a "paradigm shift", where swarm of micro robots would be launched with or without a mothership platform to deploy smaller ones in different locations across a planet. Also tethered exploration of dangerous slopes using parent-slave systems exploration strategy, where robots with conventional mobility like MER cannot access, is also expected. New missions would not only be designed to access multiple, easy locations but difficult and risky sites like deep mars "caves" where direct communication with an orbiter or lander is unfeasible. In such cases, the design of the robot's body as a whole to aid mobility as well as inter-robot communication would play a crucial role. The above mission scenarios are just a few cases, where currently proposed technologies provide a platform for exciting missions. In any case, it can be said that all missions would also follow a "high reliability, high science return and low cost" approach.

## 10. References

Borst, Ch., Fischer, M., & Hirzinger, G. (2002), Calculating hand configurations for precision and pinch grasps, *Proceedings of the 2002 IEEE/RSJ, International Conference on Intelligent Robots and Systems*, October 2002, Lausanne, Switzerland.

DLR Status Report (2009), Status Report: Scientific Results 2005-2008, Institute of Robotics and Mechatronics, Smagt, P.v.d. (ed.).

Dupuis, E., Montminy, S., & Allard, P. (2005), Hopping Robot for Planetary Exploration, *Proceedings of 8th International Symposium on Artifical Intelligence, Robotics and Automation in Space*, iSAIRAS'05, 5-8 September 2005, Munich, Germany.

Dupuis, E., Montminy, S., Farhad, M., & Champliaud, H. (2006), Mechanical design of a hopper robot for planetary exploration, *Proceedings of 9th ESA Workshop on Advanced Space Technologies for Robotics and Automation*, ASTRA 2006, 28 – 30 November 2006, ESTEC, Netherlands.

Dubowsky, S., Iagnemma, K., Boston, P. (2004), Microbots for Large-Scale Planetary Surface and Subsurface Exploration, Meeting presentation, NASA Institute for Advanced Concepts, March 2004.

Dubowsky, S., Iagnemma, K., Liberatore, S., Lambeth, D.M., Plante, J.S., Boston, P.J. (2005), A Concept Mission: Microbots for Large-Scale Planetary Surface and Subsurface Exploration, CP746, *Space Technology and Applications International Forum*, STAIF 2005.

Gibbesch, A., Krenn, R., Rebele, B., Schäfer, B., & Scharringhausen, M., Verification and Validation of the Exomars rover and Single Wheel Mobility MBS Simulations, Multibody dynamics 2009, *ECCOMAS Thematic Conference*, Arczewski, K., Fra¸czek, J., Wojtyra, M. (eds.), Warsaw, Poland, 29 June–2 July 2009.

Görner, M. (2007), *"Mechatronic Concept of Crawler from DLR-Hand II Technology"*, Diploma Thesis, Institute of Robotics and Mechatronics, German Aerospace Center.

Hirschmüller, H. (2008), Stereo Processing by Semi-Global Matching and Mutual Information, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 30(2), pp. 328-341.

Kermurjian, A. (1990), From the Moon Rover to the Mars Rover, Reprint: *The Planetary Report*, Issue: July/August 1990.

Klinker, S. (2005), Lunar Exploration using Small Satellite with Micro Rover Technology, Proceedings of *10th International Space University Annual International Symposium*, November 2005, Strasbourg, France.

Klinker, S., Lee C.G.-Y., Wagner C., Hlawatsch W., Schreyer A.-M., & Röser H.-P. (2007), Destination Moon and beyond for the Micro rover Nanokhod, *Proceedings of DGLR International Symposium "To Moon and beyond"*, 14 – 16 March 2007, Bremen.

Kubota, T., Kunii, Y., Kuroda, Y., & Working Group (2005), Japanese lunar robotics exploration by co-operation with lander and rover, *Journal of Earth System Science* 114, No. 6, 777-785, December 2005.

Lindemann, A., & Voorhees, J. (2005), Mars Exploration Rover mobility assembly design, test and performance, *Proceedings of 2005 International Conference on Systems, Man, and Cybernetics,* October 10-12, 2005, Hawaii.

Lindemann, R.A., Bickler, D.B., Harrington, B.D., Ortiz, G.M., & Voorhees, C.J. (2006), Mars Exploration Rover Mobility Development: Mechanical Mobility Hardware Design, Development, and Testing, *IEEE Robotics & Automation Magazine*, 19-26 June 2006.

Morring, F. (2007), Enabling Exploration: Rovers, reactors and "green" rockets all on, NASA's future-technology list, *Aviation Week and Space Technology*, June 4, 2007.

NASA/JPL/Caltech website (2005), Flight Projects-Mars Science Laboratory website, Robotics Section.

Roe, M. (2008), ExoMars Rover Vehicle, *European Industry Day*, 29 May 2008, Turin.

Spenneberg, D. & Kirchner, F. (2002), SCORPION: A Biomimetic Walking Robot.

Steinmetz, B.M., Arbter, K., Brunner, B., Landzettel, K. (2001), Autonomous Vision-Based Navigation of the Nanokhod Rover, 6th International Symposium on Artificial Intelligence and Robotics & Automation in Space, i-SAIRAS 2001, Canadian Space Agency, St-Hubert, Quebec, Canada, June 18-22, 2001.

Wakabayashi, S., Sato H., & Matsumoto, K. (2006), Design and Mobility Evaluation of a Crawler-type Lunar Vehicle, *Earth and Space 2006: Engineering, Constructions, and Operations in Challenging Environments*, 2006.

Wilcox, B.H., Litwin, T., Biesiadecki, J., Matthews, J., Heverly, M., Morrison, J., Townsend, J., Ahmad, N., Sirota, A., & Cooper, B. (2007), ATHLETE: A Cargo Handling and Manipulation Robot for the Moon, *Journal of Field Robotics,* 24(5), pp. 421-434.

# Part III

# 11

# Multi-Platform Atmospheric Sounding Testbed (MAST)

Meemong Lee, Richard Weidner and Kevin Bowman
*Jet Propulsion Laboratory, California Institute of Technology*
*USA*

## 1. Introduction

NASA's Earth atmospheric science missions study the physical properties of Earth's atmosphere (such as pressure, temperature, wind, humidity, aerosols, and trace gases) by employing a wide range of atmospheric sounding systems. An emerging new paradigm in Earth science missions addresses the interplay between observing systems and Earth system models where observations are assimilated to validate the models and simulated experiments are performed to optimize future observations. The new paradigm provides a bridge between scientists and engineers allowing them to collaboratively explore many questions such as

- What needs to be measured?
- When and where?
- How often and how long?
- How accurately & how precisely?

Multi-platform Atmospheric Sounding Testbed (MAST) is an end-to-end simulation environment at Jet Propulsion Laboratory that enables the new paradigm. The MAST allows scientists to formulate the exploration space parametrically, simulate observing systems realistically, and validate science impacts quantitatively. The parametric formulation of the exploration space addresses interactive design of sampling strategies and measurement quality ranges that can be represented as a sensor-web operation composed of multiple instruments and multiple platforms. The realistic observing system simulation addresses integration of numerical models of sensor-specific measurement physics, sampled atmospheric phenomena, and measurement noise. The quantitative validation of science impact addresses retrieval sensitivity as a function of measurement precision and accuracy and assimilation sensitivity as a function of sampling duration, frequency, and measurement type.

The MAST team is composed of the investigators of three independent research tasks, Observing System Simulation Experiment (OSSE) (Lee et al, 2007), and GEOS-Chem-Adjoint (GCA) (Sandu et al., 2003, Henze et al.,2007)], and Sensor-web Operations Explorer (SOX) (Lee et al., 2008). The OSSE team researches forward modeling and inverse methods for atmospheric sounding system design and validation. The GCA team researches 4D-variational assimilation methods and full-chemistry sensitivity analysis methods for the Goddard Earth Observing System-Chemistry GEOS-Chem) community. Finally the SOX team researches advanced information system technologies for an integrated air-quality

campaign that involves multiple sensors and multiple platforms. The capabilities developed in the three research tasks are seamlessly integrated to achieve the end-to-end simulation environment by developing three collaborative systems. As shown in Figure 1, scientists explore measurement scenarios primarily interacting with the SOX framework. The SOX system performs the requested exploration by interfacing with the OSSE system and the GSA system for observing system simulation and science impact validation. The heterogeneous computational environments required by the three systems (including operating systems, programming languages, processor types, and file systems) are integrated employing a two-layer network topology (Figure 2) and client-server protocols.

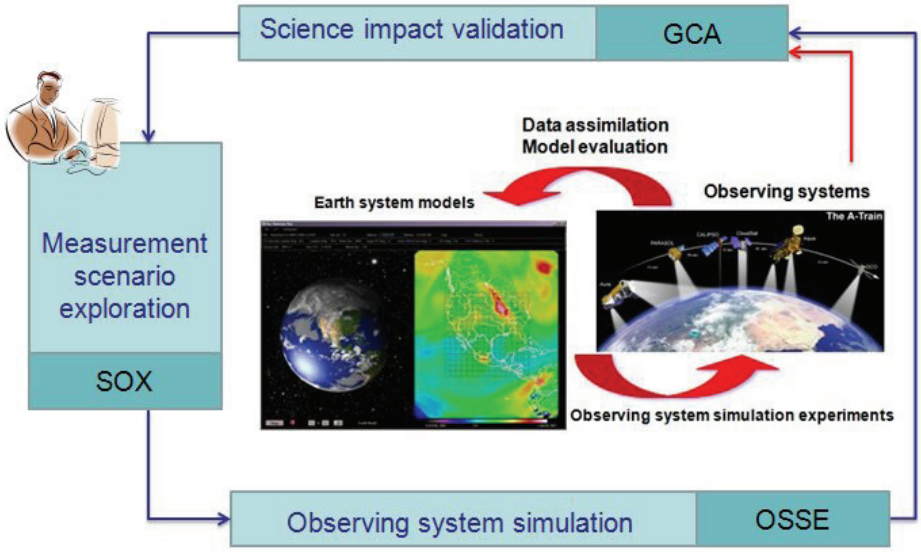



Fig. 1. New Paradigm for Earth Atmospheric Science Mission Concept Design

## 2. OSSE System

The OSSE System simulates an atmospheric sounder by integrating three types of models, a forward model for the physical properties of the atmosphere, an instrument model for measurement quality, and an inverse method for retrieving the atmospheric state variables. In the OSSE system, a radiative transfer function (Clough et al., 2005) is used as a forward model to simulate atmospheric chemistry missions, which generally exploit the spectral range between unltraviolet (UV) and mid infrared (IR). The forward radiative transfer function computes ideal signal radiance as an input to the instrument system. The instrument model distorts the ideal signal into a noisy measurement data by introducing spatial, spectral, and radiometric distortion properties of the instrument system (Figure 3). The distortion properties of an instrument system are addressed by three generic components, an imager, a spectrometer, and a radiometer. The inverse method retrieves an unknown atmospheric state variable (e.g., vertical profile of Ozone) from the noisy measurement data. The inverse method estimates the state variable employing the Tangent-Linear of the forward model function (i.e., analytic Jacobian) and a priori knowledge.

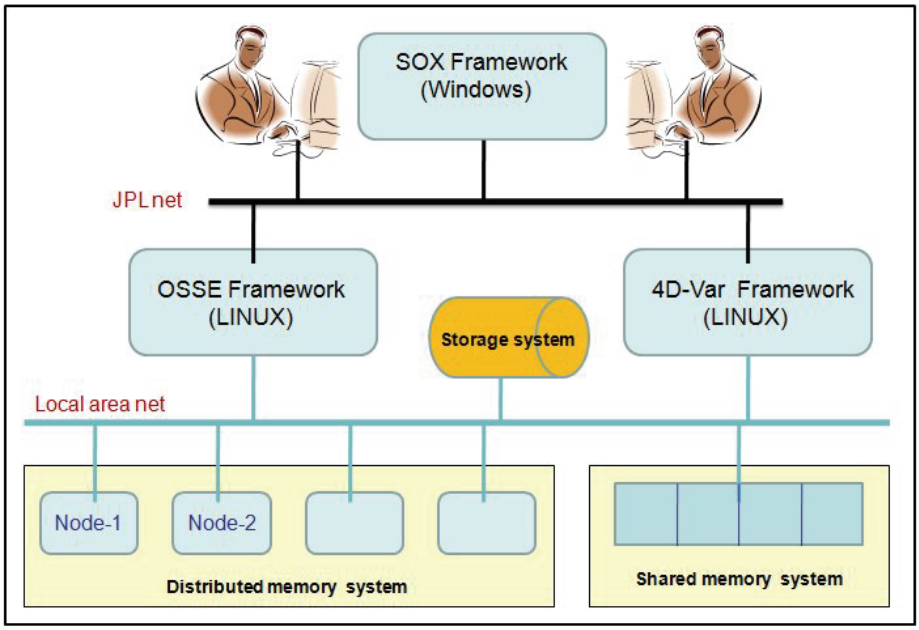



Fig. 2. The MAST System Architecture

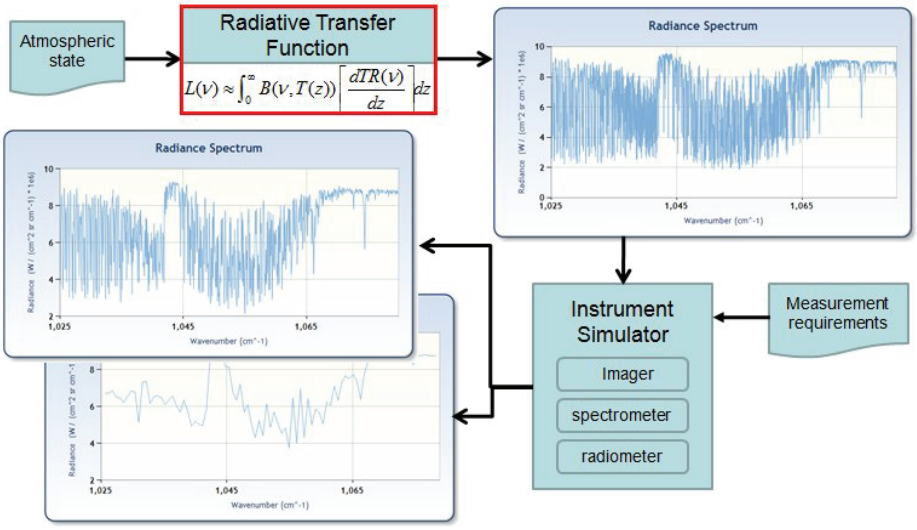



Fig. 3. Ideal Signal Radiance and Measured Radiances

The OSSE process starts with a reference atmospheric state and a measurement requirement as shown in Figure 4. The atmospheric state describes vertical profiles of temperature, humidity, pressure, and trace gas density; and the measurement requirement specifies the range of the instrument's response properties to be explored. The atmospheric state may

also include surface emission, surface reflectance and scattering, and cloud contribution when the spectral range includes the visible wavelength range. The forward radiative transfer function transforms the reference atmospheric state into a signal radiance spectrum. The instrument model simulates a set of noisy measurements where each measurement represents the output of a specific instrument response property. The inverse method estimates the atmospheric state variables from the noisy measurements and analyzes the retrieval error sensitivity with respect to instrument response properties.

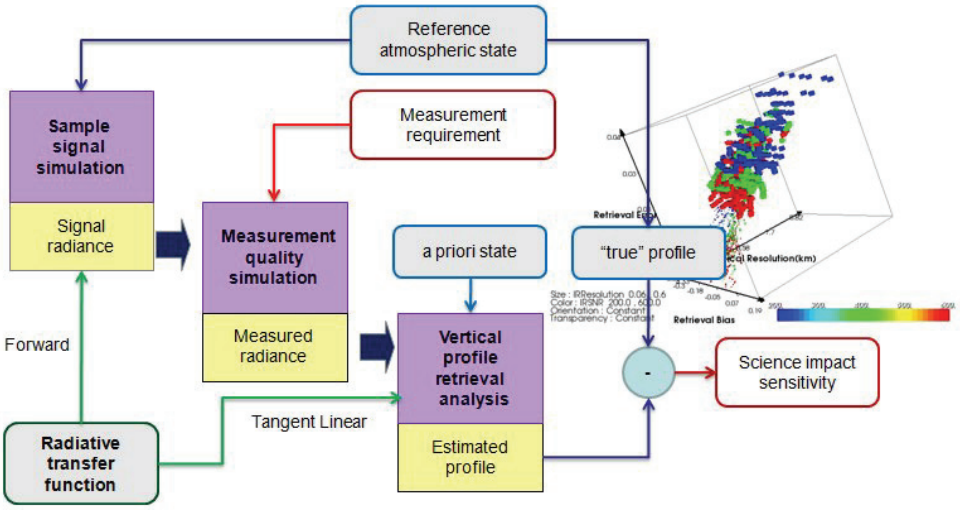



Fig. 4. Observing System Simulation Experiment Framework

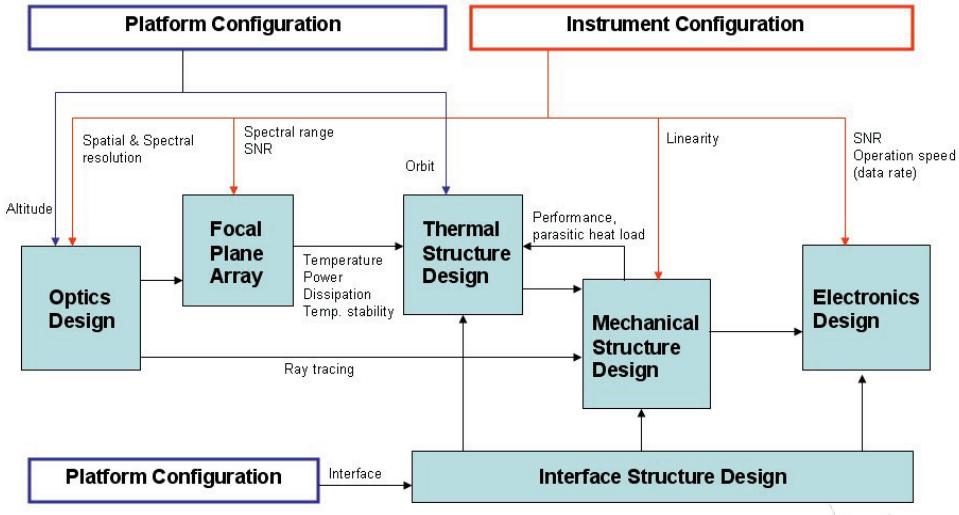



Fig. 5. Relationship between Measurement Requirements and Instrument Subsystems

The retrieval error sensitivity (Rogers, 2000) with respect to the measurement requirements provides a quantitative evaluation of science impact during the conception phase of an instrument enabling scientists to answer "what to measure", "how precisely, and how accurately". In order to establish a bridge between science and engineering, the measurement requirements is formulated as a trade space whose coordinate axes include science impacts and instrument design drivers. Figure 5 illustrates the rationale behind the instrument response properties that are represented in the measurement requirements. The platform altitude and orbit drive optics design and thermal design. The instrument signal-to-noise ratio (SNR) and spectral range requirements drive the focal-array design. The spectral linearity requirement drives mechanical design, which indirectly drives thermal design. The operation speed and date rate drives electronics design.

"Exploration" of an instrument concept refers to population of the trade space defined above where the science objective coordinate is represented with the retrieval error and retrieval error bias for a specific atmospheric component, and the instrument response property coordinate is represented with spatial and spectral resolution, SNR, spectral range, linearity, and operation speed. The evaluation of science impact, taking sampling strategies into consideration, requires a comprehensive coverage of the atmospheric states and global data assimilation.

## 3. GCA system

The GEOS-Chem-Adjoint (GCA) research team develops a standardized adjoint of the GEOS-Chem for global data assimilation and a full chemistry sensitivity analysis (Sandu et al., 2007). GEOS-Chem is a global 3D model of atmospheric composition driven by assimilated meteorological observations from the Goddard Earth Observing System (GEOS) of the NASA Global Modeling and Assimilation Office. The global data assimilation optimizes the combination of three sources of information: an a priori state, a forward model of physical and chemical processes, and observations of some state variables. The observations in this case refer to the retrieved vertical profile of the trace gas components, generally known as level-2 mission data products.

Adjoint models are powerful tools widely used in meteorology and oceanography for applications such as data assimilation, model tuning, sensitivity analysis, and the determination of singular vectors. The **GCA** System (Figure 6) provides adjoint models for chemistry, advection, convection, and deposition/emission. The adjoint model computes the gradient of a cost function with respect to control variables. Generation of adjoint code may be seen as the special case of differentiation of algorithms in reverse mode, where the dependent function is a scalar. Developing a complete adjoint of global atmospheric models involves rigorous work of constructing and testing adjoints of each of the complex science processes individually, and integrating those adjoints into a consistent adjoint model.

The mathematical formulation for calculating gradients of a model output using the adjoint method can be derived from the equations governing the forward model analytically or discretely. The GCA system employs the discrete adjoints. This provides exact resulting gradients of the numerical cost function for easier validation and allows application of automatic differentiation tools (e.g., Tangent-linear Adjoint Compiler Model (TAMC)). The Kinetic Pre Processor (KPP) module provides very effective sparse matrix computational kernels, which lead to high computational efficiency (Sandu et al., 2005).

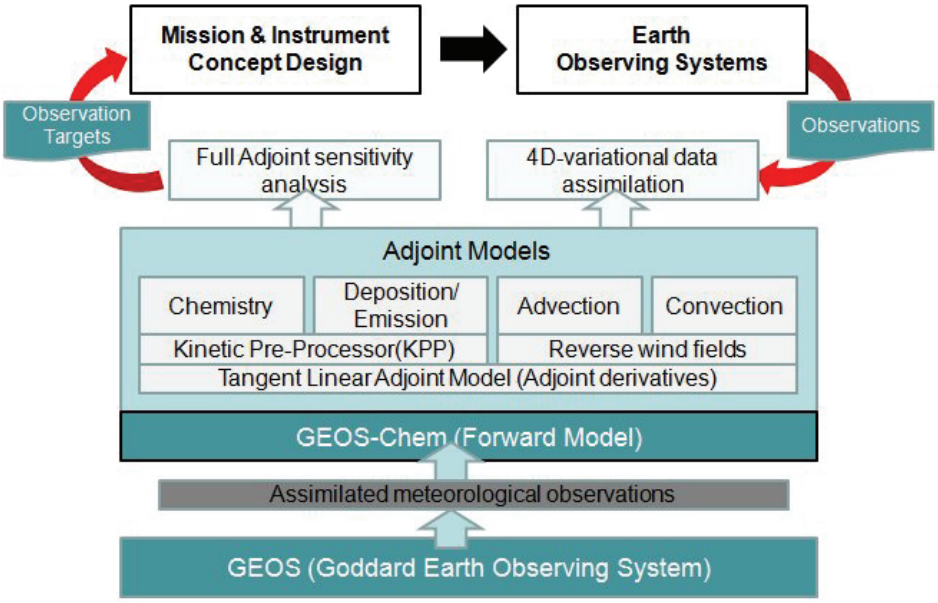



Fig. 6. GEOS-Chem-Adjoint System

The adjoint sensitivity analysis approach is receptor-oriented, and it traces backward in time for the cause of a perturbation in an output variable contrast to the forward sensitivity analysis, which propagates the initial perturbation forward in time. The sensitivity mode allows collaborative observation planning between air-borne and space-borne missions as well as targeted observation planning. Figure 7 illustrates two types of sensitivity analysis results. The top three charts show the sensitivity of an ozone profile over the **Tropospheric Emission Spectrometer** (TES) (Kulawaik et al., 2006) orbit track with respect to surrounding chemical content such as NOx, CO, and Ox. The bottom three charts show the sensitivity of ozone density in New York with respect to past presence of NOx, 7 days earlier, 0.5 day earlier, and the same day. The exploration of the sampling strategies and measurement quality ranges requires a system that integrates the OSSE and GCA capabilities.

## 4. SOX system

The SOX system enables measurement scenario exploration in collaboration with the OSSE system and the GCA system by developing complex phenomena visualization, community model integration, process streamlining, collaborative design environments, experiment database management, and computational load balancing. For exploring the measurement scenarios on multiple platforms, a set of design tools has been implemented to interactively manipulate and visualize the sampling strategies on geostationary Earth orbit (GEO), low-Earth-orbit (LEO), and air-borne platforms. The measurement scenarios are translated into a list of sample locations and times of observation, and into a list of instrument performance specifications. Each sample list represents a sampling strategy on a specific platform, while each instrument list represents the performance ranges of a single type of instrument. The

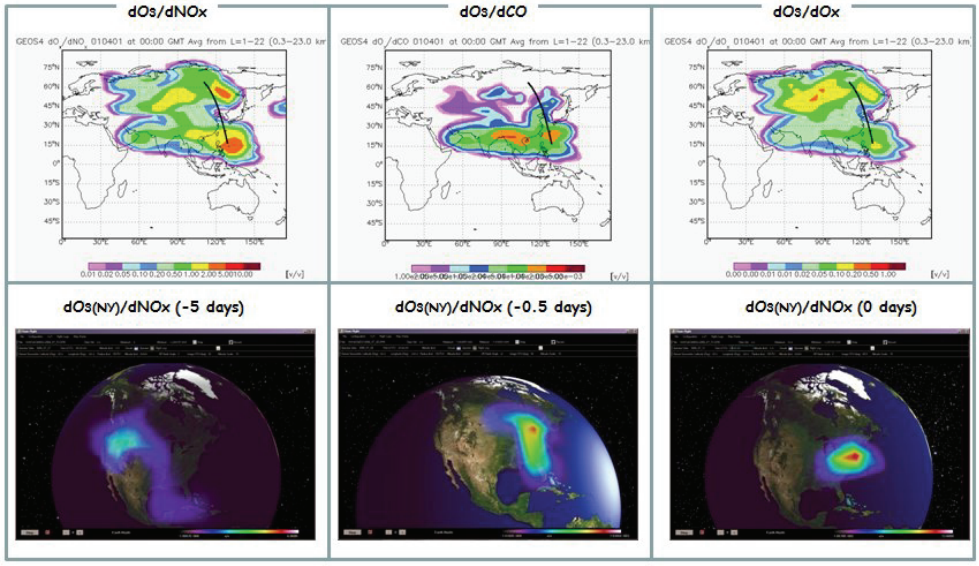



Fig. 7. Sensitivity Mode of GEOS-Chem Adjoint

sample lists can be combined to represent a sampling strategy on multiple platforms, multiple sampling strategies on a single platform, or multiple sampling strategies on multiple platforms. Similarly, the instrument lists can be combined to represent multiple sensors of varying performance ranges.

The SOX system performs the exploration by organizing the end-to-end process into four stand-alone stages, sampling strategy, instrument property, retrieval analysis, and data assimilation as shown in Figure 8. At each stage, the resulting datasets are stored in the respective database along with the request information. A sample list is submitted as an input to the sampling strategy exploration, while an instrument list is submitted as an input to the measurement quality exploration along with a signal-radiance database previously generated. The SOX on-line service facilitates the collaborative design environment among the distributed teams at JPL.

In order to manage the on-line requests, balance computational loads, and integrate community models, the SOX system has implemented four types of services, a request handler, a load distributor, an application dispatcher, and a status monitor (Figure 9). The request handler checks for the new on-line requests in the SOX database and alerts the load distributer with a list of items to be explored. The load distributer divides the list into multiple sub-lists, assigns one sub-list per available processor, and starts the application dispatcher. For example, a sample list of 1024 samples will be divided among 8 processors by composing eight sample lists with 128 samples, one per processor. Within each processor, the application dispatcher performs the exploration task activating an application-specific interface module for input dataset preparation, command script composition, and output dataset conversion. The status monitor checks the run status of the application on each processor and updates SOX database with the completion status so that it can be accessed via the on-line service. The SOX system utilizes MySQL relational database manager for distributed database service and Mono for execution of the .NET

framework on Linux platforms. A set of standard services has been also utilized including international information service (IIS), .NET framework, Active server page (ASP), Asynchronous Java script, and active XML (AJAX).

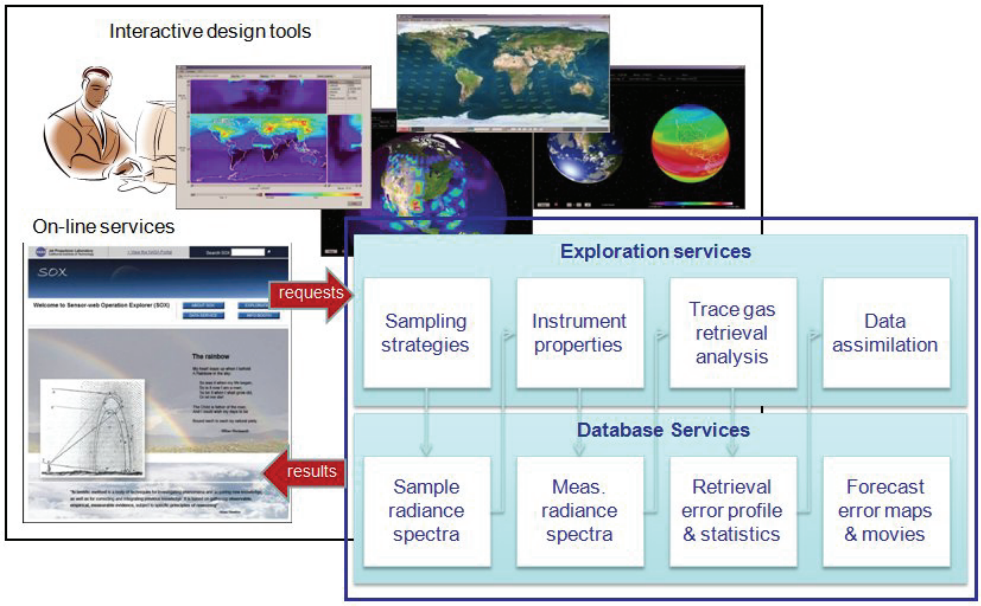



Fig. 8. Design Tools and On-line Services of SOX Framework

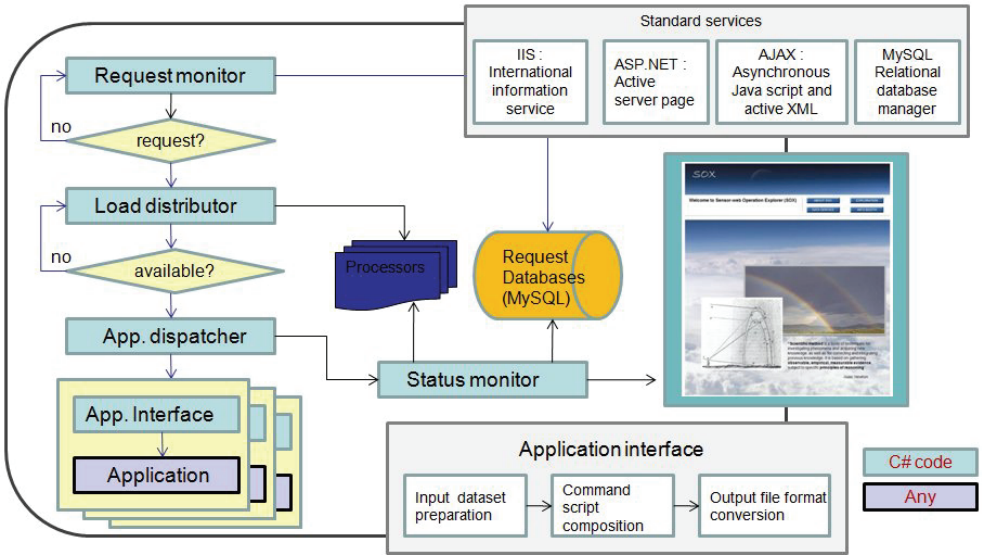



Fig. 9. On-line Service Process Flow

## 5. Conclusion

The MAST has successfully demonstrated an end-to-end simulation process that can quantitatively evaluate the science impacts of instrument concepts and sampling strategies by integrating the OSSE system, the GCA system, and the SOX system (see Figure 10). The end-to-end process is organized with four stages of (1)observation scenario exploration, (2)measurement quality exploration, (3)measurement quality evaluation, and (4)observation scenario evaluation employing forward models, system performance simulations, and inverse methods. The first two stages enable scientists to formulate and explore measurement requirements while the last two stages enable scientists to validate and optimize them against science objectives and design processes. The four stages are streamlined by infusing a wide range of information technologies including on-line services, distributed database management, and parallel computing.

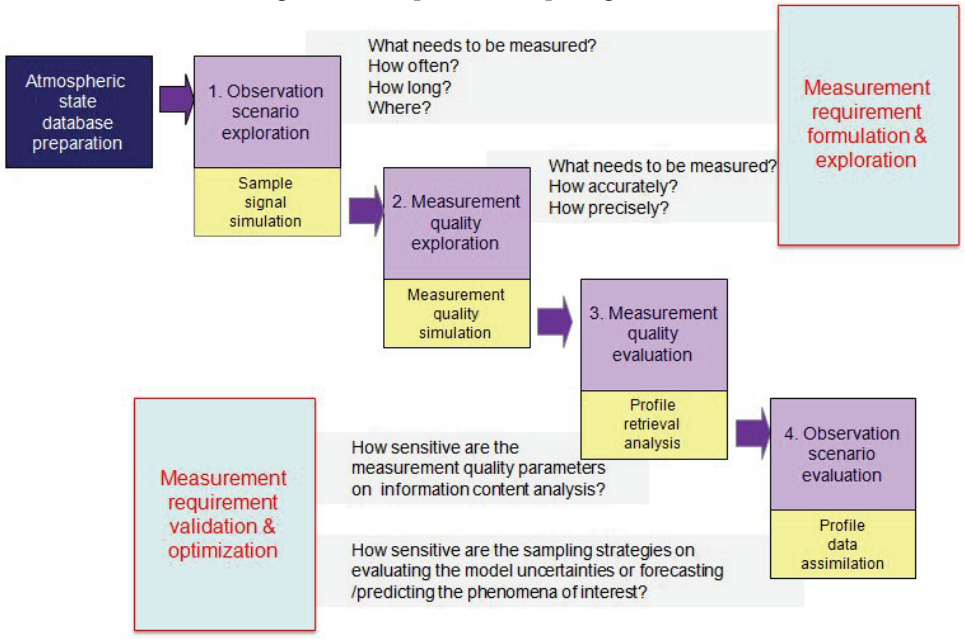



Fig. 10. End-to-end Process of the MAST and Measurement Requirements

The MAST currently supports GEOCAPE (Geostationary Coastal and Air Pollution Events) concept study (lead: Dr. Annmarie Eldering/JPL), part of Tier-2 missions recommended by the NRC decadal survey (http://nasascience.nasa.gov/earth-science/decadal-surveys). The MAST is being utilized to evaluate the advantage of geostationary orbit over low-Earth orbit and to explore the detailed science return from improved measurement capabilities including spectral coverage (IR, UV, IR+UV), spectral resolution, and signal-to-noise ratio. The science impact evaluation is with respect to chemical data assimilation for improved air quality forecasts, pollutant emission monitoring, and regional-scale to intercontinental-scale pollution transport.

The MAST capabilities will be extended to support the CLARREO (Climate Absolute Radiance and Refractivity Observatory) concept study part of Tier-1 missions recommended

by the NRC decadal survey, for mission design and virtual observation for climate model uncertainty evaluation. The largest source of uncertainty for climate prediction is climate feedbacks that are coupled radiative response of the hydrological cycle to anthropogenic forcing. The MAST will be employed to evaluate the sensitivity of the climate feedbacks which are manifested at unresolved scales for contemporary climate models and the proposed CLARREO footprint.

The future research areas of interest include a web-based model integration infrastructure that provides a dynamic coupling of global and regional phenomena models, a model-based system engineering process that comprehensively validates and verifies instrument design and mission planning, and a heterogeneous data assimilation method that can rapidly assimilate observations from multiple sensors on multiple platforms. The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the Advanced Information System Technology (AIST) program under the Earth Science Technology Office (ESTO) of the National Aeronautics and Space Administration (NASA).

## 6. References

Clough, S; Shephard, M; Mlawer, E; .Delamere, J; Iacono, M; Cady-Pereira, K; Boukabara, S; and Brown, P; "Atmospheric Radiative Transfer Modeling: a Summary of the AER Codes," *JQSRT* 91, 233–244, 2005.

Henze, D; Hakami, A; Seinfeld, J; "Development of the Adjoint of GEOS-Chem," *Atmos Chem. Phys*., 7:2413–2433, 2007.

Kulawaik, S; Worden, H; Osterman, G; Luo, M; Beer, R; Kinnison, D; "TES atmospheric profile retrieval characterization: an orbit of simulated observations," *IEEE Transactions on Geoscience and Remote Sensing*, 44:5, 2006.

Lee, M; Miller, C; Qu, Z; "Earth Science Mission Concept Design System, presented at *2007 IEEE Aerospace Conference*, Big Sky, MT, 2007

Lee, M; Weidner, R, Bowman, K; Sandu, A; Singh, K; "Sensor-web Operations Explorer (SOX) for Integrated Air Quality Campaign," presented at *2008 IEEE Aerospace Conference*, Big Sky, MT, 2008.

Rodgers, C; *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Publishing, Series on Atmospheric, Oceanic and Planetary Physics, Vol. 2, Singapore, 2000.

Sandu, A; Daescu, D; Carmichael, G; "Direct and Adjoint Sensitivity Analysis of Chemical Kinetic Systems with KPP: I - Theory and Software Tools," *Atmospheric Environment*, vol. 37 (36), pp. 5083–596, 2003.

Sandu, A; Daescu, D; Carmichael, G; Chai, T; "Adjoint Sensitivity Analysis of Regional Air Quality Models," *Journal of Computational Physics*, vol. 204, pp. 222–252, 2005.

# 12

# Low-thrust Propulsion Technologies, Mission Design, and Application

John W. Dankanich
*Gray Research Inc.*
*U.S.A.*

## 1. Introduction

Electric propulsion has been widely accepted for station-keeping and final orbit insertion of commercial satellites. NASA, JAXA, and ESA have all used primary electric propulsion systems for science missions. Electric propulsion systems have been recently developed with a significant increase in performance and ability to process large amounts of onboard solar power. While the use of electric propulsion offers significant performance gains, it is not appropriate for all missions, has limitations, and the trajectories have characteristics that may be counterintuitive to those unfamiliar with low-thrust trajectory design. This chapter describes recent U.S. technology investments in electric propulsion thrusters with emphasis on mission application and low-thrust mission design for interplanetary trajectories and geosynchronous transfer using primary electric propulsion.

## 2. Overview of electric propulsion technologies

Unlike chemical propulsion, which is limited to the energy available through the decomposition or combustion of molecular compounds, electric propulsion makes use of energy from an external source, typically solar power, to electrically accelerate the propellant to higher energies. The efficiency of momentum transfer is often described in terms of specific impulse which is proportional to the average exhaust velocity in the thrust direction.

$$I_{sp} = \frac{v_{Exhaust}}{g} \tag{1}$$

The three basic types of electric propulsion systems are electrothermal, electrostatic, and electromagnetic. The types are categorized by the method of accelerating the propellant. Resistojets, arcjets, pulsed plasma, gridded-ion and Hall thrusters have significant flight experience. Electrothermal thrusters are the most widely used electric propulsion systems to date, but electrostatic systems are the industry's state-of-the-art (SOA) with higher specific impulses. The electrostatic thruster successes are made possible through technology advancements for increased power processing capability and increased thruster life driven by an increase in spacecraft available power.

### 2.1 Electrothermal propulsion

In electrothermal propulsion thrusters, electrical energy is applied to heat a working fluid to increase the exhaust velocity. Resistojets are a form of electrothermal propulsion that operate by passing a gaseous propellant though an electric heater and then expanding it through a conventional converging diverging nozzle to create thrust. The typical flight operation is superheating catalytically decomposed hydrazine to leverage the propellant commonality of standard monopropellant chemical propulsion systems. The specific impulse of resistojets is limited by the high molecular mass of hydrazine and the maximum sustainable temperature. Specific impulse values near 350s is achievable; 40 percent higher than the conventional chemical equivalent. Arcjets are another form of electrothermal propulsion that passes propellants through an electric arc that heats the gas before it expands through a nozzle. Arcjet specific impulses are typically in the 500–600s range. Higher specific impulses are achieved because the maximum temperatures are not in contact with engine component walls, though efficiencies are less than that of resistojets.

### 2.2 Electromagnetic propulsion

Electromagnetic propulsion devices leverage magnetic fields, self-field or applied, to accelerate plasma, typically with a Lorentz force ($\mathbf{J} \times \mathbf{B}$) where the accelerating force is proportional to the cross product of the electric current density and the magnetic field.

The pulsed plasma thruster (PPT) is a form of electromagnetic propulsion that uses a capacitor to store electrical energy and when triggered creates a pulsed arc discharge across the face of a block of propellant, typically polytetrafluoroethylene (e.g. Teflon). This arc ablates and ionizes a small amount of propellant and the self-induced magnetic field acts on the ions to create a Lorentz force accelerating the plasma. The use of PPTs for East-West station keeping began in 1968 on the Lincoln Laboratory LES-6 satellite.

Pulsed inductive thrusters (PIT) and magnetoplasmadynamic (MPD) thrusters are additional forms of electromagnetic propulsion. The majority of these concepts are proposed for high power levels, >100 kWe, and have not gained any flight experience.

### 2.3 Electrostatic propulsion

While field emission electric propulsion (FEEP) and colloid thrusters fall in the electrostatic category, they are typically perceived as very low thrust devices lending themselves for disturbance force cancellation or precision control. For this reason, they are not candidates for primary propulsion. Large arrays are under investigation for primary propulsion.

Gridded-ion and Hall thrusters are the leading concepts for primary electric propulsion. Ion thrusters can achieve very high exit velocities, and have typical specific impulses in the 3000–4000s range. The first U.S. flight test with an ion engine was in 1964 with the Space Electric Rocket Test (SERT) I. Gridded-ion engines are in routine operation since the 1990s for geostationary north-south station keeping. In 1998, Deep Space 1 (DS1) was the first U.S. demonstration of primary electric propulsion. In September of 2007, NASA launched its first science mission, Dawn, using an ion engine for primary electric propulsion.

Gridded-ion thrusters operate by injecting a neutral gas in a thrust chamber. The gas is then ionized and magnetically contained within the chamber. The positively-charged ions migrate between a set of grids where the ions experience a large voltage potential. The ions are accelerated by a Coulomb force to high exhaust velocity, typically 30,000–40,000 m/s. The electrons inside the thruster chamber are then pumped by the system's power

processing unit to a neutralizing cathode to maintain a zero net charge in the plume. An operational schematic of a gridded-ion engine is shown in figure 1.

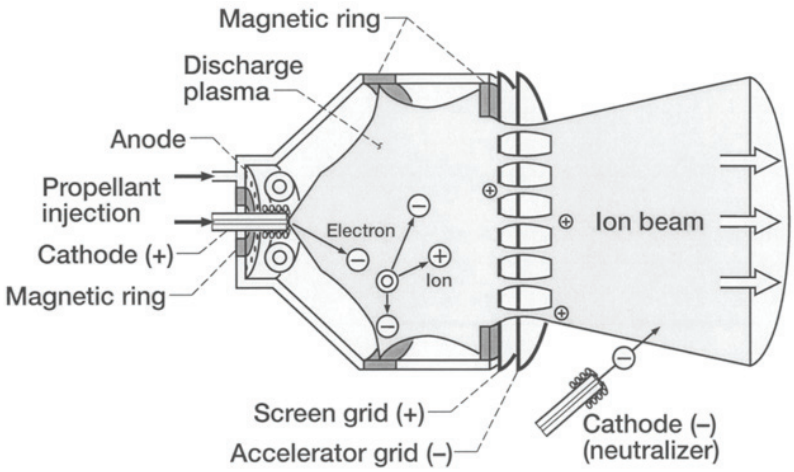



Fig. 1. Operational schematic of a gridded-ion engine.

A Hall thruster is essentially a grid-less ion engine. The thruster operates by employing magnetic fields to deflect low-mass electrons so that they are trapped under the influence of an **E** × **B** azimuthal field. The electrons are forced into an orbiting motion by the Hall effect near the exit plane of the thruster. A propellant is injected through the anode where the trapped electrons will collide and ionize the propellant. The ionized propellant will see the potential of the electron plasma and accelerate towards the thruster exit. Hall thruster exhaust velocities are typically 15,000–25,000 m/s. An operational schematic of a Hall thruster is shown in figure 2.

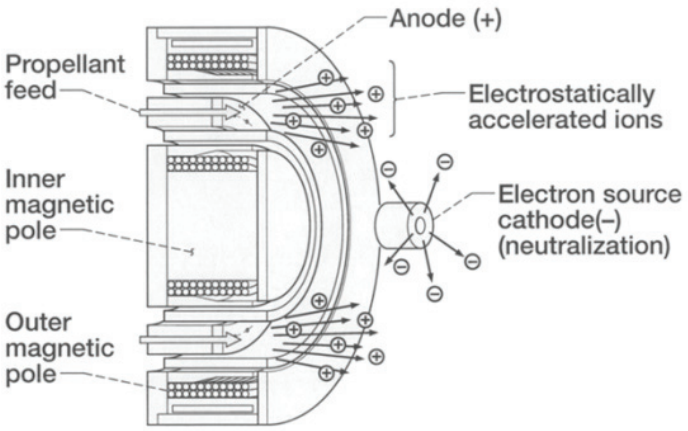



Fig. 2. Operational schematic of a Hall thruster.

## 3. U.S. advancement in electrostatic propulsion technologies

The state-of-the-art in electric propulsion thrusters suitable for primary electric propulsion is the L-3 25-cm Xenon Ion Propulsion System (XIPS), NASA's Solar Electric Propulsion Technology Application Readiness (NSTAR) thruster, and the Aerojet BPT-4000 Hall thruster. These three systems are all fully qualified with flight experience on the L-3 XIPS and NSTAR thrusters, and Aerojet's BPT-4000 is scheduled to launch in 2010. NASA is also completing the prototype development of NASA's Evolutionary Xenon Thruster (NEXT), and is completing the engineering model development of the High Voltage Hall Accelerator (HiVHAC) with Aerojet. While there are other great advancements in U.S. electrostatic thrusters, such as the BHT-200, it is not considered a candidate for primary electric propulsion and is not discussed. Representative thrust and specific impulse throttle tables of the L-3 25-cm XIPS, Aerojet BPT-4000, NSTAR and NEXT are shown in figure 3. Greater detailed performance data is available in open literature. The thrusters can operate well outside the throttle tables shown with varying thruster efficiencies.

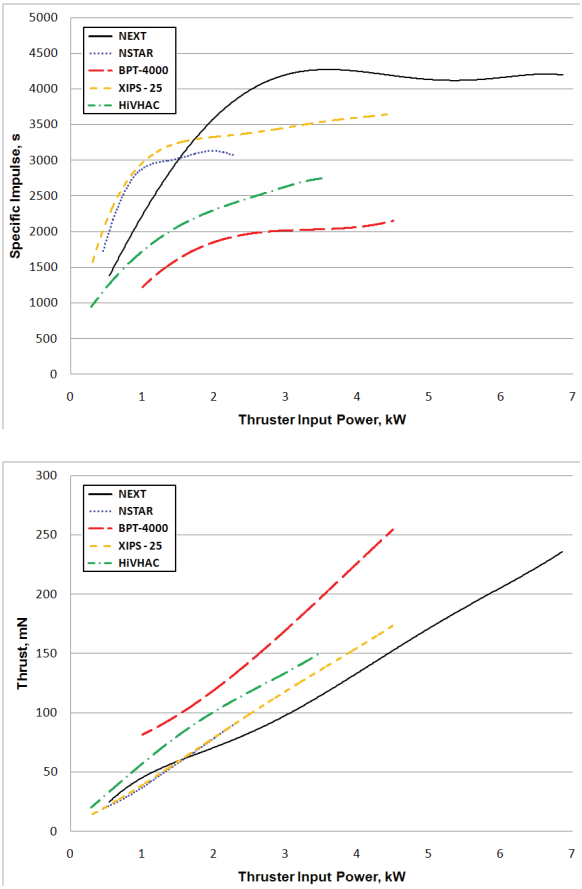



Fig. 3. Representative throttle tables for multiple commercial and NASA EP thrusters.

### 3.1 Commercial thrusters

Two commercially available thrusters are of high interest for primary electric propulsion; the L-3 XIPS and the Aerojet BPT-4000. Both commercial thrusters have a maximum operating input power near 4.5kW and primarily designed to operate over a few operating points, but have demonstrated large throttleability. The XIPS 25-cm and BPT-4000 thrusters are shown operating in figure 4.

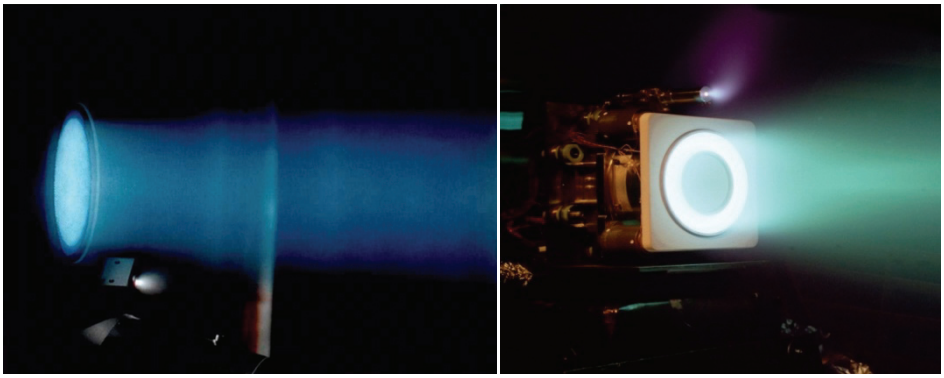


Fig. 4. The XIPS-25 firing (left) and the BPT-4000 thruster firing(right).

The L-3 25-cm XIPS thruster is used on the Boeing 702 communcation satellite for attitude control, north-south and east-west station keeping, momentum dumping, de-orbit, and augmenting orbit transfer (Tighe et al., 2006). The flight system operates in modes, 2.2 kW for attidue control and 4.4 kW for orbit transfer.

The 25-cm XIPS qualification testing demonstrated over 2,500 hours of operation at full power and over 13,000 hours at the lower power setting. While erosion and throughput capability is a function of the operating conditions of the engine, the 25-cm XIPS was projected to have similar life capability as the NSTAR thruster. Additional testing over a large throttle range was demonstrated (Goebel et al., 2006).

Aerojet completed qualification of the BPT-4000 Hall thruster in 2006 (Wilson & Smith, 2006). The BPT-4000 is a 4.5 kW multi-mode Hall thruster developed by Aerojet and Lockheed Martin Space Systems Company as part of a Hall Ion Propulsion System (IPS) for use on geosynchronous satellites. The thruster is designed to operate between 3 kW and 4.5 kW at discharge voltages between 300 volts and 400 volts. The thruster operates at lower voltage for orbit raising maneuvers and higher voltage to provide a higher specific impulse during station keeping.

Qualification life testing processed approximately 272kg for a flight operational throughput capability of 181kg. It is predicted that the thruster will have a mission throughput capability greater than 285kg of propellant that may provide science mission applicability. Multiple sources funded additional life testing of the BPT-4000 for additional erosion data, to demonstrate a larger throughput capability, and to assess performance at low powers (Welander et al., 2006).

### 3.2 NASA thrusters

NASA leads the U.S. development of primary electric propulsion thrusters. Three particular thrusters, NSTAR, NEXT, and HiVHAC were or are under development for interplanetary

science missions under NASA's science mission directorate (SMD), and all were led by the NASA Glenn Research Center (GRC).

NASA's Solar electric propulsion Technology Application Readiness thruster is the state-of-the-art electric propulsion engine for primary propulsion on NASA science missions. The thruster has a nominal full power operation of 2.3kW at 3,100 seconds of specific impulse and 94mN of thrust. The NSTAR engine flew on the (DS1) technology demonstration mission funded through the NASA New Millennium Program in 1998. The DS1 mission successfully demonstrated the capability of the ion propulsion system by processing 81 kg of xenon propellant. A DS1 flight spare thruster was part of an extended life test (ELT) that validated the thruster life up to 235kg of propellant or 157kg of operation life. Probabilistic failure analysis predicts only a one percent failure rate below 178kg (Brophy, 2007). Most recently, NASA launched the Dawn Discovery Class mission to Ceres and Vesta which makes use of multiple NSTAR thrusters to perform the first multi-rendezvous mission, significantly improving the science capability of a single spacecraft.

The NEXT project was competitively selected to develop a nominal 40-cm gridded ion electric propulsion system (Patterson & Benson, 2007) though NASA's In-Space Propulsion Technology (ISPT) project. The objectives of this development were to improve upon the experimental NSTAR system by achieving lower specific mass, higher specific impulse (4,050 seconds), greater propellant throughput (current estimates exceed 700kg of xenon) and increase the power handling capability (6.9kw), thrust (240mN), and throttle range (12:1) to enable use on Flagship Class missions. The ion propulsion system components being developed under the NEXT task include the ion thruster, the power processing unit (PPU), the xenon feed system, and a gimbal mechanism. NEXT multi-thruster testing is shown in figure 5.

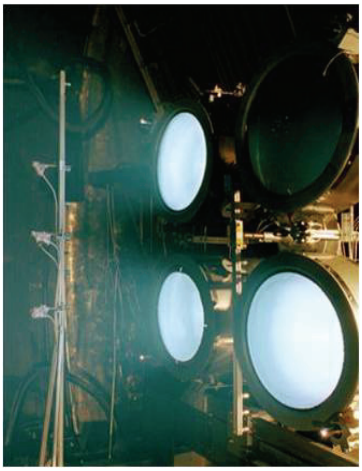


Fig. 5. Multi-thruster testing with NEXT.

The NEXT project is developing prototype-model fidelity thrusters with Aerojet. In addition to the technical goals, the project also has the goal of transitioning thruster manufacturing capability with predictable yields to an industrial source. Recent accomplishments include the production of a prototype-model NEXT thruster which has successfully completed qualification level environmental testing. As of July 1, 2009 a NEXT thruster has

demonstrated more than 425kg of xenon throughput and has exceeded 23,000 hours of operation at various throttle conditions. The NEXT wear test has demonstrated the largest total impulse ever achieved by a gridded ion thruster. Wear rates predict first failure beyond 700kg of throughout.

The ISPT project is also investing in the HiVHAC thruster (Manzella, 2007). HiVHAC, shown in figure 6, is the first NASA electric propulsion thruster specifically designed to provide a low-cost electric propulsion option optimized for Discovery and New Frontiers missions. The HiVHAC thruster does not provide as high a maximum specific impulse as NEXT, but a higher thrust-to-power ratio and lower power requirements are well suited for the anticipated demands of these smaller class science missions. The HiVHAC has a nominal full power operation of 3.6kW and a specific impulse of 2,800 seconds at 55% efficiency.


Fig. 6. HiVHAC laboratory model thruster.

Significant advancements in the HiVHAC thruster include a very large (12:1) throttle range, high discharge voltage for an increased specific impulse over conventional Hall thrusters, and a very long-life capability to allow for greater total impulse with fewer thrusters. The Hall thruster has fewer parts and low complexity with significant cost benefits over gridded-ion alternatives. A laboratory model HiVHAC thruster underwent wear testing and successfully completed its test demonstrating over 100kg of propellant throughput and over 4,750 hours of operation at a discharge voltage of 700V. The ISPT project procured the design and development of two engineering model HiVHAC thrusters with Aerojet. The engineering model thruster should be delivered and available to begin performance and life testing starting in the fall of 2009. The engineering model thrusters are projected to have thruster lifetime capability on the order of 30,000 hours of operation or having a total througput in excess of 600kg of propellant (Kamhawi et al., 2009).

## 4. Interplanetary application of electric propulsion

There is a misconception outside the low-thrust community that high specific impulse always translates directly to higher performance. This belief is founded from Tsiolkovsky's rocket equation.

$$\frac{m_{final}}{m_{initial}} = e^{\frac{-\Delta V}{g I_{sp}}} \tag{2}$$

However, its general application for high thrust, chemical propulsion, systems assumes that the mission ΔV remains relatively constant. If the ΔV remains constant, slight increases in specific impulse can have significant mass benefits to the mission. If thrust is decreased in exchange for higher specific impulse, the efficiency of the maneuver may decrease and the total ΔV requirement could rise, decreasing or negating any gain due to the increased exhaust velocity. One example is a launch vehicle whose specific impulse is increased, but its thrust-to-weight ratio is below one. The vehicle will consume all of its propellant without ever leaving the launch pad.

For electric propulsion thrusters, the thrust is inversely proportional to the specific impulse given a constant power.

$$P = \frac{g(FI_{sp})}{2\eta} \tag{3}$$

The most efficient propulsive maneuvers are impulses, or infinite thrust; though impossible to achieve. Chemical propulsion maneuvers are often treated as impulse maneuvers, but the low-thrust ΔV penalty of long finite burns can be quite severe. One example would be a simple plane change in an elliptical orbit. The ΔV of a plane change is a function of the spacecraft velocity.

$$\Delta V = 2V \sin\left(\frac{\emptyset}{2}\right) \tag{4}$$

The spacecraft velocity is slowest at apoapsis, and therefore, an impulsive maneuver at apoapsis will have a lower ΔV requirement than if the maneuver must be performed over a large arc.

The entire mission trajectory will have a decreased ΔV if the thrust arcs are smaller and centered on the most efficient locations. This will give a clear advantage to engines that can provide higher thrust. There is a trade between specific impulse and thrust. Figure 7 illustrates a Nereus sample return mission trajectory for the NSTAR thruster and the BPT-4000 (Hofer et al., 2006). The BPT-4000 operates at higher thrust, and therefore, has more efficient maneuvers to produce a lower total ΔV requirement for the mission.

Figure 7 illustrates that the higher thrust maneuvers are shorter, and the total ΔV savings is 1.3km/s. In this example mission, the NSTAR thruster requires approximately 190kg of propellant to deliver a final mass of 673kg while the BPT-4000 consumes 240kg of propellant and delivers 850kg back to Earth. It is worth noting that the higher thrust systems typically optimize to a lower launch energy; though the lower specific impulse BPT-4000 requires more propellant for a smaller ΔV, it delivers more final mass because the launch vehicle can deliver more start mass at the lower launch energy. Overall, the BPT-4000 can deliver more mass because of its higher thrust and ability to decrease the launch energy requirement of the launch vehicle. This is primarily due to the higher power processing capability of the thruster.

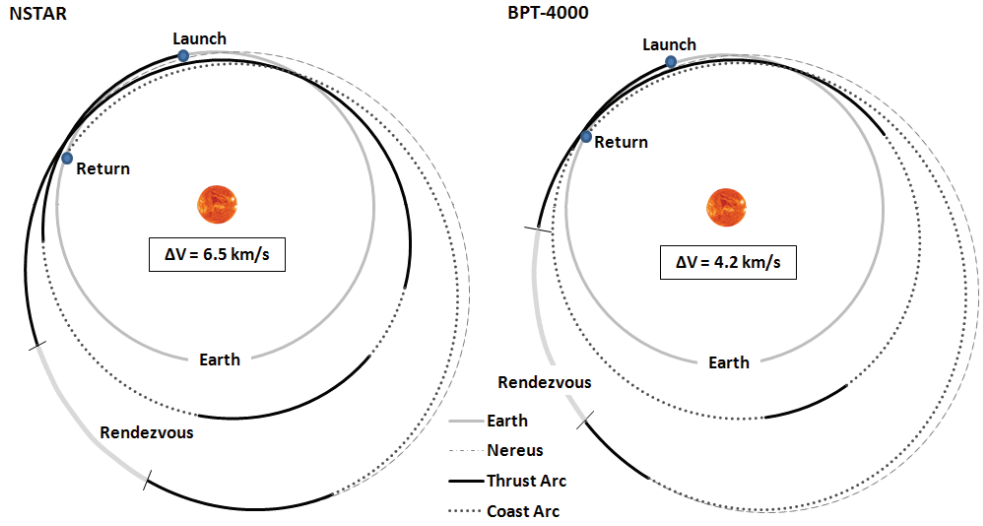



Fig. 7. NSTAR (left) and BPT-4000 (right) trajectories for a Nereus sample return mission.

The NEXT thruster performing the same mission, but de-rated to the maximum power level of the BPT-4000, can deliver 911kg consuming just 188kg of propellant. NEXT still requires a ΔV of 5.3km/s, greater than the Hall thruster, but the higher Isp results in a greater net delivered mass. It is not always obvious which thruster will have the highest performance. It is dependent on the trajectory profile, available power, mission duration, etc.

Another consideration of mission design is the ability to tolerate missed thrust periods. An advantage of higher thrust systems and the decreased thrust arcs is also the robust design of the trajectory. While the NEXT thruster delivers more mass than the BPT-4000, it is required to operate for 513 days of the 1,150 day mission. The BPT-4000 only operates for 256 days for the same mission duration. A missed thruster period, either for operations or an unplanned thruster outage, can have a negative impact on the mission. Higher thrust systems are typically more robust to missed thrust periods with their ability to makeup lost impulse in a short time period. Recalling equation 3, a higher power system can have both a higher thrust and higher specific impulse.

When power is limited, an optimal low-thrust mission will use the available power for higher thrust when small changes in thrust will create large savings in ΔV. When large changes in thrust have a small effect in ΔV, the thruster would use the remaining available power for an increased specific impulse. The trajectory is optimizing delivered mass with the ΔV term of Tsiolkovsky's equation having a strong dependency on thrust. Figure 8 is an example of optimal specific impulse for a rendezvous mission with the comet Kopff. The mission optimized to specific impulses of 2920s, 3175s, and 3420s, at power levels of 6kW, 7.5kW, and 9kW respectively.

A remaining consideration for designing low-thrust mission trajectories is the proper methodology of margin. The trajectory must account for planned and unplanned thruster outages, power margin, thrust margin, propellant margins due to trajectory errors, residuals

that cannot be expelled from the tank, or flow control accuracy, ΔV margins, etc. Though the margins are interdependent, the electric propulsion system can offer advantages with an ability to compensate for one area with additional margin in another (Oh et al., 2008).
In general, interplanetary missions with the greatest benefit of using electric propulsion are missions that do not capture into large gravity wells, and have very large total ΔV mission requirements. High ΔV missions include missions to multiple targets, large inclination changes, and deep space rendezvous with trip time limitations. Trajectory analyses were performed in Copernicus and MIDAS for chemical comparison and using SEPTOP, SEPSPOT, and MALTO for the low thrust solutions.

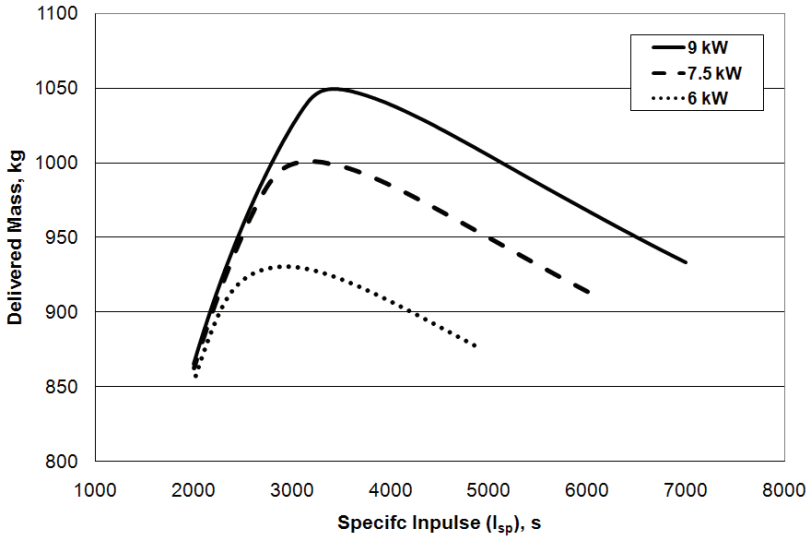



Fig. 8. Optimal specific impulse comparison for a comet rendezvous mission.

### 4.1 Multiple targets

Multi-target missions are a method to achieve considerably higher science return for a single spacecraft. Multi-target missions can range from two targets in similar orbits, several targets requiring large maneuvers, and to some extent, sample return missions.
The Dawn mission illustrates the mission enhancing capabilities of electric propulsion for just such a mission. It is the first NASA science mission to use electric propulsion. For a mission to be competitively selected and to justify new technology, the science return must be remarkably high. The Dawn mission utilizes a single spacecraft that carries an instrument suite to multiple targets, Ceres and Vesta. By traveling to multiple targets with a single spacecraft there are savings in spacecraft development, instrument development, and launch costs. The mission provides a unique opportunity to compare data from an identical sets of instruments. The Dawn mission was determined to only be viable through the use of electric propulsion. The use of chemical propulsion required significantly higher launch mass and could only feasibly reach a single target. Figure 9 illustrates the Dawn EP multi-rendezvous trajectory.

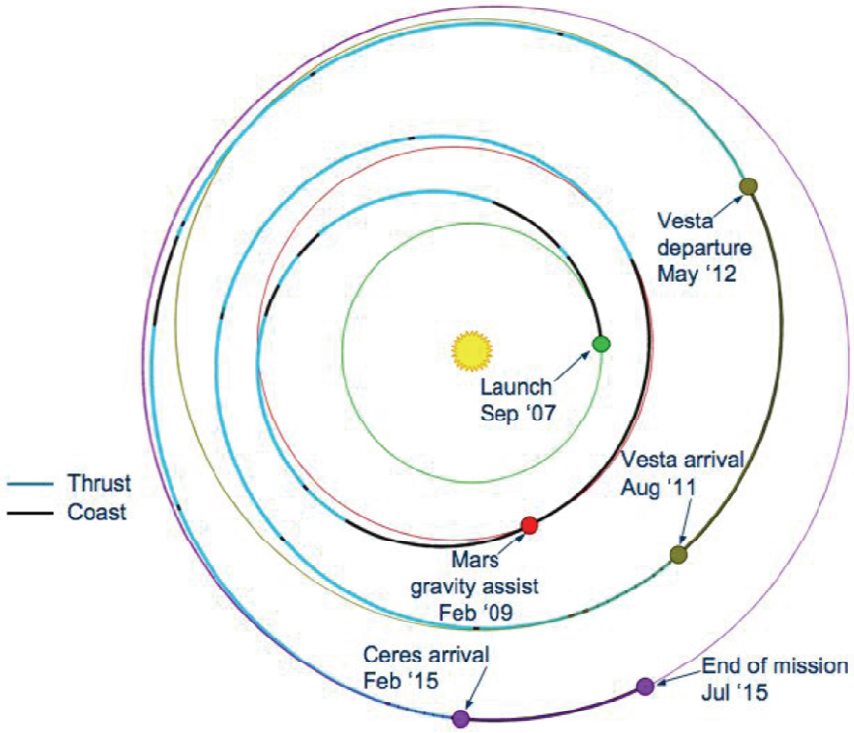



Fig. 9. Trajectory of the Dawn mission to Ceres and Vesta (Brophy et al., 2008).

For a Dawn-like mission, the use of either the NEXT or HiVHAC thruster has significant mission advantages over NSTAR. The throughput capability of a single NSTAR thruster is insufficient to complete the mission. The use of the NEXT thruster can not only deliver greater payload, but can do so using a single thruster. The use of a single thruster has system complexity and spacecraft integration advantages. The HiVHAC thruster, capable of performing the Dawn mission with a single operating thruster, is expected to have numerous advantages including greater payload and reduced cost. The Dawn mission power requirements are driven by the need to operate the thruster throughout the mission, including significant operation at relatively high AU. Figure 10 illustrates the Dawn power profile.

Using a thruster that can throttle to very low power can reduce the spacecraft power requirements, reducing the overall spacecraft cost. Because many small body rendezvous missions can benefit from higher thrust, lower specific impulse, thrusters, a low power Hall thruster can deliver greater payload than the NSTAR thruster. Finally, a Hall thruster system is expected to have lower thruster and power processing unit costs. The HiVHAC system is expected to provide significant reductions in IPS costs over SOA. Figure 11 illustrates performance and cost benefits of ISPT project technologies over SOA. Overall, the NEXT thruster can outperform the NSTAR thruster with reduced system complexity. The HiVHAC thruster, specifically designed for Discovery Class missions, has the potential for increased performance and reduced system complexity and cost (Oh, 2005).

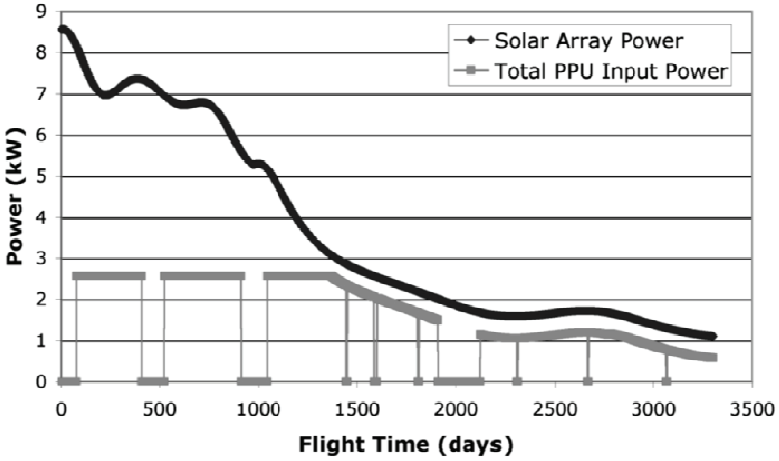



Fig. 10. Power profile for the Dawn mission (Oh, 2007).

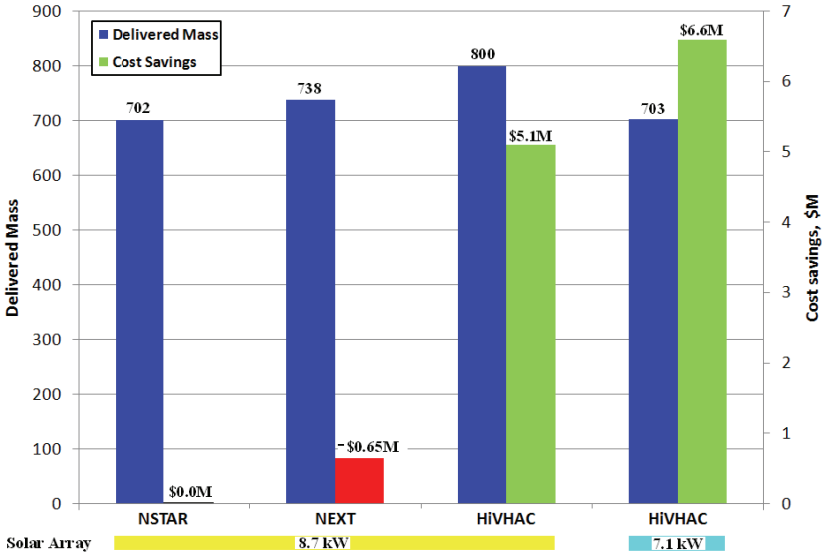



Fig. 11. Mass and cost comparison for the Dawn mission (Reference).

The extreme example of a multiple target mission would be a Super-Dawn. A "Super-Dawn" class mission is the concept of traveling to and stopping at several high interest targets with a single launch. The throughput potential of both NEXT and HiVHAC greatly opens up the trade space of achievable multi-rendezvous options. Unfortunately, most high interest targets are not necessarily co-located to allow for short transfers due to variances in inclination, eccentricity, period, etc. Sufficient throughput capability and creative mission planning could allow missions to visit multiple high interest targets and gain information from several secondary near-by targets.

The targets for a hypothetical "Super-Dawn" mission were chosen from a list of high interest targets formulated by the scientific community. Based on preliminary analysis of throughput requirements and delivered mass, a single spacecraft, with only a 5-kW array, could be used to rendezvous with four high interest near-Earth targets shown in table 1. The final delivered mass is comparable to the Dawn spacecraft. The "Super-Dawn" mission illustrates the tremendous potential of electric propulsion for these types of missions. Studies have looked at using a single spacecraft for tours of near-Earth objects, main-belt asteroids, and even Jupiter Trojans.

Sample return missions are multi-body missions because they need to return to Earth. Sample return missions are often considered high priority because of the higher fidelity science that can be performed terrestrially. Mars sample return was under investigation for many years, but the large costs of such a mission has deterred its implementation. Regolith from Phobos and Deimos are of high scientific value. The mission options offer significantly lower cost with minimal technology development required.

| Segment | Target | Start Mass, kg | Propellant Required, kg | End Mass, kg |
|---|---|---|---|---|
| 1 | Nereus | 1650 | 309 | 1341 |
| 2 | 1993 BD3 | 1341 | 52 | 1289 |
| 3 | Belenus | 1289 | 44 | 1245 |
| 4 | 1996 FG3 | 1245 | 456 | 789 |

Table 1. Table of ΔV for a "Super-Dawn" type mission.

Two concepts for a Phobos and Deimos sample return mission were evaluated using solar electric propulsion: a single spacecraft to both moons or twin spacecraft capable of returning samples from either moon. The small bodies of Phobos and Deimos, with small gravity fields (especially Deimos), make electric propulsion rendezvous and sample return missions attractive. Electric propulsion systems can be used for the transfer to Mars, and then to spiral into an orbit around the moons. Chemical systems cannot easily leverage the Oberth effect for the sample return mission from Mars' moons because of the higher altitude orbit requirement. So while the mission can be completed, it comes at a large mass penalty. Figure 12 illustrates the benefits of using electric propulsion for a Phobos and Deimos sample return mission.

Results show significant savings for using electric propulsion for Phobos and Deimos sample return missions. The baseline case uses a NEXT thruster with one operating thruster, and a spare system for redundancy (1+1). A Delta II class launch vehicle is capable of delivering enough mass for a sample return from both targets. For electric propulsion, the transfer between Phobos and Deimos has minimal mass implications. The mass and technology requirements could potentially fit within the Mars Scout cost cap.

Using an Evolved Expendable Launch Vehicle (EELV), twin electric propulsion vehicles can be sent for a low-risk approach of collecting samples from Phobos and Deimos independently. However, the use of an EELV enables a chemical solution for a sample return mission. Going to a single moon chemically remains a significant challenge and results in a spacecraft that is greater than 70 percent propellant; a mass fraction more typical of a launch vehicle stage. Launching a single chemically propelled spacecraft to retrieve samples from both moons requires staging events adding risk and complexity.

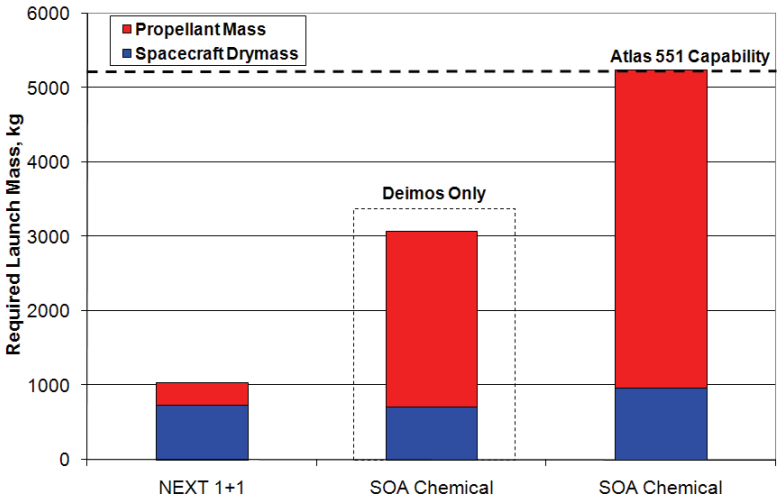



Fig. 12. Comparison of required launch mass for chemical and EP Mars' moons missions.

The use of electric propulsion was studied for various comet surface sample return (CSSR) missions. The results are highly dependant on the targets of interest. Electric propulsion compares favorably with chemical alternatives resulting in either higher performance or reduced trip times. Studies for Temple 1 (Woo et al., 2006) determined the SOA NSTAR thruster to be inadequate due to its propellant throughput capability. The mission required the use of a NEXT thruster. Studies for the comet Wirtanen (Witzberger, 2006) were conducted and determined that the NSTAR could not deliver positive payload while both the NEXT and HiVHAC thrusters can complete the mission with sufficient margin. The largest benefit is that electric propulsion enables a wide range of targets that cannot be reached using chemical propulsion systems.

In 2008, NASA GRC completed a mission design study for a multiple near-Earth asteroid sample return mission (Oleson et al., 2009). The results indicated that it is feasible to use electric propulsion to collect multiple samples from two distinct targets in very different orbits. An Earth fly-by was performed after leaving the primary target and before arriving at the second to releae the sample return capsule for a lower risk mission and mass savings to the secondary target. This mission was not feasible using chemical propulsion. The conceptual spacecraft for the multi-asteroid sample return mission is shown in figure 13.

### 4.2 Inclined targets

Other missions enabled by electric propulsion are missions to highly inclined targets. There are several Earth crossing targets that are thought to be old and inactive comets. These asteroids typically have inclined orbits. The ΔV requirement for a plane change is a function of the spacecraft velocity and angle of the plane change as shown in equation 1. With the Earth's heliocentric orbital speed near 30 km/s, a simple plane change of even 30 degrees will require a ΔV of at least 15 km/s to perform a fly-by, following equation 4.

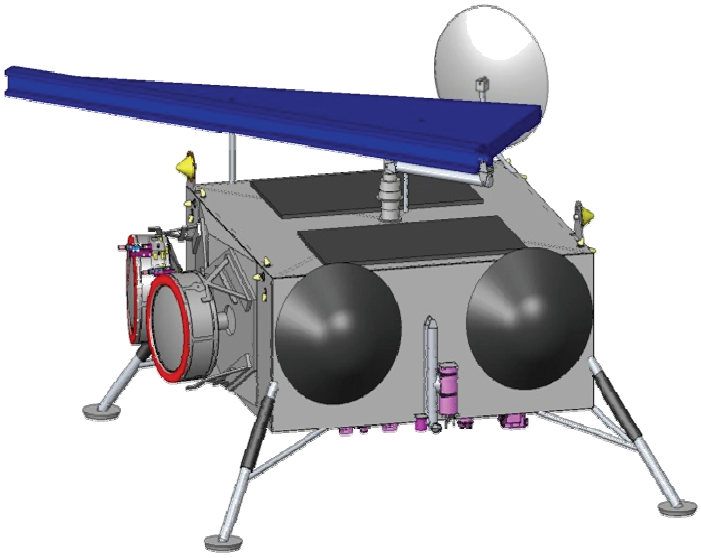


Fig. 13. NEARER spacecraft photo.

An example mission to an inclined target would be to the asteroid Tantalus. Tantalus is an Earth-crossing body with a semi-major axis of 1.29 AU, an eccentricity of 0.3, and an inclination of 64°. Because of the inclination change, the optimal chemical transfer requires a high launch velocity so that the spacecraft can perform the plane change at high AU with a lower velocity. Figure 14 illustrates the chemical transfer which requires a post-launch ΔV of 14 km/s. The Tsiolkovsky's mass fraction is only on the order of one percent dry mass, so the mission is completely infeasible with any launch vehicle using chemical propulsion.

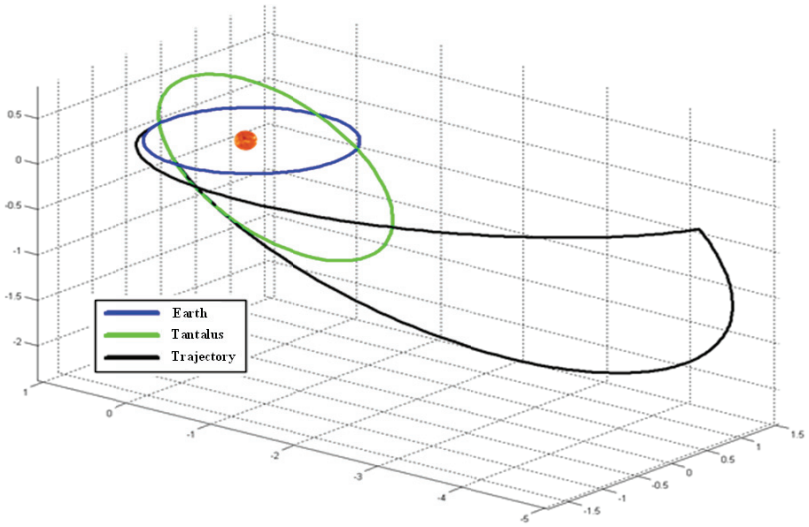



Fig. 14. Optimal chemical trajectory to Tantalus.

The electric propulsion transfer to Tantalus is also a challenging mission. The low-thrust transfer is over 30 km/s over 4.5 years, but can still deliver over 800 kg of dry mass on a rendezvous mission using an Atlas V. The mission would require two NEXT thrusters, and would not be viable with the NSTAR or Hall thruster based propulsion system. Rather than going to high AU to perform the plane change, the low-thrust transfer gradually performs the plan change through several revolutions. Figure 15 illustrates the low-thrust transfer to Tantalus. Because of the advantages of electric propulsion, efficient use of propellant and low-thrust trajectory options, scientists can plan missions to high interest targets previously unattainable.

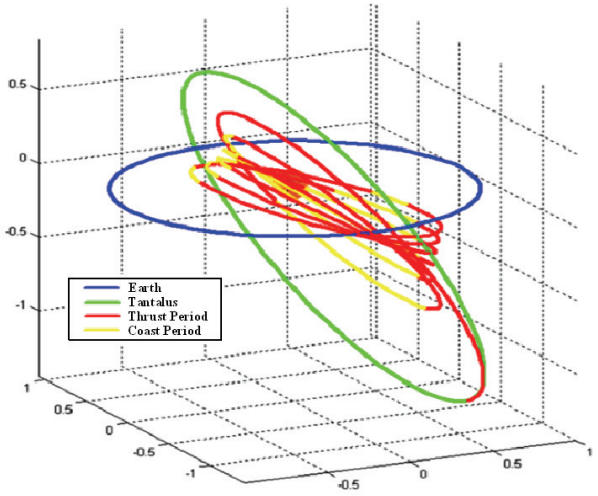



Fig. 15. Optimal low-thrust trajectory to Tantalus.

### 4.3 Radioisotope electric propulsion

Another area of interest pushing the limits of propulsion technology is the use of a radioisotope power source with an electric propulsion thruster. This achieves high post launch $\Delta V$ on deep-space missions with limited solar power. Radioisotope electric propulsion systems (REPS) have significant potential for deep-space rendezvous that is not possible using conventional propulsion options.

One example of mission that can benifit from REPS is a Centaur orbiter. The Centaurs are of significant scientific interest, and recommended by the Decadal Survey Primitive Bodies Panel as a New Frontiers mission for reconnaissance of the Trojans and Centaurs. The original recommendation was for a flyby of a Jupiter Trojan and Centaur. While a flyby mission can use imaging, imaging spectroscopy, and radio science for a glimpse at these objects, a REP mission provides an opportunity to orbit and potentially land on a Centaur. This greatly increases the science return. An exhaustive search of Centair obiter missions concluded that a wide range of Trojan flybys with Centaur Rendezvous missions are pracitical with near-term electric propulsion technology and a Stirling radioisotope generator (Dankanich & Oleson, 2008). With near-term technology, flyby missions may no longer be scientifically acceptable. Investigations are continuing using the enabling combination of electric propuslion and radioisotope power systems. On-going and recent studies include multi-Trojan landers, Kuiper-belt object rendezvous, Titan-to-Enceldaus

transfer, and Neptune Orbiters. A conceptual spacecraft design of a Centuar orbiter is shown in figure 16.

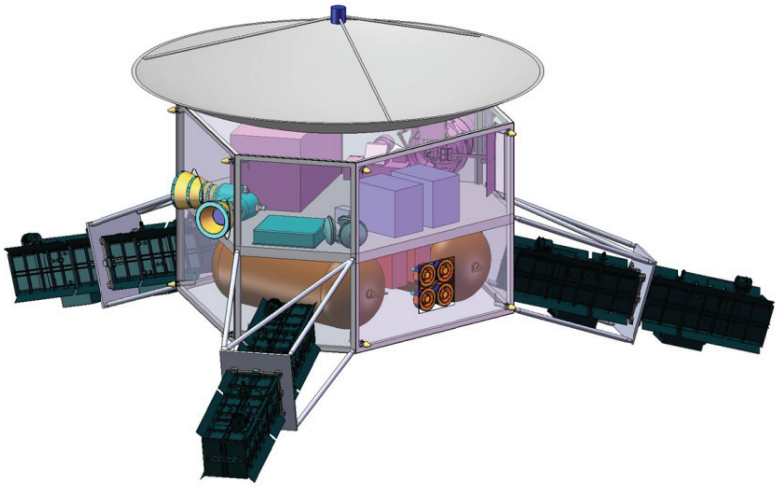


Fig. 16. Centaur orbiter spacecraft design.

## 5. Near-Earth application of electric propulsion

There are currently over 200 satellites using electric propulsion and the majority of those satellites have multiple thrusters. Only four spacecraft flew beyond geosynchronous orbit altitude, two from NASA, one from JAXA to a near-Earth object, and one from ESA to the moon. The vast majority of mission pull of electric propulsion is for application in the region between low-Earth orbit (LEO) and geosynchronous Earth orbit (GEO,) specifically commercial satellites with GEO operational orbits.

A conjunction of developments greatly increased the practicality and expectations for the use of electric propulsion for Earth-orbit transportation. Key among these advancements are trends of satellite power-to-mass ratios, required spacecraft mass, advancements in space power technology, and the broad use of electric propulsion systems. An illustration of the trends of both increasing power-to-mass ratios and spacecraft mass for commercial communication satellites are shown in figure 17 (Byers & Dankanich, 2008).

Several studies were performed to evaluate the use of electric propulsion for transfer starting at LEO. A strong deterrent is the very long transfer times. One near-term option is to leverage the launch vehicle to an eccentric orbit. The primary design capability of commercial launch vehicles is to a geosynchronous transfer orbit (GTO) with perigee and apogee altitudes of 185 km and 35,786 km respectively. Launching directly to GTO will significantly reduce the orbit transfer time while also reducing risk of orbital debris and exposure to the radiation environment within the Van Allen belts. Analyses for transfer times and $\Delta V$ requirements from equatorial launch sites (Baikonur, Kourou, and the Kennedy Space Center (KSC)) were evaluated and characterized as a function of spacecraft power-to-mass ratios (Dankanich & Woodcock, 2006). Results for KSC launches are shown in figures 18 and 19 respectively. Above 3 W/kg, the acceleration is high enough to

minimize gravity losses of the transfer. The low-thrust ΔVs are approximately 800 m/s more than a chemical GEO insertion.

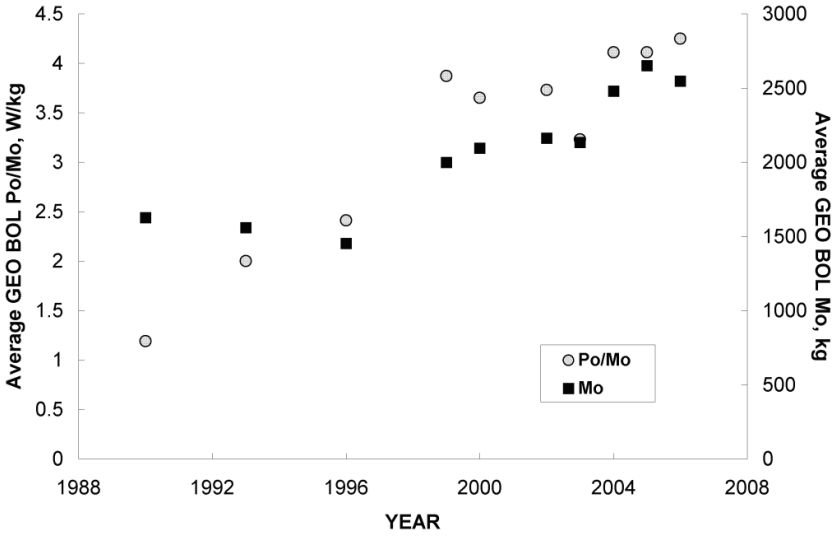



Fig. 17. Trends of commercial satellite beginning-of-life (BOL) P/M ratio and average mass.

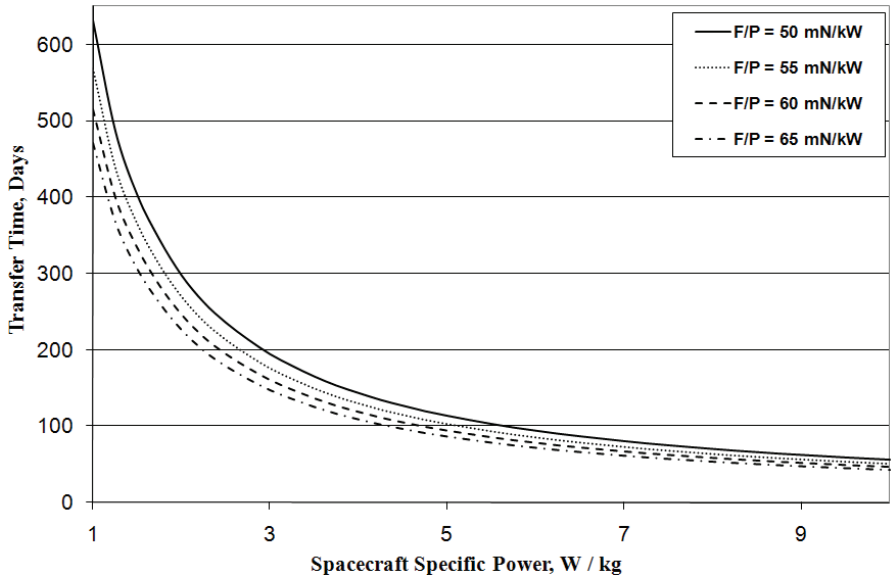



Fig. 18. GTO-to-GEO transfer times as a function of spacecraft specific power.

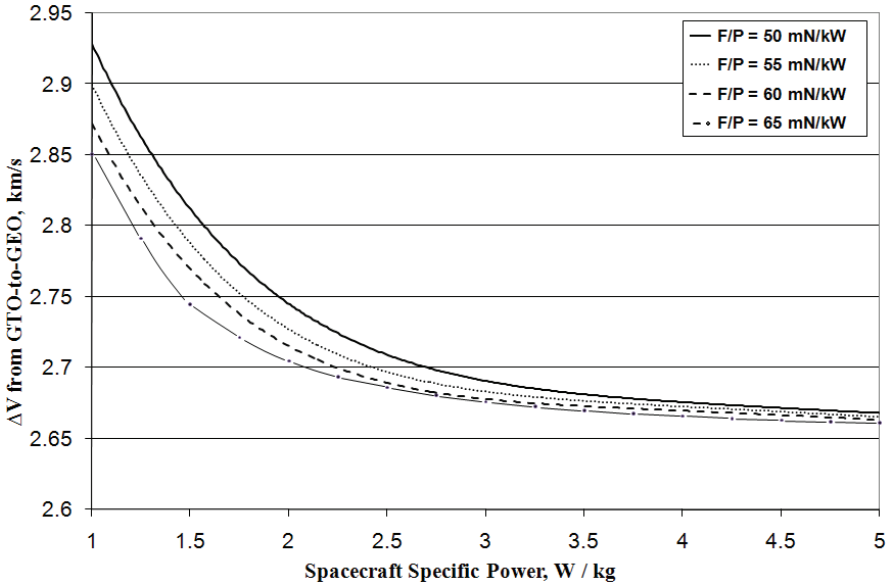



Fig. 19. Required ΔV from GTO-to-GEO as a function of spacecraft specific power.

The GTO-to-GEO transfer time and ΔV is dependant on the launch site, or initial starting inclination. Figure 20 illustrates the penalty of launch at inclined launch sites and the benefit of near-equatorial launches.

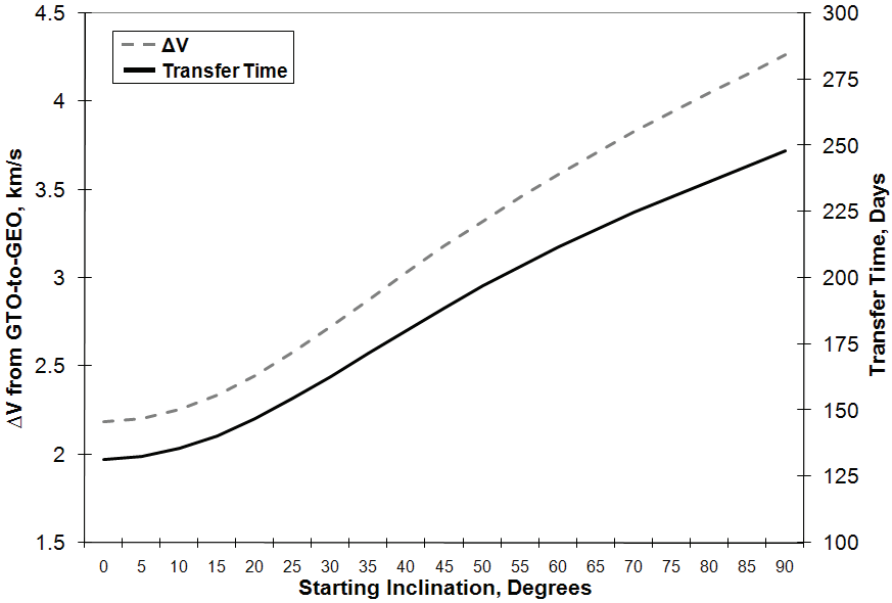



Fig. 20. Effect of starting inclination on transfer time and ΔV from GTO-to-GEO.

There were 32 commercial communication satellites launched in 2005 and 2006 as provided by the Union of Concerned Scientists database. These specific satellites were evaluated for potential to use an integrated electric propulsion system with a specific impulse of 1000 seconds, 1500 seconds, and 2100 seconds. Integrated electric propulsion systems assume the use of 95% of the onboard solar array power of the spacecraft as launched.
Using electric propulsion for the GEO insertion has significant mass benefits. Typically this is evaluated as a method to leverage the launch vehicle performance to deliver the greatest possible mass. Another perspective is to evaluate the potential for existing launch vehicles to meet the demands of the COMSAT market. Figure 21 illustrates that currently launch vehicles with GTO drop mass capabilities in excess of 7,500kg are required for a complete market capture. However, using electric propulsion, a launch vehicle with a drop mass capability of 5,500kg can have complete market capture. A low cost launcher with a capability to deliver 3,500kg to 5,500kg can create a paradigm shift in the commercial launch market. This assumes the commercial entity is willing to endure the long transfer time, ranging from 66–238 days, depending on the spacecraft power-to-mass ratio and EP thruster selected.

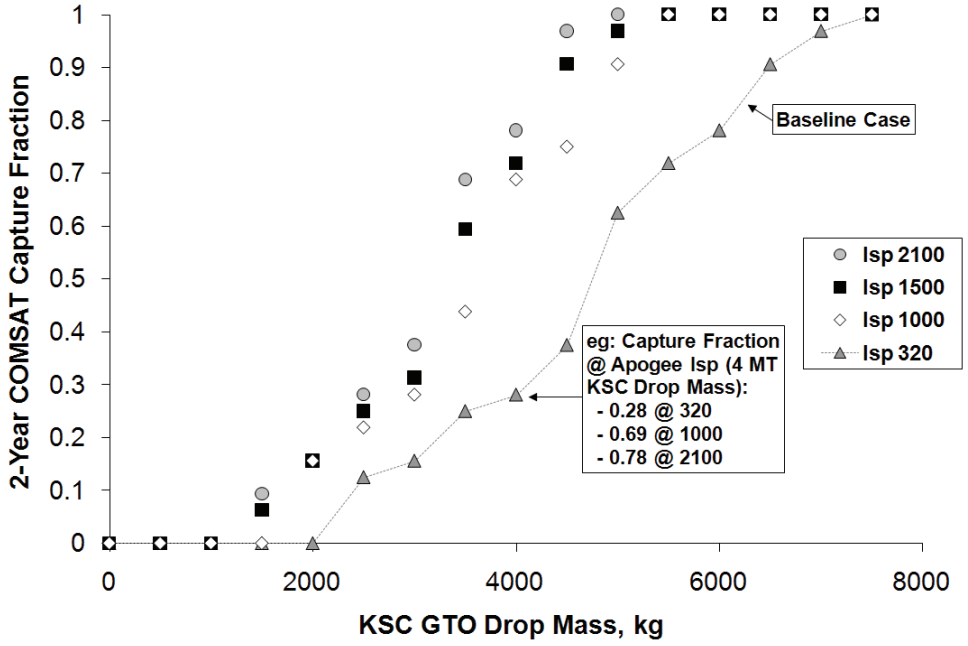



Fig. 21. Capture fraction as a function of GTO drop mass for various propulsion options.

## 6. Conclusion

Electric propulsion technology is widely used today, and multiple thrusters exist for primary electric propulsion application. NASA and the U.S. commercial market developed several thrusters suitable for primary electric propulsion on full scale spacecraft. The

technology drivers for new electric propulsion thrusters include: ability to use available power (i.e. high maximum power with large throttle range), increased total throughput capability, and lower cost systems and integration. The optimal specific impulse is limited by thrust required to minimize propulsive inefficiencies and available power. Due to power constraints, the optimal specific impulse is typically less than 5,000s and closer to 2,000s for near-Earth application. Electric propulsion is an enabling technology for a large suite of interplanetary missions. Several targets are infeasible with advanced chemical propulsion technologies, while practical with today's electric propulsion options. Electric propulsion is well suited for missions with very high post-launch ΔVs including multi-target missions, sample return missions, deep-space rendezvous, and highly inclined targets. Electric propulsion has tremendous capability to impact the commercial launch market by leveraging on-board available power. Today's commercial satellites have mass-to-power ratios for practical GTO-to-GEO low-thrust transfer. As available power and performance demand continues to rise, electric propulsion technologies will continue to supplant chemical alternatives for a wide range of missions. The technology will continue to focus on developing lower cost propulsion systems with higher power and longer lifetime capabilities.

## 7. References

Brophy, J. R. (2007). Propellant Throughput Capability of the Dawn ion Thrusters, IEPC-2007-279, 30th International Electric Propulsion Conference, Florence, Italy, September 2007.

Brophy, J., Rayman, M. D., & Pavri, B. (2008). Dawn: An Ion-propellanted Jounrey to the Beginning of the Solar System, IEEE Aerospace Conference, Big Sky, MT, March 2008.

Byers, D., & Dankanich, J. W. (2008). Geosynchronous-Earth-Orbit Communication Satellite Deliveries with Integrated Electric Propulsion. *Journal of Power and Propulsion,* Vol. 24, No. 6, November–December 2008, pp 1369–1375.

Dankanich, J. W. & Oleson, S. R. (2008). Radioisotope Electric Propulsion (REP) Centaur Orbiter Mission Design, 44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Hartford, CT, July 2008.

Dankanich, J. W., & Woodcock, G. R. (2007). Electric Propulsion Performance from GEO-Transfer to Geosynchronous Orbits, International Electric Propulsion Conference, Florence, Italy, September 2007.

Kamhawi, H., Manzella, D., Pinero, L., & Mathers, A. (2009). Overview of the High Voltage Hall Accelerator Project, AIAA 2009-5282, 45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Denver, CO, August 2009.

Manzella, D. (2007). Low Cost Electric Propulsion Thruster for Deep Space Robotic Missions, 2007 NASA Science Technology Conference, University of Maryland, MD, June 2007.

Oh, D. (2007). Evaluation of Solar Electric Propulsion Technologies for Discovery-Class Missions. *Journal of Spacecraft and Rockets*, Vol. 44, No. 2., March-April 2007, pp 399–411.

Oh, D., Witzberger, K., & Cupples, M. 2004). Deep Space Mission Applications for NEXT: NASA's Evolutionary Xenon Thruster, AIAA-2004-3806, 40th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Fort Laudersale, FL, July 2004.

Oleson, S. et al. (2009) Near-Earth Asteroid Sample Return Mission, 31st International Electric Propulsion Conference, Ann Arbor, MI, September 2009 (to be publsihed).

Patterson, M. & Benson, S. (2007). NEXT Ion Propulsion System Development Status and Performance, AIAA-2007-5199, 43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Cincinatti, OH, July 2007.

Welander, B., Carpenter, C., de Grys, K., Hofer, R. R., Randolph, T. M., & Manzella, D. H. (2006). Life and Operating Range Extension of the BPT-4000 Qualification Model Hall Thruster, AIAA 2006-5263, 42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Sacramento, CA, July 2006.

Wilson, F., & Smith, B. (2006). Hall Thruster System Qualification Provides Major Satellite Benefits, IAC-06-C4.4.09, 57th International Astronautical Congress, Valencia, Spain, October 2006.

Witzberger, K. E. (2006). Solar Electric Propulsion for Primitive Body Science Missions, NASA TM 2006-214236, March 2006.

Woo, B., Coverstone, V. L., & Cupples, M. (2006). Application of Solar Electric Propulsion to a Comet Surface Sample Return Missions. *Journal of Spacecraft and Rockets,* Vol. 43, No. 6, November–December 2006, pp 1225–1230.

# 13

# Global GNSS Radio Occultation Mission for Meteorology, Ionosphere & Climate

Nick L. Yen[1], Chen-Joe Fong[1], Chung-Huei Chu[1], Jiun-Jih Miau[1],
Yuei-An Liou[2], and Ying-Hwa Kuo[3]
[1]*National Space Organization*
[2]*Center for Space and Remote Sensing Research, National Central University*
[3]*University Corporation for Atmospheric Research*
[1,2]*Taiwan*
[3]*U.S.A.*

## 1. Introduction

The term "occultation" is widely used in astronomy when an object in the foreground optically occults objects in the background, and it refers to a geometry involving the emitter, the planet and its atmosphere if any, and the receiver changes with time.[1] Radio occultation (RO) is a remote sensing sounding technique in which a radio frequency signal emitted from a spacecraft passes through an intervening planetary atmosphere before arriving at the receiver, and is used to study the planetary atmosphere properties in the interplanetary mission (Fjeldbo & Eshleman, 1965). The atmospheric RO observations represent a planetary scale geometric optics experiment in which the atmosphere acts as a big optical lens and refracts the paths and propagation velocity of electromagnetic wave signals passing through it (Kursinski et al., 2000). The first RO experiment started with the Mars flyby by Mariner-IV in 1964 (Kliore et al., 1965). When Mariner-IV satellite passed behind and emerged from the other site of Mars, the extra carrier phase delay and amplitude variation of the microwave signals were observed. These observed data provided very first valuable atmospheric and ionospheric density information by using the inversion techniques (Melbourne et al., 2005). Since then a series of planetary experimental missions were planned to study the atmospheres and ionospheres of the planets and their moons (Yunck et al., 2000).

The limb sounding of the earth's atmosphere and ionosphere using the RO technique can be performed with any two cooperating satellites before the United States' Global Positioning System (GPS), the first Global Navigation Satellite Systems (GNSS), becoming operational (Lusignan et al., 1969). A few early RO experiments from a satellite-to-satellite tracking link had been conducted. These included the occulted radio link between ATS-6 (Applications Technology Satellite-6) and GEOS-3 (Geodetic and Earth Orbiting Satellite-3) and between the Mir station and a geostationary satellite (Liu et al., 1978; Yakovlev et al., 1996).

---

[1] http://en.wikipedia.org/wiki/Occultation [cited 1 July. 2009].

## 2. GNSS radio occultation mission

After GNSS becomes operational, substantial and significant progress has been made in the science and technology of ground-based and space-based GNSS atmospheric remote sensing over the past decade (Davis et al., 1985). The ground-based GNSS atmospheric remote sensing with upward-looking observations arose in the 1980s from GNSS geodesy. As the rapid increase of the GNSS geodetic ground networks around the world, great quantity of atmospheric integrated perceptible water (PW) were used in numerical weather prediction (NWP) for weather and climate modelling (Liou et al., 2000 & 2001; Elgered et al., 2003; Ha et al., 2003). However, one of the major limitations to the ground-based GNSS remote sensing is that it only provides integrated PW without vertical resolution, and it is restricted to land areas distributed with GNSS networks. The space-based GNSS atmospheric limb sounding offers a complementary solution to these issues (Yunck et al., 2003).

The space-based GNSS RO atmospheric remote sensing technique, which makes use of the L-band radio signals transmitted by the GNSS satellites, has emerged as a powerful approach for sounding the global atmosphere in all weather over both lands and oceans (Yunck et al., 1990 & 2003; Wu et al., 1993; & Liou et al., 2002). Figure 1 shows a schematic diagram illustrating radio occultation of GNSS signals received by a low-earth-orbit satellite. The GPS/Meteorology (GPS/MET) experiment (1995-1997) showed that the GNSS RO technique offers great advantages over the traditional passive microwave measurements of the atmosphere by satellites and became the first space-based "proof-of-concept" demonstration of GNSS RO mission to earth (Ware et., 1996; Kursinski et al., 1996; Rius et al., 1998; Anthes et al., 2000; Hajj et al., 2000; Kuo et al., 2000). For a more complete history of GNSS RO see Yunck et al. (2000) and Melbourne et al. (2005).

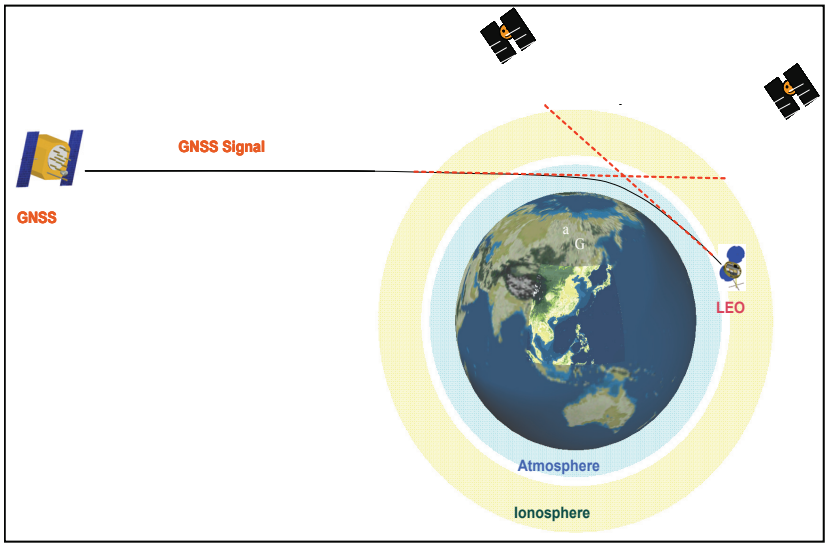



Fig. 1. Schematic diagram illustrating radio occultation of GNSS signals.

The extraordinary success of GPS/MET mission had inspired a series of other RO missions, e.g., the Ørsted (in 1999), the SUNSAT (in 1999), the Satellite de Aplicaciones Cientificas-C (SAC-C) (in 2001), the Challenging Minisatellite Payload (CHAMP) (in 2001), and the twin

Gravity Recovery and Climate Experiment (GRACE) missions (in 2002). The GNSS RO sounding data have been shown to be of high accuracy and high vertical resolution. Table 1 lists GNSS RO sounding data characteristics. All these missions set the stage for the birth of the FORMOSA SATellite mission-3/Constellation Observing Systems for Meteorology, Ionosphere, and Climate mission, also known as FORMOSAT-3/COSMIC mission Kursinski et al., 1996; Rius et al., 1998; Anthes et al., 2000; Hajj et al., 2000; Kuo et al., 2000; Lee et al., 2001).

| Characteristics of GNSS Radio Occultation Data |
|---|
| • Limb sounding geometry complementary to ground and space nadir viewing instruments<br>• Global 3-D atmospheric weather coverage from 40 km to sea level surface<br>• High accuracy temperature measurement (equivalent to <1 K; average accuracy <0.1 K)<br>• High precision temperature measurement (0.02-0.05 K)<br>• High vertical resolution (0.1 km surface – 0.1 km tropopause)<br>• Only system from space to resolve atmospheric boundary layer<br>• All weather-minimally affected by aerosols, clouds or precipitation<br>• Independent height and pressure<br>• No first guess sounding requirement<br>• Independent of radiosonde calibration<br>• No instrument drift<br>• No satellite-to-satellite bias<br>• Compact sensor, low power, low cost<br>• A typical RO sounding showing very sharp tropopause<br>• No other instrument from space provides such high vertical resolution profile |

Table 1. Characteristics of GNSS radio occultation data

## 3. FORMOSAT-3/COSMIC mission

### 3.1 Mission

The FORMOSAT-3[2] satellite constellation was launched successfully from Vandenberg Air Force Base in California 1:40 UTC on April 15, 2006 into the designated 516 km circular parking orbit altitude. Table 2 shows the mission characteristics. The FORMOSAT-3 mission is the world's first demonstration of GPS RO occultation near real-time operational constellation mission for global weather monitoring. The primary scientific goal of the mission is to demonstrate the value of near-real-time GPS RO observation in operational numerical weather prediction. With the ability of performing both rising and setting occultation, the mission provides about 1,600~2,400 atmospheric and ionospheric soundings per day in near real-time that give vertical profiles of temperature, pressure, refractivity, and water vapor in neutral atmosphere, and electron density in the ionosphere with global coverage (Anthes et al., 2000 & 2008; Liou et al., 2006a, 2006b, & 2007; Fong et al., 2008a & 2009a).

[2] In this chapter the FORMOSAT-3/COSMIC mission was referred to as the FORMOSAT-3 mission for simplicity.

The retrieved RO weather data are being assimilated into the NWP models by many major weather forecast centers and research institutes for real-time weather predictions and cyclone/typhoon/hurricane forecasts (Kuo et al., 2004; Anthes et al., 2008). The mission results have shown that the RO data from FORMOSAT-3 are of better quality than those from previous missions and penetrate much further down into the troposphere, mission results could be referred to Liou et al. (2007), Anthes et al. (2008), Fong et al. (2008a, 2008b, 2008c & 2009a), and Huang et al. (2009). In the near future, other GNSS, such as the Russian Global Navigation Satellite System (GLONASS), and the planned European Galileo system, could be used to extend the applications by the use of RO technique (Chu et al., 2008; Fong et al., 2009a & 2009b). The great success of the FORMOSAT-3 mission expected to operate through 2011, has initiated a new era for near real-time operational GNSS RO soundings (Fong et al., 2009b; Kuo et al., 2004, 2008a & 2008b).

| | |
|---|---|
| Number | Six identical satellites |
| Weight | ~ 61 kg (with payload and fuel) |
| Shape | Disc-shape of 116 cm diameter, 18 cm in height |
| Orbit | 800 km altitude, circular |
| Inclination Angle | $72^{o}$ |
| Argument of latitude | $52.5^{o}$ apart |
| Power | ~ 81 W orbit average |
| Communication | S-band uplink (32 kbps) and downlink (2 Mbps) |
| Sounding | 1,600 ~ 2,400 soundings per day |
| Data Latency | 15 minutes to 3 hours |
| Design and Mission life | 5 years |
| Launch date | April 15, 2006 |

Table 2. The FORMOSAT-3 mission characteristics

### 3.2 System architecture

Figure 2 shows the FORMOSAT-3 system architecture. After two years' in orbit operations, starting from mid-April 2008, the FORMOSAT-3 program switched and changed from two commercially operated ground stations at Fairbanks, Alaska and Kiruna, Sweden, operated by United Service Network (USN), to two new ground stations in Fairbanks and Tromso, Norway, operated by National Oceanic and Atmospheric Administration (NOAA). The constellation operation plans to use the new stations for the remaining five-year mission. The FORMOSAT-3 constellation system consists of the six microsatellites, a Satellite Operations Control Center (SOCC) in Taiwan, several tracking, telemetry and command (TT&C) ground stations, two data receiving and processing centers, and a fiducial network. There are two TT&C Local Tracking Stations (LTS), one located in Chungli, Taiwan and the other in Tainan, Taiwan, respectively. Currently there are four Remote Terminal Stations (RTS) to support the passes: Fairbanks Command and Data Acquisition Station (FCDAS), and Kongsberg Satellite Services Ground Station (KSAT), which are currently set as primary stations for the FORMOSAT-3 mission, and the Wallops station at Virginia, USA and the McMurdo station located in McMurdo, Antarctica. The latter two RTS stations provide partial support for the mission (Fong et al., 2009a; Rocken et al., 2000).

The SOCC uses the real-time telemetry and the back orbit telemetry to monitor, control, and manage the spacecraft state-of-health. The downlinked science RO data is transmitted from the RTS via National Oceanic and Atmospheric Administration (NOAA) to CDAAC (COSMIC Data Analysis and Archive Center) located at Boulder, Colorado, USA, and TACC (Taiwan Analysis Center for COSMIC) located at Central Weather Bureau (CWB) in Taiwan. The fiducial GNSS data is combined with the occulted and referencing GNSS data from the GOX payload to remove the clock errors. All collected science data is processed by CDAAC and then transferred to TACC and other facilities for science and data archive (Wu et al., 2006).

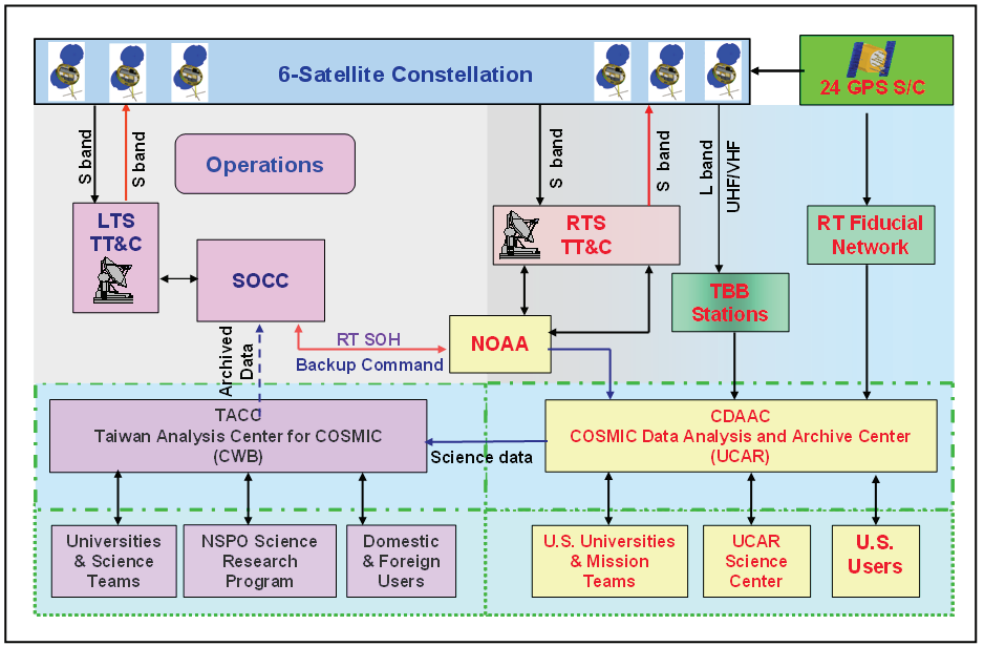



Fig. 2. The FORMOSAT-3 system architecture

The processed results are then passed to the National Environmental Satellite, Data, and Information Service (NESDIS) at NOAA. These data are further routed to the weather centers in the world including the Joint Center for Satellite Data Assimilation (JCSDA), National Centers for Environment Prediction (NCEP), European Centre for Medium-range Weather Forecast (ECMWF), Taiwan CWB, UK Meteorological Office (UKMO), Japan Meteorological Agency (JMA), Air Force Weather Agency (AFWA), Canadian Meteorological Centre (Canada Met), French National Meteorological Service (Météo France), etc. They are made ready for assimilation into weather prediction models. The data is currently provided to weather centers within 180 minutes data latency requirement in order to be ingested by the operational weather forecast model (Fong et al., 2009b).

## 4. FORMOSAT-3 satellite design

Figure 3 illustrates the FORMOSAT-3 satellite designed by Orbital Science Corporation in a deployed configuration and its major components. The FORMOSAT-3 satellite avionic block

diagram is shown in Figure 4. The major subsystem elements of the spacecraft system are Payload Subsystem, Structure and Mechanisms Subsystem (SMS), Thermal Control Subsystem (TCS), Electrical Power Subsystem (EPS), Command and Data Handling Subsystem (C&DH), Radio Frequency Subsystem (RFS), Reaction Control Subsystem (RCS), Attitude Control Subsystem (ACS) and Flight Software Subsystem (FSW). The spacecraft bus provides structure, RF power, electrical power, thermal control, attitude control, orbit raising, and data support to the instrument (Fong et al, 2008a, 2008b & 2009a). Table 3 shows the spacecraft bus key design features.

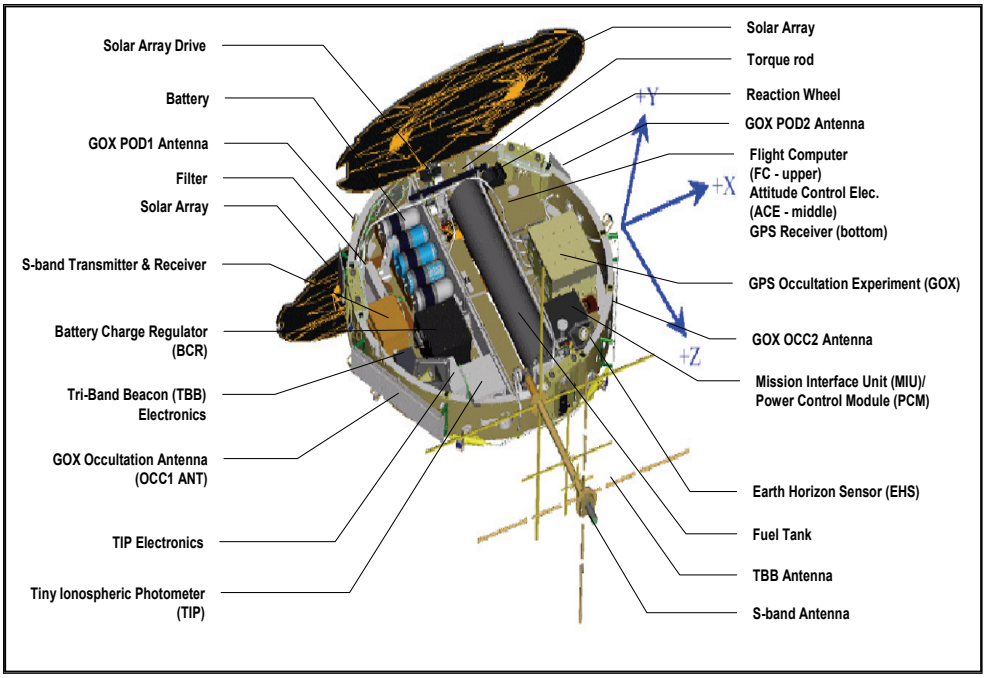



Fig. 3. The FORMOSAT-3 satellite in deployed configuration and its major components.

| | |
|---|---|
| Structure | Metal Matrix (AlBeMet) |
| Science Data Storage | 128 MB |
| Distributed Architecture | Motorola 68302 Microprocessor |
| Attitude Control | Magnetic 3-axis Control<br>Pointing Control = 5° Roll & Yaw, 2 ° Pitch |
| Propulsion | Hydrazine Propulsion Subsystem |
| S-Band Communications | HDLC Command Uplink (32 kbps)<br>CCSDS Telemetry Downlink (2 Mbps) |
| Single String Bus | Constellation Redundancy |

Table 3. The FORMOSAT-3 spacecraft bus key design.

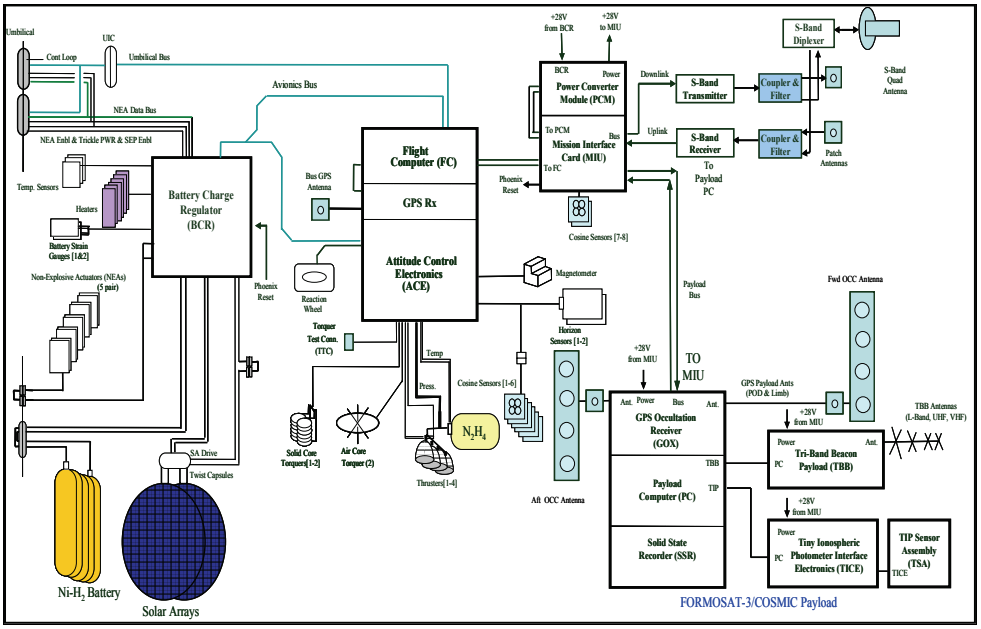



Fig. 4. The FORMOSAT-3 satellite avionic block diagram.

## 5. GNSS RO payload

### 5.1 Development of US GNSS RO receiver

In the United States, the development of space-based GPS RO receiver probing the atmospheric properties can be tracked to GPS/MET mission. Starting from 2000, the GPS occultation measurement instrument has evolved from TurboRogue geodetic receiver to a high-precision space-rated GPS receiver with dual-frequency tracking capability – "BlackJack" built by JPL. The Blackjack is an unclassified receiver, and uses a patented codeless processing technique that allows it to utilize the P-code signal without knowledge of the encryption code. The Blackjack is controlled through flexible and versatile software implementations of various receiver functions. BlackJack GPS flight receivers has been used on the following space missions such as SRTM (in 2000), SAC-C (in 2000), CHAMP (in 2000), JASON-1 (in 2000), VCL (in 2000), FEDSat (in 2001), ICESat (in 2001), and GRACE (in 2001) (Franklin et al., 2009). They have generated a lot of useful data in the areas of geophysical research.

FORMOSAT-3 carries the integrated GPS Occultation Receiver (IGOR or GOX) which is based on the NASA/JPL BlackJack space-borne GPS Receiver built by Broad Reach Engineering (BRE). The Pyxis receiver is the next generation of GNSS RO receivers based on the highly successful IGOR receiver. The Pyxis incorporates the lessons learned from the IGOR design and implements a number of improvements and upgrades. The addition of L2C and L5 frequencies and eventually Galileo frequencies provide increased Occultation Data and improved PVT resolution.

### 5.2 Development of European GNSS RO receiver

In Europe, Saab Ericsson Space (now RUAG), ESA, EUMETSAT built GRAS (GNSS RO Receiver for Atmospheric Sounding), which is an atmospheric sounding instrument carried by Metop satellites. GRAS has very low measurement noise. The mission is to provide data to operational meteorology and climate (Bonnedal, 2009). THALES develops ROSA (Radio Occultation Sounder for Atmosphere) to support international missions of OCEANSAT (India), and SAC-D (Argentina) (Fuggetta et al., 2009).

### 5.3 Development of advanced GNSS RO payload

Based on the 2007 National Research Council (NRC) publication report of "Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond," which is also referred to as Earth Sciences' Decadal Survey (SSB, 2007), and with the lessons learned and experience on GNSS RO receiver, the next generation of radio occultation instrument that will track additional new GNSS signals is under development. In addition to GPS signal, the developers also consider to include GLONASS (CDMA) and Galileo (E1, E5). The new advanced GNSS RO payload tentatively called TriG by JPL continues the evolution of hardware. The TriG hardware platform provides significantly more processing power than the IGOR platform to accommodate new signals and will have digital beam steering antenna capability, more channels, more memory available, and provide "wider" open loop tracking function (Franklin et al., 2009).[3]

## 6. Scientific contributions

Below we summarize the major scientific contributions of the current FORMOSAT-3 constellation mission using the GOX (or IGOR) payload (Yen & Fong, 2009; Anthes et al., 2008):

### 6.1 GNSS RO measurement technique

FORMOSAT-3 is the first mission that makes use of the revolutionary open-loop tracking technique. This allows more than 90% of the GNSS RO soundings to penetrate through the bottom 1 km over high latitudes, and more than 70% of soundings over the tropics. In comparison with earlier mission, such as CHAMP, only about 10% of the soundings penetrate below 1 km over the tropics.

FORMOSAT-3 is the first satellite mission that allows detection of the atmospheric boundary layer (ABL) from space. The RO soundings can be used to provide global measurements of ABL heights and their seasonal and geographical variations. These observations are crucial for the understanding of climate processes as well as tropical weather prediction.

During the early stage of constellation deployment, measurements from different FORMOSAT-3 satellites were used to determine the precision of RO measurements. Results from FORMOSAT-3 show that the precision of RO measurements is as high as 0.01%, and about an order of magnitude better than radiosonde system. Such a study was not possible before the launch of FORMOSAT-3.

[3] http://gnssro.geolinks.org/presentations

### 6.2 Operational weather prediction and meteorological research

FORMOSAT-3 data were used to support operational weather prediction by many weather centers worldwide within a year after the data were released to the public. ECMWF started operational assimilation of RO data from FORMOSAT-3 on December 12, 2006, NCEP on 1 May 2007, and UKMO on 15 May 2007, and Météo France in September 2007. All these operational centers reported significant positive impacts with the assimilation of FORMOSAT-3 data. ECMWF showed that the temperature prediction at 100 hPa over the Southern Hemisphere was improved by 10% for the first two days of forecast, and the impacts remain positive through 10 day forecast.

FORMOSAT-3 data were found to be extremely valuable for the prediction of tropical cyclogenesis. The genesis of Hurricane Ernesto (2006), that took place in August 2006 over western Atlantic, was successfully predicted only when FORMOSAT-3 data were assimilated into the model. Because GNSS RO soundings are not affected by clouds and precipitation, RO data provided valuable information on the thermodynamic structure of the hurricane environment. The assimilation of FORMOSAT-3 data produced a much more realistic analysis of low-level moisture, which was crucial for the successful prediction of the genesis of the storm.

The study of the June 2007 Mei-Yu season showed that the assimilation of FORMOSAT-3 data significantly improved the analysis of the subtropical high-pressure system over the western North Pacific. As a result, more realistic moisture fluxes are predicted, and subsequently, more accurate rainfall prediction over the Taiwan area was produced. Typhoon forecast experiments during the 2008 typhoon season showed that the FORMOSAT-3 data improved the prediction of the typhoon track by about 15% over a three-day period. The data also improved the forecast of typhoon intensity, though as not as large as the track improvement.

### 6.3 Climate research

FORMOSAT-3 is the world's first GPS RO constellation mission that provides uniform global coverage. With six satellites and a 100-min orbit, it takes 16 days for FORMOSAT-3 to provide a uniform coverage for all latitude and all local times. In contrast, it would take 6 months for CHAMP mission to obtain uniform coverage in latitudes and local time. This provides a significant advantage in climate monitoring, as we minimize the aliasing of diurnal variations into signals of climate changes.

FORMOSAT-3 data have been used to evaluate the accuracy of traditional microwave satellite sounder data and radiosonde data, allowing systematic errors of these data to be detected. The comparison of FORMOSAT-3 data with other satellite and radiosonde data allows a robust climate record to be developed for long-term climate monitoring.

The comparison of retrieval results from four different GPS RO processing centers gives essentially the same trends and changes based on CHAMP and FORMOSAT-3 data. This highlights the robustness of GPS RO measurements for climate change detection.

### 6.4 Space weather – three dimensional observation

FORMOSAT-3 is a constellation formation that observes about 2500 vertical ionospheric electron density profiles per day that are globally and uniformly distributed. Having such a global and dense set of occultation observations, a three-dimensional ionospheric electron density can be constructed in a timely manner. A new era of ionospheric space weather

study has been envisioned and numerous new ionospheric structure and application can be investigated by this unique and powerful constellation.

## 7. FORMOSAT-3 follow-on mission

### 7.1 Supporting recommendations

As addressed in the Final Report of "Workshop on the Redesign and Optimization of the Space based Global Observing System," the World Meteorological Organization (WMO) in 2007 had recommended continuing RO observations operationally and the scientific community had urged continuation of the current mission and planning for a follow-on operational mission. The WMO also calls for the international collaboration to form global constellation for RO soundings with high number of small satellites in support of the Societal Benefits Areas (SBA) of the Global Earth Observation System of Systems (GEOSS) including weather and climate (GCOS, 2003, 2004, 2006a, & 2006b; WMO, 2007). Also from the Earth Sciences' Decadal Survey (SSB, 2007), the committee on earth science and applications from space, Space Studies Board, recommended in the NRC decadal observing plan that NOAA should increase investment in identifying and facilitating the transition of demonstrably useful research observations to operational use. The committee also recommended that NOAA should transition to operations from three research observation. The three missions are vector sea-surface winds; GNSS radio occultation temperature, water vapor, and electron density soundings; and total solar irradiance. A GNSS RO mission is listed as one of the high-priority observations and missions identified by the committee.

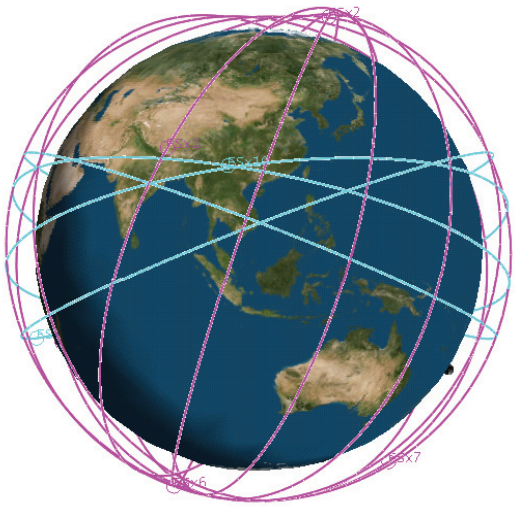
The FORMOSAT-3 Follow-on/COSMIC-II[4] mission will be a much improved constellation system for research and operation mission. The primary payload of the Follow-on satellite will be equipped with the GNSS RO receiver and will collect more soundings per receiver by adding tracking capability to receive signals from European GALILEO system and Russian's Global Navigation Satellite System (GLONASS), which will produce a significantly higher spatial and temporal density of profiles. These will be much more useful for weather prediction models and also severe weather forecasting including typhoons and hurricane, as well as for related research in the fields of meteorology, ionosphere and climate (Yen & Fong, 2009; Fong et al., 2009b).

### 7.2 Mission planning

The FORMOSAT-3 Follow-on mission is contemplated to be a 12 satellites constellation. Figure 5 shows the proposed follow-on mission spacecraft constellation configuration. The primary mission objective is to increase RO data profiles to efficiently transition into the global reliable operational constellation system for related research and operational numerical weather prediction. The effective cover area per radio occultation profile per day is expected to be 200 km x 200 km (compared to FORMOSAT-3 mission at 450 km x 450 km) when all 12 new satellites are deployed into the constellation formation. The denser RO distribution will enhance the impacts for weather/climate research and forecast in the world. The expected radio occultation profiles in the follow-on mission should be no less than 8,000 on the average per day with the averaged data latency within 90 minutes (Yen & Fong, 2009; Fong et al., 2009b).

[4] In this chapter the FORMOSAT-3 Follow-on/COSMIC-II mission was referred to as the FORMOSAT-3 follow-on mission for simplicity.

Eight satellites are at high inclination angle of 72° at 8 orbital planes (see the pink lines in Figure 5) and separated by 22.5° when complete constellation deployment. Four more satellites are at low inclination angle of 24° at 4 orbital planes (see the blue lines in Figure 5) and separated by 45° when complete constellation deployment. The satellites at high inclination angle will be launched in one or two clusters and be placed to one parking orbit. The mission operations team will then perform the spacecraft orbit raising so that their orbital plane can be separated through the differential precession rate with the differential orbit altitude. The satellite at low inclination angle will go through the similar launch and constellation deployment process. The overall deployment period will be about nineteen months for eight high inclination satellites and about seven months for four low inclination satellites, respectively.
The system will start to collect data collection once the satellite has completed the in-orbit checkout at parking orbits. Herein, we do not exclude the possibility to send the spacecraft by co-share piggyback conjunction with other mission satellites. The advantage of the proposed constellation design is that the collected data will be homogeneously distributed world-wide evenly within a 3-hour period.



| No. of SC | RO PL Capability | Occ. No. Per Day | Eq. Area Covered (km x km) |
|---|---|---|---|
| 6 | GPS | 2500 | 450 x 450 |
| 12 | GPS | 5000 | 320 x 320 |
| 12 | GALILEO | 4900 | 320 x 320 |
| 12 | GLONASS | 3800 | 370 x 370 |
| 12 | GPS + GALILEO + GLONASS | 13700 | 190 x 190 |

Fig. 5. The FORMOSAT-3 Follow-on constellation

Figure 6 shows the proposed follow-on mission system architecture that requires three launches. The mission includes space segment, launch segment, ground segment, and science segment.

### 7.3 Space segment

The primary payload of the follow-on satellite will be equipped with next-generation GNSS RO receiver to collect more soundings per receiver with GPS, Galileo and GLONASS tracking capability, which include 29 operational USA GPS satellites, several Russia's GLONASS (planning to have 18 satellites), and European GALILEO system (plan to have 30 GNSS satellites by 2013). The new GNSS RO receiver will be able to receive the USA GPS

L1/L2/L5 signals, also to receive the GALILEO E1/E5/E6 signals, and to receive GLONASS's L1/L2/L5 signals as well (Yen & Fong, 2009; Fong et al., 2009b).

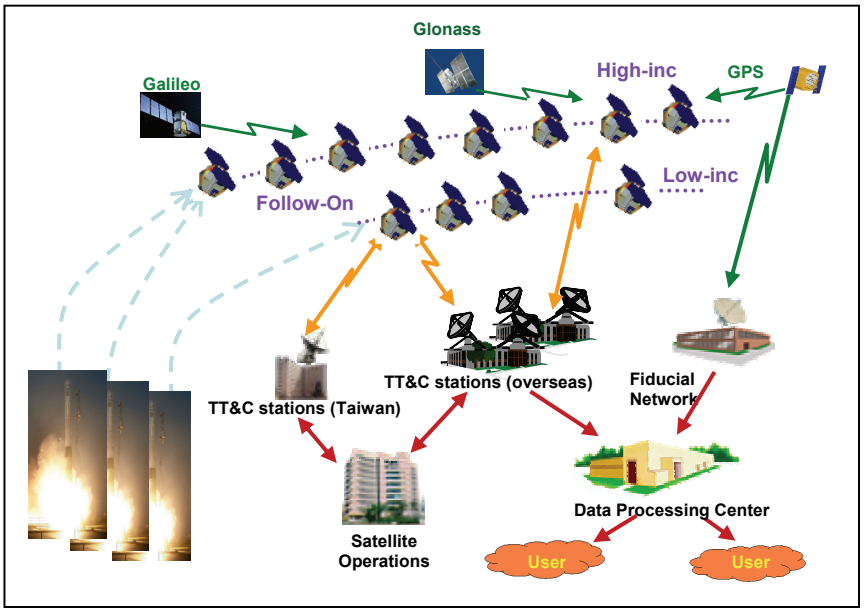



Fig. 6. The FORMOSAT-3 Follow-on mission system architecture

Figure 7 shows the proposed FORMOSAT-3 Follow-on mission spacecraft avionic diagram. The benefit and improvement for the follow-on spacecraft will improve payload performance, better attitude performance, simplified operation, simplified orbit transfer, increased data storage, and modular design for additional compatible science payloads. The follow-on spacecraft bus intended design vs. FORMOSAT-3 design is shown in Table 4.

### 7.4 Ground segment

The NSPO SOCC will be responsible for the satellites flight operations during the mission. The SOCC will take charge of up linking commands, monitoring the state of health of satellite, analyzing the trending data, scheduling the passes of the TT&C stations, planning the constellation deployment, performing orbit thrusting and maintenance, conducting anomaly resolution, etc. SOCC will use Taiwan TT&C stations as well as the RTS for satellites telemetry and commanding (Yen & Fong, 2009; Fong et al., 2009b).

More RTS stations or more antennas from the same RTS Station are desired in order to have data transmission at every orbit to reduce data latency. The current FORMOSAT-3 RTS stations at Fairbanks and Tromso are at good geo locations to cover every dump for the 72° inclination constellations, but need to add more antennas to fulfill 12 satellites data dump including 4 satellites at lower inclination angle. The EUMESAT ground stations, as example, have been considered as part of the receiving stations to reduce the data latency for the RO data retrieving. The SOCC uses the real time telemetry and the back orbit telemetry to monitor, control. The downlinked science data is transmitted from the RTS via NOAA to the two data receiving and processing centers.

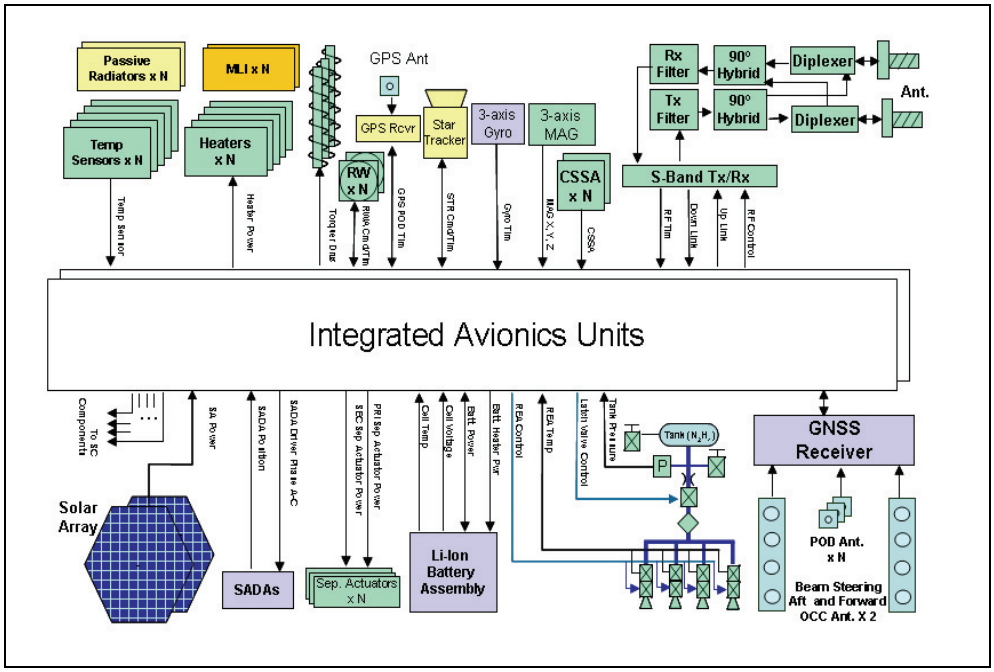



Fig. 7. Proposed spacecraft avionic diagram of the FORMOSAT-3 Follow-on mission.

| Function | Follow-on Design | FORMOSAT-3 Design | Benefit |
|---|---|---|---|
| Weight | ~60 kg | 61 kg (w/ Propellant) | Stacked or Single Launch<br>Piggy-Back Launch |
| Attitude Control Performance | 3-axis linear control<br>Roll/Yaw:+/-1° (3σ)<br>Pitch: +/- 1° (3σ)<br>3-Axis Gyro, 3-axis MAG, CSSAs, RWA x 3, Torque x 3,Star Tracker x 1 | 3-axis nonlinear control<br>Roll/Yaw: +/-5° (1σ)<br>Pitch: +/- 2° (1σ)<br>Earth Sensor x 2, CSSA x 8, RWA x 1, Torque x 3, GPS Bus Receiver PL x 1 | Improved PL Performance<br>Better Attitude Performance<br>Simplified Operation<br>Simplified Orbit Transfer |
| Science Data Storage | >1.5 G | 128 M | Increased Data Storage<br>Simplified Operations |
| Avionics Architecture | Centralized Architecture<br>Radiation - Hardness | Distributed Architecture<br>(Multiple Avionics Boxes) | Simplified Integration<br>Harnessing & Mass Reduced |
| Electrical Power | Lithium Ion Battery<br>Voltage Based Algorithm | Ni-H2 Battery<br>dM/dC Charging Algorithm | Reduced Mass & Volume<br>Simplified Operations |
| Structure | Aluminum (Al) | Metal Matrix (AlBeMet) | Cost Reduced |
| Payload Interface | Main PL: GNSS RO Rcvr<br>Science PL (Optional) | Primary PL: GOX<br>Secondary PL: TIP, TBB | Modular Design<br>Cost Reduced |

Table 4. Proposed FORMOSAT-3 Follow-on spacecraft bus design vs. FORMOSAT-3 design.

### 7.5 Science segment

The data receiving and processing centers will be CDAAC and TACC. All collected science data is processed by CDAAC and then transferred to TACC and other facilities for operations, science, and data archival. The processed atmospheric profiles are distributed in near real time to international weather centers from CDAAC through NOAA/NESDIS. These data are currently provided to weather centers within 90 minutes after satellite on-orbit science data collection in order to be ingested by the operational weather forecast models. The data from fiducial network is part of inputs to data processing center.

## 8. Conclusion

The GNSS RO mission represents a revolution in atmospheric sounding from space, with precise, accurate, and all-weather global observations useful for weather, climate, and space weather research and operations. The FORMOSAT-3 constellation has been unanimously regarded by the major user community as "the world's most accurate, precise, and stable atmospheric thermometer in space." The success of the FORMOSAT-3 mission has initiated a new era for the operational GPS RO sounding applications in the world.

With the success of FORMOSAT-3, there is a strong interest from both research and operational communities to maintain and enhance the existing FORMOSAT-3 mission. Observing system simulation experiments (OSSE) have been conducted to evaluate the impact of the GNSS RO follow-on mission on the prediction of typhoons in the vicinity of Taiwan. OSSE experiments based on Typhoon Shanshan (2006) indicated that the proposed follow-on configuration provided a significant improvement over the existing FORMOSAT-3 for typhoon prediction in terms of track, intensity and rainfall prediction. These preliminary results strongly support the use of the proposed follow-on design for the FORMOSAT-3 Follow-on mission.

## 9. References

Anthes, R. A.; Rocken, C. & Kuo, Y. H. (2000). Application of COSMIC to Meteorology and Climate. *Terrestrial, Atmospheric and Oceanic Sciences*, Vol. 11, No.1, Mar. 2000, pp. 115-156.

Anthes, R. A.; Bernhardt, P. A.; Chen, Y.; Cucurull, L.; Dymond, K. F.; Ector, D.; Healy, S. B.; Ho, S.-P.; Hunt, D. C.; Kuo, Y.-H.; Liu, H.; Manning, K.; McCormick, C.; Meehan, T. K.; Randel, W. J.; Rocken, C.; Schreiner, W. S.; Sokolovskiy, S. V.; Syndergaard, S. ; Thompson, D. C.; Trenberth, K. E.; Wee, T. K.; Yen, N. L.; & Zeng, Z. (2008). The COSMIC/FORMOSAT-3 Mission: Early Results. *Bulletin of the American Meteorological Society*, Vol. 89, No.3, Mar. 2008, pp. 313-333. doi:10.1175/BAMS-89-3-313.

Bonnedal, M. (2009). RUAG GNSS Receivers and Antennas, *Proceedings of Global Navigation Satellite System Radio Occultation Workshop*, Pasadena, California, 7-9 April, 2009, JPL, Pasadena.

Chu, C.-H.; Yen, N.; Hsiao, C.-C.; Fong, C.-J.; Yang, S.-K.; Liu, T.-Y.; Lin, Y.-C.; & Miau, J.-J. (2008). Earth observations with orbiting thermometers – prospective FORMOSAT-3/COSMIC follow-on mission, *Proceedings of Small Satellite Conference 2008*, Logan, Utah., 11-14 Aug. 2008, USU, Logan.

Fjeldbo, G. & Eshleman, V. R. (1965). The bistatic radar-occultation method for the study of planetary atmospheres. *J. Geophys. Res.*, Vol. 70, 1965, pp. 3217-3225.

Fong, C.-J.; Shiau, A.; Lin, T.; Kuo, T.-C.; Chu, C.-H.; Yang, S.-K.; Yen, N.; Chen, S. S.; Huang, C.-Y.; Kuo, Y.-H.; Liou, Y.-A.; & Chi, S. (2008a). Constellation deployment for FORMOSAT-3/COSMIC mission. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No.11, Nov. 2008, pp. 3367-3379. doi:10.1109/TGRS.2008.2005202.

Fong, C.-J.; Yang, S.-K.; Chu, C.-H.; Huang, C.-Y.; Yeh, J.-J.; Lin, C.-T.; Kuo, T.-C.; Liu, T.-Y.; Yen, N.; Chen, S. S.; Kuo, Y.-H.; Liou, Y.-A.; & Chi, S. (2008b). FORMOSAT-3/COSMIC constellation spacecraft system performance: After One Year in Orbit. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No.11, Nov. 2008, pp. 3380-3394. doi:10.1109/TGRS.2008.2005203.

Fong, C.-J.; Huang, C.-Y.; Chu, V.; Yen, N.; Kuo, Y.-H.; Liou, Y.-A.; & Chi, S. (2008c). Mission results from FORMOSAT-3/COSMIC constellation system. *AIAA Journal of Spacecraft and Rockets*, Vol. 45, No. 6, Nov.-Dec. 2008, pp. 1293-1302. doi:10.2514/1.34427.

Fong, C.-J.; Yen, N. L.; Chu, C.-H.; Yang, S.-K.; Shiau, W.-T.; Huang, C.-Y.; Chi, S.; Chen, S.-S.; Liou, Y.-A.; & Kuo, Y.-H. (2009a). FORMOSAT-3/COSMIC spacecraft constellation system, mission results, and prospect for follow-on mission. *Terrestrial, Atmospheric and Oceanic Sciences*, Vol. 20, No. 1, Jan. 2009. doi:10.3319/TAO.2008.01.03.01(F3C).

Fong, C.-J.; Yen, N. L.; Chu, C.-H.; Hsiao, C.-C.; Liou, Y.-A.; & Chi, S. (2009b). Space-based Global Weather Monitoring System – FORMOSAT-3/COSMIC Constellation and its Follow-On Mission. *AIAA Journal of Spacecraft and Rockets*, Vol. 46, No. 4, July-August 2009, pp. 883-891. doi:10.2514/1.41203.

Franklin, G. & Giesinger, B. (2009). TriG - Next Generation GNSS POD Occultation Receiver, *Proceedings of Global Navigation Satellite System Radio Occultation Worksho*p, Pasadena, California, 7-9 April, 2009, JPL, Pasadena.

Fuggetta, G.; Marradi, L.; Zin, A.; Landenna, S.; Gianeli, G.; DeCosmo, V. (2009). ROSA Instrument and Antenna, *Proceedings of Global Navigation Satellite System Radio Occultation Workshop*, Pasadena, California, 7-9 April, 2009, JPL, Pasadena.

GCOS (Global Climate Observing System). (2003). The Second Report on the Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC. GCOS-82 (WMO/TD 1143), World Meteorological Organization (WMO), Geneva, Switzerland.

GCOS. (2004). Implementation Plan for the Global Observing System for Climate in Support of the UNFCC. GCOS-92 (WMO/TD 1219), WMO, Geneva, Switzerland.

GCOS. (2006a). Systematic Observation Requirements for Satellite-based Products for Climate—Supplemental Details to the GCOS Implementation Plan. GCOS-107 (WMO/TD 1338), pp. 15-17, WMO, Geneva, Switzerland.

GCOS. (2006b) CEOS response to the GCOS Implementation Plan September 2006, Doc. 17 in Satellite Observations of the Climate System. GCOS-109 (WMO/TD 1363), WMO, Geneva, Switzerland.

Hajj, G. A.; Lee, L. C.; Pi, X.; Romans, L. J.; Schreiner, W. S.; Straus, P. R.; & Wang, C. (2000). COSMIC GPS ionospheric sensing and space weather. *Terr. Atmos. Oceanic Sci.*, Vol. 11, 2000, pp. 235–272.

Kuo, Y.-H.; Sokolovskiy, S.; Anthes, R.; & Vandenberghe, V. (2000). Assimilation of GPS radio occultation data for numerical weather prediction. *Terrestrial, Atmospheric and Oceanic Sciences*, Vol. 11, No. 1, Mar. 2000, pp. 157–186.

Kuo, Y.-H.; Wee, T.-K.; Sokolovskiy, S.; Rocken, C.; Schreiner, W.; Hunt, D.; & Anthes, R. A. (2004). Inversion and Error Estimation of GPS Radio Occultation Data. *Journal of the Meteorological Society of Japan*, Vol. 82, No. 1B, 2004, pp. 507-531. doi:10.2151/jmsj.2004.507

Kuo, Y.-H.; Liu, H.; Guo, Y.-R.; Terng, C.-T.; & Lin, Y.-T. (2008a). Impact of FORMOSAT-3/COSMIC data on typhoon and Mei-yu prediction, In: *Recent Progress in Atmopsheric Sciences: Applications to the Asia-Pacific Region*. Liou, K.-N., Chou, M.-D. (eds), pp. 458-483, World Scientific Publishing, Singapore.

Kuo, Y.-H.; Liu, H.; Ma, Z. & Guo, Y.-R. (2008b). The impact of FORMOSAT-3/COSMIC GPS radio occultation. *Proceedings of 4th Asian Space Conference and 2008 FORMOSAT-3/COSMIC International Workshop*, Taipei, Taiwan, 1-3 Oct. 2008, NSPO, Hsinchu, Taiwan.

Kliore, A. J.; Cain, D. L.; Levy, G. S.; Eschleman, V. R.; Fjeldbo, G.; & Drake, F. D. (1965). Occultation experiment: results of the first direct measurement of Mars' atmosphere and ionosphere, *Science*, Vol. 149, No. 3689, Sep. 1965, pp. 1243-1248.

Kursinski, E. R.; Hajj, G. A.; Bertiger, W. I.; Leroy, S. S.; Meehan, T. K.; Romans, L. J.; Schofield, J. T.; McCleese, D. J.; Melbourne, W. G.; Thornton, C. L.; Yunck, T. P.; Eyre, J. R.; & Nagatani, R. N. (1996). Initial results of radio occultation observations of Earth's atmosphere using the Global Positioning System. *Science*, Vol. 271, No. 5252, Feb. 1996, pp. 1107-1110.

Kursinski, E. R.; Hajj, G. A.; Leroy, S. S.; & Herman, B. (2000). The GPS Occultation Technique, *Terrestrial, Atmospheric and Oceanic Sciences*, Vol. 11, No.1, pp. 53-114.

Lee, L.-C., Kursinski, R. & Rocken, C., Ed., (2001). *Applications of Constellation Observing System for Meteorology, Ionosphere & Climate: Observing System for Meteorology Ionosphere and Climate*, Springer, ISBN 9624301352, Hong Kong.

Liu, A. S. (1978). On the determination and investigation of the terrestrial ionospheric refractive indices using GEOS-3/ATS-6 satellite-to-satellite tracking data. *NASA-CR-156848,* Nov. 1978, Jet Propulsion Laboratory, Pasadena, CA.

Liou, Y.-A. & Huang, C.-Y. (2000). GPS observation of PW during the passage of a typhoon. *Earth, Planets, and Space*, Vol. 52, No. 10, pp. 709-712.

Liou, Y.-A.; Teng,Y.-T.; Hove; T. V., & Liljegren, J. (2001). Comparison of precipitable water observations in the near tropics by GPS, microwave radiometer, and radiosondes. *J. Appl. Meteor.*, Vol. 40, No. 1, pp. 5-15.

Liou, Y.-A.; Pavelyev, A. G.; Huang, C.-Y.; Igarashi, K.; & Hocke, K. (2002). Simultaneous observation of the vertical gradients of refractivity in the atmosphere and electron density in the lower ionosphere by radio occultation amplitude method. *Geophysical Research Letters*, Vol. 29, No. 19, pp. 43-1-43-4, doi:10.1029/2002GL015155.

Liou, Y.-A. & Pavelyev, A. G. (2006a). Simultaneous observations of radio wave phase and intensity variations for locating the plasma layers in the ionosphere. *Geophys. Res. Lett.*, Vol. 33, No. L23102, doi:10.1029/2006GL027112.

Liou, Y.-A.; Pavelyev, A. G.; Wicker, J.; Liu, S. F.; Pavelyev, A. A.; Schmidt, T.; & Igarashi, K. (2006b). Application of GPS radio occultation method for observations of the internal waves in the atmosphere. *J. Geophys. Res.*, Vol. 111, No. D06104, doi: 10.1029/2005JD005823.

Liou, Y.-A.; Pavelyev, A. G.; Liu, S. F.; Pavelyev, A. A.; Yen, N.; Huang, C. Y.; & Fong, C.-J. (2007). FORMOSAT-3 GPS radio occultation mission: preliminary results. *IEEE Trans. Geosci. Remote Sensing*, Vol. 45, No. 10, Nov. 2007, pp. 3813-3824. doi:10.1109/TGRS.2007.903365.

Melbourne, W. G. (2005). *Radio Occultations Using Earth Satellites*: *A Wave Theory Treatment*, pp. 1-66, John Wiley & Sons, Inc., ISBN 0-471-71222-1, New Jersey.

Rius, A.; Ruffini, G. & Romeo, A. (1998). Analysis of ionospheric electron-density distribution from GPS/MET occultations. *IEEE Trans. Geosci. Remote Sens.*, Vol. 36, No. 2, Mar. 1998, pp. 383-394.

Rocken, C.; Kuo, Y.-H.; Schreiner, W. S.; Hunt, D.; Sokolovskiy, S., McCormick, C. (2000). COSMIC system description. *Terrestrial, Atmospheric and Oceanic Sciences*, Vol. 11, No. 1, Mar. 2000, pp. 21-54.

SSB (Space Studies Board). (2007). Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, Committee on earth science and applications from space: a community assessment and strategy for the future. Space Studies Board, National Research Council, National Academies Press. ISBN: 0-309066714-3, 456 pages.

Ware, R.; Rocken, C.; Solheim, F.; Exner, M.; Schreiner, W.; Anthes, R.; Feng, D.; Herman, B. ; Gorbunov, M.; Sokolovskiy, S.; Hardy, K.; Kuo, Y.-H.; Zou, X.; Trenberth, K.; Meehan, T.; Melbourne, W. G.; & Businger, S. (1996). GPS sounding of the atmosphere from low earth orbit: Preliminary results. *Bulletin of the American Meteorological Society*, Vol. 77, No. 1, Jan. 1996, pp. 19-40.

WMO (World Meteorological Organization) (2007). Workshop on the "Redesign and Optimization of the Space Based Global Observing System, Outcome of the OPT-2 Workshop," ETSAT/SUP3/Doc. 5(1), World Meteorological Organization, Geneva, Switzerland, 21-22 Jun. 2007.

Wu B.-H.; Fong, C.-J.; Huang, C.-Y.; Liou, Y.-A.; Yen, N.; & Chen, P. (2006). FORMOSAT-3/COSMIC mission to global earth weather monitoring, operation, and TACC/CDAAC post-processing, *Proceedings of 86th AMS Annual Meeting,-14th conf. satellite meteorology and oceanography*, Atlanta, GA, 29-2 Feb. 2006.

Wu, S. C. & Melbourne, W. G. (1993). An optimal GPS data processing technique for precise positioning. *IEEE Trans. Geosci. Remote Sens.*, Vol. 31, No. 1, Jan. 1993, pp. 146-152.

Yakovlev, O.; Matyugov, I. & Vilkov, I. A. (1996). Radio-wave phase and frequency fluctuations as observed in radio occultation experiments on the satellite-to-satellite link, *Journal of Communications Technology and Electronics*, Vol. 41, No. 11, pp. 993-998, Nov. 1996.

Yen, N. L. & Fong, C.-J., ed. (2009). FORMOSAT-3 Evaluation Report and Follow-on Mission Plan, NSPO-RPT-0047_0000, 10 May 2009, National Space Organization (NSPO), Hsinchu, Taiwan.

Yunck, T. P. & Hajj, G. A. (2003). Global navigation satellite sounding of the atmosphere and GNSS altimetry: prospects for geosciences, *Proceedings of IUGG General Assembly*, Jul. 2003, Sapporo, Japan. ISSN 0065-8448.

Yunck, T. P.; Liu, C. H. & Ware, R. (2000). A History of GPS Sounding. *Terrestrial, Atmospheric and Oceanic Sciences*, Vol. 11, No.1, Mar. 2000, pp. 1-20.

Yunck, T. P.; Wu, S. C.; Wu, J. T.; & Thornton, C. L. (1990). Precise tracking of remote sensing satellites with the global positioning system. *IEEE Trans. Geosci. Remote Sens.*, Vol. 28, No.1, Jan. 1990, pp. 108-116.

# 14

# Integrated Vehicle Health Management for Solid Rocket Motors

D.G. Luchinsky[1,2], V.V. Osipov[1,2], V.N. Smelyanskiy[1], I. Kulikov[1], A. Patterson-Hein[1], B. Hayashida[3], M. Watson[3], D. Shook[4], M. Johnson[4], S. Hyde[4] and J. Shipley[4]

[1]*NASA Ames Research Center, MS 269-3, Moffett Field, CA, 94035,*
[2]*Mission Critical Technologies Inc., 2041 Rosecrans Ave., Suite 225 El Segundo, CA 90245,*
[3]*ISHM and Sensors Branch, NASA Marshall Space Flight Center, Huntsville, Alabama 35812,*
[4]*ATK Thiokol Launch Systems R&D Labs, Large Salt Lake City Area, Utah, USA*

## 1. Introduction

Solid rocket motors (SRMs) are an integral part of human space flight providing a reliable means of breaking away from the Earth's gravitational pull. The development and deployment of an integrated system health management (ISHM) approach for the SRMs is therefore a prerequisite for the safe exploration of space with the next-generation Crew and Heavy-Lift Launch Vehicles. This unique innovative technological effort is an essential part of the novel safety strategy adopted by NASA.

At the core of an on-board ISHM approach for SRMs are the real-time failure detection and prognostics (FD&P) technique. Several facts render the SRMs unique for the purposes of FD&P: (i) internal hydrodynamics of SRMs is highly nonlinear, (ii) there is a number of failure modes that may lead to abrupt changes of SRMs parameters, (iii) the number and type of sensors available on-board are severely limited for detection of many of the main SRM failure modes; (iv) recovery from many of the failure modes is impossible, with the only available resource being a limited thrust vector control authority (TVC); (iii) the safe time window between the detectable onset of a fault and a possible catastrophic failure is very short (typically a few seconds). The overarching goal of SRM FD&P is to extract an information from available data with precise timing and a highest reliability with no "misses" and no "false alarms". In order to achieve this goal in the face of sparse data and short event horizons, we are developing: (i) effective models of nominal and off-nominal SRM operation, learned from high-fidelity simulations and firing tests and (ii) a Bayesian sensor-fusion framework for estimating and tracking the state of a nonlinear stochastic dynamical system. We expect that the combination of these two capabilities will enable in-flight (real time) FD&P.

Indeed, dynamical models of internal SRMs ballistics and many SRMs fault modes are well studied, see e.g. (Culick, 1996; Salita, 1989; Sorkin, 1967) and references therein. Examples of faults, for which quite accurate dynamical models can be introduced, include: (1) combustion instability; (ii) case breach fault, i.e. local burning-through of the rocket case; (iii) propellant cracking; (iv) overpressure and breakage of the case induced by nozzle blocking or bore choking. The combustion instabilities were studied in detail in the classical papers of (Culick & Yang, 1992; Culick, 1996) and (Flandro et al, 2004). Bore choking phenomenon due to radial deformation of the propellant grain near booster joint segments was studied numerically in (Dick et al., 2005; Isaac & Iverson, 2003; Wang et al., 2005) and observed in primary construction of the Titan IV (see the report, Wilson at al., 1992).

The FD&P system can be developed using the fact that many fault modes of the SRMs have unique dynamical features in the time-traces of gas pressure, accelerometer data, and dynamics of nozzle gimbal angle. Indeed, analysis shows that many fault modes leading to SRMs failures, including combustion instabilities (Culick,1974; Culick & Yang, 1992; Culick,1996; Flandro et al, 2004), bore choking (Dick et al., 2005; Isaac & Iverson, 2003; Wang et al., 2005), propellant cracking, nozzle blocking, and case breach (Rogers, 1986), have unique dynamical features in the time-traces of pressure and thrust. Ideally, the corresponding expert knowledge could be incorporated into on-board FD&P within a general Bayesian inferential framework allowing for faster and more reliable identification of the off-nominal regimes of SRMs operation in real time. In practice, however, the development of such an inferential framework is a highly nontrivial task since the internal ballistics of the SRMs results from interplay of a number of complex nonlinear dynamical phenomena in the propellant, insulator, and metal surfaces, and gas flow in the combustion chamber and the nozzle. On-board FD&P, on the other hand, can only incorporate low-dimensional models (LDMs) of the internal ballistics of SRMs. The derivation of the corresponding LDMs and their verification and validation using high-fidelity simulations and firing tests become an essential part of the development of the FD&P system.

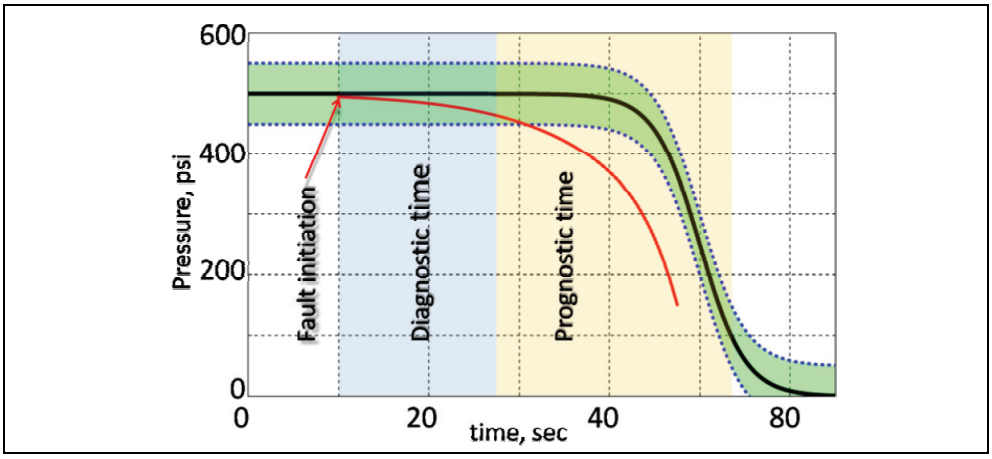



Fig. 1. Typical time-trace of pressure in the nominal regime is shown by the black line with pressure safety margins indicated by the green shading region. Fault-induced pressure time-trace in off-nominal regime is shown by the red line. Blue shading indicates diagnostic window and yellow shading indicates prognostic window.

At present the FD&P system in SRMs involves continuous monitoring of the time-traces of such variables as e.g. pressure, thrust, and altitude and setting up conservative margins on the deviation of these variables from their nominal values (see schematics in Fig. 1). However, in the absence of the on-board FD&P analysis of the SRM performance the probability of "misses" and "false alarms" is relatively high and reliability of the IVHM is reduced (see e.g. Rogers, 1986; Oberg, 2005). The goal of the on-board FD&P will be to detect the initiation time of the fault and provide its continuous diagnostic and prognostic while the performance variables are still within the safety margins to support the decision and to reduce the probability of "misses" and "false alarms".

In this chapter we report the progress in the development of such FD&P system. The main focus of our research was on the development of the: (i) model of internal ballistics of large segmented SRMs in the nominal regime and in the presence of number of fault modes including first of all case breach fault; (ii) model of the case breach fault; (iii) algorithms of the diagnostic and prognostic of the case breach fault within a general inferential Bayesian framework; and (iv) verification and validation of these models and algorithms using high-fidelity simulations and ground firing tests.

The chapter is organized as follows. In the next section we describe the low-dimensional performance model of internal ballistics of the SRMs in the presence of faults. In the Sec. III we modify this model for a subscale solid motor, analyze the axial distributions and validate the results of this model based on high-fidelity FLUENT simulations and analysis of the results of a ground firing test of the sub-scale motor faults. Developed Bayesian inferential framework for the internal SRM ballistics and FD&P algorithms is presented in the Sec. IV. FD&P for large segmented SRMs is analyzed in the Sec. V. Finally, in the Conclusions we review the results and discuss a possibility of extending proposed approach to an analysis of different faults.

## 2. Internal ballistics of the SRMs

The internal ballistics of the SRMs in the presence of the fault can be conveniently described by the following set of stochastic partial differential equations representing conservation laws for mass, momentum, and energy of the gas (Sorkin, 2005; Culick & Yang, 1992; Salita, 1989 & 2001)

$$\partial_t\left(UA_p\right)+\partial_x\left(f(U)A_p\right)=S, \tag{1}$$

where conservative variables of the gas dynamics and function $f(U)$ are given by the following equations

$$U=\begin{bmatrix}\rho\\ \rho u\\ \rho e_T\end{bmatrix},\quad f(U)=\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho u e_T+up\end{bmatrix}, \tag{2}$$

$e_T=c_VT+u^2/2$, $h_T=c_pT+u^2/2$, are the total energy and total enthalpy of the gas flow, $H=c_pT_0$ is the combustion heat of solid propellant and the source terms that include fault terms at a given location $x_0$ have the form

$$S = \begin{bmatrix} \rho_p \dot{R} l(x) - \rho_{t,h} u_{t,h} A_h \delta(x - x_0) + \xi_1(t) \\ p \partial_x A_p - \lambda \rho u^2 l(x) + \xi_2(t) \\ H \rho_p \dot{R} l(x) - h_{t,h} \rho_{t,h} u_{t,h} A_h \delta(x - x_0) + \xi_3(t) \end{bmatrix}. \quad (3)$$

This model extends the previous work (Salita, 1989 & 2001) in a number of important directions. To model various uncontrollable sources of noise (such as cracks and case vibrations) that may become essential in off-nominal conditions and may screen the variation of the system parameters a random component in the propellant density $\rho_p \to \rho_p[1+\sqrt{\sigma}\cdot\xi(t)]$ is introduced. Various faults can be modeled within the set of Eqs. (1)-(3) (including nozzle failure, propellant cracking, bore choking, and case breach) by choosing the time scale and direction of the geometrical alternations of the grain and case and the corresponding form of the sourse/sink terms. In particular, for the case breach fault two additional terms in the 1st and 3rd equations in Eqs. (3) correspond to the mass and energy flow from the combustion chamber through the hole in the rocket case with cross-section area $A_h(t)$. We now extend this mode by coupling the gas dynamics in the combustion chamber to the gas flow in the hole. The corresponding set of PDEs

$$\begin{cases} \partial_t (A_h \rho_h) = -\partial_x (A_h \rho_h u_h), \\ \partial_t (A_h \rho_h u_h) = -\partial_x (A_h \rho_h u_h^2) - A_h \partial_x p_h - f_{fr} l_h, \\ \partial_t (A_h \rho_h e_{t,h}) = -\partial_x [A_h \rho_h u_h h_{t,h}] - Q_h l_h, \end{cases} \quad (4)$$

resembles Eqs. (1). The important difference, however, is that we neglect mass flow from the walls of the hole. Instead Eqs. (4) include the term that describes the heat flow from the gas to the hole walls. The boundary conditions for this set of equations assume ambient conditions at the hole outlet and the continuity equation for the gas flow in the hole coupled to the sonic condition at the hole throat. The value of $Q_h$ is presented in Eq. (14). The dynamics of the gas flow in the nozzle is described by a set of equations similar to (4) and can be obtained from this set by substituting subscript "n" for subscript "*h*".

The model (1)-(4) allow us to include possible burning rate variations and also various uncontrolled sources of noise, such as grain cracks and case vibrations to simulate more realistic time-series data representing off-nominal SRM operation. Due to the high temperature T of combustion products in the combustion chamber, the hot mixed gas can be considered as a combination of ideal gases. As we are interested in average gas characteristics (head pressure and temperature) we will characterize the combustion products by averaged parameters using the state equation for an ideal gas:

$$\frac{p}{\rho} = (c_P - c_V)T = \frac{p_0}{\rho_0}\left(\frac{T}{T_0}\right) = \frac{c_0^2}{\gamma}\left(\frac{T}{T_0}\right) \quad (5)$$

### 2.1 Regression of propellant surface

We take into account the propellant erosion in a large segmented rocket assuming that the erosive burning rate can be presented in the form

$$\dot{R} = r_b = ap^n + \dot{r}_{er}. \quad (6)$$

The erosive burning is taken into account in the Vilyunov's approximation

$$\dot{r}_{er} = C\left(I - I_{cr}\right) \tag{7}$$

for $I > I_{cr}$ and $0$ otherwise, where $C$ and $I_{cr}$ are constants and $I$=const$(\rho u / r_b \rho_p)$Re$^{-1/8}$, where Re is the Reynolds number.

### 2.2 Model of the propellant geometry

To model the actual propellant geometry along the rocket axis the combustion chamber is divided into $N$ segments as schematically shown in the Fig. 2. For each ballistic element the port area $A_p(x_i)$ and perimeter $l(x_i)$ averaged over the segment length $dx_i$ are provided in the form of the design curves

$$A_p(x_i) = f_{Ai}(R(x_i)), \quad l(x_i) = f_{li}(R(x_i)) \tag{8}$$

(see Fig. 2). Note that the burning area and the port volume for each segment are given by the following relations

$$dV(x_i) = A_b(x_i)dx_i, \quad dA_b(x_i) = l(x_i)dx_i, \tag{9}$$

and, therefore, are uniquely determined by the burning rate $r_{bi}$ for each ballistic element.
For numerical integration each segment was divided into a finite number of ballistic elements. The design curves were provided for each ballistic segment.

### 2.3 Model of the nozzle ablation

To model nozzle ablation we use Bartz' approximation (Bartz, 1965; Hill and Peterson,1992; Handbook, 1973) for the model of the nozzle ablation (Osipov et al., March 2007, and July 2007; Luchinsky et al., 2007) in the form:

$$\begin{aligned} \dot{R}_N &= v_{t0}\left(p / p_{\max}\right)^{1-\beta}\left(R_N / R_{t,in}\right)^{-\beta}\left[\left(T - T_{abl}\right)/\left(T_t - T_{abl}\right)\right], \\ v_{t0} &= \varepsilon C_p\left(\frac{\gamma p_{\max}}{\Gamma c_0}\right)^{1-\beta}\left(\frac{2R_{t,in}}{\mu}\right)^{-\beta}\frac{\left(T_t - T_{abl}\right)}{\left[c_{ins}\left(T_{abl} - T_0\right) + q_{ins}\right]}, \end{aligned} \tag{10}$$

where $\beta \approx 0.2$ and $\varepsilon \approx 0.023$. In a particular case of the ablation of the nozzle throat and nozzle exit this approximation is reduced to

$$\dot{R}_t = v_{m,t}\left(\frac{p_0}{p_{\max}}\right)^{1-\beta}\left(\frac{R_t}{R_{t,in}}\right)^{-\beta}, \quad A_t(t) = \pi R_t^2(t), \tag{11}$$

$$\dot{R}_{ex} = v_{m,ex}\left(\frac{p_0 A_t}{p_{\max} A_{ex}}\right)^{1-\beta}\left(\frac{R_{ex}}{R_{ex,in}}\right)^{-\beta}\frac{\left(T_{ex} - T_{abl}\right)}{\left(T_t - T_{abl}\right)}, \tag{12}$$

where $R_{t,in}$=$R_t(0)$, $R_{ex,in}$=$R_{ex}(0)$ and $v_{m,t}$ and $v_{m,ex}$ are experimentally determined constants. In practice, to fit experimental or numerical results on the nozzle ablation it suffice to put $\beta$ = 0.2 and to obtain values of $v_{m,t}$ and $v_{m,ex}$ by regression.

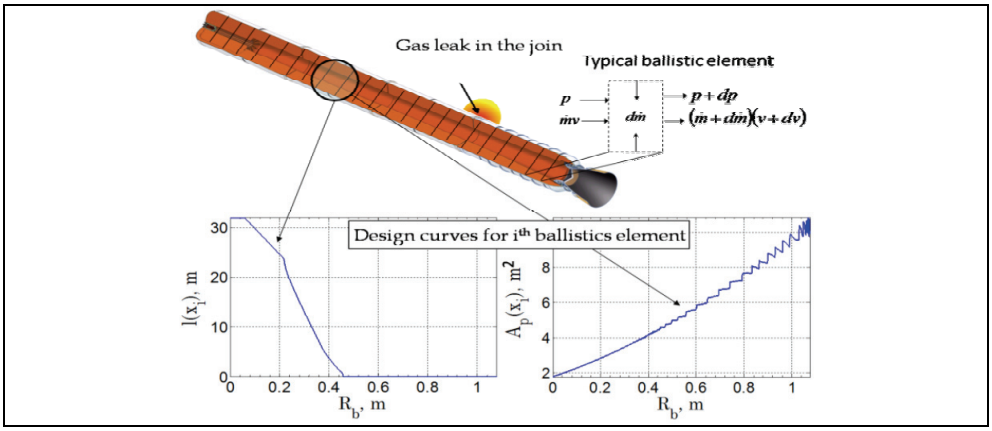



Fig. 2. Sketch of a cross-section of an idealized geometry of the multi-segment RSRMV rocket and an example of the design curves (8) for the head section.

### 2.4 Model of the burning-though of a hole

To complete the model of the case breach fault for the segmented SRMs the system of equations (1)—(12) above has to be extended by including equations of the hole growth model (Osipov et al., 2007, March and 2007, July; Luchinsky et al., 2007)

$$\dot{R}_h = v_h(p_{t,h}, T_{h,t}) = \frac{Q_c + Q_R + Q_b}{\left[q_{met} + C_{met}\left(T_{mel} - T_{m0}\right)\right]\rho_{met}}, \tag{13}$$

$$\begin{aligned} Q_R &= \sigma\left[1 - \exp\left(-\lambda p_{t,h}\right)\right]\left(T_t^4 - T_{met}^4\right), \\ Q_c &= 0.023 C_p \left(\frac{\gamma p_{t,h}}{\Gamma c_0}\right)^{0.8} \left(\frac{2R_h}{\mu}\right)^{-0.2} \left(T_t - T_{met}\right), \\ Q_b &= v_{fb}\left[q_{met} + C_{met}\left(T_{mel} - T_{m0}\right)\right]\rho_{met}. \end{aligned} \tag{14}$$

Here $Q_h = Q_c + Q_R + Q_b$.

## 3. A subscale motor

Motivated by the results of the ground firing test let us consider an application of the model (1)-(14) to an analysis of the case breach fault in a subscale motor. Note that a subscale motor can be consider as model (1)-(14) consisting of one ballistic element. In this case the velocity of the flow is small and one can neglect the effects of erosive burning, surface friction, and the variation of the port area along the motor axis.

### 3.1 SRM internal ballistics in the "filling volume" approximation

To derive the LDM of the case breach fault we integrate equations (1) along the rocket axis and obtain the following set of ordinary differential equations for the stagnation values of the gas parameters and the thickness of the burned propellant layer

$$\begin{cases} \partial_t\left(\rho_0 V\right) = -\left(\rho u A\right)\big|_L + \rho_p r_b(p_0) A_b + A_1\xi_1(t), \\ \partial_t\left(\rho_0 e_0 V\right) = -\left(\rho u A h_t\right)\big|_L + H\rho_p r_b(p_0) A_b + A_2\xi_2(t), \\ \partial_t R = r_b\left(p_0\right). \end{cases} \tag{15}$$

Here $(\rho uA)\,|_L$ and $(\rho uAh_t)\,|_L$ are the mass and the enthalpy flow from the whole burning area of the propellant including the propellant surface in the hole and $p_0$, $\rho_0$, and $e_0$ are the stagnation values of the flow parameters. The total mass flow from the burning propellant surface is equal to the sum of the mass flows through the nozzle' and hole throats. Assuming that sonic conditions hold both in the nozzle throat and the hole throat we obtain the following result

$$\begin{aligned} \left(\rho uA\right)\big|_L &= \left(\rho_h u_h A_h\right)\big|_t + \left(\rho uA\right)\big|_t = \\ &\Gamma^{-1}\sqrt{\gamma\rho_0 p_0}\,A_{t,h} + \Gamma^{-1}\sqrt{\gamma\rho_0 p_0}\,A_t = \Gamma^{-1}\sqrt{\gamma\rho_0 p_0}\left(A_{t,h} + A_t\right) \end{aligned} \tag{16}$$

Here $\Gamma=((\gamma+1)/2)^{(\gamma+1)/2(\gamma-1)}$ and $A_{et}=(A_{t,h}+A_t)$ is the effective nozzle throat area. This relation means that in the first approximation the hole is seen by the internal flow dynamics as an increase of the nozzle throat area and the dynamics of the stagnation values of the gas parameters are governed by both dynamics of the propellant burning area (related to the thickness of the burned propellant layer $R$) and by the hole radius $R_h$. Substituting results of integration (16) into (15) and using model for nozzle ablation (11), (12) and hole melting (13), (14) we obtain the low-dimensional model of the internal ballistic of a subscale SRM in the presence of the case breach fault in the form

$$\begin{aligned} &\dot\rho = -\frac{c_0\Gamma A_{et}}{V r_b}\rho\sqrt{\frac{p}{\rho}} + \frac{A_b}{V}\left(\rho_p - \rho\right)p^n + a_1\xi_1(t), \\ &\dot p = -\frac{c_0\gamma\Gamma A_{et}}{V r_b}p\sqrt{\frac{p}{\rho}} + \frac{A_b}{V}\left(\gamma\rho_p - p\right)p^n + a_2\xi_2(t), \\ &\dot R = r_b p^n, \quad A_b = f(R), \quad \dot V = A_b\dot R = f(R) r_b p^n, \\ &\dot R_t = v_{abl} = v_m\left(\frac{p_0}{p_m}\right)^{1-\beta}\left(\frac{R_t}{R_{t0}}\right)^{-\beta}, \\ &\dot R_{ex} = v_m\left(\frac{p_0 A_t}{p_m A_{ex}}\right)^{1-\beta}\left(\frac{R_{ex}}{R_{ex0}}\right)^{-\beta}\frac{\left(T_{ex} - T_{abl}\right)}{\left(T_t - T_{abl}\right)}, \\ &\dot R_h = \frac{Q_c + Q_R + Q_b}{\left[q_{met} + C_{met}\left(T_{mel} - T_{m0}\right)\right]\rho_{met}}, \end{aligned} \tag{17}$$

where the following dimensionless variables are used

$$p \to \frac{p_0}{p_m},\ \rho \to \frac{\rho_0}{\rho_m},\ t \to \frac{t r_b(p_m)}{L_0},\ R_t \to \frac{R_t}{L_0},\ A_b \to \frac{A_b}{L_0^2},\ V \to \frac{V}{L_0^3},\ R_{t,h} \to \frac{R_{t,h}}{L_0},\ R \to \frac{R}{L_0},\ A_{et} \to \frac{A_{et}}{L_0^2}. \tag{18}$$

Here subscript $m$ refers to maximum reference values of the pressure and density and $L_0$ is characteristic length of the motor. We note that two first equations in (17) correspond to the "filling volume" approximation in (Salita, 1989 & 2001). The important difference is that we have introduced noise terms and the exact dependence of the burning surface on the burn distance in the form of the design curve relation in the fourth equation in (17). We have also established an explicit connection with the set of partial differential equations (1) that helps to keep in order various approximations of the Eqs. (1), which are frequently used in practice and in our research.
The equations above have to be completed by the equations for the main thrust $F$ and lateral (side) thrust $F_h$ induced by the gas flow through the hole in the form

$$F = \Gamma^{-1}\sqrt{\gamma\rho_0 p_0}\,A_t u_{ex} + (p_{ex} - p_a)A_{ex}, \quad F_h = \Gamma^{-1}\sqrt{\gamma\rho_0 p_0}\,A_{t,h} u_{h,ex} + (p_{ex,h} - p_a)A_{h,ex} \qquad (19)$$

where $p_a$ is ambient pressure, $u_{ex}$ and $u_{h,ex}$ are gas velocities at the nozzle outlet and hole outlets respectively, and $p_{ex}$ and $p_{h,ex}$ are the exit pressure at the nozzle outlet and hole outlets respectively.

### 3.2 Axial distributions of the flow variables in a sub-scale motor

It follows from the analysis that $M_0^2 = v^2/c_0^2 \ll 1$ is small everywhere in the combustion chamber. Furthermore, the equilibration of the gas flow variables in the chamber occurs on the time scale ($t = L/c$) of the order of milliseconds. As a result, the distribution of the flow parameters follows adiabatically the changes in the rocket geometry induced by the burning of the propellant surface, nozzle ablation and metal melting in the hole through the case. Under these conditions it becomes possible to find stationary solutions of the Eqs. (1) analytically in the combustion chamber. Taking into account boundary conditions at the stagnation point and assuming that the spatial variation of the port area $A_p(x)$ is small and can be neglected together with axial component of the flow at the propellant surface $u_S(x)$, we obtain the following equations for the spatial variation of the flow parameters (Osipov et al., March 2007)

$$\left(S\rho u\right)\Big|_0^x = \int_0^x \rho_p r_b l dx, \quad \rho u^2 + p = p_0, \quad \left(S\rho h_t u\right)\Big|_0^x = H\int_0^x \rho_p r_b l dx. \qquad (20)$$

$$\rho = \rho_0\left(1+\frac{\gamma+1}{2}\frac{u^2}{c_0^2}\right)^{-1}, \; p = p_0\left(1+\frac{\gamma+1}{2}\frac{u^2}{c_0^2}\right)^{-1}\left(1-\frac{\gamma-1}{2}\frac{u^2}{c_0^2}\right)^{-1}, \; v = \frac{u_L x}{L}\left(1+\frac{3(\gamma+1)-2n\gamma)}{6c_0^2}u_L^2\frac{x^2}{L^2}\right). \qquad (21)$$

and in the nozzle area

$$p = p_0\left(1-\frac{(\gamma-1)}{2}M_0^2\right)^{\frac{\gamma}{\gamma-1}}, \quad \rho = \rho_0\left(1-\frac{(\gamma-1)}{2}M_0^2\right)^{\frac{1}{\gamma-1}}, \quad T = T_0\left(1-\frac{(\gamma-1)}{2}M_0^2\right), \qquad (22)$$

where $M_0$ is given by the solution of the nozzle equation

$$M_0\left(1-\frac{(\gamma-1)}{2}M_0^2\right)^{\frac{1}{\gamma-1}} = \frac{A_t}{\Gamma A}.$$

### 3.3 Verification and validation (V&V) of the "filing volume" model

To verify the model we have performed high-fidelity simulations using code by C. Kiris (Smelyanskiy et al., 2006) and FLUENT model (Osipov et al., 2007; Luchinsky et al., 2008). To solve the above system of equations numerically we employ a dual time-stepping scheme with second order backward differences in physical time and implicit Euler in psuedo-time, standard upwind biased finite differences with flux limiters for the spatial derivative and the source terms are evaluated point-wise implicit. For these simulations the following geometrical parameters were used: initial radius of the grain $R_0$ = 0.74 m, $R_t$ = 0.63 m, $L$ = 41.25m; $\rho$ = 1800 kg·m$^{-3}$, $H$ = 2.9x106 J·kg$^{-1}$, $r_c$ = 0.01 m·sec$^{-1}$, $p_c$ = 7.0x10$^6$ Pa. The results of integration for a particular case of the neutral thrust curve are shown in the Fig. 1(b). The fault (the nozzle throat radius is reduced by 20%) occurs at time $t_f$ = 15 sec. The comparison of the results of the simulations of the model (1) with the solution of the LDM (17) is shown in the Fig. 3(a). It can be seen from the figure that the LDM reproduces quite accurately the dynamics of the internal density in the nominal and off-nominal regimes. Similar agreement was obtained for the dynamics of the head pressure and temperature.

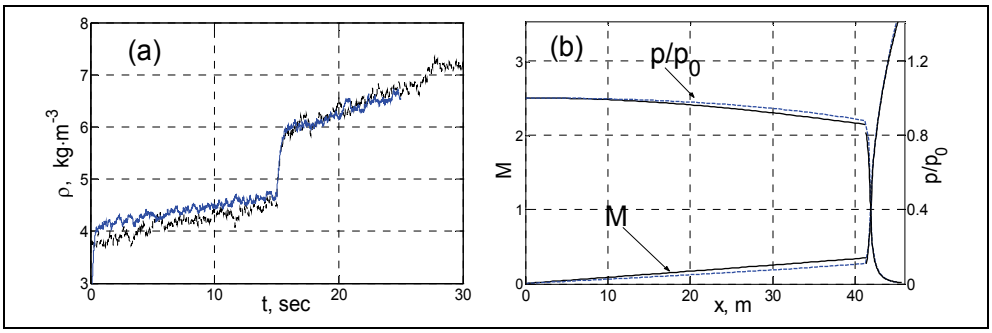



Fig. 3. (a) Comparison between the results of integration of the stochastic partial differential equations Eqs. (1), (2)(solid blue lines) and stochastic ordinary differential equations Eqs. (17)(dotted black lines) for the time evolution of the head density. (b) Comparison between the numerical (dashed blue lines) and analytical (solid lines) solutions for the gas velocity and pressure.

The comparison of the analytical solution (21), (22) for axial distribution of the pressure and velocity with the results of numerical simulation of the high-fidelity model is shown in the Fig. 3(b). It can be seen from the figure that the axial variation of the gas flow parameters is small and agrees well with the results of numerical integration. Therefore, the dynamics of the SRMs operation with small variation of the port area along the rocket axis can be well characterized by the LDM (17), obtained by integration of Eqs. (1), (2) over the length of the combustion camera.

This conclusion is also supported by the 2D high-fidelity simulations using FLUENT. To simulate time evolution of the propellant regression, nozzle ablation, and the hole burning through we have introduced the following deforming zones (see Fig. 4): (i) hole in the forward closure; (ii) nozzle ablation; and (iii) variation of the burning area as a function of time. In simulations we have used a density based, unsteady, implicit solver. The mesh was initialized to the stagnation values of the pressure, temperature, and velocity in the combustion chamber and to the ambient values of these variables in the two ambient

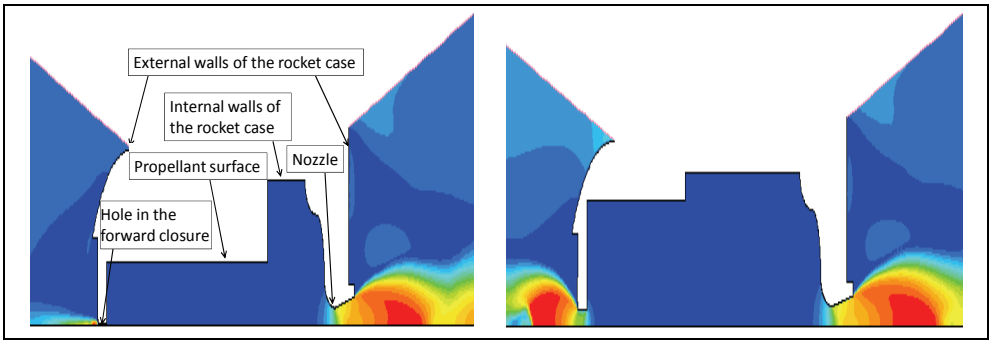



Fig. 4. 2D velocity distribution with axial symmetry obtained using FLUENT simulations after 0.14 sec (left) and t = 5.64 (right). The geometry of the model surfaces is shown in the figure. The propellant surface wall, hole wall, and the nozzle wall are deforming according to the equations (2), note the changes in the geometry of the rocket walls and the corresponding changes in the velocity distribution.

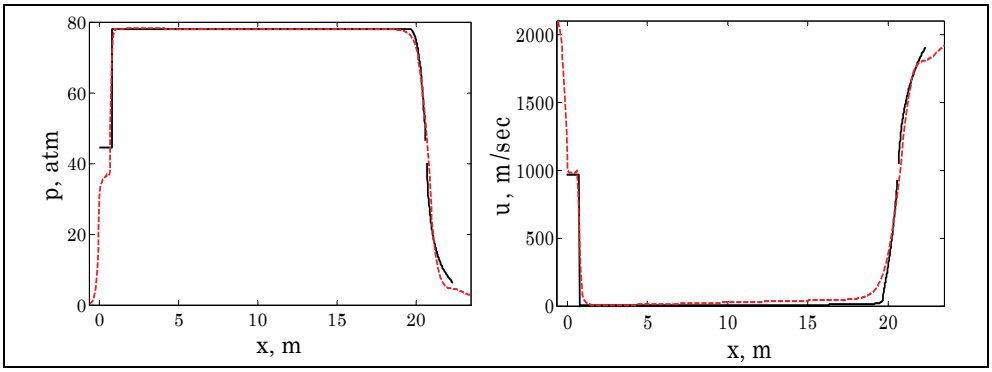



Fig. 5. Axial velocity (left) and pressure (right) profiles generated by the FLUENT model for $t$=0.05 sec (red dashed line) as compared to the analytical solutions (black solid lines) given by the (21), (22).

regions on the right and left of the chamber. The results of the comparison of the analytical distributions (21)-(22) with the axial velocity and pressure distributions obtained using FLUENT simulations are shown in the Fig. 5. It can be seen from the figure that the model (17), (21)-(22) provides a very good approximation to the results of FLUENT simulations.

Note that the difference in the time scales for dynamics of burn distance, metal erosion, and nozzle ablation as compared to the characteristic relaxation time of the distributions to their quasi-stationary values $t_{rel}$ allows us to integrate equations (1), (2) in quasi-stationary approximation as will be explained in details in Sec. 5. As a result we obtain the analytical solution for the quasi-stationary dynamics of the axial distributions of the gas parameters in the combustion chamber and in the nozzle area. The comparison of this analytical solution with the results of FLUENT simulations also demonstrates agreement between the theory and numerical solution of the high-fidelity model. The accuracy of the low-dimensional model (17) was further validated using results of a ground firing test for a subscale motor as will be described in details elsewhere.

## 4. Bayesian inferential framework for internal SRMs ballistics

We are now in a position to introduce a novel Bayesian inferential framework for the fault detection and prognostics in SRMs. Note that the effect of the case breach fault and nozzle blocking on the dynamics of the internal gas flow in SRMs is reduced to the effective modification of the nozzle throat area $A_{et}(t)$ as explained above. In a similar manner the effects of bore choking and propellant crack can be taken into account by introducing an effective burning area and by coupling the analysis of the pressure time-traces with the analysis of the nozzle and side thrust. The accuracy of the calculations of the internal SRM ballistics in sub-scale motors in nominal and off-nominal regimes based on the LDM (17) allows us to use it to verify the FD&P in numerical simulations.

### 4.1 Bayesian framework

The mathematical details of the general Bayesian framework are given in (Luchinsky et al., 2005). Here we briefly introduce earlier results in the context of fault detection in SRMs including abrupt changes of the model parameters. The dynamics of the LDM (17) can be in general presented as an Euler approximation of the set of ODEs on a discrete time lattice $\{t_k=hk; k=0,1,...,K\}$ with time constant $h$

$$x_{k+1} = x_k + hf(x_k^* \mid c) + \hat{\sigma}\sqrt{h}z_k, \tag{23}$$

where $z_k = \frac{1}{\sqrt{h}}\int_{t_k}^{t_k+h} \xi(t)dt$, $x_k^* = \frac{x_k + x_{k+1}}{2}$, $\boldsymbol{x_k} = \{p, \rho, R, V, r_h, r_t, r_i\}$ is $L$-dimensional state of the system (17), $\sigma$ is a diagonal noise matrix with two first non-zero elements $a_1$ and $a_2$, $\boldsymbol{f}$ is a vector field representing the rhs of this system, and $c$ are parameters of the model. Given a Gaussian prior distribution for the unknown model parameters, we can apply our theory of Bayesian inference of dynamical systems (Luchinsky et al., 2005) to obtain

$$D_{ij} = \frac{h}{K}\sum_{k=0}^{K-1} (\dot{x}_k - f(x_k;c))_i (\dot{x}_k - f(x_k;c))_j \tag{24}$$

$$c'_l = (A)^{-1}_{ml} w_m, \tag{25}$$

where elements $A_{ml}$ and $w_m$ are defined by the following equations

$$w_m = h\sum_{k=0}^{K-1}\left[\sum_{n,n'=1}^{L} U'_{mn}(t_k) D^{-1}_{nn'} \dot{x}_{n'}(t_k) - \frac{v_m}{2}\right] \tag{26}$$

$$A_{ml} = h\sum_{k=0}^{K-1}\left[\sum_{n,n'=1}^{L} U'_{mn}(t_k) D^{-1}_{nn'} U_{n'l}(t_k)\right]. \tag{27}$$

Here the vector field is parameterized in the form $f(x;c)=\hat{U}(x)c$, where $\hat{U}(x)$ is a block-matrix with elements $U_{mn}$ build of $N$ blocks of the form $\hat{I}\phi_n(x(t_k))$, $\hat{I}$ is $L$x$L$ unit matrix, and

$$v_m(\mathbf{x}) = \sum_{n=1}^{N} \frac{\partial U_{nm}(\mathbf{x})}{\partial x_n}.$$

To verify the performance of this algorithm for the diagnostics of the case breach fault we first assume the nominal regime of the SRM operation and check the accuracy and the time resolution with which parameters of the internal ballistics can be learned from the pressure signal only. To do so we notice that equations for the nozzle throat radius $r_t$, burn distance $R$, and combustion chamber volume can be integrated analytically for a measured time-traces of pressure and substituted into the equations for pressure dynamics. By noticing further that for small noise-intensities the ratio of dimensionless pressure and density $p/\rho \approx 1$ obtain the following equation for the pressure dynamics

$$\dot{p} = -\frac{c_0 \gamma \Gamma A_{et}}{V r_b} p + \frac{A_b}{V}\left(\gamma \rho_p - p\right) p^n + \sqrt{D}\xi_2(t), \tag{28}$$

where $A_t(t)$, $A_b(t)$, and $V(t)$ are known functions of time given by the following equations

$$\begin{gathered} R(t) = \int_0^t p^n(t')dt', \quad A_b(t) = f\left(R(t)\right), \quad V(t) = V_0 + \int_0^{R(t)} A_b(R)dR, \\ A_t(t) = \pi r_t^2(t), \quad R_t(t) = \left[ R_{t0}^{1+\beta} + v_m(1+\beta)\int_0^t p^{1-\beta}(t')dt' \right]^{\frac{1}{1+\beta}} \end{gathered} \tag{29}$$

The parameters $c_0\gamma\Gamma/r_b$, $\gamma\rho_p$, and $D$ can now be inferred in the nominal regime by applying Eqs. (23)-(27) to the analysis of equation (28). An example of the inference results is shown in the Table 1.

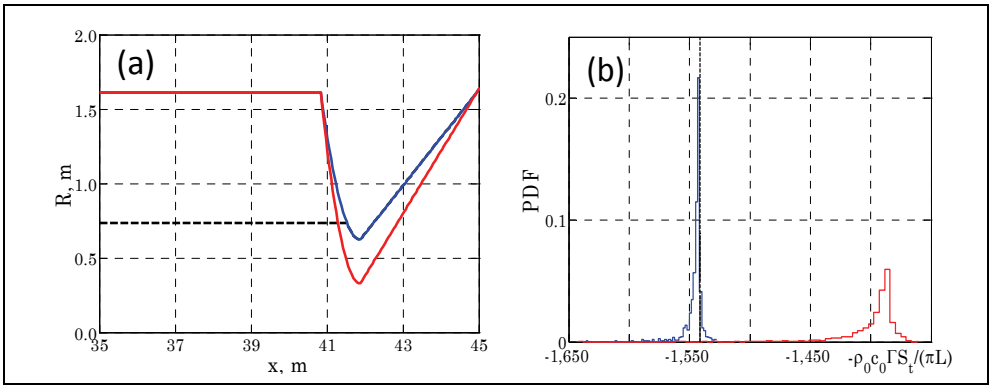



Fig. 6. (a) An example of the geometry of the simulations of the nozzle failure model using Eqs. (1), (2). The geometry of the case before and after the fault is shown by the solid blue and red lines respectively. (b) estimation of the value of the parameter $-c_0GA_t/(pL)$ before (left curve) and after (right curve) the fault. The dashed line shows the actual value of the parameter. The solid lines show the PDF of the parameter estimation with $T$=0.1 sec, $\Delta t$=0.001 sec, $N$=500 (see the caption for the Table 1).

| Parameters | Actual | Inferred | Relative error |
|---|---|---|---|
| $\gamma\rho_p$ | 248.2 | 244.7 | 1.4% |
| $-c_0\Gamma/r_b$ | -61260 | -61347 | 1.38% |
| $D$ | $2.5\times10^{-4}$ | $2.44\times10^{-4}$ | 2.4% |

Table 1. The results of the parameter estimation of the model (28), (29) in the nominal regime. The total time of the measurements in this test was $T$=1 sec, the sampling rate was 1 kHz, and the number of measured points was $N$=1000.

We conclude that the parameters of the nominal regime can be learned with good accuracy during the first few second of the flight. This result allows one to apply Bayesian algorithm for fault detection and diagnostics in SRMs.

We now provide numerical example explaining in more details how this technique can be used for in-flight FD&P in SRMs. We will be interested to verify if the Bayesian framework can provide additional information ahead of the "alarm" time about the most likely course of the pressure dynamics to reduce the probability of the "misses" and "false alarms". To model the "miss" situation a case will be considered when small pressure deviation from the nominal value persists for a few second prior to the crossing the "alarm" level and the time window between the "alarm" and "catastrophe" becomes too short. This situation is illustrated in the Fig. 7(a), where measured pressure signal (black solid line) crosses the alarm level (dashed line) initiating the alarm at approximately $t_A \approx 15$ sec. The overpressure fault occurs at $t_F \approx 17$ sec and the time window between the alarm and a "catastrophic" event becomes too short, which can be considered as a model of "miss" situation. To model the "false alarm" situation a case will be considered in which the pressure crosses the "alarm" level, but then returns to its nominal value (see Fig. 7(b)). In all the simulations presented here the overpressure fault was modeled as a reduction of the nozzle throat area. Note, however, that the results discussed below can be extended to encompass other faults, including e.g. the propellant cracking, bore choking, and case breach as will be discussed below.

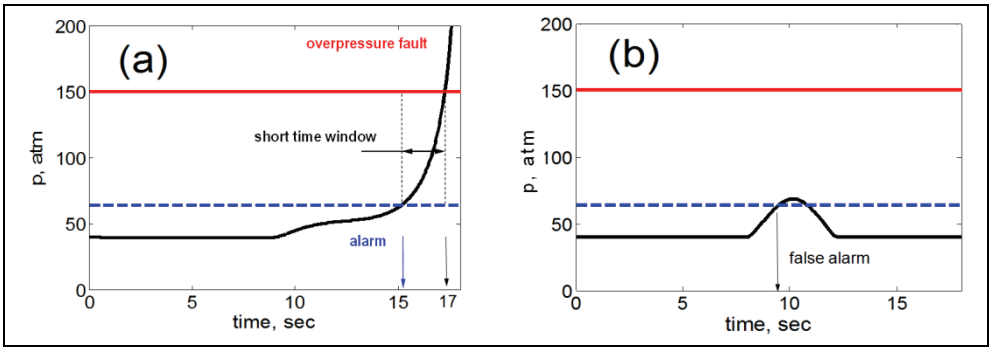



Fig. 7. (a) Example of possible time variation of the pressure fault (black line) representing a possible "miss" situation. The blue dashed and red solid lines indicate the "alarm" and the "catastrophe" levels respectively. Note that the time window between the "alarm" and the "catastrophe" is too short. (b) Example of possible time variations of the fault pressure representing a possible "false alarm" situation. The blue dashed and red solid lines are the same as in (a).

### 4.2 Modeling "misses" for the nozzle failure and neutral thrust curve

To model the "misses" we assume that the time evolution of the nozzle fault is highly nonlinear and can be described by a polynomial function

$$A_{et} = A_{t0} - \Delta A\left(\alpha\tau + \beta\tau^2 + \delta\tau^3\right) \tag{30}$$

corresponding e.g. to the slow degradation followed by the fast destruction of the nozzle walls as shown in the Fig. 7(a), where $\tau$ is the time elapsed from the fault initialization. In this case the time window between the "alarm" and the overpressure fault becomes too short and effectively the FD&P system "misses" the event. The thrust curve is chosen to be neutral. Our goal is to demonstrate that application of the Bayesian framework for the SRM FD&P allows one to extend substantially the time window between the "alarm" and the overpressure fault thereby reducing the probability of "misses". To this end we extend the model described by Eqs. (17) by including nonlinear terms from Eq.(30). The corresponding vector field of the Eq. (28) can be written as $f(x;c)=\hat{C}\phi$ with the set of the base functions given by Eq. (31) and the set of the model parameters is given in Eq.(32), where $a=(c_0\Gamma)/(\pi L r_{b0} R^*)$.

$$\vec{\phi} = \left\{ p^n, \frac{p^n}{R}, \frac{p^{n+1}}{R}, \frac{p^n\rho}{R}, \sqrt{\frac{p}{\rho}}\frac{p}{R}, \sqrt{\frac{p}{\rho}}\frac{p\tau}{R}, \sqrt{\frac{p}{\rho}}\frac{p\tau^2}{R}, \sqrt{\frac{p}{\rho}}\frac{p\tau^3}{R}, \ldots \right.$$
$$\left. \sqrt{\frac{p}{\rho}}\frac{\rho}{R}, \sqrt{\frac{p}{\rho}}\frac{\rho\tau}{R}, \sqrt{\frac{p}{\rho}}\frac{\rho\tau^2}{R}, \sqrt{\frac{p}{\rho}}\frac{\rho\tau^3}{R} \right\} \tag{31}$$

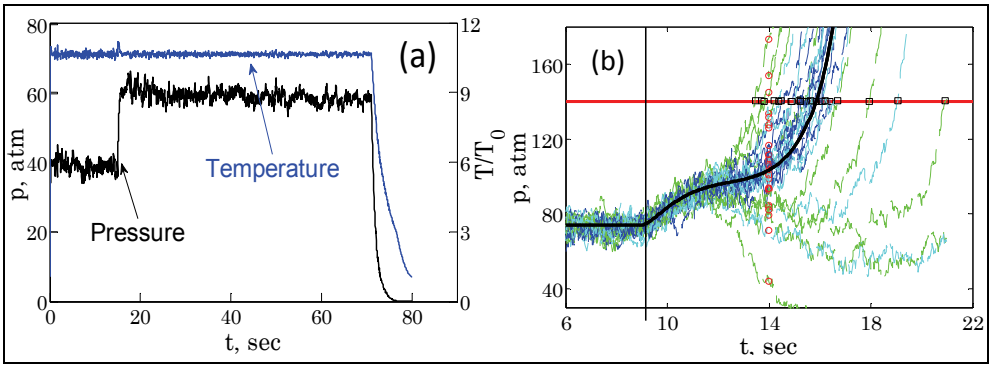



Fig. 8.(a) An example of the time-traces of temperature (blue line) and the pressure (black line) of the SRM operation with neutral thrust curve. Fault corresponding to abrupt changes of the nozzle throat area (cf Fig. 6(a)) occurs at $t$=17 sec. (b) Nonlinear time evolution of the pressure build up after the nozzle blocking fault is shown by the back solid line. Predicted dynamics of the pressure is shown by the jiggling lines. The results of the predictions build 1sec, 1.5sec, and 2.1 sec after the fault are shown by green, cyan, and blue lines correspondingly. The values of the pressure at $t$=14 sec, which are used to build the PDF of the pressure, are shown by red circles. The time moments of the predicted overpressure faults used to build the PDF of the case burst times as shown by the black squares on the red margin line. Fault occurs at $t$=9 sec.

$$\hat{C} = \begin{bmatrix} 0 & 2\gamma\rho_p & -2 & 0 & -a\gamma & -a\gamma\alpha & -a\gamma\beta & -a\gamma\delta & 0 & 0 & 0 & 0 \\ 0 & 2\rho_p & 0 & -2 & 0 & 0 & 0 & 0 & -a & -a\alpha & -a\beta & -a\delta \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (32)$$

Parameters of the system are monitored in real time. Once small deviations from the nominal values of the parameters is detected at time $t_d$ the algorithm is continuously updating the inferred values of parameters estimated on increasing intervals $\Delta t$ of time elapsed from $t_d$ . These values are used to generate a set of trajectories predicting pressure dynamics. Example of such sets of trajectories calculated for three different time intervals $\Delta t$ =1sec, 1.5 sec, and 2.1 sec are shown in the Fig. 8(b) by green, cyan, and blue lines respectively. These trajectories are used to predict the PDFs of the head pressure for any instant ahead of time. An example of such PDF for the pressure distribution at time at $t$=14 sec is shown in the Fig. 9(a). The method used to calculate PDF for the pressure distributions is illustrated in the Fig. 8(b). The same trajectories are used to predict the PDFs of the time moment of the overpressure fault as illustrated in the Fig. 8(b) and Fig. 9(b). It can be seen from the figures that the distribution of the predicted time of the overpressure fault converges to the correct value 2.1 sec after the fault thereby extending the time window between the "alarm" and the fault to 6 sec which is almost three folds of the time window obtained using standard technique.

Therefore, we conclude that the Bayesian framework provides valuable information about the system dynamics and can be used to reduce the probability of the "misses" in the SRM FD&P system. A similar analysis shows (Luchinsky et al., 2007) that the general Bayesian framework introduced above can be applied to reduce the number of "false alarms".

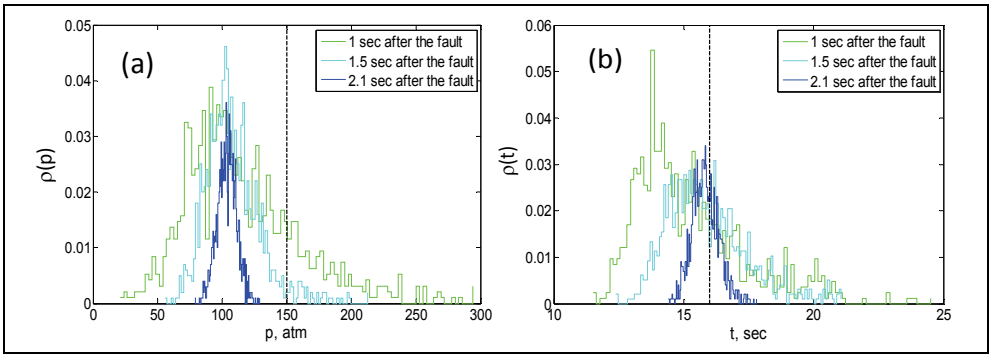



Fig. 9. (a) The PDF of the predicted values of pressure at t=14 sec build 1 sec (green line), 1.5 sec (cyan line), and 2.1 sec (blue lines) after the fault. The dashed vertical line shows the dangerous level of the pressure. (b) The PDF of the predicted times of the overpressure fault build 1sec (green line), 1.5 sec (cyan line), and 2.1 sec (blue lines) after the fault. The dashed vertical line shows the actual time when the overpressure fault is going to happen.

### 4.3 Self-consistent iterative algorithm of the case breach prognostics

In the previous section we have shown that in-flight FD&P for SRMs can be developed within Bayesian inferential framework. The introduced technique can be very useful in a wide range of contexts including in particular active control of combustion instabilities in

liquid motors (Hathout et al, 2002). In practice, however, it is often desirable (see also the following section) to further simplify the algorithm by avoiding stochastic integration. The simplification can be achieved by neglecting noise in the pressure time-traces and by considering fault dynamics in a regime of quasi-steady burning.

To illustrate the procedure of building up iterative FD&P algorithms that avoids stochastic integration let us consider the following example problem. A hole through the metal case and insulator occurs suddenly at the initial time of the fault $t_0$. The goal is to infer and predict the dynamics of the growth of the holes in the insulator layer and in the metal case, as well as the fault-induced side thrust, and changes to the SRM thrust in the off-nominal regime. In this example the model for the fault dynamics is assumed to be known. This is a reasonable assumption for the case breach faults with simple geometries. For this case the equations can be integrated analytically in quasi-steady regime and the prognostics algorithm can be implemented in the most efficient way using a self-consistent iterative procedure, which is developed below. As an input, we use time-traces of the stagnation pressure in the nominal regime and nominal values of the SRM parameters. In particular, it is assumed that the ablation parameters for the nozzle and insulator materials and the melting parameters for the metal case are known. It is further assumed that the hole radius in the metal case is always larger than the hole radius in the insulator (i.e. the velocity of the ablation of the insulator material is smaller than the velocity of the melting front), accordingly the fault dynamics is determined by the ablation of the insulator. This situation can be used to model damage in the metal case induced by an external object.

To solve this problem we introduce a prognostics algorithm of the fault dynamics based on a self-consistent iterative algorithm that avoids numerical solution of the LDM. We notice that with the limit of steady burning, the equations in (17) can be integrated analytically. Because the hole throat is determined by the radius of the hole in the insulator, we can omit the equation for the hole radius in the metal case. The resulting set of equations has the form

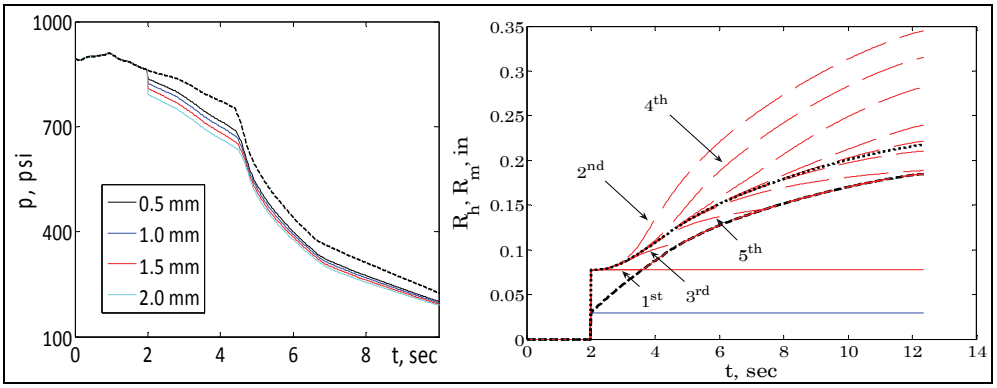



Fig. 10. (left) Results of the calculations using iterative algorithm A1. Absolute values of pressure for four different initial values of the hole in the case: 0.5, 1.0, 1.5 and 2.0 mm are shown by the black, blue, red, and cyan solid lines respectively. The nominal pressure is shown by the dashed black line. (right) Iterations of the effective hole radius in the metal case. Red solid line shows 0th approximation. Five first approximations shown by red dashed lines are indicated by arrows. Final radius of the hole in the metal case is shown by black dotted line. 0th approximation for the hole in the insulator is shown by dashed blue line. Final radius of the hole in the insulator is shown by the black dashed line.

$$\begin{cases} p(t) = p_N(t)\left[\dfrac{A_{bef}(t)A_t(t)}{A_b A_{et}(t)}\right]^{\frac{1}{1-n}}, \\ R = a\int\limits_0^t p^n(\tau)d\tau, \\ R_t^{1+\beta_t}(t) = R_{t0}^{1+\beta} + (1+\beta)v_{n0}\int\limits_{t_0}^t p^{1-\beta}(\tau)d\tau, \\ R_h^{1+\beta_i}(t) = R_{h0}^{1+\beta_i} + (1+\beta)v_{h0}\int\limits_{t_0}^t p^{1-\beta}(\tau)d\tau. \end{cases} \quad (33)$$

Here $A_{et}(t) = A_t(t) + \Delta A_t(t) + A_h(t)$ is an effective nozzle throat area where the 1st term corresponds to the nominal regime, the 2nd term corresponds to the deviation of the nozzle throat area from the nominal regime due to the fault, and the 3rd term corresponds to the area of the hole in the rocket case. Similarly, we define the effective burning area $A_{bef}(t) = A_b(t) + \Delta A_b(t)$ as a sum of the burning area in the nominal regime and a term that describes the deviation of the burning area from the nominal regime due to the fault. Using Eqs (33) the following iterative algorithm **A1** can be introduced:

1. Set initial values of the corrections to the nozzle and burning area to zero $\Delta A_t(t) = 0$ and $\Delta A_b(t)=0$. Set values of the areas of the holes in the metal and in the insulator to constant initial values $A_m(t) = \pi \cdot R_{m0}^2$ and $A_h(t) = \pi \cdot R_{h0}^2$.
2. Update time-trace of the pressure using 1st eq. in (33)
3. Update burn web distance $R$, radius of the hole in the insulator $R_h$, and nozzle throat radius $R_t$ using last three Eqs. In (33).
4. Repeat from the step (2) until convergence is reached.

The results of the application of this self-consistent algorithm to the prognostics of the case breach fault parameters are shown in Fig. 10(left). Once quasi-steady pressure and the dynamics of the hole growth in the insulator are predicted in the off-nominal regime one can determine the dynamics of the hole growth in the metal case and the dynamics of the fault-induced side thrust. To do so, we use the following self-consistent iterative algorithm **A2** for $t>t_0$ that takes into account the assumption that the velocity of the melting front is larger than the velocity of ablation in the insulator.

1. Set 0th approximation $R^{(0)}_{h0}(t)$ for the hole radius in the metal to $r_{h0}$.
2. Construct 1st approximation

$$R_m^{(1)}(t) = \begin{cases} R_m^{(0)}(t), & \text{if } R_m^{(0)}(t) > R_h(t) \\ R_h(t), & \text{if } R_m^{(0)}(t) \le R_h(t) \end{cases}$$

3. Solve Mach equation (use a maximum root)

$$M_0(t) = \begin{cases} M_0\left(1 - \dfrac{\gamma - 1}{2}M_0^2\right)^{\frac{1}{\gamma - 1}} = \dfrac{A_h}{\Gamma A_m}, & \text{if } R_m^{(0)}(t) > R_h(t) \\ \dfrac{c_t}{c_0} = \sqrt{\dfrac{2}{\gamma + 1}}, & \text{if } R_m^{(0)}(t) \le R_h(t) \end{cases}$$

4. Find the effective radius of the hole in the metal case

$$R_m^{1+\beta} = R_{m0}^{1+\beta} + v_{m0}\int_{t_0}^{t}\left(\frac{(\rho v)_m(\tau)}{(\rho v)_*^{\max}}\right)^{1-\beta}\left(\frac{T_m(\tau)-T_{mel}}{T_t-T_{mel}}\right)d\tau$$

5. Use Eqs (21)-(22) to find velocity $u_m(t)=M_{0m}(t)\cdot c_0$, mass flow $(\rho u)_m(t)$, temperature $T_m(t)$, and pressure $p_m(t)$ in the metal hole;
6. Repeat from the step (2) until convergence is reached.
7. Calculate fault-induced side thrust $F_h = \left[(\rho u)_m u_{m,ex} + (p_{m,ex} - p_a)\right] S_m$.

A similar algorithm is used to find the ablation of the nozzle and SRM thrust in the off-nominal regime in the case breach fault. The results of these calculations are shown in Fig. 10(right) and Fig. 11. We note that the fault diagnostics is achieved using the same iterative algorithm with the only exception being that the time-trace for pressure in the off-nominal regime is given by the measurements. Accordingly the first equation in the set (33) and the 2nd step in the iterative algorithm **A1** are not needed. Also note that an important feature of the algorithms introduced above is the assumption that the design curve $A_b = f(R)$ representing the relation between the burning area $A_b$ and burn web distance $R$ is known and remains invariant characteristics of the SRM in the off-nominal regime of the case breach fault.

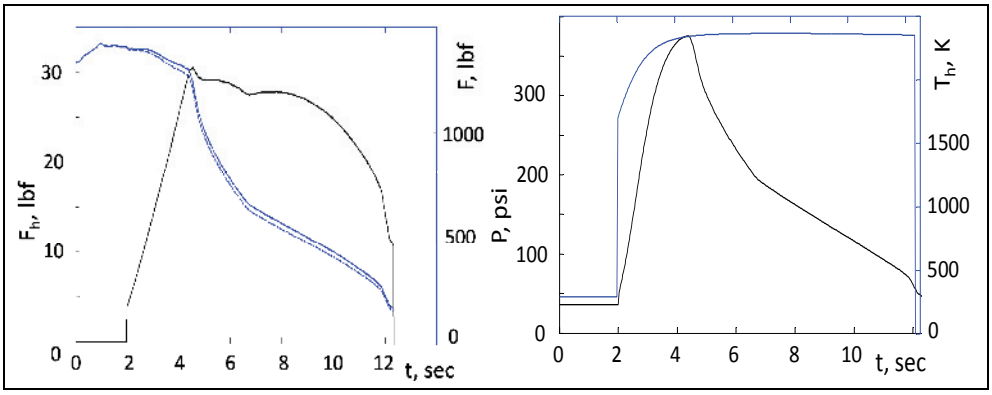



Fig. 11. (left) Fault-induced thrust (black solid line) is shown in comparison with nominal SRM thrust (blue solid line) and off-nominal SRM thrust (dashed blue line). Initial radius of the hole in the insulator is 0.75mm. (right) Pressure (black line) and temperature (blue line) in the metal hole through the case determined by the iteration algorithm A2.

The deep physical meaning of the iterative procedure introduced above rests upon the idea that the ablation of the hole walls and the case breach fault develop in a self-consistent manner. Indeed, the increase of the cross section leak, due to insulator ablation under the action of the hot gas flow, leads to decreased pressure and hence a decreased burning rate. This, in turn, decreases the hot gas flow through the hole and the ablation rate. In this way a quasi-stationary regime of burning and ablation is developed. The parameters of burning in this regime can be found in a self-consistent way using an iterative algorithm, without integration of the full system of differential equations of motion.

If the fault dynamics is determined by hole growth rate in the metal case the algorithm above is still applicable and can be substantially simplified, because the ablation in the isolator can be neglected and sonic condition holds in the hole throat. The resulting algorithm is algorithm **A1** extended using one equation from step 7 of algorithm **A2**. We now consider how the FD&P system can be extended to a large segmented motor. In what follows the fault dynamics is determined by hole growth rate in the metal case, which is the most plausible situation in practice.

## 5. FD&P for large segmented SRMs

To extend the FD&P algorithm to a large segmented motor we will first simplify the model of its internal ballistic in the nominal and off-nominal regimes introduced in Sec. 2. Combining the equations of gas dynamics with the dynamics of propellant regression, nozzle ablation, and case breach fault the performance model of the large segmented SRM in the presence of faults can be summarized in the set of Eqs. (34) with the conservative variables $U$ and function $f(U)$ given by Eq. (2) and the source terms $S$ given by Eq. (3) with neglected noise terms.

$$\begin{cases} \partial_t\left(UA_p\right)+\partial_x\left(f(U)A_p\right)=S, \\ A_p(x_i)=f_{Ai}(R(x_i)), \qquad l(x_i)=f_{li}(R(x_i)), \\ \dot{R}=ap^n+C\left(I-I_{cr}\right), \\ \dot{R}_t=v_{m,t}\left(\dfrac{p_0}{p_m}\right)^{1-\beta}\left(\dfrac{R_t}{R_{t,in}}\right)^{-\beta}, \\ \dot{R}_{ex}=v_{m,ex}\left(\dfrac{p_0A_t}{p_mA_{ex,in}}\right)^{1-\beta}\left(\dfrac{R_{ex}}{R_{ex,in}}\right)^{-\beta}\dfrac{\left(T_{ex}-T_{abl}\right)}{\left(T_t-T_{abl}\right)}, \\ \dot{R}_h=\dfrac{Q_c+Q_R+Q_b}{\left[q_{met}+C_{met}\left(T_{mel}-T_{m0}\right)\right]\rho_{met}}. \end{cases} \tag{34}$$

### 5.1 Numerical integration of the model

We notice that (34) is a system of nearly balanced PDEs with slowly varying parameters. This is an example of PDEs with multiple time scales (Knoll et al., 2003), where the slower dynamical time scale is a result of a near balance between $\partial_x(f(U)A_p)$ and $S$ in the first equation and slowly varying parameters in the last four equations in (34). The fast dynamics of (34) corresponds to the acoustic time scale. To see the multiple time scale character of the system (34) more clearly let us introduce dimensionless variables

$$p\to\frac{p}{p_0},\ \rho\to\frac{\rho}{\rho_0},\ \rho_p\to\frac{\rho_p}{\rho_0},\ T\to\frac{T}{T_0},\ R_N\to\frac{R_N}{L_0},\ t\to\frac{t}{t_0},\ R\to\frac{R}{L_0},\ l\to\frac{l}{L_0},\ A_p\to\frac{A_p}{L_0^2},\ u\to\frac{u}{c_0} \tag{35}$$

where $t_0 = L_0/ap_0^n \approx 10^{-2}$ sec; $p_0=p_0(t=0)$, $\rho_0=\rho_0(t=0)$ are the gas pressure and density near the rocket head at a start time point after the ignition, $M_0 = u/c_0$, $L_0 \approx 1$m are characteristics scales of time and length ($r_{p0} = ap_0^n$ is a typical burning rate). In dimensionless variables the first equation in (34) can be rewritten as follows

$$\partial_x A_p \begin{bmatrix} \rho M_0 \\ \gamma\rho M_0^2 + p \\ \rho M_0 \left(T + \frac{\gamma-1}{2} M_0^2\right) \end{bmatrix} = -\varepsilon \partial_t A_p \begin{bmatrix} \rho \\ \rho M_0 \\ \rho \left(T/\gamma + \frac{\gamma-1}{2} M_0^2\right) \end{bmatrix} + \begin{bmatrix} \varepsilon\rho_p \dot{R} l(x) + \rho_{t,h} M_{0,t} A_h \delta(x - x_0) \\ p\partial_x A_p - \gamma\lambda\rho M_0^2 l(x) \\ \varepsilon\rho_p \dot{R} l(x) + \rho_{t,h} M_{0,t} A_h \delta(x - x_0) \end{bmatrix} \tag{36}$$

Here we have introduced small parameter $\varepsilon = L_0/(t_0 c_0) < 10^{-5}$ corresponding to the ratio of the characteristic velocity of the propellant surface regression ($r_{p0} \approx 10^{-2}$ m/sec) to the speed of sound ($c_0 \approx 1006$m/sec). It is clear that in the first approximation at each given moment of time the axial distribution of the flow variables in a segmented rocket can be found in quasi-steady approximation neglecting a small last term proportional to $\varepsilon = 10^{-5}$. Note that two source terms in the 1st and 3rd Eqs. of (36) are also $\propto \varepsilon$ but these terms cannot be neglected, because they are proportional to $\rho_p \approx 10^2$.

To solve equations (36) one can neglect the first term on the right hand side $\propto \varepsilon$ and to complete resulting system of ODEs by a set of boundary conditions. The calculation of the axial distribution of the flow parameters in the quasi-steady approximation can be reduced to the integration of the system of ODEs with respect to spatial coordinate $x$. To this end it is convenient to write explicitly Euler approximation of Eqs. (36) in quasi-steady regime on a coarse-grained (in general non-uniform) lattice of axial coordinates $\{x_i: i=1,\ldots,N\}$

$$\begin{aligned} \rho M_0 A_p\big|_i &= \rho M_0 A_p\big|_{i+1} - \varepsilon\rho_p \dot{R} l(x) dx_i + \rho_{t,h} M_{0,th} A_h\big|_{i_h}, \\ \left(\gamma\rho M_0^2 + p\right) A_p\big|_i &= \left(\gamma\rho M_0^2 + p\right) A_p\big|_{i+1} - p\left(A_p\right)\Big|_i^{i+1} + \gamma\lambda\rho M_0^2 l(x) dx_i, \\ \rho M_0 h_T A_p\big|_i &= \rho M_0 h_T A_p\big|_{i+1} - \varepsilon\rho_p \dot{R} l(x) dx_i + h_{t,h} \rho_{t,h} M_{0,th} A_h\big|_{i_h}. \end{aligned} \tag{37}$$

where $h = \left(T + \frac{\gamma-1}{2} M_0^2\right) / H$. The dynamics of the case breach fault in this approximation is determined by the dynamics of the case breach cross-sectional area $A_h$. Note that the same model can be used to model other important fault modes in SRM. For example, the bore choking fault in the $i^{th}$ ballistics element can be modeled by introducing fault induced changes to the port area $A_p$ in this element; the crack dynamics can be modeled by introducing crack induced changes to an effective port perimeter $l(x)$ in the $i^{th}$ ballistics element; the nozzle blocking can be modeled by introducing fault induced changes to the nozzle throat area $A_t$ in the boundary conditions (38) below. The boundary conditions at the aft end (at the outlet of the grain) are defined by the choking (sonic) conditions at the nozzle throat. The boundary conditions at the rocket head are determined by the continuity conditions of the gas flow from the propellant surface and through the port area at the rocket head.

By adding to these two conditions the equation of state and the equation for the gas temperature in the combustion chamber as a function of the Mach number $M_0$ we obtain resulting boundary conditions at the rocket head (0) and aft ($A$) ends in dimensionless units as follows

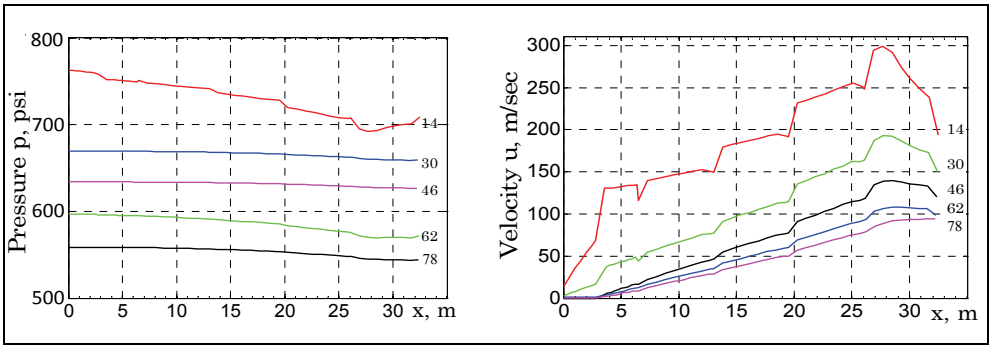



Fig. 12. Nominal regime: Results of numerical solution of Eqs. (37), (38) for axial distributions of pressure (left) and velocity (right) at different moments of time. Time after ignition: 14, 30, 46, 62, and 78 seconds. The value of $x$ is measured from the motor head.

$$M_{0,0}=\left(\frac{\varepsilon\rho_p A_{b,0}}{A_{p,0}}\right)\left(1-\frac{\gamma-1}{2}M_{0,0}^2\right)p_0^{n-1},\ M_{0,A}\left(1-\frac{\gamma-1}{2}M_{0,A}^2\right)^{\frac{1}{\gamma-1}}=\frac{A_t}{\Gamma A_A},\ p_A=\rho_A\left(1-\frac{\gamma-1}{2}M_{0,A}^2\right). \quad (38)$$

The results of the numerical solution of the problem (37), (38) for nominal regime ($A_h$=*0*) are presented in Fig. 12. This figure shows the resulting axial distributions of the pressure and velocity for five instances of time with the time step 16 sec (the time resolution of the solution was 0.2 sec). It can be seen from Fig. 12 that there is a substantial difference between the head and aft pressure due to the effect of mass addition. The difference is most significant at the initial time when the port area is the smallest and the flow velocity has the largest values along the axis. With time the port area is increasing and the difference between head and aft pressure becomes negligible. Our analysis showed that results presented in Fig. 12 coincide with those obtained by the 3rd party using high-fidelity simulation of the internal ballistics of SRM.

### 5.2 Diagnostic of the fault parameters

To be able to reconstruct fault parameters first we have to introduce a parameterization of the fault. It can be seen from the model (37) that the fault dynamics is described entirely by the dynamics of the area of the hole $A_h(t)$. The actual dynamics of the fault area can be complicated due to e.g. cracks and nontrivial geometry of the joints (see e.g. (McMillin, 2006; Smelyanskiy at al., 2008)). However, analysis of the ground test results (Smelyanskiy at al., 2008) and of the challenger accident (McMillin, 2006) shows that the case breach dynamics is sufficiently smooth, primarily determined by the burning of the metal walls of the hole in the rocket case, and can be parameterized in the form:

$$A_h(t)=a_1+a_2t+a_3t^2+a_4t^3. \quad (39)$$

This parameterization has proved to be useful in the analysis of the ground firing test (Smelyanskiy at al., 2008). The parameters of the fault dynamics $\{a_1, a_2, a_3, a_4\}$ are reconstructed from the inferred time-series data $A_h(t)$ using the least square method. The hole is most likely to be localized at one of the section joints as shown schematically in Fig. 2. As a rule, only

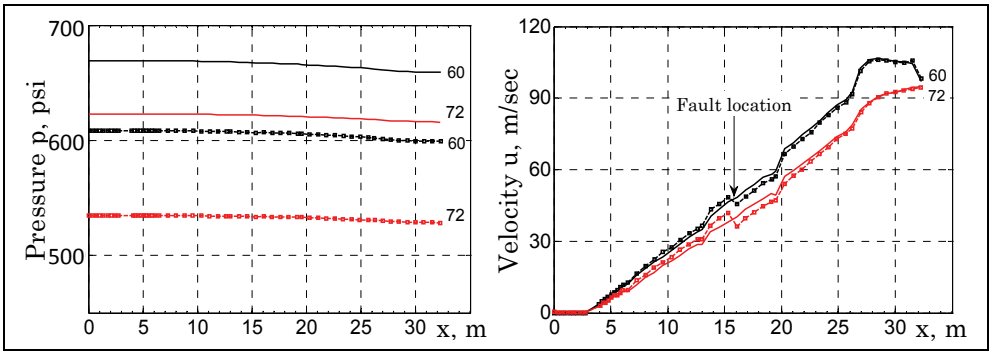



Fig. 13. (left) Comparison between spatial distribution of pressure in the nominal regime (solid lines) and off-nominal regime (squares). (right) Comparison between spatial distribution of velocity in the nominal regime (solid lines) and off-nominal regime (squares). The time instants from the top to the bottom in the figure are 60 sec and 72 sec. The time resolution of the calculations was 0.2 sec, initial radius of the hole $R_{h0}$ = 0.1 in, burning rate of the hole wall $v_m$ = 0.3 in/sec, initial time of the fault 20 sec, the fault is located in the middle section.

pressure sensor situated in the rocket head is available on-board. Therefore, we have to verify that the measurements of the head pressure can be used to infer pressure at an arbitrary location of the hole along the rocket axis. To do so we simulate the model of internal ballistics of the SRM (37), (38) in the off-nominal regime with the case breach area dynamics given by (39) at arbitrary location. The results of such simulations for the case breach at the middle of the SRM are shown in the Fig. 13. It can be seen from the figure that the pressure drop induced by the case breach is uniform along the rocket axis. This shift does not depend on the location of the burning-through hole in the case. In particular, this result allows one to determine the changes in the aft pressure from the measurements of the head pressure.

This finding allows us to use the following quasi-stationary solution for the nozzle stagnation pressure $p_{ns}$, which is hold with good accuracy for large SRMs (Salita, 1989; McMillin, 2006):

$$p_{ns} = p_c \left[ \frac{\Gamma c_0 \rho_p r_c}{\gamma p_c} \left( \frac{A_{b,eff}(t)}{A_t(t) + A_h(t)} \right) \right]^{\frac{1}{1-n}}, \tag{40}$$

We notice that burning area $A_b(R(t))$ and nozzle throat area $A_t(t)$ are determined by measured gas pressure $p_{ns}$ (see Eqs. (9) and (11)). The accuracy of the relation (40) is further improved by introducing the effective burning area $A_{b,eff}(R(t))$ in the nominal regime. Indeed, in the nominal regime $A_h(t) = 0$ while $p_{ns}(t)$ and $A_t(t)$ are well known that allows us to determine uniquely the effective burning area $A_{b,eff}(R(t))$ as a function of the burn distance $R(t)$. This functional dependence is assumed to be valid in the off-nominal regime of the case breach. Therefore, one can use data of the pressure sensor at the rocket head to estimate the deviations of the nozzle stagnation pressure $p_{ns}$ from the nominal regime and subsequently to use equation (40) to estimate the area of the case breach fault $A_h(t)$ according to the following algorithm:

1. Use the nominal regime time-traces to determine the effective burning area by Eq. (40)

$$A_{b,eff}(t) = \left[\frac{\gamma A_t(t)}{\Gamma c_0 \rho_p a}\right] p_{ns}^{1-n}(t);$$

2. Use measured time-trace of the head pressure in the off-nominal regime $p_H(t)$ to find fault-induced pressure at the aft end using the fact that the pressure changes induced by the fault are uniform along the motor axis

$$p_A^{(fault)}(t) = p_A^{(nom)}(t) + \left(p_H^{(fault)}(t) - p_H^{(nom)}(t)\right);$$

3. Use nominal time-trace of the Mach number at the aft end to determine nozzle stagnation pressure

$$p_{ns}^{(fault)}(t) = p_A^{(fault)}(t)\left(1 - \frac{\gamma - 1}{2} M_{A,0}^2(t)\right)^{-1};$$

4. Use Eq. (40) to determine time-trace of the hole area

$$A_h(t) = \left[\frac{\Gamma c_0 \rho_p a A_{b,eff}(t)}{\gamma\left(p_{ns}^{(fault)}(t)\right)^{1-n}}\right] - A_t(t).$$

The parameters of the fault dynamics $\{a_1, a_2, a_3, a_4\}$ are reconstructed from the inferred time-series data $A_h(t)$ using the least square method. We can now use the values of the parameters $\{a_i\}$ reconstructed during the diagnostic to predict fault and internal ballistics of the SRM forward in time.

### 5.3 Prognostics of the fault parameters

We note that the values of the reconstructed parameters $a_i$ of Eq. (39) depend on the diagnostics time. Therefore, the convergence of the forward predictions also depends on the diagnostic time, which is one of the key characteristics of the FD&P system. The convergence of the predicted hole area time-traces towards actual time-traces of $A_h(t)$ is illustrated in Fig. 14. In this test the hole area measurements are sampled with sampling rate 1kHz and measurement noise 0.1%. The filtering procedure is used to reduce the noise in the data. The time intervals $\Delta T_m$ used to infer fault parameters are 8 sec and 12 sec. The area of the fault is reconstructed using algorithm described in the Sec. 5.2 with the diagnostic time window 8 and 12 sec. The fault initial time is $t$ = 40 sec. The hole radius growth rate is 0.3 in/sec. The fault parameters $\{a_i\}$ (see Eq.(39)) are inferred using e.g. noise-reduction and least-square procedures. The dynamics of the fault is predicted ahead in time up to 80 sec of the flight using inferred parameters $\{a_i\}$ and Eq.(39). Note that the convergence of the predictions of the hole area is achieved approximately after 11 sec of diagnostics.

The mean values and standard deviations of the parameters $\{a_i\}$ reconstructed during diagnostics can now be used to integrate model of internal ballistics forward in time to obtain predictions of the pressure and thrust dynamics in the presence of the fault. The

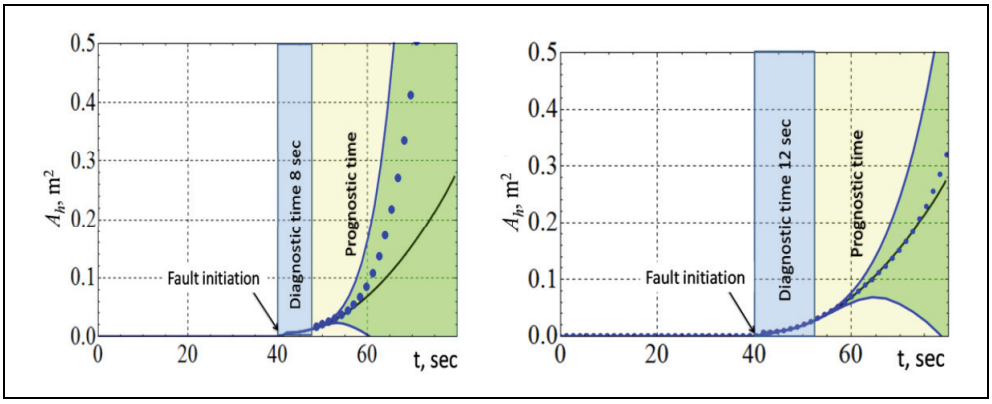



Fig. 14. Prognostic of the case breach area. The actual time series of the fault area $A_h(t)$ are shown by black lines. The mean values of predicted dynamics of the $A_h(t)$ are shown by blue dotted lines for two different values of the diagnostic time-window: (i) 8 sec (left); (ii) 12 sec (right). The green shaded areas bounded by blue lines indicate standard deviation of the predictions. The blue shading indicates the time used for diagnostic of the fault parameters. The yellow shaded region indicates the prediction time.

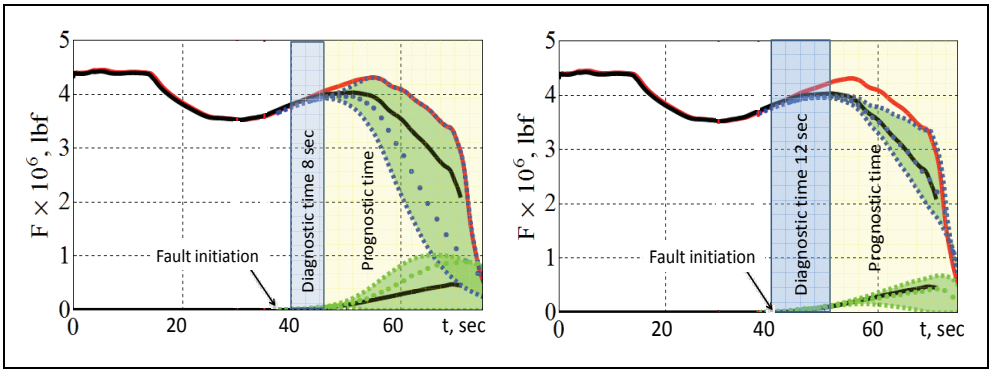



Fig. 15. The thrust in the nominal regime is shown by solid red line. The actual fault-induced time-traces of thrust are shown by black lines. The time-traces of mean predicted values of the thrust are shown by blue dots. The predicted values of the nozzle thrust and the side thrust are indicated by green shading bounded by blue dotted lines and green dotted lines respectively. In the figures the prediction is made after $\Delta T_m$=8 sec (left) and $\Delta T_m$=12 sec (right).

results of the predictions for the nozzle and side thrust are shown in the Fig. 15. In this test the fault was located at the middle of the motor and initial time $T_f$ = 40 sec. The hole radius growth rate is 0.3 in/sec. We take the initial radius of the hole equal to 0.1in. It can be seen that convergence is achieved after approximately 12 sec of diagnostics time.

We conclude that the model (34) allows for accurate calculations of the internal ballistics of SRMs in nominal and off-nominal regimes. This model can further be applied for development of the FD&P system of the case breach fault as discussed in this and previous

sections. We also note that many SRMs fault modes including e.g. nozzle blocking, bore choking, and propellant cracking can be related to the changes in the grain and/or nozzle geometry at a particular section of the SRM. Therefore these changes can be conveniently modeled using (34) and by introducing corresponding modifications into the grain geometry equations (8) and (9) in the form of additional time-dependent fault terms

$$A_p(x_i) = f_{Ai}(R(x_i)) + A_{p,fault}(t), \quad l(x_i) = f_{li}(R(x_i)) + l_{fault}(t). \tag{41}$$

The dynamics of these fault modes can be based on the corresponding failure physics (cf. with the model of the case breach in Sec. 2.4) or modeled by parametric equation similar to Eq. (39). The corresponding FD&P systems for the propellant cracking and nozzle blocking faults can be based on the algorithms similar to one discussed in this chapter.
We now briefly outline the procedures of V&V used to verify and validate the results obtained for the case breach fault.

### 5.4 Validation and verification of the FD&P system

The practical application of the FD&P algorithms requires extended V&V procedure that may take 80% and more of the total development time. The V&V of the case breach and nozzle blocking models was performed in a number of stages. The results of simulations for large motor were compared with the results of high-fidelity simulations obtained using direct integration of the model equations (Smelyanskiy et al., 2006). For a sub-scale motor the results of the simulations were verified using high-fidelity FLUENT non-stationary model (Osipov et al., 2007; Luchinsky et al., 2008) as is described in Sec. 3.3. Next the FD&P system was validated using analysis of the results of the ground firing test of a subscale motor (Smelyanskiy, 2008a & 2008b). In this section we briefly discuss verification of the FD&P system for a large segmented motor by comparison the results of prediction of the fault dynamics obtained using model (34) with the results of calculations of the fault dynamics using scaling algorithm (McMillin, 2006). This scaling algorithm (McMillin, 2006), unlike FD&P algorithm presented above, requires a knowledge of the time-trace of the pressure in the nominal regime. Since this requirement is a necessary prerequisite for the space flight one can substantially reduce the prediction time. Let us remind briefly main steps of the scaling algorithm (McMillin, 2006). The key assumption based on extensive experimental results is that the empirical coefficient $q$

$$q = \frac{\partial \ln(C^*)}{\partial \ln(p)} = const, \qquad C^* = \frac{p^{1-n} A_t}{A_b \rho_p a}. \tag{42}$$

is constant. Here $C^* = pA_t / \dot{m}$ is characteristic exhaust velocity and $\dot{m} = A_b \rho_p \dot{R} = A_b \rho_p a p^n$ is the mass flow with total burning area $A_b$, where the burning law is given by $\dot{R} = ap^n$. To see more clearly the meaning of (42) we write its solution on a discrete time lattice $\{t_i: i=1,\ldots,N\}$ chosen in such a way that every moment of time in the nominal $t_{Ni}$ and off-nominal $t_{Di}$ regimes corresponds the same fixed lattice of burned distances $\{r_i: i = 1,\ldots,N\}$ and therefore the burning area $A_{b,i}$ at any given instant of time is also the same

$$C_i^* p_i^{-q} = \frac{p_i^{1-n-q} A_{t,i}}{A_{b,i} \rho_p a} = const \tag{43}$$

The experiments and numerical simulation show that the later expression is constant with good accuracy during the steady burn out of the propellant both in nominal ($N$) and off-nominal, deviant (D), regimes. We notice that the mean value of $q$ is approximately 0.007, i.e. 2 orders of magnitude less than (1-n) = 0.7 and can be neglected in the calculations.
Using Eq. (43) for the nominal $p_{Ni}$ and off-nominal $p_{Di}$ pressures at the instant $t_i$ on a given time lattice $\{t_i\}$ we obtain

$$p_{Di} = p_{Ni} \left( A_{t,Ni} / A_{t,Di} \right)^{\frac{1}{1-n-q}}, \tag{44}$$

where $A_{t,Ni}$ and $A_{t,Di}$ are nozzle throat area in the nominal regime and the nozzle effective area in the off-nominal regime at the instant $t_i$ on the nominal time lattice $\{t_i\}$ and instant $t_{Di}$ on the deviant time lattice $\{t_{Di}\}$. It was noticed that if the lattice of burned distances $\{r_i\}$ is kept the same in nominal and off-nominal regimes the corresponding time lattices $\{t_i\}$ in these two regimes will be different. To find the scaling of the time lattice the burning law is used as follows (McMillin, 2006)

$$\Delta r_i = r_i - r_{i-1} = a p_{Ni}^n \Delta t_{Ni} = a p_{Di}^n \Delta t_{Di}$$

or

$$\Delta t_{Di} = \Delta t_{Ni} \left( p_{Ni} / p_{Di} \right)^n. \tag{45}$$

Therefore, the time scale in the presence of the fault is

$$t_{D,i} = t_{D,i-1} + \Delta t_{Di}. \tag{46}$$

Finally, substituting (44) into (45) we have

$$\Delta t_{Di} = \Delta t_{Ni} \left( A_{t,Ni} / A_{t,Di}^e \right)^{\frac{-n}{1-n-q}}. \tag{47}$$

Fault dynamics can be given in any form, in particular it is a given function of time (McMillin, 2006)

$$A_{h,i} = f(t_{D,i} - t_f). \tag{48}$$

The effective nozzle throat area for the case breach fault $A^e_{t,D}$ is the sum of the hole area $A_h$ in the case and of the slightly modified nozzle throat area $A_{t,D}$

$$A_{tD,i}^e = A_{h,i} + A_{tD,i}. \tag{49}$$

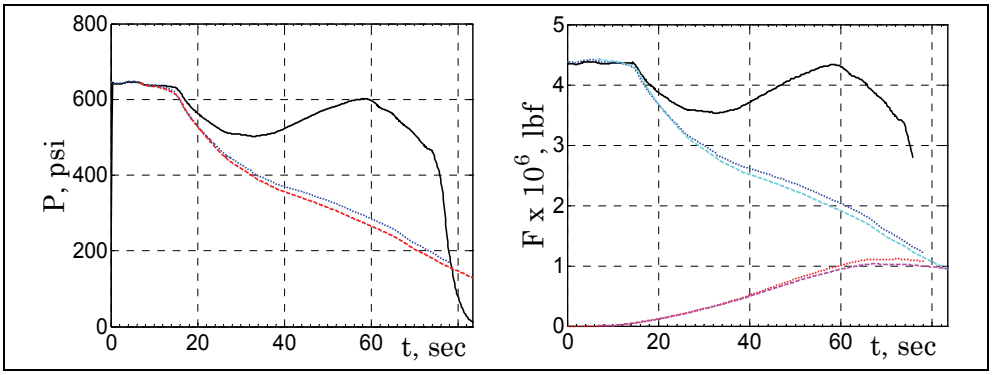



Fig. 16. (left) Nozzle pressure (black line) in the nominal regime is compared with the calculations of the fault induced pressure using 1D model (blue dotted line) and scaling algorithm (red dashed line). (right) Nozzle vacuum thrust (black line) in the nominal regime as compared to the deviation of the thrust from nominal regime calculated using scaling algorithm (cyan dashed line) and 1D calculations (blue dotted line). The time-traces of the hole' thrust calculated using scaling equations and 1D algorithm are shown by magenta dashed line and red dotted line respectively.

To calculate nozzle thrust the empirical constant thrust coefficient was used

$$C_F = F_N / pA_{t,N}, \qquad C_{F,N} = C_{F,D}. \tag{50}$$

If combined with (44) the following result for the deviant nozzle thrust $F_{Di}$ can be obtained

$$F_{Di} = F_{Ni} = \left(\frac{A_{t,Di}}{A_{t,Ni}}\right)\left(\frac{A_{t,Ni}}{A^e_{t,Di}}\right)^{\frac{1}{1-n-q}}. \tag{51}$$

Eq. (51)can be used to calculate the side breach thrust. The resulting expression for the side vacuum thrust ($p_{amb}$ = 0) can be simplified

$$F_h = p_0 A_h \left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}}(\gamma+1). \tag{52}$$

The algorithm suggested in earlier work (McMillin, 2006) consists of iteration of Eqs. (46)-(49) until convergence of $\Delta t_i$ is achieved. On substituting Eq. (48) into Eq. (49) and Eq. (49) into Eq. (47) the equations. (46)-(49) can be reduced to two equations in the form

$$\Delta t_{Di} = \Delta t_{Ni}\left(\frac{A_{t,i}}{A_{tD,i} + f(t_{D,i} - t_f)}\right)^{\frac{-n}{1-n-q}}, \quad t_{D,i} = t_{D,i-1} + \Delta t_{Di}. \tag{53}$$

where $t_{Di}$ is given by (46). There are three unknown variables $\{\Delta t_{Di}, A_{tDi}, t_{Di}\}$ in two Eqs. (53), therefore one more equation is needed to implement iterative procedure. We use assumption (McMillin, 2006) that the nozzle throat area is only a function of time $A_{tD}(t_{Di}) = A_{tN}(t_{Di})$, which holds for actual rocket parameters.

We now verify both algorithms by direct comparison of their performance in off-nominal regime with the fault initial time 10 sec and initial hole radius 0 in. The results of the calculation of pressure, nozzle and side thrusts using model integration and scaling equations are shown in the Fig. 16. It is clear from the figure that the numerical integration of the model and the results of predictions based on the scaling equations are in good agreement. The maximum relative deviation of the scaling algorithm results from the results of simulations is less than 7%. Note that the safety margins do not allow for the deviation of the nozzle thrust more than 10% of the nominal regime. Within these safety margins the agreement between the predictions based on the scaling algorithm results and on simulations is better than 1%. The reason for the small deviation of the scaling algorithm from the results of integration of the model is that the scaling algorithm overestimates nozzle throat ablation, while chamber pressure and nozzle thrust are slightly underestimated. The nozzle ablation modeling in the scaling algorithm will be improved in the next version of the code.

### 5.5 Application to the calculation of the flight safety envelopes

In this section we describe one of the applications of the case breach model in large segmented SRMs. Objective of this application is to study response of the TVC of a large segmented motor to side thrust torque and compute at which moment the vehicle will leave the flight safety window. We introduce the flight safety window in the following way: altitude difference between the flight with the side thrust and the nominal flight is not larger than 4000 ft; the vehicle's angle of attack for the flight with the side thrust is in +20/-20 degrees interval. We assume that the flight is not safe if the rocket has altitude difference larger than altitude difference margin (4000 ft) or if the angle of attack of the rocket is larger than 20 degrees in absolute value. We assume that the side thrust appears at forward and aft field joints at 0, 90, 180 and 270 degrees angle location at times from 0 to 120 seconds with 5 second interval. We use the model (34) to calculate nozzle and side thrusts, and the gas flow rate in the combustion chamber. The results of these calculations were fed into special purpose software to calculate rocket trajectory taking into account TVC. The simulations were performed about 200 times for different breach locations and different times. The breach radius growth rate for these simulations was taken 0.3 in/sec. The flight time safety envelopes were created by plotting nominal flight times against the side thrust start times for given joint position and breach location angle.

The side thrust safety envelopes were created by plotting the side thrust profiles against the side thrust start times for given joint position and breach location angle. The regions of failure were determined by plotting points of failure for each flight trajectory. It was found that for 0, 180 degrees breach location the rocket loses control due to rapid change of the altitude, for 90, 270 degrees breach location the rocket loses control mostly due to oscillations of the angle of attack. In Fig. 17 we present the flight time safety envelopes for forward and aft field joints at 0 degree angle location. The areas under the black curves restrict regions of nominal vehicle flights, while the areas outside these curves correspond to the faulty flights.

The side thrust safety envelopes describing values of side thrusts for forward and aft field joints with 0 and 90 degrees of breach locations are shown in Fig. 18. The results for 180 and 270 degrees breach locations are similar to the results for 0, 90 degrees.

The blue areas in these plots correspond to side thrusts for the nominal flights and the red areas correspond to the faulty flights. As follows from these plots, the areas of blue regions of the envelopes are more sensitive to joint location for 0 degree breach location, but still the same for 90 degrees breach locations.

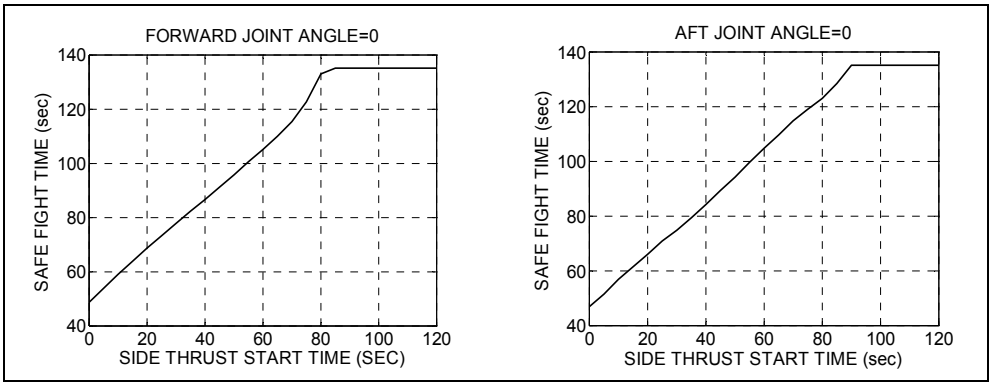



Fig. 17. The left plot corresponds to the flight time safety envelope for forward joint and the right plot corresponds to the flight time safety envelope for the aft joint at 0 deg breach location.

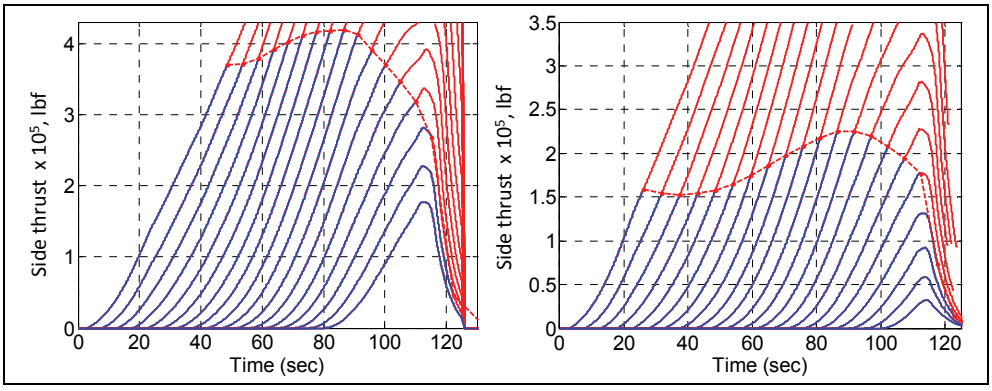



Fig. 18. (left) Side thrust values for 2nd joint, 0 deg breach location. (right) Side thrust values for 5th joint , 90 deg breach location. X-axis is start time of the side thrust. The hole radius growth rate was 0.3 in/sec.

## 6. Conclusion

In this chapter we report the work in progress on the development of hybrid probabilistic algorithms for SRMs on-board IVHM system. We introduce model that describe accurately internal ballistic of SRM in nominal and off-nominal regime. The model is verified by comparison with the results of high-fidelity simulations. A number of algorithms are derived to infer fault parameters and to predict fault dynamics. The algorithms

accommodate abrupt changes in the model parameters and can be used to develop hybrid probabilistic on-board SRM IVHM. The performance of the algorithms was tested using analysis of the experimental time-series data. It is shown that the algorithm can be successfully applied for the prognostics of the case breach fault. The model is used to develop on-board FD&P of the case breach fault and applied to predict fault-induced internal ballistics of a large segmented SRM.

The developed methods and algorithms can be used to analyze other SRM' faults, including overpressure and breakage of the case induced by nozzle blocking, bore choking and grain deformation. The bore choking phenomenon is an almost radial deformation (bulge) of the propellant near booster joint segments. This phenomenon can cause choking of the exhaust gas flow and increase the burning surface which can lead to critical overpressure in the combustion chamber. Development testing has shown that this fault was observed, for example, in primary construction of the Titan IV (Coats, 2003; Wilson, 1992). The bore choking, and also cracks and voids in the solid propellant, can result in local burning of the booster case and also in abrupt breaking off of large pieces of the propellant. These pieces can stick in the nozzle throat and block the exhaust gas flow (nozzle blocking fault). In all these cases the fault dynamics is governed by the changes in the burning area and/or effective nozzle area. Therefore the model introduced in this work can be efficiently applied to the analysis of these faults and development of the on-board FD&P.

## 7. Nomenclature

| | | |
|---|---|---|
| $\rho$ | = | gas density |
| $p$ | = | gas pressure |
| $T$ | = | gas temperature |
| $u$ | = | gas velocity |
| $\rho_p$ | = | density of the solid propellant |
| $H$ | = | combustion heat of the solid propellant |
| $e_T$ | = | total energy of the gas |
| $h_T$ | = | total enthalpy of the gas |
| $c_p$ | = | specific heat for the constant pressure |
| $c_V$ | = | specific heat for the constant volume |
| $\gamma$ | = | ration of specific heats $\gamma = c_P/c_V$ |
| $A_p$ | = | port area (cross-sectional area of combustion chamber) |
| $A_h$ | = | hole port area |
| $A_b$ | = | burning area of the propellant |
| $r_b$ | = | burning rate of solid propellant |
| $R$ | = | thickness of the burned propellant layer |
| $\dot{r}_{er}$ | = | erosive burning rate of solid propellant |
| $T_{abl}$ | = | critical temperature of the nozzle ablation |
| $q_{ins}$ | = | latent heat of insulator ablation |
| $c$ | = | sound velocity |
| $M$ | = | Mach number, $M = u/c$ |
| $F_N$ | = | normal thrust |

| | | |
|---|---|---|
| $F_h$ | = | additional thrust produced by hole gas flow |
| $Q$ | = | heat flow from the gas to the walls of the hole |
| $Q_c$ | = | heat flow due to convection |
| $Q_R$ | = | heat flow due to radiation |
| $Q_b$ | = | heat flow due to burning |
| $v_{t0}$ | = | velocity of the ablation of the nozzle throat |
| $c_{ins}$ | = | heat capacity of the insulator |
| $q_{met}$ | = | specific melting heat of the metal case |
| $\rho_{met}$ | = | metal density |
| $T_{met}$ | = | metal temperature |
| $C_{met}$ | = | heat capacity of the metal case |
| $\lambda$ | = | the emissivity of the hot gas |
| $v_{fb}$ | = | velocity of the burning front propagation |
| $V$ | = | volume of the combustion chamber |
| $in$ | = | subscript for parameters of the insulator |
| $ex$ | = | subscript for parameters at the nozzle exit |
| $h$ | = | subscript for gas parameters in the hole |
| $t$ | = | subscript for gas parameters in the nozzle and hole throat |
| $N$ | = | subscript for parameters in normal regime |
| $0$ | = | subscript for stagnation values of gas parameters |

## 8. References


Bartz, D.R. (1965). Heat Transfer from Rapidly and from Heated Air, in Advances in Heat Transfer, vol. 2, Hartnett, J. P., and Irvine, T. F. Jr., eds., New York: Academic Press.

Dick, W. A. et al. (2005, July). Advanced Simulation of Solid Propellant Rockets from First Principles", Center for Simulation of Advanced Rockets, University of Illinois, in Proceeding of 41st AIAA Joint Propulsion Conference & exhibit, Tucson, Arizona.

Congdon, P. (2001). "Bayesian statistical modelling", Wiley series in probability and statistics (Wiley, Chichester).

Coats, D.E.; Dunn, S.S. & French, J.C. (2003) "Performance Modeling Requirements for Solid Propellant Rocket Motors", 39th JANNAF Combustion Subcommittee, Colorado Springs, CO, December.

Culick, F. E. C. (1974). "Stability of Three-Dimentional Motions in a Rocket Motor", Combustion and Technology, v. 10, 109.

Culick, F. E. C. & Yang, V. (1992). Prediction of the Stability of Unsteady Motions in Solid Propellant Rocket Motors, Nonsteady Burning and Combustion Stability of Solid Propellants, edited by L. De Luca, E. W. Price, and M. Summerfield, Vol. 143, Progress in Astronautics and Aeronautics, AIAA, Washington, DC, pp. 719–779.

Culick, F. E. C. (1996). Combustion of the Stability in Propulsion Systems, Unsteady Combustion, Kluwer Academic Publisher.

Dykman, M. I. (1990). "Large fluctuations and fluctuational transitions in systems driven by colored Gaussian noise -- a high frequency noise", Phys. Rev. A, Vol. 42, 2020—2029.

Graham, R. (1973). paper in Quantum Statistics in Optics and Solid-State Physics, edited by G. Höhler, Springer-Verlag, Berlin, vol. 66, Chap. 1, p. 1.

Graham, R. (1977). "Path Integral Formulation Of General Diffusion Processes", Z. Phys. B, Vol. 26, 281-290.

Flandro, G. A. et al. (2004). "Nonlinear Rocket Motor Stability Prediction: Limit Amplitude, Triggering, and Mean Pressure Shift", in Proceeding of AIAA 2004-4054, 40th AIAA Joint Propulsion Conference & Exhibit, July, 2004, Florida.

Flandro, G. A. (1982). Approximate Analysis of Nonlinear Instability with Shock Waves, AIAA-82-120, 18th AIAA Joint Propulsion Conference, June, 1982, Ohio.

Knoll, D.A.; Chacon, L.; Margolin, L.G. & Mousseau, V.A. (2003). On balanced approximations for time integration of multiple time scale systems", J. Comput. Phys., 185(2), pp 583-611.

Hill, P. & Peterson, C. (1992). Mechanics and Thermodynamics of Propulsion, 2-rd ed., Addison-Wesley Publishing Company, Inc. New York.

Handbook of Heat Transfer Application. (1973), 2-rd ed,

Jeffreys, H. (1961). Theory of Probability, 3-rd ed. (Clarendon Press, Oxford).

Incropera, F.P. & DeWitt, D. P. (2002). Introduction to Heat Transfer, John Wiley & Sons, NY.

Isaac, D. A. & Iverson, M.P. (2003), "Automated Fluid-Structure Interaction Analysis", ATK Thiokol Propulsion, A Division of ATK Aerospace Company. L. D. Landay and E. M. Lifshits. (1987), Fluid Mechanics (Pergamon).

Luchinsky, D.G.; Millonas, M.M.; Smelyanskiy, V.N.; et al. (2005). "Nonlinear statistical modeling and model discovery for cardiorespiratory data". Phys. Rev. E, Vol. 72, 021905.

Luchinsky, D.G.; Smelyanskiy, V.N.; Osipov, V.V.; Timucin, D. A. & Lee, S. (2007). "Data management and decision support for the in-flight SRM", in Proceeding of AIAA-2007-2829 AIAA Infotech@Aerospace 2007 Conference and Exhibit, Rohnert Park, , May, California.

Luchinsky, D.G.; Osipov, V.V.; Smelyanskiy, V.N.; Timucin, D.A. & Uckun, S. (2008), "Model based IVHM system for the solid rocket booster " in Proceeding of Aerospace Conference, 2008 IEEE, 1-8, March, BigSky, Montana, Page(s):1 – 15.

Luchinsky, D.G.; Osipov, V.V.; Smelyanskiy, V.N.; Timucin, D.A.; Uckun, S.; Hayashida, B.; Watson, M.; McMillin, J.; Shook, D.; Johnson, M. & Hyde, S. (2009), "Fault Diagnostics and Prognostics for Large Segmented SRMs", in Proceeding of 2009 IEEE Aerospace Conference, Big Sky, March, Montana.

McMillin, J.E. (2006, July). AIAA 2006-5121, in Proceeding of 42nd AIAA/ASME/ SAE/ASEE Joint Propulsion Conference & Exhibit , Sacramento, CA.

Oberg, J. (2005). ''Deadly space lessons go unheeded'', NBC News, http://www.mission51l.com/disaster.htm

Osipov, V. V. ; Luchinsky, D. G.; Smelyanskiy, V.N.; Lee, S.; Kiris, C. & Timucin, D. (2007, March), "Bayesian Framework for In-Flight SRM Data Management and Decision Support", in Proceeding of IEEE 2007 Aerospace Conference, Big Sky.

Osipov, V.V.; Luchinsky, D.G.; Smelyanskiy, V.N.; Kiris, C.; Timucin, D.A. & Lee. S.H. (2007, July), In-Flight Failure Decision and Prognostic for the Solid Rocket Buster, in

Proceeding of AIAA-2007-5823, 43rd AIAA/ASME/SAE/ ASEE Joint Propulsion Conference and Exhibit, Cincinnati, OH, July.

Rhsenow, W. M.; Hartnett, J. P. & Ganic, E. N. (1973), McGraw-Hill Book Company, New York

Rogers (1986). Rogers Commission report. Report of the Presidential Commission on the Space Shuttle Challenger Accident.

Risfic, B.; Arulampalam, S. & Gordon, N. (2004), "Beyond the kalman filter - Book Review ", Aerospace and Electronic Systems Magazine, IEEE, Vol. 19, No. 7, 37- 38.

Sailta, M. (1989, January). "Verification of Spatial and Temporal Pressure Distribution in Segmented Solid Rocket Motors," in Proceeding of AIAA paper 89-0298, 27th Aerospace Science Meeting, Reno, Nevada.

Salita, M. (2001). "Modern SRM ignition transient modeling. I - Introduction and physical models", in Proceeding of AIAA-2001-3443, AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, 37th, Salt Lake City, U.

Salita, M. (1989, July). "Closed-Form Analytical Solutions for Fluid Mechanical, Thermochemical, and Thermal Processes in Solid Rocket Motors", in Proceeding of AIAA 98-3965, 34th Joint Propulsion Conference, Cleveland.

Santiago, J.C. (1995), "An experimental study of the velocity field of a transverse jet injected into a supersonic crossflow", Ph.D. thesis, University of Illinois, Urbana-Champaign.

Shapiro, A.H. (1953.), "The Dynamics and Thermodynamics of Compressible Fluid Flow", Ronald Press, NY, vol. I.

Sorkin, E. (1967), Dynamics and Thermodynamics of Solid-Propellant Rockets, Wiener Bindery Ltd., Jerusalem.

Smelyanskiy, V. N., D. G. Luchinsky, D. A. Timuçin, and A. Bandrivskyy. (2005a), "Reconstruction of stochastic nonlinear dynamical models from trajectory measurements", Phys.l Rev. E, Vol. 72, 026202.

Smelyanskiy, V. N.; Kiris, C.; Osipov, V. V.; Luchinsky, D. G.; Timucin, D. A.; & Uckun, S. (2006). "Solid Rocket Motor Health Management for Safe Space Flight", in Proceeding of AIAA Guidance, Navigation, and Control Conference and Exhibit, Keystone, Colorado. August.

Smelyanskiy, V.N.; Luchinsky, D.G.; Osipov, V.V.; Timucin, D.A.; Uckun, S.; Hayashida, B.; Watson, M.; McMillin, J.; Shook, D.; Johnson, M. & Hyde, S. (2008a), "Analysis of experimental time-traces of the ground firing test of a subscale SRM with gas leak in the forward closure", in Proceeding of 6th Modeling and Simulation/4th Liquid Propulsion/3rd Spacecraft Propulsion Joint Subcommittee Meeting, Orlando, December Florida.

Smelyanskiy, V.N.; Luchinsky, D.G.; Osipov, V.V.; Timucin, D.A.; Uckun, S.; Hayashida, B.; Watson, M.; McMillin, J.; Shook, D.; Johnson, M. & Hyde, S. (2008b) "Fault diagnostic and prognostic system for a gas leak fault in subscale SRM: modelling and verification in a ground firing test", 6th Modeling and Simulation / 4th Liquid Propulsion / 3rd Spacecraft Propulsion Joint Subcommittee Meeting, Orlando, December Florida.

Smelyanskiy, V.N.; Luchinsky, D.G.; Osipov, V.V.; Timuchin, D.A. & Uckun, S. (2008c), "Development of an on-board failure diagnostics and prognostics system for Solid Rocket Booster", in Proceeding of 44rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Hartford, CT.

Hathout, J.P.; Fleifil, M.; Annaswamy, A.M. & Ghoniem, A.F. "Active control of combustion instability using fuel-injection in the presence of time-delays," AIAA J. Propulsion and Power, vol. 18, no. 2, pp. 390-399, 2002.

Stewart, D.S.; Tang, K.C.; Yoo, S.; Brewster, Q. & Kuznetsov, I.R.. (2006). Multiscale Modeling of Solid Rocket Motors: Computational Aerodynamic Methods for Stable Quasi-Steady Burning, Journal of Prop and Power, Vol. 22, no. 6, 1382-1388.

Wilson, W.G.; Anderson, J. M. & Vander Meyden, M. (1992). "Titan IV SRMU PQM-1 Overview", AIAA paper 92-3819, in Proceeding of AIAA/SAE/ASME/ASEE, 28th Joint Propulsion Conference and Exhibi, Nashville, TN.

Young, S.; Bruck, G. & Gowrisankaran, S. (2003). "Modeling of Rocket Motor Ballistics for Functionally Graded Propellants", Proceedings of the 39th JANNAF Combustion Meeting, Colorado Springs, CO, December.

# Part IV

# 15

# Design of Low-cost Telecommunications CubeSat-class Spacecraft

Adnane ADDAIM, Abdelhaq KHERRAS and El Bachir Zantou
*Centre for Space Research and Studies, EMI*
*Morocco*

## 1. Introduction

The CubeSat-class Spacecraft hardware has volume, mass, and power limitations more extreme than in other satellites. In fact, the CubeSat-class spacecraft is a Cube with dimensions 10cm x 10cm x 10cm, mass one kilogram maximum and the available power is about 2Watt by mounting the solar cells on the CubeSat structure.
The CubeSat-class spacecrafts have the advantages of being able to perform as a test bed for new core space technologies to be applied to space programs, for much lower cost, shorter schedule, and less risk. For this reason, the world leaders in space technology, including the US and Europe, are focusing their efforts on smaller satellites under the motto of "Faster, Cheaper, Better" that can perform missions traditionally assigned to large/medium satellites in the past.
In general, the conventional design of a satellite is based on the modular architecture where each satellite subsystem has hardware and software autonomy. In the case of a CubeSat-class satellite, this kind of architecture did not prove its effectiveness since a significant number of CubeSats could not make a success of their missions. The goal of this chapter consists in developing a new design methodology of the CubeSat-class satellites.
Each satellite includes many subsystems; each subsystem has a role to play aboard the satellite. In practice, in the field of the satellites design, each satellite subsystem, which has dedicated hardware and software, is assigned to one team of experts because each satellite subsystem belongs to one branch of science. On the other hand, there are connections between the various subsystems. Therefore, each team works in coordination with other teams to achieve the tasks of its subsystem. Now, the question is: Do we have legitimacy to integrate several subsystems in only one. The answer is simple. For the satellites larger than the Microsatellite, it is not conceivable to integrate several subsystems in only one. On the contrary, to ensure the correct operation of the satellite considering the importance of the mission and the significant budget allocated, we usually proceed to the redundancy of some key equipment aboard the satellite in spite of the complexity that adds.
In the case of CubeSat-class spacecraft, considered as the other extreme of the category of the small satellites, given the imposed constraints (available electrical power is about 2Watt, mass ≤ 1kg and volume of 10x10x10 cm3) and the restricted financing (university), we find all legitimacy to integrate several subsystems in only one. This approach will save energy, will reduce the mass and will save space aboard the satellite. The key of this integration is to

replace some hardware by software. For example, two hardware modems could be replaced by programming one Digital Signal processor (DSP).
In this chapter, we will describe in details the design of Low-cost Telecommunications CubeSat-class Spacecraft by using one DSP processor with multitasking operating system which integrates all the intelligences of the satellite. Due to the environment in space and the constraints of the CubeSat-class spacecraft, different measures had to be taken when the satellite was designed.

## 2. Satellite mission description

The mission of the Telecommunication CubeSat-class satellite is to provide various services, such as mobile localization of ships, message exchange between terminals and data collection from autonomous weather stations in inaccessible.
The Telecommunications satellite will have four modes: Store-and-Forward mode, Digipeater mode, Beacon mode and Command mode. It will use only one communication protocol, the AX.25 protocol, to communicate with the earth segment (terminals and the control station). This protocol provides Connection-Oriented and Connectionless transmissions. The connectionless transmission method will be used in digipeater and beacon modes, whereas the connection-oriented method will be used in store-and-forward and Command modes.
In store-and-forward mode, the terminals use the Slotted Aloha multiple access (Addaim et al., 2008) to send their packets to the satellite and wait for an acknowledgement from the satellite, which stores the correct packets in its on-board storage system, and delivers this to the destination terminal in a later time in the same original format. The packets are position-stamped and time-stamped. Between the storage and the retrieval of the packet, the satellite moves around its orbit and the earth rotates on its axis. These movements change the location of the satellite's footprint, and the satellite effectively carries the packets from terminal to terminal situated in any place all over the world. In this mode, terminals will not use the existing AX.25 firmware. Instead, they will use a stop and wait ARQ (Automatic Repeat Request) protocol which will be implemented only at build-in-house terminals (Zantou & Kherras, 2004). When the terminal transmits its packets, it will wait for acknowledgement and will set its timer. If this timer expires without reception of an acknowledgement, the entire packet is scheduled for retransmission after a random interval time until receiving the acknowledgement.
In the Digipeater mode, which is a real-time bent-pipe relay, all the radio amateur stations, situated all over the world, could transmit their packets without waiting for an acknowledgement. The received packet will be tested if there is an error in the packet. The erroneous packet will be rejected and the correct packet will be retransmitted after adding the satellite's callsign in the Address field of the AX.25 packet. This mode is useful when both the source and destination radio amateur stations are at the same satellite's footprint. In this mode, the satellite could work with the existing APRS satellites or future ones in constellation
In the Beacon mode, the satellite sends periodically the basic telemetry (Temperatures, currents and Battery voltage) to the earth to know the life of the satellite. This telemetry is encoded in an APRS (Automatic Packet Reporting System) Telemetry format in the Information field of AX.25 packet and is identified by the character T as following:

T#SSS,111,222,333,444,555,11111111, where SSS is the sequence number followed by the five 3 digit analog values and the eight binary values. A Control station could gather the satellite APRS telemetry data from the radio amateurs all over the world via Internet. The gathered data will be archived to a database and be accessed by Internet.

In the Command mode, the control station sends telecommand (TC) to the On Board Computer aboard the satellite and receives the housekeeping data telemetry (TM) from it to know the health status of the satellite.

This Telecommunication CubeSat satellite has two communications channels (see Fig. 1), which can be multiplexed over the same transceiver. The first channel carrying the TM/TC data (the application layer) between the satellite and the control station whereas the second channel carrying the payload data (the application layer) between the satellite and the Terminals. On bottom of the application layer, we used the AX.25 protocol (Data Link layer), which is an amateur radio adaptation of the ITU-T X.25 protocol. The AX.25 supports the requirements for amateur satellite communication. The AX.25 packets are sent over the radio links (the Physical layer), which use GMSK (Gaussian Minimum Shift Keying) for the Command mode, and AFSK (Audio Frequency Shift Keying) for the payload modes, over the same 145 MHz (uplink and downlink) VHF band. The use of one frequency for both uplink and downlink keeps the system overall cost very low.

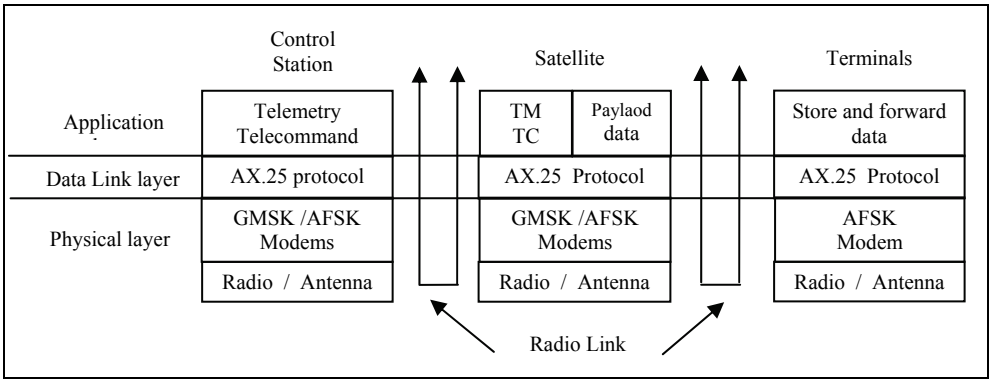



Fig. 1. Satellite Communication Layers

At VHF frequencies, the antennas, receivers and transmitters for both the ground and the space segment, are readily available and inexpensive (Paffet et al., 1998). The Doppler shift, is kept at 3 KHz or below for the satellite parameters (altitude, minimum elevation angle, frequency band) (Jamalipour, 1998), which can be ignored.

In general case, the transmitted signal is corrupted by channel noise, by Doppler and multipath effects so that the bit decoding is not easy. In our case, the Doppler effect is negligible, and the mutipath effect is also negligible using low data rate.

The link budget, shown in Tab. 1, is calculated in the worst case when the satellite, with altitude of 650 Km, is just rising above the horizon with the elevation angle E = 10°.

To enhance the radio link in Command mode, between the control station and the satellite, we added a Forward Error Correction (FEC) by using convolutional coding and Viterbi decoding (Proakis, 1989).

| *Terminal Parameters* | |
|---|---|
| Antenna gain | 0 dBi |
| Transmitted power | 5 W |
| Antenna Feed Loss | 0.5 dB |
| $T_{syst}$ | 2000 k |
| *Control Station Parameters* | |
| Antenna gain | 13 dBi |
| Antenna Feed Loss | 1,5 dB |
| Transmitted power | 5 W |
| $T_{syst}$ | 2000 k |
| *Satellite Parameters* | |
| Antenna gain | 0 dBi |
| Transmitted power | 1 W |
| Antenna Feed Loss | 0.5 dB |
| $T_{syst}$ | 5000 k |
| *Channel Parameters* | |
| Bandwidth | 15000 Hz |
| Free Space Loss | - 141 dB *1 |
| Additional Losses | 3.5 dB |
| $E_b/N_0$ required | 13.6 dB*2 |
| *Uplink Link Margin for AFSK modulation* | |
| $E_b/N_0$ estimated | 21.42 dB*3 |
| Margin | 7.82 dB |
| *Downlink Link Margin for AFSK modulation* | |
| $E_b/N_0$ estimated | 18.42 dB*3 |
| Margin | 4.82 dB |
| *Uplink Link Margin for GMSK modulation* | |
| $E_b/N_0$ estimated | 26.9 dB*3 |
| Margin | 13.3 dB |
| *Downlink Link Margin for GMSK modulation* | |
| $E_b/N_0$ estimated | 23.9 dB*3 |
| Margin | 10.3 dB |

Table 1. Budget Link

*1 Free Space Loss, FSL is calculated by (Maral & Bousquet, 2002):

$$FSL_{db} = 32.45 + 20\log\left(D_{km}\right) + 20\log\left(F_{MHz}\right) \quad (1)$$

where $F_{MHz}$ is the transmit frequency, and $D_{km}$ is the distance or range of the satellite, calculated by:

$$D_{km} = \sqrt{\left(R_T + h\right)^2 - \left(R_T.\cos\left(E\right)\right)^2} - R_T.\sin\left(E\right) \quad (2)$$

*2 Required $E_b/N_o$ gives an operating bit error rate of $2x10^{-5}$ (Proakis, 1989), assuming non-coherent demodulation of FSK.

*3 An estimation of the $E_b/N_o$ is obtained from:

$$\frac{E_b}{N_0} = \frac{P.L_i.G_t.L_s.L_a.G_r}{k.T_s.R} \quad (3)$$

Where:

- P = Transmitted Power
- $L_i$ = Feed Losses
- $G_t$ = Transmit antenna gain
- $L_s$ = Free Space Loss (FSL)
- $L_a$ = Tx Path Losses (Miscellaneous)
- $G_r$ = Receive antenna gain
- K = Boltzmans constant
- $T_s$ = System temperature
- R = Data rate of the system

## 3. Cubesat satellite architecture

### 3.1 Definitions

The satellite consists of two principal parts, namely: the payload and the platform. The payload (camera, repeater …) is defined by the satellite mission (earth observation, telecommunications …). This payload can not function correctly in the space environment without platform. The platform is made up of several subsystems; each one has its own role. The payload must be inside a structural subsystem (frame) for, on the one hand, being used as interface with the vehicle launch and, on the other hand, protecting the satellite from the space environment effects after its placing in orbit. The various electronic equipments on board the satellite need to be supplied by the electric power. Thus, the satellite needs a power subsystem which is responsible for the generation, the storage and the distribution of the electric power. The temperature in the space environment can vary between – 170 °C (in eclipse) and + 100 °C (in full sun). Thus, the satellite needs a thermal control subsystem which ensures that the temperature of each equipment on board remains within the correct operation temperatures interval. The payload components, like the antenna and the camera, must be pointed in a precise earth geographical area. The satellite will thus need an attitude control subsystem which will be responsible of controlling the orientation of the satellite in space (its attitude).

These various satellite subsystems, including the payload, must be controlled remotely. The need for data handling subsystem is then felt. Its role is to receive remote commands from earth control station and to execute them on board the satellite. For that, the data handling subsystem needs some communication equipments as transceiver and antennas.

In sum, the subsystems, which belong to the platform and offer their services to the payload, are (see Fig. 2):

- The structural subsystem (SS) ;
- The power subsystem (PS);
- The thermal control subsystem (TCS);
- The attitude control subsystem (ACS);
- The data handling subsystem (DHS).

Taking into account the miniaturization of the satellites, a new nomenclature appeared. Table 2 gives an idea of these satellite categories according to their weight.

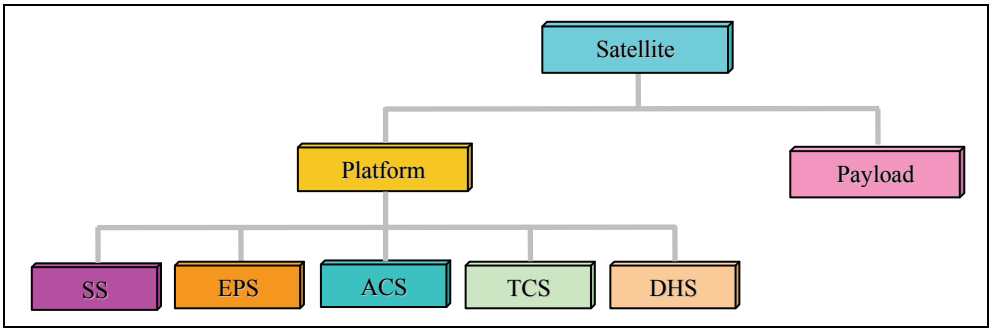



Fig. 2. Various satellite subsystems

| Type of satellite | Weight (Kg) |
|---|---|
| **Minisatellite** | **100 - 1000** |
| **Microsatellite** | **10 - 100** |
| **Nanosatellite** | **1 - 10** |
| **Picosatellite** | **< 1** |

Table 2. Nomenclature of satellites according to their weight

### 3.2 Satellite architecture overview

The Data Handling Subsystem is considered as an interface between the control station and the other satellite subsystems. The Data Handling Subsystem modules include communication subsystem (antennas, modulator, demodulator) and on board computer which controls the satellite subsystems. But how these subsystems are attached to the on board computer, two principal architectures are distinguished.

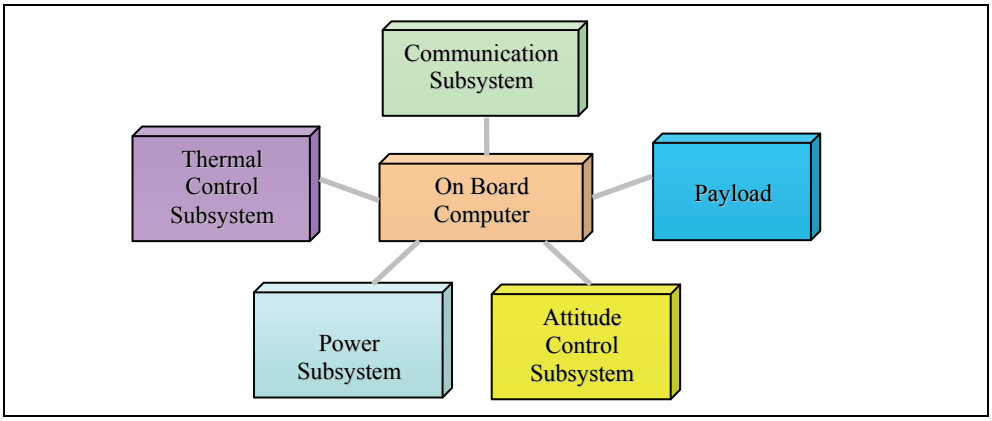



Fig. 3. Satellite Star Architecture

*Star Architecture* (see Fig. 3): the On Board computer (OBC) has to run separate data lines to each subsystem. In this architecture, the OBC should have many peripheral interfaces equal to the number of satellite subsystems on board and perhaps there will be a problem of wiring.

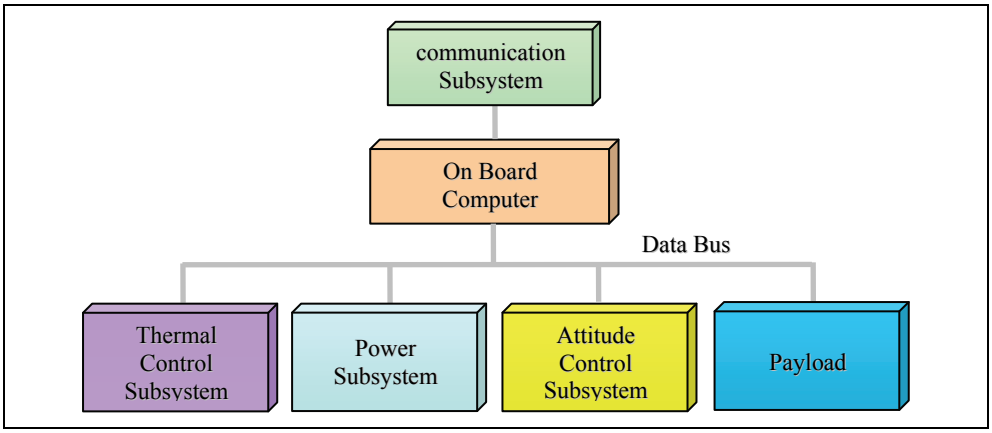



Fig. 4. Satellite Architecture with linear bus

*bus Architecture* (see Fig. 4): in this case, all satellite subsystems are connected to a bus which is similar to Local Area Network (LAN) aboard the satellite. In general, we used a Master-slave protocol. The three simple Master-slave bus commonly used in the embedded system are: CAN (Controller Area Network), SPI (Serial Peripheral Interface) and I²C (Inter Integrated Circuit).

| | CAN | SPI | I²C |
|---|---|---|---|
| Low Power | 3 | 5 | 5 |
| Availability as microcontroller peripheral | 2 | 5 | 4 |
| Slave hardware complexity | 2 | 5 | 3 |
| Slave Software complexity | 2 | 5 | 3 |
| Wiring Complexity | 4 | 3 | 5 |
| Scalability to two Slaves | 5 | 4 | 5 |
| discrete logic interfacing ease | 1 | 4 | 1 |
| Sum | 19 | 31 | 26 |
| Key: 5= best, 1= worst | | | |

Table 3. Master-Slave Buses study in case of two slaves

A trade study, comparing three buses, was conducted as seen in Table 3, in the case of two slaves. Notice that speed was not listed as a determining factor since the buses used similar data rates, all of which were plenty fast enough for CubeSat-class satellite needs. In the case of two slaves, the scalability of SPI bus is almost the same as I²C and CAN buses. According to Table 3, SPI bus won with the high score compared to CAN and I²C buses because of its simplicity. In the case of four slaves, the selected bus will be the I²C bus.
Obviously, the satellite can have a hybrid architecture made up of the two architectures (star and with a distributed linear bus). In this case, for example, payload subsystem is connected directly to the on board computer whereas the others subsystems are connected all by the means of one data bus.

### 3.3 Cubesat concept

The space age began in 1957 with the launch of Sputnik 1 weighing 80 kg. Six months later, the United States had launched an "orange" of 8 kg in space. The Pentagon was aware from the beginning that the dimensions and the weight have a great importance. However, space industry has moved towards larger, more elaborative spacecraft, because this may reduce launch costs and increase the longetivity of space investments. This was one reason for building huge launchers, like the Ariane-5 of the European Space Agency.

The prohibitively high costs of these space projects have restricted first-hand access to space to a handful of nations and international agencies. In the Nineties, research institutions are trying to go against this trend by developing microsatellites. This is mainly possible because state-of-the-art VLSI digital microelectronics enable very sophisticated functions to be achieved without having the constraints of small mass, volume and power.

Traditionally, the design of this type of satellites lasted several years and made their realization impracticable in the university world. Even in the case of realization, it appears difficult to find a launch opportunity. From this point of view, the CubeSat program was proposed in 1999 by professor Robert Twiggs from Stanford University (the USA). The goal of this program is to establish a uniform standard for the design of Nanosatellites and therefore it will increase the number of launches per year. The fundamental specification of this design requires that the satellite must be cubic with 10 cm in length (see Fig. 5) and its weight does not exceed 1 kg. The satellites fit into the CubeSat deployer, the P-POD. This carrier device holds a number of Cubesat what reduces the launch costs. In general, launch of Cubesat uses a modified SS-18 ICBM (Intercontinental Ballistic Missile) Russian or American, at a minimal cost of approximately 45.000 dollars. These days the CubeSat Initiative is a global congregation of universities and private companies from all over the world, trying to advance small satellite technology as tools of education and research.

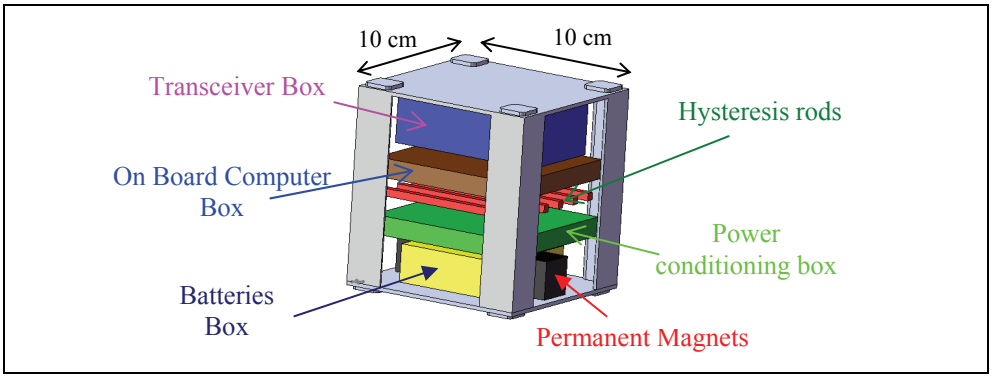



Fig. 5. Internal Arrangement of the Telecommunication CubeSat-class spacecraft

### 3.4 Cubesat subsystems requirements

The requirements of Cubesat-class spacecrafts are significantly different from their ancestors conventional satellites, and new design techniques are needed to meet these evolving requirements.

The unique requirements of the Telecommunication Cubesat satellite demonstrate this evolution. The design needs to be relatively inexpensive while at the same time

computationally robust. It should consume limited power and must support the space environment by reducing susceptibility to radiation and thermal effects.
In a previous text, we supposed that all the satellite subsystems are intelligent. i.e.; each satellite subsystem has its own microcontroller.
In the case of a Cubesat-class satellite where there are many constraints in terms of power consumption, mass and size. To design the Telecommunication Cubesat satellite, the adopted approach consisted in the integration of all the intelligences of the various satellite subsystems in only one intelligent subsystem. This intelligent subsystem that we called Data Handling Subsystem will integrate the Telemetry/Telecommand functions, the Telecommunication payload as well as the active part of the thermal control subsystem.
Fig. 6 represents the general architecture of the Telecommunication Cubesat satellite in such way that the attitude control subsystem will be designed completely with passive methods, i.e., on the one hand, it does not have any programmed microcontroller to carry out an algorithm of attitude control and, on the other hand, it does not consume any electrical power provided by the power subsystem.

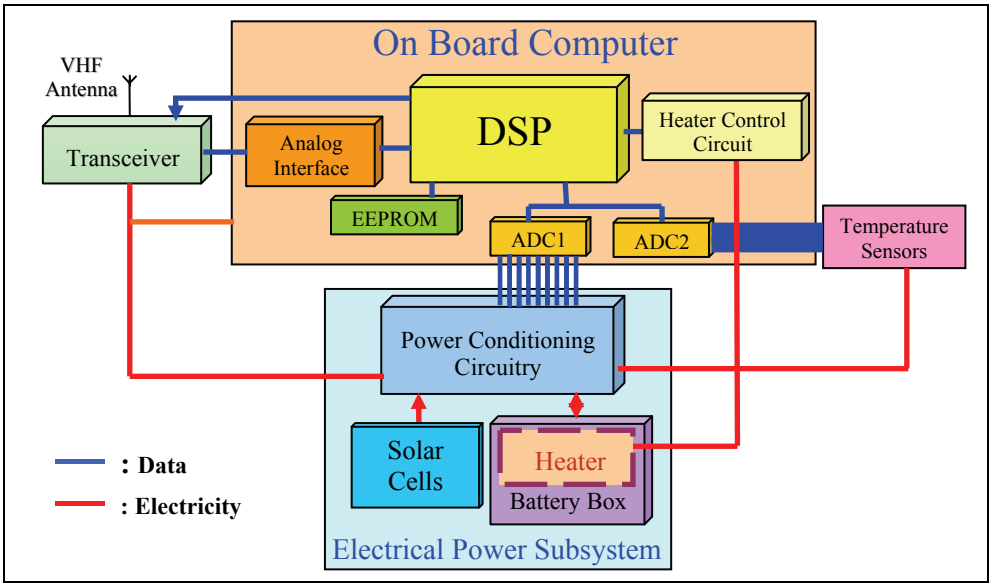



Fig. 6. General Architecture of the Telecommunication Cubesat Satellite

## 4. Cubesat satellite platform subsystems design

### 4.1 Space environment

An important issue to consider, which affects all electronic devices in space environment, is radiation. It can lead to various types of problems. These problems range from operational malfunctions to physical damage of the devices.
CMOS technology is preferred for space applications because of its high noise margins and low static power requirements. Scaling and integration are other advances CMOS technology has over other semiconductor technologies. On the other hand, CMOS is susceptive to two types of space radiation effects caused by electrons and protons trapped

by the terrestrial magnetic field: Total Ionizing Dose (TID) and Single Event Effects (SEE). TID effects are the result of accumulated exposure to ionizing radiation. SEE are the result of a single high-energy particle that strikes the device.
The total dose radiation exposure is measured in rads. The term rad (radiation absorbed dose) quantifies the total radiation exposure of a material. One rad(Si) is equal to $10 \times 10^{-6}$ W of energy absorbed per gram of silicon. The total dose radiation threshold of a device is the minimum level of rad(Si) that will cause device failure. Typical CMOS technology devices can survive nearly 5 krads before physical damage occurs (Pisacane & Moore, 1994).
However, single-event effects are significantly more hazardous to the satellite and can result in either Single Event Upset (SEU) or Single Event Transient (SET) (Poivey et al., 2003) or Single Event Latchup (SEL). SEU effects are internal device memory bit changes (0 becomes 1 and vice versa) that can cause erroneous instruction execution and SET is a transient voltage pulse that can cause erroneous data transmission over a bus. SEU and SET effects are considered soft-errors and do not cause physical damage to the devices. In contrast, SEL effect is a hard-error, which leads to a high current-flow through the device. If not remedied quickly, latch-up can cause permanent damage.

### 4.2 Satellite structural subsystem

After placing the satellite in orbit, the satellite structure materials must protect the electronic components against the ionizing radiation. The Aluminium thickness of the satellite structure determine the life time of the satellite. Many measurements made on board various satellites allow us to know the energies of the electrons and protons trapped on various altitudes. The space agencies, like ESA (the European Space Agency), established radiation models from those measurements campaigns. From this point of view, we used ESA software called SPENVES to calculate the necessary shielding to protect the electronic components in space environment. The studies determined that the satellite will experience a total radiation dose of approximately 400 rads per month when we chose Aluminium shielding of 2 mm (see Fig. 7), taking into account the constraint of the Cubesat-class satellite

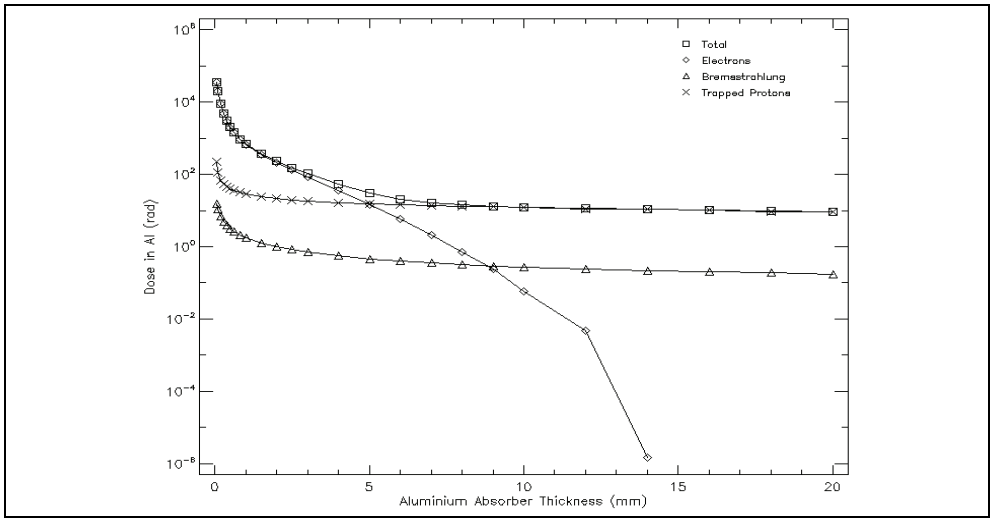



Fig. 7. Total Ionizing Dose for one month in function of the Aluminium Thickness

weight. The Aluminium shielding of 2 mm, which will be shared between the thickness of both the satellite structure and the electronic boxes, will allow the satellite to survive nearly 13 months. The studies also estimate that the satellite will experience approximately 2 bit-errors per 100 days, and only a small percentage of these will result in a latchup condition. If one in a thousand events cause latch up (Pisacane & Moore, 1994), the satellite will not observe any latch-up during its mission life.

### 4.3 Electrical power subsystem

The solar cells will be attached to the Cubesat structure, thus avoiding the need for solar panels and the necessary deployment system. The surface area for mounting solar cells is significantly reduced compared to conventional satellites, which results in less power generation ability. The solar cells chosen were spectrolab's Dual Junction Solar Cells, with an efficiency of 20% and dimension of 69 mm × 32 mm. They are not the most efficient available, but they provide the best value with low cost. These solar cells generate a power of 2.26 W which is a constraint

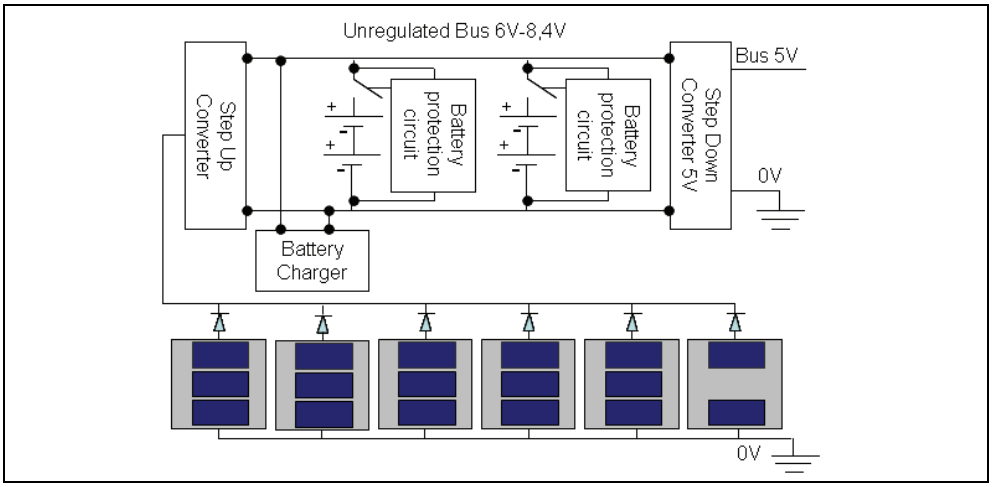



Fig. 8. Power Conditioning Circuitry Architecture

The battery Lion polymer DLP455660 from Danionics was chosen due to its extremely low weight, small size and good capacity. Four batteries were confined in a dedicated battery pressure box which will provide housing with sufficient stiffness to prevent the batteries from expanding against the high vacuum in Low Earth Orbit (LEO).

The Power Conditioning Circuitry (PCC), as shown in Fig. 8, comprises, one step-up converter based on MAX1771 circuit to provide the voltage to the unregulated bus, one battery charger based on ADP38101R-8.4 circuit with two battery protection circuits UCC3911 to protect the Li–Ion batteries from overcharging and overdischarging, and finally one step-down converter based on MAX1771 circuit to provide the voltage to the 5 V bus.

### 4.4 Thermal control subsystem

The advantage of using a limited number of electronics boards aboard the satellite is to provide a flexible internal configuration of components to arrange heat dissipating in such a

way that the excess heat is dissipated to the surroundings of the satellite as efficient as possible (Rotteveel, 2006).
Satellites that use a passive attitude control require a more homogenous thermal control where each side of the satellite body has more or less identical optical properties. As a result of the body mounted solar arrays, the optical properties of the different body panels are relatively homogeneous. This reduces the options to control passively the temperatures for the Nanosatellite. Therefore, the satellite will carry a combination of passive and active solutions for the thermal control. The passive control methods are in the form of thermal coatings (Wertz & Larson, 1999) of electronics boxes. For the active control, one heater and one temperature sensor are attached to the batteries box in order to protect them from cold temperature during the eclipse. The On Board Computer uses a timer to read periodically the batteries temperature among others telemetry data. The heater will be activated by the On Board Computer each time the batteries temperature drops below 5 °C (see Fig. 9). A polyimide heater from Minco Products has been selected. It is ideal for applications with space and weight limitations, or where the heater will be exposed to vacuum.

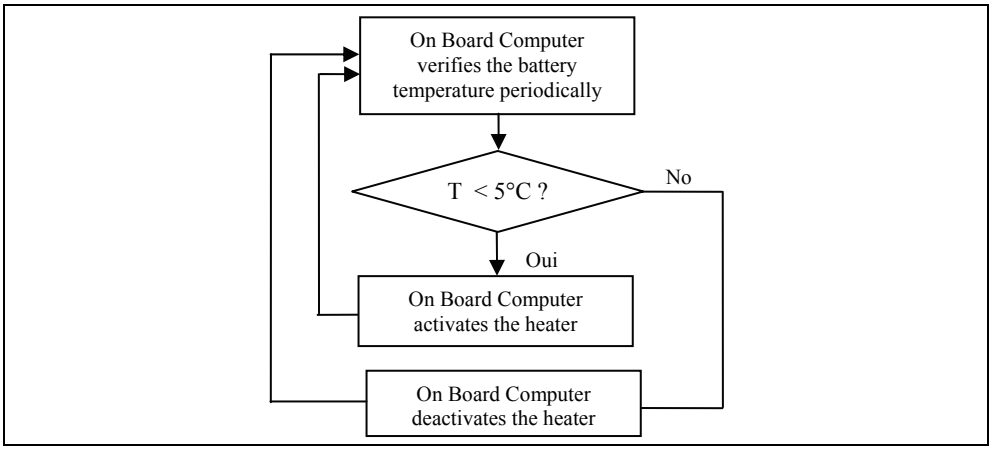



Fig. 9. Flow chart of the battery temperature control

### 4.5 Attitude control subsystem

To cut down the cost of the Telecommunication CubeSat-class satellite, all the nonessential elements is suppressed. An active attitude, which consumes the energy, is avoided by using omnidirectional antenna.
In the case of early launch CubeSats, the CubeSats that have adopted permanent magnets as passive attitude control, still alive and operated well. On the contrary, other Cubesats that incorporated active attitude control had many problems and was unstable.
The attitude control consists of two permanent magnets placed perpendicular to Six Hysteresis rods (see Fig.5). In fact, if one would think of the Earth as a giant bar magnet, then two parallel permanent magnets, using Neodymium Fer Bore (N35) material, aligned along an axis, would try to passively align the satellite according to the Earth's magnetic field lines. The satellite will spin around this axis and will be slowing down using six hysteresis rods (Alloy49) attached perpendicularly to this axis. The force caused by the hysteresis rods opposes the change in motion of the satellite. Also, it stabilizes the power

generation and provides uniform thermal distribution to the satellite. In order to point the antenna to the earth, the antenna side of the satellite should be made as heavy as possible, acting like a gravity gradient. The antenna is mounted close to the transceiver on the center of the side opposite to the magnet side. Fig. 10 shows the satellite motion on its orbit.

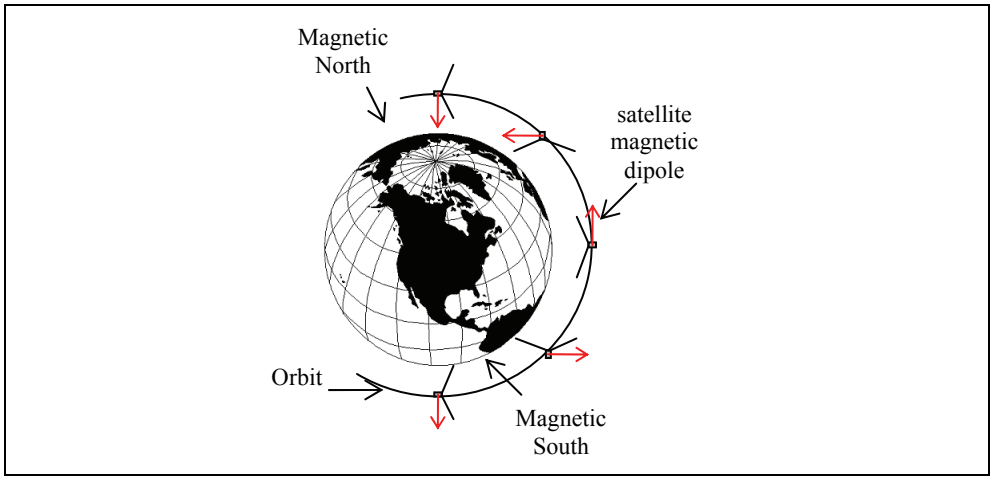



Fig. 10. Satellite motion on its orbit

## 5. Data handling subsystem design

In the satellite modular architecture (Wertz & Larson, 1999), each single subsystem has a dedicated hardware and software. The Cubesat has important constraints on cost, power and mass, and especially on size. The approach that has been taken in this research consists of the integration of the maximum subsystems within the same unit taking into account that single subsystems can be setup without modifying the operation of the remaining subsystems.

As said before, the Data Handling Subsystem will integrate the Telecommunication Payload, the Telemetry/Telecommand functions as well as the active part of the thermal control subsystem. In this section, we will describe in detail the main considerations and solutions chosen during the design of the Data Handling Subsystem based on the fixed point DSP processor. The design was split in two parts: The Hardware and Software Architectures.

### 5.1 Data handling subsystem hardware architecture

The Data Handling Subsystem hardware architecture is composed of three main parts (Fig. 11): an on board computer, Sensors and Control signals, and a VHF transceiver with associated antenna.

#### 5.1.1 On board computer

The On Board Computer board as illustrated in Fig. 11 is a hardware board in which is implemented the flight software. The flight software controls the whole operations of the satellite and is built around a 16 bits DSP TMS320C5416 from Texas Instrument (Texas

Instrument, 2002). The hardware board includes, moreover, one analogue interface connected to the Radio Frequency module (Transceiver and antenna), two analogue to digital converters for the acquisition of the Telemetry housekeeping data, a EEPROM memory containing the flight software, a JTAG port for the flight software programming and finally the control signals, by using some DSP Input/Output signals, in particular, to activate or deactivate the heater. Fig. 11 illustrates the hardware architecture of the on board computer. All the components on the on board computer logic board are chosen with SMD (Surface Mounted Device) packages for space saving and high density mounting in order to minimize the weight and dimensions of the logic board.

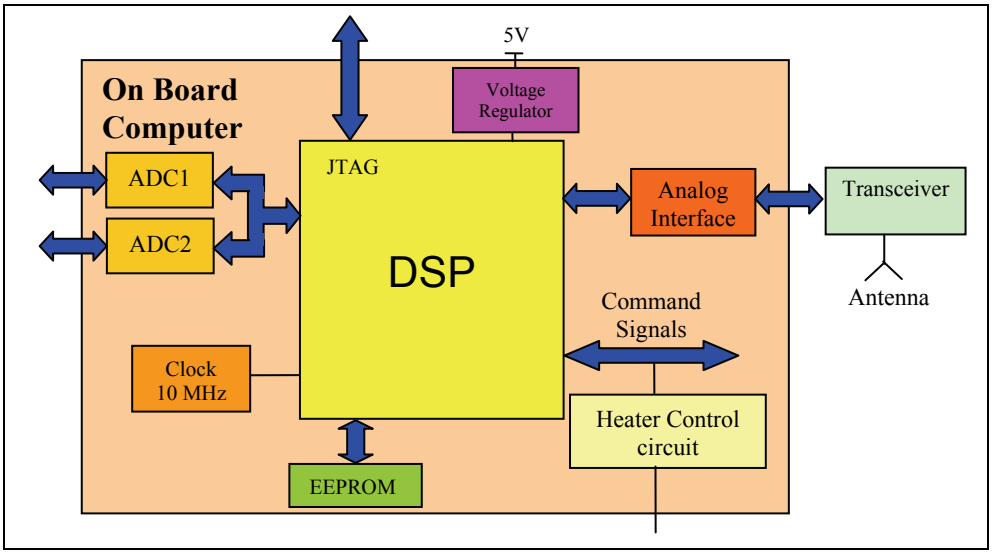



Fig. 11. The hardware architecture of the on board computer

**Processor**

The Cubesat projects generally use different general purpose processors for the on board computer. The AAU CubeSat used C161PI microcontroller from Infineon, the CUTE-1 CubeSat used an 8-bit H8/300 CPU and The RinconSat CubeSat used a PIC16C77 microcontroller from Microchip Corporation.

As on board computer for the Telecommunication CubeSat-class satellite, the TMS320C5416 DSP efficient power consumption, from Texas Instrument was chosen for its ease of implementing the physical layer (AFSK and GMSK modems) and the data link layer (AX.25 protocol), and also for the availability of many reports applications, which reduce the design time of the satellite. The advantage of implementing AFSK and GMSK modems as software is to eliminate the need for external modem hardware needed by other satellites (Hunyadi et al., 2002), and to minimize the dimensions of the on board computer logic board which must be less than 10 cm x 10 cm.

The DSP processor is connected to VHF transceiver through the AIC (Analogue Interface Circuit) interface. The TMS320C5416 contains an adjustable PLL which allow us to multiply the frequency of Quartz to reach the maximum of 160 MHz. However, the operation clock frequency of TMS320C5416 is selected to 40 MHz for the following reasons: to minimize the

energy consumption, to minimize the heat generation and also to avoid the Single Transient Event effect (SET) (Poivey et al., 2003).

**Memory**

The On Board Computer is an embedded system where the flight software is implemented in an external memory. To immunize against SEU effect, the application program is loaded in the external Programmable Read Only Memory (PROM). In general, the PROM memory is used, because it is very robust to radiation and in wide range of temperature. In the event that the APRS OBDH unit countermeasures fail, a Fail Safe Circuit is used to perform a satellite reset at regularly scheduled intervals.
The boot mode is chosen to be the SPI EEPROM boot mode using the AT25256A EEPROM memory of 32 Kwords of 8 bits which operates in the temperature range from -40 up to 125 °C. The SPI EEPROM mode is the only mode that does not require the addition of electronic components to control the transfer of the flight software into the memory program of TMS320C5416 for a faster execution. The DSP TMS320C5416 contains 128K × 16 on-chip RAM and 16K × 16 on-chip ROM.

**Analogue Interface**

The analogue interface or AFE (Analogue Front End) carries out several operations which will make it possible to adapt the signals between the DSP and Radio Frequency (RF) module. These operations are practically amplification, filtering and conversion. Fig. 12 illustrates the internal architecture of AFE interface which is chosen to be the TLC320AC021 analog interface from Texas Instrument.
The TLC320AC021 (Texas Instrument, 1996) (AIC), includes both analog-to-digital and digital-to-analog converters in the one package. This device incorporates a band pass switched capacitor antialiasing filter, a 14 bit conversion process for both the ADC and DAC, and a low pass switched capacitor output reconstruction filter. In addition, the AIC also provides a direct serial interface to the TMS320C5416. Use of this AIC greatly simplified the hardware necessary to provide an analogue interface to the DSP development system. The TLC320AC021 is specified for the −40 ◦C to 85 ◦C temperature range.

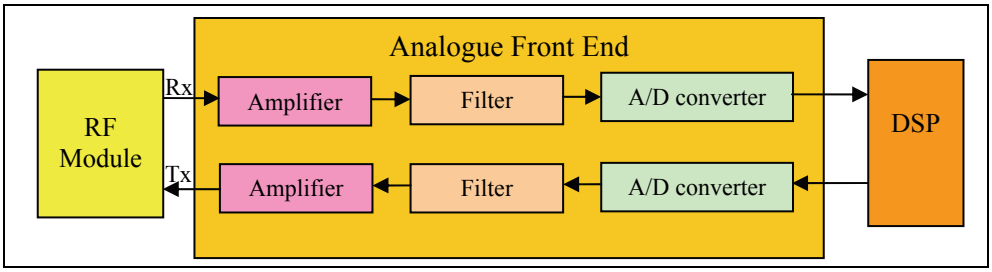



Fig. 12. the internal architecture of AFE interface

**Analogue Digital Converter**

The Interface to the different 17 sensors is done by using the TLV2556 circuit, which is a serial Analogue Digital Converter (ADC) with 11 analogue multiplexed channels. The serial bus, connected to the DSP, is based on the SPI specification.
The TMS320C5416, which is the Master, selects one of the two TLV2556, used as a slave mode. The TLV2556I is characterized for operation from −40 ◦C to 85 ◦C.

### 5.1.2 Transceiver

A COTS amateur radio transceiver was adopted to be the main flight radio due to power, weight, and time constraints (Horan, 2002; Lu, 1996). The transceiver, integrated inside the payload, operates in amateur VHF band and consists of the "guts" of a low cost Yaesu VX-1R (Hunyadi et al., 2002), arguably one of the smallest and lightest handhelds on the market. The radio is two stackable double sided PCBs measuring approximately $5 \times 5$ cm$^2$. The small size of CubeSats limits the amount of energy provided by solar arrays; therefore power availability is a constraint on both the spacecraft processor and the communications systems. Power is supplied from 5 V bus to radiate only 1 Watt RF power which achieves a positive budget link (see Table 1). Current consumption for the receiver and transmitter is 150 mA and 400 mA, respectively. Only slight modifications will be required to make this component space worthy. We use only the RF parts of the VX-1R.

The transceiver is interfaced with the DSP processor by means of AF (Audio Frequency) and PTT (Push To Talk) command signals. AF signals consist of transmit and receive signals of the AFSK and GMSK modems which carry data packet, whereas PTT command signal allows the DSP to choose the transmitting and receiving frequencies of the radio. The transmitting and receiving frequencies will be fixed on the 145.825 satellite frequency. The FM signal output from the transceiver is fed to an omnidirectional antenna.

### 5.1.3 Antenna

We have previously said that the attitude control of the satellite is based on the Earth magnetic field. As a result, the pointing direction of the satellite depends on its latitude. Therefore, the satellite antenna must be omnidirectional to cover the earth Terminals placed in all latitudes. For the satellite antenna, the turnstile antenna (Milligan, 2005) will be used. It consists of two pairs of half-wavelength dipole antenna positioned orthogonal to each other with a relative phase difference of 90 degree to achieve circular polarization. In theory of transmission line, a 90-degree phase shift can be obtained by adding a ¼ wavelength transmission line to the feed of the antenna. However, due to the limited available space in the satellite and to the long ¼ wavelength transmission line at VHF frequencies; implementation using lumped elements (Bahl, 2003) is desirable. To make easier the deployment of the antenna, the measuring tapes will be used as the four elements that consists the circularly polarized antenna.

## 5.2 Flight software architecture

The designed flight software of the Telecommunication CubeSat-class satellite is made to boot-up automatically from the external EEPROM memory and then to configure the TMS320C5416 DSP to carry out the various functions of the on board computer. The flight software is a real time software which treats the functions of the on board computer dependently of the material environment of the TMS320C5416 DSP. Its architecture was built according to two points of view:

**A static point of view:** This describes how the flight software activities are organized. All the functions to be realized by the software are broken up into modules (objects) at the same time as the definition of the relations between these modules. Those are again broken up. The software is presented in a hierarchy of modules. The modules with the lowest level are coded in the programming language. They are tested then assembled to recompose the module of the higher level and so on up to the highest level.

**A dynamic point of view:** This describes how, in time, the flight software activities will be processed (we speak about the software behaviour). These activities are organized in tasks

which share in time the resources of TMS320C5416 DSP (CPU, Memory) by assigning a priority to each task (depending on the priority of each interruption).
The assembler programming is used for the modules which require real time processing like the modules including in the communications module (COM module). In addition, the programming in C is used for the programming of the other modules of the flight software to reduce the time of their developments.

### 5.2.1 Static architecture of the flight software

Fig. 13 shows the static architecture of the flight software which is a modular architecture organized in hierarchical levels. It is designed to boot-up from a Main Control Program (MCP) acting as the principal module of the first level. Its principal function is to route the information inside the various software modules of the on board computer. The modular architecture of the flight software makes it possible to test the functionality of each module independently of the others.

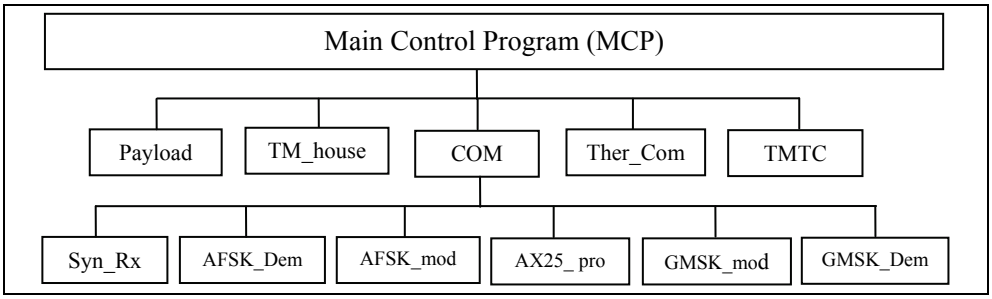



Fig. 13. Static Software Architecture of the on board computer

| | |
|---|---|
| **TM _ house** | acquires of the telemetry data of the type "housekeeping". |
| **Payload** | carries out the missions of the Nanosatellite payload. |
| **Ther_Con** | activate ou deasactivate the heater. |
| **TMTC** | executes the telecommands and sends all type of telemetries. |
| **Syn_Rx** | has as a function the detection of the beginning of the received packet. |
| **AX25_Pro** | carries out the packing and the unpacking of the AX.25 packet. |
| **AFSK_Mod** | carries out the AFSK modulation. |
| **AFSK_Dem** | carries out the AFSK demodulation. |
| **GMSK_Mod** | carries out the GMSK modulation. |
| **GMSK_Dem** | carries out the GMSK demodulation. |

Table 4. Various modules included in the flight software

### 5.2.1 Dynamic architecture of the flight software

Once the satellite is separated from the launch vehicle or after a reset, the TMS320C5416 DSP loads the software, stored in the external EEPROM memory, into its memory program for a faster execution of the software. The first part of the software contains the instructions necessary to configure the various registers and to initialize the various peripherals of the TMS320VC5416 DSP. Once these initializations are finished, the program is put on idle state waiting for an interruption either from the Timer or from the analogue interface. The flow chart of this main program is given by Fig. 14.

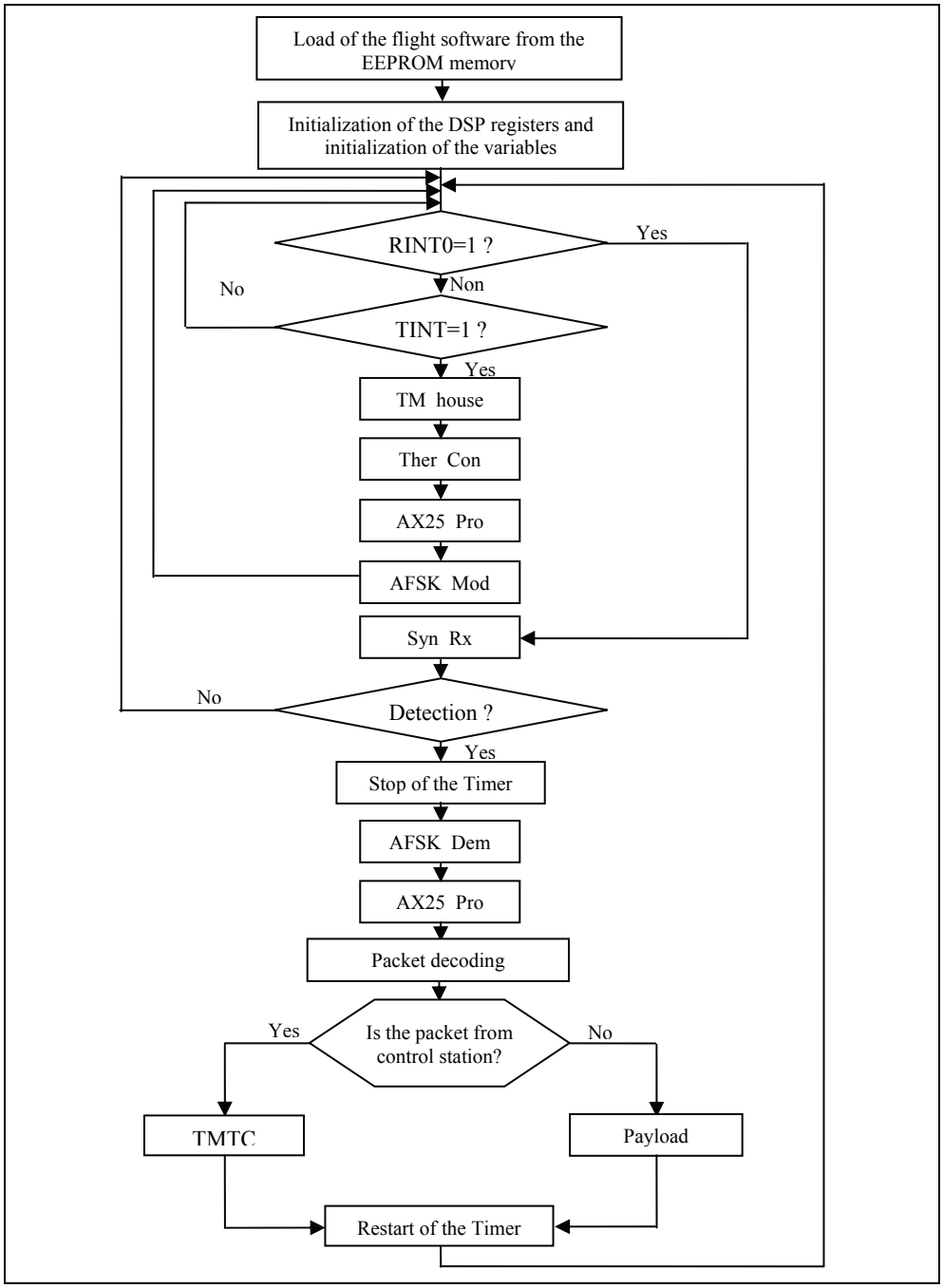



Fig. 14. The main program flow chart of the on board computer

If the interruption will come from the timer (TINT=1), the main control program (MCP) will call first the "TM_house" module, then it will control the battery temperature by the "Ther_Con" module. After that, it will use the "AX25_Pro" and "AFSK_Mod" modules for sending the APRS beacons and, at the end, it will return to its idle state.
In addition, if the interruption will come from the analogue interface (RINT0=1), the MCP will control the reception of the data by the "Syn_Rx" module. If the beginning of the packet is detected, the MCP will stop the Timer counting, and then it will call the "AFSK_Dem" module. At the end of the demodulation, the MCP will test the received packet if it contains an error by the "AX25_Pro" module. The MCP will reject the erroneous packet and will respond to the request of the correct packet. If the packet is coming from the control station, the MPC will give the command to the "TMTC" module to take over. If it is coming from terminals, the "Payload" module will take over. At the end of these tasks, the MPC will return to its idle state and will continue the counting of the Timer from its value at the moment of stop.

### 5.3 DSP implementation

In this research, we give a particular attention to software optimization due to small amount of memory available on DSP. Today, "C" compilers offer many features to facilitate DSP programming but they do not guarantee the best code optimization. We have defined a methodology approach based on two design steps: algorithmic optimization, and optimal programming using assembler language for the modules which need the real time. By an optimal software design that takes the maximum advantages from DSP architecture facilities, good real time experimental performances have been obtained for the on board computer implemented modules. The implemented modules include AX.25 protocol, AFSK modulation, AFSK demodulation, Convolutional encoding and Viterbi decoding, GMSK modulation, and GMSK demodulation. These modules were tested in real time using two DSP platforms, one for the transmission process and the other for the reception process. The used platforms are TMS320C54X DSKPlus kit (Texas Instrument, 1997) and TMS320C5416 DSK kit (Texas Instrument, 2002), which are compatible. Fig. 15 illustrates the interconnection of the two development boards DSKC542 and DSKC5416.

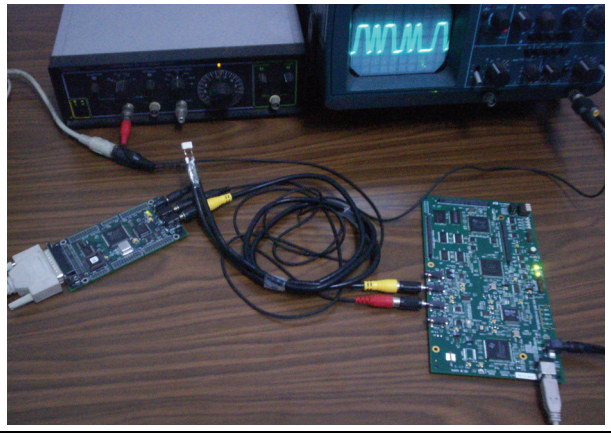


Fig. 15. The interconnection of the DSKC542 and DSK5416 boards

### 5.3.1 AX.25 communication protocol

The AX.25 protocol (TAPR, 1997) is based on HDLC format. The general frame format used for data transmission is shown in Fig. 16.

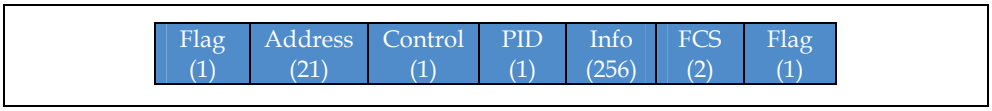



Fig. 16. The AX25 frame format where number of bytes are in parenthesize

Briefly described the functionality of each field in the frame is:

Flag: Indicates start and stop of the frame

Address: Contains sender address, receiver address and satellite address

Control: Identifies the type of frame which in this case is an information frame

PID: Identifies the type of top-level protocol

Info: Contains from zero to 256 bytes of data

FCS: The Frame Check Sequence which will be described in detail below.

The AX.25 protocol uses the standard CRC-CCITT code based on a 16-bit polynomial G(x) to calculate FCS field.

$$G(x) = x^{16} + x^{12} + x^5 + 1 \tag{4}$$

For CRC-CCITT code implementation on DSP, the truncated polynomial of 0x1021 is used with initial value of 0xFFFF.

### 5.3.2 AFSK modulation

A direct digital frequency synthesis technique (DDS or DDFS) combined with a digital-to-analog converter (DAC) provides an approach to implementation of an Audio Frequency Shift Keying (AFSK) modulator. Samples of a sine wave are generated directly from the sine look-up table. The primary function of AFSK modulator is the following: Given a stream of binary data $a_0$, $a_1$, $a_2$,.. , $a_k$ for each element $a_k$ [0, 1], generates a corresponding signal of frequency, $f_0$ = 1200Hz or $f_1$ = 2200Hz, for the duration of a bit period. The 1200Hz tone represents the binary "1" and the 2200Hz tone represents the binary "0" as indicated in the standard Bell 202. The AFSK modulator generates the two tones by stepping through the same sine table.

Suppose $f$ is the desired frequency. The signal to generate is x (t), with:

$$x(t) = \sin(2\pi ft) = \sin(\varphi(t)) \tag{5}$$

and $t = nT_s$ where $T_s$ is the sampling period. Because of Shannon's theorem, $1/T_s$ is chosen greater than $2f$. The phase φ belongs to the interval $[0,2\pi]$. It is possible to suppress the time consuming comparison tests with $2\pi$. This can be achieved by using circular buffer as intrinsic feature of TMS320C54X processors. The size $N$ of the table can be calculated in order that the step index $S$ is an integer for both frequencies. This condition can be written as:

$$Nf_0T_s = S_0 \quad and \quad Nf_1T_s = S_1 \tag{6}$$

The size of the step index determines the output signal frequency. At the bit rate $f_b$ of 1200 Hz, an interruption of AIC is sent to the DSP. To generate the frequency $f_0$ = 1200 Hz (respectively $f_1$ = 2200 Hz), the sine table of size N = 120 is read with an integer step index equal to $S_0$ = 6 (resp. $S_1$=11).

For the implementing of the AFSK modulation on DSP, we used the sampling frequency of 24 KHz and data rate of 1200 bps which corresponds to 20 samples per bit. The steps and the sine samples are represented as 16 bit integer numbers. Fig. 17 represents the output of the AFSK modulator with the following bits of inputs [-1 1 1 1 -1 1 -1 -1 1 1].

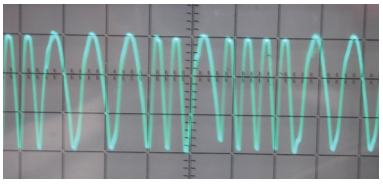


Fig. 17. The AFSK signal

### 5.3.3 AFSK demodulation

We used a bit-per-bit demodulation as the classical non-coherent demodulation scheme. The received AFSK signal is sent to DSP from the transceiver via the TDM serial port after being converted from analog to digital signal by AIC. The DSP implementation of the AFSK demodulator is illustrated in the Fig. 18.

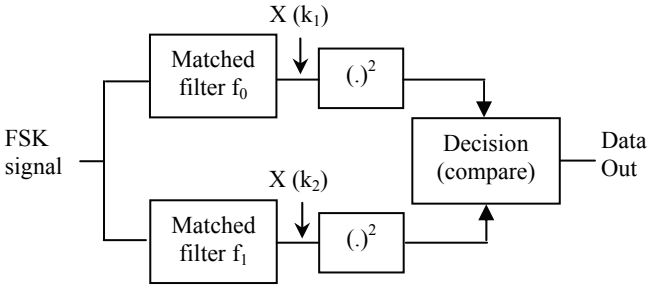



Fig. 18. General diagram of AFSK demodulation

We used the Goertzel algorithm (Oppenheim, 1999) to demodulate the AFSK signal, which can be interpreted as a matched filter for each frequency k as illustrated in Fig. 18. The transfer function $H_k(z)$ corresponds to the $k$th Goertzel filter:

$$H_k(z) = \frac{1 - e^{j2\pi k/N} z^{-1}}{1 - 2\cos(2\pi k/N) z^{-1} + z^{-2}} \tag{7}$$

A further simplification of the Goertzel algorithm is made by realizing that only the magnitude squared of X(k), which represents the energy of the received signal, is needed for tone detection. It eliminates the complex arithmetic and requires only one coefficient, $a_k$ = $\cos(2\pi k/N)$, for each $|X(k)|^2$ to be evaluated. Since there are two possible tones to be

detected, we need two filters described by (7). We conclude that the Goertzel algorithm is a Discrete Fourier Transform, calculated from a second degree recursive filter, easy to implement on DSP. In Our case, we compare only the two energies of the two AFSK frequencies to determine which AFSK tone has been received.
The synchronization is performed by detecting the first change to the received signal by using the Syn_Rx module. After processing 20 samples for each bit and calculated the energy at each of the two frequencies, the Goertzel Algorithm then decides which AFSK tone has been received. The sampling frequency is chosen to be 24 KHz because it is the highest sampling frequency available in the AIC. Also to detect the frequency 1200 Hz (resp. 2200 Hz), we used k = 1 (resp. k = 1.83). For M = 20, we have $a_1 = 0.951$ and $a_2 = 0.839$, which are corresponding to frequencies 1200 Hz and 2200 Hz respectively. The format of each variable in the algorithm was being chosen suitably taking into account that we had used a 16 bit fixed point DSP.

#### 5.3.4 GMSK modulation

The GMSK modulation is a Continuous Phase Modulation (CPM) with a modulation index $h$=0.5. A modulated GMSK signal can be expressed, over the time interval $nT_b \le t \le (n+1)T_b$, as:

$$s(t) = A\cos\left(2\pi f_0 t + 2\pi h \sum_{k=-\infty}^{n} d_k \int_{-\infty}^{t} g\left(\tau - kT_b\right) d\tau\right) \tag{8}$$

where $d_k$: sequence of data information = ±1,

and $g(t) = \frac{1}{2T_b} rect(t/T_b) * h_g(t)$ with $rect(t) = 1 \quad for \quad |t| \le 0,5$

$h_g(t)$ is the pulse of Gaussian function, $T_b$ is the symbol period, $B$ is the 3dB bandwidth of the Gaussian prefilter, and $g(t)$ is the response of the transmitted rectangular pulse to the pre-modulation filter.
By deriving the phase signal, the CPM can also be seen like Frequency Modulation (FM). The instantaneous frequency $F_i$ is given by:

$$F_i(t) = f_0 + h \sum_{k=-\infty}^{n} d_k g(t - kT_b) \tag{9}$$

In the expression (9), $h$ represents the proportionality constant of the modulator and is expressed in Hertz per volt. The baseband signal m(t) to be transmitted is written then, in the interval $nT_b \le t \le (n+1)T_b$, in the form of:

$$m(t) = \sum_{k=-\infty}^{n} d_k g(t - kT_b) \tag{10}$$

In theory, the duration of Gaussian filter is infinite, but in practice, we limit the function $h_g(t)$ to the few period bits over which it is significantly not zero. This duration is inversely proportional to $B$. For a product $BT_b = 0.5$, we consider that $h_g(t)$ is not zero over 2 bits. The convolution product of $h_g(t)$ with a rectangle function of duration $T_b$ lasts $3T_b$, which affects

the half preceding bit and the half following bit. The Fig. 19 represents the response of Gaussian lowpass filter for $BT_b$ = 0.5 over three bits to a rectangular pulse of duration $T$.
The implementation of filter convolution product requires multiple instruction processing inducing a lot of calculation time. To respect timing constraints we propose an optimized implementation code based on Lookup table of the Gaussian filter response (Fig. 19). For the implementing of the GMSK modulation on DSP, we used the sampling frequency of 24 KHz with 5 samples per bit which corresponds to data rate of 4800 bps. For data stream of [1 -1 1 1 -1 -1 1], the corresponding GMSK baseband signal is given by the Fig. 20.

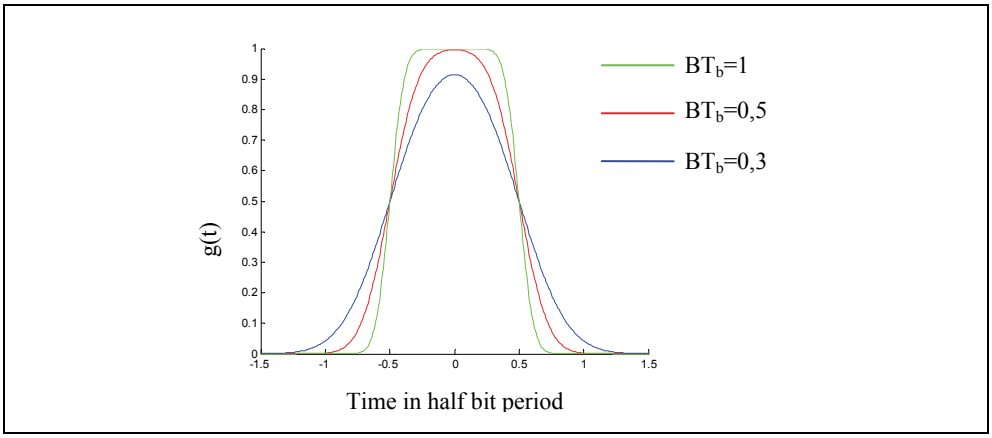



Fig. 19. Gaussian filter response in function with $BT_b$ parameter

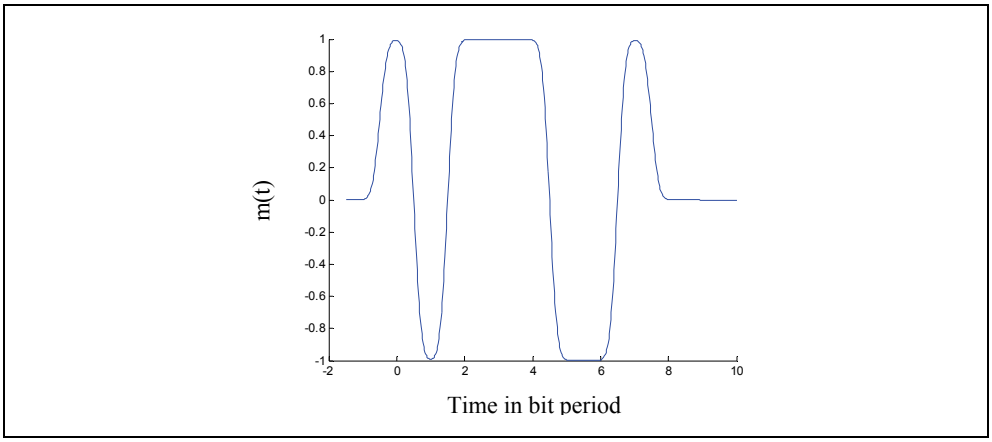



Fig. 20. Baseband GMSK output signal

### 5.3.5 GMSK demodulation

We used the classical non-coherent demodulation scheme, which performs a bit-per-bit demodulation and it does not require recovery of the carrier phase and frequency. Analysis of the GMSK baseband signal (Fig. 20) permits the identification of eight types of shapes

corresponding to binary states transition. The GMSK demodulator must extract the phase from the modulated signal and, by using a transition shape classification, decode the transmitted bit.
According to Fig. 21, we have four transition shapes for a binary "1", and four transition shapes for a binary "0". We store only two predictive transitions, (b) and (f), on the DSP memory as look-up tables. Based on the lookup tables, the demodulator uses the Absolute distance $d_e$, which shows the better performance, as matching function to classify the GMSK signal transitions, and determine the transmitted bit.

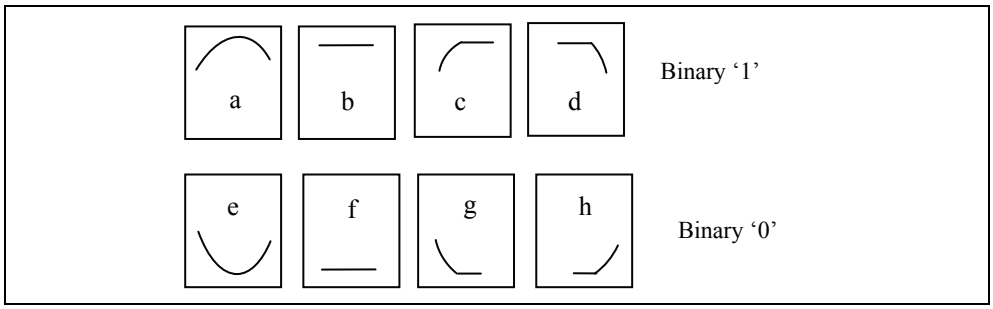



Fig. 21. Eight binary states transitions

$$d_{abs}(x,y) = \sum_{j=1}^{n} |x_j - y_j| \tag{11}$$

The demodulation of the GMSK signal is processing to perform the shape comparison of binary transition based on the look up tables. The minimum Euclidean distance $d_e$ is evaluated and the decoded bit is determined. The synchronization is performed by using the Syn_Rx module. The C54x DSP family has a dedicated instruction for faster execution of the Absolute distance.

## 6. Conclusion

As the satellite community transitions towards inexpensive distributed small satellites, new methodologies need to be employed to replace traditional design techniques. The ongoing research will contribute to the development of these cost saving methodologies. The goal of the integration of all the intelligences of the various satellite subsystems in only one intelligent subsystem is to minimize component expenditures while still providing the reliability necessary for mission success.
Associating low cost ground terminals with a low cost Telecommunication CubeSat-class satellite will allow universities to access space communications with a very economical system. The present work, dealing with the design of the Low-cost Telecommunication CubeSat-class spacecraft, shows hardware and software solutions adopted to cut down the system cost. The hardware utilizes commercial low cost components and the software is optimized using assembler language. The On Board Computer unit is small device that can be mounted on any small satellite platform to serve telecommunications applications such as mobile localization and data collection. By using a single CubeSat satellite and low-cost

communications equipments, Telecommunications systems can be kept at the extreme low end of the satellite communications cost spectrum.

## 7. References


Addaim, A.; Kherras, A. & Zantou, B. (2008). Design and Analysis of Store-and-Forward Data Collection Network using Low-cost Small Satellite and Intelligent Terminals, *Journal of Aerospace Computing, Information and Communications,* Vol. 5, No. 2, (February 2008) page numbers (35-46)

Bahl, I. (2003). *Lumped Elements for RF and Microwave Circuits*, Artech House, first ed.

Gérard, M. & Bousquet, M. (2002). *Satellite Communication Systems*, John Wiley & Sons; fourth edition

Horan, S. (2002). Preparing a COTS radio for flight – lessons learned from the 3 corner satellite project, *Proceedings of 16th Annual/USU Conference on Small Satellites*, Logan, Utah, USA

Hunyadi, G.; Klumpar, D.; Jepsen, S.; Larsen, B. & Obland, M. (2002). A commercial off the shelf (COTS) packet communications subsystem for the Montana EaRth-Orbiting Pico-Explorer (MEROPE) CubeSat, Proceedings of IEEE Aerospace Conference

Jamalipour, A. (1998). *Low Earth Orbital Satellites for Personal Communication Networks*, Norwood, MA: Arthech House

Lu, R. (1996). Modifying off-the-shelf, low cost, terrestrial transceivers for space based application, *Proceedings of the 10th Annual AIAA/USU Conference on Small Satellites*, Logan, September 1996, Utah, USA

Milligan, T. (2005). *Modern Antenna Design*, second ed., Wiley

Oppenheim, A.; Schafer, R. & Buck, J. (1999). *Discrete-Time Signal Processing*, second ed., Prentice Hall

Paffet, J.; Jeans, T. & Ward, J. (1998). VHF-Band Interference Avoidance for Next-Generation Small Satellites, *Proceedings of* 12*th AIAA/USU Conference on Small Satellites*, Logan, Utah, USA

Pisacane, V. L., & Moore, R. C. (1994). *Fundamentals of Space Systems*, New York: Oxford University Press

Poivey, C.; Buchner, S.; Howard, J. & Label, K. (2003). *Testing Guidelines for Single Event Transient*, NASA Goddard Space Flight Center, 30 June, 2003.

Proakis, J. (1989). *Digital Communications*, McGraw-Hill, (Second Edition)

Rotteveel, J. (2006). Thermal control issues for nano- and picosatellites, *Proceedings of Space Technology Education Conference*, Germany, May 2006, Braunschweig.

TAPR, (1997). AX.25 Link Access Protocol for Amateur Packet Radio, TAPR, version 2.2

Texas Instruments, (1996). TLC320AC01 data manual single-supply analog interface circuit, SLAS057D

Texas Instruments, (1997). DSKplus User's Guide, SPRU191

Texas Instruments, (2001). TMS320C54X DSP: CPU and peripherals, SPRU131G.

Texas Instrument, (2002). TMS320VC5416 DSK Technical Reference,

Wertz, R. & Larson, W. (1999). *Space Mission Analysis and Design*, Microcosm, (third ed.)

Zantou, B. & Kherras, A. (2004). Small Mobile Ground Terminal Design for a Microsatellite Data Collection System, *Journal of Aerospace Computing, Information and Communications,* Vol. 1, No. 9, (September 2004) page numbers (364–371)

# 16

# Looking into Future - Systems Engineering of Microsatellites

H. Bonyan
*Faculty of Energy Engineering and New Technologies, Shahid Beheshti University (SBU)*
*Iran*

## 1. Introduction

Space age began with the launch of Sputnik-1 in 1957, by the Soviet Union. Initially, the spacecraft, especially the western ones, were rather small due to limited capabilities of the launch vehicles. With the increasing capabilities of rocketry in the US and USSR, the limitation was soon a part of history. From 1970s, several-thousands-kilograms satellites have been placed in orbits ranging from LEOs to GEOs and to interplanetary orbits. These large satellites have been the major payloads of launch vehicles until the very last years of the Cold War, the so-called "Super-power, government-only space era". During the last two decades, however, there has been an ever-increasing interest within the private sectors in developed countries and, also, space agencies of developing countries to contribute to and take advantage of space market. It must be reminded that large satellites are not appropriate means to establish the required hardware-/software-expertise and infrastructure. Simply, the private sector is not able to afford the huge costs of large satellites and its immense complexity. This also holds true for government-funded project in many developing and third-world countries. Thus, most countries and space agencies have adopted microsatellite projects in order to initialize their space policy in order to obtain, establish and benefit from the rich space revenue. Thus, a "government/private-sector era" has been already initiated and almost established. In this methodology, microsatellites have served as "path-finders", in order to pave the way of many nations and societies (top-class universities in developed countries, space-agencies in developing countries and so on) to obtain the space technology. In the space literature of the last two decades, microsatellites have been addressed as "hands-on experience" to facilitate consolidation of space technology in order to implement some "actual large satellite" programs. Microsatellites in the next decades, however, will be employed not only as "path-finders" and/or "hands-on experience" warm-ups, but also as actual projects with considerable financial Return on Investment (ROI). This requires fundamental reconsideration of system-level characteristics of microsatellite projects, such as mission definition, subsystem performance requirements, construction, test, launch and post-launch operations. The preceding issues are addressed in this chapter.

## 2. Mission definition

Traditionally, microsatellites have served as engineering programs in order to pave the way for different communities (universities, organizations and/or nations) to acquire enough

"hands-on experience" for establishment of actual several-hundred/several-thousand kilograms satellite programs. While this approach has considerably contributed to recent advancements in satellite technologies in many developing countries and elsewhere, it still utilizes few of enormous capabilities of microsatellites. Microsatellites developed in the said paradigm, mainly serve to educate highly-qualified space engineers and managers. However, once in orbit, these vehicles are utilized to an order of magnitude less than their full capability. There are evidences that some well-designed, built and launched microsatellites have been almost abandoned after a few months in orbit. However, if properly planned, these vehicles could have been actively in service for a few years rather than a few months. It must be reminded that the owner authorities of the satellites (mostly universities and space-industry) are reluctant to officially declare the ineffectiveness of the actual products of the spaceborne system i.e. microsatellite in orbit and mostly emphasize on educational achievements of such programs. However, according to [H.Bonyan, 2010]; [E.Gill et al., 2008]; [U.Renner & M.Buhl, 2008]; [G.Grillmayer et al., 2003] & [United Nations UNISPACE III, 1998], there are evidences that there will be an enormous enhancement in actual outcomes of microsatellite programs, from a practical-application and/or economical-value point of view. The enormous enhancement of products of microsatellite programs, stated above, is briefly described in the following paragraphs.

During the last two decades, there has been an immense progress in the miniaturization of equipments incorporated in microsatellite technology. Miniaturization, in its broadest sense, is interpreted as provision of the same level of functionality via fewer resources. In satellite technology, resources are considered as mass, power and volume[1]. Today, with the increasing progress in computer technology, Commercial-Off-The-Shelf (COTS) units are accessible within the commercial space market. While these units are provided at fairly reasonable prices, they are as capable as their quite-expensive predecessors. For a given level of performance, these new units are also lighter and less power-hungry which, in turn, can be considered as extra financial benefit. Also, more efficient solar cells and battery units are now offered by suppliers of various communities. Furthermore, compact, light-weight and reliable reaction wheels and other attitude control actuators are provided by several suppliers [SSTL website, as of 2009]; [Sun Space website, as of 2009]; [Dynacon Inc. website, as of 2009] & [Rockwell Collins Deutschland website, as of 2009]. A complete list of these new components is not within the scope of this writing. It is being concluded that, at present and near future, microsatellites are and will be capable of fulfilling sophisticated missions, previously feasible only by several-hundred kilogram satellites.

The preceding advancements, to some extent, are true for every engineering field. However, they are an order of magnitude more important regarding microsatellite technology. It is being reminded that mass and power are critical issues in space technology. At the present time (as of 2009), placing a kilogram of payload into Low Earth Orbit (LEO) can be as expensive as 5000-15000 US $ [Malekan & Bonyan, 2010]; [Futron Corporation Manual, 2002]. Consequently, there is an ever-increasing interest within the satellite design community to provide the same level of functionality via lighter equipments, thus avoiding

[1] From a systems engineering point of view, all the three said items can be translated into dollars. Generally speaking, lighter, less power-hungry and smaller simply means cheaper!

high launch costs. Also, purchase of solar cells required to generate 1 watt in orbit may be as expensive as 2500-3000 US $ [Larson & Wertz, 1992]. The typical prices are given here in order to help the reader realize the desire within the space community to provide the same level of functionality via equipments consuming less power. It is being concluded that any progress within the preceding arenas can be regarded as saving millions of dollars.
Also, equally important, the unique feature of present and potential progress of microsatellite missions lies within the recent pattern of quality assurance developed within the microsatellite design community. Historically, quality programs applied in space programs have been rigorous and expensive. Also due to vastly-unknown nature of space environment, only few highly-qualified technologies have flown on space missions. Today, however, by the means of methods developed and/or established in the last two decades such as "qualification by similarity","Configuration control" and so on, much more responsive and cheaper qualification programs are available. Although these programs are not as precise as their predecessors, they still provide the required insight and confidence level required in most microsatellite programs. Also, due to the courageous microsatellite missions within the past, more components have been "space-qualified". At this step, the author would like to draw the readers' attention to the very point that, traditionally, there has been a considerable delay-gap in the technology-level utilized in space technology in comparison with commercial units available in the every-day market. As an example, in a microsatellite program, it is the ultimate wish of a Command and Data Handling (C&DH) designer to be able to incorporate a computer unit with equal capabilities as of a home-based Pentium-5. This delay-gap, however, is shrinking due to the recent missions accomplished mostly by top-class universities in US, Europe, Asia and Africa [Kitts & Lu, 1994]; [D.C.Maessen et al., 2008]; [Sabirin & Othman, 2007]; [Triharjanto et al., 2004 ]; [Kitts & Twiggs, 1994]; [Annes et al., 2002]. As a consequence, the technology-level of components employed in microsatellite technology is reaching that of hi-tech commercial market. Having considered the 10-20 years delay-gap of the space-qualified components and hi-tech COTS technologies, the importance of the new approach may be better understood.
As a conclusion, Table 1 compares the system-level capabilities of microsatellites in the past and at the present/near-future.

## 3. System and subsystem performance requirements

In this section, current status and future trends of various subsystems of microsatellites are discussed. Also, mutual effects of foreseen improvements of each subsystem on system performance are studied.

### 3.1 Payload mass ratio to total satellite mass

A satellite payload is the main reason to launch the whole vehicle. Thus, from a top level point of view, the more ratio of payload mass to total satellite mass (PM/TSM), the better. In the first years of microsatellite re-appearance, limited PM/TSM was practically achievable. Today, however, with the ever-increasing progress in microsatellite technology, PM/TSM as high as 10-25% is achievable, at the present and in near future, respectively. Furthermore, at the present, more capable payloads are being developed and supplied at reasonable prices, in a non-military, non-governmental market. Thus, for a given PM/TSM, currently-available

| | capabilities in the past | capabilities at the present and near future |
|---|---|---|
| Payload mass ratio to total satellite mass (PM/TSM) | 5-10% | 10-25% at present and near future |
| In-orbit autonomy (self-containment of the satellite) | Very poor performance | Moderate- rapid advancement expected in near future |
| Attitude knowledge and control | Less than 10 deg | Better than 1 deg at present |
| Attitude maneuverability | None-Very limited | Satisfactory- rapid advancement expected in near future |
| Resolution (in remote-sensing systems) | Tens of meters to a few hundreds of meters | Less than 10 meters feasible at present |
| Onboard available power | Less than 70 W, thus excluding power-hungry applications | 100-200 W, advancement expected in near future |
| Computer (Hardware and software) | Very limited | Fairly satisfactory for current applications- rapid advancement expected in near future |
| Communication architecture | Mostly limited to UHF/VHF/L/S-Band - Low bit rates typically less than 1 Mbit/sec | High bit rates up to 10 Mbit/sec at various frequencies |
| Propulsion | None | Very limited- practical applications not foreseen in near future |
| Thermal control | Conventional applications feasible | Conventional, more demanding applications feasible - Special applications such as cryogenic missions not foreseen in the few next years |

Table 1. System-level capabilities of microsatellites in the past and at the present/near-future

payloads offer several-times better performance in comparison with their predecessors. Having considered the combined effect of the two preceding considerations, one may appreciate the potential applicability and ever-increasing interest of various communities in microsatellite technology. As an instance, Surrey Satellite Technology Ltd (SSTL) provides light-weight optical, navigation and communications payloads at exceptionally low prices [SSTL website, as of 2009]. A few of these capable payloads will be introduced in the following paragraphs.

### 3.2 Microsatellite in-orbit autonomy

Highly-autonomous satellites are defined as those vehicles requiring minimum contact with external sources (Terrestrial and/or Spaceborne) to successfully accomplish their intended missions [H.Bonyan, 2007]. Most microsatellites are placed in LEOs, and communications gaps (time-intervals with no contact opportunity) are inherent characteristics of LEOs. Thus, logically, a given level of in-orbit autonomy must be accommodated within the orbiting vehicle to perform mission-specific tasks, when out of ground station visibility. Accommodation of a given level of onboard autonomy is a sophisticated systems engineering activity confined by inherent mass-/power-budget constraints of microsatellite missions and also by LEO characteristics. For a microsatellite mission, once in orbit, it is

required to autonomously perform various self-management and mission-specific tasks, to be utilized efficiently.
To some extent, autonomy issues of microsatellites have been ignored during the last decades. Consequently, there is little literature available on the preceding issues. However, rapid advancement is foreseen in near future. For further studies, the interested reader is referred to [Farmer & R.Culver, 1995]; [A.Kitts, 1996]; [A.Swartwout & A.Kitts, 1996]; [A.Kitts & A.Swartwout, 1997]; [E. Vicente-Vivas, 2005]; [H.Bonyan, 2007]; [H.Bonyan & A.R. Toloei, 2009].

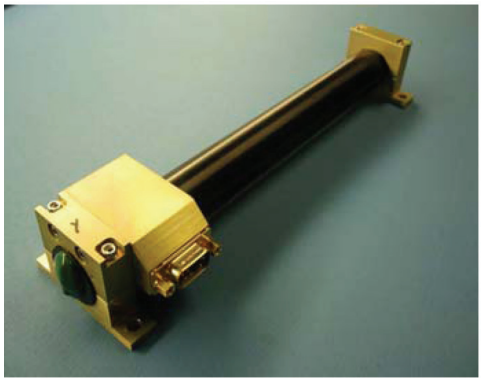
### 3.3 Attitude knowledge and control

Accurate attitude knowledge and control is a crucial requirement for most practical satellites. For most remote sensing applications, one of the most promising microsatellite applications from a financial-benefit point of view, highly-accurate Attitude Determination and Control Subsystem (ADCS) is required. Lack of accurate three-axis, stabilized control capability has been a challenging obstacle in economical profitability of microsatellites. However, with microsatellites like LAPAN-TUBSAT and a few others already in orbit and many others on their way to orbit, this obstacle is already a part of history[2]. Today, three-axis control with accuracies better than 1 degree are viable within microsatellite stringent monetary and mass/power/volume-budgets. Higher accuracies i.e. arc-min or better, are not foreseen in the near future. A few of the SSTL and Sun Space and Information Systems (Pty) Ltd. (Sun Space) attitude sensors and actuators are given below.

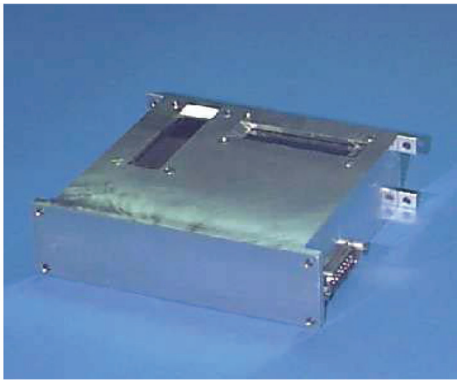
(a) (b)

Fig. 1. SSTL lightweight, yet capable Microwheel 10SP-M, for three-axis control systems (a) and Sun Space reaction wheel with built-in electronics (b)

[2] Lack of reliable three-axes ADCS has been a major reason regarding inefficient power-generation capability of microsatellites [Bonyan & Toloei, 2009]. With this problem already removed, 2-3 times enhancement is foreseen in power generation capability of microsatellites.

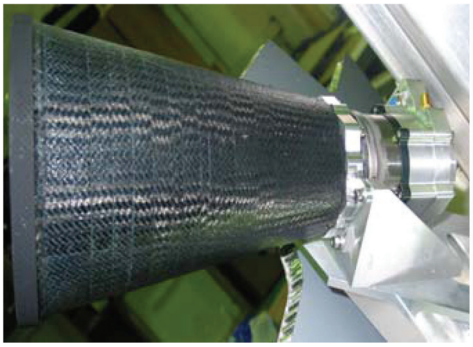


(a)



(b)

Fig. 2. The MTR-5, one of three magnetic torque rods available from SSTL (a) and SSTL 2-axis DMC sun sensor (b)



(a)

(b)

Fig. 3. SSTL star tracker (a) and Sun Space star tracker (b), both space-qualified

### 3.4 Attitude manoeuvrability

In our terminology, hereafter, attitude manoeuvrability is defined as the ability and agility of the vehicle to align itself into a new desired orientation. Attitude manoeuvrability has been traditionally one of the most demanding and challenging in-orbit activities, possible only in complicated several-hundred kilogram satellites. However, recently, microsatellites have proved their capability to accomplish demanding missions performing sophisticated Attitude manoeuvres. Thus, now, microsatellites can be scheduled to "*look*" into a certain direction, when over a desired location. This capability gives the operators much more flexibility to answer a user's requests in a more rapid and responsive fashion [Triharjanto et al., 2004]. With this in mind, another strategic shortcoming of microsatellite applications has been removed.

### 3.5 Resolution (in remote-sensing systems)

Remote sensing applications are among the most promising applications of LEO microsatellites[3]. The main requirement of such systems comes in the form of spatial resolution or GSD (Ground Sample Distance). Most practical, financially-valuable applications require GSDs on the order of (or better than) tens of meters[4], previously viable only by large satellites. Today, and/or in near future, remote sensing applications requiring resolutions as good as 5-10 meters, with frequent revisit times from a few days to a few weeks (Agriculture, Disaster monitoring, Urban planning; water resource managements, off-shore activities monitoring, to name only a few) are well within microsatellite capabilities[5][T.Bretschneider, 2003]; [U.Renner & M.Buhl, 2008].

### 3.6 Onboard available power

The general progress within all engineering fields holds true for electrical power subsystem of microsatellites, as well. Nowadays, more efficient power generation, storage and distribution hardware and software are available within the commercial space market. Thus, generally speaking, recent microsatellites are more capable compared to their predecessors, from an electrical power subsystem point of view. As an example, The SSTL high-efficiency (19.6 %) and very-high-efficiency solar panel and solar cell assembly is shown in fig 4.

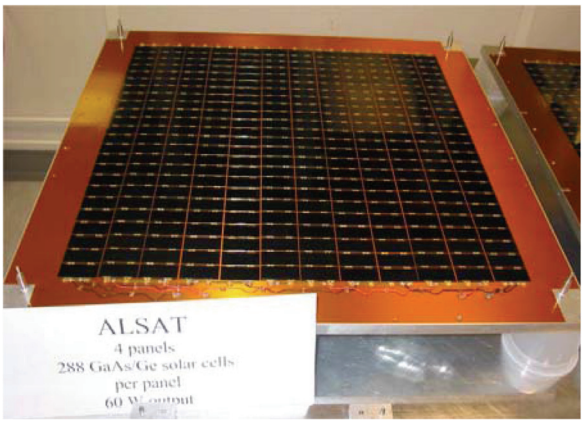



Fig. 4. SSTL solar panel and solar cell assembly

[3] LEO is the main domain of microsatellite missions. This has been due to low launch costs and limited capabilities of microsatellites. Although essential progress is foreseen in microsatellite technology, it is being anticipated that LEO will still serve as the main domain of microsatellite missions, due to its favourable characteristics.

[4] There are certain financially-valuable applications which require GSDs on the order of tens to hundreds of meters. Thus, the typical milestones are given for a basis of comparison and better understanding of current status and future trends.

[5] It must be reminded that low data rates has been an off-putting drawback in microsatellite applications. Generally speaking, whatever mission-data obtained onboard the spaceborne vehicle must be transmitted to earth with reasonable time-delay to be financially-valuable. Non-real-time communication applications, yet Mbit-order data rates are now affordable within stingy mass-power- budget of microsatellite missions.

However, Most microsatellite missions, even in recent years, have been confined to some low-power applications [United Nations UNISPACE III, 1998]; [Bonyan, 2007]. This is, in turn, due to nature of microsatellite missions, and highly-inefficient, in-orbit configuration of microsatellites i.e. a cube with solar cells attached to external facets. A study by the same author in 2009 proved the inefficient conventional in-orbit configuration of microsatellite missions, in terms of power generation, indicating that available power level of most microsatellite missions has been as low as 50-70 watts or less. There, however, are evidences that in near future power level may be enhanced by a factor of 2-3. Thus, applications such as high data-rate communications and/or sophisticated imaging techniques in various frequency bands are well within capabilities of current and near future microsatellite missions, from a power-consumption point of view. The interested reader is referred to a study by the same author, in 2009, dealing with the subject in more detail [H.Bonyan & A.R.Toloei, 2009].

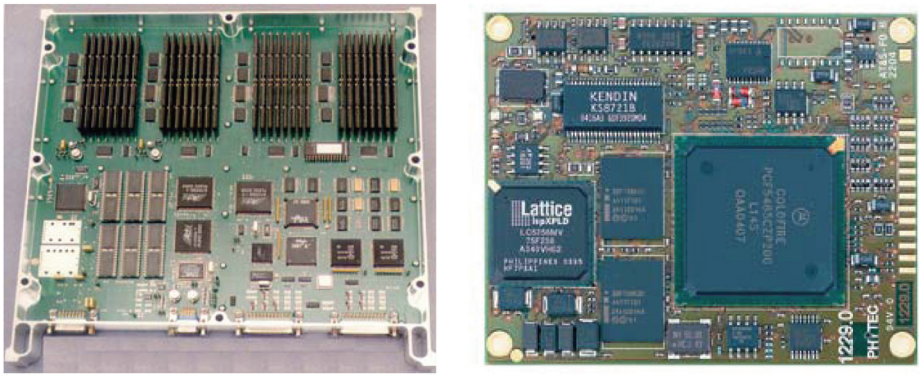
### 3.7 Command and Data Handling (C&DH)

Computer capability has been very limited in microsatellite technology. This has been partially due to the painstaking qualification process inherent in space projects, dominant in the previous century. Thus, although Personal Computer (PC) technology has experienced astonishing advancements in the last two decades, there still remains much effort to accommodate the already-available technology level into microsatellite missions. Fortunately, there are evidences of rapid progress within the field. This is mainly due to:

- Courageous hi-tech microsatellite missions accommodating more capable computer hardware components, thus space-qualifying "hi-tech" items
- A more-relaxed power-budget allocation for the C&DH subsystem
- Better understanding of space environment and maturation of software programs
- New less-demanding qualification processes established,
- Introduction of various non-governmental organizations providing hi-tech computer hardware and software items,

For further detail, see [SSTL website, as of 2009]; [PHYTEC website, as of 2009]; [Freescale semiconductor website, as of 2009]; [A.Sierra et al., 2004]; [A.Woodroffe & P.Madle, 2004]; [R.Amini et al., 2006].

The SSTL general-purpose Intel 386-based C&DH unit and phyCORE-MCF5485 SOM Module from PHYTEC are shown in fig. 5.




(a) (b)

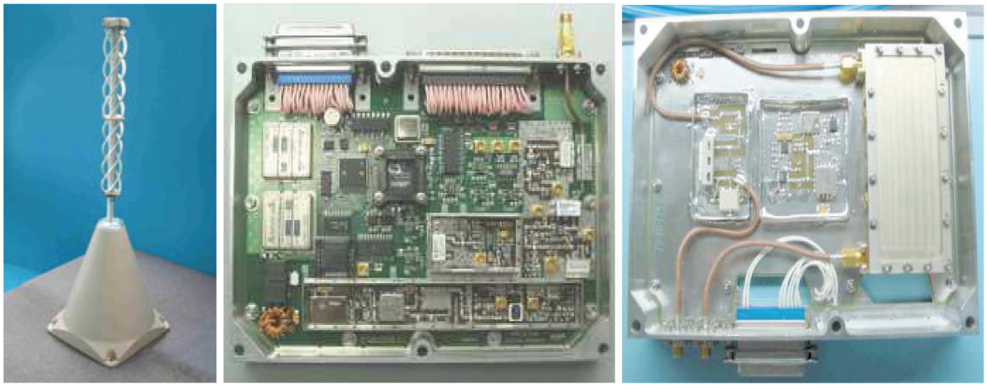
Fig. 5. SSTL OBC 386 (a) and phyCORE-MCF5485 SOM Module from PHYTEC (b)

### 3.8 Communications architecture

Generally speaking, whatever mission-data gained onboard the satellite, must be transferred to earth for further added economical-value. This can be interpreted as a requirement of high data-rate communications systems. Previous microsatellite missions have suffered much from lack of such systems. Today, and in near future, there will be order-of-magnitude improvements in such systems. This is mainly due to the following points:

- Communications systems, specifically those onboard the microsatellite, have considerably matured by thorough understandings provided by previous microsatellite missions. Also, ground-station technology regarding microsatellite applications has greatly advanced during the last few decades. At the present time, affordable ground station may be established at fairly-short time intervals, providing communications in various frequency bands [F.B.Hsiao et al., 2000]. Also, having fully comprehended the necessity of international cooperation and mutual benefits for all contributors, the number of joint projects in which several ground-stations are employed for a given microsatellite missions is greatly increasing [A.Kitts & A.Swartwout, 1998 ]; [R.H.Triharjanto et al., 2004]; [D.C.Maessen et al., 2008], [Hasbi et al., 2007]; [D.C.Maessen et al., 2009]. This issue has been studied by the same author in 2007 and 2009, [H.Bonyan, 2007]; [H.Bonyan & A.R.Toloei, 2009].
- A crucial pre-requisite of high data-rate communications is provision of a required level of electrical power. In microsatellite applications, it has rarely been possible to provide enough power to accommodate the power-hungry hardware required for such purposes. However, as mentioned previously, there is going to be a several-times enhancement in onboard available power of microsatellites. This can be interpreted as provision of much more electrical power to be fed into communications hardware, thus much higher data-rates.

SSTL S-band communications hardware is shown in fig. 6.



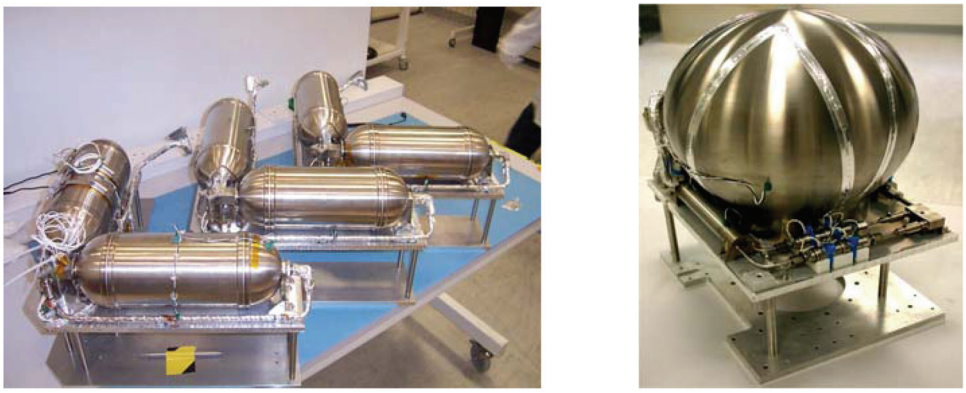
(a) (b) (c)

Fig. 6. SSTL S-Band Quadrifilar Helix Antenna (a), S-Band transmitter (b) and S-band receiver down-converter module (c)

### 3.9 Propulsion

Historically, microsatellites have not been equipped with propulsion systems. Although there have been experiences of carrying propulsion systems onboard microsatellites, these

experiences have been mainly for technology-demonstration and space-qualification purposes. Realistically, practical applications of onboard propulsion systems for microsatellites are not foreseen in near future. However, it must be reminded that lack of propulsion system does not sully unique features of microsatellite technology. By some rough calculations, microsatellites placed in orbits with altitudes higher than 700-800 km will maintain their orbit with sufficient accuracy for periods of up to 5-10 years, well covering lifetime of most microsatellite missions, typically 3-5 years[6]. Some SSTL propulsion hardware, adequate for microsatellite missions, is shown in fig. 7.



(a) (b)

Fig. 7. SSTL microsatellite Butane propulsion system (a) and SSTL Xenon Propulsion System

### 3.10 Thermal control

It is the ultimate goal of thermal-control designers to passively meet the thermal requirements of a satellite. In case of conventional applications of microsatellites, passive and/or semi-active thermal control systems have proved to be more than adequate. On the contrary, accommodating specific payloads, such as cryogenic ones which impose severe thermal requirements, are not foreseen in the next few years. However, in near future, such systems may be accommodated within courageous microsatellite missions, if economically-justified.

## 4. Construction

Compared to the first days of space age, there have been general enhancements in construction techniques of all aerospace vehicles, and microsatellites are not exceptions. For

[6] The Dutch TU Delft and Chinese Tsinghua University are planning an ambitious formation flying mission, to be launched by the end of 2011. During the 2.5-year mission, the two satellites, FAST-D (being developed in Delft) and FAST-T (being developed in Beijing), will demonstrate various new technologies such as autonomous formation flying with distributed propulsion systems and MEMS technology to optimize propellant consumption. This mission, if successful, can effectively propose "propulsion subsystem" as a feasible feature onboard microsatellites.

most aeronautical applications, national and/or international organizational construction facilities have been established in order to provide the required services and to simultaneously reduce construction costs of each individual project. This holds partially true for large space projects, as well. However, in case of microsatellites, such national/international facilities have just been or are being established. Thus, once fully deployed, these facilities may reduce construction costs of microsatellite projects, considerably. Also, with the increasing demand for microsatellite-suited applications, there will be more investment for such facilities, nationally and internationally. Having considered the two preceding issues and increasing number of such facilities and services supplied at each facility, it is being concluded that construction of microsatellites will be faster and more affordable in near future.
A few of these construction-facilities/suppliers are listed in Table 2.

## 5. Test

Test philosophies, in space applications, have been always rigorous, time-consuming and expensive. Qualification/Acceptance tests usually impose too conservative constraints on space applications. However, with the experiences gained during the several decades of space age, there seems to be a trend toward less demanding test procedures. Therefore, the "sacred" test programs which were more than mandatory for all space programs are now simply "Negotiable". It must be reminded that there exists a fragile difference between "negotiable requirements" and simply ignoring the handy qualification tools developed through many years and at the cost of many lives and dollars.
As already mentioned, during the 60s-90s, test programs were more than demanding. Any space project had to go through various time-consuming and very expensive test programs to acquire the launch "go-ahead". This was mainly due to:

- Lack or uncertainty in knowledge of space environment
- Extensive uncertainty and immaturity in analysis software
- Lack of software-in-the-loop simulation techniques
- Astronomical financial budgets supplied by the governments
- Lack of self-confidence within the space community in their products (more and more tests were desired to see if it "really works")

Today, however, due to the experience gained in previous years, with the help of software packages available at very low costs and confidence developed within space community and due to shortage in financial resources, less demanding, yet consistent, test programs have been established.

## 6. Launch

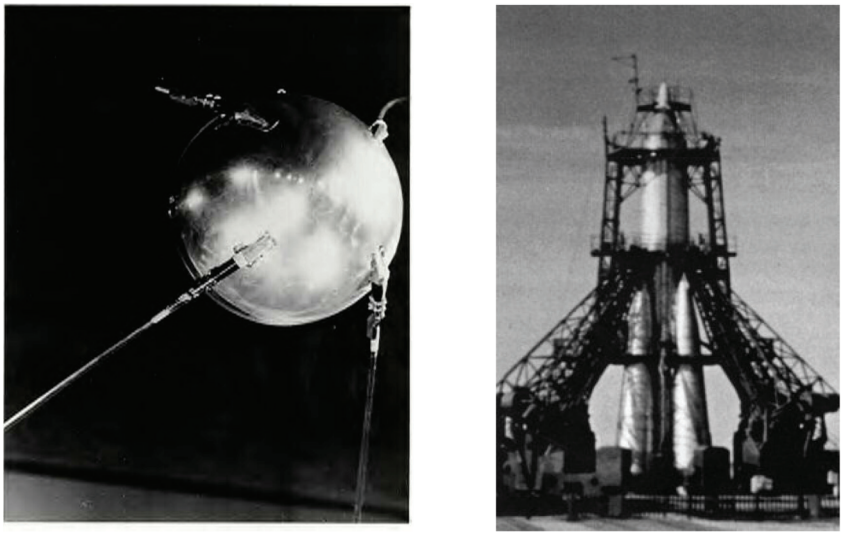
Not exaggerated, launch is what makes the space expensive. According to [H.Bonyan, 2008], launch cost can make up to 50% of the whole project cost. To be truly capable of analyzing the expensive launch phase, a brief history of the issue must be presented:
What is today known as space launch began in 1957, by the Soviet Union. However, the first engineering efforts may be traced back to the second world war (WW II), in which, the German army successfully developed the first ballistic missile i.e. the V2. After the WW II, the soviets and Americans both made all their effort to acquire the German ballistic missile technology. The Americans were the first to arrive and they got the chief designer Werner

| | Country | Products/Services |
|---|---|---|
| Actel Corporation | US | field programmable gate arrays (FPGAs) based on anti-fuse and flash technologies and embedded programmable gate arrays (EPGAs) based on SRAM technology, as well as high-performance intellectual property (IP) cores, software development tools and design services |
| AeroAstro Inc | US | Radios, Imagers, star trackers, sun sensors, communications hardware, etc |
| Atmel Wireless & Microcontrollers | US/ International | space qualified ICs |
| Carlo Gavazzi Space (CGS) | Italy | ISS utilization, instrumentation and processors for scientific and application payloads, instruments for microgravity experiments, small satellites, attitude control and modeling, mission analysis, on-board automation |
| CSEM | Switzerland | systems engineering and design, mechanical engineering, wavefront engineering, micro-imagers, robotic systems, optical systems for high accuracy pointing, smart structures for artificial damping of vibrations, micro/nano technologies for small satellites |
| Fokker Elmo | Netherlands/ International | wiring harnesses, magnetic torque rods, printed circuit boards, multi layer insulation |
| Honeywell Inc | US/ International | guidance, navigation and control systems for launch vehicles and spacecraft; data handling management and processing systems; space instrumentation and sensors; on board data processors; display systems; flight and engine controls for manned space flight; satellite mechanisms; precision inertial instruments; radiation hardened memories; re-entry vehicles; reaction wheel assemblies; antenna pointing systems |
| Matra Marconi Space | International | communications, earth observation and scientific satellites; structures and mechanisms; attitude and orbit control systems; robotics; optical communications; propulsion systems; sterling cycle cryogenic coolers; on board data handling; antennas; optics; sensors; power systems; microwave and millimetre wave components; hardened ASICs; radiometers |
| Megsat | Italy | microsatellite subsystems; high efficiency solar arrays; power subsystems; mechanical devices; actuators; antennas for microsatellites |
| SpaceDev Inc | US | design, manufacture, marketing and operation of sophisticated micro and nano satellites, hybrid rocket-based orbital Maneuvering and orbital Transfer Vehicles (MTVs) as well as safe sub-orbital and orbital hybrid rocket-based propulsion systems; commercial hybrid rocket motors and small high performance space vehicles and subsystems |
| Spectrum Astro | USA | small spacecraft buses, space electronics and power management systems, electrical ground support and ground control equipment |
| SunSpace & Information Systems | South Africa | high performance micro-satellites and imagers |
| SSTL | UK | manufacturer of mini and microsatellites |
| Swedish Space Corporation (SSC) | Sweden | system integrator and supplier of small and micro-satellites; develops attitude orbit & control systems, onboard data handling units, software, propulsion systems and ground systems for testing and mission control; extensive operational experience |

Table 2. some construction-facilities/suppliers in various countries[7]

[7] This table is not intended to provide a complete list of construction-facilities/suppliers and is only meant to name a few. A comprehensive list can be found at EPPL (European Preferred Part List ), Issue 13; Issue Date: 2008-09-12

Von Braun with his most outstanding men. Von Braun was secretly sent to US and was soon the most influential designer in the US ballistic missile program. The soviets, on the other hand, got most of the technicians and the hardware and soon moved them to the soviet territories. The reader's attention is drawn to the point that the US had already acquired the atomic-bomb (A-bomb) technology during the WW II, under the project named The Manhattan Project in which they tested an A-bomb in the US remote territories and dropped two on the Japanese cities. The soviets, thanks to their covert intelligence network within the US territory, had their hands on the A-bomb technology as soon as a year after the WW II. Thus, with the A-bomb technology available, both nations looked for the necessary delivery-means to launch their A-bombs over great distances. The necessity for the delivery-system for A-bombs was the main driver for the InterContinental Ballistic Missile (ICBM) programs, in both east and west. The soviets, however, made much more effort due the fact that they had almost no means of A-bomb delivery to the US territory while the US had its A-bomb-equipped aircrafts in several Air-Force-Bases (AFB) located in lands near to or neighbouring the soviet territory[8]. Thus the soviets made all their effort to counterbalance the condition and in 1957, the efforts came to reality. They, however, picked up a much more dramatic approach to prove their ICBM capability. Since a ballistic missile has much in common with a LEO launch vehicle, the soviets placed the first man-made satellite Sputnik-1 to prove their ICBM capability, in 1957. Spunik-1 and R-7, the first ICBM/launch-vehicle in the world, are shown in figure 8.



(a) (b)

Fig. 8. First artificial satellite Sputnik-1 (a) and soviet R-7 launch vehicle with Sputnik 2 as its payload (b)

[8] This, in turn, was because of due-east policy of the NATO (North Atlantic Treaty Organization), during the cold war. US, by the first years of 1960's, had its NATO AFBs in several European countries, near to or neighbouring the soviet union.

During the first two decades after the first space flight, Launch-campaign was an issue mainly influenced by the political and military drivers. From 1980s and afterwards, specifically after collapse of the Soviet Union, cost has been considered as the main parameter in space launch community. In 1990s and after the Cold War, LEO launch was as expensive as 30,000 US $/Kg which was still quite high for many nations and organizations[9]. The following table summarizes launch cost per pound (kilogram) for different small launch vehicles (5,000 lbs. or less to LEO), as of 1990-2000.

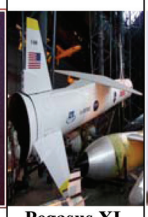
| |  Athena 2 |  Cosmos |  Pegasus XL |  Rockot |  Shtil[10] |  Start |  Taurus |
|---|---|---|---|---|---|---|---|
| **Country/ Region of Origin** | USA | Russia | USA | Russia | Russia | Russia | USA |
| **Reference Site and Inclination** | CCAFS 28.5 deg. | Plesetsk 62.7 deg. | CCAFS 28.5 deg. | Plesetsk 62.7 deg. | Barents Sea 77-88 deg. | Svobodny 51.8 deg. | CCAFS 28.5 deg. |
| **Reference LEO Altitude mi (Km)** | 115 (185) | 249 (400) | 115 (185) | 186 (300) | 124 (200) | 124 (200) | 115 (185) |
| **LEO Capacity lb (Kg)** | 4,520 (2,065) | 3,300 (1,500) | 976 (443) | 4,075 (1,850) | 947 (430) | 1,392 (632) | 3,036 (1,380) |
| **Estimated LEO payload cost per lb (Kg)** | $5,310 ($11,622) | $3,939 ($8,667) | $13,832 ($30,474) | $3,313 ($7,297) | $211 ($465) | $5,388 ($11,687) | $6,258 ($13,768) |

Table 3. Launch cost per kilogram for different Small (5,000 lbs. or less to LEO) launch vehicles, as of 1990-2000

At this moment, it must be reminded that after collapse of the Soviet Union, there has been a general tendency within the Russian space sector to utilize the inherited, already-available InterContinental Ballistic Missile (ICBM) infrastructure for commercial space launches in order to raise extra funding for space activities and partially avoid high maintenance costs

[9] This is, in turn, one of the main reasons why microsatellites have gained more and more attention during the last 2-3 decades. Microsatellites are obviously much cheaper and quite affordable to be launched into orbit, compared to conventional large satellites.

[10] Shtil launch costs are partially subsidized by the Russian Navy as part of missile launch exercises

of such systems. These new launchers, also known as converted-ICBMs, offer inexpensive and frequent launch opportunities to various space communities. It is being anticipated that in the next decade, there will be frequent and affordable launch opportunities provided by the Russian space-launch market. The following derivatives of the soviet ICBMs now serve as launch vehicles:

1. Rockot (Based on SS-19 ICBM; flight proven more than 140 times)
2. Shtil (a derivative from R-29-family of submarine-launched ballistic missiles)
3. Dnepr ( based on RS-20 ICBM; SS-18 Satan by NATO designation)
4. Start (based on RT-2PM Topol, NATO reporting name: RS-12M Topol ICBM)
5. Strela (based on UR-100 ICBM, NATO reporting name SS-11 Sego)
6. Tsyklon (based on R-36 ICBM, NATO reporting name SS-9 Scarp)
7. Volna (based on R-29R submarine-launched ballistic missiles)

Current Status and Future Trends of Russian Space-Launch Market is being addressed by the same author in a separate paper [M.Malekan & H.Bonyan, 2010].

Table 4 summarizes launch cost per pound (kilogram) for different medium (5,001-12,000 lbs. to LEO) and intermediate (12,001-25,000 lbs. to LEO) launch vehicles, as of 1990-2000.

| |  Ariane 44L |  Atlas 2AS |  Delta 2 (7920/5) |  Dnepr |  Long March 2C |  Long March 2E |  Soyuz |
|---|---|---|---|---|---|---|---|
| **Country/ Region of Origin** | Europe | USA | USA | Russia | China | China | Russia |
| **Reference Site and Inclination** | Kourou 5.2 deg. | CCAFS 28.5 deg. | CCAFS 28.5 deg. | Baikonur 46.1 deg. | Taiyuan 37.8 deg. | Taiyuan 37.8 deg. | Baikonur 51.8 deg. |
| **Reference LEO Altitude mi (Km)** | 124 (200) | 115 (185) | 115 (185) | 124 (200) | 124 (200) | 124 (200) | 124 (200) |
| **LEO Capacity lb (Kg)** | 22,467 (10,200) | 18,982 (8,618) | 11,330 (5,144) | 9,692 (4,400) | 7,048 (3,200) | 20,264 (9,200) | 15,418 (7,000) |
| **Estimated LEO payload cost per lb (Kg)** | $5,007 ($11,029) | $5,136 ($11,314) | $4,854 ($10,692) | $1,548 ($3,409) | $3,192 ($7,031) | $2,467 ($5,435) | $2,432 ($5,357) |

Table 4. Launch cost per kilogram for different medium (5,001-12,000 lbs. to LEO) and intermediate (12,001-25,000 lbs. to LEO) launch vehicles, as of 1990-2000

Finally, Launch cost per kilogram for different heavy (more than 25,000 lbs. to LEO) launch vehicles is shown in table 5, as of 1990-2000.

| |  Ariane 5G |  Long March 3B |  Proton |  Space Shuttle |  Zenit 2 |  Zenit 3SL |
|---|---|---|---|---|---|---|
| **Country/ Region of Origin** | Europe | China | Russia | USA | Ukraine | Multination al |
| **Reference Site and Inclination** | Kourou 5.2 deg. | Xichang 28.5 deg. | Baikonur 46.1 deg. | KSC 28.5 deg. | Baikonur 46.1 deg. | Odyssey Launch Platform 0 deg. |
| **Reference LEO Altitude mi (Km)** | 342 (550) | 124 (200) | 124 (200) | 127 (204) | 124 (200) | 124 (200) |
| **LEO Capacity lb (Kg)** | 39,648 (18,000) | 29,956 (13,600) | 43,524 (19,760) | 63,443 (28,803) | 30,264 (13,740) | 34,969 (15,876) |
| **Estimated LEO payload cost per lb (Kg)** | $4,162 ($9,167) | $2,003 ($4,412) | $1,953 ($4,302) | $4,729 ($10,416) | $1,404 ($3,093) | $2,431 ($5,354) |

Table 5. Launch cost per kilogram for different heavy (more than 25,000 lbs. to LEO) launch vehicles, as of 1990-2000

The price-per-pound (kilogram) figures in the previous tables vary significantly from a launch vehicle to another. From the preceding tables, it is concluded that the non-western (Russian/Ukrainian, Chinese) vehicle offer lower prices than their western counterparts (American and European), primarily because of lower labour and infrastructure costs. The following table shows that these differences in average price-per-pound can be significant [Futron Corporation Manual, 2002].

| Vehicle Class | Average Price-per-pound @ LEO ($/lb) | |
|---|---|---|
| | Western | Non-Western |
| Small | 8,445 | 3,028 |
| Medium/Intermediate | 4,994 | 2,407 |
| Heavy | 4,440 | 1,946 |

Table 6. Average Price-per-pound for Western and Non-Western Launch Vehicles[11], as of 1990-2000

[11] The Zenit 3SL is considered a non-Western launch vehicle because of its Ukrainian and Russian heritage.

## 7. Post-launch operations

Post-launch (in-orbit) operation of microsatellites has been vastly ignored, both in practice and in the literature, until very recently. During the last decade, however, the significance of the issue has been highlighted by various communities and is evolving rapidly [R.Annes et al., 2002], [Hardhienata et al., 2005], [H.Bonyan 2010]. There, however, still remain certain shortcomings regarding in-orbit operations of microsatellites. In fact, most involved-parties are reluctant to officially declare inefficient in-orbit utilization of their microsatellites. Without referring to any specific project, it is being highlighted that according to the author's studies, there are several cases in which fully-operational microsatellites have been almost abandoned in orbit due to poor in-orbit operations strategy. These crafts could have provided invaluable services, with considerable financial benefit, if adequate short- /long-term in-orbit operations strategy had been carefully planned. It is being reminded that in the next decades, microsatellites will not only serve as hands-on experience to train university students and to be financially-valuable, much attention must be paid to the in-orbit operations of such vehicles.

## 8. Reference

R.Amini, Gerard Aalbers, Rob Hamann, Wim Jongkind (2006), New Generations of Spacecraft Data Handling systems Less Harness more Reliability, 57th International Astronautical Congress, Valencia, Spain

Annes et al., (2002), Operation Delfi – A Space Mission Development Project, 17th AMSAT-UK Colloquium, Guildford, Surrey, England, UK

H.Bonyan (2010). An in-depth Analysis of the Ambiguity of Economical-profitability of Microsatellite Missions, accepted for publication in IEEE Aerospace conference Big Sky, MT USA

H.Bonyan (2010), Efficient In-orbit Operations of LEO Polar/Sun-synchronous Satellites; Southern Hemisphere Revisited?, accepted for publication in IEEE Aerospace conference Big Sky, MT USA

H.Bonyan and A.R.Toloei (2009), Systems Engineering Analysis of Required Level of On-orbit Autonomous Operation of a LEO Student-microsatellite Mission, Recent Advances in Space Technology conference (RAST 2009), Istanbul, Turkey

H.Bonyan and A.R.Toloei (2009), Systems Engineering Approach toward the Problem of Sunlight Collection of a Light-micro Satellite, Recent Advances in Space Technology conference (RAST 2009), Istanbul, Turkey

H.Bonyan (2008), Investigation and Utilization oF the Low-earth Equatorial Orbits for Missions Concerning the African Continent, AIAA & IEEE joint conference, Big Sky, USA

H.Bonyan (2007), System engineering approach toward the problem of battery depth-of-discharge of a LEO satellite, International Conference on Complex Systems (ICCS) Quincy MA USA

H.Bonyan (2007), Systems Engineering Approach toward the Problem of Required Level of In-orbit Autonomous-operation of a LEO Microsatellite Mission, International Conference on Complex Systems (ICCS) Quincy MA USA

T. Bretschneider, S.H. Tan, C.H. Goh, K. Arichandran, W.E. Koh, E. Gill (2003), X-SAT Mission Progress, 5th IAA Symposium on Small Satellites for Earth Observation IAA-B5-0504, Berlin, Germany

Farmer, Mike and Randy Culver (1995), The Challenges of Low-Cost, Automated Satellite Operations, Proceedings of the 31st International Telemetering Conference, Las Vegas, Nevada, pp. 203-209

Futron Corporation Manual (2002), Space Transportation Costs: Trends in Price per Pound to Orbit 1990-2000

E. Gill et al. (2008), Atmospheric Aerosol Characterisation with the Dutch-Chinese FAST Formation Flying Mission, IAC-08-B1.I.1, In Proc. of the 59th IAC, Glasgow, Scotland

G. Grillmayer et al, (2003), ILSE – First Laboratory Model of the Small Satellite Program at the University of Stuttgart 54th International Astronautical Congress, Bremen, Germany

S. Hardhienata, A. Nuryanto, R. H. Triharjanto, U. Renner (2005), Technical Aspects and Attitude Control Strategy of LAPAN-TUBSAT Micro Satellite, 5th IAA Symposium on Small Satellites for Earth Observation, Berlin, Germany

W.Hasbi, E.Nasser, A.Rahman (2007), Spacecraft Control Center of Lapan-Tubsat Micro Satellite, 3rd Asian Space Conference, NTU@one-north campus, Singapore

Fei-Bin Hsiao ,Hui-Ping Liu , Chung-Cheng Chen (2000), The Development of a Low-Cost Amateur Microsatellite Ground Station for Space Engineering Education, Global J. of Engng. Educ., Vol.4, No.1 Printed in Australia

http://www.freescale.com/webapp/sps/site/homepage.jsp?nodeId=0162468rH3 , last visited July 2009

http://www.ic.gc.ca/app/ccc/srch/nvgt.do?lang=eng&prtl=1&sbPrtl=&estblmntNo=123456114317&profile=cmpltPrfl&profileId=1421&app=sold , last visited July 2009

http://www.phytec.com/ , last visited July 2009

http://www.rockwellcollins.com/about/locations/deutschland/index.html , last visited July 2009

http://www.sstl.co.uk/ , last visited July 2009

http://www.sunspace.co.za/ , last visited July 2009

Kitts, Christopher A. and Michael A. Swartwout (1998), Experimental Initiatives in Space System Operations, In Proceedings of the Annual Satellite Command, Control and Network Management Conference, Reston, VA Presented by Kitts in the Systems and Mission Analysis Session. Also presented by Michael A. Swartwout at the 1998 INFORMS Conference, Monterey, CA, January, 1998, Spacecraft Automation Session

Kitts, Christopher A (1996)., A Global Spacecraft Control Network for Spacecraft Autonomy Research, In Proceedings of SpaceOps '96: The Fourth International Symposium on

Space Mission Operations and Ground Data Systems, Munich, Germany. Presented by Kitts in the Operations Automation Session

Kitts, Christopher A., and Robert J. Twiggs (1994), the Satellite Quick Research Testbed (SQUIRT) Program, In Proceedings of the 8th Annual AIAA/USU Conference on Small Satellites, Logan, Utah

Kitts, Christopher A., and Richard A. Lu (1994), The Stanford SQUIRT Micro Satellite Program, In Proceedings of the AMSAT-NA 12th Space Symposium and AMSAT Annual Meeting, Orlando, Florida, Presented by Robert J. Twiggs at the 1994 AMSAT-NA Space Symposium, Orlando, Florida

Larson Wiley J., James R.Wertz (1992) Space Mission Analysis and Design, pp. 393 & 397

Maessen D.C. et al. (2009), Mission Design of the Dutch-Chinese FAST Micro-Satellite Mission, 7th IAA Symposium on Small Satellites for Earth Observation, Berlin, Germany

Maessen D.C. et al. (2008), Preliminary Design of the Dutch-Chinese FAST Micro-Satellite Mission, In Proc. of the 4S Symposium, Rhodes, Greece

M.Malekan & H.Bonyan (2010), On the Current Status and Future Trends of Russian Space-Launch Market, accepted for publication in IEEE Aerospace conference Big Sky, MT USA

U.Renner, Matthias Buhl (2008), High Precision Interactive Earth Observation with LAPAN-TUBSAT, Proceedings of the 4S Symposium Small Satellites, Systems and Services, Rhodos, Greece

A.Sabirin, M.Othman (2007), Razaksat- High resolution imaging satellite for near equatorial orbit (Neqo) COSPAR/IAF Symposium, "Use of the equatorial orbit for space science and applications: Challenges and opportunities", Vienna, Austria

A.Sierra, Juan J. Quiroga, Roberto Fernández and Gustavo E. Monte (2004), An Intelligent Maintenance System for Earth-based Failure Analysis and Self-repairing of Microsatellites, Acta Astronautica Volume 55, Issue 1, Pages 61-67

Swartwout, Michael A., and Christopher A. Kitts (1997), Automated Health Operations for the SAPPHIRE Spacecraft, In Proceedings of ITC/USA '96: The 33rd Annual International Telemetering Conference, Las Vegas, NV. Presented by Swartwout in the Space Systems Session

Swartwout, Michael A., and Christopher A. Kitts (1996), A Beacon Monitoring System for Automated Fault Management Operations, In Proceedings of the Tenth Annual AIAA/USU Small Satellite Conference, Logon, UT. Presented by Swartwout in the University Student Session. Honourable Mention Winner in Student Paper Competition

Third United Nations Conference on the Exploration and Peaceful Uses of Outer Space, UNISPACE III, 1998

Triharjanto, H.R., et al, (2004), LAPAN-TUBSAT: Micro-satellite platform for surveillance and remote sensing, 4S symposium, Paris, France

E. Vicente-Vivas, Fabián García-Nocetti and Francisco Mendieta-Jiménez (2005), Automatic maintenance payload on board of a Mexican LEO microsatellite, Acta Astronautica Journal, Elsevier Science, Volume 58, Issue 3, Pages 149-167

A. M. Woodroffe and P. Madle (2004), Application and experience of CAN as a low cost OBDH bus system, MAPLD 2004, Washington D.C. USA

# Part V

# 17

# An Aircraft Separation Algorithm with Feedback and Perturbation

White, Allan L.
*NASA Langley Research Center*
*USA*

## 1. Introduction

An open question in air traffic management is whether or not algorithms can be realistically shown to meet Federal Aviation Administration (FAA) safety requirements.
This current work shows such a demonstration is possible for a separation algorithm with perturbations from feedback control and atmospheric turbulence in a local setting. This project can accept a three to four magnitude increase in complexity and still remain viable, but clearly this is not enough of a margin to include every detail in a global analysis. Future work is needed in sensitivity analysis to determine what must be included in the simulation.
Section two contains a probability and statistical analysis of a recent FAA requirement, and it is this analysis that most distinguishes this paper from other efforts. Section three presents the assumptions about the aircraft and flight space. Section four shows that the minimum-distance point determines the relative angle of approaching aircraft, and section five gives a pictorial description of the separation maneuver. Section six gives the precise description of the maneuver and a proof that it maintains separation if no perturbations are present.
The aircraft proceed along their flight paths by means of feedback control, and section seven presents the control equations. The algorithm is simple and generic and needs more development. In its current state, it is intended to represent either control-with-pilot-in-the-loop or future-automatic-control. Once feedback control is introduced, it is possible to include perturbations in a realistic manner, and section eight describes its stochastic nature while section nine offers more commentary. The approach to perturbation in this paper is to examine distributions of increasing severity. If the algorithm can survive these distributions, then it can survive the real world perturbations. The severity of the examined perturbations can be seen in figure 10 in section nine.
This paper does not include a test of any decision algorithm since such a test should include the uncertainty due to instrumentation error where the position and heading of the aircraft are not precisely known.

## 2. Representative FAA requirement

### 2.1 Probability

The stated FAA goals are changing, but this paper addresses a recently expressed goal which was stated in terms of a moving average: no more than three incidents (of all types) over the last three years. Since there are about ten million flights per year, this translates into

one or fewer incidents per 10 million ($10^7$) flights. The examination compares this moving average to a goal stated in terms of one year. There are two comments. First, future FAA goals or their interpretation may be different from the ones examined below, but the examination below outlines an approach to analyzing any stated goal. Second, as they stand, these goals are not probability statements, and they require interpretation.

In the absence of information and for simplicity, the typical assumption is that all flights are equivalent and independent, and the typical interpretation of the goal is that the expected number of incidents for 10 million flights be equal to one. Using the expectation does not require any more information from the FAA, but it does have a disadvantage as will be seen below.

The disadvantage of this interpretation appears when we consider the probability of more than one incident during 10 million flights. It's reasonable to want the probability of more than one incident to be low, but it will be shown that using the expectation-interpretation does not guarantee this. On the other hand, the low-probability approach raises the question of how low the FAA wishes the probability to be.

With the assumption that the flights are equivalent and independent, the distribution is binomial with the probability of an incident equal to $10^{-7}$ per flight. The binomial distribution with parameter p gives the probability of zero or one incidents during 10 million flights as

$$Q = (1-p)^{10000000} + 10000000\, p\, (1-p)^{9999999} = 0.7358 \text{ if } p=1e\text{-}7. \quad (1)$$

The probability of two or more incidents for p = 1e-7 is 1-Q = 0.2642. Hence, if the probability of an incident is equal to $10^{-7}$ per flight, then the probability of more than one incident during 10 million flights is greater than 1/4.

If the goal is a less than one in a hundred chance of more than one incident per ten million flights, then a little numerical work gives that for p = 1.5e-8, Q = 0.9898 and 1-Q = 0.0102.

The moving average reduces the likelihood of not achieving the goal provided the probability of an incident during a flight is smaller than required.

Suppose the probability of an incident is equal to $10^{-7}$ per flight. Then

Prob{more than one incident in a year | p=1e-7} = 0.26.

Prob{more than three incidents in three years | p=1e-7} = 0.35.

Whereas

Prob{more than one incident in a year | p=1e-8} = 0.0047.

Prob{more than three incidents in three years | p=1e-8} = 0.0003.

The crossover point appears to be p=7e-8.

Prob{more than one incident in a year | p=7e-8} = 0.16.

Prob{more than three incidents in three years | p=7e-8} = 0.16.

Returning to the interpretation of the FAA goal as a probability statement, one possibility is that the FAA would desire there is only 1 in N chance the goal not be met. A reasonable choice for N is some number between 10 and 100. Looking at the extremes, the computations below give values for p = probability of an incident during a flight if the requirement is a 1 in N chance the goal not be met.

For N=10:

The Prob{more than one incident in a year } = 0.10 requires p = 5.3e-8.

The Prob{more than three incidents in three years} = 0.10 requires p = 5.8e-8.

For N=100:

The Prob{more than one incident in a year} = 0.01 requires p = 1.5e-8.

The Prob{more than three incidents in three years} = 0.01 requires p = 2.7e-8.

### 2.2 Probabilities and confidence levels for the simulation

A problem in establishing that a loss-of-separation-algorithm meets the FAA goal is that loss-of-separation is one incident among many. Hence, showing that the probability of loss-of-separation during a flight is less than 1e-7 may not be sufficient since there are other incidents and their probabilities accumulate.

The problem is compounded since when studying incidents, especially the prevention of incidents, it is useful to distinguish between the potential for an incident and the incident itself. For instance, two aircraft on a collision course is a potential for an incident, but successful maneuvering will result in no incident. In addition, there may be multiple causes for an incident or an incident may require multiple causes. There may be no cause for alarm if two aircraft are on a collision course unless some malfunction prevents successful maneuvering.

Hence, a precise probability analysis for loss-of-separation requires an encyclopedic knowledge of incidents and their causes which the author, at least, does not currently posses. Nevertheless, an elementary, incomplete analysis can offer some guidance. One approach in such an analysis is to be conservative: in the absence of complete information, use probabilities that overestimate the likelihood of dire events.

We begin with a simplified scenario and then generalize it. Suppose there are K types of incidents. Let $C_i$ be the set of causes for incident i. Let B (for benign) be the set no causes for an incident. The initial simplifying assumption is that the $C_i$ and B partition the set of flight conditions. That is, the intersection of two different sets is empty, and their union is the entire set. This initial simplifying assumption is justified if incidents are rare and flights with more than one incident are rare enough to be ignored.

With this approach, the study of an incident i consists of the study of the effect of the set $C_i$. For instance, for this study of loss-of-separation, the causes are deviations from the flight paths due to feedback control and external perturbations. The realism of the simulation is increased by adding more causes.

Let $P(A_i | C_i)$ be the conditional probability of an incident given that its causes appear. Then we want

$$P(A_1 \mid C_1)\,P(C_1) + P(A_2 \mid C_2)\,P(C_2) + \ldots + P(A_K \mid C_K)\,P(C_K) \le p. \qquad (2)$$

Based on the assumption that there is a positive probability that a flight is routine (no cause for an incident appears), we have

$$P(C_1) + \ldots + P(C_K) < 1. \qquad (3)$$

Using this assumption, one way to accomplish this is to have $P(A_i \mid C_i) \le p$ for all i since this gives

$$P(A_1 \mid C_1)\,P(C_1) + P(A_2 \mid C_2)\,P(C_2) + \ldots + P(A_K \mid C_K)\,P(C_K) \le p\,P(C_1) + p\,P(C_2) + \ldots + p\,P(C_K) \le p\,[\,P(C_1) + \ldots + P(C_K)\,] \le p. \quad (4)$$

The generalization of the above eliminates the partition requirement. That is, different $C_i$ can have a non-empty intersection, allowing for more than one incident per flight. The reasoning above still holds if $P(C_1) + \ldots + P(C_K) \le 1$, which this paper will assume.

There are two cases where the approach above requires modification. First, if the sets $C_i$ have significant overlap, then the probabilities can sum to greater than 1. If a bound for the sum of probabilities is known and it is less than M, then it is sufficient to demonstrate $P(A_i \mid C_i) \le q$ where $q\,M \le 1$, although if there is significant overlap, then the studies will have to examine the probability that a single set of causes produces several incidents.

Second, a scenario that would require a different type of analysis is if a set of causes had a high probability of producing an incident. That is, for some j, $P(A_j \mid C_j)$ cannot be made small. In this case, the alternative is to arrange things so that $C_j$ is small.

### 2.3 Confidence levels for the simulation

The driver for Monte Carlo is the required confidence level which is a quantitative statement about the quality of the experiment. The frequency interpretation is that a confidence level of 100( 1 - h )% means there is a 100h % or less chance that the experiment has misled us. This paper takes the point of view that the quality of the experiment should match the quality of the desired results. That is, if the probability to be established is p, then the confidence level should be at least 100( 1 - p )%. Hence, this paper will seek confidence levels of at least 100( 1 – 1e-7 )%. The confidence level may need to be even higher because loss-of-separation is only one incident among many. The final confidence level must combine the confidence level of a number of experiments. A result in combining confidence levels is the following.

Theorem: Suppose $(a_j , b_j)$ is a $100( 1 - h_j )$% confidence interval for $\theta_j$ for $1 \le j \le n$, then $[ (a_1 , b_1), \ldots (a_n , b_n) ]$ is a $100( 1 - h_1 - \ldots - h_n )$% confidence interval for $(\theta_1 , \ldots , \theta_n)$.

For example, if there are 10 parameters to be estimated with a desired overall confidence level of 100( 1 – 1e-7 )%, then it is sufficient to estimate each of the parameters at the 100( 1 – 1e-8 )% level. In general, the individual confidence intervals do not need to be the same although the lack of confidence must have a sum less than or equal to 1e-7. Assuming all the trials are successful and given a desired probability p and confidence level h, the formula for computing the number of trials is

$$( 1\text{-}p )^n = h . \quad (5)$$

The reasoning is that ( 1-p ) is the probability of success (equivalently the non-occurrence of a failure) and repeated successes (n of them) implies that p is small. The probabilities (values of p) that appear in table 1 are those computed in section 2.1.

### 2.4 Baseline for simulation effort

Since the primary concern of this paper (and future efforts) is introducing realism while maintaining enough efficiency to establish the algorithms at the required probability and confidence levels, it is worthwhile to state what this study says about such efforts.

The case chosen is that the requirement is the probability of more than 3 incidents in 3 years is 0.10 and there are 100 types of incidents. This requires 370,000,000 trials. Using a desktop

| Requirement | Value of p per flight | Types of incidents is 100<br><br>Confidence level is<br>1 - p × (1e-2)<br><br>Number of trials | Types of incidents is 1000<br><br>Confidence level is<br>1- p × (1e-3)<br><br>Number of trials |
|---|---|---|---|
| Expected number of incidents per year is 1 | 1.0e-7 | 2.1e+8 | 2.3e+8 |
| Probability more than 1 incident a year is 0.10 | 5.3e-8 | 4.0e+8 | 4.5e+8 |
| Probability more than 1 incident a year is 0.01 | 1.5e-8 | 1.5e+9 | 1.7e+9 |
| Probability more than 3 incidents in 3 years is 0.10 | 5.8e-8 | 3.7e+8 | 4.1e+8 |
| Probability more than 3 incidents in 3 years is 0.01 | 2.7e-8 | 8.2e+8 | 9.0e+8 |

Table 1. Number of trials given requirement and number of types of incidents

computer, it took 25 hours to run this many trials. Assuming that it is feasible to run the program for half a year, that it is feasible to use 10 to 100 desktop computers, and that more efficient programs and faster computers are available, this implies it would be possible to run a simulation that is three or four orders of magnitude more complex.

## 3. Assumptions

As an early effort (for the author), there are a number of assumptions. (1) Only two aircraft at a time are considered. (2) All aircraft have the same speed and maintain this speed. (3) All scenarios are two-dimensional: all maneuvering is at a constant altitude. (4) The position and heading of all aircraft is precisely known. (5) Both aircraft know which one will make the collision-avoidance maneuver and what the maneuver will be. (6) Only approaching aircraft are considered for the reason below.

This study restricts itself to approaching aircraft since for aircraft on nearly coincident courses, it is possible that a simple jog will not prevent loss-of-separation as illustrated in figure 1

Fig. 1. Two aircraft on nearly coincident course

The solution is either trivial: have one aircraft perform a circle for delay, or it is global: have one aircraft change altitude or arrange traffic to avoid such circumstances. Hence, the examination of nearly-coincident flight paths is postponed to a later study.

## 4. The minimum point and flight angles

For algebraic and geometric convenience, we establish the coordinate system such that the first aircraft travels along the x-axis and the two aircraft are their minimum distance apart when this first aircraft is at the origin. Initially, we let the second aircraft approach the first at any angle although we later restrict the study to aircraft with opposite headings.
The first result is that the minimum-point for the second aircraft determines its flight angle except for the special case where the minimum point is (0,0). Let the minimum point for the second aircraft be (a,b).
The graph is

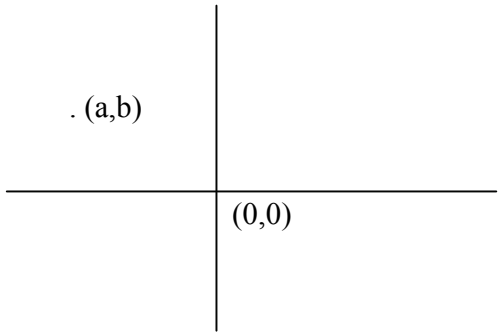



Fig. 2. The minimum points (0,0) and (a,b) for the two aircraft

The parametric equations for the original paths for the first and second aircraft are

$$\begin{aligned} x_1 &= t \\ y_1 &= 0 \\ x_2 &= a + t\cos\alpha \\ y_2 &= b + t\sin\alpha \end{aligned} \tag{6}$$

The distance-squared and its first and second derivative are

$$\begin{aligned} s^2 &= \left[a + t\cos\alpha\text{-}t\right]^2 + \left[b + t\sin\alpha\right]^2 \\ \frac{d\left(s^2\right)}{dt} &= 2\left[a + t\cos\alpha\text{-}t\right]\left[\cos\text{-}1\right] + 2\left[b + t\sin\alpha\right]\left[\sin\alpha\right] \\ \frac{d^2\left(s^2\right)}{dt^2} &= 2\left[\cos\alpha\text{-}1\right]^2 + 2\left[\sin\alpha\right]^2 \end{aligned} \tag{7}$$

Since the second derivative is positive, the zero-value of the first derivative gives a minimum. Skipping some algebraic steps, setting the first derivative equal to zero gives

$$0 = a\left[\cos\alpha - 1\right] + b\sin\alpha \tag{8}$$

Placing sine and cosine on opposite sides of the equation, squaring, substituting, and solving the quadratic gives

$$\cos \alpha = \frac{a^2 - b^2}{a^2 + b^2}, 1 \tag{9}$$

The value 1 corresponds to the two aircraft flying in parallel a constant distance apart. That case will not be considered in this study.
Hence, except for the point (0,0), the minimum-distance point determines the flight path of the second aircraft with

$$\begin{aligned} \cos \alpha &= \frac{a^2 - b^2}{a^2 + b^2} \\ \sin \alpha &= \frac{2ab}{a^2 + b^2} \end{aligned} \tag{10}$$

Since we are considering approaching aircraft, the cosine for the flight path of the second aircraft is negative. Hence, for the minimum-point (a,b), b > a.

## 5. Description of modified flight paths

The trigonometric result in the last section makes it natural to divide the region containing the minimum=distance points into four sectors where the angles range from π/4 to π/2, from π/2 to 3π/4, from 5π/4 to 3π/2, and from 3π/2 to 7π/4.
The sine of the trajectory is positive in the first sector, negative in the second, positive in the third, and negative in the fourth. This change is illustrated in figure 3.

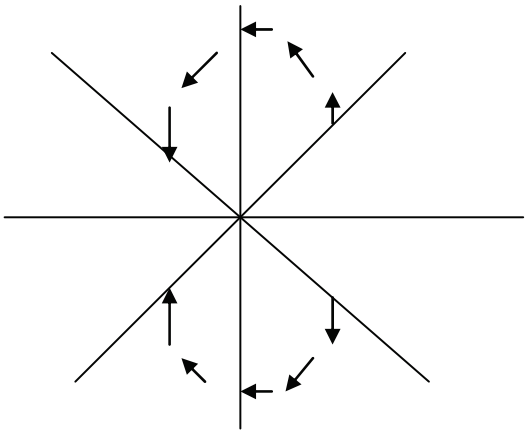


Fig. 3. The changes in flight-path angle according to the location of the minimum-distance point

The basic algorithm is that the second aircraft to reach the point of path-intersection turns into the path of the other aircraft. This algorithm does not cover parallel paths when the minimum-point lies on the y-axis, but for this study, this event has probability zero and is temporarily ignored since more robust algorithms must handle uncertainty due to instrumentation error.

As an example, consider two aircraft whose initial minimum distance point is in the first sector which is displayed in figure 4.

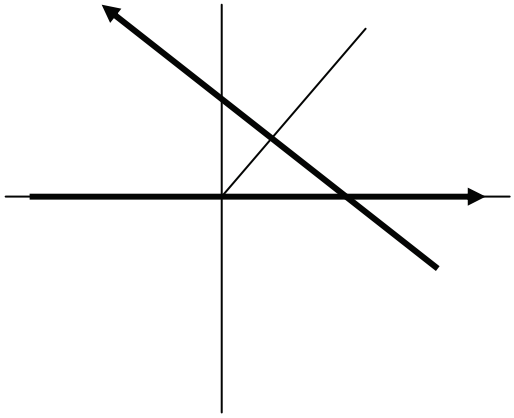


Fig. 4. Flight paths when the minimum-distance point lies in the first sector.

The burden of maneuver falls on the aircraft moving along the x-axis. This is illustrated in figure 5.

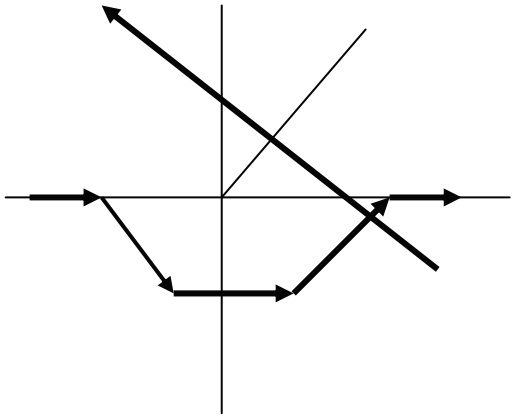


Fig. 5. The separation maneuver when the minimum-distance point lies in the first sector

The maneuvers for sectors 2, 3, and 4 are given in figures 6, 7, and 8.

Because of the symmetrical nature of the separation maneuvers, it is sufficient to examine the case where the minimum-distance point lies in the first sector and the maneuver is given in figure 5.

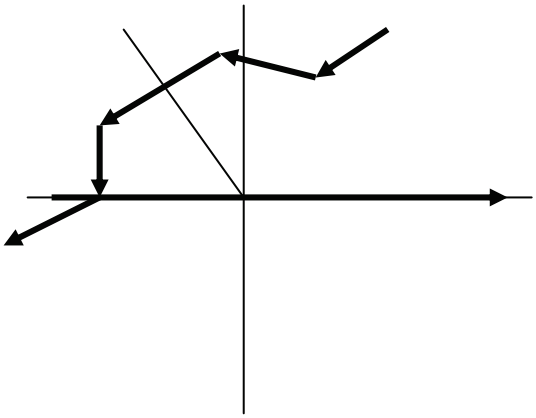


Fig. 6. The separation maneuver when the minimum-distance point lies in the second sector

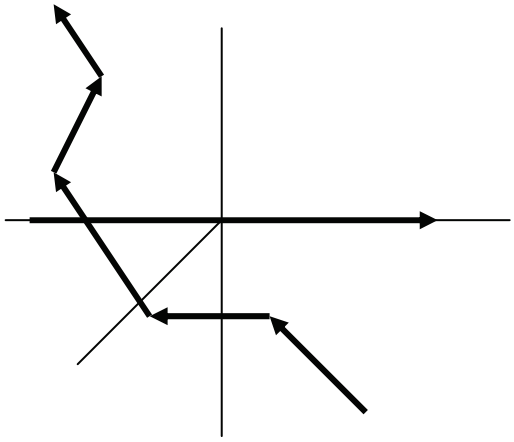


Fig. 7. The separation maneuver when the minimum-distance point lies in the third sector

## 6. Analytical demonstration of separation

We will scale the required minimum distance to 1.
Showing the two paths maintain separation is an exercise in calculus. The idealized paths are either a single straight line or a sequence of straight lines. The demonstration pivots on the path that is a sequence of straight lines. Each segment is examined at its endpoints and zero value of the derivative of the distance when traversing the straight line segment.
First, there are the endpoints and parametric equations for each of the segments.
The first segment goes from (-4,0) to (-2,-2) along the path

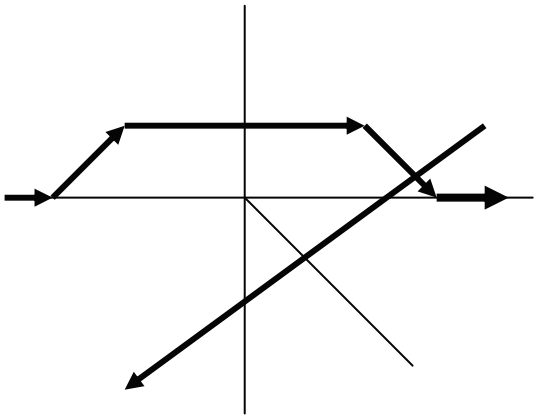


Fig. 8. The separation maneuver when the minimum-distance point lies in the fourth sector

$$\begin{aligned} x_1(t) &= a - 4\cos\alpha + t\cos\alpha \\ y_1(t) &= b - 4\sin\alpha + t\sin\alpha \\ x_2(t) &= -4 + \frac{\sqrt{2}}{2}\, t \\ y_2(t) &= 0 - \frac{\sqrt{2}}{2}\, t \\ &\text{for } 0 \le t \le 2\sqrt{2} \end{aligned} \tag{11}$$

The second segment goes from (-2,-2) to (+2,-2) along the path

$$\begin{aligned} x_1(t) &= a - (4 - 2\sqrt{2})\cos\alpha + t\cos\alpha \\ y_1(t) &= b - (4 - 2\sqrt{2})\sin\alpha + t\sin\alpha \\ x_2(t) &= -2 + t \\ y_2(t) &= -2 \\ &\text{for } 0 \le t \le 4 \end{aligned} \tag{12}$$

The third segment goes from (+2,-2) to (+4,0) along the path

$$\begin{aligned} x_1(t) &= a + 2\sqrt{2}\cos\alpha + t\cos\alpha \\ y_1(t) &= b + 2\sqrt{2}\sin\alpha + t\sin\alpha \\ x_2(t) &= +2 + \frac{\sqrt{2}}{2}\, t \\ y_2(t) &= -2 + \frac{\sqrt{2}}{2}\, t \\ &\text{for } 0 \le t \le 2\sqrt{2} \end{aligned} \tag{13}$$

We can first check the distances at the endpoints, which correspond to the corners for the path of the aircraft making the maneuver.

The first endpoint and distance-squared from equations (11) are

$$\begin{aligned} &(-4, 0) ; (a - 4\cos\alpha, b - 4\sin\alpha) \\ &s^2 = [a - 4\cos\alpha + 4]^2 + [b - 4\sin\alpha]^2 \end{aligned} \tag{14}$$

Since cosine is less than or equal to zero, the expression inside the first bracket is greater than or equal to 4, which implies the distance is greater than or equal to 4.
The second endpoint and distance-squared from equations (12) are

$$\begin{aligned} &(-2, -2) ; (a - (4 - 2\sqrt{2})\cos\alpha, b - (4 - 2\sqrt{2})\sin\alpha) \\ &s^2 = [a - (4 - 2\sqrt{2})\cos\alpha + 2]^2 + [b - (4 - 2\sqrt{2})\sin\alpha + 2]^2 \end{aligned} \tag{15}$$

Since cosine is negative, the term inside the first bracket is greater than or equal to 2, which implies the distance is greater than or equal to 2.
The third endpoint and distance-squared from equations (13) are

$$\begin{aligned} &(-2, -2) ; (a + 2\sqrt{2}\cos\alpha, b + 2\sqrt{2}\sin\alpha) \quad \text{for } t = 0 \\ &s^2 = [a + 2\sqrt{2}\cos\alpha - 2]^2 + [b + 2\sqrt{2}\sin\alpha + 2]^2 \end{aligned} \tag{16}$$

Since sine is positive, the term inside the second bracket is greater than or equal to 2, which implies the distance is greater than or equal to 2.
The fourth endpoint and distance-squared from equations (13) are

$$\begin{aligned} &(+4, 0) ; (a + 4\sqrt{2}\cos\alpha, b + 4\sqrt{2}\sin\alpha) \\ &s^2 = [a + 4\sqrt{2}\cos\alpha - 4]^2 + [b + 4\sqrt{2}\sin\alpha]^2 \end{aligned} \tag{17}$$

Since cosine is negative, the expression inside the first bracket is less than or equal to a-4, and since a is less than or equal to 1, the expression is less than or equal to -3, which implies the distance is greater than or equal to +3.
Next we consider the distances while traversing the segments between the endpoints.
The distance while traversing the first segment in terms of the parametric equations (11) is

$$\begin{aligned} s^2 = &\left[ a - 4\cos\alpha + t\cos\alpha + 4 - \frac{\sqrt{2}}{2} t \right]^2 \\ &+ \left[ b - 4\sin\alpha + t\sin\alpha + \frac{\sqrt{2}}{2} t \right]^2 \\ &\text{for } 0 \le t \le 2\sqrt{2} \end{aligned} \tag{19}$$

Now, ( -4 + t ) cosα is greater than or equal to zero, and $4 - \frac{\sqrt{2}}{2} t$ is greater than or equal to 2. Hence, the distance is greater than or equal to 2.
The distance while traversing the second segment in terms of the parametric equations (12) is

$$s^2 = \left[a - (4-2\sqrt{2})\cos\alpha + t\cos\alpha + 2t - t\right]^2 + \left[b - (4-2\sqrt{2})\sin\alpha + t\sin\alpha + 2\right]^2 \quad (20)$$

The derivative is

$$\frac{d(s^2)}{dt} = 2\left[a - (4-2\sqrt{2})\cos\alpha + t\cos\alpha + 2t - t\right]\left[\cos\alpha - 1\right] + 2\left[b - (4-2\sqrt{2})\sin\alpha + t\sin\alpha + 2\right]\left[\sin\alpha\right] \quad (21)$$

The second derivative is

$$\frac{d^2(s^2)}{dt^2} = 2\left[\cos\alpha - 1\right]^2 + 2\left[\sin\alpha\right]^2 \quad (22)$$

which is positive. Hence the zero value for the first derivative gives a minimum.
Setting the derivative equal to zero and solving for t gives

$$t = 3\text{-}\sqrt{2} + \frac{\sin\alpha}{\text{-}1 + \cos\alpha} \quad (23)$$

Substituting this value for t into the original distance formula and some algebra give

$$s^2 = \left[a + (-1+\sqrt{2})\cos\alpha \text{ -}1 + \sqrt{2} + \sin\alpha\right]^2 + \left[b + (-1+\sqrt{2})\sin\alpha\text{-}\cos\alpha + 1\right]^2 \quad (24)$$

Since all the terms in the expression in the second bracket are positive, the distance is greater than or equal to 1.
The distance while traversing the second segment in terms of the parametric equations (13) is

$$s^2 = \left[a + 2\sqrt{2}\cos\alpha + t\cos\alpha - 2 - \frac{\sqrt{2}}{2}t\right]^2 + \left[b + 2\sqrt{2}\sin\alpha + t\sin\alpha + 2 - \frac{\sqrt{2}}{2}t\right]^2 \quad (25)$$

for $0 \le t \le 2\sqrt{2}$

Now, $2\sqrt{2}\cos\alpha + t\cos\alpha - 2 - \frac{\sqrt{2}}{2}t \le 0 - 2 - \frac{\sqrt{2}}{2}t \le$ -2. Since a is less than or equal to 1, the expression inside the first bracket is less than or equal to -1, which implies the distance is greater than or equal to 1.

## 7. The feedback equations

Both aircraft are guided by onboard digital controllers. Hence, the desired flight path is given by a sequence of points: the positions the aircraft should be at the end of a control cycle.

The feedback control law is a modified PID (Position-Integral-Derivative) that pivots on the velocity vector. If it pivoted on distance, the aircraft would be constantly lurching forward to the next point on its path. Hence, for this controller, the velocity vector plays the role of P, the acceleration plays the role of D, and the error between the actual point of the aircraft and the desired point plays the role of I.
The aircraft is treated as a constant point of mass 1. Hence, the acceleration is proportional to the applied force. The distance between points is also normalized to 1. This last normalization is different from the normalization used in the analytic derivations for the idealized paths. This last normalization, of control-distance equal to 1, will be the one used for the rest of this paper and in the computer simulation.
At the time interval k, let

a(k) = current acceleration at time k

v(k) = current velocity at time k

s(k) = current position at time k

$v_d(k)$ = desired velocity at time k

$s_d(k)$ = desired position at time k

$\tau$ = control time interval

The equations for the next time interval are

$$a(k+1) = \alpha\, a(k) + \beta\left[\, v(k)\text{-}v_d(k)\right] + \delta\left[\, s(k)\text{-}s_d(k)\right]$$

$$\begin{aligned} v(k+1) &= v(k) + \tau\, a(k) \\ &= v(k) + \tau\left\{\, \alpha\, a(k) + \beta\left[\, v(k)\text{-}v_d(k)\right] + \delta\left[\, s(k)\text{-}s_d(k)\right]\right\} \end{aligned} \tag{26}$$

$$\begin{aligned} s(k+1) &= s(k) + \tau\, v(k) + \frac{\tau^2}{2} a(k) \\ &= s(k) + \tau\, v(k) \\ &\quad + \frac{\tau^2}{2}\left\{\, \alpha\, a(k) + \beta\left[\, v(k)\text{-}v_d(k)\right] + \delta\left[\, s(k)\text{-}s_d(k)\right]\right\} \end{aligned}$$

The acceleration is constant throughout the interval [k, k+1] while v(k+1) and s(k+1) are the values at the end of the interval.
For the second equation, subtract $v_d(k)$ from both sides and use $v_d(k) = v_d(k+1)$. For the third equation, write

$$\begin{aligned} s(k) + \tau\, v(k) &= s_d(k) + \tau\, v_d(k) + \left(s(k) - s_d(k)\right) + \tau\left(v(k) - v_d(k)\right) \\ &= s_d(k+1) + \left(s(k) - s_d(k)\right) + \tau\left(v(k) - v_d(k)\right) \end{aligned} \tag{27}$$

This gives

$$
\begin{aligned}
a(k+1) &= \alpha\, a(k) + \beta\left[v(k)\text{-}v_d(k)\right] + \delta\left[s(k)\text{-}s_d(k)\right] \\
v(k+1)\text{ - }v_d(k+1) &= \alpha\tau\, a(k) + (1+\beta\tau)\left[v(k)\text{-}v_d(k)\right] + \delta\tau\left[s(k)\text{-}s_d(k)\right] \\
s(k+1)\text{ - }s_d(k+1) &= \frac{\tau^2}{2}\alpha\, a(k) + \left(\tau + \frac{\tau^2}{2}\beta\right)\left[v(k)\text{-}v_d(k)\right] \\
&\quad + \left(1 + \frac{\tau^2}{2}\delta\right)\left[s(k)\text{-}s_d(k)\right]
\end{aligned}
\tag{28}
$$

This gives the coefficient matrix

$$
\begin{bmatrix}
\alpha & \beta & \delta \\
\alpha\tau & 1+\beta\tau & \delta\tau \\
\alpha\dfrac{\tau^2}{2} & \tau+\beta\dfrac{\tau^2}{2} & 1+\delta\dfrac{\tau^2}{2}
\end{bmatrix}
\tag{29}
$$

For the time and gain values

$$
\begin{aligned}
\tau &= 1 \\
\alpha &= -0.001 \\
\beta &= -0.5 \\
\delta &= -0.2.
\end{aligned}
\tag{30}
$$

The eigenvalues are 0, 0.1995+0.2433i, and 0.1995-0.2433i, which are less than one. Hence the system is stable. A more realistic study would include a more detailed model of the aircraft and base the control parameters on the aircraft's performance.

Continuing to choose numbers convenient for scaling, this study assumes an aircraft speed of 600 knots and a control-time interval of one second, which implies the five nautical mile separation requirement translates into a distance of 30 units in the simulation.

The control law above can take the aircraft through the required turn of $\pi/4$ radians. As expected, there is a small overshoot at the corners. The effect of the overshoot is displayed and examined in section nine.

## 8. The turbulence model

This section describes the external turbulence applied to the aircraft. The assumption is that turbulence is the accumulation of small effects which implies it has a normal distribution. Since the flight path is two-dimensional, the turbulence has an x-component and a y-component. In this study, the x-component and y-component are independent.

For both components, the force is a constant over the control interval of one second. The generality of this assumption will be studied later in this section.

The variants are the variance of the normal distribution, the correlation between the aircraft, and the correlation in time. This paper uses standard deviations of 1/10 and 1 for the normal distribution. For correlation between the aircraft, the two aircraft experience the same turbulence or experience independent turbulence. Time correlation is more complicated.

Time correlation is introduced by having the mean of the distribution for step k+1 depend on the values chosen for step k. If $x_k$ is the value for the turbulence in step k, then the value for the turbulence in step k+1 is chosen from a normal distribution with mean $c\,x_k$ for some positive constant c. All the distributions have the same variance $\sigma^2$.
Suppose $x_k$ is the value for step k, and suppose the $z_i$ 's are from a normal distribution with mean zero and variance $\sigma^2$. Then $x_{k+1}$ is given by

$$\begin{aligned} x_{k+1} &= c\,x_k + z_{k+1} \\ &= c\left(c\,x_{k-1} + z_k\right) + z_{k+1} \\ &= c^k z_1 + c^{k-1} z_2 + \ldots + c\,z_k + z_{k+1} \end{aligned} \tag{31}$$

Using the standard results for the means and variances of independent variables, the means and variances for this stochastic process are

$$\begin{aligned} \text{mean}\left(x_k\right) &= \text{mean}\left(c^{k-1} z_1 + c^{k-2} z_2 + \ldots + z_k\right) = 0 \\ \text{var}\left(x_k\right) &= \text{var}\left(c^{k-1} z_1 + c^{k-2} z_2 + \ldots + z_k\right) \\ &= c^{2k-2}\sigma^2 + c^{2k-4}\sigma^2 + \ldots + c^2\sigma^2 + \sigma^2 \\ &= \left(\frac{1-c^{2k}}{1-c^2}\right)\sigma^2 \qquad \text{if } c \neq 1 \\ &= k\sigma^2 \qquad \text{if } c = 1 \end{aligned} \tag{32}$$

Using the result on expectation that $E[\,z_i\,z_j\,] = 0$ if the z's are independent variables with zero means, the covariance of the process is

$$\begin{aligned} &E\left[\left(x_{n+k} - 0\right)\left(x_n - 0\right)\right] \\ &= E\left[\left(c^{n+k-1} z_1 + \ldots + c^k z_n + c^{k-1} z_{n+1} \ldots + z_{n+k}\right)\right. \\ &\qquad \left.\times\left(c^{n-1} z_1 + \ldots + z_n\right)\right] \\ &= E\left[\,c^{2n+k-2} z_1^2 + \ldots + c^k z_n^2\,\right] \\ &= \left[c^k \sum_{j=1}^{n} c^{2n-2j}\right]\sigma^2 \\ &= c^k \frac{1-c^{2n}}{1-c^2}\,\sigma^2 \quad \text{if } c \neq 1 \\ &= n\,\sigma^2 \quad \text{if } c = 1 \end{aligned} \tag{33}$$

It can be seen that if c < 1, then the covariance goes to zero as k, the distance between the two points, becomes large. Also, as n becomes large, the process approaches a stationary process: the covariance depends only on the distance between the variables.
We next look at the effect of the turbulence (the stochastic process) on the control states of acceleration, velocity, and distance. There is a slight shift in notation since we follow the control theory convention of indexing the initial parameters with zero. Let y be the state vector and x the stochastic input.

$$
\begin{aligned}
y(1) &= A\,y(0) + B\,x(0) \\
y(2) &= A\,y(1) + B\,x(1) \\
&= A^2\,y(0) + A\,B\,x(0) + B\,x(1) \\
&\vdots \\
y(n) &= A^n\,y(0) + \sum_{k=0}^{n-1} A^{n-1-k}\,B\,x(k)
\end{aligned}
\tag{34}
$$

Since the system is stable, the first term converges to the desired acceleration, velocity, and position(s). Since the x(k)'s are normal with mean zero, the error from turbulence (the stochastic process) is normal with mean zero. No attempt has been made to derive the variance of the error although it could be obtained empirically from simulation. Section nine estimates the mean and variance for one set of turbulence parameters.

The final topic in this section asks if the turbulence model technique of applying a constant force for a discrete interval is completely general. Can some type of average duplicate the effect of any function of force over that interval? The answer is no. Consider a point of mass 1, a time interval of 1, and the force function

$$
f(t) = \begin{cases} 2 & \text{for } 0 \le t \le 1/2 \\ 0 & \text{for } 1/2 \le t \le 1 \end{cases}
\tag{35}
$$

At time t=1, the velocity and position are

$$
\begin{aligned}
v &= f(t)\left[\frac{1}{2}\right] = 1 \\
s &= \frac{f(t)}{2}\left[\frac{1}{2}\right]^2 + (1)(1/2) = 3/4
\end{aligned}
\tag{36}
$$

To duplicate the result, a constant force $\alpha$ over the interval must satisfy the equations

$$
\begin{aligned}
v &= 1 = \alpha(1) = \alpha \\
s &= 3/4 = \frac{\alpha}{2}[1]^2 = \alpha/2
\end{aligned}
\tag{37}
$$

which is not possible.

## 9. Miscellany on perturbations

The graph in figure 9 shows the aircraft making the separation-maneuver moving under feedback control with no perturbation present. Note the overshoot at the corners.

The mean and standard deviation of the error for distance were estimated. The path in figure 9 lasts 320 seconds which is 320 discrete control intervals. At the end of each control interval the difference between actual position and desired position was computed. This computation was performed for 3000 flight paths or 960,000 points. The results are mean = 3.7298e-4 and std = 0.1531.

The graph in figure 10 shows the aircraft making the separation-maneuver moving under feedback control with perturbation present. This perturbation is one of the more extreme

ones with the variables initially chosen from the normal with standard deviation 1 and then given time correlation with c = 1/2.

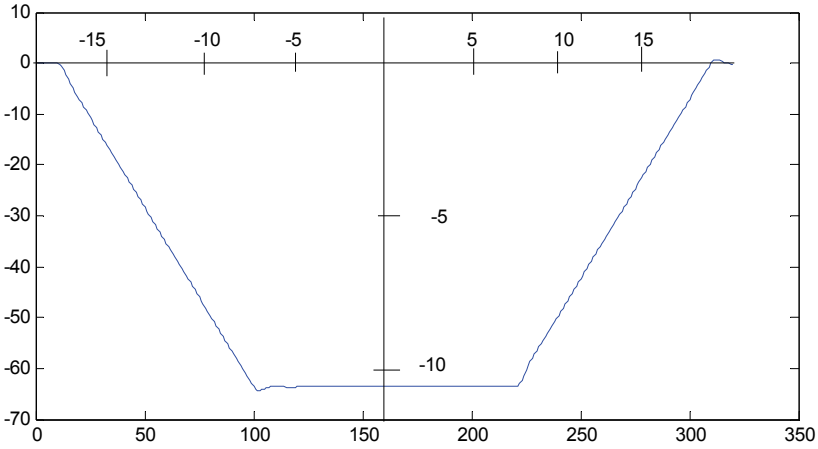



Fig. 9. Path of maneuvering aircraft moving with feedback. The marks on the axes indicate five-mile intervals

Once again, there were 3000 trials with 320 points per trial. For this case, the error had mean = 0.0068 and std = 4.3898.

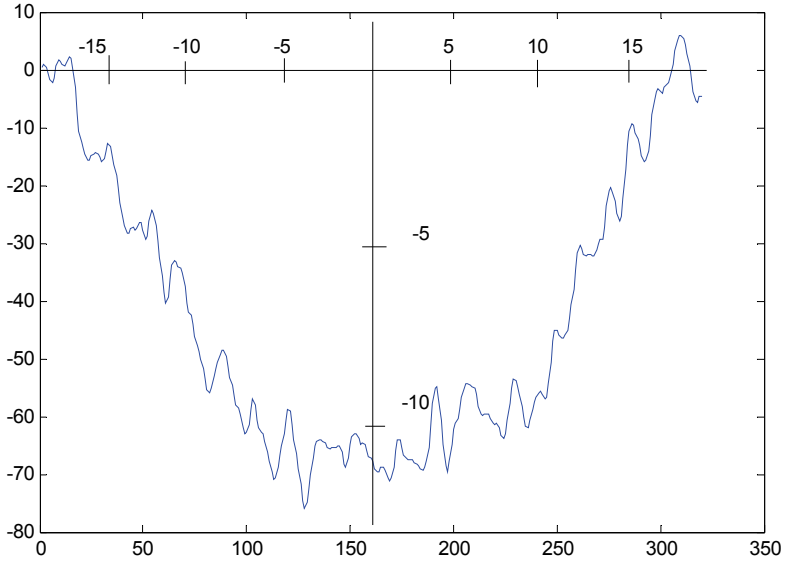



Fig. 10. Path of maneuvering aircraft in the presence of perturbations. The marks on the axes are five-mile increments

There are additional areas for investigation for perturbations. One item not covered in this study is when the perturbation has a drift, a constant non-zero value for its mean.
Another item is whether or not it is possible to establish a high probability envelope for the path of the aircraft in the presence of perturbations. Such an envelope would imply that a separation algorithm that allowed for the envelope could be proved, and the need for studying the separation in the presence of perturbations would be eliminated.

## 10. Choosing the minimum distance point in a sector

Randomly choosing the minimum-distance point in the sector for a trial brings up two problems. The first problem is what is the actual distribution for air traffic. This is not known to the author. Since the usual response to such lack of information is to use the uniform distribution, the second problem is obtaining a distribution and demonstrating it is uniform.
Suppose the sector is part of a circle of radius 1. Consider an arbitrary subsector as in figure 11.

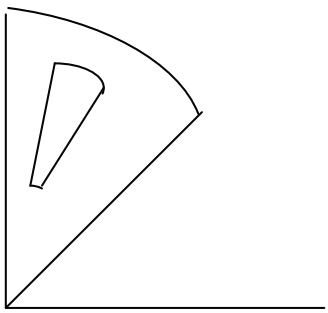


Fig. 11. Subsector inside a sector

Suppose the subsector lies between angles $\beta_1$ and $\beta_2$ and radii $r_1$ and $r_2$. The area of the subsector divided by the area of the one-eighth circle is

$$\begin{aligned} A &= \frac{\left(\pi r_2^2 - \pi r_1^2\right)\left(\frac{\beta_2 - \beta_1}{2\pi}\right)}{\frac{\pi}{8}} \\ &= \frac{4}{\pi}\left(\beta_1 - \beta_2\right)\left(r_2^2 - r_1^2\right) \end{aligned} \tag{38}$$

We have a uniform distribution in the sector if the probability of a point being in the subsector equals its proportional area.
To this end, choose $n_1$ and $n_2$ independently from a uniform distribution, and form the point with angle $\frac{\pi}{4} + \frac{\pi}{4} n_1$ and radius $\sqrt{n_2}$. Then

$$
\begin{aligned}
&\mathrm{Prob}\left\{\beta_1 \le \frac{\pi}{4} + \frac{\pi}{4} n_1 \le \beta_2 \text{ and } r_1 \le \sqrt{n_2} \le r_2\right\} \\
&= \mathrm{Prob}\left\{\beta_1 \le \frac{\pi}{4} + \frac{\pi}{4} n_1 \le \beta_2\right\} \\
&\quad \times \mathrm{Prob}\left\{r_1 \le \sqrt{n_2} \le r_2\right\} \\
&= \mathrm{Prob}\left\{\frac{4}{\pi}\beta_1 - 1 \le n_1 \le \frac{4}{\pi}\beta_2 - 1\right\} \\
&\quad \times \mathrm{Prob}\left\{r_1^2 \le n_2 \le r_2^2\right\} \\
&= \frac{4}{\pi}\left(\beta_1 - \beta_2\right)\left(r_2^2 - r_1^2\right)
\end{aligned}
\tag{39}
$$

Hence, the probability that the point is in the subsector is equal to the relative area of the subsector which gives a uniform distribution for points in the sector.

## 11. The simulation cases

Eight Monte Carlo experiments were performed, each with 370,000,000 trials.

The aircraft had identical perturbations or independent perturbations.

The initial random variable was chosen with std = 1/10 or std = 1.

The correlation constant was c = 0 or c = ½.

There was no loss of separation in any of the trials.
For each case, the experiment showed that the probability of more than 3 incidents in three years is less than 0.10 at the 100( 1 - 1e-9 ) confidence level, which is the confidence level appropriate if there are 100 types of incidents.

## 12. Minimum distance estimation

It's satisfying that the separation algorithm handled all the perturbation cases, but it gives no insight with respect to the effect of the perturbations. This section attempts to remedy this by estimations of minimum distance. For each set of conditions, one million (1e+6) trials were conducted, and the minimum distance of the two aircraft was recorded for each trial. For each set of conditions, the mean, standard deviations, and 99% confidence interval for the minimum distance were computed.
Since this comparison of conditions is not related to any FAA requirement, the investigation does not require meeting the confidence level of incidents, and the 99% confidence interval is sufficient to compare results.
To make the results more comparable, each experiment used the same set of (1e+6) points in the sector. The points were chosen according to the uniform-criteria in section ten with the angles and radii given by

$$
\begin{aligned}
&\alpha_k = \pi/4 + k\,\pi/4000 \quad \text{for } k = 1, \ldots, 1000 \\
&r_k = \mathrm{sqrt}\,(\,k/1000\,) \quad \text{for } k = 1, \ldots, 1000\,.
\end{aligned}
\tag{40}
$$

The parameters that were varied were

The perturbations for the two aircraft were either identical (pert2=pert1) or independent (Ind).

The initial random variable was chosen from a normal distribution with mean = 0 and std = 1/10 or std = 1.

The correlation constant was 0 or ½.

The results for the eight cases are given in table 2.
From an examination of the table, the major factor is the standard deviation of the original normal distribution. Once this standard deviation is large, the correlation factor becomes significant.

| Conditions | mean | std | 99% int |
|---|---|---|---|
| Pert1=pert2<br>Std = 1/10<br>C = 0 | 67.3279 | 1.5910 | 67.3238<br>67.3320 |
| Pert1=pert2<br>Std = 1<br>C = 0 | 65.9497 | 1.7335 | 65.9452<br>65.9542 |
| Ind<br>Std = 1/10<br>C = 0 | 67.1959 | 1.6200 | 67.1917<br>67.2001 |
| Ind<br>Std = 1<br>C = 0 | 63.0055 | 2.8300 | 62.9981<br>63.0129 |
| Pert1=pert2<br>Std = 1/10<br>C = 1/2 | 67.2951 | 1.5941 | 67.2910<br>67.2992 |
| Pert1=pert2<br>Std = 1/10<br>C = 1/2 | 65.3106 | 1.8880 | 65.3057<br>65.3155 |
| Ind<br>Std = 1/10<br>C = 1/2 | 67.0579 | 1.6499 | 67.0536<br>67.0622 |
| Ind<br>Std = 1<br>C = 1/2 | 60.5988 | 3.6930 | 60.5892<br>60.6084 |

Table 2. Estimated mean, standard deviation, and 99% confidence interval for minimum distance

## 13. Conclusions

The primary objective of this paper is to establish that an air traffic algorithm meets the FAA requirements in the presence of perturbations.
There is a discussion of a representative FAA requirement, its probabilistic interpretation, and the number of trials needed in a Monte Carlo simulation. The algorithm chosen is flight

separation for approaching aircraft. It is first shown that the minimum-distance point of the aircraft determines their flight angles. With this result in hand, the algorithm is described and an analytical proof of the algorithm is given for idealized flight paths.
Next, two items are introduced for realism. A modified PID control law is proposed and shown to be stable. Once the control law is in place, perturbations can be introduced as external forces. The parameters of the perturbation are the standard deviation, the correlation between aircraft, and the time correlation. The stochastic properties of the perturbation are derived. Monte Carlo simulation shows the separation-algorithm meets a chosen FAA requirement for eight cases of the perturbation. A minimum-distance study shows the different perturbations have different effects. The natural extensions to this effort are to consider more than two aircraft, a more global setting, wind, and instrumentation errors. Wind, for instance, can be included by letting the random variables have a nonzero mean. How much can be established by realistic simulation remains an open problem. The method presented in this paper depends on the ability to execute the algorithm 100,000 times faster than real time. Whether such speeds can be achieved for realistic algorithms is an open question.

## 14. References

There is no lack of publications about separation assurance or, when things get more desperate, collision avoidance. In comparing this effort to other publications, the salient points are (1) examination of the FAA requirements as probability conditions, (2) integrating the combined realism of feedback control and perturbations, and (3) determining and performing the necessary Monte Carlo trials.
The paper [Kuchar & Yang 2000] sets up a framework to review 68 conflict detection and resolution methods, but there is no discussion of their verification. The paper [Christodoulou & Kodazakis 2006] examines a three dimensional algorithm for collision avoidance, but there are no perturbations, and only a few test cases. The paper [Sullivan & Malsam 2002] includes human in the loop with the intent of moving from sector to global implications to improve efficiency, but there is no mention of FAA safety requirements. The paper [Sweet et al 2002] identifies possible areas to investigate for improvement of efficiency, but there is no mention of FAA safety requirements. The paper [Farley & Erzbereger 2007] proposed a conflict resolution algorithm and tested it using real data where it appeared to be about 97% successful. The paper [McNally & Gong] develops and tests an operating concept and a laboratory analysis methodology to examine how four-dimensional trajectory analysis methods could support higher levels of automation for separation assurance while recognizing that trajectory prediction uncertainties can be on the same order of magnitude as safe separation criteria, which implies that uncertainty is an important consideration, but there is no reference to the FAA requirements. The paper [Paielli] looks at vertical maneuvers when loss of separation is imminent, and these methods were tested on archived data from 100 actual operational errors. The paper [Alam & Abbass 2008] attempts a global simulation of the Australian airspace to study the feasibility of free flight. It contains numerous realistic elements, but there is no quantitative study of safety with confidence levels. The paper [Pallottino et al 2002] applies integer programming to solve conflict resolution for multiple aircraft, and several examples are presented, but the program is not established at any confidence level.

There is a school of thought that believes simulation alone will not establish safety requirements and proof by formal methods is necessary. The papers [Dowek & Munoz] and [Geser & Munoz] are representative.

Alam, S. & Abbass, H. ATOMS: Air Traffic Operations and Management Simulator, *IEEE Transactions on Intelligent Systems*, Vol. 9, No. 2, June 2008, pp. 209-225.

Butler, R. & Munoz, C. NASA *Langley's Formal Methods Research in Support of the Next Generation Air Transport System*, http://shemesh.larc.nasa.gov/LFM2008/slides/InvitedTalks/Butler.pdf.

Christodoulou, M. & and Kodazakis, S. Automatic Commercial Aircraft-Collision Avoidance in Free Flight: The Three Dimensional Problem, *IEEE Transactions on Intelligent Systems*, Vol. 7, No. 2, June 2006, pp. 242-249.

Dowek, G. & Munoz, C., Conflict Detection and Resolution for 1,2, …, N Aircraft, *7th AIAA Aviation*
*Technology, Integration and Operations Confrence.*

Farley, T. & Erzberger, H. Fast-Time Simulation Evaluation of a Conflict Resolution Algorithm Under High Air Traffic Demand, *Proceedings of 7th USA/Europe ATM 2007 R&D Seminar*, Barcelona, Spain, 2-5 July 2007.

Geser, A. & Munoz, C., A Geometric Approach to Strategic Conflict Detection and Resolution, *Proceedings*
*of the 21st Digital Avionics Systems Conference, 2002.*

Kuchar J. & Yang, L. A review of conflict Detection and resolution Modeling Methods, *IEEE Transactions on Intelligent Systems*, Vol. 1, No. 4, December 2000, pp. 179-189.

McNally, D., & Gong, C. Concept and Laboratory Analysis of Trajectory-Based Automation for Separation Assurance, *Air Traffic Control Quarterly*, Vol. 15(1), pp. 35-63.

Pallottino, L., Feron, E. & Bicchi, A. Conflict Resolution Problems for Air Traffic Management Systems Solved With Mixed Integer Programming, *IEEE Transactions on Intelligent Systems*, Vol. 3, No. 1, March 2002, pp. 3-11

Sullivan, B. & Malsam, S. Development of a Real-Time Virtual Airspace Simulation Capability for Air Traffic Management Research, *Proceedings of AIAA Modeling and Simulation Technologies Conference and Exhibit*, 5-8 August 2002, Monterey.

Sweet, D., Manikonda, D., Aronson, V., Roth & Blake, M. Fast-Time simulation System for Analysis of Advanced Air Transportation Concepts, *Proceedings of AIAA Modeling and Simulation Technologies Conference and Exhibit*, 5-8 August 2002, Monterey, California.

# 18

# Modelling of the Wireless Propagation Characteristics inside Aircraft

Carl James Debono, Reuben A. Farrugia and Keith Chetcuti
*University of Malta*
*Malta*

## 1. Introduction

Advances in wireless communications technology and more sophisticated portable devices have led to a drastic increase in wireless services and applications. This advancement was made possible through hardware improvements which allow more functions to be implemented in smaller sized devices. The demand for more wireless services has pushed the industry and the research community to increase the communication data rates, connectivity, and availability. The increase in the user base has also brought a decrease in the cost of the services.

Although wireless access is becoming available in all public places around the world, this is still not the case for the air passengers who are cut off from wireless services during the duration of the flight. Deployment of wireless technology inside aircraft is still a hot issue due to uncertainties related to interference. Yet, the European Commission has prepared the legal framework for wireless connectivity inside aircraft in April 2008 (Commission, 2008). Aircraft manufacturers will benefit from this technology by exploiting wireless networks to reduce the cable complexity, hence the weight of the aircraft, and by providing new in-flight services, such as online passenger meal selection, service request and video on demand. On the other hand, provisioning of wireless service during flight would profit passengers, since they gain seamless access to common wireless services, such as phone, Internet, and multimedia communications.

A wireless network can always be designed using a measurement campaign, whereby transmitters are placed in strategic locations inside the aircraft and measurements are taken at all points inside the aircraft. This solution is very expensive and requires repetitive readings to measure the propagation losses at all the frequencies lying within the frequency band of interest. These measurements have to be done on all the aircraft models and for every different furniture configuration. While the latter is not much of concern in commercial aircrafts, since these are quite similar in terms of furniture location, this creates a sensible issue in the business jet market, where each aircraft's configuration is tailor-made for the customer. With such requirements, it is not possible to find the ideal antenna locations using measurement techniques.

This calls for the development of an accurate computer model that can estimate the propagation characteristics inside any aircraft. This model must be capable of estimating the power level that can be received at any point inside the enclosed structure, thus creating a

propagation map. Therefore, any changes in aircraft model and/or its configuration can be easily accommodated and the propagation map recomputed. This map can then be used to determine the parameters associated with the quality of the wireless channel, hence making the characterisation of the wireless channel possible.

In this chapter we present the application of ray tracing techniques, which are based on geometric optics, to model an IEEE 802.11a wireless system propagation map. This technique has become viable with the increased computational power and speeds available in today's computer technology. The developed propagation model can be used to maximise the coverage within the aircraft cabin and to minimise the power emitted from each access point, by determining the optimum location of its antenna. The cabin presents an interesting scenario because the space within it is limited and there is a high density clutter, because of the furniture and seats, resulting in merely any line-of-sight transmission between the server and the client locations. This implies that propagation inside aircraft is mainly due to reflection, transmission and scattering. Also the cabin does not have any attenuation caused by walls as present in normal office environments. We will also discuss how multipath, coherence bandwidth and time dispersion parameters can be obtained, from the computer model, and used to model the channel. These parameters indicate whether inter-symbol and inter-carrier interference will be experienced by the channels.

## 2. IEEE 802.11a model

### 2.1 Overview of the IEEE 802.11 standards

The IEEE 802.11 standards deployed today are a result of technological advancements, both in hardware and software. The first IEEE 802.11 standard was deployed in 1997 with a maximum throughput of 2Mbps (Crow, B.P. et al, 1997). This has gone a long way and current amendments are looking at high speeds of more than 100Mbps (Paul, T.K. & Ogunfunmi, T. 2008). The IEEE 802.11 standard has a mandatory throughput of 1Mbps which can be extended to 2Mbps. The standard permits three physical layer implementations, namely, frequency hopping spread spectrum (FHSS), direct sequence spread spectrum (DSSS), and Infra red (IR). The FHSS technique uses Gaussian frequency shift keying (GFSK) for modulation, while DSSS and IR utilise differential binary shift keying (DBPSK) / differential quadrature phase shift keying (DQPSK) and pulse position modulation (PPM) respectively (Crow, B.P. et al, 1997).

IEEE 802.11a presents data rates of up to 54Mbps and uses the 5GHz unlicensed national information infrastructure (U-NII) frequency band. The physical layer uses orthogonal frequency division multiplexing (OFDM) multicarrier transmission. The modulation schemes used are binary phase shift keying (BPSK), quadrature phase shift keying (QPSK), and quadrature amplitude modulation (QAM) (IEEE standard, 2003).

### 2.2 Orthogonal frequency division multiplexing

The application of OFDM to wireless communications has gained ground in the last decade. The main features offered by this technology are: (i) OFDM is spectrally efficient (Can De Beek, J.J. et al, 2002) and (ii) it increases robustness against frequency selective fading (Yomo, H. et al, 2005). The latter is because in a multi-carrier system, it is highly improbable that a fade will affect all the subcarriers and the system can still function.

The orthogonal subcarriers overlap one another making the system bandwidth efficient when compared to non-overlapping multi-carrier systems. This characteristic makes sure that the subcarriers do not interfere with each other. The input stream is divided into *N* data streams, corresponding to the number of subcarriers, and each subcarrier modulates one of these low rate data streams. A multiplex of these modulated subcarriers results in an OFDM symbol.
One of the problems with OFDM is that orthogonality can be lost if the signal is transmitted over a dispersive channel causing inter-carrier interference (ICI). This problem can be solved by introducing a cyclic prefix (Peled, A. & Ruiz, A., 1980). This solution also prevents inter-symbol interference (ISI) between consecutive OFDM symbols. This means that simple one-tap equalizers, implemented at the receiver, are enough to offer protection against ICI and ISI, making the use of OFDM possible in wireless environments.

### 2.3 IEEE 802.11a PHY layer

The physical (PHY) layer interfaces the medium access control (MAC) layer to the wireless medium. The table below (IEEE standard, 2003) gives a summary of the permissible data rates and their characteristics. A block diagram illustrating how a symbol is generated in the PHY layer is presented in Figure 1.

| Mode | Data rate (Mbps) | Modulation | Coding rate (R) | Coded bits per subcarrier | Coded bits per OFDM symbol | Data bits per OFDM symbol |
|---|---|---|---|---|---|---|
| 1 | 6 | BPSK | 1/2 | 1 | 48 | 24 |
| 2 | 9 | BPSK | 3/4 | 1 | 48 | 36 |
| 3 | 12 | QPSK | 1/2 | 2 | 96 | 48 |
| 4 | 18 | QPSK | 3/4 | 2 | 96 | 72 |
| 5 | 24 | 16QAM | 1/2 | 4 | 192 | 96 |
| 6 | 36 | 16QAM | 3/4 | 4 | 192 | 144 |
| 7 | 48 | 64QAM | 2/3 | 6 | 288 | 192 |
| 8 | 54 | 64QAM | 3/4 | 6 | 288 | 216 |

Table 1. IEEE 802.11a physical layer parameters.

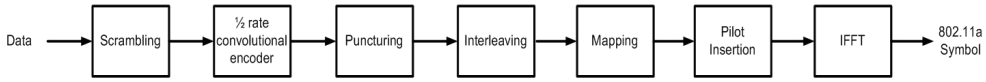



Fig. 1. Block diagram of IEEE 802.11a symbol generation

The data coming from the higher layers is first scrambled and protected by the convolutional encoder. Puncturing is applied to these symbols, when necessary, to obtain the output symbol. Symbol interleaving is applied to protect the sequence from burst errors, by introducing sparsity between consecutive data elements. The new data sequence is mapped using the operating modulating scheme and a serial-to-parallel converter divides the data stream into 48 different streams, corresponding to the number of carriers. Pilot signals are then inserted to aid in the demodulation and the channel estimation process at the receiver side. The inverse Fourier transform (IFFT) converts the signal back into the time

domain. A cyclic prefix is added to each symbol before re-converting the streams to a single data stream. This cyclic prefix helps in preventing ISI between adjacent symbols and ICI between adjacent carriers. The length of this prefix is a design parameter which depends on the application, since it must be larger than the delay spread of the channel to guarantee that no ISI or ICI distortion effects are introduced in the system.
The OFDM PHY layer divides the information signal across 52 subcarrier frequencies. From these, 48 are used for data transmission and 4 are used by the pilot signals. Another 12 subcarriers are used as cyclic prefix. The available bandwidth is 20MHz and each subcarrier is separated by 312.5kHz. With this configuration, each symbol has a duration of 4μs, where the IFFT information has a period of 3.25μs while the cyclic prefix uses the remaining 0.75μs.

## 3. Propagation model

The geometry inside an aircraft contains high density clutter, very different from typical office and home scenarios. This means that severe wireless propagation conditions prevail. The advances in computer computation power has made ray tracing modelling possible. Ray tracing techniques offer an adequate representation of the wireless propagation characteristics inside closed, high clutter environments (Diaz, N.R. & Achilli, C., 2003).
The method uses the theory of geometric optics (GO) and is a brute-force solution. Rays are launched from a transmitter at a fixed power level and an estimate of the power levels along the path is found. These rays will undergo reflection, refraction and diffraction as they impinge on the structure of the cabin and the furniture found along the path. The result is a three-dimensional map of the power levels inside the aircraft. This map therefore represents an estimate of the field strength that would be received by mobile terminals at any position inside the aircraft.

### 3.1 Ray tracing technique

The two most popular ray tracing techniques are: (i) the method of images, and (ii) the method of shoot and bounce. The first solution gives the exact radiation pattern when the environment is smooth, infinite, or made up of semi-infinite perfect electrically conducting surfaces arranged in a limited set of canonical geometries (Diaz, N.R. & Achilli, C., 2003). The second method is more suitable for complex environments and is therefore used for the propagation modelling inside the aircraft.
In the cabin environment, the signal wavelength, i.e. 5.7cms, is much smaller than the dimensions of the clutter inside the structure and therefore the approximation of the GO hypothesis holds. This implies that ray tracing can be safely used to model the system. A ray is associated with a local plane wave which can be represented by (Diaz, N.R. & Achilli, C., 2003):

$$\nabla^2 \psi + k^2 n^2 \psi = 0 \tag{1}$$

where $\psi$ is the waveform function which governs the scalar wave propagation, $k$ is the wave number, and $n$ is the reflection index of the medium. Solving for $\psi$ gives (Diaz, N.R. & Achilli, C., 2003):

$$\psi = Ae^{-jkS} \tag{2}$$

where $A$ is a function which determines the amplitude of the wave and the function $S$ determines the direction and phase. Using equations (1) and (2), we can conclude that if

$$\frac{\nabla^2}{k^2} << n^2 \tag{3}$$

a solution which does not depend on the frequency of operation can be obtained and therefore geometric optics can be used.

As the rays pass through the medium they will experience reflection, refraction and diffraction. This occurs because of the obstacles found in the path between the transmitters and the receivers. If we assume that the obstacles found in the cabin environment are made up of homogeneous material, we simplify the model as each surface can now be described through its dielectric constant, magnetic permittivity and conductivity. This assumption implies that diffraction is not taken into consideration.

If two media possessing different conductivity and permittivity characteristics are assumed to be separated by an infinite plane, then we can obtain equations that relate the reflected electromagnetic wave to the incident wave and the properties of the media. Polarization effects must also be considered and this is done by splitting the electric field into two components; a component which lies parallel to the incident surface and another one which lies perpendicular to it. The refracted and reflected rays can be estimated by multiplying each component to the corresponding Fresnel's coefficient, given by (James, G.L., 1986):

$$R_{\perp} = \frac{\cos\theta - \sqrt{\hat{\varepsilon}_r - \sin^2\theta}}{\cos\theta + \sqrt{\hat{\varepsilon}_r - \sin^2\theta}} \tag{4}$$

$$R_{\parallel} = \frac{\hat{\varepsilon}_r \cos\theta - \sqrt{\hat{\varepsilon}_r - \sin^2\theta}}{\hat{\varepsilon}_r \cos\theta + \sqrt{\hat{\varepsilon}_r - \sin^2\theta}} \tag{5}$$

$$T_{\perp} = \frac{2\cos\theta}{\hat{\varepsilon}_r \cos\theta + \sqrt{\hat{\varepsilon}_r - \sin^2\theta}} \tag{6}$$

$$T_{\parallel} = \frac{2\sqrt{\hat{\varepsilon}_r}\sin\theta}{\hat{\varepsilon}_r \cos\theta + \sqrt{\hat{\varepsilon}_r - \sin^2\theta}} \tag{7}$$

where $\theta$ is the angle of incidence, and $\hat{\varepsilon}_r = \varepsilon_r - j(\eta/2\pi)\sigma\lambda$ is the relative complex dielectric constant. In this equation, $\varepsilon_r$ is the relative dielectric constant, $\sigma$ is the conductivity of the materials, $\lambda$ is the wavelength in meters, $R$ is the reflection coefficients, and $T$ is the refraction coefficients. These relations are valid in this case as the medium is air, which presents a dielectric constant of $\varepsilon_0$ (Diaz, N.R. & Achilli, C., 2003).

The Fresnel reflection coefficients will account for reflections coming from smooth surfaces. However, this is not the case for the environment being considered. The Rayleigh criterion (Bothias, L., 1987) can be used as a roughness test. When a surface is rough the incident's

ray energy will be diffused in angles other than the main angle of reflection, and therefore there is a reduction in the energy of the main reflected ray (Landron, O. et al, 1993). The critical height, $h_c$, in meters, defined by the Rayleigh criterion is given by:

$$h_c = \frac{\lambda}{8\cos\theta_i} \tag{8}$$

where $\lambda$ is the wavelength of the signal and $\theta_i$ is the angle of incidence.
A surface is considered to be rough if the protuberances exceed $h_c$. In such cases the reflection coefficients ($R_\perp$ and $R_\parallel$) have to be modified by the scattering loss factor (Hashemi, H., 1993):

$$\rho_s = \exp\left[-8\left(\frac{\pi\sigma_h\cos\theta_i}{\lambda}\right)^2\right] I_o\left[8\left(\frac{\pi\sigma_h\cos\theta_i}{\lambda}\right)^2\right] \tag{9}$$

where $\sigma_h$ is the standard deviation of the surface height. The reflection coefficients thus become:

$$\left(R_\perp\right)_{rough} = \rho_s\left(R_\perp\right)_{smooth} \tag{10}$$

$$\left(R_\parallel\right)_{rough} = \rho_s\left(R_\parallel\right)_{smooth} \tag{11}$$

Within the IEEE 802.11a frequency band, the critical height varies from 7mm at an angle of 1° to 40.9cm at an angle of 89°. Inside the cabin, we will not find large roughness in the materials used. Figure 2 shows the profile of the scattering roughness factor for an IEEE 802.11a system. From this figure, it can be seen that the ray can loose up to 75% of its energy when it hits a surface.

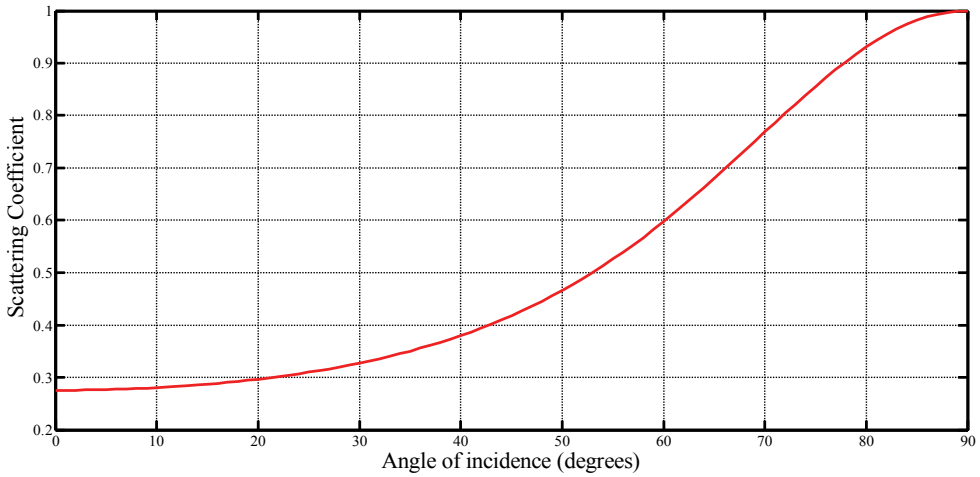



Fig. 2. Scattering roughness factor against incident angle at 5.25GHz

### 3.2 Signal strength propagation model

Considering the layout of a typical A340-600 cabin, an aircraft model which includes the furniture is developed. The signal strength at each location inside the aircraft is determined by placing the access points at fixed locations inside the aircraft. Rays are then launched from each transmitting antenna. By vectorially adding all the rays passing through all the points inside the cabin, we obtain an estimate of the signal strength at each position within the aircraft. Therefore, a propagation map which indicates the radio coverage is created.

A ray leaving the transmitter will travel in free space until it impinges on a surface. At this point it is reflected or reflected and refracted as illustrated in figure 3. The rays that result from this interaction are launched again with the new power level from the point of collision. This process will be repeated until the power of the ray falls below a pre-determined threshold.

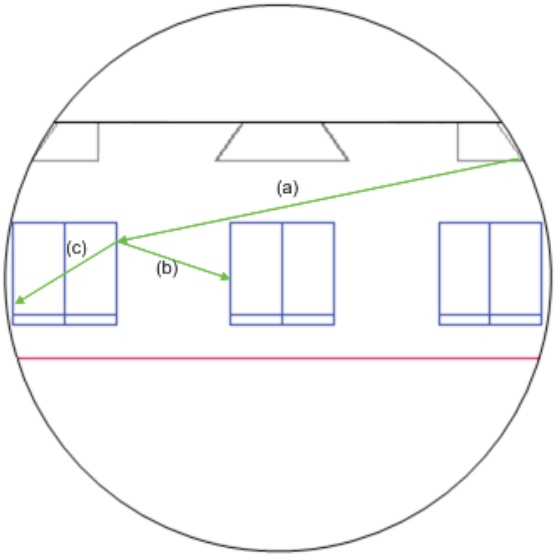



Fig. 3. The ray tracing technique (a) main ray, (b) reflected ray, and (c) refracted ray

The power that is received by the path of the k$^{th}$ ray that reaches a single point, is given by (Diaz, N.R. & Achilli, C., 2003):

$$P_k = P_T \left( \frac{\lambda}{4\pi r} \right)^2 G_T G_R \prod_i R_i \prod_j T_j \tag{12}$$

where $P_T$ is the transmit power in Watts, $G_T$ and $G_R$ are the transmitter and receiver gains respectively, $\lambda$ is the wavelength in meters, $r$ is the total unfolded path length in meters, $R_i$ and $T_j$ are the reflection and refraction coefficients respectively (determined by equations (4) to (7)), and $i$ and $j$ are the indexes that increment over reflection and refraction respectively.

The polarization model can be simplified using techniques found in (Chizhik, D. et al., 1998). The phase, $\phi_k$, of the received field is computed from the fast fading prediction, where $\phi_k$ is a function of the unfolded path length and the number of reflections. The signal strength at a point in the aircraft can thus be evaluated using:

$$P_R = \left| \sum_k \sqrt{P_k} e^{-j\phi_k} \right|^2 \tag{13}$$

and

$$\phi_k = kr + R_{sh} \tag{14}$$

where $k$ is the wave number per meter, $r$ is the length of the path in meters, and $R_{sh}$ is the phase shift due to the reflection in radians.

### 3.3 The reflection model

Reflection is implemented according to Fermats principle. The direction of the reflected ray is found using:

$$\vec{r} = \vec{v} - 2(\vec{n} \cdot \vec{v})\vec{n} \tag{15}$$

where $\vec{v}$ is the incident ray, $\vec{n}$ is the normal to the plane of incidence, and $\cdot$ is the dot product operator. In the cabin environment we will experience a large number of reflections, from every surface. Using Fresnels coefficients, defined above, the rays will experience a phase shift of $\pi$ radians every time there is a reflection. The field power of the reflected signal becomes:

$$P_r = \sqrt{P_{r\|}^2 + P_{r\perp}^2} \tag{16}$$

where

$$P_{r\|} = P_{i\|} \cdot \left| R_{\|} \right|^2 \tag{17}$$

$$P_{r\perp} = P_{i\perp} \cdot \left| R_{\perp} \right|^2 \tag{18}$$

$$P_{i\|} = P_i \cos\theta \tag{19}$$

and

$$P_{i\perp} = P_i \sin\theta \tag{20}$$

The subscripts $r$ and $i$ represent the reflected ray and the incident ray respectively.
The GO principle can also be used to model the propagation of the waves as they hit the curved walls of the aircraft. This can be done because the radius of curvature of this surface is very large compared to the wavelength of the signal. Therefore, the incident ray is reflected at the tangential plane of the surface at the point where the incident ray impinges on the cabin wall.

### 3.4 The refraction model

The propagating signal experiences refraction whenever the ray hits an obstacle. The rays which are refracted in a direction of travel which lies outside the aircraft are assumed to be

absorbed within the material. This is because we are not interested in the rays which leave the aircraft. The direction of the refracted ray can be calculated using (Diaz, N.R. & Achilli, C., 2003):

$$\vec{t} = n\vec{v} - \left( n(\vec{v},\vec{n}) + \sqrt{1 - n^2(1-(\vec{v},\vec{n})^2)} \right)\vec{n} \tag{21}$$

where $n = n_1/n_2$ and $n_1$ and $n_2$ are the refraction indexes of the two different media.
The refracted power is calculated using the following:

$$P_t = \sqrt{P_{t||}^2 + P_{t\perp}^2} \tag{22}$$

where

$$P_{t||} = P_{i||} \cdot \left| T_{||} \right|^2 \tag{23}$$

$$P_{t\perp} = P_{i\perp} \cdot \left| T_{\perp} \right|^2 \tag{24}$$

where the subscripts $i$ and $t$ represent the incident ray and refracted ray respectively. The model assumes that the obstacles encountered by the rays have constant dielectric properties.

### 3.5 The access point model

Each IEEE 802.11a access point is assumed to have an omni-directional antenna. This simplifies things as the access point can be modelled as a point source which radiates the rays uniformly in the three-dimensional space. The Monte Carlo stochastic launching model (Diaz, N.R. & Achilli, C., 2003) is then used to model each transmitter deployed on the aircraft. This will generate rays having random directions within the cabin with equal probability. This ensures that no region within the cabin will contain more rays than another, something which would otherwise skew the results.
The one-dimensional probability density functions are given by (Diaz, N.R. & Achilli, C., 2003):

$$p_\theta(\theta) = \int_0^{2\pi} p_{\theta,\phi}(\theta,\phi)d\phi = \frac{\sin\theta}{2} \tag{25}$$

$$p_\phi(\theta) = \int_0^{\pi} p_{\theta,\phi}(\theta,\phi)d\theta = \frac{1}{2\pi} \tag{26}$$

where $\phi$ and $\theta$ are the spherical coordinates, with $0 \le \phi \le 2\pi$ and $0 \le \theta \le \pi$. These two randomly distributed variables are generated using:

$$\theta = \arccos(1 - 2\xi_1) \tag{27}$$

and

$$\phi = 2\pi\xi_2 \tag{28}$$

where $\xi_1$ and $\xi_2$ are random variables which are distributed in [0,1] and in [0,1) respectively.

### 3.6 Multipath characteristics

We know that the aircraft layout has a very high object density. These objects produce a lot of reflections and refractions as the signals propagate. Therefore, the signal strength arriving at a receiver is highly affected by the large number of multipath signals arriving at that location. These signals will be added vectorially by the receiver hardware. The impulse response can be used to obtain the channel characteristics in these scenarios.

The impulse response of the channel can be modelled using (Hashemi, H., 1993) and (Saleh, A.A.M. & Valenzuela, R.A., 1987). The time invariant impulse response is expressed as a sum of $k = 1 \ldots N$ multipath components, each having a random amplitude $a_k$, delay $\tau_k$, and phase $\theta_k$:

$$h(t) = \sum_{k=1}^{N} a_k \partial(t - \tau_k) e^{j\theta_k} \tag{29}$$

The three main distributions that are used in communications theory to model multipath effects are the (i) Nakagami, (ii) Rician, and (iii) Rayleigh distributions. At a point inside the aircraft, the signal will experience different fading characteristics as the number of multipath components reaching that point varies. The Rician distribution is more appropriate to scenarios having a dominant line-of-sight signal, which is not the case for the cabin environment. The choice is therefore between the Nakagami and the Rayleigh distribution models. The study in (Can De Beek, J.J. et al, 2002) shows that the multipath distribution of an indoor channel can be better represented by a Nakagami distribution. This distribution is characterised by the cumulative density function:

$$p(x, \mu, \omega) = \frac{2\mu^{\mu}}{\Gamma(\mu)\omega^{\mu}} x^{2\mu-1} \exp\left(-\frac{\mu}{\omega} x^2\right) \tag{30}$$

where $\mu$ is a shape parameter and $\omega$ controls the spread of the distribution.

These distribution parameters have to be extracted from the simulation model. In order to do this, the total number of multipath rays and the maximum and minimum delay times must be recorded for each location inside the cabin. The area inside the cabin is divided into areas, called cells. All the data that is located within the same cell number, which represents the cell distance from the transmitter, is clustered together. The Nakagami model is then fit to this data. Hence, this will give a list of fit parameters that model the multipath propagation inside the cabin.

### 3.7 Time dispersion parameters

The IEEE 802.11a channel parameters are characterised by the time dispersion parameters. The main components are composed of: (i) the mean excess delay, and (ii) the root-mean-square (rms) delay spread. These parameters give an estimate of the expected performance that the wireless system will achieve if deployed in the cabin environment. These parameters are then input to the top level IEEE 802.11a system model to obtain the bit error

rate (BER) of the channel. The BER results are then used to get an estimate of the quality of service (QoS) and other data transmission characteristics.
The mean excess delay, $\tau_m$, is defined as:

$$\tau_m = \frac{\sum_k \left\{ \left( \tau_k - \tau_1 \right) a_k^2 \right\}}{\sum_k a_k^2} \tag{31}$$

while the rms delay spread, $\tau_{rms}$, is:

$$\tau_{rms} = \left( \frac{\sum_k \left\{ \left( \tau_k - \tau_m - \tau_1 \right)^2 a_k^2 \right\}}{\sum_k a_k^2} \right)^{\frac{1}{2}} \tag{32}$$

where $\tau_k$ is the delay of the k$^{th}$ multipath ray with a normalised amplitude of $a_k$, and $\tau_1$ is the delay of the line-of-sight signal.

### 3.8 Coherence bandwidth

The coherence bandwidth, $B_c$, is a measure of the range of frequencies over which two frequency components are likely to have amplitude correlation (Rappaport, T.S., 2002). This bandwidth is related to the rms delay spread. Two signals centered at frequencies that have a separation which is less than or equal to $B_c$ will have similar channel impairments. Otherwise the signals can experience frequency selective fading.
For a frequency correlation function of 0.9 or above, $B_c$ can be approximated by (Lee, W.C.Y., 1989):

$$B_c \approx \frac{1}{50\tau_{rms}} \tag{33}$$

while for a frequency correlation function above 0.5, this approximation becomes:

$$B_c \approx \frac{1}{5\tau_{rms}} \tag{34}$$

Results within a business jet can be found in (Debono, C.J. et al, 2009). For the system to guarantee that the receiver does not experience inter-symbol interference and/or inter-channel interference, the guard interval at any location within the cabin must be less than 800ns (as specified in the IEEE 802.11a standard). Moreover, the system will only introduce flat fading if the coherence bandwidth is greater than the bandwidth of the subcarriers, which is equal to 312.5kHz.

## 4. Simulation results

### 4.1 The cabin model

A three-dimensional model of the cabin can be developed using any computer aided design (CAD) software. This model can then be imported in the simulation software, which for this

work was developed in Matlab®. The propagation characteristics presented here relate to an Airbus A340-600 but further results on a Dassault Aviation business jet can be found in (Debono, C.J. et al, 2009) and (Chetcuti, K. et al, 2009). The structure of the aircraft is modelled through a cylinder which represents the fuselage and a horizontal plane to model the floor. Furniture and the stowage bins were also included as shown in Figure 4.

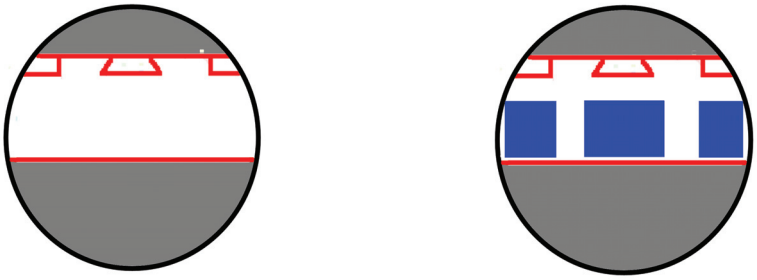



Fig. 4. Cross-section of cabin without seats (a), and with seats (b).

The seats have a specific thickness and are modelled as two intersecting planes. The dielectric constant, permittivity and conductivity depend on the material used. Typical values of the materials used inside the cabin of the aircraft are given in Table 2.

| Material | Electric Conductivity | Relative Permittivity |
|---|---|---|
| Aluminium | 4E7 | Inf |
| Leather | 1E-2 | 3 |
| Wood | 1E-2 | 3 |

Table 2. Electrical characteristics of materials used inside a cabin

### 4.2 Simulation environment

The ray tracing algorithm is implemented in Matlab®. A flow chart of the main algorithm is shown in Figure 5. The geometry file used by the developed simulator represents the typical environment of the aircraft under test. The signal strength propagation map is determined by launching 200,000 rays from each transmitter antenna. The equivalent isotropic radiated power from each access point is 30dBm. At any particular cell inside the cabin, the signal strength is determined by summing the power levels of all the rays passing through that point. This implies that the received signal is a distorted version of the transmitted signal.

As discussed in section 3.5, the starting direction of each ray is determined using Monte Carlo techniques, where two random numbers, representing the angles in spherical coordinates, in the range 0 to $2\pi$ and 0 to $\pi$ respectively are generated. Each ray is traced one cell size at a time, where at each cell position, the simulator assesses whether the ray is still inside the aircraft. If it is found to lie outside the aircraft, then the trace ends there and the

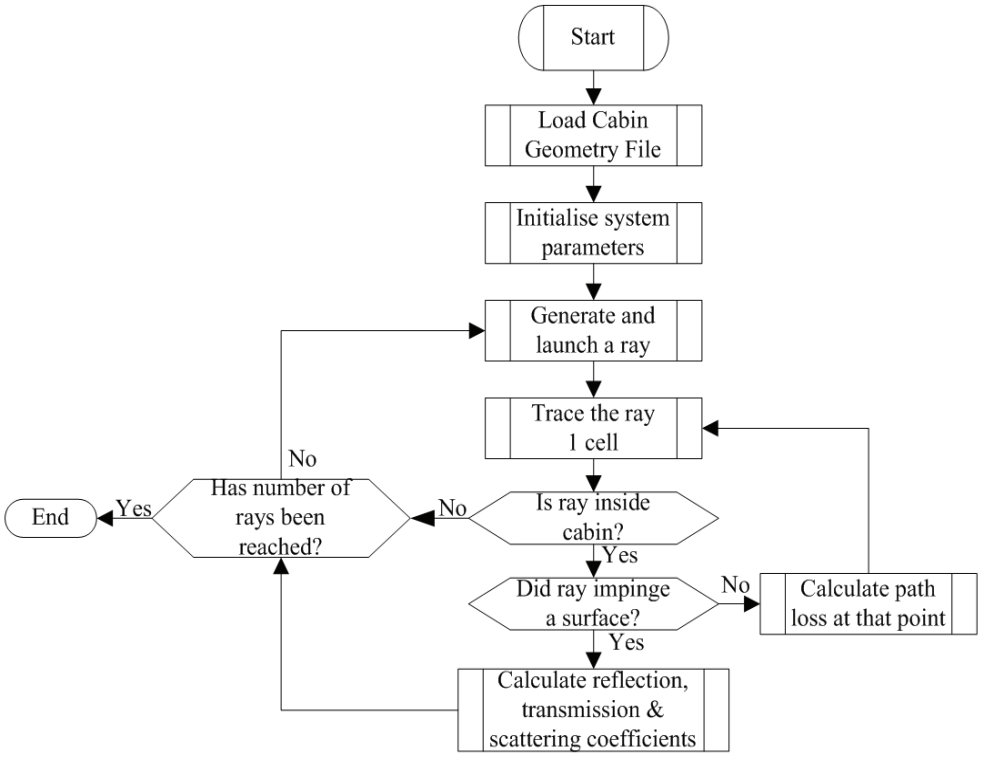



Fig. 5. Flowchart of the ray tracing method

simulator goes back to the antenna to start a new ray trace. If the ray is still inside the aircraft, the propagation loss is calculated. The new power level is compared to the predetermined threshold, which in our case is set to -120dBm, and if it is above this threshold a check is performed to test whether the ray has impinged on a surface. The -120dBm level is well below the minimum detectable signal for IEEE 802.11a, but because of the multipath effects some margin is required to allow for the eventuality that the vectorially summed power level could still exceed the -100dBm limit defined in the standard. The received signal strength at the receiver affects the signal-to-noise ratio (SNR) posing a limit on the maximum useable data rate for error free communication.

### 4.3 Results

Placing just one access point inside the aircraft limits the number of users that can access the network. This occurs because of the limited capacity that would be offered and the radio propagation coverage that can be obtained with reasonable transmit power levels. The higher the power emanating from the access point, the more interference it is likely to cause to the aircraft's electronics. The simulator developed can be used to determine the optimum number of access points and their position within the aircraft. An analysis for the optimum

antenna locations for a Universal Mobile Terrestrial System (UMTS) is found in (Debono, C.J. & Farrugia, R.A., 2008).

The resulting propagation map for the A340-600, using four IEEE 802.11a access points, is shown in figures 6 to 9. Figure 6 presents the view from the antenna plane, Figure 7 shows the top view at the middle of the aircraft, Figure 8 shows the side view, while Figure 9 shows cross-sections looking from the front of the aircraft.

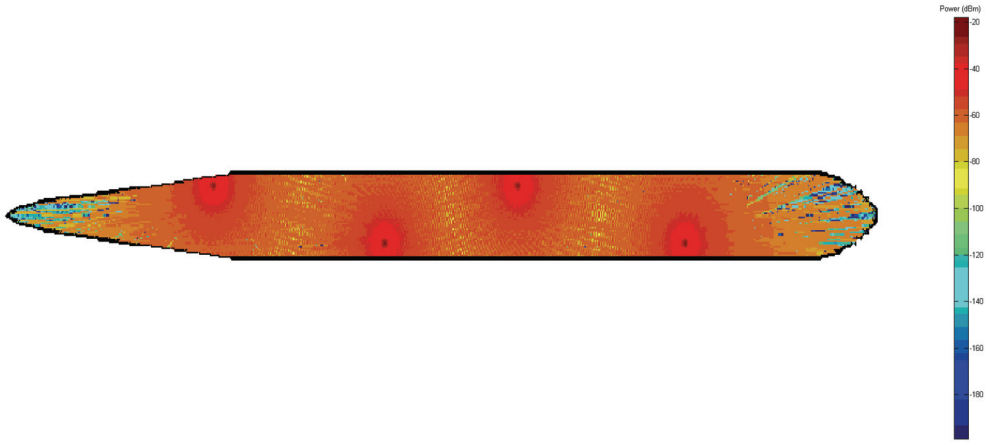



Fig. 6. Propagation map at the antenna plane. The four access points are shown by the areas of maximum signal concentration.

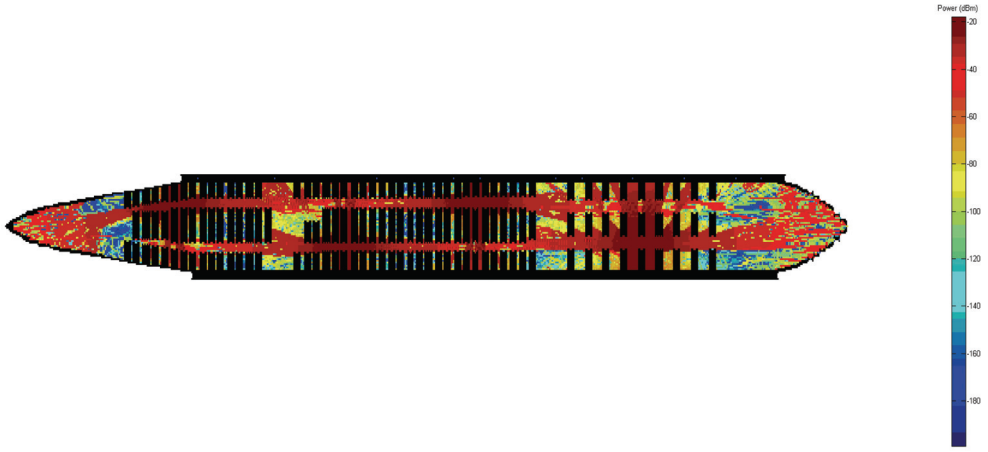



Fig. 7. Propagation map at the middle of the aircraft as seen from the top.

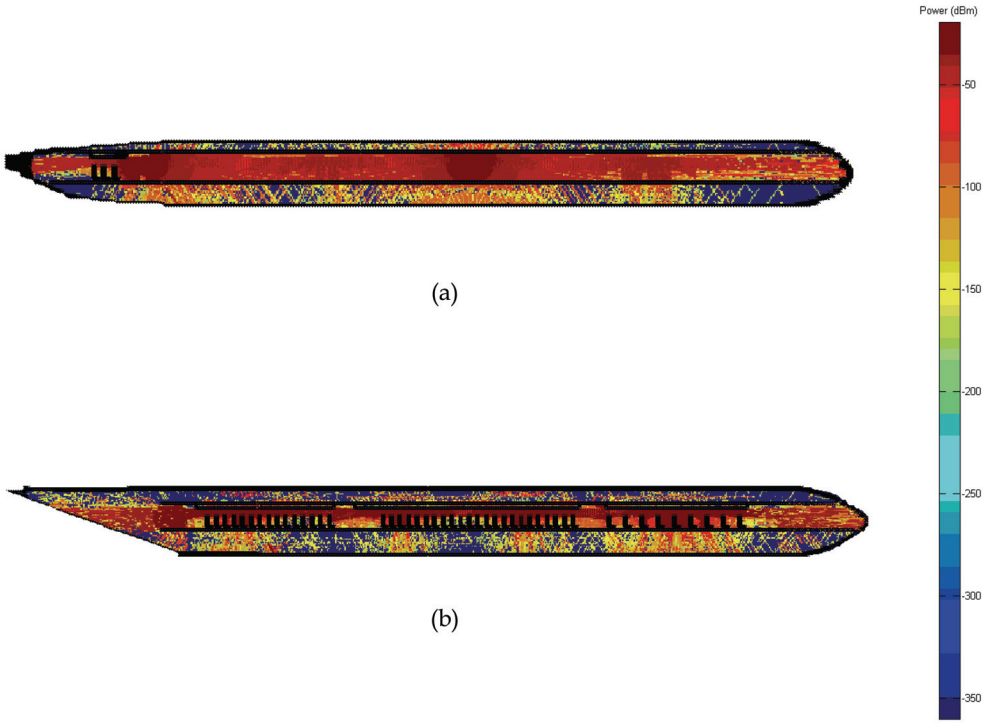



Fig. 8. The propagation map as seen from the side; (a) aisle, and (b) across a column of seats.

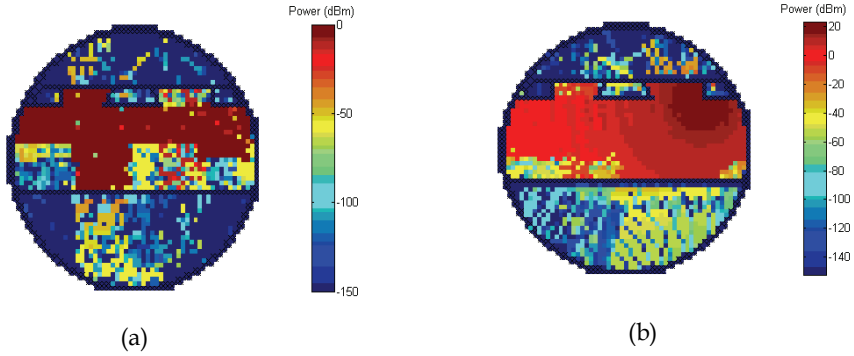



Fig. 9. The propagation map as seen from the aisle; (a) at the front row, and (b) near one of the access points.

The simulator has also been used to model the propagation inside a business jet. A measurement campaign was done in this case to compare the results obtained and determine the confidence level of the simulations. The results have been presented in (Chetcuti, K. et al, 2009) and show that the model is reasonably accurate, especially within the cabin area.

## 5. Conclusion

This chapter has given the theoretical background necessary to develop a radio propagation map for an IEEE 802.11a system. The method is based on ray tracing techniques which rely on the theory of geometric optics. This solution can be applied to both commercial aircraft, like the A340-600 used in this case, and business jets. The flexibility of the simulator allows easy modifications to simulate different frequency bands and different furniture location and material. This makes the simulator very attractive especially in the business jet environment, where the interior furniture of each aircraft is specifically designed for each customer.

The simulator allows the user to insert the number of access points required and their location. Using an intelligent optimisation technique, such as neural networks, genetic algorithms and support vector machines, one can find the optimum number of access points and their optimum location within the aircraft. This can be done given some constrains imposed by the wiring system of the aircraft.

Moreover, the propagation map gives an idea of the electromagnetic radiation field strength hitting the fuselage of the aircraft. A similar method can be used for each portable device held by each passenger in the aircraft to simulate the uplink. Therefore, the designer can estimate the electromagnetic interference that is generated by the system. Through optimum design of the system the electromagnetic interference can be kept within acceptable limits and thus ensure that no interference occurs with the aircraft's navigation and control system.

## 6. Acknowledgements

The authors would like to thank Mr. Serge Bruillot from Dassault Aviation for providing us with the model file of their Falcon X7 business jet and for the measurement campaign referenced in the text.

This work forms a small part of the project E-Cab which is financially supported under the European Union 6th Framework Programme (FP6) (E-Cab Website, 2008). The E-Cab consortium is made up of 30 partners from 13 countries across Europe. The authors are solely responsible for the contents of the chapter which does not represent the opinion of the European Commission.

## 7. References

Bothias, L. (1987), *Radio Wave Propagation,* McGraw Hill Inc., New York, USA.

Can De Beek, J.J., Odling, P., Wilson, and S.K., Bojesson, P.O. (2002), Orthogonal Frequency-Division Multiplexing, *International Union of Radio Science,* 2002.

Chetcuti, K., Debono, C.J., Farrugia, R.A., and Bruillot, S. (2009), Wireless Propagation Modelling Inside a Business Jet, *Proceedings of Eurocon 2009,* May 2009, pp. 1640-1645.

Chizhik, D., Ling, J., and Valenzuela, R.A. (1998), The Effect of Electric Field Polarization on Indoor Propagation, *IEEE 1998 International Conference on Universal Personal Communications,* October 1998, pp. 459-462.

Commission Decision of […] on harmonised conditions of spectrum use for the operation of mobile communication services on aircraft (MCA services) in the Community – Commision of the European Communities, April 2008.

Crow, B.P., Widjaja, I., Kim, J.G., and Sakai, P.T. (1997), IEEE 802.11 Wireless Local Area Networks, *IEEE Communications Magazine,* September 1997, pp. 116-126.

Debono, C.J., and Farrugia, R.A. (2008), Optimization of the UMTS Network Radio Coverage On-board an Aircraft, *Proceedings of the 2008 IEEE Aerospace Conference,* March 2008.

Debono, C.J., Chetcuti, K. and Bruillot, S. (2009), 802.11a Channel Parameters Characterization on board a Business Jet, *Proceedings of the 2009 IEEE Aerospace Conference,* March 2009.

Diaz, N.R., and Achilli, C. (2003), Cabin Channel Characterization for Personal Communications via Satellite, *Proceedings of the 21st International Communications Satellite Systems Conference and Exhibit,* 2003.

E-Cab Consortium Website, Online: http://www.e-cab.org

Hashemi, H. (1993), The Indoor Radio Propagation Channel, *Proceedings IEEE,* vol.81, July 1993.

IEEE Std 802.11a-1999(R2003), *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications High-speed Physical Layer in the 5 GHz Band*, 2003.

James, G.L. (1986), *Geometric Theory of Diffraction for Electromagnetic Waves,* Peter Peregrinus Ltd., London, UK.

Landron, O., Feuerstein, M.J., and Rappaport, S. (1993), In Situ Microwave Reflection Coefficient Measurements for Smooth and Rough Exterior Wall Surfaces, *Proceedings of the IEEE 43rd Vehicular Technology Conference,* May 1993, pp. 77-80.

Lee, W.C.Y. (1989), *Mobile Cellular Telecommunications Systems,* McGraw Hill Publications, New York, USA.

Paul, T.K., and Ogunfunmi, T. (2008), Wireless LAN Comes of Age: Understanding the IEEE 802.11n Amendment, *IEEE Circuits and Systems Magazine*, First Quarter 2008, vol. 8, no. 1, pp 28 – 54.

Peled, A., and Ruiz, A. (1980), Frequency Domain Data Transmission using Reduced Computational Complexity Algorithms, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing,* 1980, pp. 964-967.

Rappaport, T.S. (2002), *Wireless Communications Principles and Practice,* Prentice Hall, New Jersey, USA.

Saleh, A.A.M., and Valenzuela, R.A. (1987), A Statistical Model for Indoor Multipath Propagation, *IEEE Journal on Selected Areas Communications,* February 1987, pp. 128-137.

Yomo, H., Nguyen, C.H., Kyritsi, P., Nguyen, T.D., Chakraborty, S.S., and Prasad, R., PHY and MAC Performance Evaluation of IEEE 802.11a WLAN over Fading Channels, *Institution of Electronics and Telecommunications Engineers (IETE),* vol. 51, no. 1, January 2005, pp. 83-94.

# 19

# Air Traffic Control Tracking Systems Performance Impacts with New Surveillance Technology Sensors

Baud Olivier, Gomord Pierre, Honoré Nicolas, Lawrence Peter,
Ostorero Loïc, Paupiah Sarah and Taupin Olivier
*THALES*
*FRANCE*

## 1. Introduction

Nowadays, the radar is no longer the only technology able to ensure the surveillance of air traffic. The extensive deployment of satellite systems and air-to-ground data links lead to the emergence of other means and techniques on which a great deal of research and experiments have been carried out over the past ten years.
In such an environment, the sensor data processing, which is a key element of an Air Traffic Control center, has been continuously upgraded so as to follow the sensor technology evolution and, at the same time, ensure a more efficient tracking continuity, integrity and accuracy.
In this book chapter we propose to measure the impacts of the use of these new technology sensors in the tracking systems currently used for Air Traffic Control applications.
The first part of the chapter describes the background of new-technology sensors that are currently used by sensor data processing systems. In addition, a brief definition of internal core tracking algorithms used in sensor data processing components, is given as well as a comparison between their respective advantages and drawbacks.
The second part of the chapter focuses on the Multi Sensor Tracking System performance requirements. Investigation regarding the use of Automatic Dependent Surveillance – Broadcast reports and/or with a multi radars configuration, are conducted.
The third part deals with the impacts of the "virtual radar" or "radar-like" approaches that can be used with ADS-B sensors, on the multi sensor tracking system performance.
The fourth and last part of the chapter discusses the impacts of sensor data processing performance on sub-sequent safety nets functions that are:

- Short term conflict alerts (STCA),
- Minimum Safe Altitude Warnings (MSAW), and
- Area Proximity Warnings (APW).

## 2. Air traffic control

Air Traffic Control (ATC) is a service provided to regulate the airline traffic. Main functions of the ATC system are used by controllers to (i) avoid collisions between aircrafts, (ii) avoid

collisions on maneuvering areas between aircrafts and obstructions on the ground and (iii) expediting and maintaining the orderly flow of air traffic.

### 2.1 Surveillance sensors

Surveillance sensors are at the beginning of the chain: the aim of these systems is to detect the aircrafts and to send all the available information to the tracking systems.

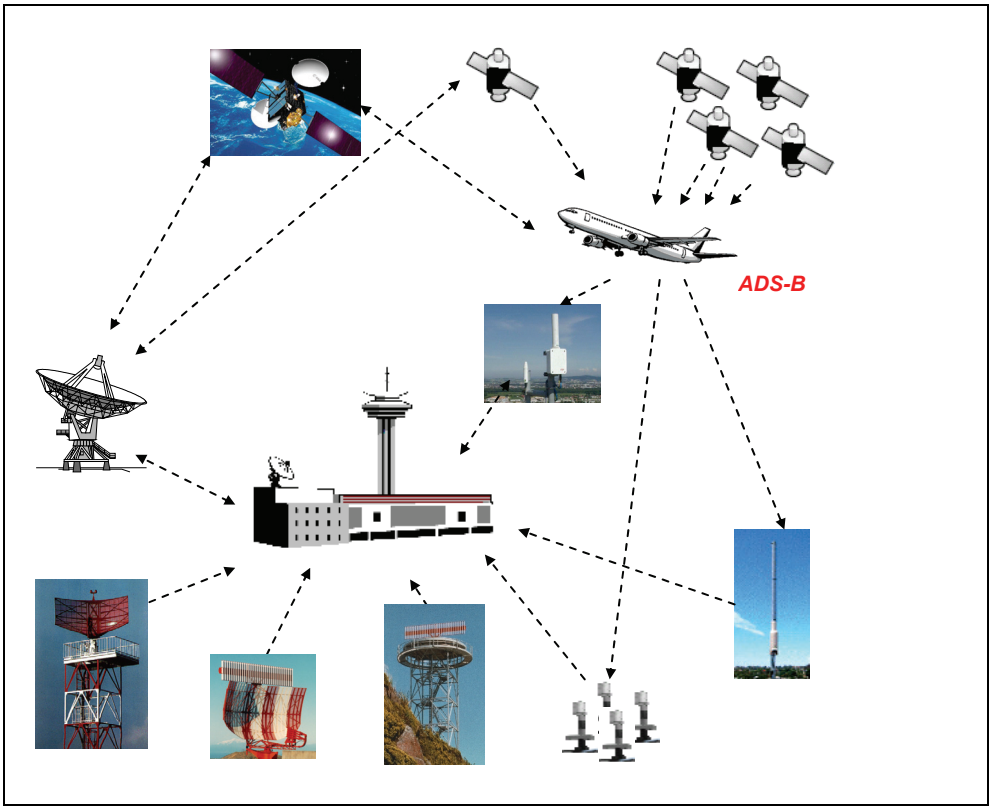



Fig. 1. Surveillance sensor environment

Current surveillance systems use redundant primary and secondary radars. The progressive deployment of the GPS-based ADS systems shall gradually change the role of the ground based radars. The evolution to the next generation of surveillance system shall also take into account the interoperability and compatibility with current systems in use.
The figure 3 above shows a mix of radar, ADS and Multilateration technologies which will be integrated and fused in ATC centers in order to provide with a high integrity and high accuracy surveillance based on multiple sensor inputs.

#### 2.1.1 Primary Surveillance Radar (PSR)

Primary radars use the electromagnetic waves reflection principle. The system measures the time difference between the emission and the reception of the reflected wave on a target in

order to determine its range. The target position is determined by measuring the antenna azimuth at the time of the detection.
Reflections occur on the targets (i.e. aircrafts) but unfortunately also on fixed objects (buildings) or mobile objects (trucks). These kind of detections are considered as parasites and the "radar data processing" function is in charge of their suppression.
The primary surveillance technology applies also to Airport Surface Detection Equipment (ASDE) and Surface Movement Radar (SMR).

#### 2.1.2 Secondary Surveillance Radar (SSR)

Secondary Surveillance Radar includes two elements: an interrogative ground station and a transponder on board of the aircraft. The transponder answers to the ground station interrogations giving its range and its azimuth.
The development of the SSR occurs with the use of Mode A/C and then Mode S for the civil aviation.
Mode A/C transponders give the identification (Mode A code) and the altitude (Mode C code). Consequently, the ground station knows the 3-dimension position and the identity of the targets.
Mode S is an improvement of the Mode A/C as it contains all its functions and allows a selective interrogation of the targets thanks to the use of an unique address coded on 24 bits as well as a bi-directional data link which allows the exchange of information between air and ground.

#### 2.1.3 Multilateration sensors

A multilateration system is composed of several beacons which receive the signals which are emitted by the aircraft transponder. The purpose is still to be able to localize the aircraft. These signals are either unsolicited (squitters) or answers (SSR or Mode S) to the interrogations of a nearby interrogator system (can be a radar). Localization is performed thanks to the Time Difference Of Arrival (TDOA) principle. For each beacons pair, hyperbolic surfaces whose difference in distance to these beacons is constant are determined. The aircraft position is at the intersection of these surfaces.
The accuracy of a multilateration system depends on the geometry of the system formed by the aircraft and the beacons as well as the precision of the measurement time of arrival.
Nowadays, multilateration is used mainly for ground movement's surveillance and for the airport approaches (MLAT). Its use for en-route surveillance is on the way of deployment (Wide Area Multilateration (WAM)).

#### 2.1.4 Automatic Dependant Surveillance – Broadcast (ADS-B)

The aircraft uses its satellite-based or inertial systems to determine and send to the ATC center its position and other sort of information. Aircraft position and speed are transmitted one time per second at least.
ADS-B messages (squitters) are sent, conversely to ADS-C messages which are transmitted via a point-to-point communication. By way of consequence, the ADS-B system is used both for ATC surveillance and on-board surveillance applications.

### 2.2 Sensor data processing

As shown in figure 5 hereunder, a sensor data processing is composed generally of two redundant trackers. Radar (including Surface Movement Radar) data are received directly by the trackers while ADS-B and WAM sensor gateways help in reducing the data flow as well as checking integrity and consistency.

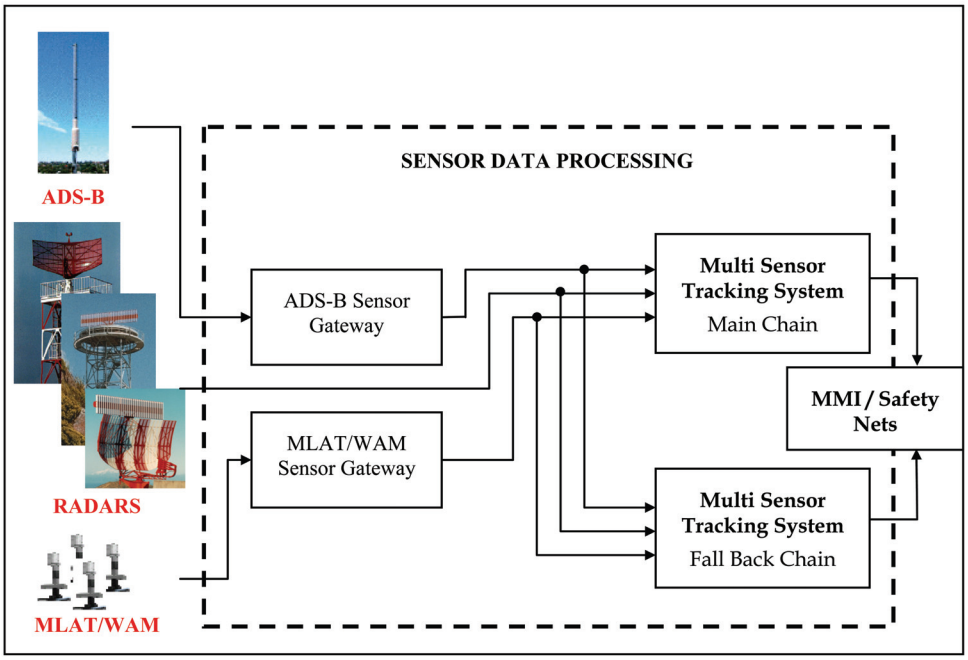



Fig. 2. Sensor Data Processing

As shown in figure 5 above, trackers are potentially redundant in order to prevent from sub-systems failure.

Sensor Data Processing architectures have been shown and discussed in details in (Baud et al., 2009).

## 3. Multi sensor tracking performance

### 3.1 Sensor characteristics and scenarios

Radar sensor characteristics are available in table 1.

ADS-B sensor characteristics are available in table 2.

Scenarios that are used to compare the horizontal tracking performance among all possible sensor configurations are composed of straight line motion followed by a set of maneuvers including turn with different bank angles.

These scenarios are mainly derived from the EUROCONTROL performances described in (EUROCONTROL 1997). They have been used to provide relative comparisons. Results extrapolation to live data feeds must take into account the sensor configuration, the traffic repartition over the surveillance coverage and specific sensor characteristics.

| RADAR CHARACTERISTICS | PSR | SSR | PSR + SSR |
|---|---|---|---|
| Range | Up to 250 NM | Up to 250 NM | Up to 250 NM |
| Antenna rotation time | 4 up to 12 s | 4 up to 12 s | 4 up to 12 s |
| Probability of detection | > 90 % | > 97 % | > 95 % |
| Clutter density (number of plots per scan) | 40 | | |
| Nominal measurement accuracy:<br>- Range (m)<br>- Azimuth (deg) | <br>40<br>0.07 | <br>30<br>< 0.06 | <br>30<br>< 0.06 |
| Measurement quantization (ASTERIX standard):<br>- Range (NM)<br>- Azimuth (deg) | <br>1/256<br>0.0055 | <br>1/256<br>0.0055 | <br>1/256<br>0.0055 |
| SSR false plots (%):<br>- Reflection<br>- Side lobes<br>- Splits | | <br>< 0.2<br>< 0.1<br>< 0.1 | <br>< 0.2<br>< 0.1<br>< 0.1 |
| Mode A code detection probability | | > 98 % | > 98 % |
| Mode C code detection probability | | > 96 % | > 96 % |
| Mode C measurement accuracy (m) | | 7.62 | 7.62 |
| Time stamp error | <= 100 ms | <= 100 ms | <= 100 ms |
| Nominal time stamp error (time disorder) | 50 ms | 50 ms | 50 ms |

Table 1. Radar sensor characteristics

| ADS-B CHARACTERISTICS (1090ES) | NOMINAL VALUE |
|---|---|
| Range | 250 NM |
| Refresh period | 1s |
| Probability of detection | > 95% |
| Nominal Position Standard Deviation | 10 m |
| Figure Of Merit | 7 |
| Altitude Standard Deviation | 25 fts |
| ADS-B transponder consistency | 100% |

Table 2. ADS-B sensor characteristics

### 3.2 Simulation results

Multi sensor tracking accuracy has been evaluated among 5 sensor configurations that are:

- PSR only: radar with 4s revolution period,
- SSR only: radar with 4s revolution period,
- Multi radars configuration including 1 PSR radar, 1 SSR radar and 1 PSR + SSR radar,
- ADS-B only: one ADS-B ground station at 1s update rate,
- Multi sensors configuration that includes both multi radars configuration and the ADS-B ground station.

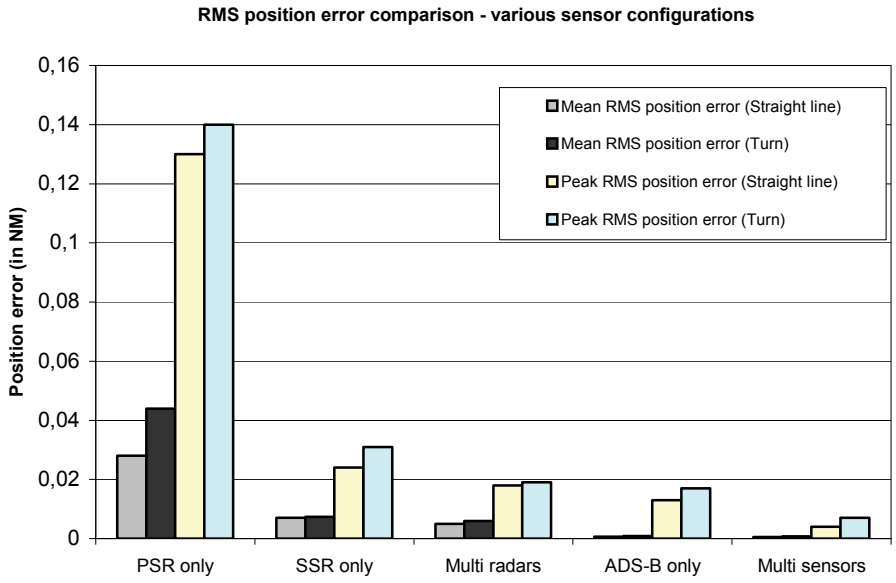



Fig. 3. RMS position error comparison

Multi sensor tracking coverage helps to globally improve the tracking performance in term of:

- Latency metrics: Latency reduced in update/broadcast modes to several hundreds of milliseconds instead of several seconds thanks to:
  - the update rate of new technology sensors (1s) compared to radar sensors (at least 4s and up to 12s),
  - the variable update technique used which does not make any bufferisation of new technology sensors data.
- Continuity/integrity metrics:
  - Possible reduction of multi sensor tracks broadcast cycle thanks to the update rate of new-technology sensors,
  - Quicker track initiation.
  - Bigger coverage areas including airport areas (MLAT) and desert areas (ADS-B) where no radar data are available,
- Accuracy metrics:
  - Improved accuracy even if the multi sensor configuration relies on one ADS-B ground station only, as can be seen on figure 3.

## 4. Virtual radar emulation – "radar like" solutions

As can be seen in the previous paragraph, introduction of new technology sensors in the tracking systems that are used for Air Traffic Control applications improves the global

performance of the systems compared to what is used at the current time (multi radar tracking systems).
Use of these new technology sensors require an evolution that leads from multi-radar tracking systems to multi-sensor tracking systems.

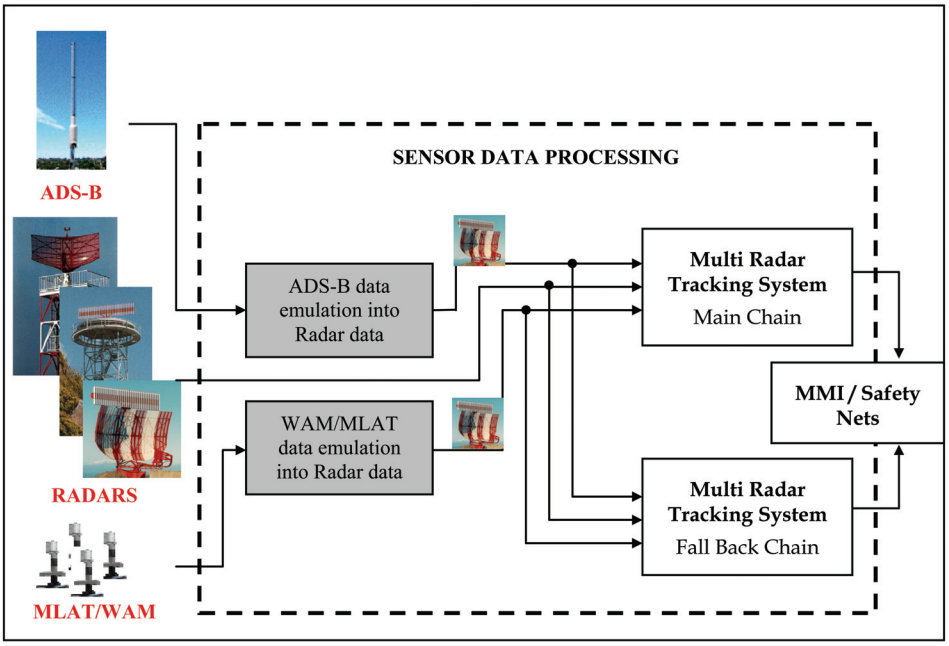



Fig. 4. Virtual radar concept

However, in most cases, the transition from the existing radar based surveillance means (network, radar data processing, …) cannot be done straight away, and the Air Navigation Service Providers mainly ask for an integration of these new sensors into the existing system by a "radar-like" or "virtual radar" approach. Then, decisions could be done to have the WAM/MLAT reports or ADS-B reports appearing as if they are from any radars. This process is explained in details in (Thompson et. Al). This concept can be shown in figure 5.
Most of the advantages of the "radar-like" or "virtual radar" approaches are discussed in (Thompson et. Al) and in (Whitman et. Al).

| | **"Radar like" approach with new technology sensors as ADS-B and WAM** | **Multi radars tracking system** |
|---|---|---|
| Multi sensor coverage | Multi sensor coverage allowed: provides coverage where none currently exist. | Only multi radars coverage. When an area is covered by ADS-B only, no control can be done. |
| Transition from former to new technology | Allow transition and test environment | New technology sensors not used in existing systems |

Table 3. "Radar like" solution main advantages

A comparison between a "radar like" approach and an integrated multi sensor fusion with Variable Update technique is done in the following table.

| | **"Radar like" approach** | **Integrated multi sensor fusion with Variable Update technique** |
|---|---|---|
| Existing radar data network impacts | Degrade the quality of the ADS-B / WAM report by introducing an additional latency (at least 1s) to buffer the reports.<br>The refresh rate is increased to typically 4s (3 report ignored upon 4) or 12s (11 reports ignored upon 12). | No latency introduced by any radar data network.<br>The refresh rate is the one provided by the sensor itself. |
| Time stamping | Depending on the radar data format, the time stamping is sometimes not available. | Time stamping available in the ADS-B and WAM standard. |
| Fitting accurate data into useless radar format impacts | This approach is not able to associate a correct standard deviation to the polar radar coordinates. For Radar, the error standard deviation in range and azimuth are fixed. For ADS-B / WAM report, the standard deviation is not constant and mainly depends either on the satellite configuration / Inertial Navigation System precision/bias or on the geometry of the receivers. | Information available in the ADS-B / WAM standards. |
| Down-linked Aircraft Parameters (DAPs) | Does not allow the transmission of DAPs information including Mode S data if CD2 or ASTERIX Category 001 /002 is used to transmit ADS-B / WAM data | Information available in the ADS-B / WAM standards. |

Table 4. "Radar like" solution discussion

Figure 5 provides a comprehensive comparison of the RMS position error accuracy between three configurations:

- ADS-B data are fitted into a multi sensor tracking system using Multiple Report Variable Update technique,
- ADS-B data fitted into standard radar data and multi sensor tracking system makes use of these ADS-B data as they are radar ones,
- ADS-B data fitted into useless radar data format (introducing high quantization in range and in azimuth: Common Digitizer 2 format) and multi sensor tracking system makes use of these ADS-B data as they are radar ones.

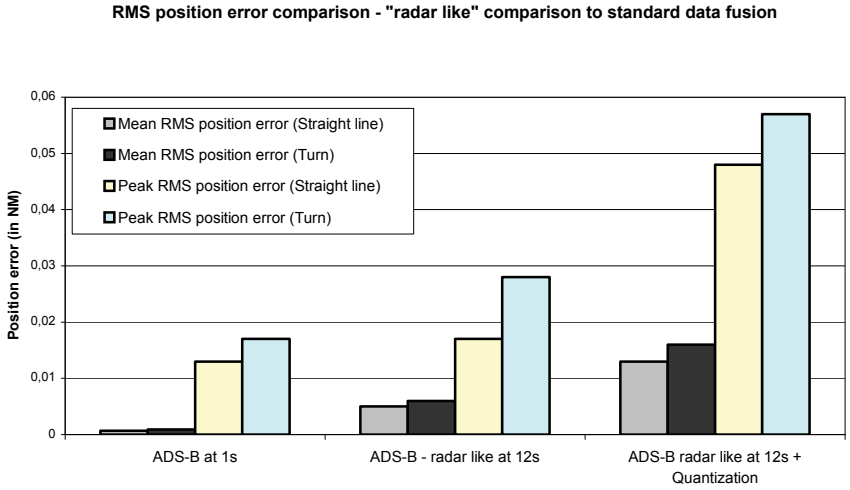



Fig. 5. RMS position error comparison between "radar like" and standard data fusion

By way of conclusion, we can say that:

- the "radar like" solution is interesting, whatever the kind of coverage:
  - when the tracking system is based on a track-to-track data fusion technique, and
  - when the ADS-B data has a high level of integrity.
- the "radar like" solution is interesting only when the area to cover is not yet covered by other kind of sensors when the existing tracking system uses a multiple report variable update technique,
- the accuracy of "radar-like" solution is worse than if we use the available ADS-B standards (see figure 5),
- the gain in term of accuracy is very low when the area is covered by multiple radars.

## 5. Safety Nets impacts

Safety Nets are functions intended to alert air traffic controllers to potentially hazardous situations in an effective manner and with sufficient warning time so that they can issue appropriate instructions to resolve the situation.
Safety Nets monitoring systems typically include:

- Short term conflict alerts (STCA),
- Minimum safe altitude warnings (MSAW),
- Area proximity warnings (APW).

### 5.1 Definitions

STCA (Short Term Conflict Alert) checks possible conflicting trajectories in a time horizon of about 2 or 3 minutes and alerts the controller prior the loss of separation. The algorithms used may also provide in some systems a possible vectoring solution, that is, the manner in which to turn, descend, or climb the aircraft in order to avoid infringing the minimum safety distance or altitude clearance.

Minimum Safe Altitude Warning (MSAW) is a sub-system that alerts the controller if an aircraft appears to be flying too low to the ground or will impact terrain based on its current altitude and heading.
Area Penetration Warning (APW) is a tool that informs any controller that a flight will penetrate a restricted area.

### 5.2 Performance impacts discussion

The most widely used safety net is STCA which is mandatory in many areas and appreciated by air traffic controllers. STCA requires short term trajectory predictions of up to 2 minutes. This is the maximum time over which it is considered valid to predict aircraft paths based solely on surveillance data. The trajectory data are
The utility of safety nets depends on both the reliability of conflict detection and the false alert rate. The false alert rate tends to be highest in the areas where such tools are most needed i.e. in the Terminal Major Areas and particularly during the approach and climb out phases of flight.
Safety nets function directly benefits from the more accurate state vector (position and velocity for both horizontal and vertical axis) provided by any multi sensor tracking system. Indeed, the use of more accurate information and Down-linked Aircraft Parameters such as ADS-B or MLAT/WAM, specifically in Terminal Major Areas, improves the tracking in term of accuracy.
These enhancements of the safety nets ensure safer and more efficient operations, by taking into account the development of new approach and climb procedures and by generalizing the use of user defined routes and closely spaced route networks.
The possibility of using additional information (such as Aircraft Derived Data) for improving prediction (with regard to safety issues) needs to be mentioned, as well as the technical feasibility of adapting safety nets separation parameters to aircraft types.

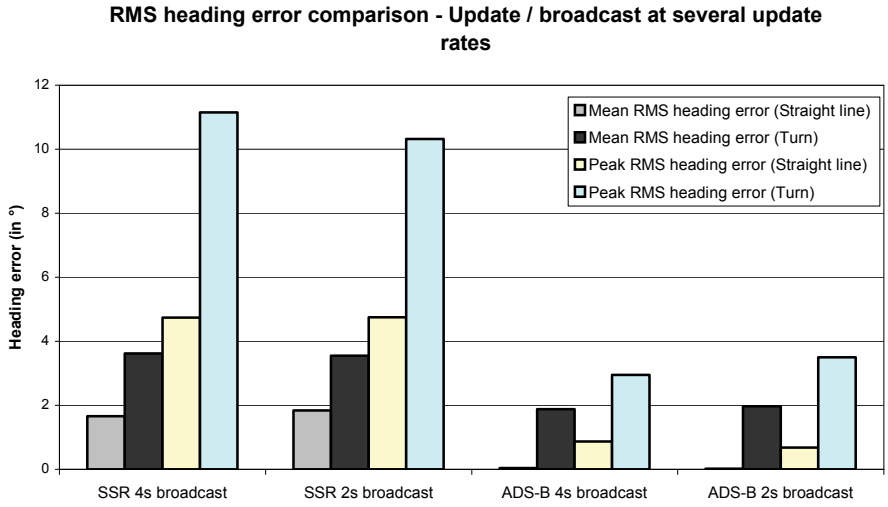



Fig. 6. RMS heading error comparison between update and broadcast at several update rates

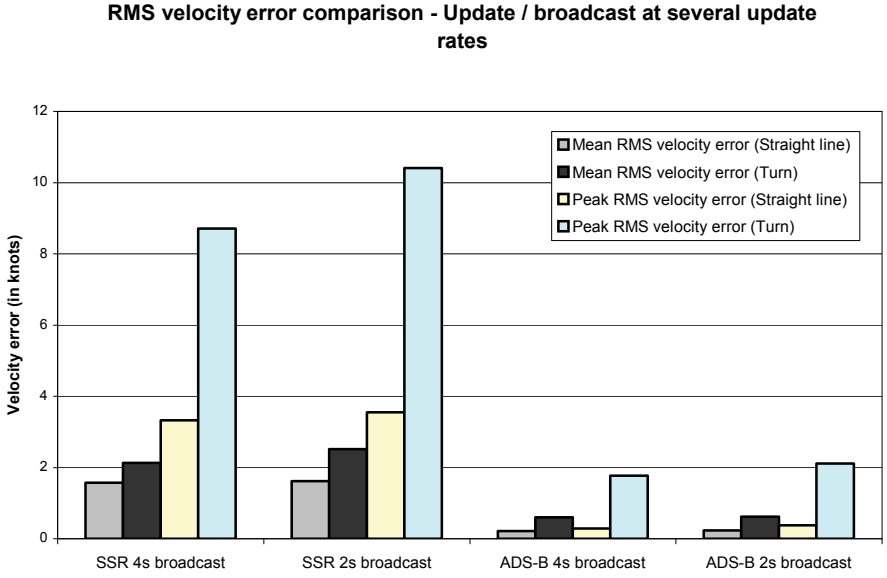



Fig. 7. RMS velocity error comparison between update and broadcast at several update rates

Multi sensor tracking performance helps to globally improve the STCA sub-system performance in term of:

- Quicker STCA detection thanks to the reduction of multi sensor tracks broadcast rate:
  - the update rate of new technology sensors (1s) compared to radar sensors (at least 4s and up to 12s),
  - the variable update technique used which does not make any bufferisation of new technology sensors data.
- Reduction of tolerances required for STCA,
- More accurate multi sensor track velocity vector as can be seen on figures 6 and 7 that leads to less false STCA's, especially for maneuvering aircrafts,
- Transmission of down-linked parameters including rate of turn and trajectory intent information that helps the STCA to enhance and predict the track state vector more accurately.

## 6. Conclusion

Nowadays, the development of advanced ATM systems is realized by the implementation of advanced means of communication, navigation and surveillance for air traffic control (CNS/ATM).

The definition of a new set of surveillance standards has allowed the emergence of a post-radar infrastructure based on data-link technology. The integration of this new technology into gate-to-gate architectures has notably the following purposes:

- fluxing air traffic which is growing continuously,
- increasing safety related to aircraft operations,

- reducing global costs (fuel cost is increasing quickly and this seems to be a long-term tendency), and
- reducing radio-radiation and improving the ecological situation.

In this context, sensor data processing will continue to play its key rule and its software as well as its hardware architecture is expected to evolve in the meantime. In a previous paper (see (Baud et. al., 2009)), we investigated the past and future of the sensor data processing architectures. In this paper, we have demonstrate the interest to integrate new technology sensors either in existing centers through the use of "radar-like" solutions (suitable for Non Radar Area only) or in future ATC centers in order to improve the global performance of the system.

The accuracy performances that can be seen in this paper have been achieved under the hypothesis that the new technology sensors are really accurate and have a high level of integrity. However, it's not completely the "real world" and we propose to discuss the ways to integrate inaccurate or inconsistent sensor data into multi sensor tracking systems for ATC applications in a future paper.

## 7. References

Bar Shalom, Y. (1989). *Multitarget-Multisensor Tracking: Advanced Applications*, Artech House, ISBN 978-0964831223, Norwood

Bar Shalom, Y. (1992). *Multitarget-Multisensor Tracking: Applications and Advances*, Artech House, ISBN 978-0890065174, Norwood

Bar Shalom, Y. & Fortmann, T. E. (1988). *Tracking and Data Association*, Academic Press, ISBN 978-0120797608, San Diego

Baud, O.; Gomord, P.; Honoré, N.; Ostorero, L.; Taupin, O. & Tubery, P. (2009). *Multi sensor data fusion architectures for Air Traffic Control applications*, In-tech publishing, ISBN 978-3902613523, Vienna

Baud, O.; Honoré, N. & Taupin, O. (2006). *Radar / ADS-B data fusion architecture for experimentation purpose*, ISIF'06, 9th International Conference on Information Fusion, pp. 1-6, July 2006.

Baud, O.; Honoré, N.; Rozé, Y. & Taupin, O. (2007). *Use of downlinked aircraft parameters in enhanced tracking architecture*, IEEE Aerospace Conference 2007, pp. 1-9, March 2007.

EUROCONTROL, *EUROCONTROL standard document for radar surveillance in en-route airspace and major terminal areas*, SUR.ET1.ST01.1000-STD-01-01 Ed. 1.0, March 1997.

Daskalakis, A. & Martone, P. (2005). *Assessing Wide Area Multilateration and ADS-B as alternative surveillance technology,* AIAA 5th ATIO and 16th Lighter-Than-Air Sys. Tech. and Balloon Systems Conference, Sep. 26-28, 2005, Arlington, Virginia.

Thompson, S. D. & Sinclair, K. A. (2008). *Automatic Dependent Surveillance – Broadcast in the Gulf of Mexico,* Massachusetts Institute of Technologies Publishing, Lincoln Laboratory, 2008.

Whitman, D. (2008). *Virtual Radar – Emulating Long Range ARSRs with ADS-B,* 2008 ICNS Conference, 2008.

# 20

# A Regime Recognition Algorithm for Helicopter Usage Monitoring

David He[1], Shenliang Wu[1] and Eric Bechhoefer[2]
[1]*The University of Illinois at Chicago*
[2] *Goodrich*
*USA*

## 1. Introduction

The importance of regime recognition for structural usage monitoring of helicopters cannot be overemphasized. Usage monitoring entails determining the actual usage of a component on the aircraft. This allows the actual usage/damage from a flight to be assigned to that component instead of the more conservative worst-case usage. By measuring the actual usage on the aircraft, the life of the components can be extended to their true lifetime. Usage monitoring requires an accurate recognition of regimes, where a regime is the flight profile of the aircraft at each instant of the flight. For each regime, a damage factor is assigned to each component that has usage. These damage factors are assigned by the original equipment manufactures (OEMs) based on measured stresses in the aircraft when undergoing a given maneuver. Therefore, it is important that the regimes can be recognized correctly during the flight of the aircraft to avoid either underestimated or overestimated damages for the aircraft. Another important aspect of regime recognition is related to the certification of health usage and monitoring system (HUMS). As outlined in a document of the Federal Aviation Administration (FAA) HUMS R&D Initiatives (Le, 2006), regime recognition and monitoring has been identified as a high priority HUMS R&D short-term task in the area of structural usage monitoring and credit validation. The certification readiness and the aircraft applicability of regime recognition and credit validation are lower in comparison to the overall HUMS assessment (12% to 18% and 59% to 82%, respectively). These cited assessment results clearly show the weakness of current regime recognition methods in HUMS.

Although important, not much work on regime recognition has been published. Two recent research papers are worth mentioning here. The first paper (Teal *et al*., 1997) described a methodology for mapping aircraft maneuver state into the MH-47E basic fatigue profile flight regimes in a manner which ensures a conservative, yet realistic, assessment of critical component life expenditure. They also presented the use of wind direction and magnitude estimation and inertial/air data blending to obtain high fidelity airspeed estimation at low speeds. An accuracy rate of 90% based on time was reported. This method basically is a logical test. The system firstly identifies the maneuver based on flight dynamic data and general principles of tandem rotor helicopter flight which are derived from flight experience and mathematical models correlated with flight test results, then the aircraft maneuver state is mapped directly into one of the basic fatigue profile flight regimes. The method is subject

to the main weakness of logical test in dealing with the noisy measurement. If the measured parameters were free of noise, logical tests would give accurate results. The second paper (Berry *et al*., 2006) presented a regime recognition scheme implemented as a hierarchical set of elliptical function (EBF) neural networks. Motivated to develop an automatic regime recognition capability as an enhancement to the US Army's Vibration Management Enhancement Program (VMEP), the EBF neural networks were devised to simplify neural net training and to improve the overall performance. The idea of using a hierarchical set of neural networks is to group individual regimes into regime groups, including an unknown regime group (regimes that cannot be classified as any regimes in one of the known regime groups). Regimes in each group are classified by an individual net in the hierarchical set. Regime recognition is carried out through a hierarchical process, e.g., if a regime cannot be classified as the first regime group by the top net in the hierarchy it will be passed to the lower level nets for further classification. In the paper, a total of 141 regimes of Sikorsky's S-92 helicopter were grouped into 11 groups, including "level flight", "auto", "climb", "dive", and etc. As shown in the paper, the EBF neural network regime recognition scheme gave near perfect classification results for "level flight" regimes. However, the results for classification of all regime groups didn't show a consistent effectiveness of the scheme. For example, for "level flight" group the classification rate is 97.85% but it is 33.18% for "turns" group. Because of the low classification rates for some groups, the scheme gave an overall rate of 76.21%. In addition to the requirement for a large amount of data to train a neural network, one variable that could also affect the performance of the scheme is the way by which the regimes are grouped. Another limitation of neural network is that as it is a black-box methodology, little analytical insights can be gained to enhance the regime recognition process.

Regime recognition is basically a data mining problem, i.e., mining the measured parameter data and mapping them to a defined flight profile. In this paper, the philosophy of data mining is adopted for regime recognition. In particular, a regime recognition algorithm developed based on hidden Markov model (HMM) is presented.

## 2. Regime recognition algorithm

Before presenting the data mining based approach for regime recognition, we first describe the regime recognition problem from a data mining perspective as following. Suppose we have $Q$ regimes, denoted as $\omega = \{\omega_1, \omega_2, ..., \omega_i, ..., \omega_Q\}$. By taking into account the time factor in regime recognition, each individual regime at time $t$ is expressed as $s(t) = \omega_i$. Given an observation sequence $R = R_1 R_2 ... R_t ... R_T$, where $T$ is the length of observations in the sequence and each observation $R_t$ is a $1 \times O$ vector, denoted as $R_t = \{f_{t1}, f_{t2}, ..., f_{tj}, ... f_{tO}\}^T$ with $f_{tj}$ being the value of feature $j$ of the $t$th observation and $O$ the number of the features, the objective is to identify regime sequence denoted as $\Omega = \{s(1), s(2), ..., s(T)\}$.

Accordingly, a hidden Markov model $\lambda = (\pi, A, B)$ for regime recognition could be characterized as follows:

1. The initial regime distribution $\pi = \{\pi_i\}$, where $\pi_i = P[s(0) = \omega_i]$, $1 \le i \le Q$.
2. The regime transition probability distribution $A = \{a_{ij}\}$, where $a_{ij} = P[s(t+1) = \omega_j \mid s(t) = \omega_i]$, $1 \le i, j \le Q$.
3. The observation probability distribution in regime $\omega_j$, $B = \{b_j(t)\}$, where $b_j(t) = P[R_t \mid s(t) = \omega_j]$, $1 \le j \le Q$, $1 \le t \le T$.

The estimation of $\lambda = (\pi, A, B)$ is a crucial step if we want to compute the probability of a system in regime $\omega_i$ based on the estimated HMM model $\hat{\lambda}$. Generally, there are two main stages in regime recognition using HMM. The first stage is the training stage. The purpose of training stage is to estimate the three parameters of the HMM. The estimation of $\lambda = (\pi, A, B)$ is carried out through an iterative learning process of adjusting the model parameters to maximize the probability $P(R_{train} \mid \lambda)$ of producing an observation sequence $R_{train} = R_1 R_2 ... R_{T_1}$, given model $\lambda$. Therefore, at the end of training process, we could obtain an estimated HMM model $\hat{\lambda}_i = (\hat{\pi}_i, \hat{A}_i, \hat{B}_i)$ for each regime $\omega_i$. The second stage is the testing stage. The purpose of testing is to calculate the probability of generating the unknown observation sequence, given the estimated model $\hat{\lambda}_i$, $1 \le i \le Q$. Given a testing observation sequence $R_{test} = R_1 R_2 ... R_{T_2}$ and a set of estimated models $\lambda = \left\{ \hat{\lambda}_i = (\hat{\pi}_i, \hat{A}_i, \hat{B}_i),\ 1 \le i \le Q \right\}$, log-likelihood $LL$ of $R_t$ from the observation sequence $R_{test} = R_1 R_2 ... R_{T_2}$ can be computed. Note that, in general, $T_1 \ge T_2$.

### 2.1 The training stage

The training stage is a process to determine model parameters from a set of training data. *A priori* values of π, *A*, and *B* are assumed and observations are presented iteratively to the model for estimation of parameters. Likelihood maximization is the basic concept behind this estimation procedure. In each iteration, the goal is to maximize the expected log-likelihood, i.e., logarithm of the probability that the model generates the observation sequence. This iterative process continues until the change in log-likelihood is less than some threshold and convergence is declared.

In an HMM, the observation probability is assumed to follow a Gaussian distribution. Although, all of the classical parametric densities are unimodal, many practical problems involve multimodal densities. In our algorithm, a Gaussian mixture model (GMM) is used. Let $X = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k, ... \mathbf{x}_T]$ be the sample dataset and $\mathbf{x}_k$ is a $O \times 1$ vector. Let *M* be the number of mixture components. There is no definition for the number of mixture components per output distribution and there is no requirement for the number of mixture components to be the same in each distribution. If $M = 1$, it is the unimodal density case. When $M > 1$, a mixture model can be expressed as:

$$p(X) = \sum_{i=1}^{M} w_i h(X \mid \mu_i, \sigma_i) \tag{1}$$

where $p(X)$ is the modeled probability distribution function and $w_i$ is the mixture weight of component *i*. Clearly, $w_1 + w_2 + ... + w_M = 1$, and $0 < w_i < 1$ for all $i = 1, 2, \cdots, M$. $h(X \mid \mu_i, \sigma_i)$ is a probability distribution parameterized by $\mu_i, \sigma_i$, and can be computed as:

$$h(X \mid \mu_i, \sigma_i) = \frac{e^{-\frac{1}{2}d_i^2}}{(2\pi)^{\frac{O}{2}}\sqrt{|\sigma_i|}} \tag{2}$$

where $d_i^2$ can be computed as:

$$d_i^2 = \left[d_{i1}^2, d_{i2}^2, ..., d_{ik}^2, ..., d_{iT}^2\right]^t$$

$$d_{ik}^2 = (\mathbf{x}_k - \mu_i)^t \sigma_i^{-1} (\mathbf{x}_k - \mu_i), \; k = 1,2,...,T \tag{3}$$

Once $w$, $\mu$ and $\sigma$ are determined, $p(X)$ is defined, i.e., the observation probability distribution $B$. So the estimation problem of an HMM model $\lambda = (\pi, A, B)$ is converted to estimate $\lambda = (\pi, A, \mathbf{\mu}, \mathbf{\sigma}, \mathbf{w})$, where $\mathbf{\mu} = \{\mu_1, \mu_2, ..., \mu_M\}$, $\mathbf{\sigma} = \{\sigma_1, \sigma_2, ..., \sigma_M\}$, $\mathbf{w} = \{w_1, w_2, ..., w_M\}$. Here, the GMM parameters of each HMM model can be split into two groups: the untied parameters that are Gaussian-specific and the tied parameters that are shared among all the Gaussians of all the HMM states.

Although there is no optimal way of estimating the model parameters so far, local optimal is feasible using an iterative procedure such as the Baum-Welch method (or equivalently the expectation-modification method) (Rabiner, 1989; Levinson *et al*., 1983), or using gradient techniques (Dempster, 1977). In order to facilitate the computation of learning, three forward-backward variables are defined in the forward-backward algorithm:

1. The probability of the partial observation sequence, $R_1 R_2 ... R_t$, and regime $\omega_i$ at time *t*, given model $\lambda$: $\alpha_t(i) = P[R_1 R_2 ... R_t, s(t) = \omega_i \mid \lambda]$
2. The probability of the partial observation sequence from *t*+1 to the end, given regime $\omega_i$ at time *t* and model $\lambda$: $\beta_t(i) = P[R_{t+1} R_{t+2} ... R_T \mid s(t) = \omega_i, \lambda]$
3. The probability of being in regime $\omega_i$ at time *t*, and regime $\omega_j$ at time *t*+1, given model $\lambda$ and observation sequence $R_t$, i.e.,

$$\xi_t(i, j) = P[s(t) = \omega_i, s(t+1) = \omega_j \mid R_t, \lambda]$$

We initialize forward variable as $\alpha_{t=0}(i) = 0$, for all $1 \le i \le Q$ at time $t = 0$, then in the forward iteration, we calculate the forward variable $\alpha_t(j)$ by the following equations from *t* = 1 to *t* = *T*:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{Q} \alpha_t(i) a_{ij}\right] b_j(R_{t+1}), \text{ where: } 1 \le t \le T-1, \; 1 \le j \le Q \tag{4}$$

In the backward iteration, we compute the backward variable $\beta_t(i)$ and $\xi_t(i, j)$ after initialization $\beta_T(i) = 1$ at time *t* = *T*:

$$\beta_t(i) = \sum_{j=1}^{Q} a_{ij} b_j(R_{t+1}) \beta_{t+1}(j), \text{ where: } t = T-1, T-2, ..., 1, \; 1 \le i \le Q \tag{5}$$

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(R_{t+1}) \beta_{t+1}(j)}{P(R_{train} \mid \lambda)} = \frac{\alpha_t(i) a_{ij} b_j(R_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{Q} \sum_{j=1}^{Q} \alpha_t(i) a_{ij} b_j(R_{t+1}) \beta_{t+1}(j)}, \text{ where: } 1 \le i, j \le Q \tag{6}$$

We define $\gamma_t(i)$ as the probability of being in regime $\omega_i$ at time *t*, given the observation sequence and the model. Therefore, $\gamma_t(i) = \sum_{j=1}^{Q} \xi_t(i,j)$. So, $\hat{a}_{ij}$, the estimated probability of a regime transition from $s(t-1) = \omega_i$ to $s(t) = \omega_j$, can be calculated by taking the ratio between the expected number of transitions from $\omega_i$ to $\omega_j$ and the total expected number of any transitions from $\omega_i$ can be computed as:

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)},\ 1 \le i \le Q,\quad 1 \le j \le Q \tag{7}$$

Then the estimated observation probability distribution $(\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j, \mathbf{w}_j)$ in regime $\omega_j$ can be computed as:

$$\hat{\boldsymbol{\mu}}_j = \frac{\sum_t \mathbf{w}_t^j E_{tj}[R_{train}]}{\sum_t \mathbf{w}_t^j} \tag{8}$$

$$\hat{\boldsymbol{\sigma}}_j = \frac{\sum_t \mathbf{w}_t^j E_{tj}[R_{train} R_{train}^{'}]}{\sum_t \mathbf{w}_t^j} - \hat{\boldsymbol{\mu}}_j \hat{\boldsymbol{\mu}}_j^{'} \tag{9}$$

Note in equations (8)- (9), weights $\mathbf{w}_t^j = \sum_{t=1}^{T-1} \gamma_t(j)$ are posterior probabilities, $E_{tj}[R_{train}]$ is the mathematical expectation of observation $R_{train}$ at time *t* under regime $\omega_j$, $\hat{\boldsymbol{\mu}}_j, \hat{\boldsymbol{\sigma}}_j$ are both untied parameters of GMM, and the covariance type is full.
Similarly, the estimated initial regime distribution can be computed as:

$$\hat{\pi}_i = \gamma_1(i),\ 1 \le i \le Q \tag{10}$$

### 2.2 The testing stage

As mentioned before, testing is a stage to evaluate the likelihood of an unknown observation belonging to a given regime. Since model $\hat{\lambda}_i = (\pi_i, A_i, \boldsymbol{\mu}_i, \boldsymbol{\sigma}_i, \mathbf{w}_i)$ has been built up in the training process, it can be used to calculate the log-likelihood *LL* of a testing observation $R_{test}$ based on model $\hat{\lambda}_i$ for regime $\omega_i$. This log-likelihood can be calculated efficiently using the forward algorithm.
The probability that model $\hat{\lambda}_i$ produces observation $R_{test}$ is computed as:

$$P(R_{test} \mid \hat{\lambda}_i) = \alpha(i)$$

By definition, $\alpha(i)$ is the probability of generating $R_{test}$ and ending in regime $\omega_i$, therefore,

$$P(R_{test} \mid \hat{\lambda}_i) = \alpha(i) = P[s(t) = i] P[R_{test} \mid s(t) = i] = \pi_i p_i(R_{test}) \quad 1 \le i \le Q \tag{11}$$

In (11), $p_i(R_{test})$ can be calculated from equation (1), and $R_{test}$ in regime recognition is the unknown signal. So, the log-likelihood value is computed as:

$$LL_i = \mathrm{Ln}[P(R_{test} \mid \hat{\lambda}_i)] \tag{12}$$

To classify a testing observation into one of $Q$ regimes, train $Q$ HMMs, one per regime, and then compute the log-likelihood that each model gives to the testing observation, a set of log-likelihood value $LL = \{LL_1, LL_2, ..., LL_i, ..., LL_Q\}$ will be obtained.

## 3. Algorithm validation

In this section, the developed regime recognition algorithm was validated using the Army UH-60L flight card data.

The Army UH-60L flight card data was collected during a flight test and provided by Goodrich. The intent of the flight test was to provide flight data which could be used to refine and revise a preliminary set of regime recognition algorithms. The test pilots annotated detailed flight cards with actual event times as maneuvers were conducted during the UH-60L regime recognition flights. The on-board pilots maintained a detailed log of the maneuvers, flight conditions, and corresponding event times encountered during the mission flight. A total of 50 regimes were conducted with annotation in the flight test. A limited amount of non-annotated actual flight data was used prior to the flight test to check the functionality of the HUMS system. The recorded data was downloaded and processed after the flight test.

For the Army UH-60L helicopter, a total of 90 preliminary regimes were defined by original equipment manufacturer (OEM). Data of 22 basic aircraft parameters were collected from sensors mounted on the aircraft, or sensors added to the Goodrich IMD-HUMS system for regime recognition. These parameter data is used for the identification of events, control reversals, and regimes. The parameter monitoring is performed during the whole ground-air-ground (GAG) cycles, from rotor start to rotor shutdown, and takeoff to landing. Table 1 provides the list of parameters with their description collected from IMD-HUMS system for regime recognition.

During the validation process, the dataset was randomly divided into two subsets: 70% of data was used for training and 30% for testing. By using the training data, an HMM model was built for each regime. Then the testing data was input into the trained HMM models to compute the log-likelihood values. The maximum log-likelihood value indicates the identified regime. The confusion matrix generated during the testing is provided in Table 2. From the results in Table 2, we see that the overall accuracy of the regime recognition is 99%.

Note that in solving regime recognition using HMM, the training set is dependent on the time sequence of maneuvers. Thus, it is able to find the regime or very complex grouped maneuvers. On the down side, the training set is really too small to capture all of the various maneuvers sequences that could be encountered. For example, it likely that from straight and level, you could go in to a left turn, or a right turn. From a right turn, you can go back to level, climbing right turn, diving right turn, or a higher angle of back turn. In reality, a flight card should contain all of the mixed mode maneuvers.

| Parameter No. | Parameter Name | Parameter Description |
|---|---|---|
| 1 | WowDly | WOW Delayed |
| 2 | LngFlg | Landing Flag |
| 3 | TkOFlg | Takeoff Flag |
| 4 | RollAt | Roll Attitude |
| 5 | PtchAt | Pitch Attitude |
| 6 | RdAlt | Radar Altitude |
| 7 | YawDt | Yaw Rate |
| 8 | AltDt | Altitude Rate |
| 9 | LatDt2 | Lateral Acceleration |
| 10 | VertAccl | Vertical Acceleration |
| 11 | MrRpm | RPM of Main Rotor |
| 12 | CrNz | Corrected Normal Acceleration |
| 13 | CalSpd | Calibrated Airspeed |
| 14 | Vh | Airspeed Vh Fraction |
| 15 | TGT | Turbine Gas Temperature |
| 16 | RMS_Nz | RMS Normal Acceleration |
| 17 | TEngTrq | Torque 1/Torque 2 |
| 18 | AOB | Angle of Bank |
| 19 | CR_Pedal | Control Reversal Flag |
| 20 | Cr_Colct | Corrected Collective Rate |
| 21 | Cr_Lat | Corrected Latitude |
| 22 | Cr_Lon | Corrected Longitude |

Table 1. Monitored parameters in IMD-HUMS system

| Regime | 2 | 3 | 4 | 5 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 0.99 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 5 | 0.05 | 0.66 | 0.26 | 0.97 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 1 | 0.99 | 0.5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 | 0.95 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 | 0.97 | 1 | 0.98 | 1 | 0.93 | 0.99 | 1 | 1 |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.93 | 0.91 | 0.98 | 1 | 1 |
| 12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.99 | 0.87 | 0.96 | 1 | 1 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.96 | 1 | 1 |
| 15 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.98 | 1 |
| 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.93 | 0.9 |
| 17 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.98 | 1 |
| 19 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 20 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 21 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 22 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 23 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 24 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 25 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 26 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 27 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 28 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 36 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 37 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 40 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 41 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 42 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 43 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 44 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 45 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 46 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 48 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 49 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 2. The confusion matrix of validation test

| Regime | 2 | 3 | 4 | 5 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 51 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 52 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 53 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 54 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 55 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 56 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 57 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 59 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 60 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 61 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 63 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 64 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 65 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 2. The confusion matrix of validation test (Continued 1)

| Regime | 41 | 42 | 43 | 44 | 45 | 46 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 15 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 17 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 19 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 20 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 21 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 22 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 23 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 2. The confusion matrix of validation test (Continued 2)

| Regime | 41 | 42 | 43 | 44 | 45 | 46 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 24 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 25 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 26 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 27 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 28 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 36 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 37 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 40 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 41 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 42 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 43 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 44 | 1 | 1 | 1 | 0.98 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 45 | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 46 | 1 | 1 | 1 | 1 | 1 | 1.00 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 48 | 1 | 1 | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 49 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 50 | 1 | 1 | 1 | 1 | 1 | 1 | 0.97 | 1 | 0.97 | 1 | 1 | 1 | 1 | 1 |
| 51 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 |
| 52 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.93 | 1 | 1 | 1 |
| 53 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.96 | 1 | 1 |
| 54 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.97 | 1 |
| 55 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.98 |
| 56 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 57 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 59 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 60 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 61 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 63 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 64 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 65 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 2. The confusion matrix of validation test (Continued 3)

| Regime | 56 | 57 | 59 | 60 | 61 | 63 | 64 | 65 | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 15 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 17 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 19 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 20 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 21 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 22 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 23 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 24 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 25 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 26 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 27 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 28 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 36 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 37 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 40 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 41 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 42 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 43 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 44 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 45 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 46 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 48 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 49 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 50 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |

Table 2. The confusion matrix of validation test (Continued 4)

| Regime | 56 | 57 | 59 | 60 | 61 | 63 | 64 | 65 | |
|---|---|---|---|---|---|---|---|---|---|
| 51 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 52 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 53 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 54 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 55 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 56 | 0.97 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| 57 | 1 | 0.98 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 59 | 1 | 1 | 1.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| 60 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 | 1 |
| 61 | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 | 1 |
| 63 | 1 | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 |
| 64 | 1 | 1 | 1 | 1 | 1 | 1 | 0.98 | 1 | 1 |
| 65 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.98 | 1 |
| **Overall Accuracy** | | | | | | | | | **0.99** |

Table 2. The confusion matrix of validation test (Continued 5)

In addition to the validation test, the performance of the HMM based regime recognition algorithm was compared with a number of data mining methods. These data mining methods included: neural network, discriminant analysis, K-nearest neighbor, regression tree, and naïve bayes. The results of the performance comparison test are provided in Table 3. In this test, to be consistent, data with the same regimes were used for all the data mining methods. From Table 3, we can see that the HMM based regime recognition algorithm outperforms all other data mining methods.

Note that in Table 3, the names of methods are defined as: HMM = hidden Markov model; NN = neural network (back propagation); DA = discriminant analysis; KNN = k-nearest neighbor; RT = regression tree; NB = naïve bayes.

## 4. Conclusion

In this paper, a data mining approach is adopted for regime recognition. In particular, a regime recognition algorithm developed based on HMM was presented. The HMM based regime recognition involves two major stages: model learning process and model testing process. The learning process could be implemented off-board. In this process, Gaussian mixture model (GMM) was used instead of unimodal density of Gaussian distribution in HMM. Once the learning process is completed, new incoming unknown signal could be tested and recognized on-board. The developed algorithm was validated using the flight card data of an Army UH-60L helicopter. The performance of this regime recognition algorithm was also compared with other data mining approaches using the same dataset. Using the flight card information and regime definitions, a dataset of about 56,000 data points labeled with 50 regimes recorded in the flight card were mapped to the health and usage monitoring parameters. The validation and performance comparison results have showed that the hidden Markov model based regime recognition algorithm was able to accurately recognize the regimes recorded in the flight card data and outperformed other data mining methods.

| Regime No. | HMM | NN | DA | KNN | RT | NB |
|---|---|---|---|---|---|---|
| 2 | 0.01% | 32.00% | 16.00% | 0.00% | 12.00% | 46.00% |
| 3 | 0.10% | 17.65% | 23.53% | 41.18% | 100.00% | 94.12% |
| 4 | 0.06% | 0.00% | 84.62% | 69.23% | 100.00% | 84.62% |
| 5 | 10.30% | 3.03% | 0.00% | 0.00% | 39.39% | 75.76% |
| 7 | 2.51% | 10.20% | 5.10% | 10.20% | 3.06% | 0.00% |
| 8 | 0.26% | 10.00% | 6.67% | 43.33% | 0.00% | 53.33% |
| 9 | 0.04% | 18.75% | 62.50% | 56.25% | 100.00% | 100.00% |
| 10 | 0.66% | 50.00% | 50.00% | 35.71% | 100.00% | 100.00% |
| 11 | 1.02% | 0.00% | 26.67% | 20.00% | 13.33% | 66.67% |
| 12 | 0.92% | 0.00% | 0.00% | 0.00% | 100.00% | 88.89% |
| 13 | 0.05% | 0.00% | 0.00% | 3.45% | 100.00% | 68.97% |
| 14 | 0.21% | 13.33% | 6.67% | 46.67% | 100.00% | 80.00% |
| 15 | 0.09% | 33.33% | 33.33% | 66.67% | 100.00% | 50.00% |
| 16 | 0.88% | 40.00% | 100.00% | 100.00% | 100.00% | 80.00% |
| 17 | 0.12% | 100.00% | 100.00% | 100.00% | 100.00% | 100.00% |
| 19 | 0.11% | 10.00% | 0.00% | 50.00% | 100.00% | 80.00% |
| 20 | 0.09% | 0.00% | 0.00% | 0.00% | 0.00% | 70.83% |
| 21 | 0.09% | 9.09% | 36.36% | 36.36% | 100.00% | 100.00% |
| 22 | 0.01% | 21.43% | 14.29% | 50.00% | 100.00% | 100.00% |
| 23 | 0.07% | 0.00% | 9.09% | 9.09% | 100.00% | 90.91% |
| **Overall** | 0.88% | 12.79% | 15.58% | 21.16% | 47.44% | 57.44% |

Table 3. Results of performance comparison of various data mining methods (regime recognition error rate)

## 5. References

Berry, J., Vaughan, R., Keller, J., Jacobs, J., Grabill, P., and Johnson, T. (2006). Automatic Regime Recognition using Neural Networks. *Proceedings of American Helicopter Society 60th Annual Forum,* Phoenix, AZ, May 9 – 11, 2006.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, Vol. 39, No. 1, pp. 1 - 38.

Le, D. (2006). Federal Aviation Administration (FAA) Health and Usage Monitoring System R&D Initiatives. http://aar400.tc.faa.gov/Programs/agingaircraft/rotorcraft/index.htm.

Levinson, S. E., Rabiner, L. R., and Sondhi, M. M. (1983). An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition. *Bell System Technical Journal*, Vol. 62, No. 4, pp. 1035 - 1074.

Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, Vol. 77, No. 2, pp. 257 - 286.

Teal, R. S., Evernham, J. T., Larchuk, T. J., Miller, D. G., Marquith, D. E., White, F., and Deibler, D. T. (1997). Regime Recognition for MH-47E Structure Usage Monitoring. *Proceedings of American Helicopter Society 53rd Annual Forum,* Virginia Beach, VA, April 29 – May 1, 1997.

# Part VI

# 21

# A New Method of High Temporal Resolution Remote Sensing Imaging: Moving Target Detection with Bionics Compound Eye

Lei Yan, Pengqi Gao, Huabo Sun and Hongying Zhao
*Peking University*
*China*

## 1. Introduction

Moving target detection has always been a significant interest in temporal image analysis field since it rose up to the present. It has widespread application prospect in dealing the emergency event, detecting moving object in big square, and rescuing in natural disaster. Based on the common single lens camera imaging, for the restraint of the photographable screen and the expression of plane image, we have to change the actual spatial information from three-dimension to two-dimension, and meanwhile we need the excellent handling ability to protect image high frequency detail. When we detect moving target on computer, we have to transform the two-dimensional information to the three dimensional attribute to reappear the actual target. Such "3-2-3" dimensional conversion has large computation, low real time and bad precision, so it hinders the development of moving target detection. Insect 's compound eye system has the unique superiority in detecting moving target. It can detect moving goal with little information by distinguishing outline characteristic of object. And it does not have the visual dimension transformation question.

The traditional visible light and infrared target detection systems mostly use the existing imaging sensors (for example camera and so on) to gain the image, and then use computer signal processing method to achieve recognition and localization for target. The precision of system has a positive correlation with the computation payload. The precision is higher and the need of the computation load is larger. However, the existing processor cannot completely satisfy the demand of the system for the time request. Therefore, many researchers want to develop new methods for high temporal resolution remote sensing imaging while they continue to research the original processing mode. Moving target detection, with less computation, is an important development direction in future. In this aspect, some insects, just like dragonfly and honeybee with compound eye, have provided good example for us.

Because insect's compound eye has the characteristics of small volume, light weight and large viewing field, it has the unique superiority in the moving object detection and track. According to the research, the honeybee's reaction time to the object which appears suddenly only needs 0.01s, but the human eye actually needs 0.05s. Dynamic target detection ability of dragonfly reaches 100%, and capture ability reaches 99%. It is definite to establish one kind of new movement object detection system by simulating this kind of

characteristic of insect. In this aspect, many foreign researchers have done a lot of work, and have studied many outstanding algorithms and system. American Martin Marietta Company and the Texas instrument company developed the sensor automatic target recognition module, also developed the miniaturized sensor module which is used for performance flight. American ID Image Company, Missouri/ Columbia University have united to study the staring algorithm (GATOR) system which has been used to face the track recognition. Britain and Norway have united to develop the image formation infrared sensor and the programmable phantom processor. In 1991, B.Moslehi, W.Parkyn et al. of the American Ghali Furry Asia Physical optics Company invented one kind bionic compound eye vision system.

Some research has already been done in this domain in China. Li Wenyuan of Tianjin University used the computer to simulate the image effect that real image pass through the compound eye optical system. Lu Liming and Wang Guofeng et al. of the Northwestern Industrial University used two modulators to simulate the function of the compound eye localizing the target. They integrate the image formation detector to construct more than one mold to guide the missile homing head. This research has proven the feasibility and value application of the movement object detection using bionic compound eye. Tian Weijian and Yao Shengli et al. of Chinese Academy of Science Xi'an Optics Precision machinery Research institute proposed one kind multichannel imaging system which has been used in the moving target detection. Huang Zhuxin, Xu Guili et al. of Nanjing University of Aeronautics and Astronautics have designed one kind bionic compound eye measurement system which is composed of 7 cameras. This system has big field of view, but its volume is also huge. The combination mode has not jumped out of the ordinary single aperture camera image thought completely.

## 2. Research on moving target detection based on compound eye

The traditional video moving target detection technology mainly refers to detect moving goal with static camera photographs. When the camera is loaded on moving carrier, there will be relative movement between camera and moving goal. So the electronic image stabilization is need for accurate detection moving goal.

Target monitoring scope is limited with single aperture camera. Image is much smaller after the cutting out visible scope with electronic image stabilization, and it is not favor for the wide range detection. Compound eye has unique superiority in moving target detection. Based on compound eye image mechanism and method, a bionic system that mimics large field of view and spherical imaging characteristics of compound eye at structure level is proposed. In view of high temporal resolution remote sensing, our research include: moving target outline is rapidly set with changeable resolution detection mode; the geometrical structure between cameras is given out; method for panoramic mosaic and calculation of overlap ratio of adjacent cameras is given out, and the synchronization of different cameras is discussed.

The authors intend to research new method for moving target detection with imitating insect vision system. We do lots of research on bionic compound eye mechanism and sum up the basic principle of moving target detection based on insect compound eye. The principle is shown in Fig. 1.

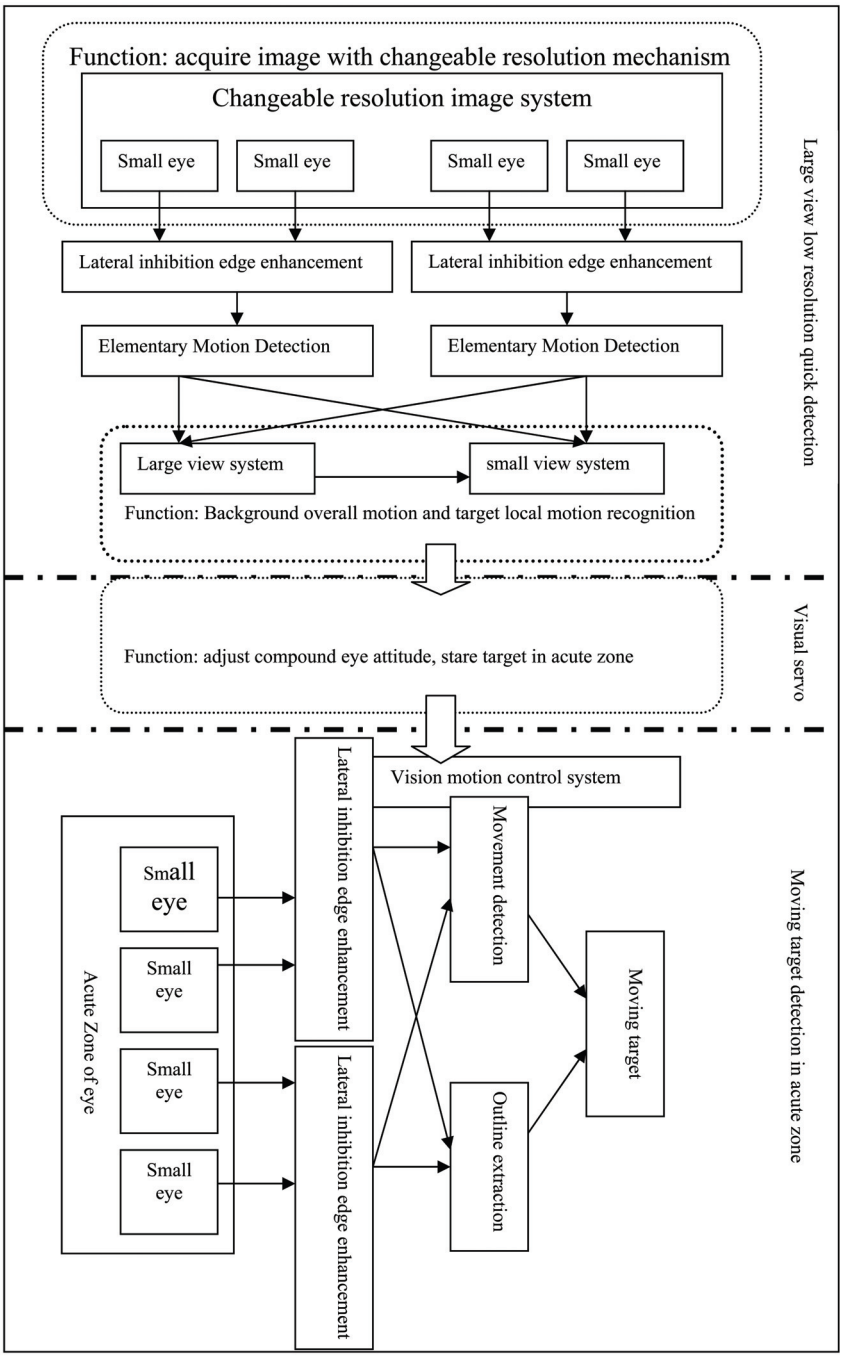



Fig. 1. Principle of bionic compound eye moving target detection

Based on anatomy of insect compound eye and electrophysiological data, the similar experimental apparatus is researched on. By comparing insect compound eye with experimental data acquired from the apparatus, the mechanism of moving target detection is revealed in theory to establish the corresponding mathematical model. In final, the new algorithm of moving target detection is modeled to improve temporal resolution of remote sensing imaging. The framework of research is shown in Fig. 2.

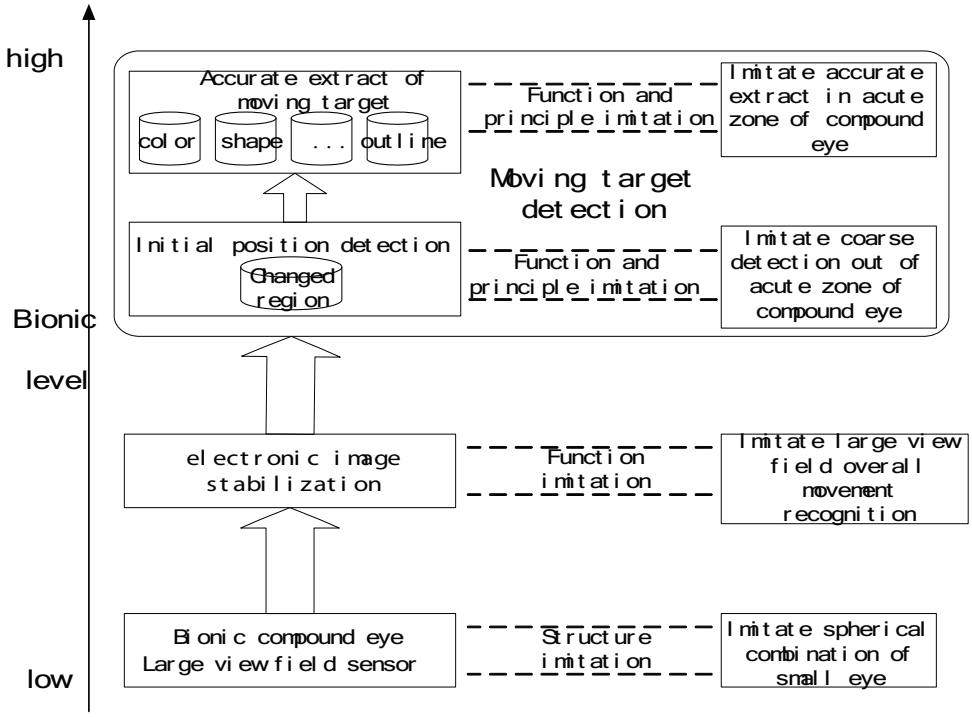



Fig. 2. Framework of bionic compound eye moving target detection

Moving target detection by compound eye is imitated at the two levels of function and principle. A bionic moving target detection algorithm is proposed, which consists of two main stages. In the first stage, principle of coarse detection of moving targets by low resolution area that out of acute zone of compound eye is mimicked. A motion window construction method is proposed by merging minimal bounding rectangle of adjacent changing area to roughly point out positions of targets. In the second stage, principle of accurate target extraction in acute zone of compound eye is mimicked. A target segmentation method is proposed, which in the motion window uses the complementarities of edge detection and symmetrical difference to get a coherent outline at first and then calculates convex hull of the outline to extract accurate target. The algorithm decreased the probability of occurrence of more than one target in a single motion wind. It also makes the targets detected more integrated. As a result of coarse-to-fine detection, high temporal efficiency is achieved.

## 3. Bionic compound eye hardware equipment

We research bionic compound eye equipment in reference of dragonfly eye structure. The equipment comprises multi-channel ray system and each ray system is similar to an aperture of compound eye. The ray peak is small lens which correspond an independent photosensitive part. The photosensitive part equipped with cpu process the data and connect with the next advanced treatment equipment. Equipment processing mode simulate two levels of processing pattern of compound eye.After the aperture data is sampled and extracted by the first treatment unit, it is submitted to the second treatment unit. The flow is shown in Fig. 3.

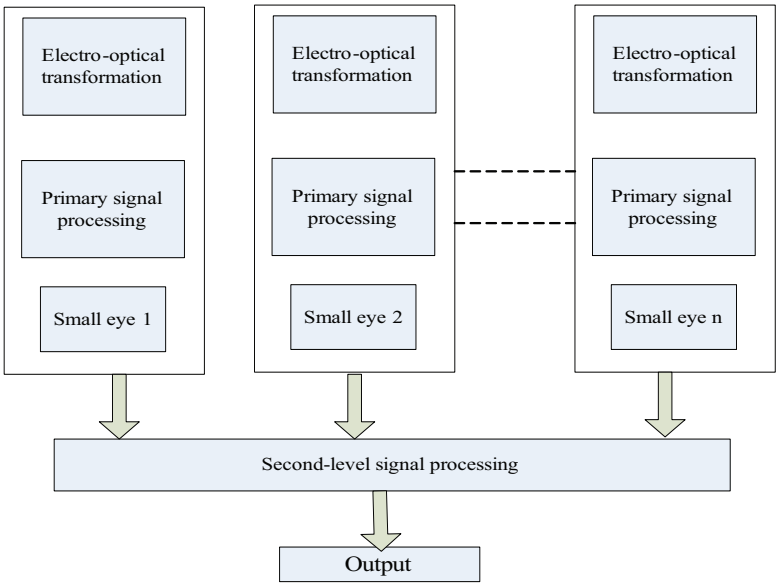



Fig. 3. Compound eye two levels of signal processing schematic drawing

We imitate dragonfly eye structure to set up bionic compound eye equipment. The equipment is used to test moving target detection algorithm, and its schematic drawing is shown in Fig. 4.

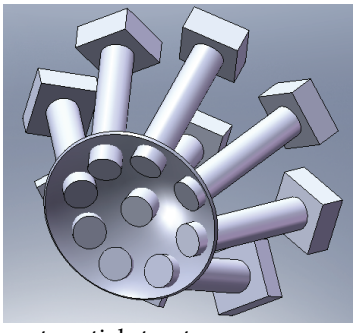


Fig. 4. Compound eye equipment spatial structure

## 4. Electronic image stabilization of bionic compound eye equipment

The rocking of image will reduce detection precision when aiming the moving target. So the electronic image stabilization before detection is crucial for the following automatic detection. The main research of electronic image stabilization includes three sections: identify overall motion parameter of moving target; forecast active movement value; compensation and splicing of image sequence. The electronic image stabilization goal is to eliminate rocking noisy and make image clear and stable. The key step of electronic image stabilization is calculating movement value between images and its flow is shown in Fig. 5.

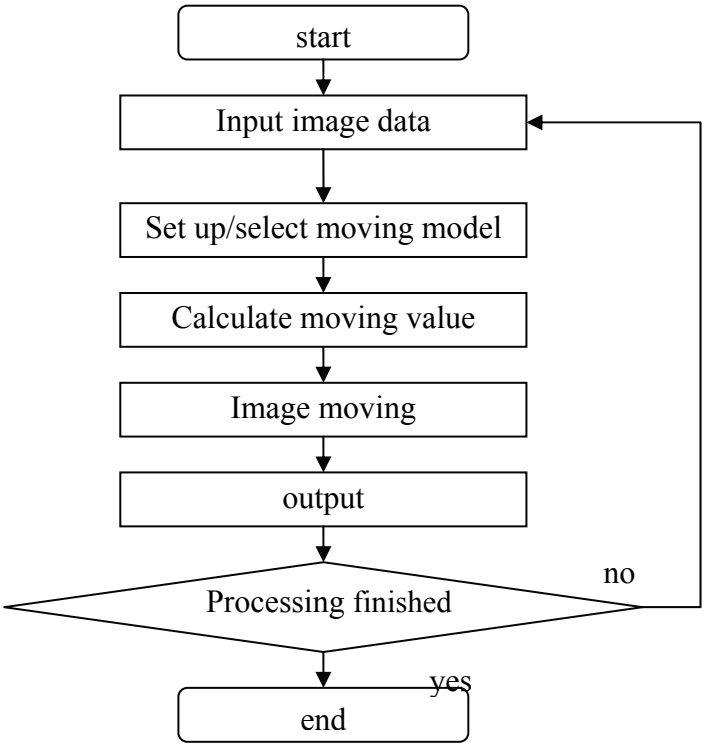



Fig. 5. Flow of electronic image stabilization

Behind electronic image stabilization, image mosaic of adjacent cameras is done. Its destination is to expand the view field of compound eye system. Image mosaic software researched by us can splice several images to a complete image, and the result is shown in Fig. 6.

## 5. Simulate compound eye visual function to construct moving target detection model

Small and medium-sized compound eyes of insects are not evenly distributed, and sensitive vision areas lie in the regions with dense compound eye. The authors research sensitive vision areas to find its effect for detecting moving target and to give qualitative analysis.

Fig. 6. Automatic image mosaic by three high resolution cameras

With restriction of the number of small eyes, spatial resolution is not too high on compound eye imaging.But this is advantagement that compound eye structure is favorable to detect moving target. In most cases, moving target detection only needs information of targets outline. Compound eye could quickly extract objects profile with little information.In this case, high resolution will not only reduce the efficiency of contour extraction, but also limit object in a small area and can not recognize the overall targets, resulting in the error detection. Therefore, it is an important part that quickly extracts moving object contour mechanism with low resolution imaging based on insect compound eye.

Compound eyes structure and bionics research have shown that there is a disparity that adjacent compound eyes test on the same point. According to the relative position of small eyes, they detect the same point time difference parameters which can be quickly calculated the value of changes in external optical flow and the moving targets location and speed. When compound eye movement with background fixed, compound eyes could also be detected insect own speed and relative position. Therefore the second part of this section is exploring insect to utilize compound eye which could detect changes in the external optical flow and locating moving target position, and to find the main impact factors on moving targets detection, such as velocity, distance and small eye visual muscle movement.

In the 1950s, Hassenstein and Reichardt proposed compound eye elementary motion detection model (HR model), shown in Fig. 7. The model describes process of inter-related operations on two adjacent photoreceptor cells (or small eyes),which the HPF represent high-pass filter, LPF on behalf of low-pass filter, $\pi$ stand for multiplication, $\Sigma$ represent addition and subtraction. The model has better explanation on compound eye which receives the primary optical signal, but the model is not still perfect now. In addition, HR model aim at part process. In the basis of these studies, on the one hand we will combine with the existing experimental data on the HR model to improve it further; on the other hand we will focus on simulation of the overall signal processing neurons, and analysis function on moving target recognition in sensitive areas. We also want to establish compound eye imaging mathematical model and give calculative method of parameters values in model.

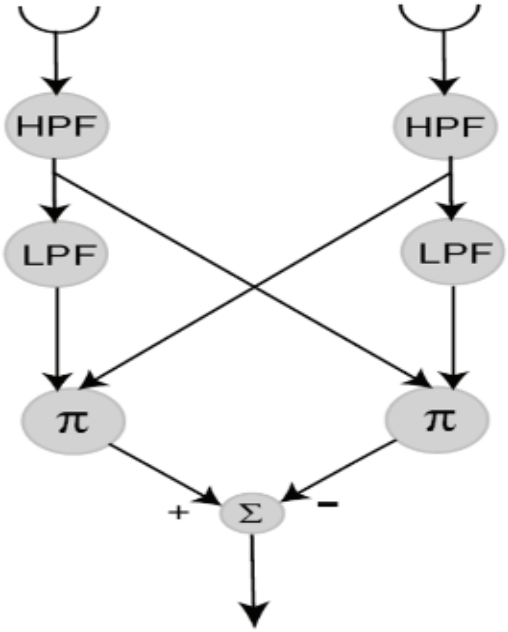



Fig. 7. Hassenstein-Reichardt elementary motion detection model

Based on insect visual perceptive mechanism, we establish the corresponding mathematical model by simulating bionic compound eye experiments with experimental device. For common single aperture image, large amount of useful experimental data (such as three-dimensional information, speed information etc.) are almost completely lost in the imaging process with multi-dimensional conversion. During the phase of recognizing moving target, two-dimensional image need to be reverted multi-dimensional information. The process needs large volume of calculation which is difficult to ensure real-time. In order to make up lost information, we introduce additional knowledge which will create additional noise. An important part of research is to find how to image directly which utilize the target multi-dimensional information at the same time. In addition, subject to structural constraints, resolution of compound eye imaging is far lower than the human eye and the existing high-resolution imaging. However, low-resolution imaging is precisely conformed to its limited brain processing capacity of insect which could detect targets with small computation.

## 6. Design compound eye moving target detection algorithm

To design and simulate moving targets detection algorithm of compound eye imaging, we establish models with computer simulation software. By changing different settings and different parameters of the handler functions, we can get different simulation results. Compared simulation results with the actual measurement data of compound eyes, we improve algorithm further. We do some experiment to test the algorithm based on the completion of installation. We select two types of indoor and outdoor scenes to verify effectiveness of algorithm. The algorithm, combined indoor with outdoor scenes, will form a complete system for moving target detection. The flow of algorithm is shown in Fig. 8.

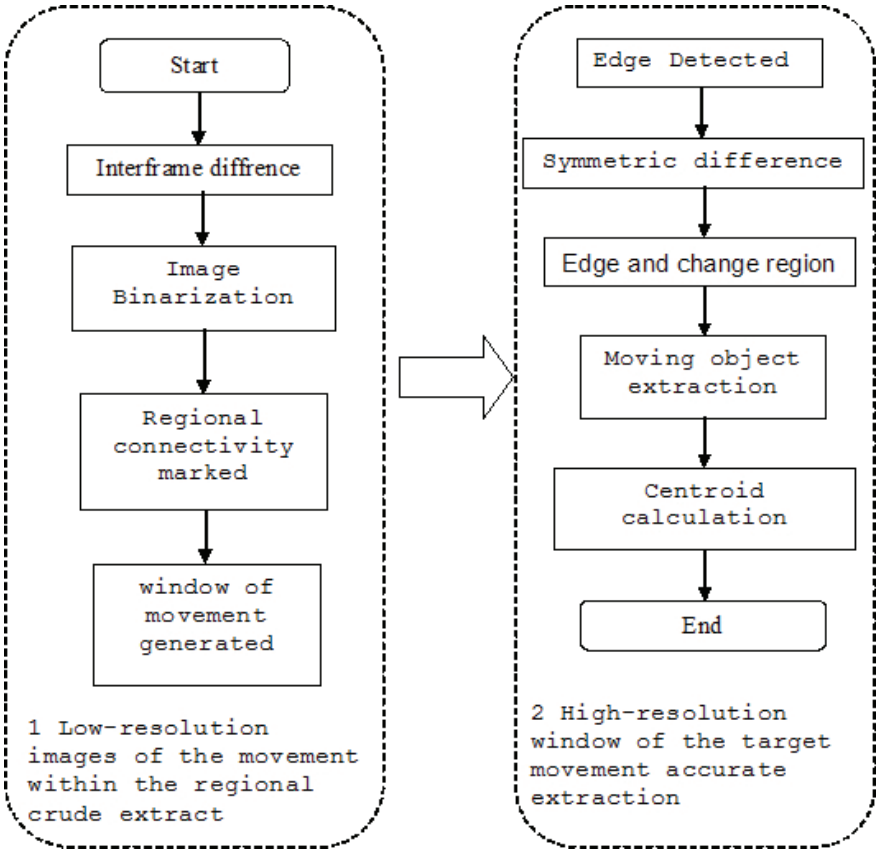



Fig. 8. Compound eye moving target detection algorithm flow

During the first phase of the algorithm, we simulate insect compound eye to detect moving target outline and extract targets in a rough area.In the second phase of algorithm, we simulate compound eye mechanism to accurately extract moving object.
Based on the above research, we developed electronic image stabilization and moving target detection software, shown in Fig. 9.

## 7. Conclusion

Generally speaking, the former research mostly focused in imitation of equipment structure, rather than in compound eye mechanism, which cause the research to be relatively slow. This project team has done some foundation work in insect morphology, pattern recognition and digital imaging domain. Based on former foundation, this project continues to do some work, such as simulating dragonfly compound eye imaging system, establishing elementary mathematical model developing compound eye equipment and introducing the electronic image stabilization. The work will obviously enhance system performance. With the breakthrough in this project, it will lead to any transformation in methods and means for the intelligence vision system.

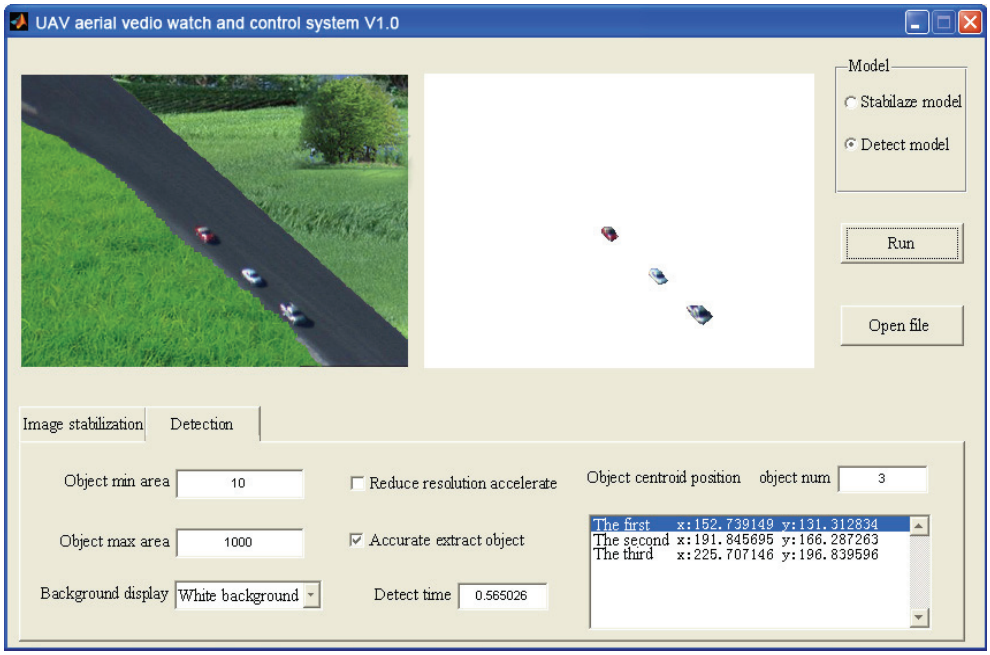



Fig. 9. Software interface of aerial image stabilization and moving target detection

## 8. Acknowledgment

1. Project supported by the National Key Technologies R & D Program of China
**Project Name:** Development of aerial remote sensing system of multipurpose Unmanned Aerial Vehicle
Project No.: 2004BA104C
2. Project supported by: National Natural Science Foundation
**Project Name:** Study on video image stabilization mechanism of UAV aerial remote sensing
Project No.: 60602042

## 9. Reference

[1] Pengqi Gao, Lei Yan, Hongying Zhao. A new method for stabilization of video images with large moving object. EIMTV, 2007.
[2] Pengqi Gao, Lei Yan, Hongying Zhao. Aerial remote sensing automatic control system for UAV. ASSAIII, 2007.
[3] Pengqi Gao, Lei Yan, Hongying Zhao, Shuqiang Lu. Study on shooting control algorithm of remote sensing control system for UAV. IGARSS2007.
[4] Rakesh Kumar, Harpreet Sawhney, Supun Samrasekera, Steve Hsu, Hai Tao, Yanlin Guo, Keith Hanna, Arthur Pope, Richard Wildes, David Hirvonen, Michael Hansen, Peter Burt. Aerial video surveillance and exploitation[C]. Proceeding of IEEE, October 2001, 89(10): 1518-1539

[5] Jun Wu and Guoqing Zhou. Real-time UAV video processing for quick-response to natural disaster[C]. Proceedings of IEEE International Conference on Geoscience and Remote Sensing Symposium, July 31-Aug. 4, 2006, Denver, USA. 2006: 976-979.
[6] Suman, Srinivasan, Haniph Latchman, John Shea. Airborne traffic surveillance systems – video surveillance of highway traffic[C]. Proceedings of the ACM 2nd International Workshop on Video Surveillance & Sensor Networks, October 2004, New York, USA. 2004: 131-135.
[7] A. Shastry, R. Schowengerdt. Airborne video registration for visualization and parameter estimation of traffic flows[C]. Proceedings of Pecora 15/Land Satellite Information IV/ISPRS Commission I/FIEOS Conference, 2002.
[8] Daiyong Wei and Guoqing Zhou. Real-time UAV ortho video generation[C]. Proceedings of IEEE International Conference on Geoscience and Remote Sensing Symposium, Boston, USA, July 2008, 5: 510-513.
[9] A.C. Shastry, R.A. Schowengerdt. Airborne video registration and traffic-flow parameter estimation[J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(4): 391- 405.
[10] http://www.cs.cmu.edu/~vsam/
[11] Hongying Zhao, Pengqi Gao, Lei Yan, Shuqiang Lu. The real time image merge method for the remote sensing image acquired from the UAV. IGARSS2005.
[12] Hongying Zhao, Zhiyang Gou, Pengqi Gao, Yinqian Cheng. No ground control point making the orthophoto for the UAV remote sensing system. ISPDI, 2007
[13] Lei Yan, Pengqi Gao, Hongying Zhao, Kai He, Jingjing Liu, Shaowen Yang, Qingxi Tong. Data collection of UAV remote sensing system in cloudy and rainy area. APCATS'2006.
[14] Lei Yan, Jie Ding, Qiming Qin, Yuefeng Liu, Hongying Zhao, Daping Liu, Shihu Zhao, Pengqi Gao, Zhouhui Lian. A ground-based teaching and experimental simulation system for earth observing digital aerial remote sensing. ISITAE'07, Nov. 2007: 388-390.
[15] Arambel Pablo, Silver Jeff, Krant Jon, Antone Matthew, Strat Thomas. Multiple-hypothesis tracking of multiple ground targets from aerial video with dynamic sensor control[C]. Signal processing, sensor fusion, and target recognition XIII, Orlando FL, 12-14 April 2004. 2004, 5429: 23-32.
[16] http://www.sarnoff.com/sarnoff.com/press-room/news/2007/03/12/pyramid-vision-receives-videodetective-order-from-dhs
[17] http://www.pyramidvision.com/products/terrasight
[18] Yu-chia Chung, Zhihai He. Low-complexity and reliable moving objects detection and tracking for aerial video surveillance with small UAVs[C]. Proceeding of IEEE International Symposium on Circuits and Systems, 2007.5, Volume 27, pp2670–2673.
[19] Jongrae Kim, Yoonsoo Kim. Moving ground target tracking in dense obstacle areas using UAVs[C]. Proceeding of the 17th IFAC World Congress, Seoul, Korea, 2008.7.
[20] Heiko Helble, Stephen Cameron. OATS: Oxford aerial tracking system[C]. Proceeding of IEEE International Conference on Robotics and Automation, 2007
[21 Isaac Cohen, Gérard Medioni. Detecting and tracking moving objects for video surveillance[C]. IEEE Proceeding of Computer Vision and Pattern Recognition, Fort Collins CO, June 23-25, 1999.

[22] Isaac Cohen, Gérard Medioni. Detection and tracking of objects in airborne video imagery[C]. CVPR'98, 1998

[23] Pedro A. Rodriguez, William J. Geckle, Jeffrey D. Barton, John Samsundar, Tia Gao, Myron Z. Brown, Sean R. Martin. An emergency response UAV surveillance system[C]. Proceedings of AMIA 2006 Symposium, 2006: 1078

[24] Thomas Loveard, Paul Boxer. Automated target detection and tracking for UAV platforms. http://www.auvsi.org/unmannedscience/newsletter/attachments/86/Loveard_T.pdf

[25] Pitu Mirchandani, Mark Hickman, Alejandro Angel, Dinesh Chandnani. Application of aerial video for traffic flow monitoring and management[C]. Proceedings of Pecora 15/Land Satellite Information IV, Denver, USA. on CD-ROM.

[26] Robert T. Collins, Alan J. Lipton, Takeo Kanade, Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt and Lambert Wixson. A system for video surveillance and monitoring[R]. Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000

[27] Charles M. Higgins, John K. Douglass, Nicholas J. Strausfeld. The computational basis of an identified neuronal circuit for elementary motion detection in dipterous insects[J]. Visual Neuroscience, 2004, 21: 567-586.

[28] Vivek Pant. Modular Neuromorphic VLSI architecture for visual motion and target tracking[D]. The University of Arizona, USA, 2003

# 22

# Adaptive Beamforming Algorithm Using a Pre-filtering System

Omar Abu-Ella[1] and Bashir El-Jabu[2]
[1]*Assistant Lecturer, Seventh October University,*
[2]*Associate Professor, Higher Industrial Institute, Misurata,*
*Libya*

## 1. Introduction

As depicted in Fig. 1, adaptive beamforming algorithms can be classified into two categories: non-blind adaptive algorithms and blind adaptive algorithms [1]. Non-blind adaptive algorithms rely on statistical knowledge about the transmitted signal in order to converge to a solution. This is typically accomplished through the use of a pilot training sequence sent over the channel to the receiver to help it identifying the desired user. On the other hand, blind adaptive algorithms do not require prior training, and hence they are referred to as "blind" algorithms. These algorithms attempt to extract salient characteristic of the transmitted signal in order to separate it from other users in the surrounding environment [2]. In this chapter we propose a technique that could be used with the blind or non blind algorithms to enhance their performance. This technique acts on the input signal vector **x**(k) as a band pass filter but in spatial domain, so it minimizes the noise and interference effects as a function of the Direction of Arrivals (DOAs).

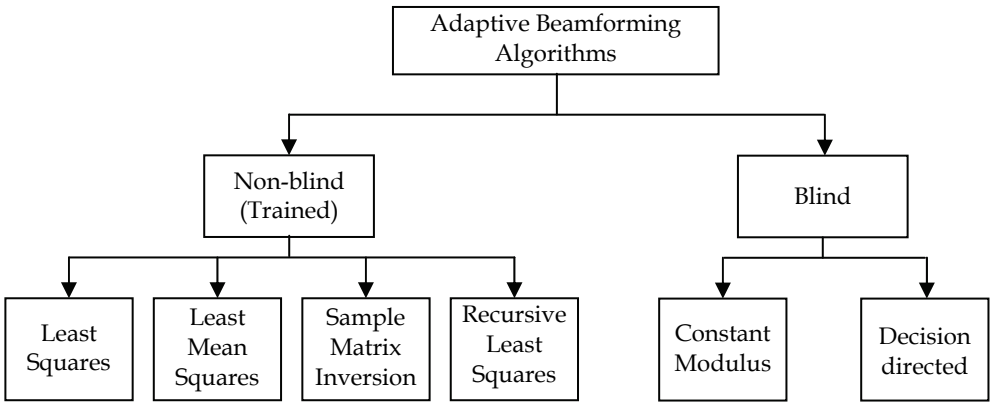



Fig. 1. General classifications of adaptive array algorithms

## 2. Proposed hybrid adaptive beamforming technique

The proposed technique which introduced in [3] aims to increase the Signal-to-Interference and Noise Ratio (SINR) of the beamforming system by reducing the interference and noise effects on the desired user signal using filtering in spatial domain, or extracting the desired signal from the instantaneous input signal vector **x**(k) of the beamformer, as can be seen in Fig. 2. In this context, it is worth pointing out that in image processing, especially in image compressing techniques, one can find an abundance of techniques that can reconstruct the original image with acceptable performance, without using all transformation components, but rather using only the lower component coefficients of the image transform matrix [4]. We will exploit this fact with some modifications and employ it in the antenna array processing to obtain a new beamforming technique.

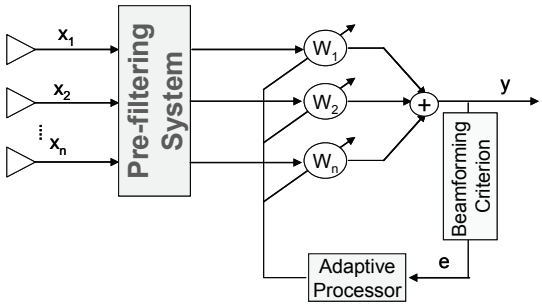



Fig. 2. Proposed blind adaptive beamformer with pre-filtering system

Since the interfering signals are in the same frequency band of the desired signal, we investigate representing them in another domain other than frequency domain so as to distinguish between the mixed signals that form the input signal. Therefore, the proposed technique is based on the idea that the desired and interfering signals arrive at the antenna array from different directions. Thus, we exploit these differences between arriving signals. The distinction can be obtained by converting the input signal to the spectrum of the spatial domain (this domain is the sine of the direction of arrival, or $\sin\theta$ domain). The desired signal can be extracted from the input system signals simply by making a band-pass filter in the spectrum of the spatial domain, i.e., in the $\sin\theta$ spectrum. This filtering process is shown in Fig. 3, and can be explained generally as follows:

- The original input signal is passed through a Fast Fourier Transform (FFT) stage to obtain its coefficients in the spectrum domain.
- The Most Significant Coefficient (MSC) of the transformed signal is selected. This is ranked as the largest sample of the transformed desired signal.
- The most significant coefficient is placed at its rank in the M zeros element vector (zero padding).
- The Inverse Fast Fourier Transform (IFFT) is applied to the filtered vector of the previous step to reconstruct an alternative input signal that contains a reduced amount of interference and noise.
- The reconstructed data vector is used as input signal to the adaptive beamforming system.

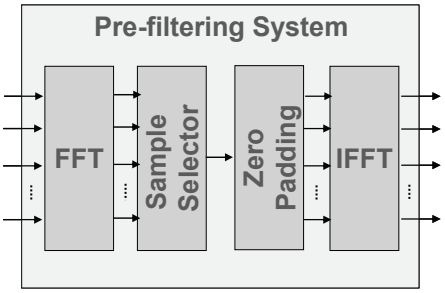



Fig. 3. Pre-filtering system

Mathematically, assuming that the narrow band array propagation vector for the $\theta$ direction of arrival is given by

$$\mathbf{a}(\theta) = e^{-j2\pi\frac{d}{\lambda}\sin\theta(m-1)},\ m = 1,2,\cdots,M \tag{1}$$

where $M$ is the number of array elements, $d$ is the spacing distance between any two adjacent elements, and $\lambda$ is the wavelength of the operating carrier frequency. Applying the FFT on the array propagation (Equation 1) gives

$$\begin{aligned}\mathbf{A}(\Theta) &= FFT\left(\mathbf{a}(\theta)\right)\\ &= \frac{1}{M}\sum_{m=1}^{M} e^{-j2\pi\frac{d}{\lambda}\sin\theta(m-1)} e^{-j2\pi K(m-1)/M}\\ &= \frac{1}{M}\sum_{m=1}^{M} e^{-j2\pi\left(\frac{d}{\lambda}\sin\theta(m-1)+\frac{K}{M}(m-1)\right)}\end{aligned} \tag{2}$$

Let $d = \lambda/2$, then

$$\begin{aligned}\mathbf{A}(\Theta) &= \frac{1}{M}\sum_{m=1}^{M} e^{-j2\pi\left(\frac{K}{M}+\frac{\sin\theta}{2}\right)(m-1)}\\ &= \frac{1}{M}\frac{e^{j\pi(2K+M\sin\theta)/M} - e^{-j2\pi(2K+M\sin\theta)(M-1)/M}}{e^{j\pi(2K+M\sin\theta)/M} - 1}\\ &= \frac{1}{M}\left\{\frac{e^{j\pi\left(\frac{2K+\sin\theta}{M}\right)\left(\frac{3-2M}{2}\right)}\left[e^{j\pi\left(\frac{2K+M\sin\theta}{M}\right)\left(M-\frac{1}{2}\right)} - e^{-j\pi\left(\frac{2K+M\sin\theta}{M}\right)\left(M-\frac{1}{2}\right)}\right]}{e^{j\pi\left(\frac{2K+\sin\theta}{2M}\right)}\left[e^{j\pi\left(\frac{2K+M\sin\theta}{2M}\right)} - e^{-j\pi\left(\frac{2K+M\sin\theta}{2M}\right)}\right]}\right\}\\ &= e^{-j\pi\left(\frac{2K+\sin\theta}{M}\right)(M-1)}\left[\frac{\sin\left(\pi\left(\frac{2K+M\sin\theta}{M}\right)\left(M-\frac{1}{2}\right)\right)}{M\sin\left(\pi\left(\frac{2K+M\sin\theta}{2M}\right)\right)}\right]\end{aligned} \tag{3}$$

By differentiating Equation (3) and equating the result to zero, the following formula gives the index $K_{MSC}$ (or the order) of the most significant coefficient as a function of the direction of arrival $\theta$ and the number of array elements $M$

$$K_{MSC} = \left[ -\frac{1}{2} \left( \frac{M\left[ 2\pi M \sin\theta - \pi \sin\theta + j2 \tanh^{-1}\left( \frac{M-\frac{1}{2}}{M-1} \right) \right]}{\pi(2M-1)} \right) \right]_{\text{int}} \tag{4}$$

where $[x]_{\text{int}}$ means the integer part of $x$.

Since the index $K_{MSC}$ must be a real integer, the formula in Equation (4) is modified to take the form

$$K_{MSC} = \underset{M}{\text{mod}} \left\{ 1 + M \left[ -\frac{1}{2} \left( \frac{M(2\pi M \sin\theta - \pi \sin\theta)}{\pi(2M-1)} \right) \right]_{\text{int}} \right\} \tag{5}$$

where $\underset{M}{\text{mod}}$ is the modulus notation, performed on M points.

Equation (5) can be simplified to

$$K_{MSC} = \underset{M}{\text{mod}} \left( \left[ 1 - \tfrac{1}{2} M \sin\theta \right] \right) \tag{6}$$

To illustrate the idea of the pre-filtering system consider, as an example, an array of 4 elements, with desired DOA = –45˚and signal sample $s_d$= - 0.0035 - 0.1000i. Then, from Equation (1) the array propagation factor at the desired direction is

$$\mathbf{a}(\theta_d) = \begin{bmatrix} 0.9990 - 0.0000\text{i} \\ -0.6049 + 0.7958\text{i} \\ -0.2663 - 0.9623\text{i} \\ 0.9280 + 0.3715\text{i} \end{bmatrix}$$

Thus, the desired signal vector $\mathbf{s} = s_d \mathbf{a}(\theta_d)$ is given by

$$\mathbf{s} = \begin{bmatrix} -0.0035 - 0.0999\text{i} \\ 0.0817 + 0.0577\text{i} \\ -0.0953 + 0.0300\text{i} \\ 0.0339 - 0.0941\text{i} \end{bmatrix}$$

And assuming that there is an undesired signal vector u, such that:

$$\mathbf{u} = \begin{bmatrix} -0.0241 - 0.7054\text{i} \\ -0.7078 + 0.0253\text{i} \\ 0.0239 + 0.7065\text{i} \\ 0.7069 - 0.0256\text{i} \end{bmatrix}$$

Thus, by adding the weighted desired and the desired signals the total input signal is given as:

$$\mathbf{x} = \begin{bmatrix} -0.0276 - 0.8053i \\ -0.6261 + 0.0830i \\ -0.0714 + 0.7365i \\ 0.7408 - 0.1197i \end{bmatrix}$$

By passing this input data vector through the FFT stage, we get the following transform coefficients vector

$$\mathbf{X} = \begin{bmatrix} 0.0156 - 0.1055i \\ 0.2465 - 0.1749i \\ -0.2137 - 0.0322i \\ -0.1589 - 2.9087i \end{bmatrix}$$

Appling Equation (6) on the given DOA and the number of array elements we find that the most significant coefficient is the second coefficient. Replacing the other coefficients by zeros, we get the following vector

$$\mathbf{X}_z \begin{bmatrix} 0 \\ 0.2465 - 0.1749i \\ 0 \\ 0 \end{bmatrix}$$

Now, applying the IFFT on the zero padded vector we get the reconstructed input vector as follows

$$\tilde{\mathbf{x}} = \begin{bmatrix} 0.0616 - 0.0437i \\ 0.0437 + 0.0616i \\ -0.0616 + 0.0437i \\ -0.0437 - 0.0616i \end{bmatrix}$$

To take advantage of the pre-filtering technique to reduce the mean square error (MSE), we determine the fluctuations of the original input data vector $\tilde{\mathbf{x}}$ and the reconstructed input data vector $\mathbf{x}$ from the desired signal vector $\mathbf{s}$

$$\mathbf{s} - \mathbf{x} = \begin{bmatrix} 0.0241 + 0.7054i \\ 0.7078 - 0.0253i \\ -0.0239 - 0.7065i \\ -0.7069 + 0.0256i \end{bmatrix}$$

$$\mathbf{s} - \tilde{\mathbf{x}} = \begin{bmatrix} -0.0652 - 0.0562i \\ 0.0379 - 0.0039i \\ -0.0338 - 0.0137i \\ 0.0776 - 0.0324i \end{bmatrix}$$

MSE of the conventional technique = $20\log\left(\left|\mathbf{s}-\mathbf{x}\right|\right)$ = 9.0306 dBv

MSE of the proposed technique = $20\log\left(\left|\mathbf{s}-\tilde{\mathbf{x}}\right|\right)$ = -12.2258 dBv.

It is clear that the proposed technique reduces the MSE in this example by more than 21 dB; this improvement is significant in communications systems.

As explained in the general steps shown in Fig. 4, the output of the pre-filtering stage is applied as an input to the beamforming algorithms. The flowchart of the whole process using the constant modulus algorithm CMA(1,2) is as indicated in the figure. Other algorithms (other than the CMA) can also be used.

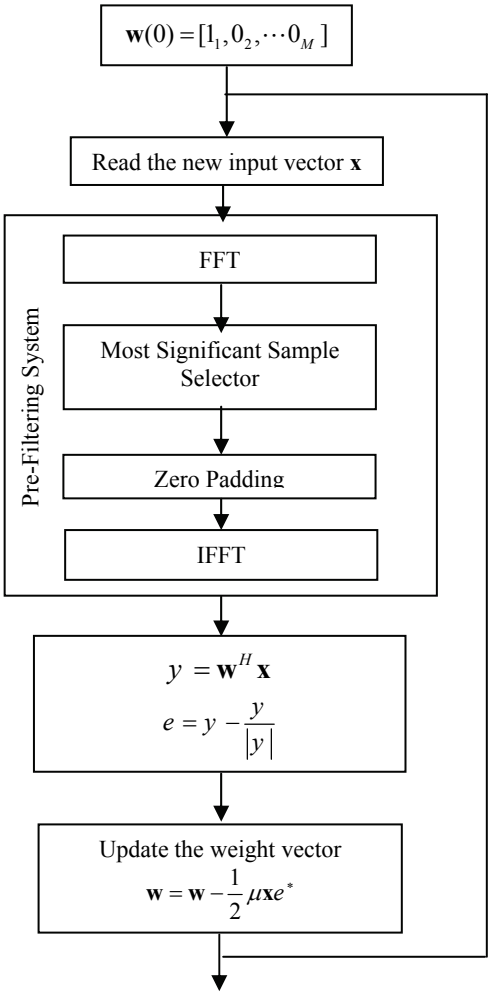



Fig. 4. Flowchart of the proposed hybrid technique

## 3. Quadratic diagram and the SIR improvement in the beamformer input signal using the proposed hybrid technique

The quadratic (i.e. the complex plane) diagram of the input signal vector to the beamformer system is represented in Fig.5, where $\mathbf{s}_i, \mathbf{u}_i$ and $\mathbf{x}_i$ represent, respectively, the desired, undesired and the total input signal on the ith element of the array . For clarity, only the angles of $\mathbf{s}_1, \mathbf{u}_1, \mathbf{x}_1$ are shown.

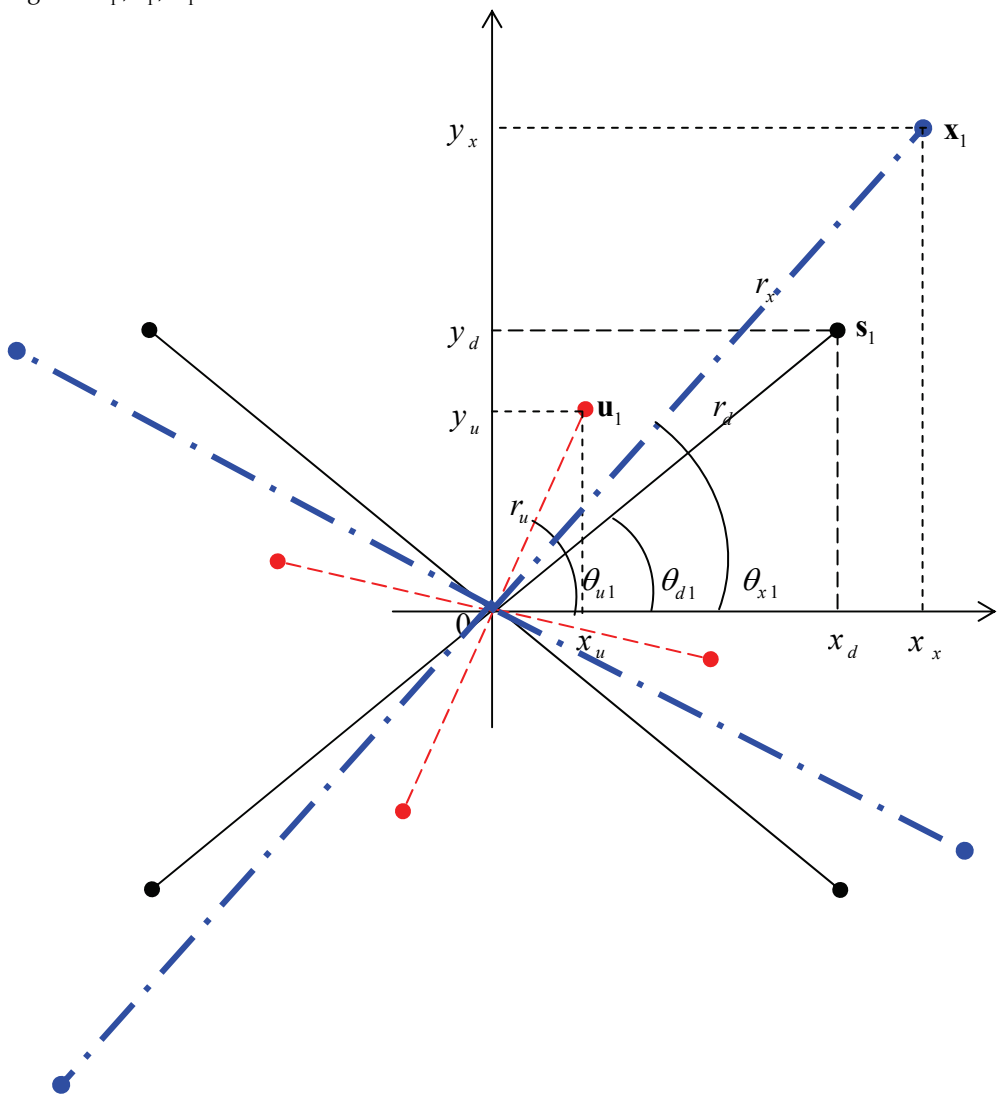



Fig. 5. The desired, interfering and the total input signal vectors in the quadratic plane

The desired signal vector can be represented as

$$\mathbf{s} = s\mathbf{a}(\theta_d) \tag{7}$$

where $\mathbf{a}(\theta_d)$ is the array propagation vector corresponding to the desired direction and it is given as,

$$\mathbf{a}(\theta_d) = e^{-j\frac{2\pi}{\lambda}(m-1)d\sin\theta_d}\ , m = 1,2,\cdots,M$$

and $s = r_d e^{j\phi_d}$ is the desired signal sample. Equation (7) can be put in the form

$$\mathbf{s} = r_d e^{j\phi_d} e^{-j\frac{2\pi}{\lambda}(m-1)d\sin\theta_d} \tag{8}$$

Assuming $d = \lambda/2$, the desired signal vector can be written in the form

$$\mathbf{s} = r_d e^{j(\phi_d - \pi(m-1)\sin\theta_d)} \tag{9}$$

The complex rectangular coordinates of the vectors can be expressed as

$$x_d = r_d\cos\left(\phi_d - \pi\left(m-1\right)\sin\theta_d\right) \tag{10a}$$

$$y_d = r_d\sin\left(\phi_d - \pi\left(m-1\right)\sin\theta_d\right)\ m = 1,2,\cdots,M \tag{10b}$$

It is clear that the resulting signal is an array containing vectors with a constant magnitude $r_d$, and each vector has a specific phase $\phi_d(m) = -\pi(m-1)\sin\theta_d$, $m = 1,2,\cdots,M$. Similarly, the undesirable signal vector can be put in the form

$$\mathbf{u} = r_u e^{j(\phi_u - \pi(m-1)\sin\theta_u)} \tag{11}$$

the complex rectangular coordinates of which can be expressed as

$$x_u = r_u\cos\left(\phi_u - \pi\left(m-1\right)\sin\theta_u\right)$$

$$y_u = r_u\sin\left(\phi_u - \pi\left(m-1\right)\sin\theta_u\right) \tag{12}$$

In case of interference, the total received signal can be put as

$$\mathbf{x} = \mathbf{s} + \mathbf{u} \tag{13}$$

$$\mathbf{x} = r_d e^{j(\phi_d - \pi(m-1)\sin\theta_d)} + r_u e^{j(\phi_u - \pi(m-1)\sin\theta_u)} \tag{14}$$

for which, the complex rectangular coordinates may be written as:

$$x_x = r_d\cos\left(\phi_d - \pi\left(m-1\right)\sin\theta_d\right) + r_u\cos\left(\phi_u - \pi\left(m-1\right)\sin\theta_u\right)$$

$$y_x = r_d\sin\left(\phi_d - \pi\left(m-1\right)\sin\theta_d\right) + r_u\sin\left(\phi_u - \pi\left(m-1\right)\sin\theta_u\right) \tag{15}$$

The resulting overall vector magnitude can be obtained as follows

$$\begin{aligned} r_x &= \sqrt{{x_x}^2 + {y_x}^2} \\ &= \sqrt{\left(x_d + x_u\right)^2 + \left(y_d + y_u\right)^2} \\ &= \sqrt{\left({x_d}^2 + {y_d}^2\right) + \left({x_u}^2 + {y_u}^2\right) + 2\left(x_d x_u + y_d y_u\right)} \\ &= \sqrt{\left({r_d}^2 + {r_u}^2\right) + 2 r_d r_u \psi} \end{aligned} \tag{16}$$

where

$$\begin{aligned} \psi = &\cos\left(\phi_d - \pi(m-1)\sin\theta_d\right)\cos\left(\phi_u - \pi(m-1)\sin\theta_u\right) \\ &+ \sin\left(\phi_d - \pi(m-1)\sin\theta_d\right)\sin\left(\phi_u - \pi(m-1)\sin\theta_u\right) \end{aligned} \tag{17}$$

is the variation in the total signal magnitude (desired and undesired signals) caused by the interference (undesirable signals). Here, $\psi(m)$ is not a constant value for all values of $m$ except when $\theta_d$ and $\theta_u$ are equal; in this case the proposed technique will not improve SIR; but generally most of beamforming algorithms fail in this case. Since we have assumed in this chapter that the interfering signal arrive at a direction different from the desired signal direction, $r_x(m)$ can not to be a constant value, i.e., the total input signal vector $\mathbf{x}(m)$ will appear scattered in the complex plane.

To show the Signal-to-Interference Ratio (SIR) improvement in the input signal vector using the proposed hybrid beamforming technique we determine the gained signal-to-interference ratio in the reconstructed signal vector using the pre-filtering technique of the hybrid beamforming system. Assuming a linear system, the total input signal vector $x(m)$ is the sum of the desired signal **s**(m) and the undesired one $\mathbf{u}(m)$ –i.e.-

$$\mathbf{x}(m) = \mathbf{s}(m) + \mathbf{u}(m) \tag{18}$$

Similarly, for the reconstructed input data vector $\tilde{x}(m)$ ,

$$\tilde{\mathbf{x}}(m) = \tilde{\mathbf{s}}(m) + \tilde{\mathbf{u}}(m) \tag{19}$$

where $\tilde{\mathbf{s}}(m)$, and $\tilde{\mathbf{u}}(m)$ are the desired and undesired reconstructed signals vector respectively. To generalize the procedure, we will assume that an arbitrary signal sample $v = re^{j\phi}$ arrives to the array at $\theta$ direction. Therefore, the array signal vector of this signal will be in the form

$$\mathbf{v}(m) = re^{j\phi}e^{-j\pi\sin\theta(m-1)}, \quad m = 1, 2, \cdots, M \tag{20}$$

The fast Fourier transform of the resulting signal vector $\mathbf{v}(m)$ is

$$\begin{aligned} \mathbf{V}(k) &= FFT\left[\mathbf{v}(m)\right] \\ &= \sum_{m=1}^{M} \mathbf{v}(m) \cdot e^{j2\pi(k-1)(m-1)/M} \\ &= \sum_{m=1}^{M} re^{j\phi} \cdot e^{j2\pi(k-1)(m-1)/M} \end{aligned}$$

$$= re^{j\phi} \cdot e^{-j\frac{\pi}{2}\left[M\sin\theta+2(k-1)\right](M+1)/M} \cdot \left[ \frac{\sin\left(\frac{\pi}{2}\left[M\sin\theta+2(k-1)\right]\right)}{\sin\left(\frac{\pi}{2}\left[M\sin\theta+2(k-1)\right]/M\right)} \right] \tag{21}$$

Now, the value of the specific $k$th coefficient from the transformed signal vector $\mathbf{V}(k)$ can be calculated. This coefficient will be selected to reconstruct the signal vector $\tilde{\mathbf{v}}(m)$. Substituting for the $k$th index in Equation (6) into Equation (21) gives the value of the most significant coefficient as

$$\mathbf{V}(k_{MSC}) = re^{j\phi} \cdot e^{-j\frac{\pi M}{2}\left[\sin\theta-\sin\theta_d\right](M+1)/M} \cdot \left[ \frac{\sin\left(\frac{-\pi M}{2}\left[\sin\theta-\sin\theta_d\right]\right)}{\sin\left(\frac{\pi}{2}\left[\sin\theta-\sin\theta_d\right]\right)} \right] \tag{22}$$

This value will be put in a zeros vector in the rank of the $k_{MSC}$ so that we can express the resulting signal vector as

$$\mathbf{V}(k_{MSC}) = re^{j\phi} \cdot e^{-j\frac{\pi M}{2}\left[\sin\theta-\sin\theta_d\right](M+1)/M} \cdot \left[ \frac{\sin\left(\frac{-\pi M}{2}\left[\sin\theta-\sin\theta_d\right]\right)}{\sin\left(\frac{\pi}{2}\left[\sin\theta-\sin\theta_d\right]\right)} \right] \cdot \delta(k-k_{MSC}), \quad k = 1,2,\cdots M \tag{23}$$

where

$$\delta(k-k_{MSC}) = \begin{cases} 1, k = k_{MSC} \\ 0, \text{elsewhere} \end{cases}$$

The reconstructed signal vector can be obtained by applying inverse Fourier transformation to Equation (23) as follows

$$\tilde{\mathbf{v}}(m) = \frac{1}{M}\sum_{k=1}^{M} \mathbf{V}(k_{MSC}) e^{j2\pi(k-1)(m-1)/M}$$

$$= \frac{1}{M}\mathbf{V}(k_{MSC}) e^{-j\pi\sin\theta_d(m-1)}$$

$$= re^{j\phi} e^{-j\pi\sin\theta_d(m-1)} e^{-j\frac{\pi M}{2}\left[\sin\theta-\sin\theta_d\right](M+1)/M} \cdot \frac{1}{M}\left[ \frac{\sin\left(\frac{-\pi M}{2}\left[\sin\theta-\sin\theta_d\right]\right)}{\sin\left(\frac{\pi}{2}\left[\sin\theta-\sin\theta_d\right]\right)} \right] \tag{24a}$$

Therefore, we can put the signals in the form

$$\tilde{\mathbf{v}} = G_1(r,\phi,\theta)\cdot G_2(\theta) \tag{24b}$$

where

$$G_1(r,\phi,\theta) = re^{j\phi}\cdot e^{-j\pi\sin\theta_d(m-1)}\cdot e^{-j\frac{\pi M}{2}[\sin\theta-\sin\theta_d](M+1)/M} \tag{24c}$$

It is clear from Equation (24 c) that the magnitude of $G_1$ equals the signal magnitude $r$ and

$$G_2(\theta) = \frac{1}{M}\left[\frac{\sin\left(\frac{-\pi M}{2}[\sin\theta-\sin\theta_d]\right)}{\sin\left(\frac{\pi}{2}[\sin\theta-\sin\theta_d]\right)}\right] \tag{24d}$$

The variation of $G_2$ as a function of $\theta$ is shown in Fig. 6 from which it is clear that its value equal unit only if and only $\theta = \theta_d$. Fig. 7 shows the effects of the gains on the magnitude of the desired, undesired and the total input signal.
In Fig. 7, it is assumed that $r_d > r_u$ without loss of generality.
Note that, in case of $\tilde{\mathbf{s}}$

$$G_1(r_d,\phi_d,\theta_d) = r_d e^{j\phi d}\, e^{-j\pi\sin\theta_d(m-1)} \tag{25a}$$

and from Equation (2d)

$$G_2(\theta_d) = 1 \tag{25b}$$

Note that, in case of $\tilde{\mathbf{u}}$

$$G_1(r_u,\phi_u,\theta_u) = r_u e^{j\phi_u}\cdot e^{-j\pi\sin\theta_d(m-1)}\cdot e^{-j\frac{\pi M}{2}[\sin\theta_u-\sin\theta_d](M+1)/M} \tag{25c}$$

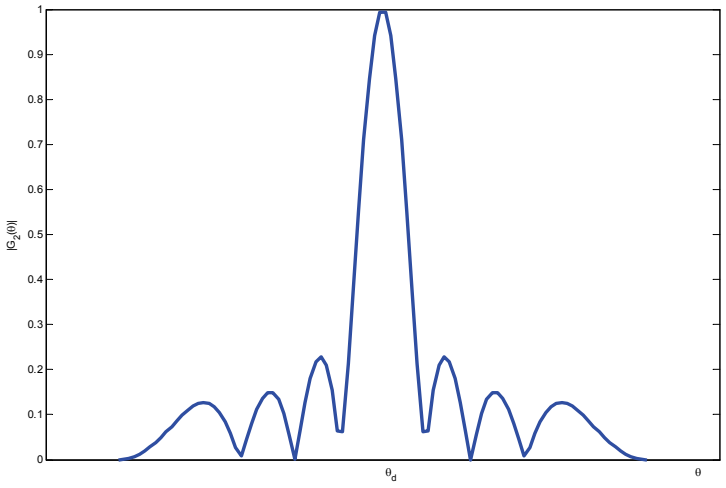



Fig. 6. Magnitude of $G_2$ versus DOA

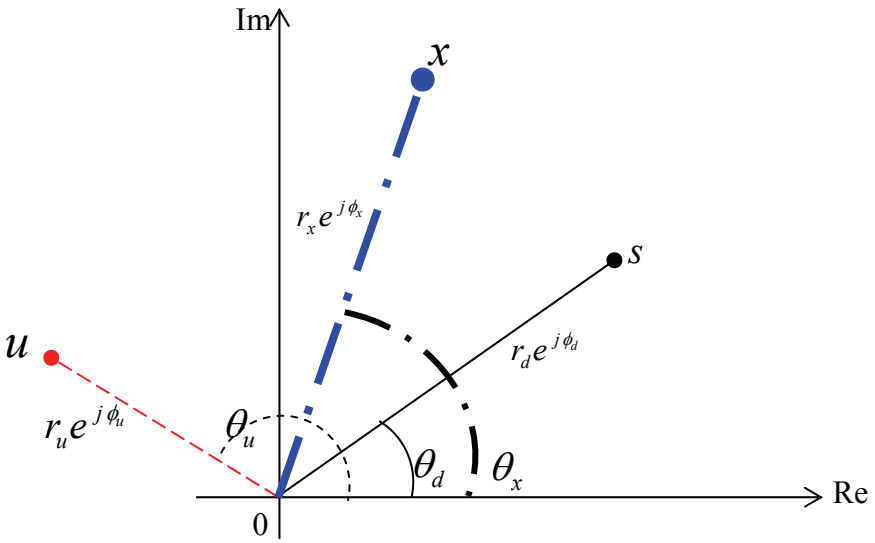



Fig. 7. Desired, interference and the total input signal in the quadratic plane

and

$$G_2(\theta_u) = \frac{1}{M}\left[\frac{\sin\left(\frac{-\pi M}{2}\left[\sin\theta_u - \sin\theta_d\right]\right)}{\sin\left(\frac{\pi}{2}\left[\sin\theta_u - \sin\theta_d\right]\right)}\right] \tag{25d}$$

Note that in case of the desired signal (i.e. $\theta = \theta_d$) $G_2$ will be equals one i.e. the maximum value. However, in case of the undesired signal (i.e. $\theta = \theta_u$) $G_2$ will be attenuated i.e. will be less than one.

The overall reconstructed input signal vector $\tilde{\mathbf{x}}(m)$, which consists of the desired signal vector and attenuated interfering signal vector can be given by

$$\tilde{\mathbf{s}} = G_1(r_d, \phi_d, \theta_d) \cdot G_2(\theta_d)$$

and

$$\tilde{\mathbf{u}} = G_1(r_u, \phi_u, \theta_u) \cdot G_2(\theta_u) \tag{26}$$

Substituting Equation (26) into Equation (19) gives

$$\begin{aligned}\tilde{\mathbf{x}}(m) &= r_d e^{j\phi_d} e^{-j\pi \sin\theta_d (m-1)} \\ &+ r_u e^{j\phi_u} e^{-j\pi \sin\theta_d (m-1)} e^{-j\frac{\pi M}{2}\left[\sin\theta_u - \sin\theta_d\right](M+1)/M} \\ &\cdot \frac{1}{M}\left[\frac{\sin\left(\frac{-\pi M}{2}\left[\sin\theta_u - \sin\theta_d\right]\right)}{\sin\left(\frac{\pi}{2}\left[\sin\theta_u - \sin\theta_d\right]\right)}\right]\end{aligned} \tag{27}$$

Note that the overall reconstructed $\tilde{\mathbf{x}}(m)$ (overall reconstructed signal on the $m$th element) has a constant magnitude, because its derivative with respect to m is zero (i.e. $\frac{d}{dm}\left(\left|\tilde{\mathbf{x}}(m)\right|\right)=0$), which means that the hybrid technique eliminates the magnitude fluctuation which is described in Equations (16) and (17).

As a result, we can evaluate the signal-to-interference ratio SIR improvement in the reconstructed input vector $\tilde{\mathbf{x}}(m)$. Since the magnitude of the desired signal vector stays constant without any attenuation and the interfering signal vector magnitude is decreased by the factor $\left|\frac{1}{M}\left[\frac{\sin\left(-\pi M/2[\sin\theta_u-\sin\theta_d]\right)}{\sin\left(\pi/2[\sin\theta_u-\sin\theta_d]\right)}\right]\right|$, therefore, the maximum possible gained $SIR_G$ in dB can be expressed as

$$SIR_G = 20\log\left\{\left|\frac{M\sin\left(\frac{\pi}{2}\left[\sin\theta_u-\sin\theta_d\right]\right)}{\sin\left(\frac{-\pi M}{2}\left[\sin\theta_u-\sin\theta_d\right]\right)}\right|\right\} dB \tag{28}$$

A simulation setup will be used to show the improvements in the reconstructed input signal vector. Fig. 8 shows a quadratic diagram representation of a desired signal, arriving at DOA of 60˚, an interfering signal with arbitrary DOA of 40˚, the original received signal without using the pre-filtering technique, and the reconstructed signal after the pre-filtering process. From these results, one can observe that the reconstructed input data vector is closer to the desired signal than the original received input vector which is scattered by the interference effects. For this reason, the pre-filtering process is performed directly before the beamforming process in the proposed hybrid algorithm. This will help the beamformer in extracting the desired signal in easier and faster way.

Fig. 9 illustrates the quadratic diagram of a similar scenario but with the desired signal arriving at DOA of 45˚ and 6 undesired signals arriving at DOAs of (19˚, 60˚, 120˚, 240˚, 300˚, 341˚) with respect to the DOA of the desired signal. In this figure, one can observe the extensively scattering of the compound interfering signal which causes the scattering of the original input data vector. However, the same figure shows that the reconstructed input signal is not scattered and nearer to the desired signal. Thus, pre-filtering effectively reduces the noise and interfering.

In summary, the pre-filtering technique generates a new input data signal with lower undesired signal than the original one. This reduction is obtained previously by taking the most significant coefficient of the desired signal (and ignoring the other coefficients) (similar to bandpass filtering). To see the possibility of reducing the undesired signal by discarding the most significant coefficient of the undesired signal (similar to band reject filtering), the same scenario is used in Fig. 8 gave the results shown in Fig. 10 which illustrates lower performance.

## 4. Simulation results

This section compares the performance of the proposed hybrid adaptive algorithms with some of conventional adaptive beamforming algorithms. The factors used for this comparison are:

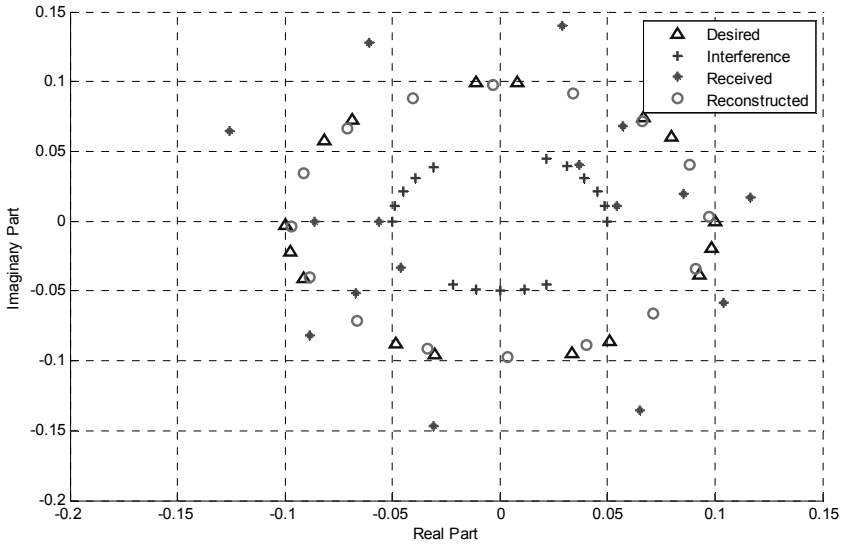



Fig. 8. Quadratic diagram in case of single interferer signal

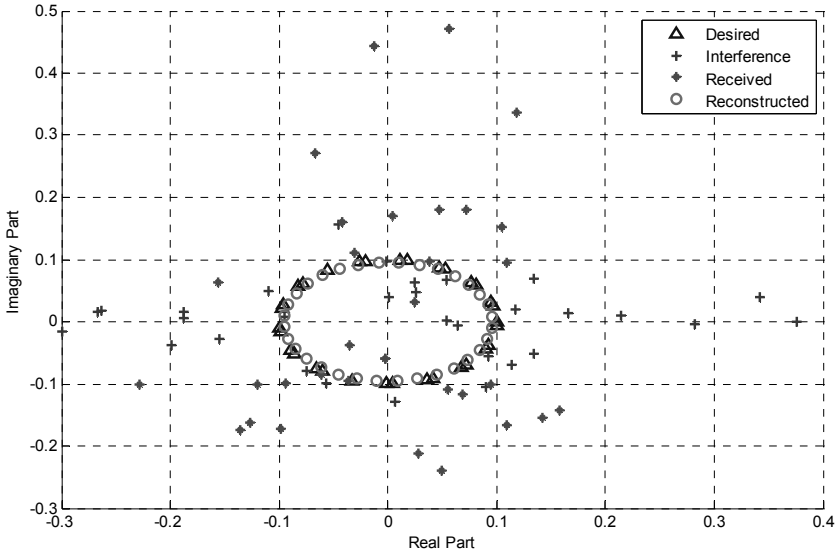



Fig. 9. Quadratic diagram in case of multiple interferer signals

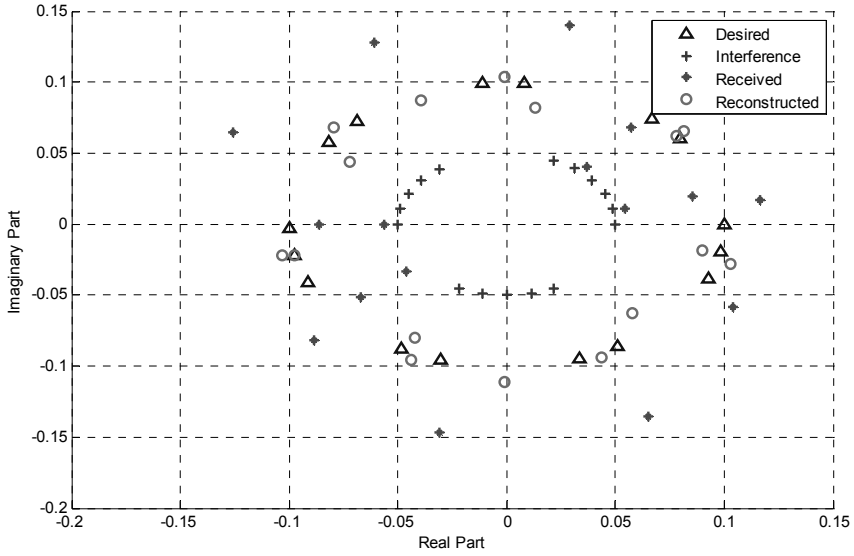



Fig. 10. Quadratic diagrams (complex plane) in case of discarding the interferer signal

- Beam Pattern Characteristics (Beam gain in small angular separation between desired and undesired signals).
- Resulting Signal-to-Interference Ratio SIR.
- Bit Error Rate (BER) behavior with Respect to SINR.
- System Capacity (Number of user in the system).
- The Outage Probability of the System.
- Computational Complexity.
- Convergence Speed.
- Tracking performances

In each case, simulation environment parameters are assumed to realize the system in the case. Overall simulation environments are written using MATLAB code.

### 4.1 Beam pattern performance

The array response to the signal with direction of arrival $\theta$ is given by [5]

$$G(\theta) = \sum_{i=1}^{M} w_i^* e^{-j\frac{2\pi}{\lambda}(i-1)d\sin\theta} \tag{29}$$

where $w_i$ is the $i$th value of the weight vector. This array response, $G(\theta)$, can also be represented in a vector form as

$$G(\theta) = \mathbf{w}^H \mathbf{a}(\theta) \tag{30}$$

where **w** is the weight vector and $H$ is the Hermitian notation. The beam pattern $G_{BP}$ defined as the magnitude of $G(\theta)$ has the form

$$G_{BP}(\theta) = \left| \mathbf{w}^H \mathbf{a}(\theta) \right| \tag{31}$$

Initially, a simulation is carried out assuming arrival direction of the desired signal at 60˚ interfering signal arrives from 50˚(small separation angle and equal power of the desired signal, which represents a tight condition) and the initial weights vector is taken as the first column of the identity matrix, i.e. $\mathbf{w}(0) = [1, 0, 0, \cdots, 0_M]^T$ . The results obtained after 1000 iterations, using an adaptive array system with eight antennas. The beam pattern obtained for the conventional Constant Modulus Algorithm (CMA), which is considered as one of the most efficient blind algorithms and the proposed technique. A comparison between these two results is displayed in Fig.11. It is clear that the maximum gain of the main lobe directed toward the desired direction for the proposed hybrid technique is greater than that for the conventional CMA by about 2 dB, and the minimum gain (null) at undesired direction for the proposed technique is less than that for the conventional CMA by 7 dB (better performance). This means that the proposed hybrid algorithm is more efficient in forming the beam pattern. The results are as displayed in Fig. 12 from which it is clear that the directivity of the hybrid algorithm is higher than that of the conventional one, i.e., the former is more capable in distinguishing between neighboring users (a serious problem for conventional blind adaptive beamforming algorithm).

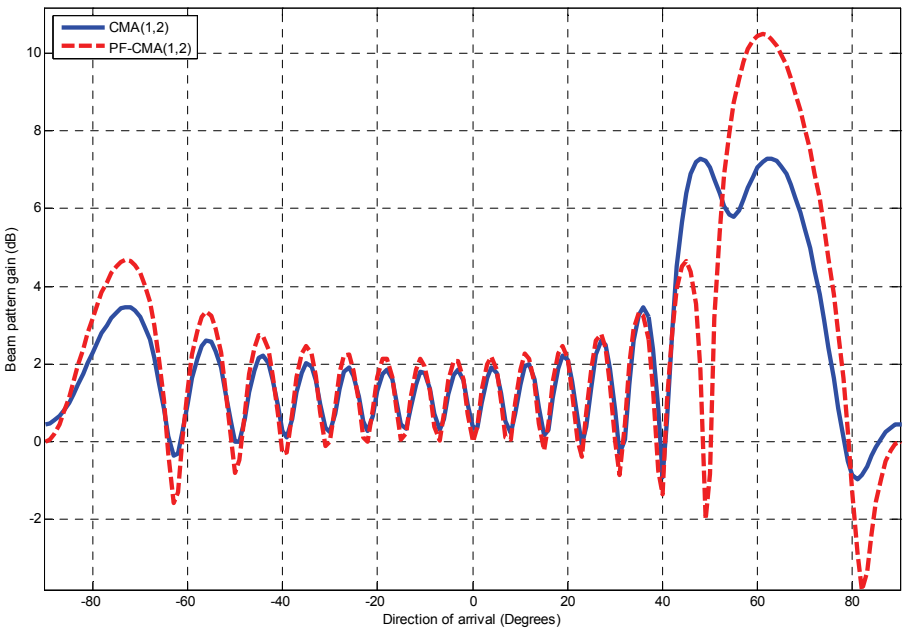
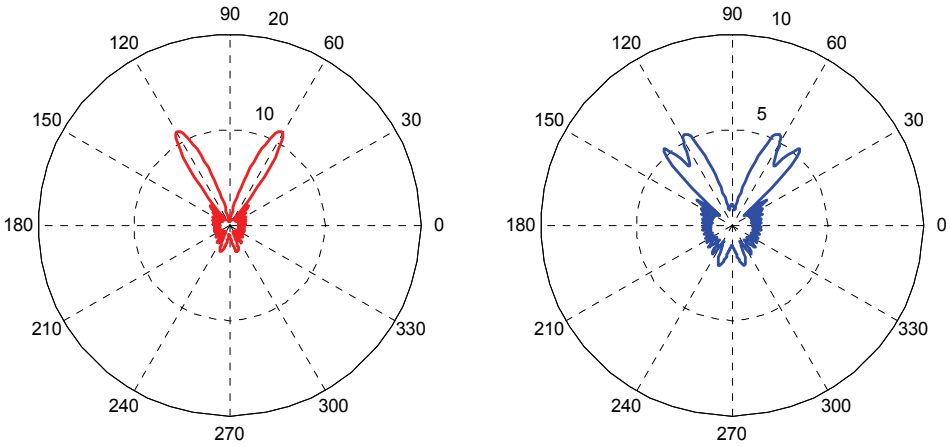



Fig. 11. Beam patterns of CMA and hybrid technique in (dB)




(a) Polar beam pattern of hybrid technique (b) Polar beam pattern of CMA(1, 2)

Fig. 12. Polar beam patterns of CMA and hybrid technique for small angular separation between desired and interfering DOAs

### 4.2 Signal-to-Interference Ratio (SIR) performance and power budget

The SIR for adaptive array can be written to have the form

$$SIR = \frac{\mathbf{w}^H s_d{}^2 \mathbf{a}(\theta_d)\mathbf{a}^H(\theta_d)\mathbf{w}}{\mathbf{w}^H \mathbf{R}_I \mathbf{w}} \tag{32}$$

where $\mathbf{a}(\theta_d)$ is the array propagation vector at the desired direction of arrival, $s_d$ is the desired signal sample, and $\mathbf{R}_I$ is the covariance matrix of the interference signal.

The simulation is carried out under the same previous conditions, but with lower noise level. Fig. 13 shows the results. The steady state of the hybrid technique is at about 23 dB while that of the conventional algorithm is less than 6 dB (improvement of more than 17 dB in favor of the proposed technique). The SIR for both algorithms is evaluated as a function of interfering DOA assuming the desired user is at 0˚. Fig. 14 indicates a better performance for the hybrid technique in most of the time.

To obtain the power budget in this tight case, i.e., at a smaller angular separation of 10 degrees, we determined the ratio between the gain beam pattern at desired direction, $G_d$, and the maximum possible gain $G_{\max}$, i.e. $G_d/G_{\max}$. This ratio is 11/16 for the hybrid CMA technique and 5/16 for the conventional CMA, another advantage of the proposed technique (Gain of the power budget more than 3 dB).

### 4.3 BER behavior with respect to SINR

This part studies the effects of the proposed technique on the required SINR to achieve a specific bit error rate. Through out this chapter the required BER is assumed to be $10^{-8}$ because of simulation considerations, since a BER of $10^{-k}$ needs $10^{(k+1)}$ samples for the simulation results to lie within a confidence interval of 90% [2]. The conventional CMA

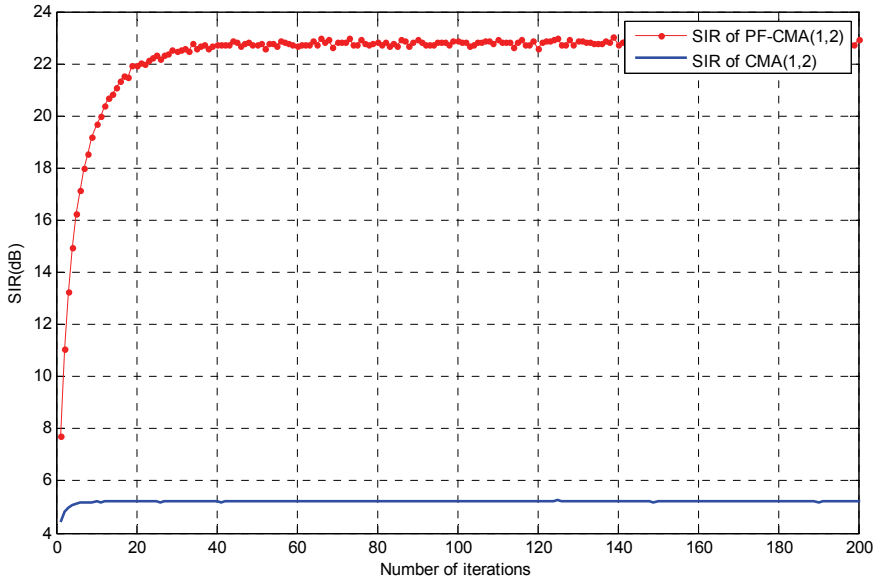



Fig. 13. SIR versus the number of iterations for the conventional CMA and hybrid technique

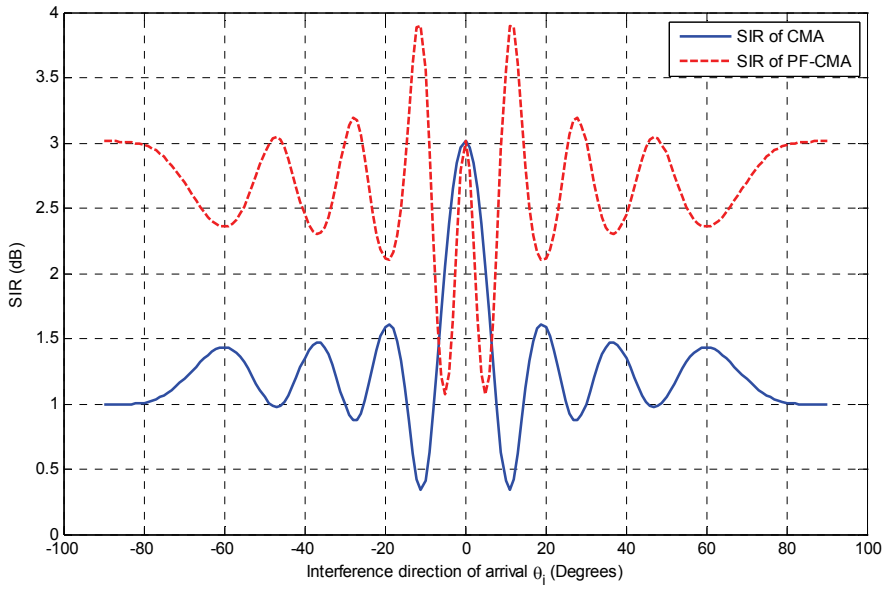
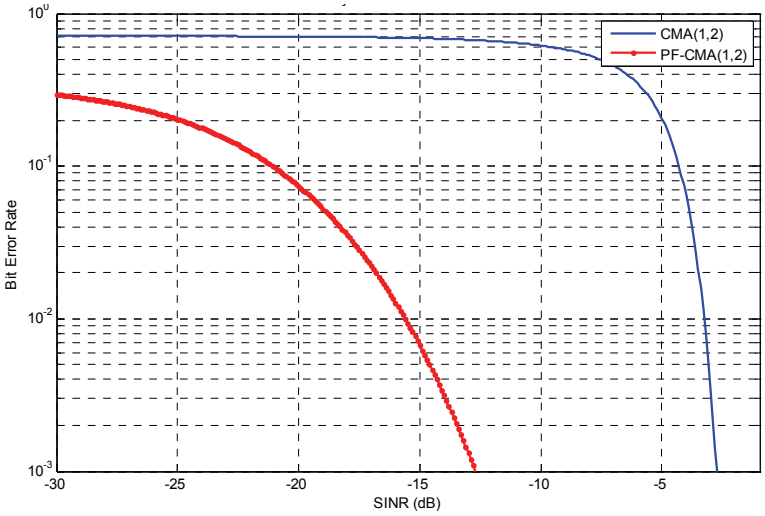
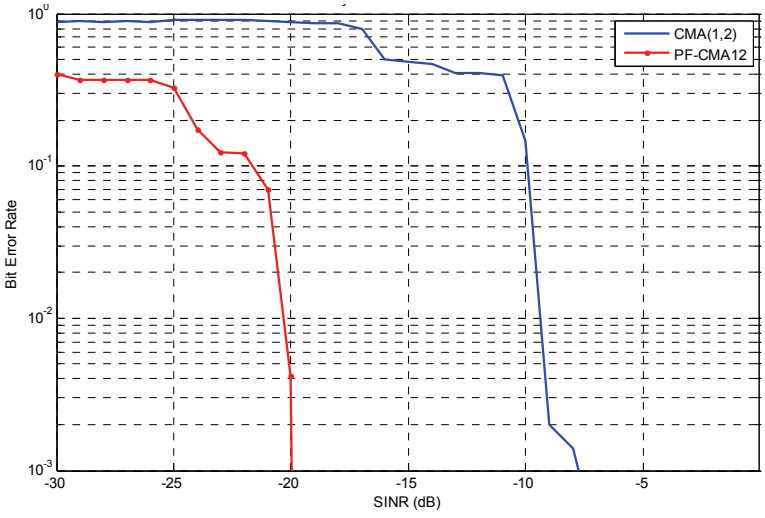



Fig. 14. SIR variations against interference direction of arrival for desired direction of arrival = 0˚

algorithm will be used as a reference. The results shown in Fig. 15 are obtained using an array receiving a desired signal at 0˚ and 6 interfering signals arriving at 19˚, 60˚, 120˚, 240˚, 300˚ and 341˚. The array uses 4 elements. These curves show that in case of SNR = 6 dB, the proposed hybrid technique needs -12.7 dB SINR to achieve $10^{-8}$ BER while the conventional CMA algorithm needs -2.7 dB to achieve the same BER (a difference of 10 dB in favor of the proposed technique). When the SNR is changed to 20 dB the difference becomes 12 dB.




(a) SNR = 6dB




(b) SNR = 20dB

Fig. 15. BER versus SINR using CMA, hybrid technique, using 4-antenna array

### 4.4 System capacity

Beamforming often affects system capacity. Here, effect of the proposed technique on the capacity of cellular CDMA system is considered. System capacity (in terms of the number of users M) can be determined using [7]

$$M = \frac{\left(\frac{E_b}{I_o}\right)\left(\frac{R_b}{B_c}\right)+1}{(1+\gamma)\left(\frac{E_b}{I_o}\right)\left(\frac{R_b}{B_c}\right)} \tag{33}$$

where $E_b$ is the energy per bit, $I_o$ is the interference power spectral density (PSD) in Watts/Hertz, $R_b$ is the message data rate in bits per second, $B_c$ is the radio channel bandwidth in Hertz, where $B_c >> R_b$, and

$$\gamma = \frac{\sum_{j=1}^{J} G_{I_j}}{G_d} \tag{34}$$

is the inverse of the total SIR experienced by the mobile in the cell under considerations from the cochannel cells assuming one mobile per cochannel cell, $G_d$ is the desired beam gain evaluated at direction $\theta_d$, $G_{Ij}$ is the beam gain of the $j$th interfering cell evaluated at $\theta_i$. The aim of this proposed hybrid technique is to suppress the $\gamma$ value by decreasing the interference beam gain at $\theta_i$ direction. The factor $\gamma$ is evaluated by simulating the beamforming system with 4- and 8-element array antenna, in environment with 20 dB SNR assuming each interfering user generates interference power equal to the desired user power. A CDMA cellular system is assumed with service bit rate of 8000 and 32000 bps for each user, and with chip rate of 4 Mcps. The results are displayed in Fig. 16. The results shows that for the case of 8000bps rate the systems capacities for $E_b/I_o$ = -6dB are 213 and 228 users for the hybrid system using 4- and 8-element array respectively whereas the conventional system capacity are 80 and 89 users for 4 and 8 elements array system respectively. Thus, the proposed algorithm increases the system capacity by 1.5 folds comparing with the conventional system capacity. The same trend applies for the case of 32000 bps service bit rate as shown in Fig. 16 (b).

### 4.5 The outage probability

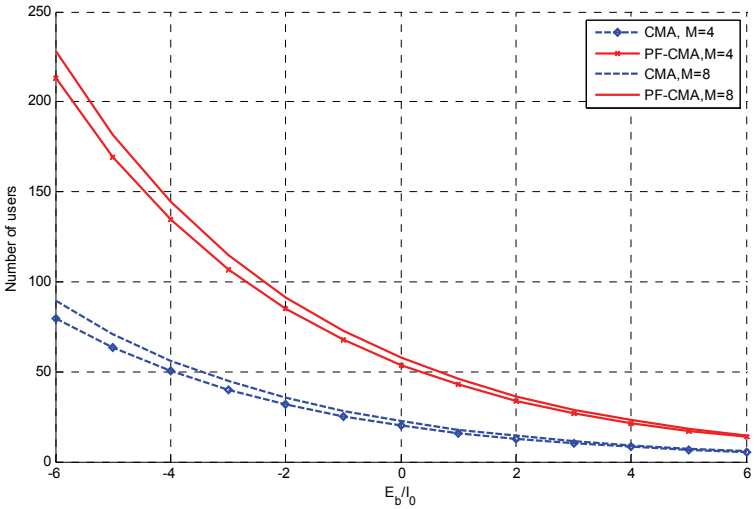
Here, the effect of the proposed technique on the outage probability for both up link and down link will be consider.

#### 4.5.1 Outage probability of the down-link system

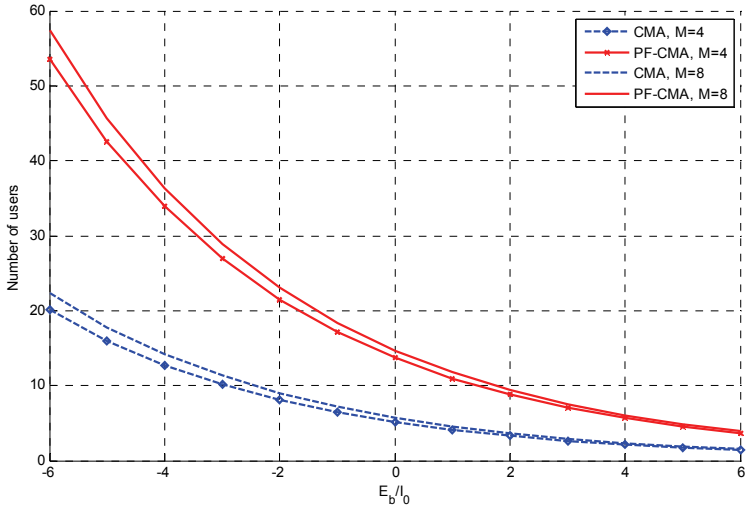
The performance of the system can be expressed in terms of the outage probability, i.e., the probability that the bit error rate (BER) exceeds a certain threshold required BER, ($BER_{\max}$), normally$10^{-8}$, i.e.

$$P_o = \Pr\left(BER > BER_{\max}\right) \tag{35}$$

The outage probability can be defined as the probability that the SIR falls below minimum required SIR [6]. In cellular CDMA the system outage probability could be determined by the following [7]




(a) Service bit rate =8000bps




(b) Service bit rate =32000bps

Fig. 16. Number of users versus Eb/Io

$$p_o = Q\left(\frac{\chi - E\left[\frac{I_{out}}{S}\right]}{\frac{M}{GP}\sqrt{\mathrm{var}\left[\frac{I_{out}}{S}\right]}}\right) \tag{36}$$

where $\chi = \frac{1}{\sigma_{req}} - \frac{I_{in}}{S} - \frac{\eta}{S}$, $E\left[I_{out}/S\right]$ is the ratio of the intercellular interference-to-signal ratio, $\sigma_{req}$ is the required SINR for the BER to be less than $BER_{\max}$, $M$ is the number of

users, $G_P$ is the processing gain, $\frac{\eta}{S}$ is the ratio of received thermal noise to user signal power, and

$$E\left[\frac{I_{out}}{S}\right] = \frac{(M+1)}{G_P} E[\gamma] = \frac{(M+1)}{G_P} \sum_{\substack{j=1 \\ j\neq k}}^{J} E\left[\frac{G(\theta_j)}{G(\theta_k)}\right] \tag{37}$$

where $E[\gamma]$ is the intercellular interference factor. Therefore, the outage probability will be given as

$$P_o = Q\left(\frac{\frac{1}{\sigma_{req}} - \frac{M+1}{G_P}(1+E[\gamma]) - \frac{\eta}{S}}{\frac{M}{G_P}\sqrt{\mathrm{var}[\frac{I_{out}}{S}]}}\right) \tag{38}$$

The system with seven cells (desired one and six interfering using 8-element array system), which is represented in Fig. 17, is used to simulate the down-link of the cellular system in this section. The results of this simulation model are tabulated in Table 1.

Now, assuming M>>1, and using the simulation results in Table 1 we can expressed the outage probability using CMA(1,2) (conventional algorithm) for 8-element array system as follows

$$p_o = Q\left(\frac{\frac{1}{\sigma_{req}} - \frac{M\,7.7026}{G_P} - \frac{\eta}{s}}{\frac{M\,0.4520}{G_P}}\right) \tag{39}$$

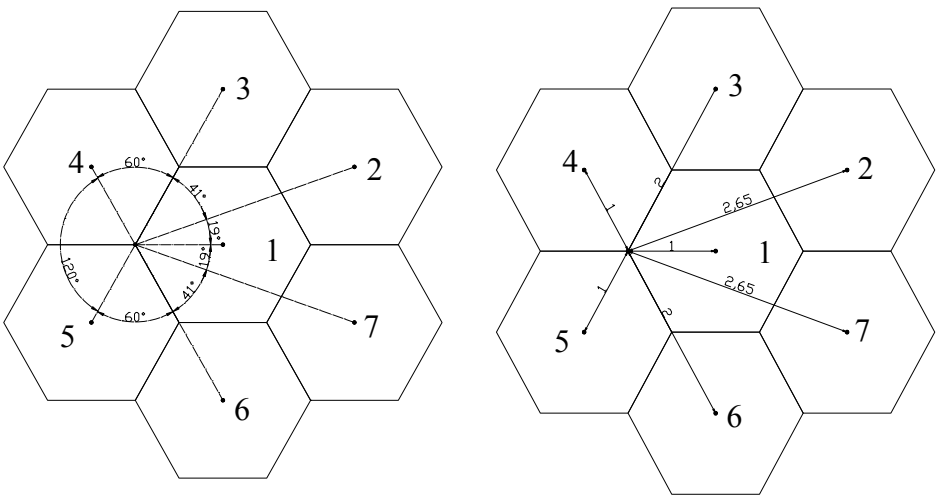



Fig. 17. Down-link cellular system with seven cells

| Technique | Item | Cell number (j) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| CMA(1,2) | $G(\theta_j)$ | 1.8304 | 1.5027 | 2.3538 | 2.2899 | 2.3533 | 2.3457 | 1.4232 |
| | $G(\theta_j)/G(\theta_1)$ | 1.0000 | 0.8209 | 1.2859 | 1.2510 | 1.2857 | 1.2815 | 0.7775 |
| Hybrid Technique | $G(\theta_j)$ | 9.8264 | 1.5895 | 1.0812 | 0.9990 | 1.0812 | 1.0098 | 1.1439 |
| | $G(\theta_j)/G(\theta_1)$ | 1.0000 | 0.1618 | 0.1100 | 0.1017 | 0.1100 | 0.1028 | 0.1164 |

Table 1. Down-link beam gains at different directions of arrival

while that of the proposed technique is given by

$$p_o = Q\left(\frac{\frac{1}{\sigma_{req}} - \frac{M1.7027}{G_P} - \frac{\eta}{s}}{\frac{M0.2214}{G_P}}\right) \tag{40}$$

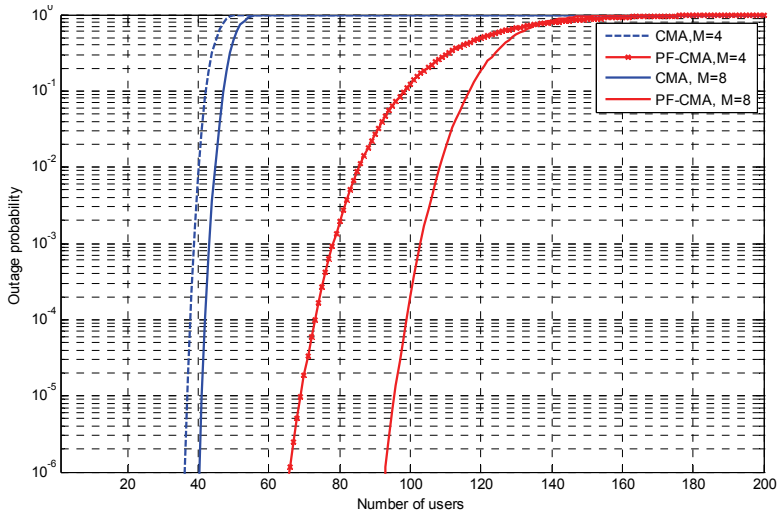
The results of Equation (39) and (40) are displayed in Fig. 18 for SNR = 0 dB and 20 dB. From this figure, it is clear that the hybrid technique decreases the outage probability of the system. This decrease will help in increasing the number of users that can be accommodated in the system (System Capacity). As an example, assuming that the acceptable outage probability is 1% and SNR = 0 dB, the system capacity will be 40 and 86 users for the conventional and hybrid system respectively, in case of 4-element array, and 45 and 108 users in case of 8-element array. Therefore, doubling the array size improves the system capacity by 11% and 26% for the conventional and hybrid systems respectively. In case of SNR = 20dB, the system capacity will be 45 and 95 users for the conventional and hybrid system respectively, in case of 4-element array, and 50 and 120 users in case of 8 elements array. Therefore, doubling the array size, again improves the system capacity by 11% and 26% for the conventional and hybrid systems respectively.

#### 4.5.2 The outage probability of the up-link system

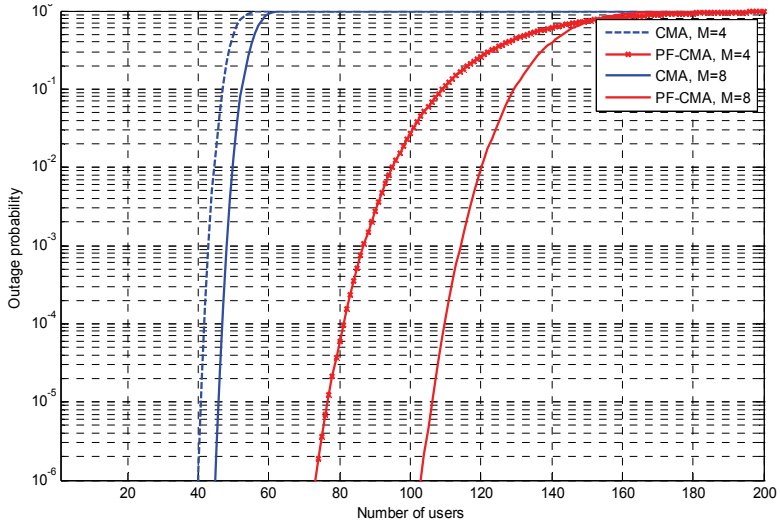
To check up-link performance improvement using the hybrid technique, we simulate the uplink cellular system shown in Fig. 19 using 6 interferer sources distributed uniformly around a central base station with a single desired user communicates with this station. We assumed that the interference sources located outside of the central base station cell. For simplicity, we assumed the distance of each interferer from the base station is double that of the desired user. The results of this simulation model are tabulated in Table 2.

Using Equation (38) and Table 2, the outage probability of the up-link cellular system using 16-element array for the conventional technique given by

$$p_o = Q\left(\frac{\frac{1}{\sigma_{req}} - \frac{M6.9501}{G_P} - \frac{\eta}{s}}{\frac{M0.0303}{G_P}}\right) \tag{41}$$




SNR= 0dB




(b) SNR= 20dB

Fig. 18. The outage probability of cellular down link system with 4- and 8-element array while that of 16-element array for the propose hybrid technique is

$$p_o = Q\left(\frac{\frac{1}{\sigma_{req}} - \frac{M3.4715}{G_P} - \frac{\eta}{s}}{\frac{M0.1202}{G_P}}\right) \tag{42}$$

| Technique | Item | Cell number (j) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| CMA(1,2) | $G(\theta_j)$ | 2.4981 | 2.4392 | 2.4998 | 2.4981 | 2.4967 | 2.4998 | 2.4394 |
| | $G(\theta_j)/G(\theta_1)$ | 1.0000 | 0.9764 | 1.0007 | 1.0000 | 0.9995 | 1.0007 | 0.9765 |
| Hybrid Technique | $G(\theta_j)$ | 10.000 | 0.9167 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.2663 |
| | $G(\theta_j)/G(\theta_1)$ | 1.0000 | 0.0917 | 0.1000 | 0.1000 | 0.1000 | 0.1000 | 0.1266 |

Table 2. Up-link beam gains at different directions of arrival

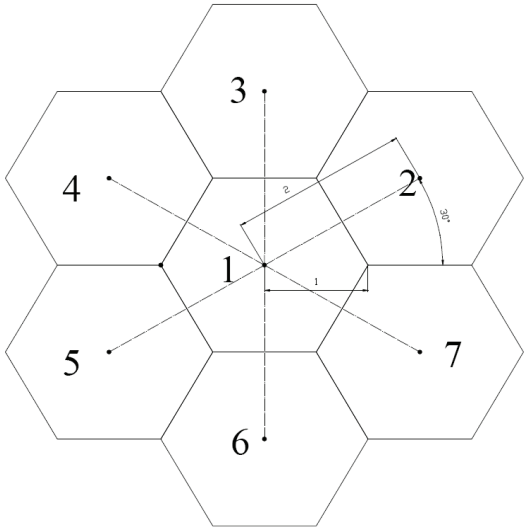
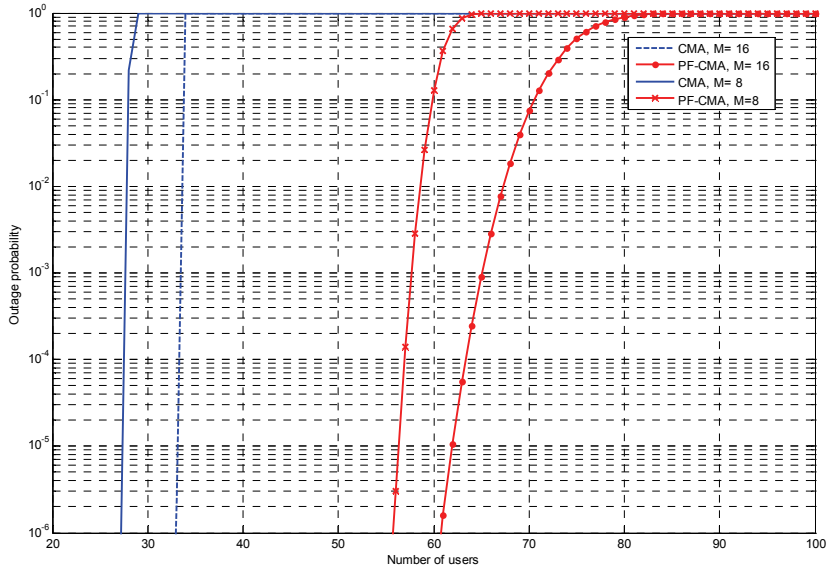
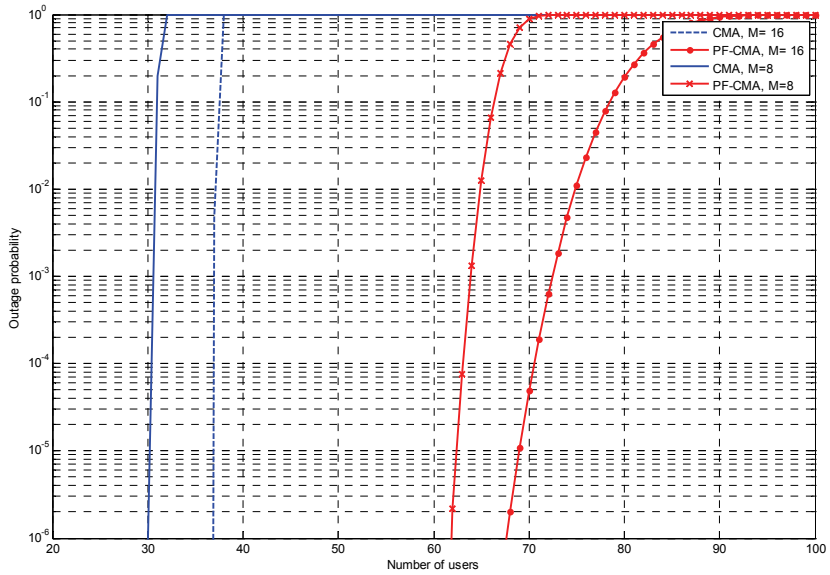



Fig. 19. Up-link cellular system with seven cells

The results are as shown in Fig.20 It is clear that the hybrid technique decreases the outage probability of the system. This decrease will help in increasing the number of users that can be accommodated in the system (System Capacity). As an example, assuming that the acceptable outage probability is 1% and SNR = 0dB, the system capacity will be 28 and 59 users for the conventional and hybrid system respectively, in case of 8-element array, and 34 and 67 users in case of 16-element array. Therefore, doubling the array size improves the system capacity by 21% and 14% for the conventional and hybrid systems respectively. In case of SNR = 20dB, the system capacity will be 31 and 65 users for the conventional and hybrid system respectively, in case of 8 elements array, and 37 and 75 users in case of 16 elements array. Therefore, doubling the array size, improves the system capacity by 19% and 15% for the conventional and hybrid systems respectively. Here we can observe that the capacity improvement percentage of the conventional algorithm is slightly greater than the improvement percentage of the overall hybrid technique, but in all of above cases the hybrid system capacity is greater.




(a) SNR=0 dB




(b) SNR=20dB

Fig. 20. Outage probability of up-link system with 8- and 16- element array

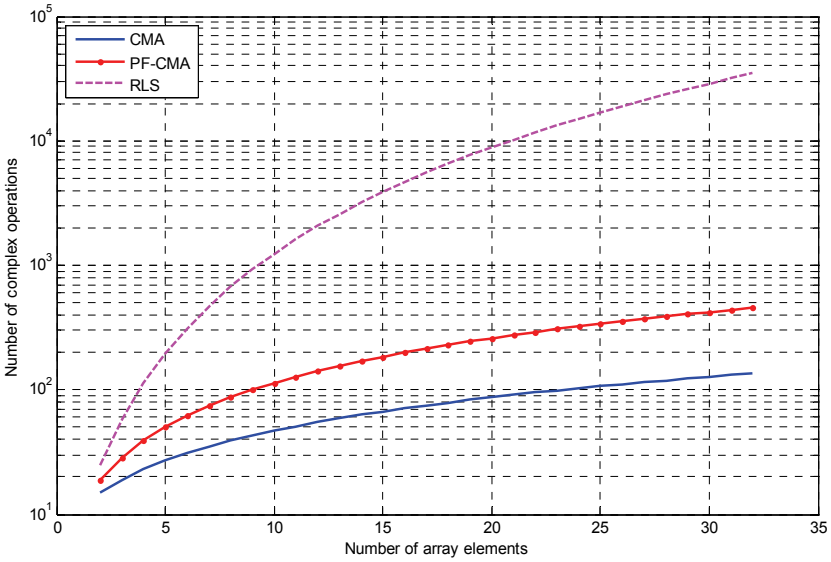



Fig. 21. Computation complexity for the hybrid algorithm and some of the conventional types of adaptive beamforming algorithms

### 4.6 Convergence speed of the proposed algorithm

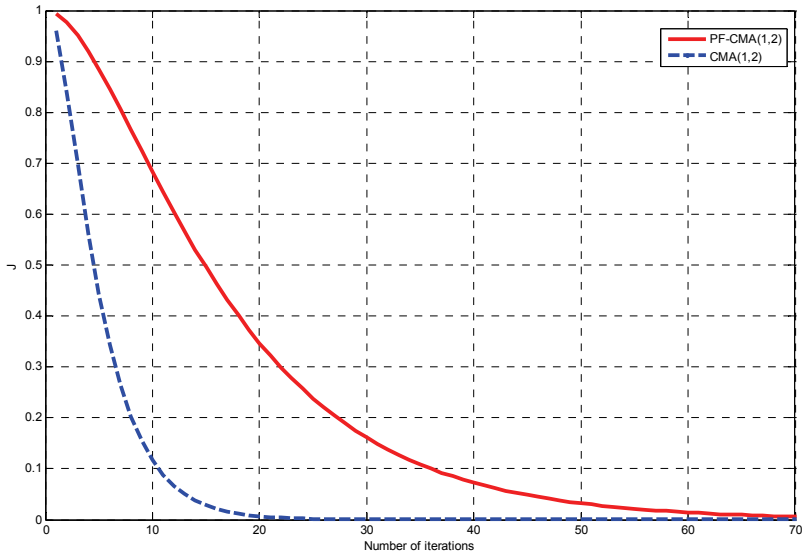
Here, a comparison between the convergence speed of the conventional adaptive beamforming algorithms and the proposed hybrid technique is carried out. This speed is determined by measuring the error behavior of the algorithms versus the used samples in the training period i.e. measuring the value of the cost function (the mean square error) at each sample time. The cost functions of LMS and CMA (1, 2) algorithms are respectively as follows:

$$J_{LMS}(k) = E\left[ \left| d(k) - y(k) \right|^2 \right] \tag{43}$$

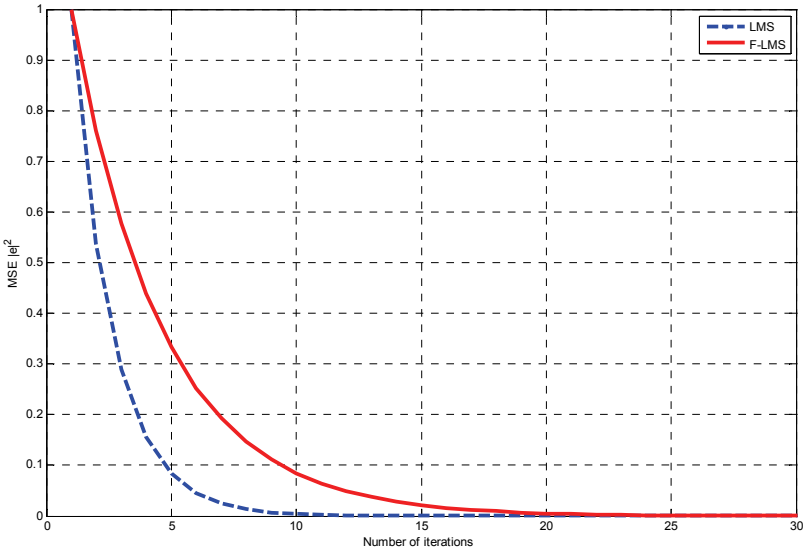
and

$$J_{CM(1,2)}(k) = E\left[ \left| \left| y(k) \right| - 1 \right|^2 \right] \tag{44}$$

where $d(k)$ and $y(k)$ are the desired and output signal samples respectively. Using the same environment used for the beam pattern performance, we plotted the result of error curves in Fig. 22 by taking into account only the first significant iterations (samples) of both algorithms. From the figure it is clear that the convergence speed of the conventional algorithms is slightly faster than the hybrid technique, but this difference is very small.




(a) Hybrid technique and CMA algorithm with $\mu = 0.001$



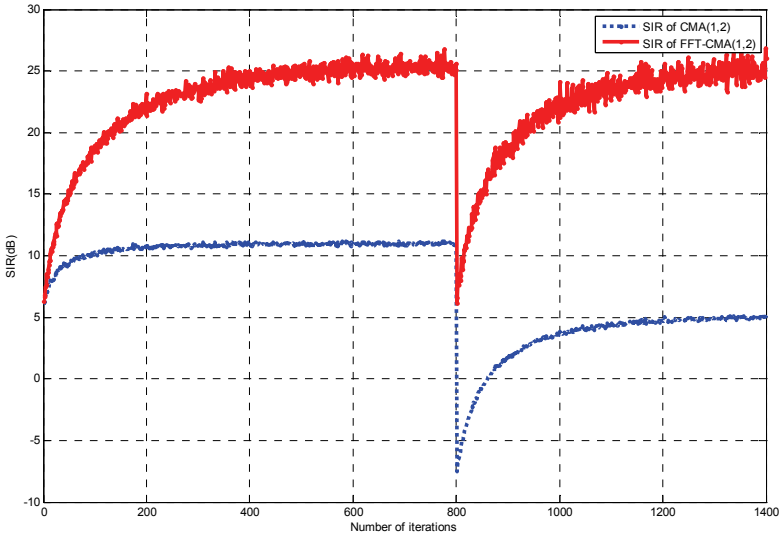

(b) LMS and hybrid LMS algorithm with $\mu = 0.001$

Fig. 22. Convergence of the hybrid and conventional algorithms

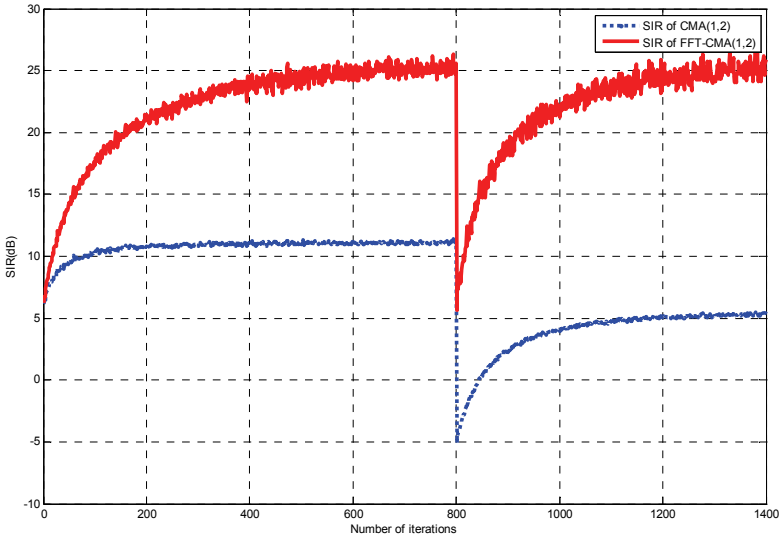
### 4.7 Tracking performances

Although the computational cost of the hybrid technique was slightly higher comparing with computational cost of the conventional technique, the simulation results have shown

that the hybrid technique is more capable of tracking the targets with varied directions of arrival. The results in Fig. 23 show that the hybrid technique can track the signal arriving from the desirable source, which has the DOA of 60˚in the initial 800 iterations; the DOA of the desired signal is changed to 30˚, where the interfering DOA remain unchanged at 0˚ in two tracking periods. We can observe that the hybrid technique is more effective in tracking




(a) Tracking from 60˚to 30˚




(b) Tracking from–20˚ to 60˚

Fig. 23. Tracking capabilities of the hybrid and conventional techniques References

the desired target. Also in case of the desired target is flipped to the other side of the interfering source (from –20˚ to 60˚) the hybrid tracking capability is still better than the tracking capability of the conventional technique. In fact the reason of this good tracking capability is the hybrid technique use of the initial reconstructed input signal which has less interference.

## 5. Acknowledgment

The authors express their gratitude to Professor Otman Basir, University of Waterloo, for reviewing and editing this chapter, and for his valuable remarks.

## 6. References

[1] K. Shetty, "A novel Algorithm for Up-link Interference Suppression Using Smart Antennas in Mobile Communications," A thesis submitted to the Department of Electrical and Computer Engineering, Florida State University in partial fulfillment of the requirements for the degree in Master of Science, Spring 2004.

[2] J. Litva, T. k. Lo, "Digital Beamforming in Wireless Communications," 1996.

[3] O. Abu-Ella and B. El-Jabu, "Capacity Improvement of Blind Adaptive Beamforming Algorithms Using Pre-filtering Technique," in press in IET Microwave, Antenna & Propagation Journal.

[4] Paul A. Wintz, "Transform Picture Coding," Proceedings of the IEEE, vol. 60, No. 7, July 1972.

[5] Y. Chwn, T. Le-Ngoc, B. Champagne, C. Xu, "Recursive Least Squares Constant Modulus Algorithm for Blind Adaptive Array," IEEE Transaction on Signal Processing. Vol. 52. No. 5. May 2004.

[6] A. S. Sawant, D. K. Anvekar, "Capacity Improvement in CDMA and FDMA Cellular Mobile Communication Systems Using Adaptive Antenna," 0-7803-4912-1/99, 1999 IEEE.

[7] B. El-Jabu, "Cellular Communications using Aerial Platforms," A doctoral thesis submitted in partial fulfillment of the requirements for the award of Doctor of philosophy of the university of Southampton, September 1999.

# 23

# Improved Cloud Detection Technique at South China Sea

Ng Hou Guan, Mohd.Zubir Mat Jafri and Khiruddin Abdullah
*Universiti Sains Malaysia*
*Malaysia*

## 1. Introduction

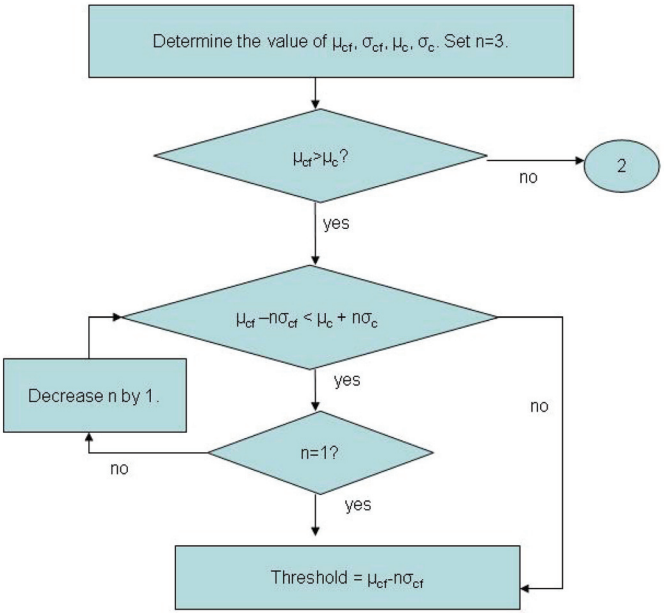
The sea surface temperature (SST) algorithm was only valid for cloud free water pixels. The cloudy pixels should be separated before the SST algorithm could be applied. The cloud masking algorithm was used to separate the cloudy pixels from non-cloudy pixels. The cloud surface, ocean surface and vegetated, arid or snow covered land surfaces have different response to reflectance, brightness temperature and emissivity. The cloud detection or masking tests were based on the different response patterns of the earth surfaces or clouds to the reflection or emission of the wave radiation. The threshold values were different for the different seasonal and regional areas. Therefore the threshold values for each test would be determined before cloud masking test were performed.

Krieble(1989) had proposed a procedure to derive suitable temperature thresholds for new areas of application. The land and sea areas which seen likely to be the coldest but cloud free were identified visually by users. However this method is subjective and quite time consuming. The results were varying with the users. Sauders (1986) had determined the threshold for local uniformity test with SD value less than 0.2 K for cloud free pixels over the sea in Northeastern Europe. France and Cracknell (1994, 1995) found SD values less than 0.4 K for cloud free pixels over the sea in northeastern Brazil.

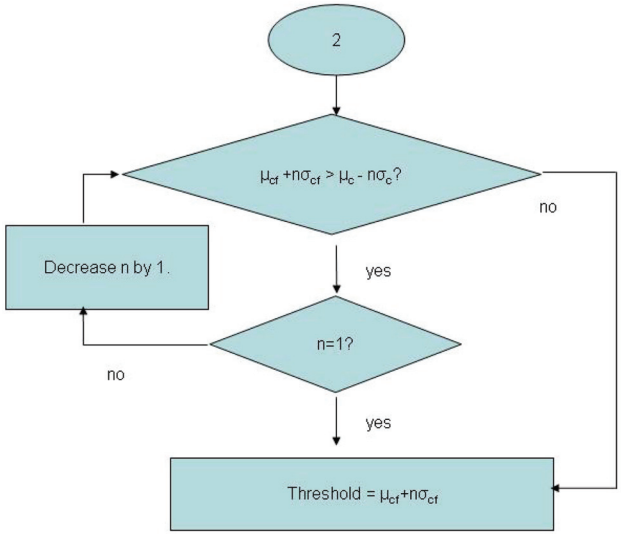
In this study, histograms of the cloud over land, cloud free land, cloud over sea, cloud free sea areas would be utilized. It was different with the suggested method that utilized whole sea area. The histogram was expected to be bimodal, a clear separation between the digital number for the colder clouds and the warmer sea surface (Cracknell, 1997). However, in practical, it was difficult to get the clear bimodal histogram for whole sea area. Therefore the new method generated the four histograms by using the ROI tool of software Envi V.4.4 to select the four separate areas.

## 2. Methodology

The images of calibrated reflectance for channel 1 and 2 were created, and the brightness temperature for channel 4 and 5 were calibrated. After that, the function of Band Math in software ENVI was used to create the image of $R_2/R_1$. The ROI (Region of Interest) tool in ENVI was used to choose the region of cloud, sea and land. The selected regions were then applied to the image of brightness temperature for channel 5 and channel 4.




(a) The case when $\mu_{cf} > \mu_c$




(b) The case when $\mu_{cf} < \mu_c$.

Fig. 1. The value of threshold was determined from the mean and standard deviation of cloud free and cloudy water pixels.

To determine the threshold for the cloud masking techniques, the mean, μ and standard deviation, σ for the cloud free water pixels ($\mu_{cf}$, $\sigma_{cf}$) and cloudy water pixels ($\mu_c$,$\sigma_c$) were determined. The value of n was set to three. Then the mean of the cloudy water pixels and

cloud free water pixels were compared. If the mean of cloud free water pixels greater than cloudy water pixels, then we compared the values of $\mu_{cf}$-n$\sigma_{cf}$ and $\mu_c$+n$\sigma_c$. The value of threshold was assigned the value of $\mu_{cf}$-n$\sigma_{cf}$ if $\mu_{cf}$-n$\sigma_{cf}$> $\mu_c$+n$\sigma_c$ or n=1, otherwise the value of n was decreased by one until $\mu_{cf}$-n$\sigma_{cf}$> $\mu_c$+n$\sigma_c$ or n=1 (Figure 1(a)). However, if the mean of cloud free water pixels less than the mean of cloudy water pixels, then the values of $\mu_{cf}$+n$\sigma_{cf}$ and $\mu_c$-n$\sigma_c$ were compared. The value of threshold was assigned the value of $\mu_{cf}$+n$\sigma_{cf}$ if $\mu_{cf}$+n$\sigma_{cf}$< $\mu_c$-n$\sigma_c$ or n=1, otherwise the value of n was decreased by one until $\mu_{cf}$+n$\sigma_{cf}$< $\mu_c$-n$\sigma_c$ or n=1 (Figure 1(b)).

The concept of determining the threshold for separating the cloudy from non-cloudy water was based on the six sigma techniques. There were 99.9996% of data lie between $\mu$-3$\sigma$ to $\mu$+3$\sigma$ and 99.38% data lie between $\mu$-2$\sigma$ to $\mu$-2$\sigma$. Therefore, if more than 99% of the data for cloudy and cloudy pixels were not intersected, the value of $\mu\pm n\sigma$ could be selected as the threshold value.

## 3. Result and discussion

### (i) Test: Gross cloud check

Image: Channel 5 brightness temperature

A histogram of channel 5 brightness temperature was generated. The brightness temperature for cloudy pixels and brightness temperature for cloud free water pixels were significant different. A threshold value was determined to separate the cloudy pixels from the non-cloudy pixels.

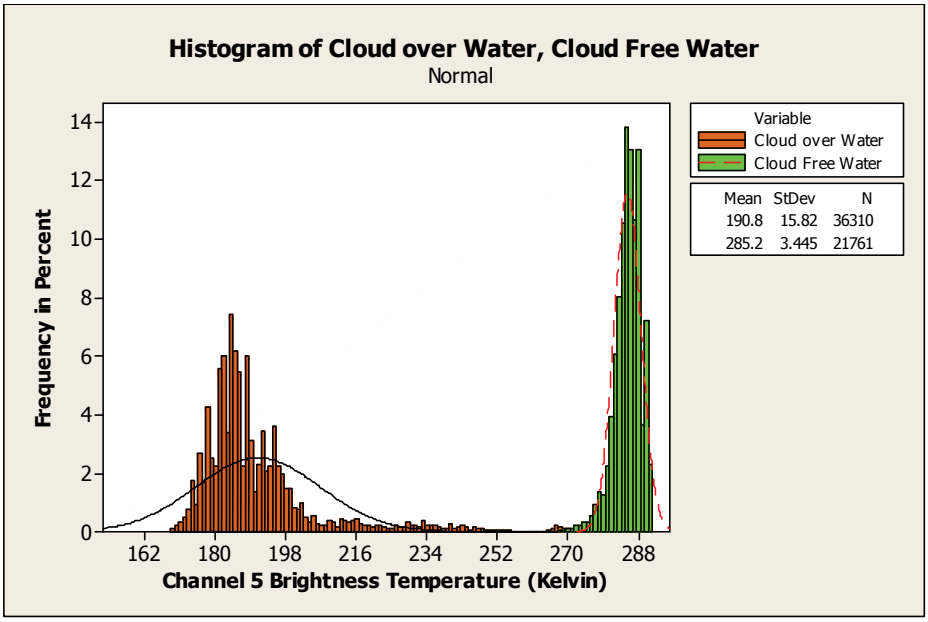



Fig. 2. The histogram of channel 5 brightness temperature for cloud over water and cloud free water pixel.

Weilbull Distribution was used as the fitted distribution in Figure 2. It was used instead of Normal Distribution because the data was not distributed normally. The Weilbull distribution was also more suitable on showing the peak value and shape of the histogram in this case. There is a significant difference between cloud free water pixels and cloudy water pixels from Figure 2. Therefore, the clear water pixels could be separated from cloudy pixels if a proper threshold value was selected. There is no significant different between clouds free water pixel and cloud free land pixels from Figure 3. This indicates that we could not discriminate between land and sea by using the image of brightness temperature. The Figure 3 also shows that the cloud tend to have the lower channel 5 brightness temperature compared to land and sea.

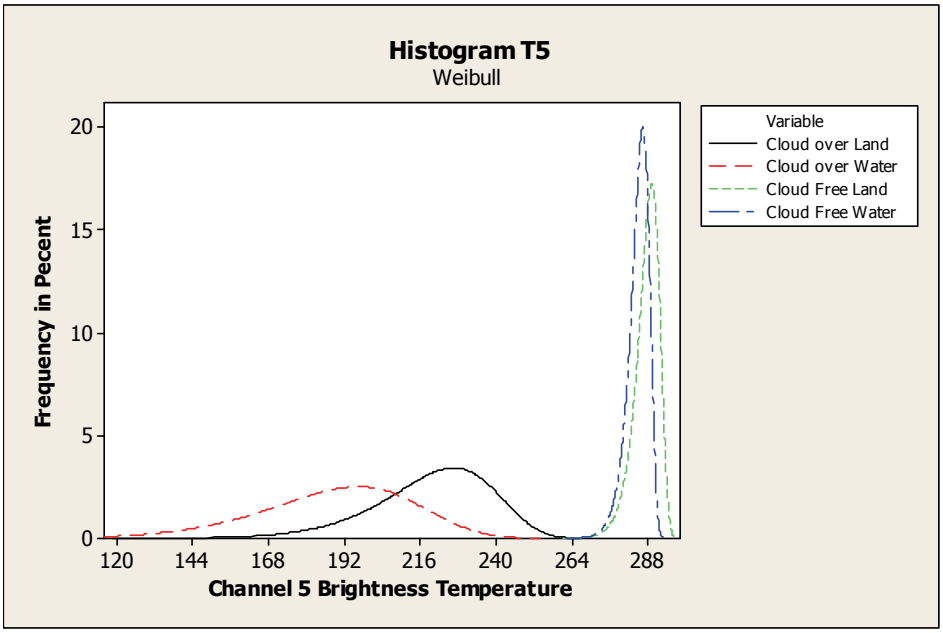



Fig. 3. The histogram of Channel 5 brightness temperature for cloud over land, cloud over water, cloud free land and cloud free water pixels. The Weilbull distribution was used as the fitted distribution for the histogram.

After that, a box-plot with median inter-quartile range box was generated to give an overview of the distribution of channel 5 brightness temperature for land, sea and cloud.
The value of threshold for separating the cloudy and cloud free water pixels was then determined by using the mean and standard deviation of these pixels. The methodology had been discussed in previous section.

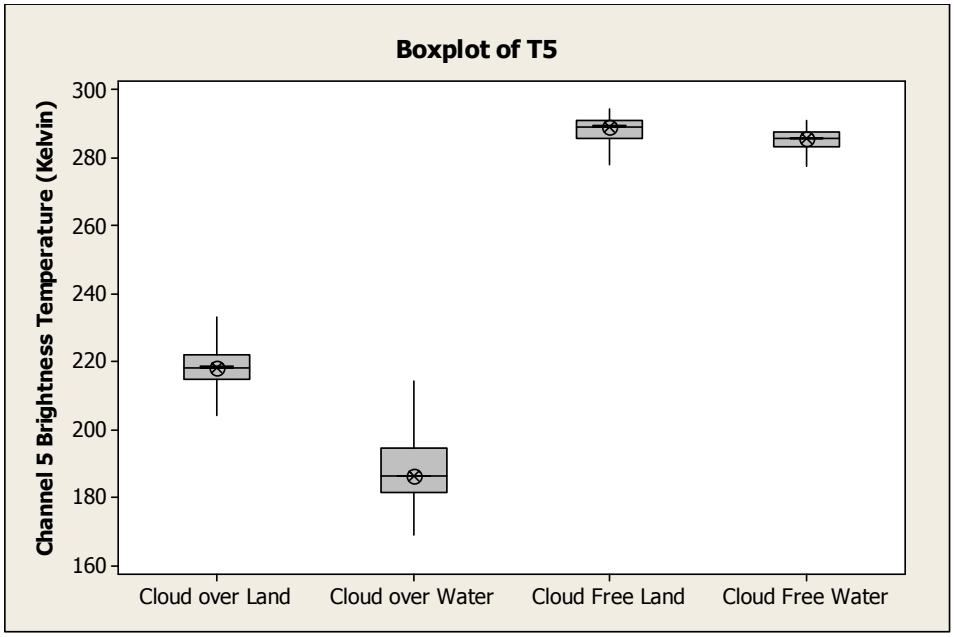



Fig. 4. The boxplot of channel 5 brightness temperature for cloud over land, cloud over water, cloud free land and cloud free water pixels.

| | Cloud Over Water | Cloud Free Water |
|---|---|---|
| **Mean** | 190.8 | 285.2 |
| **Std. Dev.** | 15.82 | 3.445 |

Table 1. The statistics of the channel 5 brightness temperature for cloudy and cloud free water pixels.

$\mu_c$=190.8, $\mu_{cf}$= 285.2, $\sigma_c$=15.82, $\sigma_{cf}$=3.445
$\mu_{cf} > \mu_c$ and $\mu_{cf} - 3\sigma_{cf} > \mu_c + 3\sigma_c$
Therefore, thereshold = 274.87 K
The pixels were masked as cloudy pixels if the channel 5 brightness temperature was less than 274.865K. The same procedure was repeated for the image of channel 4 brightness temperature.

**(ii) Test: Minimum channel 4 temperature**

Image: Channel 4 brightness temperature
The cloudy water pixels showed a wide variation in terms of $T_4$ compared with the cloud free water pixels. The cloudy pixels also had a lower value of $T_4$ compared with the cloud free sea pixels. Therefore if a pixel value less than a certain value of $T_4$, it can be masked as a cloudy pixel.

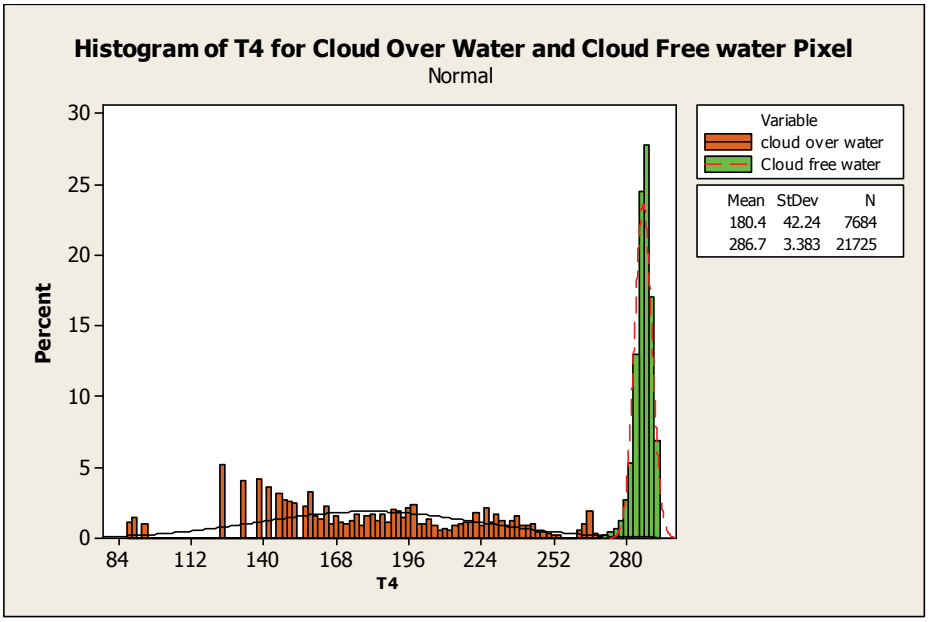



Fig. 5. The histogram of channel 4 brightness temperature for cloud over water and cloud free water pixel.

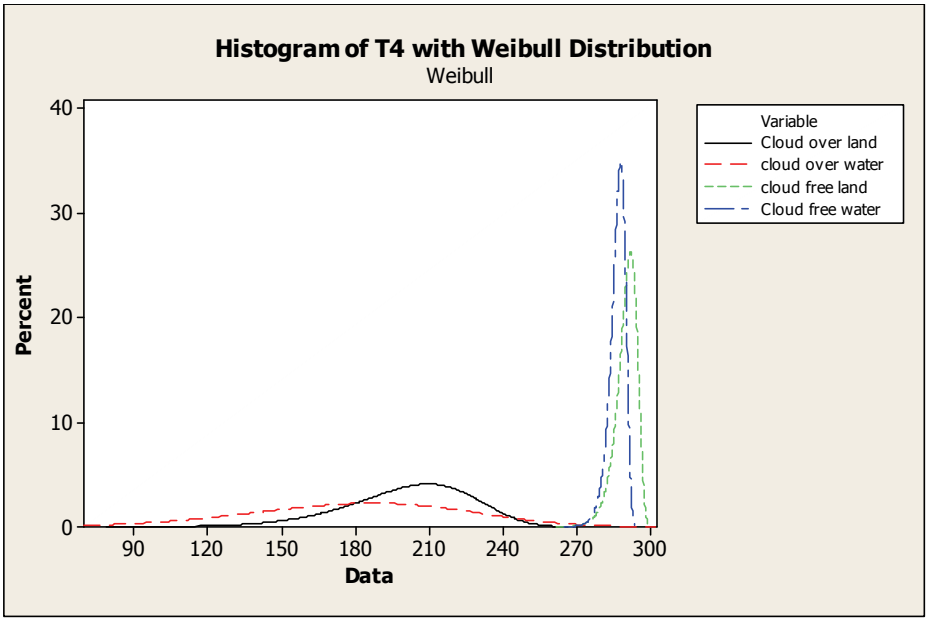



Fig. 6. The histogram of Channel 4 brightness temperature for cloud over land, cloud over water, cloud free land and cloud free water pixels. The Weilbull distribution was used as the fitted distribution for the histogram.

The cannel 4 brightness temperature of land and sea was not significant. The land and sea area could not be discriminated by using brightness temperature, but the land and sea areas cloud be discriminated from cloud by using $T_4$.
The median of $T_4$ for cloud over land and cloud over water were 199.5K and 177.7 K respectively. However, the median of $T_4$ for cloud free land and cloud free water were 291.4K and 287.0K respectively. The cloud was significantly colder than the land and sea.

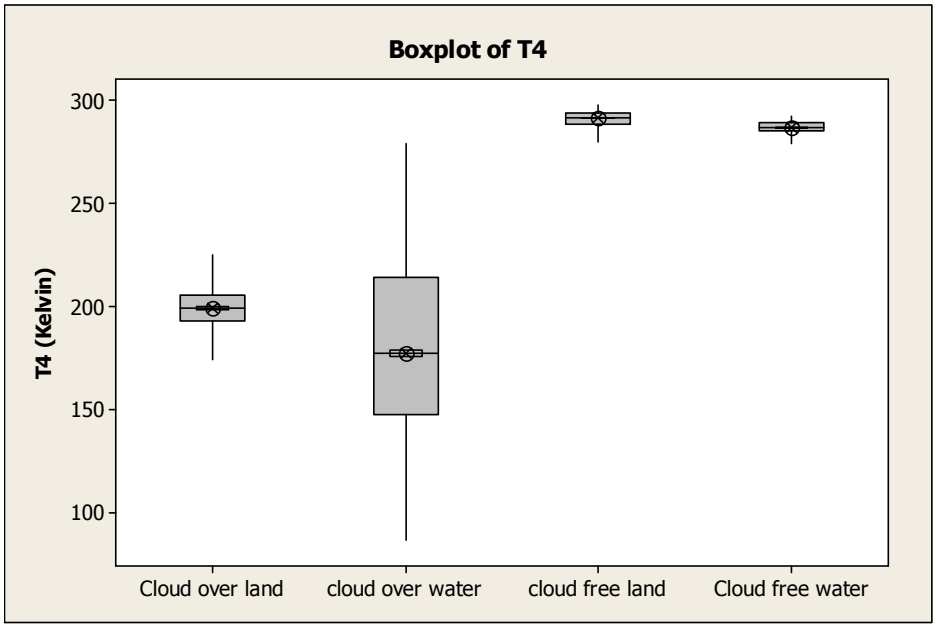



Fig. 7. The boxplot of channel 4 brightness temperature for cloud over land, cloud over water, cloud free land and cloud free water pixels.

| | Cloud Over Water | Cloud Free Water |
|---|---|---|
| **Mean** | 203.2 | 286.7 |
| **Std. Dev.** | 18.82 | 3.383 |

Table 2. The statistic of the channel 4 brightness temperature for cloudy and cloud free water pixels.

$\mu_c$=203.2, $\mu_{cf}$= 286.7, $\sigma_c$=18.82, $\sigma_{cf}$=3.383
$\mu_{cf} > \mu_c$ and $\mu_{cf} - 3\,\sigma_{cf} > \mu_c + 3\,\sigma_c$
Therefore, thereshold = 276.55

**(iii) Test: Dynamic Visible Threshold Test**
(a) Image: Channel 1 albedo/ reflectance
A histogram of channel 1 reflectance was generated. The reflectance of sea was around 0 to 18%, but the reflectance of cloud was around 36% to 100% (Figure 8).

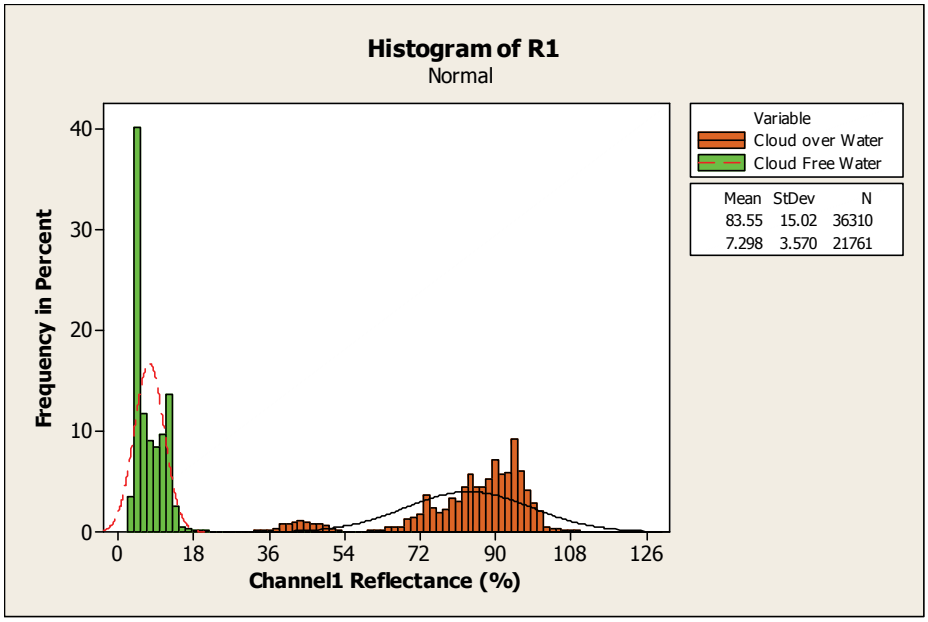



Fig. 8. The histogram of channel 1 reflectance for cloud over water and cloud free water pixel.

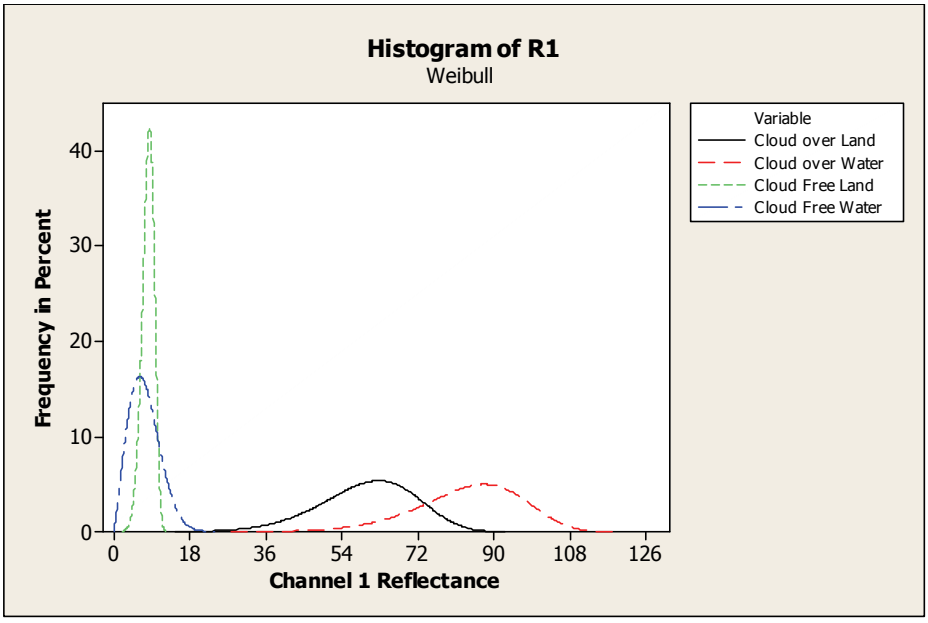



Fig. 9. The histogram of Channel 1 reflectance for cloud over land, cloud over water, cloud free land and cloud free water pixels. The Weilbull distribution was used as the fitted distribution for the histogram.

The reflectance of land and sea at channel 1 was lower than the cloud reflectance. Majority of the land and sea area had the reflectance lower than 18%. However, majority of the cloudy area had the reflectance greater than 36% (Figure 9).

There was a clear discrimination between cloudy and cloud free water pixels, but no significance difference between cloud free water and land pixels (Figure 10). Therefore the cloud free water pixels could be separated from cloudy pixels but the cloud free water pixels cannot be separated from the cloud free land pixels from channel 1 reflectance.

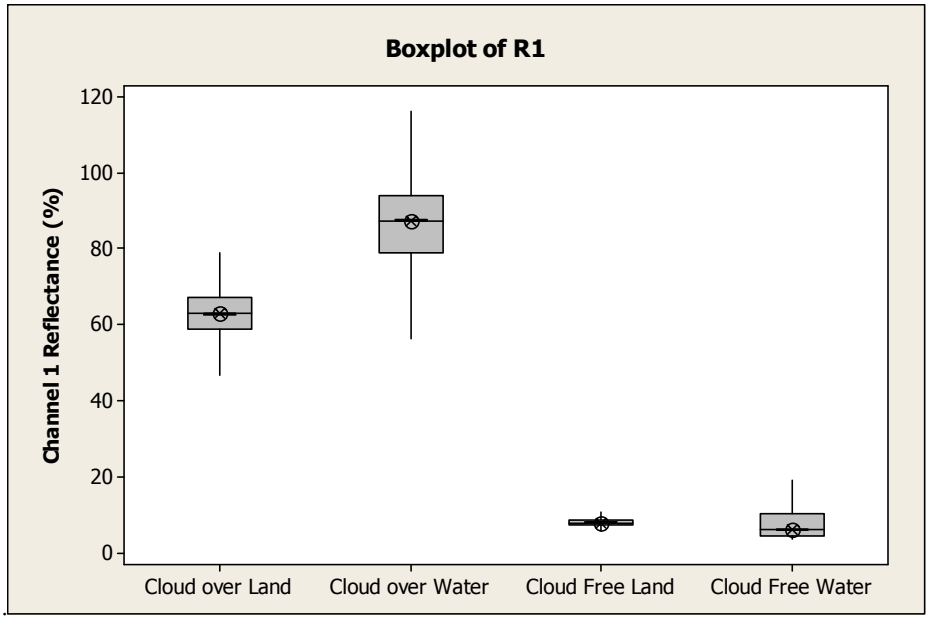



Fig. 10. The boxplot of channel 1 reflectance for cloud over land, cloud over water, cloud free land and cloud free water pixels.

| | Cloud Over Water | Cloud Over Land | Cloud Free Land | Cloud Free Water |
|---|---|---|---|---|
| **Mean** | 60.359 | 83.778 | 8.137 | 7.426 |
| **Std. Dev.** | 12.161 | 15.024 | 1.133 | 3.568 |

Table 3. The statistic of the channel 1 reflectance for difference surface cover.

$\mu_c$=60.359, $\mu_{cf}$=7.426, $\sigma_c$=12.161, $\sigma_{cf}$=3.568

$\mu_{cf} < \mu_c$ and $\mu_{cf}+3\,\sigma_{cf} < \mu_c - 3\,\sigma_c$

thereshold = 18.13

Therefore, the pixel was masked as cloudy water pixels if the reflectance was greater than 18.13%.

(b) Image: Channel 2 Albedo

There was a clear discrimination between cloudy and cloud free water pixel (Figure 11). The mean of channel 2 reflectance for cloudy water pixels is significant higher than the mean of the channel 2 reflectance for cloud free water surface. The higher value of standard deviation of cloud over water pixels was due to the inhomogeneous of cloud surface.

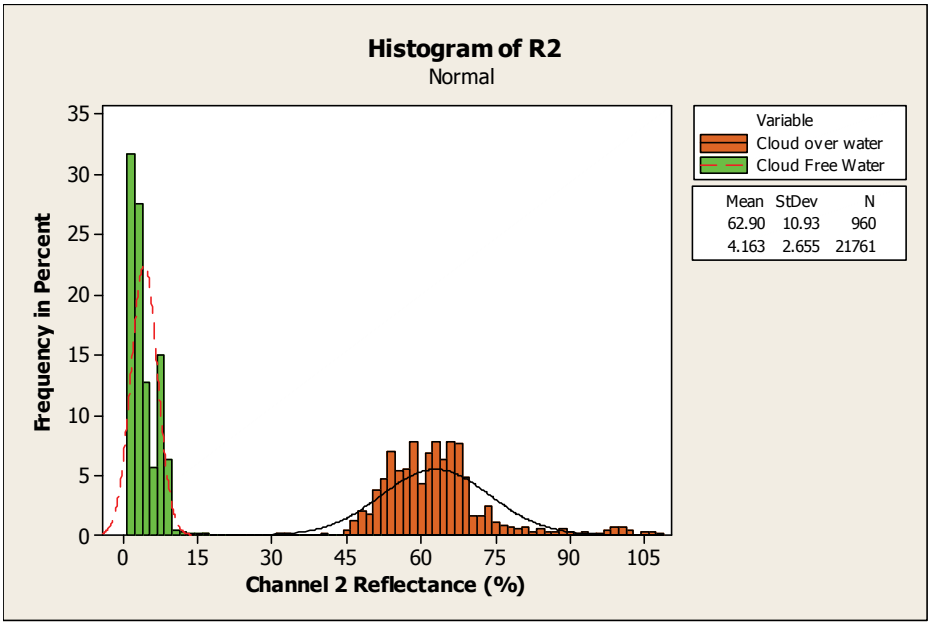



Fig. 11. The histogram of channel 2 reflectance for cloud over water and cloud free water pixel.

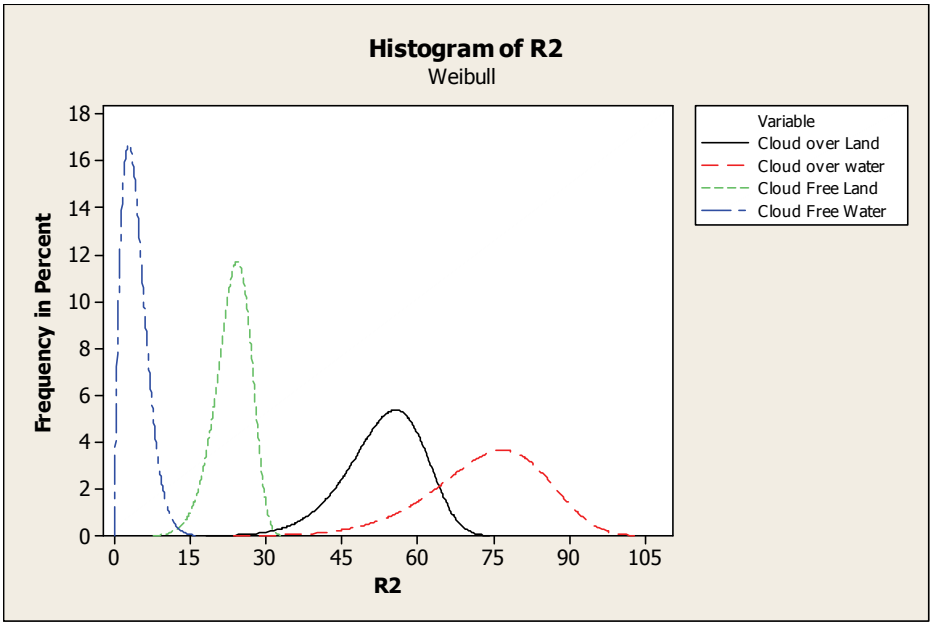



Fig. 12. The histogram of Channel 2 reflectance for cloud over land, cloud over water, cloud free land and cloud free water pixels. The Weilbull distribution was used as the fitted distribution for the histogram.

There was a clear separation between the mean of cloud free land, sea and cloudy surface (Figure 12). Therefore the land, sea and cloud could be separated if a proper value of threshold was selected.
The channel 2 albedo of the cloud over sea, cloud over water, cloud free land and cloud free water pixels was significant different among each other. 75% or majority of cloud over land pixel had the channel 2 albedo between 52.1 % and 58.4%, and the cloud over sea pixels was between 67.8 %and 82.7%. However the albedo for cloud free water pixels and was between 23.4%and 24.5%, and cloud free land pixels was between 2.2% and 5.9% (Figure 3.13).

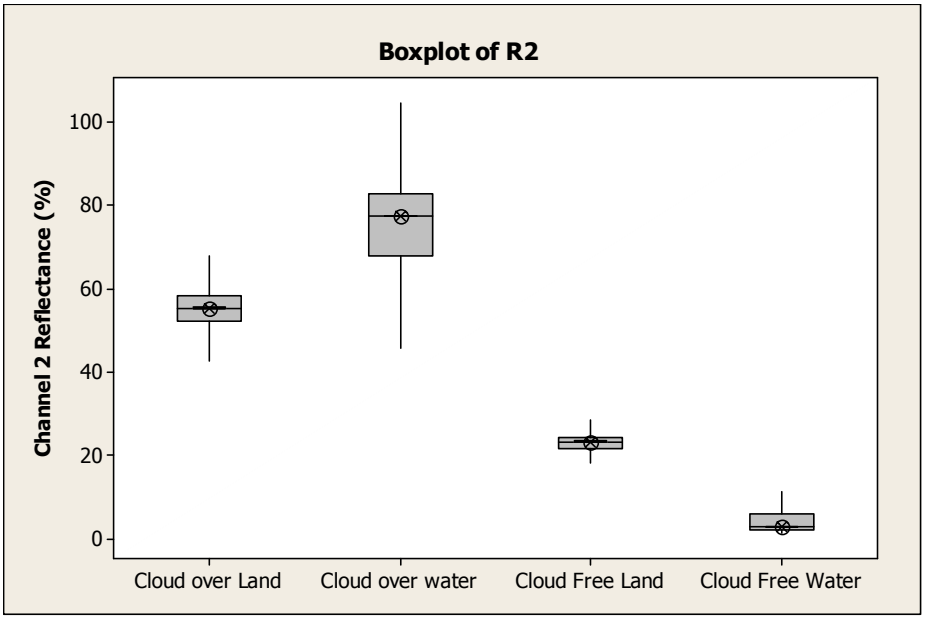



Fig. 13. The box-plot of channel 2 reflectance for cloud over land, cloud over water, cloud free land and cloud free water pixels.

| | Cloud Over Land | Cloud Over Water | Cloud Free Land | Cloud Free Water |
|---|---|---|---|---|
| Mean | 62.9 | 83.778 | 8.137 | 4.163 |
| Std. Dev. | 10.93 | 15.024 | 1.133 | 2.655 |

Table 4. The statistic of the channel 2 reflectance for difference surface cover.

Cloud free and cloudy water pixels:
$\mu_c$=60.358545, $\mu_{cf}$=7.426386, $\sigma_c$=12.161196, $\sigma_{cf}$=3.568147
$\mu_{cf} < \mu_c$ and $\mu_{cf}+3\,\sigma_{cf} < \mu_c - 3\,\sigma_c$
Therefore, thereshold = 12.228
Therefore, the pixel was masked as cloudy if the reflectance was greater than 12.228%.
Sea and land pixels:
$\mu_l$=8.137, $\mu_s$=4.163, $\sigma_l$=1.133, $\sigma_s$=2.655

**(iv) Test: Ratio of near infrared to visible reflectance test**
Image: Ratio of Channel 2 albedo and Channel 1 Albedo, $R_2/R_1$.

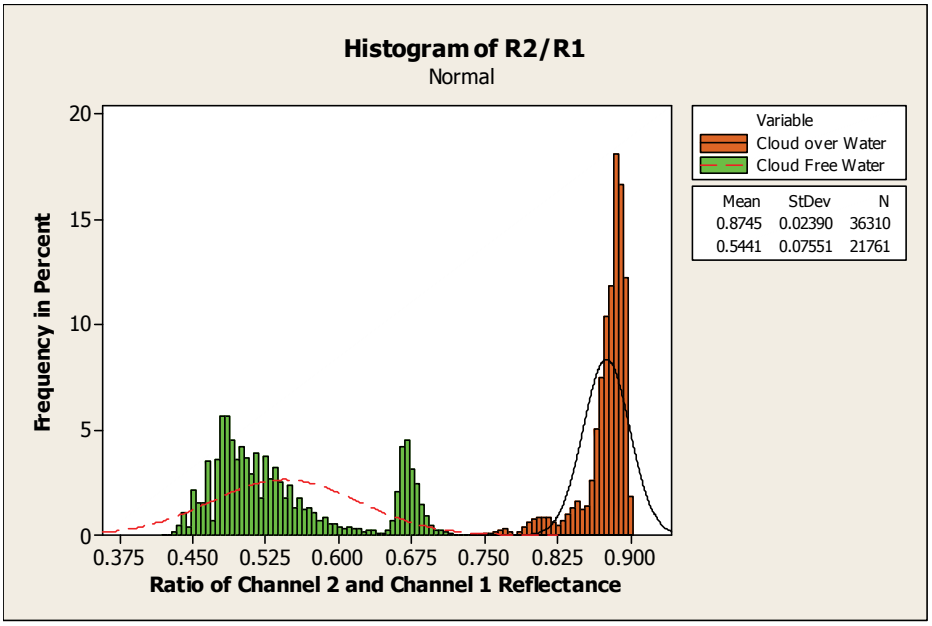



Fig. 14. The histogram of reflectance ratio for channel 2 and channel 1 for cloud over water and cloud free water pixel.

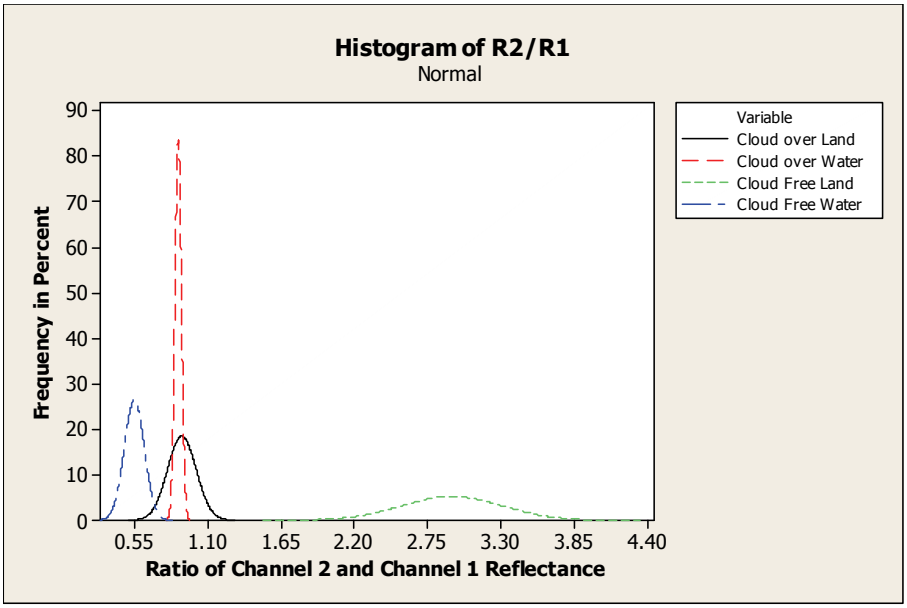



Fig. 15. The histogram of ratio reflectance for channel 2 and channel 1 for cloud over land, cloud over water, cloud free land and cloud free water pixels. The Weilbull distribution was used as the fitted distribution for the histogram.

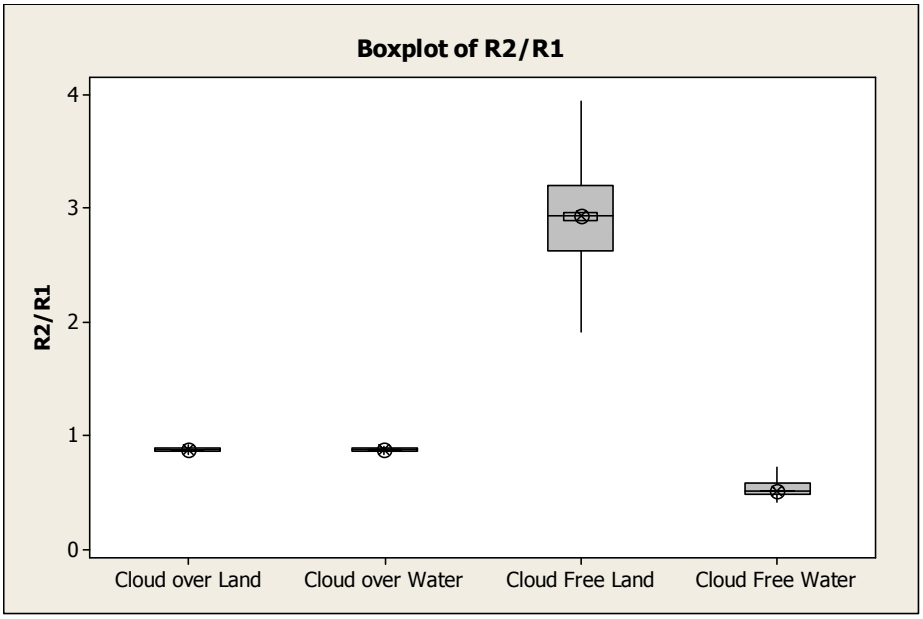



Fig. 16. The box plot of reflectance ratio for channel 2 and channel 1 for cloud over land, cloud over water, cloud free land and cloud free water pixels.

| | Cloud Over Land | Cloud Over Water | Cloud Free Land | Cloud Free Water |
|---|---|---|---|---|
| **Mean** | 0.9019 | 0.8745 | 2.931 | 0.5441 |
| **Std. Dev.** | 0.1079 | 0.0239 | 0.3802 | 0.0755 |

Table 5. The statistic of reflectance ratio for channel 2 and channel 1 for difference surface cover.

$\mu_c$=0.8745, $\mu_{cf}$=0.5441, $\sigma_c$=0.0239, $\sigma_{cf}$=0.0755
$\mu_{cf} < \mu_c$ and $\mu_{cf}+3\,\sigma_{cf} < \mu_c - 3\,\sigma_c$
Therefore, thereshold = 0.7706
The pixels were classified as cloud free water pixels if the ratio of reflectance was less than 0.7706.
Overall, the threshold values for all of the cloud masking tests were summarized as table below:

| Test | The threshold value for cloud masking |
|---|---|
| Gross Cloud Check | T5<274.87 K |
| Minimum Channel 4 Temperature | T4<276.55K |
| Dynamic Visible Threshold Test | R1>18,13%, R2>12.23% |

Table 6. The Threshold values for Cloud Masking Tests

**The cloud masking algorithm**
First of all, we had to determine whether the daytime algorithm or night time algorithm was used. We check the solar zenith angle and channel 2 albedo. The entire solar zenith angle for the image was below 56.61˚. Almost all of the pixels' reflectance was greater than 1%, and

only 0.0079% of the pixels' reflectance was less than 1%. Therefore the daytime algorithm was used.

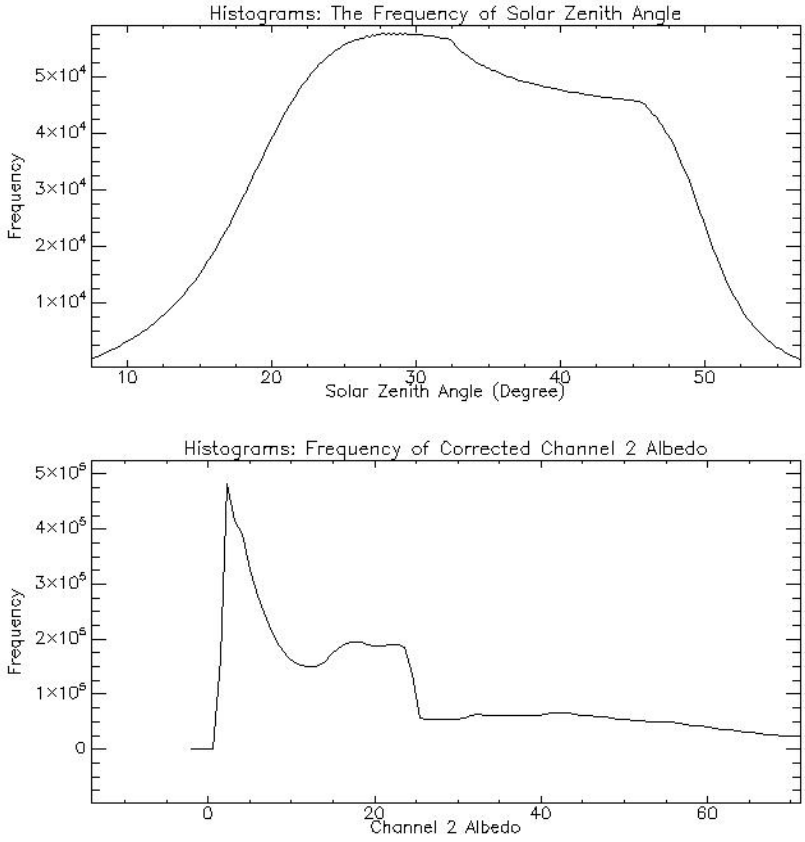



Fig. 17. The frequency distribution for the solar zenith angle and channel 2 albedo viewed with the image processing software.

**Daytime algorithm**

Step 0. If Satellite zenith angle<53˚, then go to step 1. Otherwise, reject or mask the pixel.

Step 1. If solar zenith angle<1˚, then mask the pixel, end.

Step 2. If $T_{B5}$<274.87 or $T_{B4}$<276.55K, then mask the pixel.

Step 3. For land, if corrected albedo channel 1, $R_{corr1}$>0.1813, mask the pixel ($R_{corr1}$= $R_1/\cos\theta_s$). For sea water, if corrected albedo channel 2, $R_{corr2}$>0.1223, then mask the pixel, end.

Step 4. If the vegetation index (ratio of channel 2 albedo and channel 1 albedo, $R_2/R_1$) >0.7706, then mask the pixel, end.

Step 5. Accept the pixel.

The image after geo-referenced and cloud masking was shown in the figure below. The cloud masking area was represented by the black colour (Figure 18).

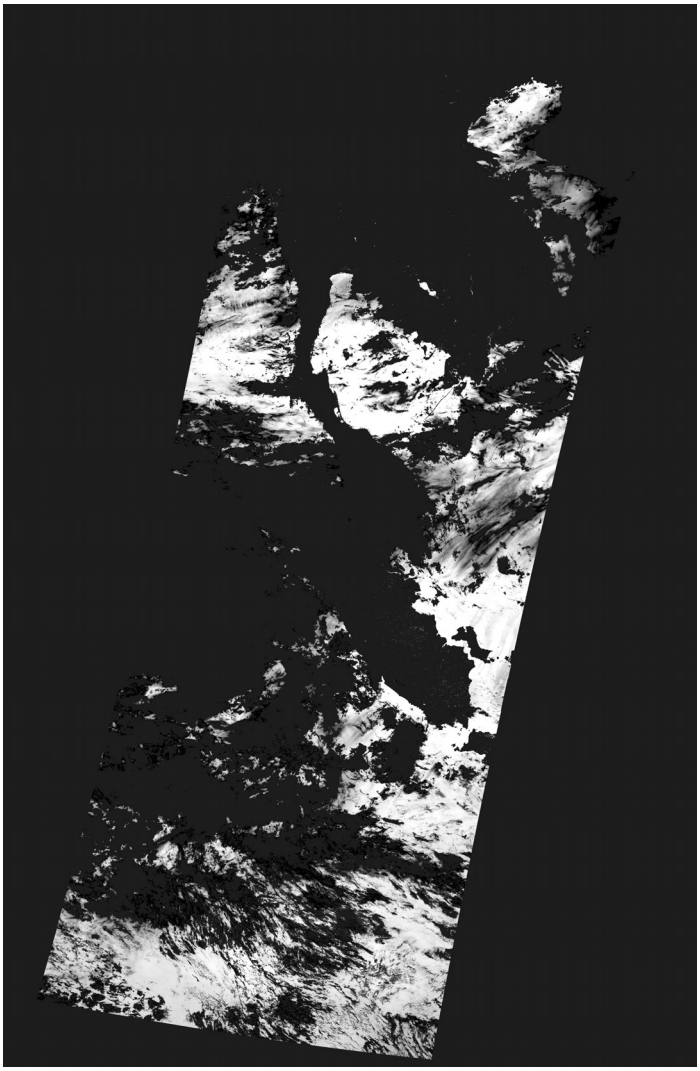


Fig. 18. The SST image after cloud masking.

## 4. Conclusion

Although the cloud masking tests suggested were not able to be used for cloud classification or did not provide the good quality of cloud detection, but it gives an easier and practical way to separate the cloudy pixels from clear water pixels. The albedo of visible channel (channel 1 and channel 2) and brightness temperature of thermal infrared channels were good enough to be used for filtering the cloudy pixels in the application of sea surface temperature calibration application. Besides of that, the study also provided the database for determining the thresholds values at the South China Sea.

## 5. References

Coakley; J.A. and Bretherton, F.P. (1982) Cloud Cover from high-resolution scanner data;detecting and allowing for partially filled of view. J*ournal of Geophysical Research*, 87, 4917-4932.

Cracknell, A.P.(1997). *The Advance Very High Resolution Radiometer*, Taylor & Francis, London.

Franca, G.B. and Cracknell, A.P. (1994) Retrieval of Land and Sea Surface Temperature using NOAA-11 AVHRR data in northeastern, Brazil. *International Journal of Remote Sensing*, 15, 1695-1712.

Franca, G.B. and Cracknell, A.P. (1995) A simple cloud masking approach using NOAA AVHRR daytime time data for tropical areas. *International Journal of Remote Sensing*, 16, 1697-1705.

G.D'Souza et al.(eds.).(1996) Advances *in the Use of AVHRR Data for Land Applications*, 195-210, Kluwer Acameic Pubohers, The Netherlands.

Kriebel,K.T., Saunders,R.W., Gesell, G. (1989)Optical Properties of Clouds Derived from Fully-Cloudy AVHRR Pixels. *Beitr. Phys. Atmosph.*,62, 165-171.

Saunders, R.W. (1986) An automated scheme for the removal of cloud contamination from AVHRR radiances over western Europe. *International Journal of Remote Sensing*,7 867-886.

# 24

# MEMS Tunable Resonant Leaky-Mode Filters for Multispectral Imaging Applications

Robert Magnusson and Mehrdad Shokooh-Saremi
*University of Texas at Arlington, Department of Electrical Engineering*
*USA*

## 1. Introduction

Multispectral imaging refers to a combination of spectroscopy and photography. By using rapidly tunable filters and two-dimensional (2D) image planes such as those provided by charge-coupled device (CCD) detectors, data sets containing spatial (x, y) and spectral information are acquired. The resulting spectral image cubes contain intensity and wavelength (λ) data at each pixel in the 2D image (Gat, 2000). Under time-varying conditions, the data cube would be multidimensional in (x, y, λ, t) space. Hyperspectral imaging is a similar concept principally differentiated from multispectral imaging in that many more wavelengths and narrower spectral passbands are employed. Thus, in multispectral imaging, relatively few wavelengths are used to carry the spatial information, whereas in hyperspectral imaging, the number of wavelength channels may reach ~100 (Vo-Dinh et al., 2004). Each of these methods is connected with a plethora of useful applications. Examples include spatio-spectral diagnostics in agricultural crop management, true-color night vision, forensics, and archaeology and art (Gat, 2000). In medicine, hyperspectral in-vivo diagnostics of tissue may avoid excision and permit in situ analysis (Vo-Dinh et al., 2004). Its application to real-time guidance in surgery is promising (Vo-Dinh et al., 2004).
The capability of the tunable filters central to these spectral imaging methods defines the quality of the data sets collected. Gat lists principal attributes of ideal tunable filters and describes examples of filters employed to date (Gat, 2000). Among these, tunable liquid-crystal and acousto-optical filters represent two prominent device classes (Gat, 2000; Vo-Dinh et al., 2004). The former is based on stacks of birefringent liquid-crystal plates integrated with polarizers, whereas the latter is diffractive in nature.
In the present contribution, we introduce a new tunable filter concept for potential application in multispectral and hyperspectral imaging systems. In short, we employ a resonant waveguide grating supporting leaky modes that is tuned by micro-electro-mechanical (MEMS) means. We begin this chapter by summarizing the physical basis for this class of tunable filters. Then, we provide numerical spectral characteristics of resonance elements based on exact electromagnetic models of the devices with representative materials. We investigate theoretically the operation of MEMS-tunable resonant elements.

---

Based on "Tunable Leaky-Mode MEMS Filters for Multispectral Imaging Applications," by R. Magnusson and M. Shokooh-Saremi, which appeared in IEEE Aerospace Conference Proceedings, March 1-8, 2008. 

In particular, we provide numerical results for a fixed transmission filter, a tunable reflection filter mounted on a low-index substrate, and then contrast its tuning capability with that of a classical Fabry-Perot filter in the LWIR band. Further examples of guided-mode resonance (GMR) tunable devices for multispectral imaging applications quantify their tunability relative to the mechanical displacement as well as spectral bandwidths and associated sideband levels. We envision these tunable filters finding use in aerospace multispectral imaging applications such as multi-channel thermal imaging, landscape temperature mapping, remote sensing, multispectral IR target recognition, and in other areas.

## 2. Resonance principle and context

Subwavelength periodic elements are presently of immense interest owing to their applicability in numerous optical systems and devices including biosensors, lasers, and filters. When the lattice is confined to a layer, thereby forming a periodic waveguide, an incident optical wave may undergo a guided-mode resonance (GMR) on coupling to a leaky eigenmode of the layer system. The external spectral signatures can have complex shapes with high efficiencies in both reflection and transmission. Computed examples in the optical spectral region show that subwavelength periodic leaky-mode waveguide films provide diverse spectral characteristics such that even single-layer elements can function as narrow-line bandpass filters, polarized wideband reflectors, wideband polarizers, polarization-independent elements, and wideband antireflection films (Ding & Magnusson, May 2004; Ding & Magnusson, November 2004). The relevant physical properties of these elements can be explained in terms of the structure of the second (leaky) photonic stopband and its relation to the symmetry of the periodic profile. The interaction dynamics of the leaky modes at resonance contribute to sculpting the spectral bands. The leaky-mode spectral placement, their spectral density, and their levels of interaction decisively affect device operation and associated functionality (Ding & Magnusson, May 2004; Ding & Magnusson, November 2004). In this paper, we investigate the tuning properties of a grating resonance element in which mechanical motion alters the structural symmetry. The chief properties of example tunable micro-electro-mechanical (MEMS) devices are summarized. This work initiates development of multispectral pixels operating in spectral regions where no comparable studies have been conducted to date.

GMR device parameters, including refractive index of grating layers or surrounding media, thickness, period, and fill factor, can all be applied to implement tunability. In past publications, a tunable laser using a rotating resonance element (i.e., angular tuning) and a photorefractive tunable filter were described (Wang & Magnusson, 1993). Furthermore, tuning can be accomplished by changing layer thickness or material refractive index, a method of significance in resonant sensor operation (Magnusson & Wang, 1993). Suh et al. reported analysis of a tunable structure consisting of two adjacent photonic-crystal films, each composing a two-dimensional waveguide grating, which could be displaced laterally or longitudinally by a mechanical force (Suh et al., 2003). Each periodic waveguide admitted guided-mode resonances whose coupling could be mechanically altered for spectral tuning. Additionally, numerous other tunable structures not inducing leaky modes have been described in the literature. As an example of a device in this class, Nakagawa and Fainman presented a structure in which a subwavelength grating was placed between planar dielectric mirrors, composing a Fabry-Perot cavity (Nakagawa & Fainman, 2004). Lateral and longitudinal motion yielded effective tuning via associated near-field coupling mechanisms. Park and Lee presented a tunable nanophotonic grating layer that was placed

on a flexible substrate (Park & Lee, 2004). By mechanically stretching the lattice, thereby changing the grating period, a variation in the angle of refraction of an incident beam of light was achieved.

Previously, we presented the characteristics of MEMS-tunable guided-mode resonance structures in the telecommunications spectral band and explained their operational principles. It was shown that such systems are highly tunable with only nanoscale displacements needed for wide-range tuning. Working with a single-example materials system, namely silicon-on-insulator (SOI), and fixed parameters, we quantified the level of tunability per unit movement for an example resonant structure. It was found that effective MEMS-based tuning can be accomplished by variation of grating profile symmetry, by changing the waveguide thickness, or both (Magnusson & Ding, 2006). Clearly, analogous tunable devices can be made in numerous other materials systems and made to operate in arbitrary spectral regions. As the operational wavelength diminishes to the visible region, the associated finer-feature patterning demands stricter tolerances in fabrication. Conversely, for the MWIR and LWIR bands, the structural features increase in size, relaxing fabrication tolerances.

## 3. Resonance device classes

In this section, we present examples of optical filters with distinct features and performance. A fixed guided-mode resonance element provides a narrow bandpass filter centered at 10 μm wavelength. A tunable bandstop filter fashioned with substrate-mounted complementary gratings is MEMS-tuned in the same spectral region. Finally, the tuning capability of a classical multilayered Fabry-Perot cavity is assessed for comparison and contrast with the GMR MEMS filters.

*Narrow-line bandpass filter for the LWIR band* – Bandpass filters are widely used to filter spectra into narrow wavelength bands typically in transmission geometry. Here, a narrowband filter based on leaky-mode resonance is designed with the particle swarm optimization (PSO) technique and its transmittance is determined (Shokooh-Saremi & Magnusson, 2007). Figure 1 shows the structural details of the device. This device consists of a subwavelength (namely, there exists no higher-order, freely propagating diffracted waves) grating whose period has been divided into four parts. The fraction of the period occupied by each medium is defined by the corresponding fill factor $F_i$. Figure 2 shows the

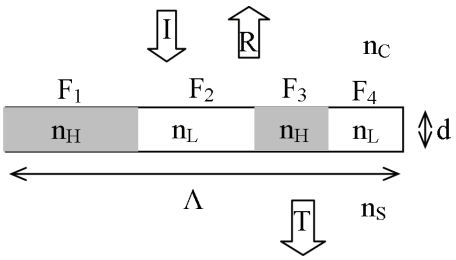



Fig. 1. Structure of a four-part GMR device used for designing a narrow bandpass filter. Λ, d denote the period and thickness of the grating, respectively, whereas $n_C$ and $n_S$ define the refractive indices of the cover and substrate media. Also, $n_H$ and $n_L$ are the refractive indices of materials in the grating region ($n_H > n_L$). The fractions $F_i$ (i=1,2,3,4) denote the associated fill factors. R is reflectance, and T is transmittance.

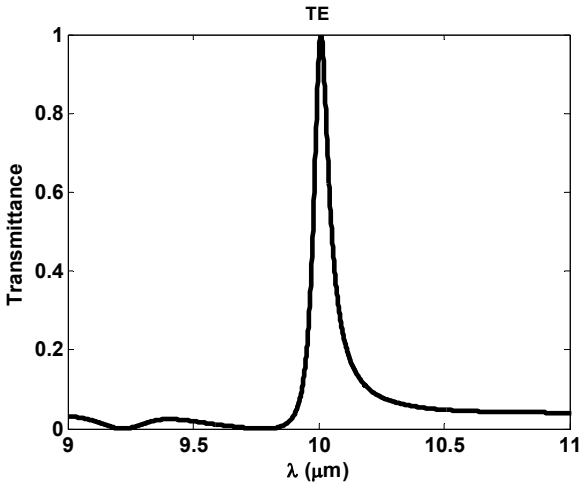



Fig. 2. Transmittance spectrum of a narrow bandpass filter designed by PSO for TE polarization (electric field vector normal to the plane of incidence). The period is $\Lambda$ = 6.57 μm, thickness d = 5.93 μm and $\{F_1,F_2,F_3,F_4\}$ = {0.137,0.177,0.372,0.314}. Also, $n_C = n_S = n_L = 1.0$ and $n_H$ = 3.42 (Si).

transmittance spectrum of the PSO-designed filter for normal incidence and TE polarization. The final design parameters are: $\Lambda$ = 6.57 μm, d = 5.93 μm, and $\{F_1,F_2,F_3,F_4\}$ ={0.137,0.177,0.372,0.314}. Also, $n_C = n_S = n_L$= 1.0 (membrane structure) and $n_H$ = 3.42 (Si). This filter has a transmission band of ~ 0.1 μm around the $\lambda$ = 10.0 μm central resonance wavelength. In the examples, silicon is used due to its high refractive index in the IR band; however, other applicable materials with high and low refractive indices in the LWIR band can be used in practical applications like Ge (n = 4.0), GaAs (n = 3.27), ZnSe (n = 2.4), NaCl (n = 1.5) and KCl (n = 1.46) (Janos).

In fabrication of elements of this class, the aspect ratio, namely the height-to-width ratio of each grating block is of key importance. In this example, the smallest aspect ratio is $d/F_1\Lambda$~6.6. Fabrication of this device would be possible with optical lithography and deep reactive-ion etching.

*Tunable LWIR bandstop filter*—Figure 3 shows a schematic diagram of a tunable structure that can be constructed with two silicon single-layer waveguide gratings, one of which would be mobile. The period, $\Lambda$, of the resonance structure in Fig. 3 is selected to implement tunability in the 8–12 μm wavelength range for TE polarization. Other parameters are selected such that an appreciable range of motion is available. The tuning parameters studied here are limited to the separation of the two binary Si blocks along the horizontal direction denoted by $F_{tune}$ (dimensionless fill factor) and the separation of the two gratings along the vertical direction denoted by $d_{tune}$. The tuning with horizontal motion varies the symmetry of the grating profile by shifting a Si block within the period (Ding & Magnusson, May 2004; Magnusson & Ding, 2006). This alters the spatial configuration of the localized resonant fields, including relative position of standing-wave peaks and grating materials. The vertical motion changes the net thickness and also affects the resonance wavelength and leaky mode distribution. The horizontal and vertical translational parameters $F_{tune}$ and $d_{tune}$ can be applied simultaneously or independently. The simulation results show that the

tuning by horizontal movement is more effective than the vertical movement (Magnusson & Ding, 2006). MEMS technology and actuation methods can be applied to implement these tunable elements.

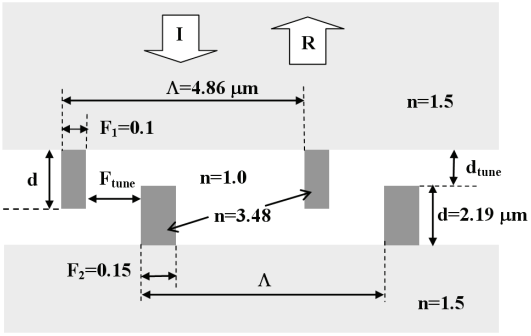



Fig. 3. An example tunable double-grating resonant structure. The gratings are made with silicon and supported on lower-index substrates. The incident wave is taken to be TE polarized.

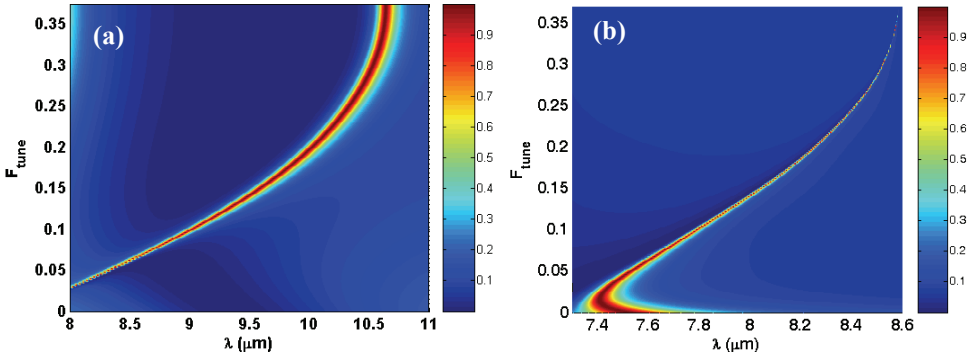



Fig. 4. Color-coded map illustration of resonance tuning $R_0(\lambda,F_{tune})$ by modulation of the profile symmetry while holding $d_{tune} = 0$ for (a) TE , and (b) TM polarizations. The incident angle is $\theta = 0°$.

Figure 4 provides a color-coded map of the reflectance of the zero-order wave $R_0(\lambda,F_{tune})$ that quantifies the spectral shift, lineshape, and linewidth of the resonance reflectance peak as a result of horizontal profile tuning for TE and TM polarizations. As seen, the tuning map for TM polarization falls outside of the 8–12 μm range; however, these two polarizations can provide a total tuning range of ~7.6–10.5 μm. Therefore, by utilizing a polarization switching method, wider spectral tuning range can be achieved. Figure 5 shows snapshots of the reflectance spectra for selected values of $F_{tune}$ for TE polarization. In general, the peak shift is accompanied by linewidth change; in this case, the resonance linewidth increases as the peak shifts to longer wavelength within the range shown. It is seen that the resonance wavelength can be shifted by ~2.5 μm, from 8.0 μm to 10.5 μm, with a horizontal movement of ~1.7 μm. At $F_{tune} = 0.375$, the structure is symmetric, accounting for the reversal in wavelength shift at that point. Thus, for example, the physical situation for $F_{tune} = 0.05$ is the

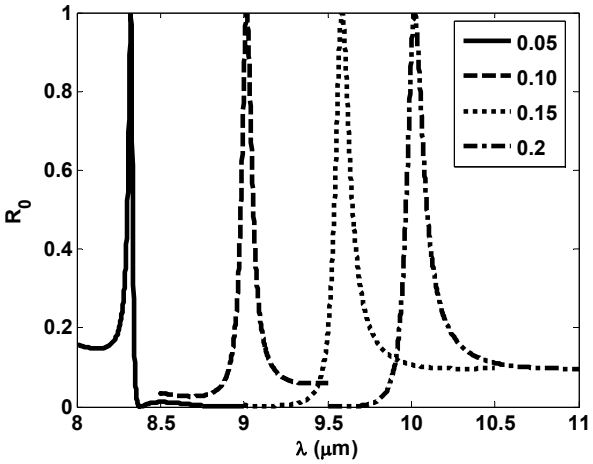



Fig. 5. Examples of reflectance spectra of the silicon double grating tunable filter for various values of $F_{tune}$ for TE polarization. The zero-order reflectance is denoted by $R_0$.

same as that for $F_{tune}$ = 0.70. Figure 6 shows the distribution of the total electric field inside the device and also the surrounding media at resonance for a given set of parameters. It is seen that the field amplitude in the Si blocks increases by ~×10 ($F_{tune}$ = 0.20) over the input wave amplitude, which is one unit. Varying the symmetry tuning parameter, $F_{tune}$, alters the internal field distributions and their amplitudes as seen in Fig. 6.

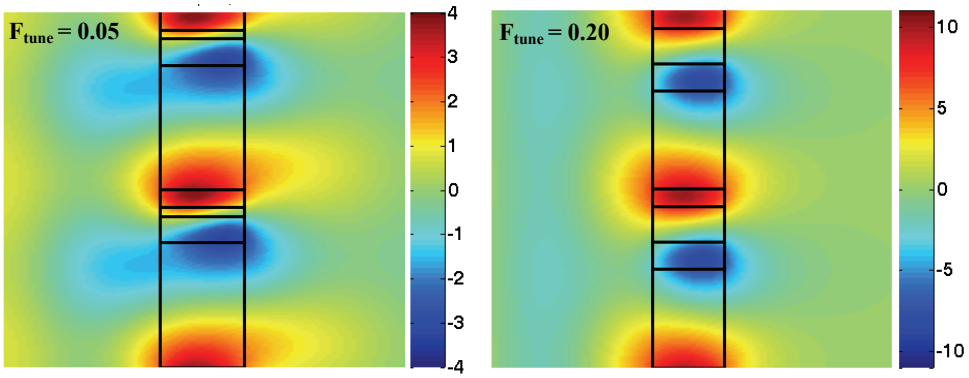



Fig. 6. Total electric field distribution patterns at resonance for two values of the symmetry parameter (TE polarization). The two counterpropagating leaky modes form a standing wave with a $TE_0$ mode shape at resonance. The incident wave has unit amplitude.

The spectral and modal results shown are obtained by rigorous coupled-wave analysis (RCWA) (Gaylord & Moharam, 1985) and modal analysis technique (Peng et al., 1975), respectively. Using these exact electromagnetic methods, the interaction of the incident light plane wave with general multilayered periodic devices is efficiently modeled. We have developed computer codes that handle general combinations of periodic and homogeneous

layered structures. Because of the plane-wave assumptions used, these codes run extremely fast and are found to be highly reliable as verified by repeated comparisons with experimental results. Additionally, coupled-wave field distributions, including resonant leaky-mode amplitudes as illustrated in examples above, can be conveniently and efficiently computed with RCWA and related methods.

*Tunable Fabry-Perot filters* — For context and to connect and contrast our methods with better known technology, we address briefly the properties of MEMS-tunable Fabry-Perot (FP) filters. Figure 7 shows the device details consisting of two quarter-wave Bragg stacks with 8 layers each surrounding a variable gap. Figure 8 shows the performance of the FP filter with

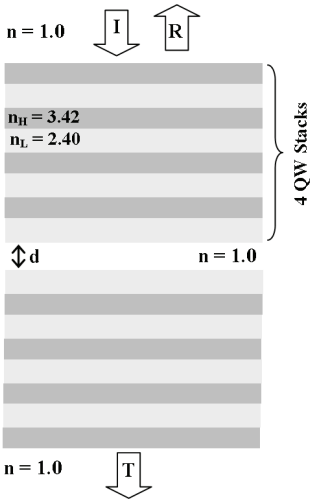



Fig. 7. A Fabry-Perot MEMS-tunable thin-film filter with variable gap operating in the in 8–12 μm band.

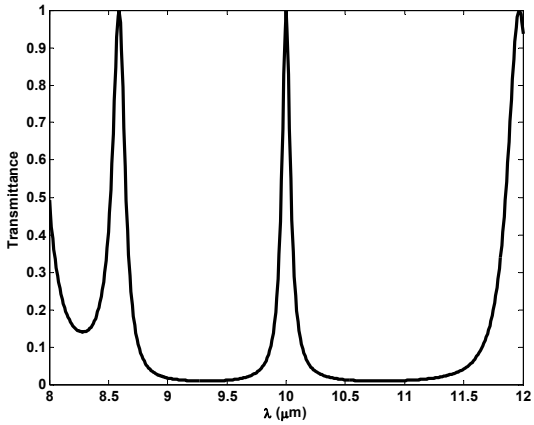



Fig. 8. FP filter transmission curve for example parameters that are $\theta = 0°$, $\lambda_0 = 10.0$ μm, $d_H = \lambda_0/4n_H = 0.731$ μm, $d_L = \lambda_0/4n_L = 1.04$ μm, and fixed air gap width of $d = 5.0$ μm.

representative parameters. Finally, Fig. 9 displays the tuning properties of the FP filter. Note that for a given gap width, say d = 5 μm, two transmission peaks arise in the 8–12 μm region. Thus, to eliminate extraneous transmissions, additional blocking (edge) filters are needed. The net result is that tuning is restricted by the parasitic neighboring resonance transmission channels as seen in the figure. In this example, spectral tuning across ~1 μm with gap change of ~5 μm is possible with proper blocking filters. This is to be compared with the tuning capability shown in Fig. 4 where a single resonance is encountered across a wide spectral band. This yields resonance wavelength change of ~2.5 μm with a movement of ~1.7 μm, which is considerably more effective.

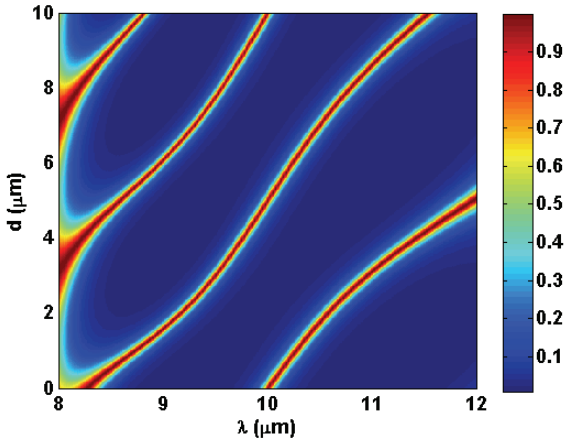



Fig. 9. FP filter performance under tuning by varying the gap dimension, d. The red bands define (d, λ) loci where the filter is highly transmissive.

## 4. Tunable membrane filter

In this section, a freestanding, tunable reflective pixel is introduced as a potential candidate for multispectral imaging applications. The device has a membrane structure in which the incident and substrate media are assumed to be air. The grating has four parts per period like the structure in Fig. 1. Figure 10 shows the structure of this tunable element. For simulating the action of the MEMS system for tuning the reflectance spectrum of the device, the air part with filling factor of $F_2$ is considered as being variable. This imitates the movement of the silicon part with filling factor $F_3$ by MEMS actuation as indicated in Fig. 10. The tunable imaging pixel has been designed to operate in the 8–12 μm band. The parameters of the device are as follows: $\Lambda$ = 6.0 μm, d = 2.4 μm, $F_1$ = 0.15, $F_3$ = 0.1, and $n_H$ = 3.42 (Si). Considering these parameters, Fig. 11 displays a color-coded map of $R_0(\lambda,F_2)$ illustrating the tuning of the resonance reflection spectrum. As seen in this figure, the pixel is tunable over the 9–12.4 μm range while the mechanical displacement needed for this tuning is ~0.373$\Lambda$ = 2.24 μm. Therefore, the rate of tuning is ~1.52 (wavelength shift per mechanical shift). Also, Fig. 12 shows example snapshots of the spectrum for various values of $F_2$. This figure quantifies the resonance peak line shape, line width, and side lobe levels associated with this particular pixel.

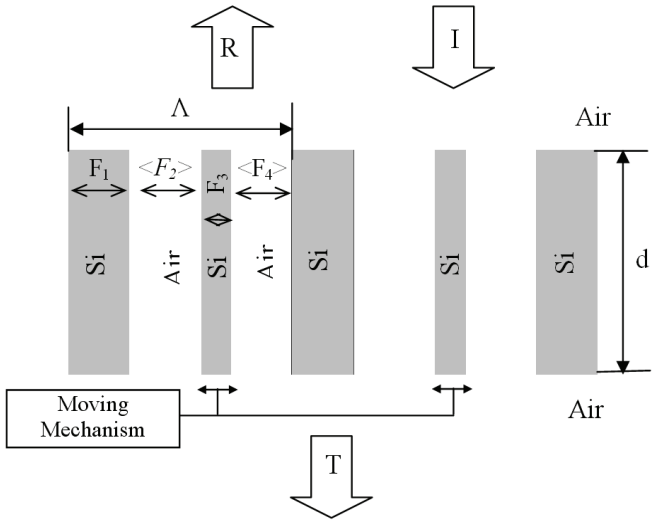



Fig. 10. Structure of a four-part GMR tunable membrane device. Λ, d are the period and thickness of the grating, respectively.

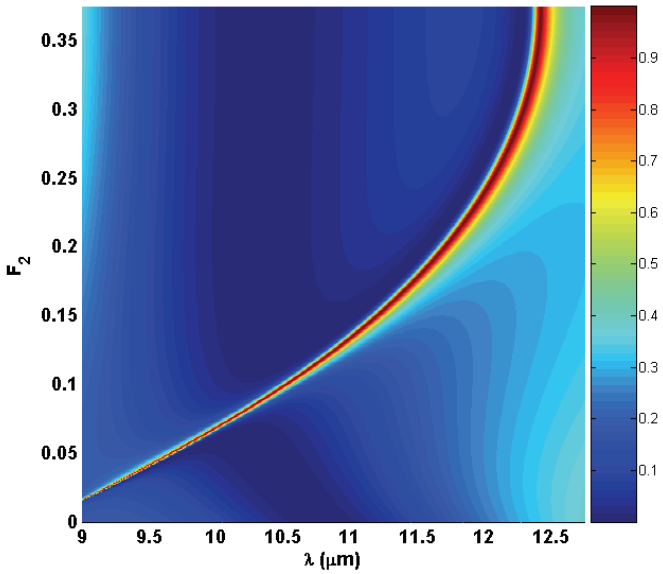



Fig. 11. Color-coded map of $R_0(\lambda,F_2)$ for the tunable MEMS pixel made with a silicon membrane. The parameters of the device are as follows: Λ = 6.0 μm, d = 2.4 μm, $F_1$ = 0.15, $F_3$ = 0.1, and $n_H$ = 3.42 (Si).

To study the angular response of the tunable elements, the variation of the resonance peak reflectance versus angle of incidence has been calculated and the result is shown in Figure 13. The center wavelength is 10.52 μm, and $F_2$ is chosen to be 0.1. It is seen that a favorable

numerical aperture is available for these devices. At ±2.5° angular deviation, the reflectance of the resonance exceeds 0.9 and the FWHM of the spectrum is ~10°.
Since these elements work in reflection mode, practical arrangements are needed to suitably direct the reflected beam to the detection system (for example, detector arrays). Figure 14 illustrates two possible schematic detection arrangements. In Fig. 14(a), a beamsplitter cube is utilized to direct the reflected beam from the pixel element to the detector array. This arrangement is useful if the element is designed to work under normal incidence conditions. On the other hand, for pixel elements designed to work at oblique incidence, the arrangement in Fig. 14(b) is more appropriate.

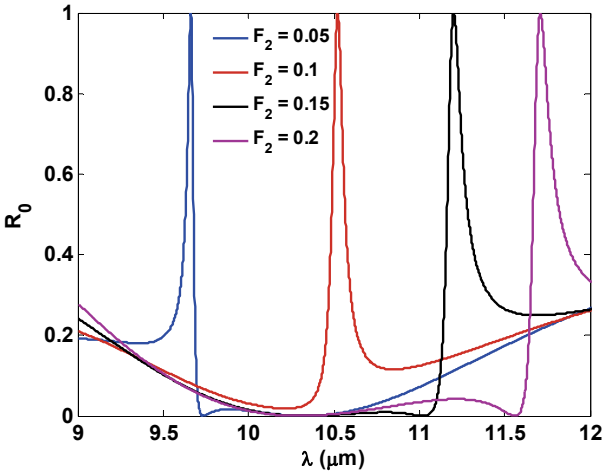



Fig. 12. Snapshots of reflection spectra for various values of $F_2$.

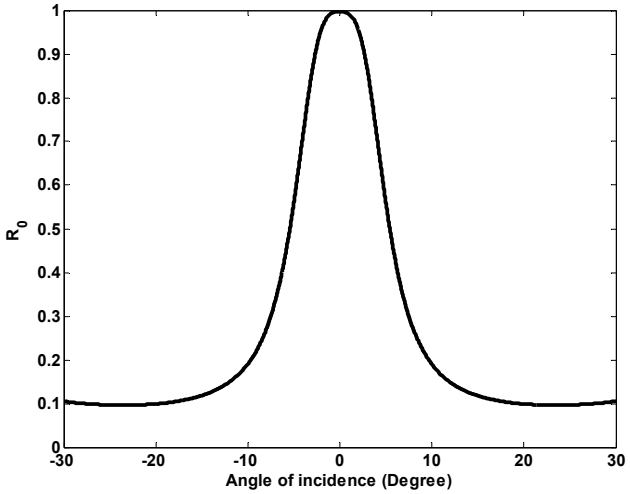



Fig. 13. Angular spectrum of the pixel element at λ = 10.52 μm and $F_2$ = 0.1.

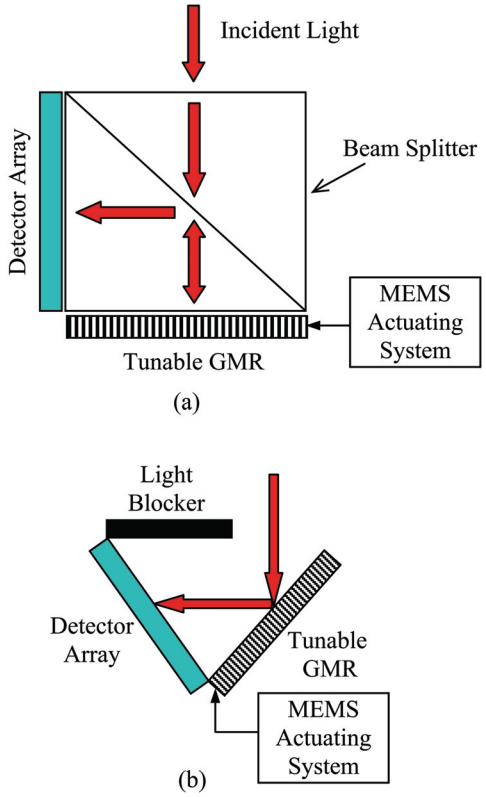



Fig. 14. Arrangements for reflected light detection from the tunable pixels under, (a) normal incidence and (b) oblique incidence.

## 5. Conclusions

In this paper, MEMS-tunable leaky mode structures have been investigated for applications in multispectral and hyperspectral imaging. It has been shown that high degrees of tunability can be achieved without parasitic neighboring spectral channels. Numerous computed examples of these devices have quantified their tunability relative to the mechanical displacement as well as spectral bandwidths and associated sideband levels. Particular example results for a silicon grating element with 6.0 μm period and 2.4 μm thickness show MEMS tuning of ~3.4 μm in the ~9–12 μm band and ~100 nm spectral resonance linewidth. We have previously studied analogous devices in the telecommunications region around 1.55 μm wavelength (Magnusson & Ding, 2006) and in the visible spectral region for use as display pixels (Magnusson & Shokooh-Saremi, 2007). For resonance devices operating in the MWIR and LWIR bands, the structural features increase in size relative to those in the short-wave regions, thereby relaxing fabrication tolerances to some degree. Using photolithography and deep reactive-ion etching, these filters can be fabricated in many common materials systems including silicon. Nevertheless, the high aspect ratios encountered in some cases demand high precision in fabrication.

High aspect ratios are particularly associated with small filling factors in the basic resonance gratings. Optimization in design to minimize aspect ratios while retaining high degrees of tuning remains a chief challenge. Experimental realization and characterization of MEMS-tuned LWIR multispectral elements is another interesting, future prospect.

## 6. References


Ding, Y. & Magnusson, R. (2004). Use of nondegenerate resonant leaky modes to fashion diverse optical spectra. *Opt. Express*, Vol. 12, No. 9, (May 2004) pp. 1885-1891, ISSN # 10944087

Ding, Y. & Magnusson, R. (2004). Resonant leaky-mode spectral-band engineering and device applications. *Opt. Express*, Vol. 12, No. 23, (November 2004) pp. 5661-5674, ISSN # 10944087

Gat, N. (April 2000). Imaging spectroscopy using tunable filters: A review, In: *Wavelet Applications VII*, Harold H. Szu, Martin Vetterli, William J. Campbell, James R. Buss, Eds., (Vol. 4056), pp. 50-64, SPIE, 0819436828, Bellingham, Wash

Gaylord, T. K. & Moharam, M. G. (1985). Analysis and applications of optical diffraction by gratings. *Proc. IEEE*, Vol. 73, No. 5, (May 1985) pp. 894-937, 00189219

Janos Technology, http://www.janostech.com

Magnusson, R. & Ding, Y. (2006). MEMS tunable resonant leaky mode filters. *IEEE Photonics Technol. Lett.,* Vol. 18, No. 13-16, (July 2006) pp. 1479-1481, 10411135

Magnusson, R. & Shokooh-Saremi, M. (2007). Widely tunable guided-mode resonance nanoelectromechanical RGB pixels. *Opt. Express*, Vol. 15, No. 17, (August 2007) pp. 10903-10910, ISSN # 10944087

Magnusson, R. & Wang, S. S. (1993). Optical guided-mode resonance filter. U.S. patent number 5,216,680, June 1, 1993

Nakagawa, W. & Fainman, Y. (2004). Tunable optical nanocavity based on modulation of near-field coupling between subwavelength periodic nanostructures," *IEEE J. Select. Topics Quantum Electron.*, Vol. 10, No. 3, (May/June 2004) pp. 478–483, 1077260X

Park, W. & Lee, J. B. (2004). Mechanically tunable photonic crystal structures. *Appl. Phys. Lett.*, Vol. 85, (November 2004) pp. 4845-4847, ISSN # 00036951

Peng, S. T.; Tamir, T. & Bertoni, H. L. (1975). Theory of periodic dielectric waveguides. *IEEE Trans. on Microwave Theory Tech.,* Vol. 23, No. 1, (January 1975) pp. 123-133, 00189480

Shokooh-Saremi, M. & Magnusson, R. (2007). Particle swarm optimization and its application to the design of diffraction grating filters. *Opt. Lett.*, Vol. 32, No. 8, (April 2007) pp. 894-896, 01469592

Suh, W.; Yanik, M. F.; Solgaard, O. & Fan, S. (2003). Displacement-sensitive photonic crystal structures based on guided resonances in photonic crystal slabs. *Appl. Phys. Lett.*, Vol. 82, (March 2003) pp. 1999-2001, ISSN # 00036951

Vo-Dinh, T.; Stokes, D. L.; Wabuyele, M. B.; Martin, M. E.; Song, J. M.; Jagannathan, R.; Michaud, E.; Lee, R. J. & Pan, X. (2004). A hyperspectral imaging system for in vivo optical diagnostics. *IEEE Engineering in Medicine and Biology Magazine*, Vol. 23, No. 5, (September/October 2004) pp. 40-49, 07395175

Wang, S. S. & Magnusson, R. (1993). Theory and applications of guided-mode resonance filters. *Appl. Opt.*, Vol. 32, No. 14, (May 1993) pp. 2606-2613, 00036935

# 25

# A Real Options Approach to Valuing the Risk Transfer in a Multi-Year Procurement Contract

Scot A. Arnold and Marius S. Vassiliou
*The Institute for Defense Analyses*
*The United States of America*

## 1. Introduction

The purpose of this paper is to develop methods to estimate the option value inherent in a multi-year government procurement (MYP), in comparison to a series of single-year procurements (SYP). This value accrues to the contractor, primarily in the form of increased revenue stability. In order to estimate the value, we apply real options techniques[1].

The United States government normally procures weapons systems in single annual lots, or single year procurements (SYP). These procurements are usually funded through a Congressional Act (the annual National Defense Authorization Act or NDAA) one fiscal year at a time. This gives Congress a great deal of flexibility towards balancing long and short term demands. For defense contractors, however, the Government's flexibility results in unique difficulties forecasting future sales when demand is driven by both customer needs and global politics.

Defense contractors face risks and advantages that set them apart from commercial businesses. Within a contract, the contractor faces a range of execution cost risk: from none in a cost plus fixed fee contract to high risk in a firm fixed price contract. The government also provides interest-free financing that can greatly reduce the amount of capital a contractor a contractor must raise through the capital markets. Additionally the government provides direct investment and profit incentives to contractors to invest in fixed assets. The net effect is that defense contractors can turn profit margins that may appear low when compared to other commercial capital goods sectors, into relatively high return on invested capital.

However, contractors have always faced high inter-contract uncertainty related to the short term funding horizon of the government. While the United States Department of Defense (DoD) has a multiyear business plan, in any given year, generating a budget entails delaying acquisition plans to accommodate changing demands and new information. At the end of the cold war, defense firms were allowed unprecedented freedom to consolidate. The resulting industrial base is composed of five surviving government contractors: Boeing, General Dynamics, Lockheed, Northrop Grumman, and Raytheon. By diversifying across a large number of government customers, these giants with thousands of contracts each have taken a giant step towards reducing inter-contract risk—no one contract is large enough to

[1] E.g., Amram & Howe (2003)

seriously harm the companies if it were canceled for convenience. However, the uncertainty around the likelihood of getting the next contract or how large it will be is still there and it is particularly important for large acquisition programs. For example, while Lockheed is the sole source for the F-22A, they always faced uncertainty in the number of units they will sell in the future. For example both the F-22A and the B-2 were originally expected to sell many more airplanes to the government than the actual number the government eventually purchased.

Under Title 10 Subtitle A Part IV Chapter 137 § 2306b, the military services can enter into multi-year procurement (MYP) contracts upon Congressional approval. There are six criteria that must be satisfied, listed in Table 1. The chief benefit for the government has been the "price break", criterion 1, afforded through the operating efficiencies of a long term contract. This benefit is readily passed to the government because it funds the necessary working capital investments needed to optimize production. It is still possible for the government to cancel the MYP contract; however, significant financial barriers such as a cancellation or termination liability that make it undesirable to do so.

| | Criteria Descriptions |
|---|---|
| 1 | That the use of such a contract will result in substantial savings of the total anticipated costs of carrying out the program through annual contracts. |
| 2 | That the minimum need for the property to be purchased is expected to remain substantially unchanged during the contemplated contract period in terms of production rate, procurement rate, and total quantities. |
| 3 | That there is a reasonable expectation that throughout the contemplated contract period the head of the agency will request funding for the contract at the level required to avoid contract cancellation. |
| 4 | That there is a stable design for the property to be acquired and that the technical risks associated with such property are not excessive. |
| 5 | That the estimates of both the cost of the contract and the anticipated cost avoidance through the use of a multiyear contract are realistic. |
| 6 | In the case of a purchase by the Department of Defense, that the use of such a contract will promote the national security of the United States. |

Table 1. The Six Criteria for a Multi-Year Procurement[2]

The government reaps operational savings by negotiating a lower up-frontprocurement price. These savings are achieved through more efficient production lot sizes and other efficiencies afforded through better long-term planning not possible with SYP contracts. The government can explicitly encourage additional savings by using a cost sharing contract. It can implicitly encourage additional savings with a fixed price contract. In the latter case the longer contract encourages the contractor to seek further efficiencies since it does not share the savings with the government. In fact some might propose this last reason is the best reason for a contractor to seek an MYP.

In addition to the cost savings achieved through more stable production planning horizon, we see that the MYP provides the contractor with intrinsic value through the stabilization of its medium term revenue outlook. Thus an MYP is also coveted by defense contractors because it provides lower revenue risk. What about the possibility that a longer term firm

[2] United States Code, Title 10, Subtitle A, Part IV, Chapter 137, Section 2306b

fixed price contract exposes the contractor to higher cost risk? This risk is often eliminated through economic pricing adjustment (EPA) clauses that provide a hedge against unanticipated labor and material inflation. Furthermore, from the criteria in Table 1, MYP contracts are only allowed for programs with stable designs that have low technical risk. As stated above, it is more likely that the MYP offers the contractor the opportunity to exploit the principle-agent information asymmetry and make further production innovations unanticipated at contract signing[3].

We believe that the lower risk MYP contract will allow investors to discount contractors' cash flow with a lower cost of capital creating higher equity valuations. From the contractors' perspective, the MYP contract provides a hedge against revenue risk. We can estimate the incremental value of the MYP versus the equivalent SYP sequence using option pricing methods. Presently the government does not explicitly recognize this risk transfer in its contracting profit policy. The government profit policy is to steadily increase the contract margin as cost risk is transferred to the contractor. For example a cost plus fixed fee contract might have a profit margin of 7% while a fixed price contract, where the contractor is fully exposed to the cost-risk, of similar content could have a margin of 12%[4]. By limiting some of the contractor's cost-risk exposure, an EPA clause might result in a lower profit margin; however, the profit policy makes no mention of an MYP contract, which reduces the contractor's inter-contract risk. And while most of the profit policy is oriented towards compensating the contractor for exposing its capital to intra-contract risk and entrepreneurial effort, there are provisions designed to provide some compensation for exposing capital to inter-contract risk—e.g. the facilities capital markup. The implication is that as long as the government does not explicitly price the reduction in cost-risk going from a fixed price SYP contact to an MYP contract, the contractor is able to keep the "extra" profit. In this paper we present a method to estimate the value an MYP creates for a defense contractor in its improved revenue stability. The contractor can use this information in two ways. First, the information provides guidance for how much pricing slack the contractor can afford as it negotiates an MYP with the government whether or not the latter recognizes that better revenue stability has discernable value. Second, if the government tries to reduce the contractor's price based on this transfer of risk, the contractor has a quantitative tool to guide its negotiation with the government.

## 2. Financial structure and valuation of an MYP

In this paper, we will present how to estimate the value imbedded in the risk transfer from the contractor to the government in an MYP contract using real options analysis. Table 2 lists recent MYP contracts. Note that while the table mostly shows aircraft the contract type can be applied to other acquisitions. Since FY2000, MYP contracts have declined from about 18 percent of defense procurement to about 10 percent; however, over this period they have totaled to about $10 billion per year. These contracts are 3 to 5 times larger than SYP contracts and can represent an important portion of the contractor's revenue.

---

[3] Rogerson, W. P, *The Journal of Economic Perspectives* ,V. 8, No. 4, Autumn 1994, pp. 65-90

[4] Generally the project with a cost plus contract has higher technical uncertainty than the project with the fixed price contract. The government does not expect contractors to accept high technical risk projects using a fixed price contract.

| Program | Period | Amount ($ Billions) | Type of System |
|---|---|---|---|
| Virginia Class[5] | 2009-2013 | $ 14.0 | Submarine |
| CH-47F[6] | 2008-2013 | 4.3 | Aircraft |
| V-22[7] | 2007-2012 | 10.1 | Aircraft |
| F-22A[5] | 2007-2010 | 8.7 | Aircraft |
| F-18 E/F[5,7] | 2005-2009 | 8.8 | Aircraft |
| DDG-51[8] | 2002-2005 | 5.0 | Ship |
| AH-1 Apache[5,7] | 2001-2005 | 1.6 | Aircraft |
| C-17A[5,9] | 1997-2003 | 14.4 | Aircraft |

Table 2. Recent Major Multi-Year Procurement Contracts

As an acquisition programmatures, the contractor implicitly receives an option on an MYP that is not executable until authorized by the Congress and negotiated by the relevant military service. If conditions are met and the option is exercised, the contractor transfers the SYP revenue risk to the government, which commits to buying the predetermined number of units. There are two financial instruments that approximate this transaction: a put and a cash flow swap or exchange option. Both structures provide the protection buyer, i.e. the contractor, insurance against losses in the underlying asset, i.e. the net present value of the cash flow derived from the sales. For the duration of the MYP contract, the contractor receives predictable revenue while the government forgoes the flexibility to defer or cancel the procurement by agreeing to pay substantial cost penalties for canceling the MYP contract. To value the MYP, we will employ the exchange option of Margrabe[10]. From this analysis the government will be able to estimate the contractor's value of transferring revenue risk to the government as a function of the size of the contract and the volatility of the contract's value. Since the option is not actively traded, the ultimate negotiated price could be heavily influenced by the government and contractor attitudes towards risk.

## 3. Real options

A put option is a common financial contract that gives the owner the right to sell an asset, such as a company's stock, for a pre-determined price on or before a predetermined date. Non-financial contingent pay-offs that behave like financial options, but are not traded as separate securities are called *real options*. Real options provide the holder of the asset similar risk management flexibility though they are not yet sold separately from the underlying asset. For example, oil drilling rights give the holder the option, but do not require, exploring, drilling, or

[5] Internal publication from Northrop Grumman, "Navy Awards $14 Billion Contract for Eight Virginia Class Submarines", *Currents*, January 5-9, 2009

[6] Graham Warwick, "Boeing Signs CH-47F Mulityear Deal", *Aviationweek.com*, August 26, 2008

[7] United States Government Accountability Office, *Defense Acquisitions DoD's Practices and Processes for Multiyear Procurement Should be Improved*, GAO-08-298, February, 2008, p. 9

[8] U.S. Department of Defense Press Release, Office of the Assistant Secretary of Defense (Public Affairs), No. 470-02, September 13, 2002.

[9] Second of two multi-year contracts.

[10] Margrabe, W., *Journal of Finance*, 33, 177-86 (1978)

marketing the oil to customers. Patents are another example that can be viewed the same way: the holder of the patent has the option but is not obliged to deploy the technology. Usually these investment flexibilities come into play as contingent pay-offs: they allow the investor to delay committing cash until positive pay-off is better assured. Real options capture the capability of investors or managers to make valuable decisions in the future.

More generally, real options analysis captures some of the value of management's capability to make dynamic programmatic changes, based on new and better information, within the levers and construct of a given business project. The real-options approach explicitly captures the value of management's ability to limit downside risk by stopping poorly performing programs. It also captures the value inherent in the possibility that management will exploit unexpected successes.

An MYP contract contains a real option allowing the contractor a choice to abandon the uncertainty associated with relying on sequential SYP contracts to implement the government's acquisition strategy for a weapon system. For example an aircraft manufacturer who is the single source for an air vehicle, such as the F-16 or F/A-18, has the exclusive option to negotiate an MYP contract to sell the next four lots to the Air Force or Navy. Given that most weapons acquisition programs buy fewer units than planned, the contractor will exercise the option by entering into an MYP contract.

The contractor implicitly owns the MYP option as the sole source for the procurement. Unlike a financial option which the buyer can choose from a selection of the strike prices and tenors, an MYP option does not explicitly exist until the government and contractors negotiate the terms of the contract. In negotiating the terms of the MYP, the contractor and government are negotiating the option's strike price—and up to that point it appears as though the contractor received the option for free. Once negotiated it is usually executed which is like exercising an at-the-money put option. We will define the option parameters below, recognizing that they may not be explicitly defined until the option is exercised.

There are a number of techniques that may be used to value a real option. One way is to adapt the framework developed by Black and Scholes[11] (BS) for financial options. Real-options investments are not often framed as neatly as puts and calls on corporate equities traded on the Chicago Board Options Exchange. However, if we can describe the real options embedded in an MYP contract along the lines of the appropriate standard options framework, we can try to employ the BS option pricing framework. Other alternatives include the binomial method[12], dynamic programming, simulation, and other numerical methods to name a few.

## 4. Are real options really used by managers?

Real options have been a topic of vigorous academic research for decades. The published literature abounds with theoretical papers, and with applications to a wide variety of domains. These domains include, for example: the aerospace[13,14], telecommunications[15],

[11] Black & Scholes (1973)

[12] E.g., Copeland & Tufano (2004)

[13] Richard L. Shockley, *J. of Applied Corporate Finance*, 19(2), Spring 2007

[14] Scott Matthews, Vinay Datar, and Blake Johnson, *J. of Applied Corporate Finance*, 19 (2), Spring 2007

[15] Charnes et al. (2004)

oil[16], mining[17], electronics[18], and biotechnology[19] industries; the valuation of new plants and construction projects[20]; real estate[21]; the analysis of outsourcing[22]; patent valuation[23]; the analysis of standards[24]; and the valuation of R&D and risky technology projects[25].

There is some evidence that real-options thinking has permeated the real world in some niches. The technique does appear to be used seriously in the oil industry, for example,[26] to analyze new ventures. Perhaps one reason is that it is easier to track the value of the underlying asset in that industry than in others. Reportedly, real options analysis has been used at Genentech in all drug development projects since 1995, and Intel has used it to value plant expansion[27]. Hewlett-Packard reportedly uses a set of risk management tools, including real options analysis, in its procurement practices[28]. It is perhaps not surprising that real options analysis has taken root in engineering and R&D-intensive industries engaging in large and risky capital expenditures. The fact that many of these companies have relatively high proportions of engineers and scientists in their management structures may also be a contributing factor. There appears to be a perception that real options analysis is inherently more "difficult" than other valuation methods, although this is not necessarily the case[29].

Real-options analysis is not as pervasive as conventional discounted cash flow analysis in most corporate and government capital budgeting decisions. This alone does not invalidate the analysis; it takes decades for analytical tools to take hold or to be changed. Financial engineering has become entrenched in the financial services and consulting industries[30]. As these tools evolve it will be natural to apply them to non-financial business problems. Indeed the tools are not unique to the financial sector but were adapted from the mathematical sciences. The relatively slow penetration of real-options analysis reflects the difficulty for most organizations in articulating the risks faced in capital decisions.

The remainder of this paper will focus on explaining and applying options pricing methods to valuing the portion of the MYP contract this is a risk management proposition.

## 5. Options theory

We will use closed form BS-type option pricing methods to estimate the contractor's value in an MYP contract. Financial options fit into the larger domain of derivatives or contingent

[16] Cornelius et al. (2005)
[17] Colwell et al. (2003)
[18] Duan et al. (2003)
[19] Ekelund (2005); Remer et al. (2001)
[20] Ford et al (2004); Rothwell (2006)
[21] Fourt (2004); Oppenheimer (2002)
[22] Nembhard et al. (2003)
[23] Laxman & Aggarwal (20030
[24] Gaynor & Bradner (2001)
[25] Paxson (2002); MacMillan et al. (2006)
[26] Cornelius et al. (2005); IOMA (2001)
[27] IOMA (2001)
[28] Maumo (2005)
[29] Amram & Howe (2003); Copeland & Tufano (2004)
[30] Although with mixed results in structured finance and credit default swap applications.

claims: financial instruments whose value derives from claims on pay-offs from event-driven changes in the value of an underlying asset. There are two types of derivatives buyers: hedgers who are naturally exposed to the underlying asset volatility and speculators who seek exposure to this risk.

A simple example is an equipment manufacturer with occasional large foreign exchange exposures when its machines are exported. The manufacturer could hedge the foreign exchange risk by buying put options on the foreign currency he expects to receive upon the sales transaction. The put option allows the manufacturer to exchange foreign currency for dollars at a predetermined date and exchange rate and thus eliminates profit volatility. The manufacturer is the hedger and the bank could be a speculator[31].

Insurance is another example where the insurer (the speculator), sells coverage to insureds (hedgers) for a premium. The insurer mitigates its position through many risk management tools: setting up loss reserve accounts which are based on detailed loss histories; diligent underwriting (i.e. pricing the coverage according to specific risks); avoiding certain risks (i.e. correlated high exposure risks such as asbestos, floods, or mold damage); limiting correlated risks (i.e. wind damage in Florida or earthquakes in California); hedging through reinsurance; etc. The government is actually one of the largest insurers providing many types of coverage against risks that many private insurers avoid: flood, nuclear; commercial space launch, terrorism, aviation war and hijacking, etc.

Compared to most risks to which the government is exposed, absorbing a few years of SYP volatility is a relatively tame risk transfer particularly in the context of the statutory "underwriting" that must occur before Congress will authorize such a contract. In the MYP contract, the defense contractor is the hedger, while the government is "speculating" that by meeting the MYP criteria it should be able to benefit by accepting the contractor's risk. The MYP criteria in Table 1 are an effective underwriting tool for the government. By passing the criteria, the government is actually absorbing little risk since by criteria 2 and 3 they would have acquired all of the units even without the MYP.

It is important to note that not all hedges make good business sense. The rules as whether or not to hedge are based entirely on the cost and benefits to shareholders who are free to diversify some of the idiosyncratic risk away from their investment portfolio. The options pricing models will not discern this trade-off for the contractors but it is likely to be the basis for the contractor's perspective in negotiating with the government. Regardless of the contractors' risk aversion, our goal is to elucidate the value created by the risk transfer. The government is taking on new risk by entering into the MYP contract—this risk transfer creates a significant benefit for the contractor counterparty whether or not they want to pay for it.

## 6. MYP option analysis

A put option has the desired insurance-like structure of an MYP contract: with the embedded risk transfer component of the MYP contract the contractor gains the right to sell a fixed number of units at a pre-set price. However, the MYP, like many real options, does not strictly eliminate the SYP risk; there is some risk that the government could cancel the

[31] The bank may also hedge its foreign exchange exposure.

contract or change the number of units[32]. Thus an exchange option, which gives the holder the right to exchange one cash flow for another on or before a given date, has advantages over a put option since its cash flow corresponds more closely to the way an MYP would be structured. The put and exchange options are closely related.

The key difference between the put and the exchange option is that on exercise, a put buyer receives a certain cash settlement while with an exchange option the buyer obtains a "cash flow" with different volatility. This property is ideal when in fact the MYP contract usually has a flexibility clause for variations in quantity (VIQ).

Consider a put option for the sake of the simplicity of its properties. A put provides a payoff to the option holder when it is exercised before the expiry and the exercise price is greater than the market or spot price of the underlying asset. An option holder can buy the asset at spot price S and sell it at the strike price X and receive a payoff X-S. Alternatively, an option holder having a long position in an asset canhedge against losses with puts, much like an insurance policy.

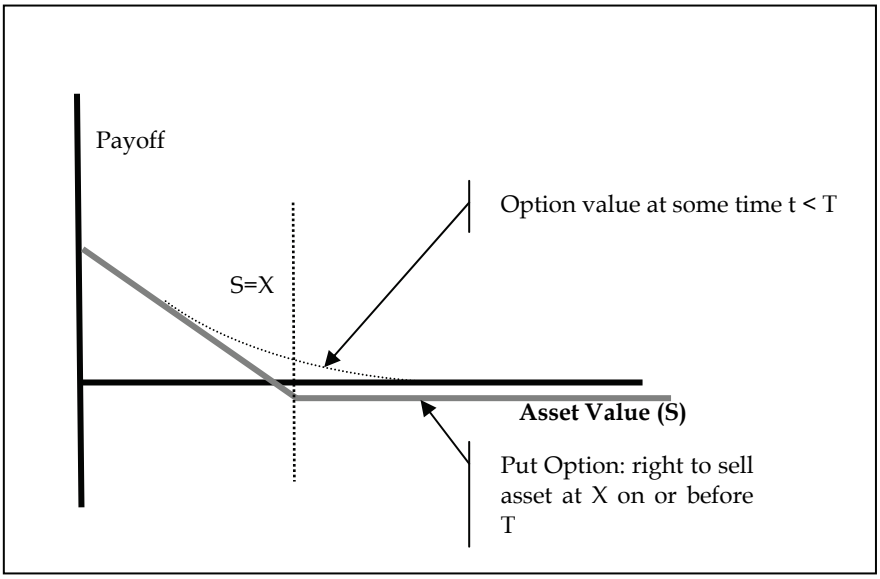



Fig. 1. Put Pay-off Diagram

Figure 1 depicts the payoff of a put option on or prior to the expiry. Once exercised, options are zero-sum contracts: the writer "loses" and the holder gains or vice versa. If the option expires unexercised, the holder's only loss is the premium paid to the writer. If the put option is held as a 1:1 hedge against a long position in the underlying stock, however, the net pay-off is nil, ornegative once the option premium is included. In the same way a contractor with an MYP contract is hedging against the uncertainty in the government's procurement decisions. The contractor net gain is neutral since the payoff depicted in Figure 1 is offset by the underlying losses in sales that would have happened if there were no MYP. The MYP option pay-off is the protection against losses and the contractor will only observe

[32] Canceling the contract usually would come with considerable cost to the government.

that it has stable, predictable cash flows. However, more predictable cash flows allow investors to value the contractor's equity higher. The government, on the other side, faces the risk that it will be forced to manage future budget uncertainties by increasing taxes or debt, cutting programs other than the MYP, or paying a higher termination fee if it cuts the MYP.

## 7. Extending financial options to the MYP option

Ideally we would like to be able to use a formula, such as that of Black and Scholes, to estimate the value of a MYP contract option. However, this is only reasonable if the contingent pay-offs behave within the constraints and assumptions behind the BS model. Though the basic BS formula applies to dividend protected European options in an arbitrage free market, it could be applied to a real option if its value depends on: the underlying asset value (S); the asset's volatility (σ); and whether the option time frame resembles that of a European option[33].

The worth of the MYP contract option depends on the value of the underlying asset—i.e. the net present value of future cash flow implied by the procurements. The uncertainty around the size of these cash flows is also a key value driver: low risk SYP contracts have less risk to be transferred to the government and lower the contractor's need for an MYP. Later we will discuss in more detail how to assess the volatility(the standard deviation of the market price of an asset) of the value of a series SYP contracts. Unlike equity stocks, currencies, and other traded securities, volatility in the case of a real option is difficult if not impossible to observe so we need to find a suitable tracking asset. The option pricing models can still be used to value the real option using the tracking asset's volatility if there is sufficient correlation between the tracking asset and the real option underlying asset valuation fluctuations.

The time frame of the MYP contract option is reasonably close to a European option, since it can be exercised only when the contract is executed. Also inherent in the BS model is that the return process of the underlying asset follows a Brownian motion process where the returns have a lognormal distribution.

## 8. The Black-Scholes model

The value of the put option p on Company A's stock at time t until expiration at time T can be estimated using the BS model:

$$p(S,t) = Xe^{-r(T-t)}N(-d_2) - SN(-d_1) \quad (1)$$

S and X are A's stock spot price at valuation and strike (at expiry T) per share respectively. $N(d_1)$ and $N(d_2)$ are the cumulative normal distributions of $d_1$ and $d_2$:

$$d_1 = (\ln(S/X) + (r + \sigma^2/2)(T-t))/ (\sigma (T-t)^{1/2})$$

$$d_2 = d_1 - \sigma (T-t)^{1/2}$$

[33] European options can only be exercised on the expiration data while American options can be exercised on or before expiry.

σ is the standard deviation or volatility of A's stock price over the span of the option life[34]. r is the interest rate of a risk-free bond with the tenor of the option expiry. Note that the thin dotted curve in Figure 1 never goes below zero; an option has value until expiry even if it is out of the money (i.e. for a put, S>X). This value is derived from the "time value" or asymmetric opportunity value of the option which allows the holder the possibility that it will come into the money prior to expiry without any risk of negative payoff.

The BS model assumes that the stock price changes are log-normally distributed, such that over time, the logarithm of the price changes follows a Weiner process. With the use of Ito's theorem and several more assumptions, the put option price *p,* as a function of S, is calculated using (1)[35]. In contrast to no-dividend European options assumed in (1) American options can be exercised up to or on the expiry date greatly complicating the mathematics behind their valuation. Most single equity options are American, while options on indices, such as the S&P 500, are European.

Applying (1) to the MYP, S is the net present value (NPV) of the cash flow expected from a series of SYP contracts; X is the price of the NPV of the MYP contract cash flows; and T is last day the final lot could be changed under an SYP. σ would ideally be the volatility of the NPV of the SYP cash flows, but since this volatility is virtually impossible to observe, it will be estimated using the contractor's stock as a tracking asset.

## 9. Exchange option

The exchange option allows the holder the right to swap cash flow $x_2$ (the risky SYP profit stream) for cash flow $x_1$ (the less risky MYP profit stream). This option is more general and better captures some of the flexibility the government has with actual MYP contract terms. The BS-based formula to value an exchange option is:

$$w = e^{-r(T-t)}x'_1N(d_1) - e^{-r(T-t)}x'_2N(d_2) \quad (2)$$

Again r is the risk-free rate, x'1 the strike price of asset 1 (MYP), x'2 the strike price of asset 2 (SYP), and $N(d_1)$ and $N(d_2)$ are the cumulative normal distributions of $d_1$ and $d_2$:

$$d_1 = (\ln(x_1/x_2)+(\sigma'^2/2)(T-t))/\ (\sigma'\ (T-t)^{1/2})$$

$$d_2 = d_1 - \sigma'\ (T-t)^{1/2}$$

$$\sigma' = (\ \sigma_1^2 + \sigma_2^2 - 2\ \rho\sigma_1\ \sigma_2)^{1/2}$$

[34] Technically it is the instantaneous volatility – something that is hard to measure.

[35] p(S,t) is found by solving the following partial differential equation:

$$p_t = \tfrac{1}{2}\ \sigma^2S^2p_{SS} + rSp_S - rp$$

The equation is subject to the terminal condition: p=max[0,X-S], and to upper and lower boundary conditions: $p=Xe^{-rT}$ for S=0 and p=0 for $S\to\infty$. S follows the Wiener process through the following stochastic differential equation: $dS = \mu Sdt+\sigma Sdz$. Here μ is the average growth rate; σ is the standard deviation of this growth process; and r is the risk free interest rate.

Where $\sigma_1^2$ is the variance of $x_1$, $\sigma_2^2$ is the variance of $x_2$, and $\rho$ is the correlation between $x_1$ and $x_2$. Here $\rho$ is likely to be close to 1 since $x_1$ and $x_2$ are essentially the same assets whose risks are derived from the same source. In our base analysis, $x_1$ is assumed to be certain, i.e. the MYP units are fixed in each lot and the government has no flexibility to cancel the MYP. Thus $\sigma' = \sigma_2$ since $\sigma_1 = 0$. If however, the MYP contract has some uncertainty, e.g. from a VIQ clause or a low termination fee, $\sigma_1$ could be adjusted to reflect the relative risk between $x_1$ and $x_2$.

The exchange option can also be thought of as a simultaneous call option on asset 1 with strike price $x_2$ and a put option on asset 2 with a strike price $x_1$. A call option is a contract that gives the owner the right to buy an asset at a predetermined price on or before a predetermined time. The main difference between the put and exchange options is that the latter allows both assets to have price volatility. Furthermore the exchange option allows for the upside volatility in the MYP, i.e. that more units than the original plan could be purchased.

## 10. Estimating option pricing parameters

Consider as an example a major acquisition weapon system, "Program G,", executed by the contractor Company A. Program G and Company A do not correspond directly to any real-life program or company, although the numbers discussed in this paper are constructed from real examples. Program G's base SYP net cash flows can be derived from the relevant military service's Selective Acquisition Report (SAR). Table 1 lists the profits associated with Program G system lots 6 through 10[36]. Since lot 6 is the first year of both contract scenarios its profits are omitted from the analysis since they will not depend on whether the MYP is executed. The SYP uncertainty is only in lots 7 through 10. The profits are stated in "then year" (nominal) terms and the net present value of the flows is discounted at Company A's cost of capital.

| | | |
|---|---|---|
| Lot 6 | $ | - |
| Lot 7 | | 200 |
| Lot 8 | | 200 |
| Lot 9 | | 250 |
| Lot 10 | | 175 |
| Total Profit | $ | 825 |
| | | |
| Present Value | $ | 630 |

Table 3. Contractor SYP Profit ($ Mils)

The present value total of $630 resents the projected total asset value ($x_2$) of the last four lots of the SYP. We initially restrict $x_2 = x_1$, or that the option be "at the money"[37].

[36] We assume a dollar for dollar profit cash flow conversion.

[37] This is a realistic assumption since the number of units in the MYP and SYP are assumed to be the same in the standard business case analysis.

## 11. Volatility

For most non-traded assets, such as the profits of Program G, even the historical volatility is difficult to measure[38]. To properly use the BS model to value Program G MYP option it is imperative to find a traded tracking asset whose volatility is highly correlated to the implied volatility of the asset underlying the embedded real option.

Fortunately, Company A's equity is publicly traded. Company A is a moderately diversified government contractor with two divisions, Defense and Non-Defense, that serve different government sectors:. We find from their financial statements that Program G represents a substantial share of the Defense Division's earnings before interest and tax (EBIT). The EBIT breakout by division is presented in Table 2. The Defense Division has contributed a significant portion of the total profits, particularly in recent years. Comparing Tables 1 and 2 we can see that Program G represents over half of the Defense Division's historical EBIT.

| Year | Non-Defense | Defense | | Total | Stock Price |
|---|---|---|---|---|---|
| | | EBIT | % Total | | |
| 2001 | $ 758 | $ 242 | 24% | $ 1,000 | $ 13 |
| 2002 | 564 | 92 | 14% | 656 | 7 |
| 2003 | 522 | 128 | 20% | 650 | 13 |
| 2004 | 652 | 123 | 16% | 775 | 20 |
| 2005 | 679 | 167 | 20% | 846 | 19 |
| 2006 | 552 | 257 | 32% | 809 | 19 |
| 2007 | 443 | 335 | 43% | 778 | 21 |
| 2008 | 742 | 370 | 33% | 1,113 | 26 |

Table 2. A's EBIT Breakdown by Division ($ Mils except for Stock Prices - $/Share)

Company A is a large enterprise, and while Program G contributes significant profits towards to total corporate profit, it is not necessarily enough to drive the overall equity performance. Before we can assign Company A's equity volatility as a tracking asset for Program G, we need to establish a closer linkage. Table 3 shows Company A's earnings growth and volatility by division as well as the market performance of its equity from 2000 to 2007. We see that the Defense Division tracks the overall stock performance better than the Non-Defense Division, and better than the company as a whole. This may be because Company A is often identified as a defense company and its stock price, which is forward looking, trades on the trends in the overall defense industry.

| | Non-Defense | Defense | Total | Stock |
|---|---|---|---|---|
| Growth | 0% | 6% | 2% | 10% |
| Volatility (σ) | 29% | 36% | 22% | 37% |

Table 3. A's EBIT Growth and Volatility by Division 2001 to 2008

[38] A crude estimate could be constructed by collecting the annual Selected Acquisition Report estimates for the number of units funded through the life of the program.

One more indication that Company A's stock is a good tracking asset for Program G is the correlation between the division's EBIT and the stock price, as shown in Table 4. Defense Division EBIT has a 72% correlation to the stock price—even higher than the company's total EBIT. Note that this is not to imply that the stock price drives Program G profit volatility; but rather that the stock price mirrors the EBIT volatility of the Defense division which is strongly driven by the program G business. Since we cannot measure Program G EBIT volatility directly, we will use the stock price volatility as a proxy. We could use the Defense division's historical EBIT volatility (Table 3) to track Defense division volatility, instead we prefer to use the forward-looking implied volatility estimated in Table 5.

| | $\rho_{Division,Stock\ Price}$ |
|---|---|
| Non-Defense | 18% |
| Defense | 72% |
| Total | 59% |

Table 4. Correlation between Stock Price and Division EBIT from 2001 to 2008

## 12. Time horizon

We have already hinted at the time horizon for the MYP option. It starts when congress gives the services authority to enter into an MYP with A. It expires at the beginning of the last year or lot of production (assuming one lot per fiscal year) since that would be the last point at which the government could have reduced the number of units in an SYP contract. Assume that the MYP authority is granted six months prior to negotiation. The total life of the MYP is then five years and six months.

## 13. Interest rate

The risk free interest rate used in the analysis is the rate on a Treasury bill whose maturity ties roughly to the expiry of the MYP option.

## 14. Option valuation

First we estimate the implied volatility of a Company A call option that expires close to the MYP expiry. Unfortunately options beyond two years are rare, even for established companies like A. Thus we use the Jan '10 call option to estimate the implied volatility. The parameters to estimate the implied volatility are listed in Table 6. S*, X*, T*, and c* are the stock price, strike price, expiry, and option price for the A Jan '10 $25 call. Using these values in the BS call option formula we can calculate the implied asset volatility[39]. The asset volatilities are then used in (2) to estimate the exchange option price for the MYP.

Table 6 summarizes the valuation of the MYP structured as an at-the-money exchange option. Setting the strike value equal to the spot value gives an option value of $127 million which the contractor would need to pay the government upon executing the MYP contract.

[39] We use an algorithm based on the Newton-Raphson method to solve for the implied volatility of a European option.

Much of this value is in the time to expiration or "time premium". Just to illustrate, if the option were for one month it would be worth $20 million and worth $4 million if it was for one day—all else equal.

| | |
|---|---|
| Risk Free Rate (r) | 4.73% |
| Stock Price (S*) | $ 26.15 |
| Exercise Price (X*) | $ 25.00 |
| Expiry (years) (T*) | 1 2/3 |
| Option Price (c*) | $ 5.40 |
| Asset Volatility | 29% |

Table 5. BS Parameters for A ($/Share)

| | ($ Millions) |
|---|---|
| Present Value SYP ($x_2$) | $ 630 |
| Strike Value ($x_1$) | 630 |
| Real Option Price | $ 127 |
| Expiry (yr) | 5.0 |

Table 6. MYP Evaluated as an Exchange Option – Risk on SYP Cash Flow Only

The analogy between MYP and insurance is a good one because, as anyone who has made a claim might have discovered, the insurance pay-off is not certain. The MYP can have a variation-in-quantity clause that allows the government to reduce the number of units by a pre-determined number. For example, if the EPA clause is activated by unanticipated labor and materials inflation, the government might reduce the quantity purchased to maintain its bottom line budget. Thus there is some uncertainty around the MYP that must be considered in our risk transfer pricing. This is where the exchange option framework has an advantage over the plain put option structure. It can be used to value cash flow trades that have different levels of uncertainty. For the valuation in Table 6 we set $\sigma' = \sigma_1$ and $\sigma_2 = 0$. Assume now the government and the contractor agree that the former could reduce the number of Program G units by 2 each year or 10% of the number of units in each lot. We use the exchange option structure to value the right to swap the SYP cash flow with volatility $\sigma_1$ for the MYP volatility of volatility $\sigma_2$ –see Table 7 for the valuation.

| | $ Millions |
|---|---|
| Present Value SYP ($x_2$) | $ 630 |
| Strike Value ($x_1$) | $ 630 |
| Real Option Price | $ 112 |
| Combined Volatility ($\sigma$) | 26% |
| SYP Volatility ($\sigma_2$) | 29% |
| MYP Volatility ($\sigma_1$) | 10% |
| SYP / MYP Volatility Correlation ($\rho$) | 50% |

Table 7. MYP Evaluated as an Exchange Option with Risk on Both Cash Flows

The price of the option falls from $127 million to $112 million. It would drop to $84 million with 100% correlation; however, if there were no correlation between the two cash flows, the

price would have increased to $134 million. This is due to the upside potential of the MYP and SYP. The exchange option is essentially a put option with a stochastic strike price which allows the protection buyer to capture more payoff if the MYP turns out to yield more units. This assumes that the risk of the MYP is symmetric. There is no reason to believe otherwise, since the government can always buy more units than planned, if they are needed.

## 15. Other real options embedded in an MYP

Within this chapter, we only have the scope to focus on a single real option example within the MYP contract. However, there is at least one other real option available to the contractor with a sole source production franchise such as a major aircraft, missile, ship, etc. This is because defense contracts are incomplete leaving the contractor with residual control of cost reduction innovations. While we will not estimate the value of this real option here, we mention it because in some cases it is potentially worth far more than the revenue stabilization discussed here.

Regulatory lag is an incentive concept that emerged from explicitly regulated industries such as utilities. These industries' profits are regulated directly through rate setting, e.g. $/kWhr, or through rate of return settings by a regulatory authority. Between rate settings, the utility is free to innovate and achieve higher profits. Upon the next regulatory oversight review, the regulator discovers the new cost structure and adjusts the new rate accordingly to a lower profit level-presumably slightly above the weighted average cost of capital for the utility. Longer periods between regulatory oversight periods (i.e. higher regulatory lag), mean greater opportunities for higher profits.

Similarly, a defense contractor with a sole source series of production contracts for a weapon system has the incentive to achieve greater than expected efficiency innovations even if the savings are passed on to the government in subsequent contracts. It turns out that there is a substantial regulatory lag in defense contracts due to the length of time it takes for cost reports to be submitted to the government. The regulatory lag increases substantially in a MYP contract.

These innovations are real options since the contractor is not obligated to make the necessary investment to achieve the cost savings. They can use real options valuation tool to estimate the worth of these options before a program is executed by looking at prior history of achieving cost reduction innovations as well as a forward looking assessment of the opportunities in a specific weapon system. Unlike the revenue stabilization option, there is considerable information asymmetry between the government and contractor with the regulatory lag options. However, the government could look at prior programs and assess the degree of regulatory lag driven innovation that occurred in past programs and roughly estimate the value of this type of incentive on a new program. This valuation can provide important insight into how aggressively contractors will compete to win a large sole source program.

## 16. Conclusion: the cost implications of the MYP option

Options pricing analysis offers a way to systematically estimate value from the MYP contract earned by the Government for which they have not previously been explicitly compensated. This incremental value is the revenue risk transferred to the Government from the contractor upon signing an MYP. The MYP does not eliminate the revenue risk for

the contractor associated with SYP contracts; rather it transfers it to the government and it becomes budget risk. The Congress clearly values its budget flexibility, as evidenced by the statutory criteria to judge the worth of an MYP proposal.

MYP cost savings are usually through operational efficiencies earned through process and purchasing improvements funded by the Government's "economic order quantity" advance funding. The transfer of revenue risk to the Government is a cash flow hedge that provides real value to the contractor just as any financial hedge does for currency, commodity, or interest rate risks--or property and casualty insurance does for operational risks. Lockheed and Raytheon, for example, carry interest rate swaps that hedge interest rate risk for notional $1 billion and $600 million respectively[40,41]. General Dynamics reported a currency swap to hedge a Canadian denominated loan with a fair value of $42 million[42]. It also reported embedded options in the terms of its long term labor and commodity contracts. One can argue that just as public companies are expected to incur expenses as they pay for insurance and financial hedges, they should pay the government when it reduces the contractor's risk.

The option methodology helps the government objectively quantify some of the cost in relinquishing its budget flexibility with a relatively simple tool that has widespread use in the financial community. We do not try to value the cost of transferring the risk from the Government's side because there is not a readily available tracking asset to estimate the volatility of the revenue risk. It is possible to estimate the actuarial loss history of certain procurements by looking at the Selected Acquisition Report over the span of past programs. If such data were available, it might be desirable to use it in lieu of the equity volatility of the contractor. One benefit of using the contractor's volatility, however, is that it is more closely coupled to the risk the contractor might be willing to hedge.

The option value of the MYP has not been explicitly paid to the government in the past. Thus any method that helps rationalize the cost of this risk transfer is a benefit to the government. Furthermore, the contractor will likely see the value of the MYP option if it is evaluated in its own financial terms.

Strategically, the MYP option value represents a significant reduction in the contractor's profits. Given the skill and sophistication that contractors employ to manage their government customers, they will likely argue that the MYP real option has limited value as an earnings hedge. They could contend that financial hedges are only appropriate for risks that are outside of managers' control, such as interest and exchange rates, and cannot be offset within the business. They might also contend that not only is their portfolio of business well diversified among a broad scope of government elements but that they have enough support on Capitol Hill to ensure that they will sell all the units in the SYP plan. They would be arguing that the program is less risky than their business in total (i.e. their equity volatility). This would be a difficult argument for most businesses. However, initially it is unlikely the contractors will proactively volunteer to pay for it.

---

[40] Lockheed Martin Corporation, Securities and Exchange Commission Form 10-K, Commission file number 1-11437, Fiscal Year December 31, 2006, p.71.

[41] Raytheon Company, Securities and Exchange Commission Form 10-K, Commission file number 1-13699, Fiscal Year December 31, 2006, p. 74.

[42] General Dynamics Corporation, Securities and Exchange Commission Form 10-K, Commission file number 1-13671, Fiscal Year December 31, 2006, p. 49.

However, the fact is that the lower earnings risk from an MYP has tangible value whether or not the contractors wish to pay for it. The option has the same value no matter what the contractors' risk preference. If there is no risk hedge in an MYP, why do the contractors routinely enter into this type of contract? In fact Lockheed readily acknowledged that the value of the MYP is its long term stability[43].

The options methodology allows the Government to build a logical business case for reducing the profit on cost paid to contractors when switching from an SYP series to an MYP contract. The exchange option model in particular allows the Government to quickly estimate changes in the value of the contract as the details, e.g. the EPA and VIQ clauses, become more complete.

## 17. References

Amram, M., and K. N. Howe (2003), "Real Options Valuations: Taking Out the Rocket Science," Strategic Finance, Feb. 2003, 10-13.

Baldi, F. (2005), "Valuing a Leveraged Buyout: Expansion of the Adjusted Present Value by Means of Real Options Analysis," J. Private Equity, Fall 2005, 64-81.

Barnett, M. L. (2005), "Paying Attention to Real Options," R&D Management 35, 61-72.

Black, F. and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," J. Political Economy 81, 637-654.

Charnes, J. M., and B. R. Cobb (2004), "Telecommunications Network Evolution Decisions: Using Crystal Ball and Optiquest for Real Options Valuation," Proc. 2004 Crystal Ball User Conference.

Colwell, D., T. Henker, J. Ho, and K. Fong (2003), "Real Options Valuation of Australian Gold Mines and Mining Companies," J. Alternative Investments, Summer 2003, 23-38.

Copeland, T., and K. M. Howe (2002), "Real Options and Strategic Decisions," Strategic Finance, April 2002, 8-11.

Copeland, T., and P. Tufano (2004), "A Real-World Way to Manage Real Options," Harvard Business Review, March 2004, 90-99.

Cornelius, P., A. Van de Putte, and M. Romani (2005), "Three Decades of Scenario Planning at Shell," California Management Review 48, 92-109.

Duan, C. W., W. T. Lin, and C. F. Lee (2003), "Sequential Capital Budgeting as Real Options: The Case of a New DRAM Chipmaker in Taiwan," Rev. Pacific Basin Financial Markets and Policies, 6, 1, 87-112.

Ekelund, A. (2005), "Valuating Biotech Project Portfolios Using Crystal Ball and Real Options-Case: Natimmune," Proc. 2005 Crystal Ball User Conference.

Ford, D. N., D. M. Lander, and J. Voyer (2002), "A Real Options Approach to Valuing Strategic Flexibility in Uncertain Construction Projects," Construction Management and Economics 20, 343-351.

Fourt, R. (2004), Risk and Optimal Timing in a Real Estate Development Using Real Options Analysis, Proc. 2004 Crystal Ball User Conference.

[43] LMT-Q3 2006 Lockheed Martin Earnings Conference Call, Preliminary Transcript, Thompson StreetEvents, Thompson Financial, October 24, 2006, 11:00AM ET

Gaynor, M. and S. Bradner (2001), "Using Real Options to Value Modularity in Standards," Knowledge, Technology and Policy 14, 2, 41-66.

Herath, H., and C. S. Park (2002), "Multi-Stage Capital Investment Opportunities as Compound Real Options," The Engineering Economist 47, 1-27.

IOMA (2001), "Real Options Analysis Creeps into Use at Intel, Genentech, and Texaco," IOMA's Report on Financial Analysis, Planning & Reporting 01-12, 5-11.

Janney, J. and G. G. Dess (2004), "Can Real-Options Analysis Improve Decision-Making? Promises and Pitfalls," Academy of Management Executive, 18, 4, 60-75.

Kayali, M. M. (2006), "Real Options as a Tool for Making Strategic Investment Decisions," J. American Academy of Business 8, 282-286.

Laxman, P. R. and S. Aggarwal (2003), "Patent Valuation Using Real Options," IIMB Management Review, Dec. 2003, 44-51.

MacMillan, I., A. B. van Putten, R. G. McGrath, and J. D. Thompson (2006), Using Real Options Discipline for Highly Uncertain Technology Investments," Research-Technology Management, Jan.-Feb. 2006, 29-37.

Maumo, D. A. (2005), "HP Applies Risk Management to Procurement," Manufacturing and Business Technology, Nov. 2005, 26-27.

Nembhard, H. B., L. Shi, and M. Aktan (2003), "A Real Options Design for Product Outsourcing," The Engineering Economist, 41, 3, 199-217.

Oppenheimer, P. H. (2002), "A Critique of Using Real Options Pricing Models in Valuing Real Estate Projects and Contracts," Briefings in Real Estate Finance 2, 3, 221-233.

Paxson, D. (2002) (ed.) "Real R&D Options," Butterworth Heinemann, 333pp.

Remer, S., S. H. Ang, and C. Baden-Fuller (2001), "Dealing with Uncertainty in the Biotechnology Industry: The Use of Real Options Reasoning," J. Commercial Biotechnology 8, 2, 95-105.

Rigby, D. (2001), "Management Tools 2001-Global," Bain and Co., Boston, Massachusetts, June, 2001.

Rigby, D. and B. Bilodeau (2005), "Management Tools and Trends 2005," Bain and Co., Boston, Massachusetts.

Rothwell, G. (2006), "A Real Options Approach to Evaluating New Nuclear Power Plants," Energy Journal 27, 37-53.

Synergy Partners (2003), "Real Options Primer," Synergy Partners, Greensboro, North Carolina.

Teach, E. (2003), "Will Real Options Take Root?" CFO, July 2003, 73-76.

van Putten, A. B., and I. C. MacMillan (2004), "Making Real Options Really Work," Harvard Business Review, Dec. 2004, 134-141.