Proceedings of the
14th Australian Space Research Conference
Adelaide
September 29 - October 1, 2014

Australian Space Research Conference Series

1st Edition
Published in Australia in 2015 by
National Space Society of Australia Ltd
GPO Box 7048
Sydney NSW 2001
Fax: 61 (02) 9988-0262
email: nssa@nssa.com.au
website: http://www.nssa.com.au

ISBN 13: 978-0-9775740-8-7
Editors: Wayne Short and Iver Cairns

Distributed on DVD

Preface to the Proceedings

A large number of the presenters at the conference later submitted completed written papers which form the basis of this Conference Proceedings.

All papers published in these proceedings have been subject to a peer review process whereby a scholarly judgement by at least two suitable individuals endorsed by the Program Committee determined if the paper was suitable to be published. All papers not rejected were revised until deemed suitable. Responsibility for the content of each paper lies with its author(s). The publisher retains copyright over the text. Papers appear in the Conference Proceedings with the permission of the authors.

The Editors would like to give special thanks to the Program Committee and those scholars who participated in the peer review process:
Brad Alexander, Simone Atzori, Jeremy Bailey, Phil Bland, Russell Boyce, Graham Brooker, Melrose Brown, Ian Bryce, Graziella Caprarelli, Sarah Chamberlain, Brad Carter, Brett Carter, Iain Cartwright, Jon Clarke, Lucyna Chudczer, Michael Chang, Andrew Dempster, Boyin Ding, Alina Donea, Graham Dorrington, Brian Fraser, Bruce Forster, Eamonn Glennon, James Gilmore, Li Guo, Jason Held, Kasper Johansen, Charlie Ji, Jon Kim, David Lingard, Xi Li, Samsung Lim, Patryk Sofia Lykawka, Ken Lynn, Sarah Maddison, Jonathan Marshall, Fred Menk, Marcello de Michele, Dennis Odijk, Jonathan Paxman, Lily Qiao, Peter Rayner, Stuart Ryder, Samira Tasnim, Matt Tetlow, Paul Tregoning, John Trinder, Matthew Trinckle, Sean Tuttle, Gordon Xu, Colin Waters, Robert Wittenmyer, Xiaofeng Wu, Joel Younger, Kegen Yu, Kefei Zhang.

Finally we would like to thank our sponsor University of South Australia for its support in funding the venue; the Organising Committee, the Program Committee and colleagues Graziella Caprarelli, Alice Gorman, Trevor Harris and Garry Newsam for giving generously of their time and efforts.

We trust that you will find the 2014 Conference Proceedings enjoyable and informative.

Wayne Short and Iver Cairns
Editors, 14ASRC Conference Proceedings, May 2014

Conference Background

The Australian Space Research Conference (ASRC) is the focus of scientific cooperation and discussion in Australia on research relating to space. It is a peer reviewed forum for space scientists, engineers, educators, and workers in Industry and Government.

The conference is of relevance to a very broad cross section of the space community, and therefore generates an enlightening and timely exchange of ideas and perspectives. The scope of the conference covers fundamental and applied research that that can be done from space and space-based platforms, and includes the following:

- **Space science**, including space and atmospheric physics, remote sensing from space, planetary sciences, astrobiology and life sciences, and space-based astronomy and astrophysics ;
- **Space engineering**, including communications, navigation, space operations, propulsion and spacecraft design, testing, and implmentation ;
- **Space industry ;**
- **Space archeology and Indigenous sky knowledge;**
- **Space situational awareness**
- **Current and future Australian space projects ;**
- **Government, International relations and law ;**
- **Education and outreach.**

The 14th ASRC was held at the University of South Australia (City-West Campus) during September 29 to October 1, 2014. The Conference was opened by Pro Vice-Chancellor for Research of University of South Australia (Research), Professor Richard Head.

The conference included a comprehensive program of plenary talks and special sessions on the national context for space the foci and programs of Australian Government units with interests in space. In addition, the program contained a special planetary science session, with a strong preponderance of projects involving the Mars Society of Australia. The program also ontained multiple sessions of invited and contributed presentations, both oral and poster, on Propulsion, Planetary Science, Earth Observation and GNSS, Space Capabilities, Space Physics, Space Technology, Space Archeology, Education and Outreach and Indigenous Sky Knowledge.

Appendix A lists all abstracts accepted for presentation at the conference.

The 14th ASRC was organised by the National Space Society of Australia (NSSA), the Academy of Sciences National Committee for Space and Radio Science (NCSRS) and University of South Australia. The Mars Society Australia (MSA) and Australian Space Research Institute (ASRI) also helped significantly with organising abstract submissions.

A call for papers was issued in March 2014 and researchers were invited to submit abstracts for presentation at the conference. Following the conference itself, a call for written papers was issued in October 2014: this invited presenters to submit a formal written paper for this Proceedings that covered their conference presentations.

Table of Contents

List of Proceedings Papers

List of Proceedings Papers

Welcome to the 14^{th} Australian Space Research Conference

formerly the Australian Space Science Conference (ASSC), and to the University of South Australia, Adelaide! This was the eighth conference jointly sponsored and organised by the National Committee for Space and Radio Science (NCSRS) and the National Space Society of Australia (NSSA). The ASRC is intended to be the primary annual meeting for Australian research relating to space science. It welcomes space scientists, engineers, educators, and workers in Industry and Government.

The 14th ASRC had over 130 accepted abstracts across Australian space research, academia, education, industry, and government.

We would like to thank the University of South Australia for their sponsorship. Special thanks also go to the Mars Society Australia (MSA) and Australian Space Research Institute (ASRI) for their support.

Iver Cairns
Co Chair ASSC 2014
University of Sydney

Wayne Short
Co Chair ASSC 2014
President, NSSA

The National Space Society of Australia is the coming together of like-minded space enthusiasts who share a vision for the future in which there is an ambitious and vigorous space program leading to eventual space settlement.

To this end the National Space Society (worldwide) promotes interest in space exploration, research, development and habitation through events such as science and business conferences, speaking to the press, public outreach events, speaking engagements with community groups and schools, and other pro-active events. We do this to stimulate advancement and development of space and related applications and technologies and by bringing together people from government, industry and all walks of life for the free exchange of information.

As a non-profit organisation, the National Space Society of Australia draws its strength from an enthusiastic membership who contributes their time and effort to assist the Society in pursuit of its goals.

For more information, and to become a member:

http://www.nssa.com.au

Ad Astra!
Wayne Short
NSSA President

AUSTRALIAN ACADEMY OF SCIENCE

The National Committee for Space and Radio science Science (NCSRS) is chartered by the Australian Academy of Science to foster space science, to link Australian space scientists together and to their international colleagues, and to advise the Academy's Council on policy for science in general and space and radio science in particular. The NCSRS was formed in 2012 by combining the former National Committee for Space Science (NCSS) and the National Committee for Radio Science (NCRS). The NCSRS web page can be reached at

https://www.science.org.au/committee/space-and-radio-science

NCSRS believes that ASRC meetings provide a natural venue to link Australian space scientists and foster the associated science, two of its core goals. As well as ASRC, NCSRS is also sponsoring the VSSEC – NASA Australian Space Prize.

This is the fourth ASRC meeting following launch of the first Decadal Plan for Australian Space Science. NCSRS encourages people to work together to accomplish the Plan's vision: "Build Australia a long term, productive presence in Space via world-leading innovative space science and technology, strong education and outreach, and international Collaborations."

2014 ASRC Program Committee

Russell Boyce (University of New South Wales, Canberra)

Iver Cairns (Chair, University of Sydney)

Graziella Caprarelli (University of South Australia)

Jonathan Clarke (Mars Society of Australia)

Andrew Dempster (University of New South Wales)

Alice Gorman (Flinders University)

Duane Hamacher (University of New South Wales)

Trevor Harris (Defence Science Technology Organisation)

Isabelle Kingsley (Powerhouse Museum)

Alexey Kondyurin (University of Sydney)

John Le Marshall (Bureau of Meteorology)

Carol Oliver (University of New South Wales)

Chris Rizos (University of New South Wales)

Richard Samuel (Australian Space Research Institute)

Murray Sciffer (University of Newcastle)

Michael Smart (University of Queensland)

Collette Snowden (University of South Australia)

Martin Westwell (Flinders University)

Kefei Zhang (RMIT University)

2014 ASRC Organising Committee

Iver Cairns
Co Chair ASRC 2014
University of Sydney
Chair, ASRC Program Committee

Jeremy Bailey
University of New South Wales

Graziella Caprarelli
University of South Australia

Jon Clarke
Mars Society Australia

Andrew Dempster
University of New South Wales

Alice Gorman
Flinders University

Trevor Harris
Defence Science and Technology Organisation

Garry Newsam
Defence Science and Technology Organisation

Wayne Short
Co Chair ASSC 2014
President, NSSA

Kefei Zhang
RMIT University

Conference Plenary Speakers

Philip Bland
Curtin University
"The Desert Fireball Network: A continent-scale observational facility serving the planetary science, space debris, and astronomy research communities"

Andrew Brawley
Silanna
Silanna: Manufacturing in Australia High-Reliability Integrated Circuits for Space Applications

Brett Carter
RMIT University
"Severe space weather events and their impact on our technology-dependent society"

Linda Davis
Institute of Telecommunications Research (ITR), UniSA
"6S Success: Satellite Services, Systems, Spectrum, Software and Signals"

Conference Plenary Speakers

Lucyna Kedziora-Chudczer
University of NSW
"Diversity of Planetary Atmospheres: New results and Open Questions"

Helen Maynard-Casely
Bragg institute, ANSTO
"Probing the planets with crystallography"

Peter Ward
University of Adelaide
"Astrobiology and planetary science: New results and Open Questions"

Kefei Zhang
RMIT University
"CRC SEM - a new horizon of Australian space tracking research"

14th ASRC Conference Program
September 29 – October 1

Time	Monday		Tuesday		Wednesday	
7:30	Registration					
8:00			Registration		Registration	
9:00	Opening / Plenaries *Alan Scott Auditorium*		Plenaries *Bradley Forum*		Plenaries *Bradley Forum*	
10:30	Break		Break		Break	
11:00	National Context *Alan Scott Auditorium*		Space materials, Structures & Devices *Bradley Forum*	Education, Outreach & Culture *Room H6-12*	Space Engineering *Bradley Forum*	Space Physics 3 *Room H6-12*
13:00	Lunch		Lunch	Workshop - H6-03	Lunch	
14:00	Space Physics 1 *Bradley Forum*	Space Engineering & Projects *Room H6-12*	Space Physics 2 *Bradley Forum*	Mars - Planetary Science & Engineering *Room H6-12*	Space Situational Awareness *Bradley Forum*	Planetary Science & Astrobiology *Room H6-12*
16:00	Break		Break		Break	
16:30	Posters & Cocktail Reception *Bradley Forum*		Remote Sensing *Bradley Forum*	Education, Outreach & Space Engineering *Room H6-12*	ATV-5 re-entry workshop *Bradley Forum*	GPS & GNSS *Room H6-12*
18:00			Close		Closing Remarks	
18:15	Close				Close	
18:30						

Analysis of Greenhouse Gas Concentrations Retrieved from Observations with a Ground-based Fabry-Perot Spectrometer in the Near-infrared

Champlain Kenyi, Daniel V. Cotton and Jeremy Bailey

School of Physics, The University of New South Wales, NSW 2052, Australia.

Summary: The remote sensing of atmospheric greenhouse gas concentrations from space platforms is a vital measurement technique applied in a number of space-based missions. However, obtaining accurate measurements from space is complicated by clouds and aerosols, which attenuate an already weak reflected signal. Ground-based total column measurements can largely avoid such difficulties by direct pointing of the instrument at the Sun; and can thus complement and help to validate the space instruments. Nevertheless, coverage of current ground-based networks is still poor in the developing countries and over the ocean.

We have built a compact, portable and low-cost fibre Fabry-Perot spectrometer that could be deployed in remote locations and used to establish more extensive network of greenhouse gas monitoring surface sites. This is a ground-based near-infrared instrument for measuring the absorption lines of target trace gases in the atmosphere, which attenuate the solar spectrum. For data reduction we apply a nonlinear least-squares fitting algorithm in conjunction with an efficient line-by-line radiative transfer model to retrieve the atmospheric properties of interest. Our main goal here is to report progress on the development of the software to retrieve estimates of the atmospheric constituents (e.g. XCO_2 – the column-averaged dry air mole fraction of CO_2). We identify some of the issues that remain to be overcome.

Introduction

Carbon dioxide (CO_2) and methane (CH_4) are the most abundant anthropogenic greenhouse gases in the Earth's atmosphere; increases in their concentrations are driving global climate change [1]. To help arrest this trend, we need to understand and characterise the global carbon budget – in particular the sizes of the sources and sinks need to be established. The ability to infer the surface carbon fluxes depends on interpreting spatial and temporal variations of atmospheric gases and relating them back to surface fluxes [2]. Currently the CO_2 fluxes, for instance, are inferred from in-situ observations made at a network of ground-based monitoring platforms [3]. Studies by Heimann and Kaminski [4] however demonstrate that inverse modelling based on in-situ data suffers drawbacks due to its sparse network and limited coverage, and because errors in the model prediction of advective transport and mixing are aliased into flux estimates [2, 3].

Space missions and instruments such as OCO-2, MOPITT, AIRS, IASI, and GOSAT, can provide the necessary global coverage. However achieving the necessary signal quality, especially over the ocean, is a challenge [5, 6] and probably unlikely to be achieved in the near future. This is because satellites that depend on reflected light (such as OCO-2 and GOSAT) receive only a weak signal over the ocean.

Ground-based total column measurements, such as via the near-infrared high resolution Fourier Transform Spectrometers (FTSs), common at the TCCON (Total Carbon Column Observing Network) locations are a useful part of observing systems. The accurate gas column densities measured with such instruments provide useful constraints for inverse modelling when combined with simultaneous co-located in-situ surface measurements [3]. The principal challenge is that the FTSs are bulky and expensive to maintain; they also require a stable laboratory platform and dedicated air-conditioned space [7, 8]. Attempts to address these issues by a number of authors including Kobayashi et al.[9], Wilson et al. [10] and Petri et al. [11] are reported by us elsewhere [7, 12, 13]. Petri et al. [11] and Gisi et al. [14] have also described the use of smaller and hence lower resolution Fourier-Transform spectrometers, while Kawasaki et al. [8] describes the use of an Optical Spectral Analyser (OSA) based on a grating spectrometer.

The CO_2 column measured by such an instrument varies with changes in topography and atmospheric pressure, as well as with changes in the actual CO_2 concentration. These effects can be corrected by making a simultaneous measurement of the O_2 column. Dividing the measured CO_2 column by the O_2 column, and using the known fixed mixing ratio of O_2 in dry air, provides a corrected measurement of the CO_2 concentration. This method of correction is used with both space [15] and ground based [16] instruments.

The focus of this paper is to demonstrate progress toward making such measurements reliably with a low cost, compact and portable instrument [7]. Such an instrument could be deployed in remote locations, and/or be used for Ground Validation (GV) – i.e. help to calibrate trace gas column densities/measurements from space missions. The instrument investigated here is built locally from commercial optical devices; it is composed of two fibre Fabry-Perot (FP) spectrometers, one designed for CO_2 detection, and the other for measuring the O_2 a-X band at 1.27 μm. These spectrometers measure total column concentrations by recording absorption lines in the solar spectrum due to the respective gas (as dictated by the filter chosen for the application). While all greenhouse gases are of interest, the main focus here is on atmospheric CO_2, with H_2O retrieved as a contaminant. The CO_2 and O_2 are measured at near infrared bands centred on 1.58 and 1.27 μm.

TCCON measures near-infrared solar absorption in the region 4000-14000 cm^{-1} at 0.02 cm^{-1} resolution. Spectra are analysed to retrieve CO_2, CO, CH_4, N_2O, HF and O_2 column densities and column average mixing ratios [3, 17]; the measurements are highly precise and accurate and can be used to calibrate and validate satellite measurements [2]. TCCON measurements can also provide useful constraints to calibrate other instruments when measurements are made concurrently in the same location [3]. We have made observations with our instrument at the TCCON site at the University of Wollongong (UoW) simultaneously with that instrument. This paper builds on our previous work published in these proceedings [12, 13], by including simultaneous O_2 measurements and a direct comparison with the TCCON data from the same site. We present comparisons between TCCON retrievals, results obtained with our instrument and software, as well as retrievals calculated from TCCON spectra using our software.

Modelling of measured spectra

Our forward model and procedure for the inverse modelling are described elsewhere [12, 13, 18]. Briefly, the inverse modelling software – ATMOF (ATMOspheric Fitting) [18] – uses a Levenberg-Marquardt non-linear least squares fitting routine in conjunction with an efficient line-by-line radiative transfer model. The molecular absorption line list for calculating the atmospheric transmission is taken from the HITRAN 2008 database [19]. As an *a priori* model we use the mid-latitude summer profile which is one of the standard atmospheric profiles from the ICRCCM (InterComparison of Radiation Codes in Climate Models) [20] and is one of half a dozen built-in to the VSTAR (Versatile Software for the Transfer of Atmospheric Radiation) radiative transfer code [21, 22]. However, these are not the same as the TCCON *a priori* profile; TCCON uses the European Centre for Medium-range Weather Forecasting (ECMWF) a priori for their model [17].

Input parameters of the inverse model

Our inversion algorithm has an input text file, where we specify each spectrum to be fitted (one at a time) and the variable parameters of the fit. The parameters include those listed in Table 1, together with a pre-determined numerical filter function in the case of the Fabry-Perot measurements. Each parameter may either be fit or held constant in the fitting procedure. The greater the number of parameters fit, the longer the code takes to run and the greater the potential for an unrealistic solution. The choice of which parameters to fit is made based on a number of factors, including how well each parameter is known, whether it is likely to have deviated from default values and the likely influence it has on the spectrum and the other retrieved parameters. Below and in Table 1 we detail our approach, through the listed initial conditions, in fitting the various different types of spectra.

Table 1: Input parameters for the inverse modelling.

Input Parameter	FP		FTS		Comments
	CO_2 band	O_2 band	CO_2 band	O_2 band	
Wavelength Shift (m)	-3.7600E-13	3.7600E-15	2.8989E-13	1.8697E-14	Held constant.
Doppler Shift (m.s^{-1})	100	100	100	100	Fitted.
SZA (º)	variable	variable	variable	variable	Fixed for each spectrum.
CO_2 (ppm)	395	393.5	395	393.5	Fitted only in CO_2 band.
H_2O (pp10,000)	70	70	70	70	Fitted.
O_2 (pp10,000)	2095	2095	2095	2095	Fitted only in O_2 band.
IRS: Scaling Factor	18.702	18.702	2.5	2.5	Fitted.
IRS: Slope (m^{-1})	-3.5454E-05	1.6720E-03	1.9528E-03	1.5715E-03	Investigated then fixed.
~R ($\lambda/\Delta\lambda$)	70000	90000	382500	382500	Held constant.
Angular Radius (mm)	NA	NA	1.2	1.2	Held constant.

"Held constant" means the parameter assigned for either spectrum is not allowed to vary. The spectral resolution of the Fabry-Perot tuneable filter was initially calculated based on information provided by the manufactures and then investigated further by comparing real spectra with those modelled at different resolutions. The values differ significantly for the FP-O_2 band, where we find a performance better than that given by the manufacturer. Once determined, the resolutions were fixed at a constant value across the whole spectrum. If fit they can fluctuate between different spectra and render the retrievals senseless. The spectral resolving power of the FTS is given by multiplying the maximum optical path length (45 cm) of the interferometer by the highest scanned wavenumber (8500 cm^{-1}).

The Solar Zenith Angles (SZAs) are determined using a FORTRAN software package that takes account of locality and time of the observation. In contrast to the FTS spectra, which take about 3 minutes each to acquire, the scan time for the FP spectra presented here is about 13 minutes, which is significant, and can contribute to airmass dependent artefacts at large SZA. Consequently, we are careful to use the average time of each scan rather than simply the start time.

The observer-Sun Doppler shift can be calculated precisely for a given time and location. However, the pointing of the telescope is not precise enough to always pick out the centre of the Sun, thus variability can be introduced through the contribution of stellar rotation. Consequently, we also fit this parameter.

The Instrument ReSponse (IRS) Scaling Factor and Slope are essential inversion parameters of the model. The scaling factor is fitted individually to each spectrum. The slope is adjusted by examining the residuals of fits to solar spectra, and once determined is set to a fixed value for each wavelength region observed. The physical interpretation of these parameters is the varying sensitivity of the detector with wavelength. A wavelength shift correction is determined in the same way; in essence here we are using the well-known positions of the telluric atmospheric lines to make a correction to the wavelength calibration of the instrument.

Measurement Errors

Errors in these measurements originate from random, known and unknown systematic sources. The random effects manifest in unpredictable variability of the measurements. Systematic errors result in measurements consistently being too large or too small – they cannot be removed with averaging [23]. Such errors emerge when there is a problem with the measuring instrument (e.g. calibration issues) or its data handling system, or through incorrect use of the instrument. Some of the systematic effects arise specifically from the *a priori* profile as a result of errors in estimating the pressure, temperature, as well as the volume mixing ratios of the gas abundances (CO_2, O_2, and H_2O). The possibility of solar-tracker pointing offset and distortion of the instrument line shape caused by shear or angular misalignments (as for example, discussed by Wunch et al. [17]) cannot be ruled out of the systematic contributions. Some of these effects are common to both spectrometers and cancel out in ratioing the column concentrations [17, 24] as described in the Introduction. A detailed breakdown of the error sources are beyond the scope of this paper and are the subject of on-going work which we plan to present in the future.

Our uncertainty estimates for each fitted parameter are given, in the first instance, by the square root of the diagonal entries of the respective fitted parameter in the inverse vector as calculated by ATMOF. For CO_2 and H_2O this gives the uncorrected column densities in ppm and pp10,000 respectively. But for the normalised column concentrations XG (also known as column-averaged dry-air mole fraction) for gas G, the error is given as:

$$(XG)_{err} = XG \times \sqrt{\left(\frac{\Delta G}{G}\right)^2 + \left(\frac{\Delta O_2}{O_2}\right)^2} \tag{1}$$

where $XG = \frac{[G]}{[O_2]} \times 0.2095$

Here ΔG is the error uncertainty of the target gas while ΔO_2 the uncertainty in the O_2 column density; all being the square roots of the diagonal entries. The global average O_2 concentration is 0.2095.

It should be noted that the errors we calculate represent the random errors for the measurement, not the systematic ones. By comparing our measurements with the TCCON measurements, we were able to detect systematic effects that were not otherwise showing up in our observing system.

Retrieved Columns

Retrieval Software

In this section we present column densities retrieved with ATMOF/VSTAR and GFIT. ATMOF [18] is an atmospheric fitting routine that makes use of VSTAR [7, 21] atmospheric models. Since our last communication we have made some improvements to our VSTAR model and our ATMOF fitting routine. We are now able to fit O_2 concentration (in addition to CO_2 and H_2O). We can also now use the instrumental line shape of a FTS (for fitting TCCON spectra) as well as the Airy function line shape needed for fitting our own FP spectra. GFIT [17] is the spectral fitting routine commonly used by the TCCON team to fit their high resolution measurements recorded with the FTSs.

Observations

The spectra used here were measured with our FP instrument and with the TCCON FTS instrument on the 12th and 17th of July 2013 at the TCCON observatory located at the UoW, where our spectrometers were connected to the TCCON solar tracking heliostat for simultaneous independent measurements. Wollongong is a coastal city, south of Sydney in New South Wales, Australia.

Results and Discussion

Three types of retrievals are presented in this section: (1) Those obtained through analysis of FP spectra with ATMOF/VSTAR; (2) Those obtained through analysis of FTS spectra with ATMOF/VSTAR; and (3) Those produced by GFIT from analysis of FTS spectra.

Fig. 1 and Fig. 2 show, respectively, changes in CO_2 column concentrations in ppm with time of observation without (Fig. 1) and with (Fig. 2) normalisation by the simultaneously observed O_2 column concentrations. In Fig. 2, the normalised CO_2 concentrations (XCO_2) are obtained by ratioing the CO_2 retrieval obtained from the CO_2 band, shown in Fig. 1, by the O_2 concentration retrieved from the corresponding O_2 band spectra. The agreement between the FP retrievals by ATMOF/VSTAR and the GFIT retrievals of the FTS spectra is reasonable. Though there is a clear trend in the FP retrievals from the 12th, and a less obvious trend on the 17th in Fig. 3, investigation of different slopes leads us to believe that this trend may be caused by the long time taken – 13 minutes – to complete a scan. An appropriate slope was chosen for the data set as a whole, but the airmass change over the duration of the scan is different at different times of day. At overhead sun the net airmass change is zero, but at 5 pm (the latest scan on the 17th) the change is 0.27. Another possible cause of gradients in the retrieved column densities, particularly for spectra observed at the beginning of the day, is the "creep" effect in the piezoelectric transducers of the FP as discussed by Bailey [7].

The main area of disagreement in the presented retrievals is in the XCO_2 values obtained by ATMOF/VSTAR from the FTS spectra – the retrieved values are much too low. This comes about as a result of a high O_2 retrieval (Fig. 3) – 5% higher – from the TCCON raw data. We have, as yet, been unable to properly take account of the FTS's instrumental transmission in this band and believe this to be the cause of the discrepancy. Specifically, the FTS uses at least one optical filter for which we do not have the transmission function. However, we do not have full details of the TCCON data reduction procedure and there may be other causes we are unaware of.

In Fig. 4 the water vapour column concentrations increase over the course of observing in all the retrieval formats. We see no significant difference between the corrected and uncorrected concentrations in the FP-VSTAR fitted spectra as opposed to FTS-VSTAR fits where the corrected values are clearly lower. This is due to the inaccurate O_2 column densities retrieved for reasons stated earlier.

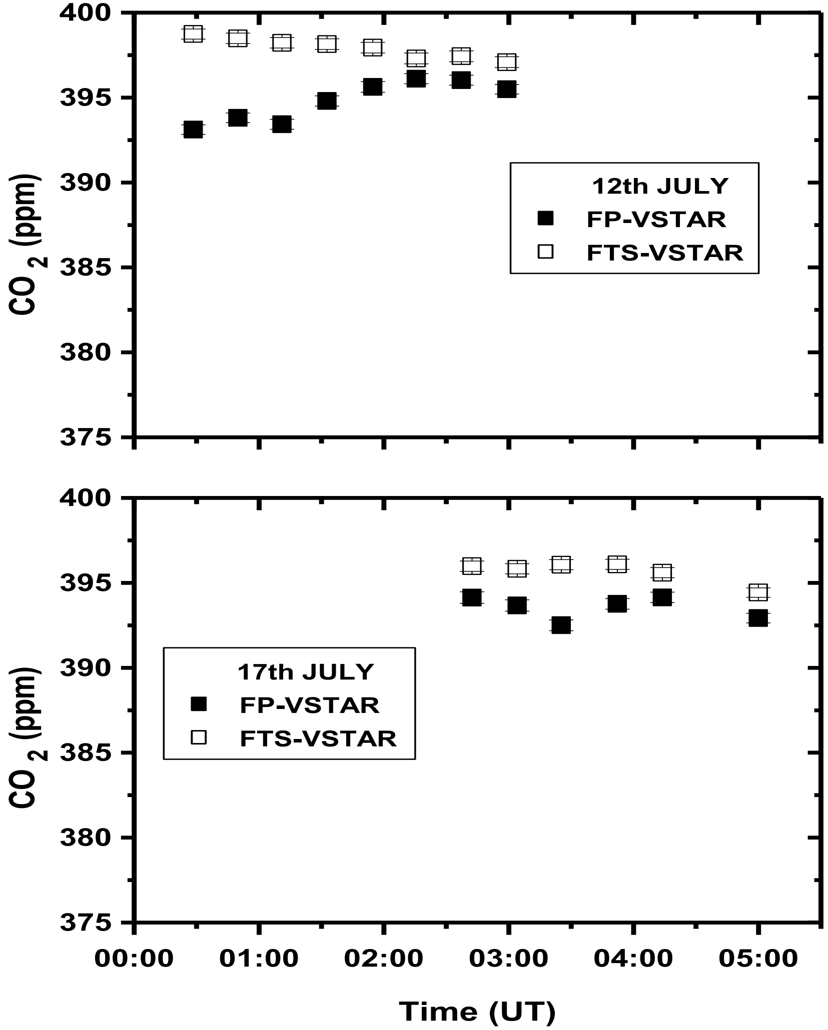


Fig. 1: The CO_2 mixing ratios retrieved before normalising with the O_2 column density for the observations made on the 12th (top) and 17th (bottom) of July 2013. Gap between the horizontal bars indicate the error in CO_2.

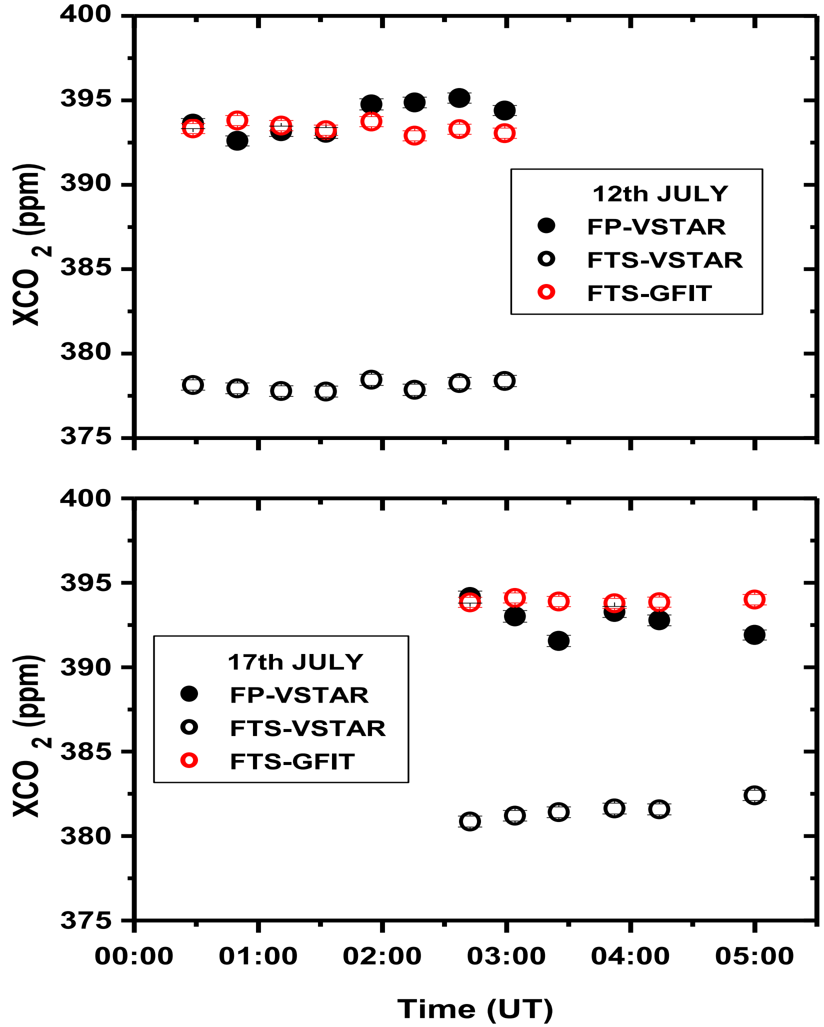


Fig. 2: Retrieved CO_2 mixing ratios (XCO_2) after normalising with the O_2 column density for the observations made on the 12th (top) and 17th (bottom) of July 2013. Gap between the horizontal bars indicate the error in XCO_2.

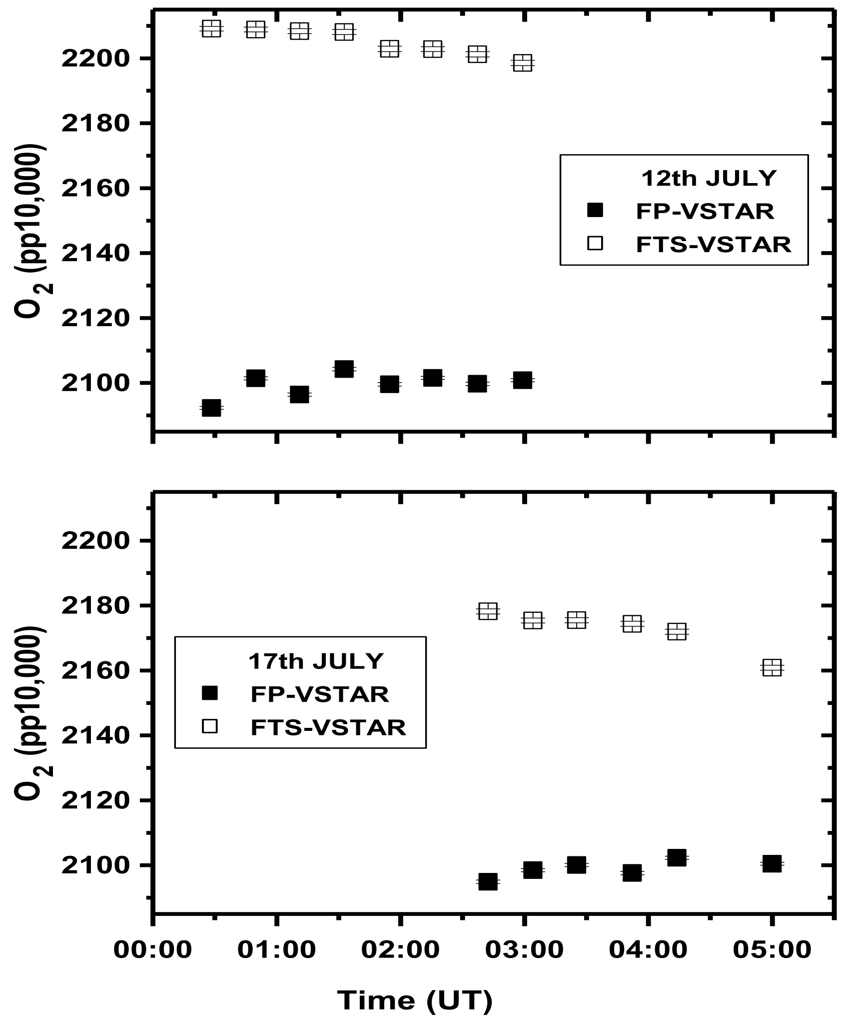


Fig. 3: The O_2 mixing ratios retrieved from the O_2 X-a band for observations made on the 12th (top) and 17th (bottom) of July 2013. Gap between the horizontal bars indicate the error in O_2.

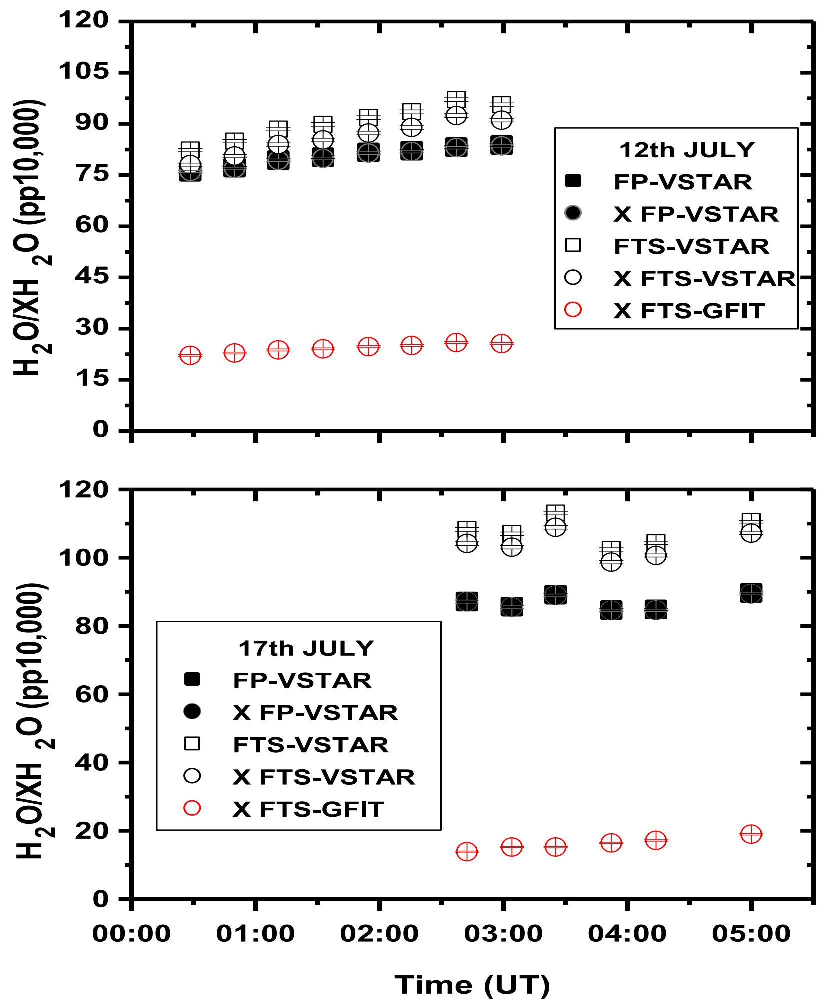


Fig. 4: A comparison of the corrected (XH_2O) and uncorrected (H_2O) mixing ratios retrieved from fitting in the CO_2 band for observations on the 12th (top) and 17th (bottom) of July 2013. Gap between the horizontal bars indicate the error in H_2O/XH_2O.

Aside from the GFIT determined XH_2O, large biases exist between the VSTAR calculated columns for spectra of different resolutions. This seems to occur because the FP and FTS spectra modelled have different assigned inputs, some of which were fixed (see Table 1). These VSTAR normalised retrievals, especially H_2O, are by far larger than the corresponding GFIT retrievals. Our *a priori* model (mid-latitude summer profile) is fairly simple and consists of 50 levels from 0 to 120 km high. The H_2O *a priori* volume mixing ratios (VMRs) are scaled from 50 km altitude to the surface, while for the CO_2 scaling is done for all levels; this is not as sophisticated as the (ECMWF) *a priori* profiles used in the TCCON GFIT in terms of temperature, pressure and water vapour profiles [19]. Because our instrument primarily measures the total column of an absorber and is not sensitive to the altitude where the absorption occurs, and because there is very little overlap between H_2O and CO_2 lines in our spectra the retrievals are not sensitive to the adopted H_2O profile. Our error budget analysis showed uncertainties for XCO_2 due to this error source contribution were all under 0.0035 ppm for spectra observed on the 12th and 0.0118 ppm for those observed on the 17th.

The gradients in the retrieved columns are an indication that the atmosphere was changing on both days of the observations – which always is the case – and that has an effect on atmospheric pressure as well as the uncorrected columns. So these could be some of the reasons accounting for the noticeable biases.

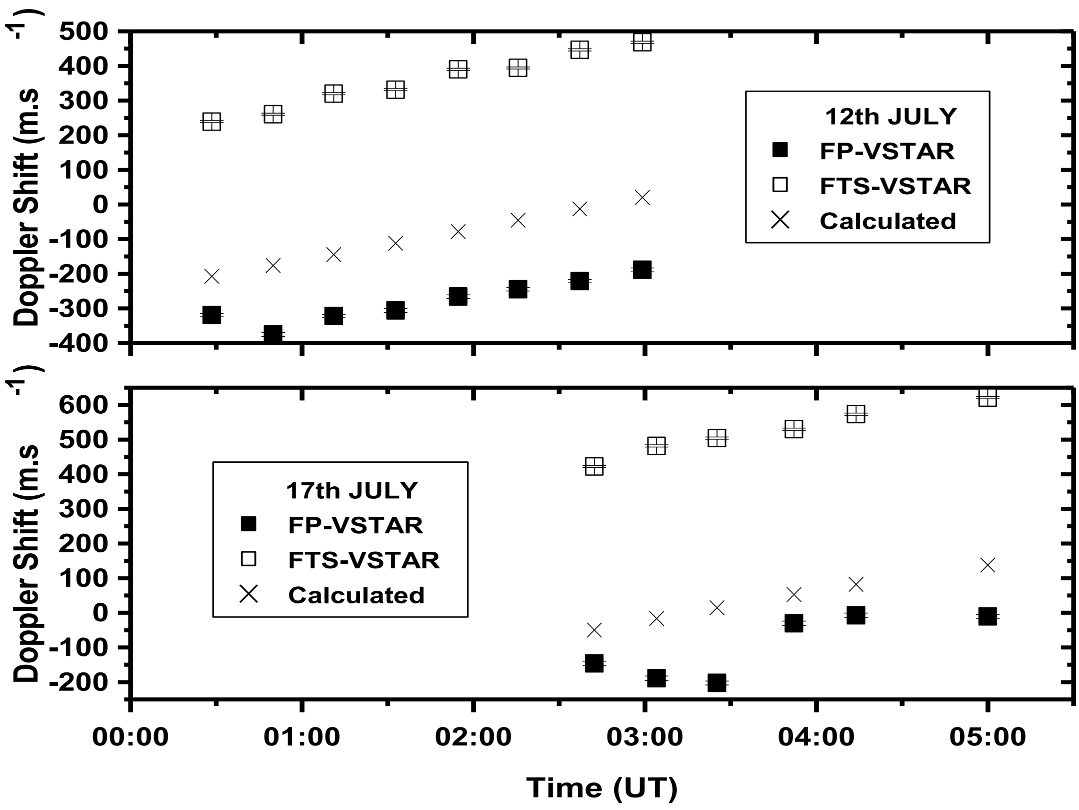


Fig. 5: Variation of retrieved and calculated Doppler shifts for observations made on the 12th (top) and 17th (bottom) of July 2013. Gap between the horizontal bars indicate the error in Doppler shift.

Fig. 5 shows how the fitted solar Doppler shift varies with the time of the observations. The calculated Doppler shifts represent the precise values for the Sun-observer system at the mean time of observation. The trend in Doppler shift is approximately linear over the duration of the observations. The Doppler shift retrievals from the FTS spectra mirror the linear trend, and indeed have a very similar slope, though a significant offset. The FP retrieved Doppler shift trend for the observations made on the 12th are also linear, after the first hour. The trend is not so linear for the observations made on the 17th however.

Offsets and variations in the retrieved Doppler shift could be due to deficiencies in the wavelength calibration of the spectra. In general this is not a serious problem, as it will be corrected in the fitting process. A more serious problem is that additional Doppler shifts could be caused by a pointing error that would introduce a velocity component due to solar rotation. In this case our SZA would also be affected and this could introduce errors in the retrieved CO_2 column. At present there is no obvious correlation between Doppler shift deviations and retrieved CO_2 column that would indicate such a problem.

Conclusion and Future Directions

We have made improvements to the VSTAR version for the Earth atmospheres transmission, which when used in conjunction with the atmospheric fitting routine can retrieve atmospheric properties from measurements. The XCO_2 estimates from the Fabry-Perot instrument differ from the TCCON GFIT calculated benchmark by less than 2.0 and 2.4 ppm for measurements recorded on the 12th and 17th of July 2013, respectively, which overall is a reasonable agreement.

This paper improves on our previous work published in this proceedings series [12, 13] by making use of simultaneous CO_2 and O_2 measurements. This procedure minimises systematic and correlated errors present in both the retrieved CO_2 and O_2 columns (e.g. pressure errors, instrumental line shape, sun tracker pointing errors, [17, 24]). Though observations from two days are used for convenience, we find the normalised (XCO_2) column densities comparing well as opposed to the uncorrected CO_2 columns. The main issue is with the O_2 high-resolution spectra (TCCON raw data) fitted with our software, predicting O_2 ~5% higher than what the current atmosphere is composed of (~21%). Consequently, the resulting XCO_2 are too low relative to the TCCON GFIT calculations. We attribute this to the lack of an instrumental filter response for the FTS in the ATMOF model. This will be addressed accordingly in the near future so we can locate any limitations within our software.

Comparison with TCCON data has highlighted a number of other issues of significance. In particular, the long duration of our scans is not appropriately accounted for with a single value for the SZA. We also need to investigate the use of more realistic *a priori* profiles.

Our previous tests of the sensitivity of CO_2 retrievals to temperature and H_2O demonstrate the ability of our model to handle profile retrieval [13]. However, construction of a complete averaging kernel for use with actual profiles (e.g. ECMWF, National Centres for Environmental Prediction –NCEP), rather than the standard atmospheres, is the next step of the project. With improvements to the current instrument these tests are planned for a number of other trace gases in addition to CO_2 and O_2.

Acknowledgements

We thank Nick Jones and David Griffith (University of Wollongong) for allowing us to observe with our instrument at their excellent TCCON-FTIR observatory site, and for helping us with data for this intercomparison.

References

1. IPCC (Intergovermental Panel on Climate Change), "Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller, Eds. Cambridge, U.K. and New York, U.S.A.: Cambridge University Press, 2007.

2. Kuai, L., Worden, J., Kulawik, S., Bowman, K., Lee, M., et al., "Profiling tropospheric CO_2 using Aura TES and TCCON instruments," *Atmospheric Measurement Techniques*, vol. 6, pp. 63-79, 2013.
3. Sherlock, V., Connor, B., Wood, S., Gomez, A., Toon, G., Wennberg, P. and Washenfelder, R., "Preliminary results from the Lauder site of the Total Carbon Column Observing Network TCCON," presented at International TOVS Conference XIV, 2005.
4. Heimann, M. and Kaminski, T., "Inverse modelling approaches to infer surface trace gas fluxes from observed atmospheric mixing ratios," *Developments in Atmospheric Science*, vol. 24, pp. 277-295, 1999.
5. Crisp, D. and Johnson, C., "The orbiting carbon observatory mission," *Acta Astronautica*, vol. 56, pp. 193-196, 2005.
6. Yokota, T., Yoshida, T., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H. and Maksyutov, S., "Global concentrations of CO2 and CH4 retrieved from GOSAT: first preliminary results," *SOLA*, vol. 5, pp. 160-163, 2009.
7. Bailey, J., "A low-cost portable fibre-optic spectrometer for atmospheric absorption studies," *Atmospheric Measurement Techniques Discussions*, vol. 6, pp. 1067-1092, 2013.
8. Kawasaki, M., Yoshioko, H., Jones, N.B., Macatangay, R., Griffith, D.W.T., et al., "Usibility of optical spectrum analyzer in measuring atmospheric CO_2 and CH_4 column densitites: inspection with GTS and aircraft profiles in situ," *Atmospheric Measurement Techniques*, vol. 5, pp. 2593-2600, 2012.
9. Kobayashi, N., Inoue, G., Kawasaki, M., Yoshioko, H., Minomura, M., et al., "Remotely operable compact instruments for measuring atmospheric CO_2 and CH_4 column densities at surface monitoring sites," *Atmospheric Measurement Techniques*, vol. 3, pp. 1103-1112, 2010.
10. Wilson, E.L., Georgieva, E.M. and Heaps, W.S., "Development of a Fabry-Perot interferometer for ultra-precise measurements of column CO_2," *Measurement Science and Technology*, Iss. 5, vol. 18, pp. 1495-1502, 2007.
11. Petri, C., Warneke, T., Jones, N., Ridder, T., Messerschmidt, J., Weinzierl, T., Geibel, M. and Notholt, J., "Remote sensing of CO_2 and CH_4 using solar absorption spectrometry with a low resolution spectrometer " *Atmospheric Measurement Techniques*, vol. 5, pp. 1627-1635, 2012.
12. Kenyi, C., Cotton, D.V. and Bailey, J., "Retrieval software for total column greenhouse gas measurements from ground and space," presented at 12th Australian Space Science Conference, Melbourne, 13: 978-0-9775740-6-3, 2013.
13. Kenyi, C., Bailey, J. and Cotton, D.V., "Retrieval of greenhouse gas concentrations from observations with a ground-based spectrometer in the near-infrared," presented at 13th Australian Space Science Conference, UNSW, Sydney, 13:978-0-9775740-7-0, 2014.
14. Gisi, M., Hase, F., Dohe, S., Blumenstock, T., Simon, A. and Keens, A., "XCO2-measurements with a tabletop FTS using solar absorption spectroscopy," *Atmospheric Measurement Techniques*, vol. 5, pp. 2969-2980, 2012.
15. Georgieva, E., Wilson, E. and Heaps, W.S., "Precise measurement of CO2 from space using Fabry-Perot based optical setup – current status and development," *Proceedings of the SPIE*, vol. 6296, pp. 62961G-1, 2006.
16. Crisp, D., Boesch, H., Brown, L., Castano, R., Christi, M., et al., "Orbiting Carbon Observatory (OCO)-2; Level 2 full Physics Algorithm-Theoretical Basis Document, Version 1.0 Rev 4," OCO, 2010, pp. D-65488.
17. Wunch, D., Toon, G.C., Blavier, J.-F., Washenfelder, A., Notholt, J., Conner, B.J., Griffith, D.W.T., Sherlock, V. and Wennberg, P.O., "The Total Carbon Column

Observing Network (TCCON)," *Philisophic Transactions of the Royal Society of London, Series A*, vol. 369, pp. 2078-2112, 2011.

18. Cotton, D.V., Bailey, J. and Kedziora-Chudczer, L., "Atmospheric modelling for the removal of telluric features from infrared planetary spectra," *Monthly Notices of the Royal Astronomical Society*, Iss. 1, vol. 439, pp. 387-399, 2014.
19. Rothman, L.S., Gordan, L.E., Barbe, A., Benner, D.C., Bernath, P.F., et al., "The HITRAN 2008 molecular spectroscopic database," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 110, pp. 533-572, 2009.
20. Ellingson, R.G. and Fouquart, Y., "The intercomparison of radiation codes in climate models: An overview," *Journal of Geophysical Research*, Iss. D5, vol. 96, pp. 8925-8927, 1991.
21. Bailey, J., "VSTAR - A new high-spectral-resolution atmospheric radiative transfer code for Mars and other planets," presented at Second workshop on Mars atmosphere modelling and observations, Granada, Spain, 2006.
22. Bailey, J., "Methane and deuterium in Titan's atmosphere," presented at 11ASSC, ANU Canberra, 13:978-0-9775740-5-6, 2012.
23. Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P., "Numerical Recipies in FORTRAN: The Art of Scientific Computing," 2nd ed: Cambridge University Press, 1992.
24. Washenfelder, R.A., Toon, G.C., Blavier, J.-F., Yang, Z., Allen, N.T., Wennberg, P.O., Vay, S.A., Matross, D.M. and Daube, B.C., "Carbon dioxide column abundances at the Wisconsin Tall Tower site," *Journal of Geophysical Research*, vol. 111, pp. D22503, 2006.

Atmospheric Modelling for Neptune's Methane D/H Ratio – Preliminary Results

Daniel V. Cotton, Lucyna L. Kedziora-Chudczer, Kimberly Bott and Jeremy A. Bailey

School of Physics, University of New South Wales, NSW 2052, Australia

Summary: The ratio of deuterium to hydrogen (D/H ratio) of Solar System bodies is an important clue to their formation histories. Here we fit a Neptunian atmospheric model to Gemini Near Infrared Spectrograph (GNIRS) high spectral resolution observations and determine the D/H ratio in methane absorption in the infrared H-band (~1.6 μm). The model was derived using our radiative transfer software VSTAR (Versatile Software for the Transfer of Atmospheric Radiation) and atmospheric fitting software ATMOF (ATMOspheric Fitting). The methane line list used for this work has only become available in the last few years, enabling a refinement of earlier estimates. We identify a bright region on the planetary disc and find it to correspond to an optically thick lower cloud. Our preliminary determination of CH_3D/CH_4 is $3.0\text{x}10^{-4}$, which is in line with the recent determination of Irwin et al. [1] of $3.0\text{x}10^{-4}$ ($^{+1.0}/_{-0.9}\text{x}10^{-4}$), made using the same model parameters and line list but different observational data. This supports evidence that the proto-solar ice D/H ratio of Neptune is much less than that of the comets, and suggests Neptune formed inside its present orbit.

Keywords: Neptune, radiative transfer, atmospheres, Solar System planets, methane, deuterium.

Introduction

Deuterium as an indicator of planet formation location

The amount of deuterium in the universe is decreasing. Deuterium was produced in the Big Bang, and is destroyed in stars [2]. The amount of deuterium present during the formation of the Solar System is referred to as the *proto-solar D/H ratio*. The atmospheres of Jupiter and Saturn are considered representative of the proto-solar D/H ratio [1]; they are thought to have formed first of the giant planets at the *water ice-line*, representing the distance from the Sun at which water freezes – from the accretion, predominantly of water ice and clathrate hydrates [1]. Although much of the water ice in the proto-solar nebula is derived from the dense interstellar medium, its initial D/H ratio is not wholly retained; reprocessing takes place as part of the formation of the solar nebula, altering the D/H ratio [3]. Beyond the water ice-line the temperature controls the D/H ratio in water ice through an equilibrium reaction [2]:

$$H_2O + HD \rightleftharpoons HDO + H_2 \qquad (1)$$

Further away from the Sun the colder temperatures result in more deuterium being found in heavy-water ice rather than deuterated molecular hydrogen. The giant planets are thought to have accreted from the proto-solar nebula around icy cores formed predominantly from water ice [4]. Consequently, the D/H ratio of a planet – either in water, the accreted (predominantly Hydrogen) gas envelope or products evolved from those – is indicative of where it formed in relation to the Sun. Where the atmospheric constituents are evolved from the originally accreted material, measurement of these can be used to determine the D/H ratio of the accreted material that formed the planet. Determinations made of the D/H ratio of Jupiter and Saturn rely on an assumption – backed by models of the interiors of the planets – that deuterium present in the core ices have not significantly enriched the outer gas envelope [5]. In contrast, Uranus and Neptune are thought to have gas envelopes enriched through mixing during the formation by volatiles from their ice cores or infalling planetesimals; thus the comparison with Jupiter and Saturn provides an estimate of the deuteration of the ices in the local environment of their formation [5]. In the case of Neptune, methane is deuterated to its present level through an isotopic exchange reaction with molecular hydrogen [5]:

$$CH_4 + HD \rightleftharpoons CH_4D + H_2 \qquad (2)$$

Recently it has been proposed that Uranus and Neptune formed at the carbon monoxide ice-line, and thus were formed predominantly from carbon monoxide ice [6]. If this is the case, then the water in the interior of these planets will have CO as its origin, with a D/H ratio representative of that origin. CO in the core would react with H_2 according to [6]:

$$CO + 3H_2 \rightleftharpoons CH_4 + H_2O, \qquad (3)$$

and then as a greater portion of deuterium in the atmospheres of Uranus and Neptune would have come from the hydrogen envelope, the D/H ratio would be expected to be lower than for formation from water ices the same distance from the Sun. Recent measurements of the D/H ratios of Uranus and Neptune are lower than those of comets, which are thought to have formed in the same region of the Solar System, and so Ali-Dib et al. suggest that a CO ice origin reconciles this position [6]. Ali-Dib et al.'s initial calculations of the D/H ratio reveal a discrepancy with the measured ratio in that they infer a H_2/He ratio different to the Solar ratio during their formation [6]. A definitive measurement of the D/H ratio, for both Neptune and Uranus (which we are also investigating [7]) is very important to the viability of their hypothesis.

Recent measurements of the Neptunian D/H ratio

The most recent measurement of Neptune's D/H ratio was made by Irwin et al. [1]. They used the Near-infrared Integral Field Spectrometer (NIFS) on Gemini to obtain R = 5290 infrared H-band spectra to measure the CH_3D/CH_4 ratio from the CH_3D $3\nu_2$ lines. Only recently have sufficiently detailed line lists become available to accurately model methane at cryogenic temperatures [8]. Previously Feuchtgruber et al. [9] used the Photodetector Array Camera and Spectrometer (PACS) instrument on Herschel to determine the D/H ratio in molecular hydrogen from its R(0) and R(1) HD lines, supplemented by previous measurements of the R(2) rotational line using the Short Wavelength Spectrometer on board the Infrared Space Observatory [4]. To compare measurements of the D/H ratio in H_2 to those of the CH_3D/CH_4 ratio requires a knowledge of the fractionation factor, which describes how the D/H ratio differs in chemical products of the original material. The fractionation factor has been estimated by Lecluse et al. [5], and used by Irwin et al. to make the comparison: the results are summarised in Table 1.

Table 1: Recent Measurements of the Neptunian D/H ratio. Italicised entries represent calculations based on the fractionation factor estimated by Lecluse et al. (1.61+/-0.21) [5].

Reference	Instrument	D/H (H_2) ($x10^{-5}$)		CH_3D/CH_4 ($x10^{-4}$)	
Feuchtgruber et al. 2013 [9].	Herschel-PACS	4.1	+0.4	*2.6*	*+0.4*
			-0.4		*-0.4*
Irwin et al. 2014 [1].	Gemini/NIFS	*4.7*	*+1.7*	3.0	+1.0
			-1.5		-0.9

Neptunian Clouds

Neptune has a dynamic atmosphere with the highest zonal wind speeds in the Solar System (400 m/s) and a variety of cloud features [10]. In contrast to Neptune's uniform appearance in the visible, at infrared wavelengths there can be a striking contrast between light and dark features, where the lighter features correspond to haze in the stratosphere (~0.01 – 0.1 bar) and clouds near the tropopause (~2 – 3 bar) [10, 11]. What constitutes the haze and cloud layers seen in the (1.6 μm) H-band is not well understood [1] and although various trace hydrocarbons detected in the mid-infrared would condense out at pressures of around 0.002 – 0.01 bar their optical depths are calculated to be negligible in the H-band [10].

Observations and Data Reduction

Observations

We made observations of Neptune on the 18th of August 2011 using the Gemini Near Infrared Spectrograph (GNIRS) instrument on the Gemini-North 8m telescope, using a 0.1 arc-second slit. H-band (1.46 – 1.63 μm) spectra were obtained using both the 32 and 110.5 lines/mm gratings, having a resolution of R = 5100 and R = 17800 respectively at 1.65 μm. The higher resolution data represents a narrower spectral window (1.525 – 1.57 μm).

An acquisition image of Neptune was taken before each long-slit observation, with the cloud pattern appearing nearly identical. As such, we have indicated in Fig. 1 the slit positions and extracted pixels of both observations on one of the acquisition images. The slit was orientated N-S on the sky, an acute angle to planetary north (as shown by reference to the north-pointing arrow in Fig. 1). The 1-D spectra are obtained from the average of 5 pixels along the 0.1 arcsec slit corresponding to the maximum signal area, which overlaps with the brightest patch of cloud in the acquisition image of the planet taken with the narrowband H-G0516 filter (shown in Fig. 1).

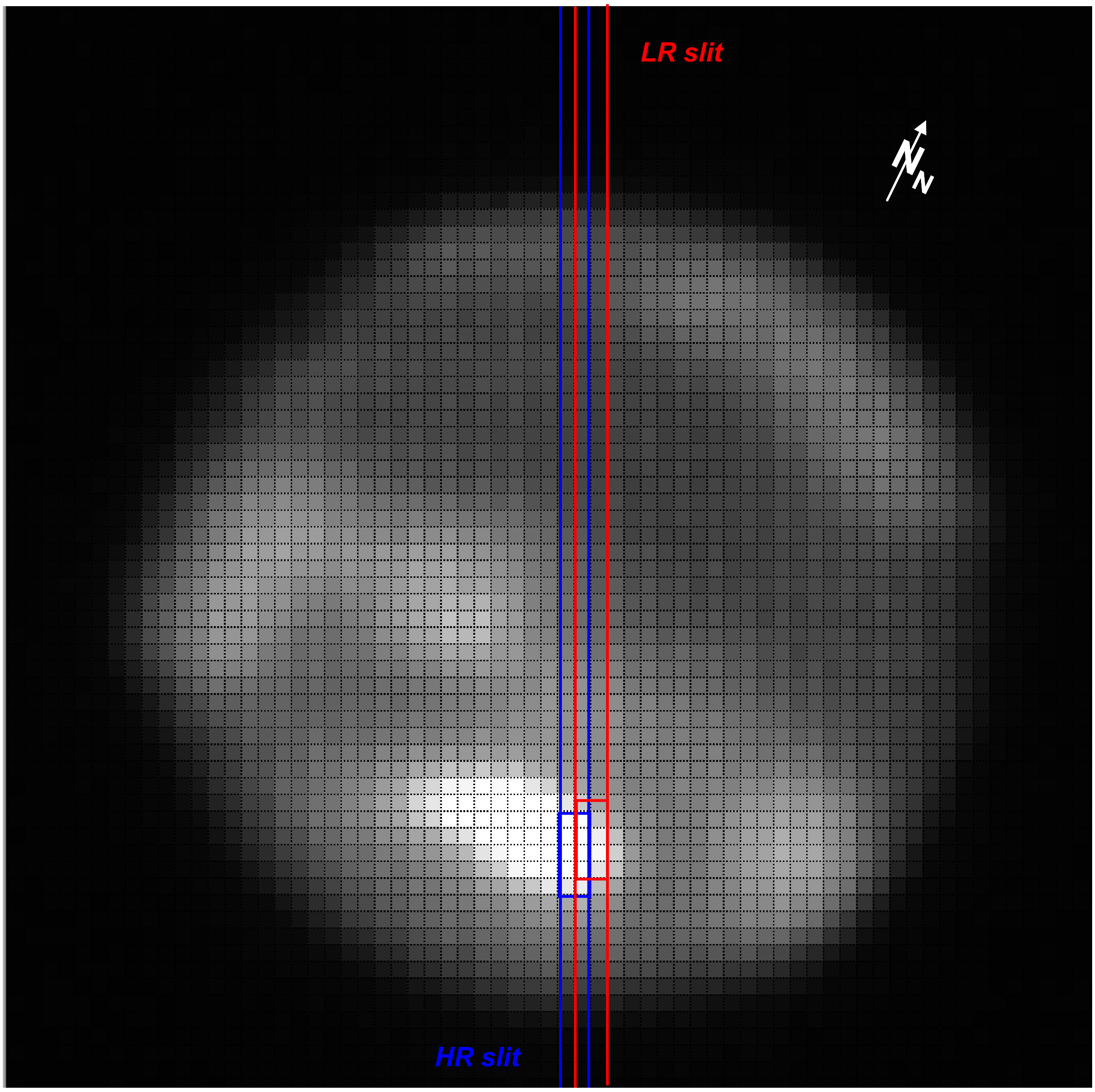


Fig. 1 Acquisition image for the spectroscopic observations obtained; made with the narrowband H-G0516 filter. The positions of the slit and the extracted pixels are indicated for both the lower resolution (LR) and high resolution (HR) spectra. Image is scaled in greyscale, with white areas being the brightest, and black darkest. Planetary North is indicated in the top right.

We fit an atmospheric model to both low-resolution* and high-resolution spectra. However in this work we present results of the fitting and retrieval of the lower resolution spectra only. The larger wavelength span of such a spectrum allows a robust determination of an appropriate cloud model that can be compared with previously published data. The best fitting model will then be applied to our high-resolution spectrum which will be reported on elsewhere.

At the time of the observation (07:23:27 UT) the mean airmass was 1.891, and the relative humidity 8%. For the low resolution spectra, our observing strategy involved twice obtaining four 350 sec exposures with Neptune moved between two slit positions (A and B) in an

* We refer to the R=5100 spectra as low-resolution only in comparison to the R=17800 spectra. R=5100 is higher resolution than was possible for this work prior to the availability of the current methane line list.

ABBA sequence for efficient sky subtraction, and averaging these. Wavelength calibration lamps (Ar-Xe) and spectra of the telluric standard star HIP 105734 (G1V type) were acquired after each sequence

Standard Data Reduction

The Gemini IRAF data reduction technique was applied to our spectra as follows: After visual examination of 2D spectra, all the data frames were aligned, corrected for bias level, non-linearity and bad pixels on the Aladdin III InSb detector array. The spectra of both Neptune and HIP 105734 were then flat fielded and wavelength calibrated using an interactive procedure to ensure correct identification of the Argon and Xenon lines present in the spectral region. The sky subtraction was performed by combining the A position frames, and subtracting the B position frames. The final step was to apply a dispersion correction and extract the 1D straightened spectra from the region of highest intensity along the 2D frame; this corresponded to the 5 pixel wide area (0.1x~0.25 arc-seconds) shown in Fig. 1.

Telluric Removal

The removal of the telluric absorption from our spectra was carried out using the ATMOF (ATMOspheric Fitting) code [12] as follows: We first take a high-resolution solar spectrum, and pass it through models of the Earth's atmosphere and the instrumental response. A number of model parameters are free, and ATMOF fits these to match the standard star data. The Earth atmosphere model used in this instance was the inbuilt VSTAR (Versatile Software for the Transfer of Atmospheric Radiation) [13] for Mauna Kea Observatory; its free parameters were the CO_2 and H_2O content. CO_2 is varied as a multiplication of the whole altitudinal profile, whilst H_2O is only varied for the lower layers of the atmosphere as detailed in Cotton et al. [12]. In the H-band water is a strong absorber at the short end of the band, and there is a smaller contribution from CO_2 near the centre of the band.

The instrumental response consisted of an augmented filter function, a filter response, a scaling factor and slope, as well as a wavelength shift correction. The augmented filter response was obtained from an average of the low frequency residuals from other standard star fits in the same wavelength window observed by us as part of the same observing run; the process is described in more detail by Bailey [14].

Once the parameters for the Earth atmosphere model have been retrieved, the model is re-run with the zenith angle corresponding to the Neptune observation. The Neptune spectrum is then multiplied by the Earth atmosphere transmission, and the instrumental response with the retrieved parameters applied. The only parameter retrieved from the standard star fitting not applied in correcting the Neptune model is the wavelength shift: this needs to be determined independently for each observation, as errors around 0.1Å can have a significant impact, and the quality of arc lamp calibration depends on the number of identified features in different spectral regions.

It is our preference to determine the wavelength shift by fitting a Doppler shifted solar spectrum to the planetary observational data; where the reflected solar lines are prominent in the spectrum this works well. However, this is not the case for Neptune in the H-band. Instead we used a basic model for Uranus from our earlier work [7] appropriately Doppler shifted. Neptune and Uranus have very similar spectra in the H-band, both dominated by methane absorption features, and as the purpose of this fitting is to retrieve only quadratic coefficients for wavelength shift, differences in the peak intensities or the overall flux are unimportant.

Atmospheric Model and Fitting

To determine the D/H ratio we fit a Neptune VSTAR model to the spectrum with telluric lines removed using ATMOF.

ATMOF Fitting

Although we have previously demonstrated the ability of the ATMOF software and associated procedures to remove telluric features from our data [12], this is the first time it has been used as a retrieval scheme to derive model parameters for a planetary atmosphere. Here we begin with the processed (tellurics and instrumental response removed) planetary spectrum. We then modify parameters of an atmospheric model to produce a model spectrum and compare it with the data at the same spectral resolution (R) at which the data was collected. For this purpose we used the published values of R from the Gemini website[†].

It is customary to present planetary spectra as a *radiance factor*, i.e. the planet radiance divided by the incident stellar flux; we have not done this when fitting with ATMOF. Commonly the removal of telluric lines is approximated by division by a standard star spectrum. As we have more precisely removed the telluric lines using ATMOF, to present our spectra as radiance factor, we in fact have to divide by a solar spectrum. Any model we fit to spectra presented as radiance factor would also have to be divided by a solar spectrum (since it also includes light from the Sun as a direct beam source) which is a complication without benefit.

Initial Neptune Model

The parameters of the atmosphere model applied to our Neptune data primarily come directly from the most recent work of Irwin et al. [1]. Pressure and temperature profiles were taken from the results of the *Voyager 2* radio occultations through the planetary atmosphere [15]. The methane profile used is that by Irwin et al. [1]: the deep methane mole fraction is set to 4%, the maximum mole fraction limited to 60% saturation vapour pressure in the troposphere, but limited to $6\text{x}10^{-4}$ in the stratosphere; a value used previously by Karkoschka and Tomasko [16], being intermediate of those derived by Lellouch et al. [17] and Baines et al. [18]. The N_2 mixing ratio was set to a constant 0.3% [1], and the remainder at each level divided amongst H_2 and He in the ratio 0.823/0.177 [15]. The atmospheric temperature, pressure, and gas mixing ratio profiles are shown in Fig. 2; they cover a pressure range between $3.5\text{x}10^{-4}$ to 6.3 bar. The VSTAR, radiative-transfer calculations are performed by subdividing this range of pressures into 45 levels. VSTAR requires altitudes corresponding to pressure values, these were calculated according to,

$$\Delta z = \frac{T_z}{L_b}\left(\Delta P^{\frac{RL_b}{gM_z}} - 1\right) \qquad (4)$$

where Δz is the change in altitude; T_z the temperature at that altitude; ΔP the change in pressure; M_z is the molecular mass at height z, calculated using the mixing ratios of the gases in the model; R the universal gas constant = 8.31432 $\text{N.m.mol}^{-1}\text{K}^{-1}$; g the gravity of Neptune, taken to be 11.1046 m.s^{-2}[19]; and L_b the adiabatic lapse rate, having a value of $8.53\text{x}10^{-4}$ K.m^{-1} [19]. The highest pressure in the model was then taken as zero altitude for reference.

† http://www.gemini.edu/sciops/instruments/gnirs/spectroscopy

The CH_4 line data used for the model are from laboratory measurements by the Grenoble group [20] made at cryogenic and room temperatures allowing the temperature dependence to be reliably determined. These line data have been shown to provide excellent models of the spectrum of Titan in the same wavelength region [2, 8, 21].

For the CH_4 lines broadened by H_2 we use the sub-Lorentzian line shape of Hartmann et al. [22], which has a far-wing χ-factor of 0.05882. We include collision-induced absorption between both H_2-H_2 and H_2-He molecules, as well as Rayleigh scattering from H_2, He and N_2. Irwin et al. [1, 10] also included contributions from H_2-CH_4 and CH_4-CH_4 collision induced absorption, but the contributions from these should be less, and we have neglected them for the time being.

Our cloud model was based on Irwin et al.'s thin cloud model, which they found to provide a very similar fit to their more sophisticated continuous distribution of cloud [1]. The cloud model is a bi-layer cloud model based on their earlier work [10] with a lower cloud layer, and an upper cloud layer, referred to as a haze, with the pressure of the upper haze being consistent with pressures retrieved by Gibbard et al. for their clouds [11]. Both layers are calculated with a Henyey-Greenstein phase function with asymmetry factor 0.7 and a modified-Gamma log-normal size distribution of particles with an effective radius of 1 μm, $\sigma = 0.05$, and the refractive index of the particles is set to $1.4+0i$; this value is typical of giant planet condensates [1, 7]. In the upper haze the single scattering albedo was set to 0.45. For the lower cloud the single scattering albedo and optical depth were given the wavelength dependant profile shown by Irwin et al. [1], and reproduced in Fig. 3.

For comparison, we also carried out preliminary cloud layer retrievals using Irwin et al.'s bright cloud model variant, in which the single scattering albedo of the upper cloud is 0.85 [1]. The bright cloud model produced a lower altitude upper haze, as well as much greater lower cloud opacities than anything previously reported by Irwin et al. [10]. Together with the propensity of ATMOF to retrieve a thicker lower cloud rather than an upper haze regardless of which of the two models is used, this allowed us to conclude that the standard model was the more appropriate.

The model was calculated at a spectral resolution of 0.01 cm^{-1} before convolving to the instrumental resolution. Upon completion of modelling we checked our model spectra against one calculated with a spectral resolution of 0.001 cm^{-1} and found no discernable difference.

Preliminary cloud layer fitting

Our primary goal for this work was to obtain appropriate cloud base pressures and opacities to be used for the high resolution data (with obtaining an initial D/H ratio a secondary objective), to that end we fit these cloud parameters first with the initial values set to those given by Irwin et al. for their bright spot; lower cloud: 2 bar, 0.7; upper haze: 3.5×10^{-2} bar, 0.5. As our data has not been flux calibrated we also fit a scaling factor.

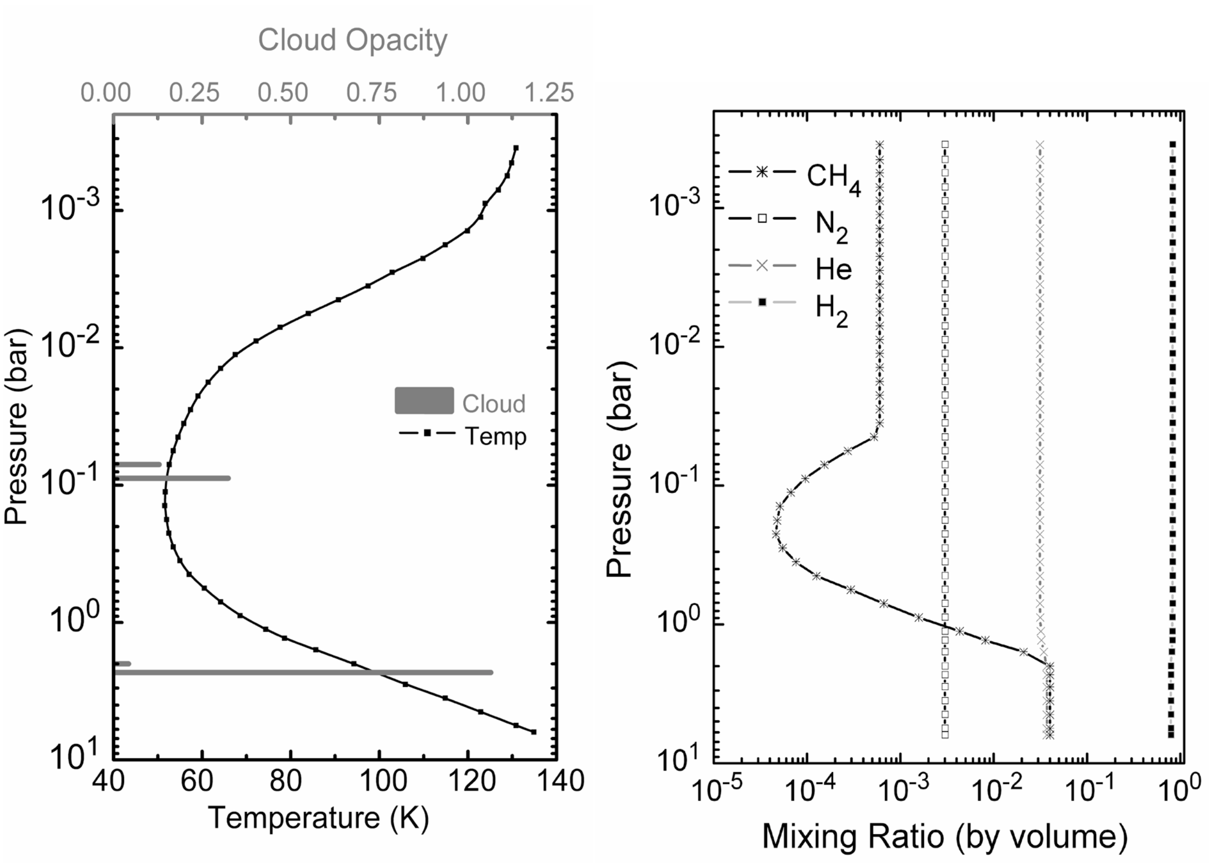


Fig. 2 The temperature-pressure profile (black) and retrieved cloud heights and optical depths (grey) shown together but represented by separate abscissa (left panel) and the gas mixing ratio profile (right panel) used in the Neptune model.

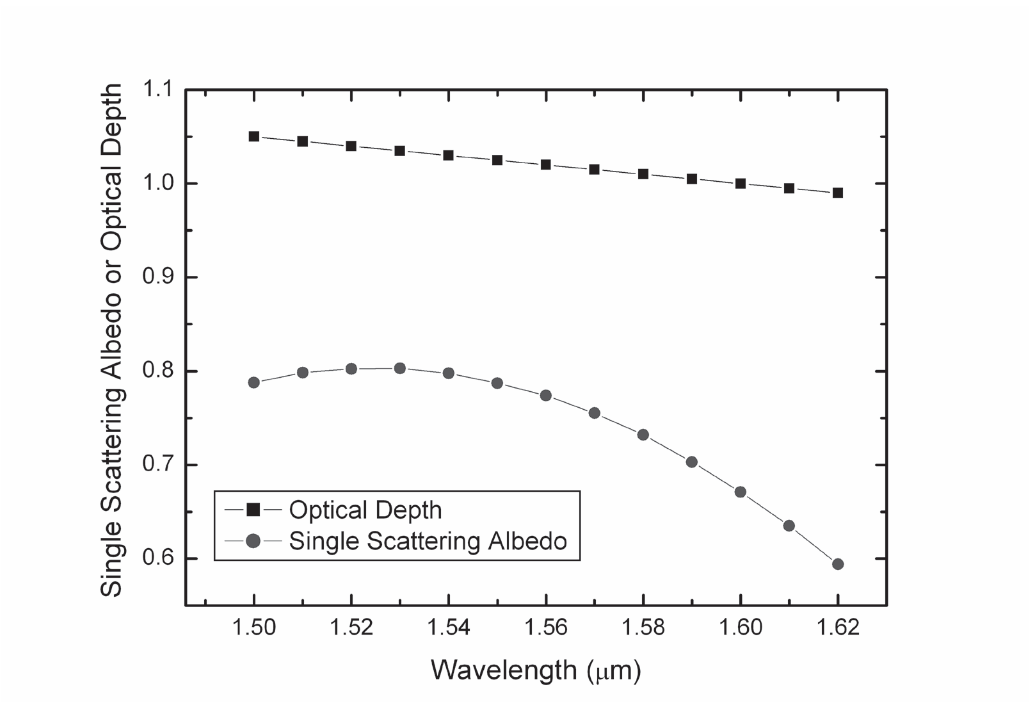


Fig. 3 Lower cloud properties that vary with wavelength from Irwin et al. [1].

Our modification to the Irwin et al. cloud model was to split the opacity of both clouds across two of our model layers[‡] in a ratio determined based on the nominal pressure of the cloud layer as follows: where the pressure corresponds to the centre of the model layer all the opacity is placed in that model layer; if the pressure is less then some of the opacity is placed in the model layer above, if the pressure is more, some of the opacity is placed in the layer below; the fraction of opacity placed in the model layer above or below is determined by the pressure difference with that of the centre of the layer. This was done in order to supply ATMOF's fitting routine with a continuously varying function, rather than a step-wise one – with which it deals poorly. A base pressure corresponding to the centre of the layer results in all of the opacity being placed in that layer, whereas a lower base pressure resulted in an increasing portion (up to half) of the opacity being placed in the lower layer.

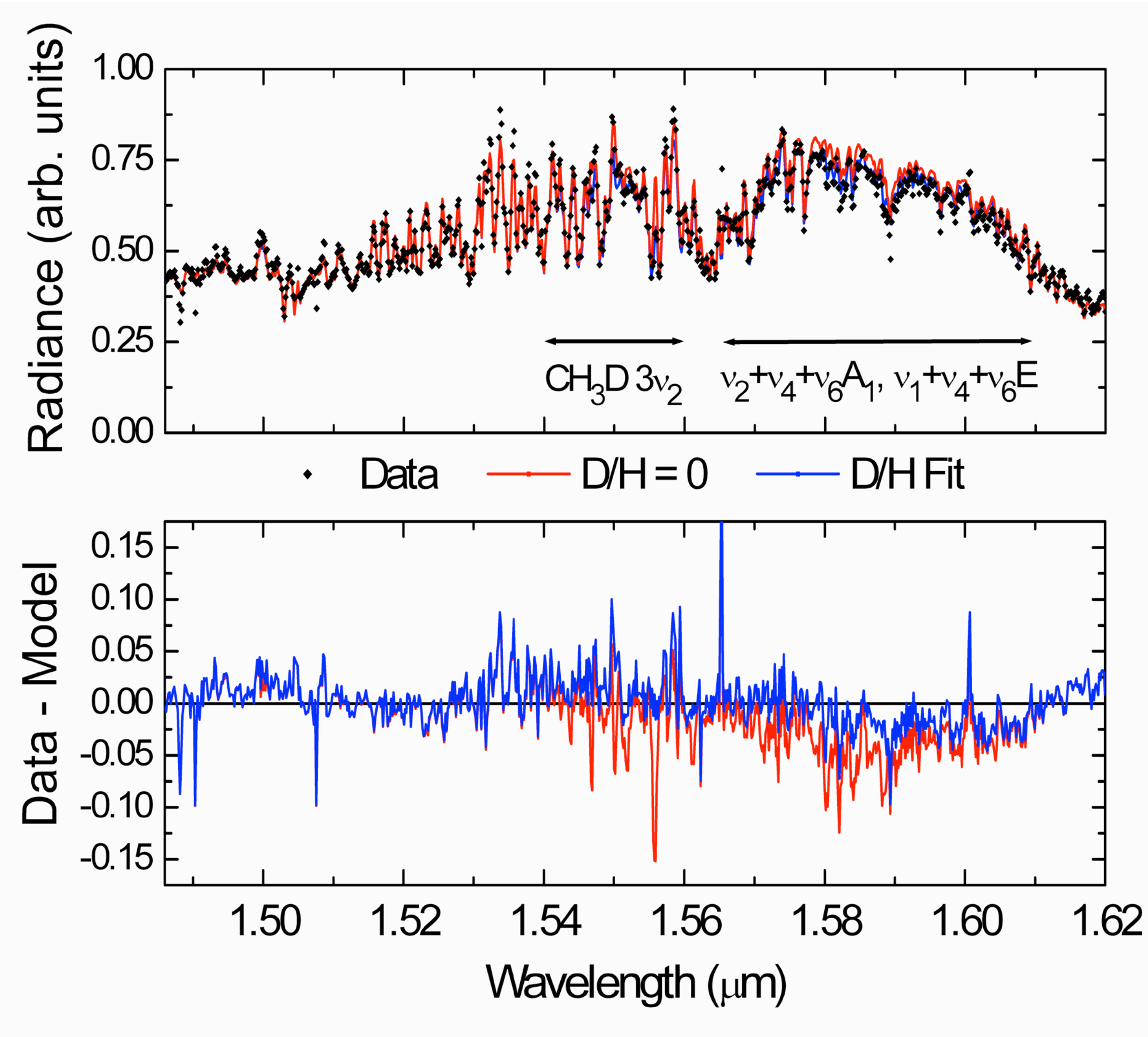


Fig. 4 Model fit to the data showing the contribution made by CH_3D. Marked in the top panel are the regions of absorption due to the CH_3D $\nu_2+\nu_4+\nu_6A_1$, $\nu_1+\nu_4+\nu_6E$ and $3\nu_2$ bands.

Cloud Layer and CH_3D/CH_4 Retrievals

Only after the initial cloud parameters have been retrieved and the goodness of fit assessed do we fit for the initial D/H ratio as the sole parameter. Then, once ATMOF has achieved a

‡ We therefore have two double-layer clouds.

solution, we refit the D/H ratio, along with all the cloud parameters and the scaling factor starting from the previously retrieved values. We find this to be most efficient and less prone to finding an erroneous local minimum as solution.

The cloud parameters, of central pressure of the cloud layer and opacity respectively, retrieved for the upper haze were 8.7×10^{-2} bar, 0.45; and for the lower cloud 2.0 bar, 1.11, where the break-down of the optical depth into the layers of the model is shown pictorially in the left panel of Fig. 2. The retrieved CH_3D/CH_4 was 3.0×10^{-4}. Fig. 4 shows the fit obtained by the model to our observational data, along with an otherwise identical model devoid of CH_3D.

Discussion and Further Work

Most of the bright features seen by Irwin et al. were the result of thick upper hazes (around 0.1 bar). On the 9th of September, 2011, however, they did observe one feature, like that we report here, that appears to have been due to a thick lower cloud [10]; this cloud was seen at similar latitudes to that which we report here.

This is the first time we have used ATMOF to retrieve parameters for a planet other than Earth. The excellent agreement between our determination of CH_3D/CH_4 with that of Irwin et al. [1], and the sensible cloud parameters retrieved by assuming only one unknown parameter – the flux scaling factor, show that ATMOF can be adapted to work well to retrieve basic atmospheric parameters. Further development of this option will include a proper assessment of uncertainties in the derived parameters, that will help in assessing the uniqueness of models under consideration.

As noted by Feuchtgruber et al. [9], recent determinations of the D/H ratio for Neptune (see Table 1), with which our value agrees, correspond to D/H ratios for the proto-planetary ices that are significantly less than those obtained so far for comets, which implies a formation location for Neptune well inside its current orbit [23].

Despite a good agreement between our model and observed spectrum (Fig. 4) the most notable discrepancy is in the failure of the model to match the extremes of depth or height of the finer spectral features, a characteristic shared by Irwin et al.'s fits [1]. To test ATMOF we deliberately limited the number of model parameters we investigated for this work, which limited our ability to precisely fit the data. An obvious candidate for an extra parameter is the deep methane concentration. Here we adopted Irwin et al.'s recommended value of 4%, but the lower value of 2.2% has often been used in the literature based on the work of Baines et al. [1, 18] . Another possibility is an inaccurate line-shape model, as the one we are using was not specifically developed for the conditions of Neptune's atmosphere.

It should also be noted that our model does not contain CO. The value determined in the upper atmosphere, 1 ppm, has previously been used throughout the atmosphere [1]. CO lines if present will show up as distinct P and R branches between 1.570 and 1.582 μm. However, the spectrum is most sensitive to CO around the level of the lower cloud, and Irwin et al. [1] found best agreement in their dark band data (where smaller cloud opacities were retrieved) with no CO at all (but noted that this was affected by the modelled deep methane concentration). Whether this is also the case in a region with a thick reflective lower cloud is something to investigate.

The next step is to flux calibrate our data and determine the uncertainties before re-determining the cloud parameters using the method laid out here, and using these to fit our higher resolution spectra. It is anticipated that doing so will allow us to place tighter constraints on the D/H ratio.

Acknowledgements

Based on observations obtained at the Gemini Observatory (and processing using the Gemini IRAF package), which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciência, Tecnologia e Inovação (Brazil) and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina). The observations were obtained under the program GN-2011B-Q-1.

References

1. Irwin, P.G.J., Lellouch, E., De Bergh, C., Courtin, R., Bézard, B., et al., "Line-by-line analysis of Neptune's near-IR spectrum observed with Gemini/NIFS and VLT/CRIRES," *Icar*, vol. 227, pp. 37-48, 2014.
2. Kedziora-Chudczer, L., Bailey, J. and Horner, J., "Observations of the D/H ratio in Methane in the atmosphere of Saturn's moon, Titan – where did the Saturnian system form?," presented at 12th Australian Space Science Conference, Melbourne, 13:978-0-9775740-6-3, 2013.
3. Cleeves, I.L., Bergin, E.A., Alexander, C.M.O.D., Du, F., Graninger, D., Öberg, K.I. and Harries, T.J., "The ancient heritage of water ice in the solar system," *Sci*, Iss. 6204, vol. 345, pp. 1590-1593, 2014.
4. Feuchtgruber, H., Lellouch, E., Bézard, B., Encrenaz, T., de Graauw, T. and Davis, G.R., "Detection of HD in the atmospheres of Uranus and Neptune: a new determination of the D/H ratio," *A&A*, vol. 341, pp. L17-L21, 1999.
5. Lecluse, C., Robert, F., Gautier, D. and Fuirand, M., "Deuterium enrichment in giant planets," *P&SS*, vol. 44, pp. 1579-1592, 1996.
6. Ali-Dib, M., Mousis, O., Petit, J.-M. and Lunine, J.I., "The measured compositions of Uranus and Neptune from their formation on the CO ice line," *ApJ*, Iss. 9, vol. 739, art. no. doi:10:.1088/0004-637X/793/1/9, 2014.
7. Bott, K., Kedziora-Chudczer, L. and Bailey, J., "VSTAR modelling of the infrared spectrum of Uranus," presented at ASSC12, RMIT Melbourne, 13:978-0-9775740-6-3, 2013.
8. Bailey, J., "Methane and deuterium in Titan's atmosphere," presented at 11ASSC, ANU Canberra, 13:978-0-9775740-5-6, 2012.
9. Feuchtgruber, H., Lellouch, E., Orton, G., De Graauw, T., Vandenbussche, B., et al., "The D/H ratio in the atmopsheres of Uranus and Neptune from Herschel PACS observations," *A&A*, vol. 551, pp. 126-135, 2013.
10. Irwin, P.G.J., Teanby, N.A., Davis, G.R., Fletcher, L.N., Orton, G.S., Tice, D., Hurley, J. and Calcutt, S.B., "Multispectral imaging observations of Neptune's cloud structure with Gemini-North," *Icar*, vol. 216, pp. 141-158, 2011.
11. Gibbard, S.G., de Pater, I., Roe, H.G., Martin, S., Macintosh, B.A. and Max, C.E., "The altitude of Neptune cloud features from high-spatial-resolution near-infrared spectra," *Icar*, vol. 166, pp. 359-374, 2003.
12. Cotton, D.V., Bailey, J. and Kedziora-Chudczer, L., "Atmospheric modelling for the removal of telluric features from infrared planetary spectra," *MNRAS*, Iss. 1, vol. 439, pp. 387-399, 2014.
13. Bailey, J. and Kedziora-Chudczer, L., "Modelling the spectra of planets, brown dwarfs and stars using VSTAR," *MNRAS*, vol. 419, pp. 1913-1929, 2012.
14. Bailey, J., "A low-cost portable fibre-optic spectrometer for atmospheric absorption studies," *AMTD*, vol. 6, pp. 1067-1092, 2013.

15. Lindal, G.F., "The atmosphere of Neptune: an analysis of radio occultation data acquired with Voyager 2," *AJ*, Iss. 3, vol. 103, pp. 967-982, 1992.
16. Karkoscha, E. and Tomasko, M., "The haze and methane distributions of Neptune from HST-STIS spectroscopy," *Icar*, vol. 211, pp. 780-797, 2011.
17. Lellouch, E., Hartogh, P., Feuchtgruber, H., Vandenbussche, B., De Graauw, T., et al., "First results of Herschel-PACS observation of Neptune," *A&A*, Iss. L152, vol. 518, art. no. doi:10.1051/0004-6361/201014600, 2010.
18. Baines, K.H., Mickelson, M.E., Larson, L.E. and Ferguson, D.W., "The abundances of methane and ortho/para hydrogen on Uranus and Neptune: Implications of new laboratory 4-0 H_2 quadrapole line parameters," *Icar*, vol. 114, pp. 328-340, 1995.
19. Huber, L. and Beebe, R., "NASA PDS: The Planetary Atmospheres Data Node – Adiabatic Lapse Rate, Dry," 2013.
20. Wang, L., Kassi, S., Liu, A.W., Hu, S.M. and Campargue, A., "The 1.58 μm transparency window of methane (6165-6750 cm^{-1}): Empirical line list and temperature dependence between 80K and 296K," *JQSRT*, vol. 112, pp. 937-951, 2011.
21. Bailey, J., Ahlsved, L. and Meadows, V.S., "The near-IR spectrum of Titan modeled with an improved methane line list," *Icar*, vol. 213, pp. 218-232, 2011.
22. Hartmann, J.-M., Boulet, C., Brodbeck, C., van Thanh, N., Fouchet, T. and Drossart, P., "A far wing lineshape for H2 broadened CH4 infrared transitions," *JQSRT*, Iss. 2, vol. 72, pp. 117-122, 2002.
23. Kavelaars, J.J., Mousis, O., Petit, J.-M. and Weaver, H.A., "On the formation location of Uranus and Neptune as constrained by dynamical and chemical models of comets," *ApJL*, Iss. L30, vol. 734, pp. 1-5, art. no. doi:10.1088/2041-8205/734/2/L30, 2011.

The role of Jupiter in driving Earth's orbital evolution: An update

Jonathan Horner [1,2], James B. Gilmore [2], and Dave Waltham [3]

[1] *Computational Engineering and Science Research Centre, University of Southern Queensland, West St, Toowoomba QLD 4350, Australia*
[2] *Australian Centre for Astrobiology, UNSW Australia, Sydney, NSW 2052, Australia*
[3] *Department of Earth Sciences, Royal Holloway, University of London*

Summary: In the coming decades, the discovery of the first truly Earth-like exoplanets is anticipated. The characterisation of those planets will play a vital role in determining which are chosen as targets for the search for life beyond the Solar system. One of the many variables that will be considered in that characterisation and selection process is the nature of the potential climatic variability of the exoEarths in question.

In our own Solar system, the Earth's long-term climate is driven by several factors – including the modifying influence of life on our atmosphere, and the temporal evolution of Solar luminosity. The gravitational influence of the other planets in our Solar system add an extra complication – driving the Milankovitch cycles that are thought to have caused the on-going series of glacial and interglacial periods that have dominated Earth's climate for the past few million years.

Here, we present the results of a large suite of dynamical simulations that investigate the influence of the giant planet Jupiter on the Earth's Milankovitch cycles. If Jupiter was located on a different orbit, we find that the long-term variability of Earth's orbit would be significantly different. Our results illustrate how small differences in the architecture of planetary systems can result in marked changes in the potential habitability of the planets therein, and are an important first step in developing a means to characterise the nature of climate variability on planets beyond our Solar system.

Keywords: Astrobiology, Exoplanets, Exo-Earths, Habitability, Jupiter, Milankovitch cycles

Introduction

The question of whether we are alone in the universe is one that has long fascinated mankind. In the past twenty years, we have made the first steps toward being able to answer that question. In 1995, astronomers announced the discovery of 51 Pegasi b [1] – the first planet found orbiting a sun-like star. In the first years after that discovery, news of new exoplanets came infrequently, as planets very different to those in our own Solar system were discovered (e.g. [2][3][4]). As time has passed, the techniques by which we detect exoplanets have been refined (as detailed in [11]), and the temporal baseline over which observations have been carried out has expanded. As a result, astronomers are now beginning to discover planetary systems that more closely resemble our own – such as those hosting "Jupiter Analogues", giant planets moving on near-circular orbits that take of order a decade to complete (e.g. [5][6][7]). At the same time, improvements in the our ability to monitor the minute variations in a star's brightness that reveal the transit of an unseen planet across its host star have resulted in the detection of ever smaller planets. A thorough analysis of the data obtained by the *Kepler* spacecraft ([8][9]) has already resulted in the discovery of just over 1,000 planets[1], including several that are significantly smaller than the Earth (e.g. [10]).

It is now increasingly apparent that small planets are far more common than larger ones. This result is clearly seen in data taken by Kepler (e.g. [9][46][47]) and that obtained by radial-velocity search programs (e.g. [48][49]). Building on this work, a number of new exoplanet search programs will soon begin that should further bolster the number of small exoplanets known. These feature both ground-based programs (including dedicated facilities such as MINERVA [50] and NRES [51]) and space-based surveys (the Kepler K2 mission [52], TESS [53] and PLATO [54]).

With the rate at which new planets are being detected, and given the exciting new detection programs due to begin in coming years, it is likely that the first truly Earth-like planets (exoEarths) will soon be discovered. At that point, the search for life beyond the Solar system will become a key focus of astronomical research. However, the observations needed to characterise those planets, and search for any signs of life, will be incredibly challenging. The recent detection of water vapour in the atmosphere of a Neptune-sized exoplanet, HAT-P-11b [12] is a case in point. Despite the fact the planet is ~4 times the radius of Earth, and orbits its host star at around 1/20th the Earth-Sun distance, the detection still required the combination of 212 observations, totalling approximately 7 hours of integration time, using the Hubble Space Telescope's WFC3 G141 grism spectrometer. To gather spectroscopic evidence for life in the atmosphere of an exoEarth will be a significantly greater challenge, and would require more detailed and lengthy observations – making it a costly and time consuming process[2]. For that reason, it is imperative that astronomers select the most promising targets for the search for life before beginning their observations. But how will those targets be selected?

It is now thought that there are many factors that come together to render a given planet more or less habitable (e.g. [14], and references therein). One of the key factors to be considered is the stability of the climate of the planets in question. If all else were equal, it is reasonable to assume that the planet with the most stable climate would be the most promising target to search for life. It seems likely that otherwise ideal candidates could be rendered uninhabitable, should their climates vary too dramatically or rapidly for life to adjust and survive.

In our Solar system, the Earth's orbit is perturbed by the gravitational interaction with the other planets. This interaction results in Earth's eccentricity, inclination, and tilt varying over time. These orbital changes cause the radiation received at Earth's polar regions to vary over astronomical timescales, with periods of ~20 kyr and ~100 kyr. These cycles are known as Milankovitch cycles (e.g. [15][16][17]), and are thought to drive the glaciation/degalciation cycles observed over the last few million years (e.g. [18][19]).

Despite the role played by the Milankovitch cycles in driving Earth's climate and recent glaciations, the scale of the variations in Earth's orbital elements is actually relatively small. The same is not true elsewhere in the Solar system, however. For example, Mercury exhibits Milankovitch cycles with far greater amplitude than those of the Earth – with an orbital eccentricity (currently ~0.21) that can reach values as high as 0.45 (e.g. [20][21]). Were the

[1] As of 18 Feb 2015, the number of confirmed planets discovered by the *Kepler* spacecraft stands at 1013 (http://kepler.nasa.gov), with a further 4,175 candidates awaiting confirmation.
[2] To illustrate the difficulty performing such observations, we note that recent simulations (as detailed in [13]), suggest that even the James Webb Space Telescope, due to be launched in 2018, will find it '*nearly impossible to measure spectra of terrestrial analogs*' (section 2.1.3).

Milankovitch cycles on the Earth equally extreme, it would sometimes move on an orbit that brought it closer to the Sun than Venus, at perihelion, and take it out to almost the orbit of Mars, at aphelion. The effects on Earth's climate might prove to be catastrophic, and it seems questionable whether life could survive, let alone thrive, on the surface of such a world.

In this paper, we present the preliminary results of a new study that aims to bring together *n*-body dynamical methods and climate modelling to assess the influence of planetary architecture on the climate of Earth-like planets. Here, we study the influence of the orbit of the planet Jupiter on the amplitude and frequency of Earth's orbital oscillations. Our work builds on that presented at the 13^{th} Australian Space Science Conference [22] by using a greatly expanded suite of dynamical integrations (both in terms of number and duration), and using a corrected and improved methodology to perform the relativistic corrections required to model the orbit of the planet Mercury.

In the next section, we detail our methods, before presenting our preliminary results, and then conclude with a discussion of our plans for future work.

Method

In order to study how the architecture of the Solar system influences Earth's orbital evolution, we use a modified version of the Hybrid integrator within the *n*-body dynamics package MERCURY ([23]). The standard version of the package models the dynamical interaction of test particles (both massive and massless) in a purely Newtonian sense. In addition, the modified version, developed for this work through the implementation of an additional user-defined force, takes account of the first-order post-Newtonian relativistic corrections [24]. This allows the code to accurately model the evolution of the orbit of the planet Mercury, when using a solely Newtonian method would yield erroneous results (e.g. [25][26]).

Using our modified version of MERCURY, we performed 159,201 simulations following the dynamical interaction of the eight planets in our Solar system for a period of 10 Myr. The initial orbits of all planets, except Jupiter, were held fixed throughout the suite of simulations, using the NASA DE431 [26] ephemeris to provide their current best-fit orbits. The simulation start epoch was taken as JD2450985.5, which corresponds to 00:00 21^{st} June 1998, UT.

In each of the simulations, we placed the giant planet Jupiter on a different orbit. The inclination and rotation angles of Jupiter's orbit were identical in all cases, taken from the DE431 ephemeris. The semi-major axis and eccentricity of Jupiter's orbit, however, were varied from one run to the next. 399 unique values of semi-major axis were simulated, evenly spread across a 4 au region, centred on Jupiter's best-fit orbit (i.e. spanning the range 5.203102 ± 2.000000 au). At each of the semi-major axes tested, 399 discrete values of initial orbital eccentricity were considered. These ranged from a circular orbit (i.e. e = 0.0) to one with moderate eccentricity (e = 0.4), again in even steps. As such, the tested Jovian orbits spanned a 399 x 399 grid in *a*-*e* space, giving us our 159,021 simulations.

To maximise the accuracy of the orbital solution, a time-step of 1 day is used in each simulation. The orbital elements of each of the planets were output at 1,000-year intervals throughout the integrations. As a result of the large range of orbital elements considered for Jupiter, it was anticipated that at least some of the systems considered would prove dynamically unfeasible (as has often been observed when examining the dynamics of recently proposed exoplanetary systems – e.g. [27][28][29][30]). If any of the planets collided with one another, or with the central body, the simulation was stopped, and the time at which the collision occurred was recorded. Similarly, if any planet reached a barycentric distance of 40 au, that planet was considered ejected from the Solar system, and that integration was stopped, with the time recorded[3].

The results were used to create maps of the variability of the Earth's orbital elements as a function of Jupiter's initial semi-major axis and eccentricity. These maps, which build on earlier work studying the stability of proposed exoplanetary systems (e.g. [31][32]), give a quick visual guide to the degree to which the Earth's Milankovitch cycles are influenced by small changes in the orbit of Jupiter, and we present a number of examples of such plots in the next section. We are currently in the process of taking the numerical results of our simulations (the orbital elements for the Earth across the various runs) and using them as input for simple climate models (e.g. [33]), to examine how the observed variability in Earth's orbit might affect its climate. We anticipate that this analysis will be complete in the coming year.

Preliminary Results

Figure 1 shows the variation in the Earth's orbital eccentricity (top panels, in red) and inclination (lower panels, in blue) for two exemplar simulations. The only difference between the two scenarios plotted was the initial semi-major axis chosen for Jupiter's orbit. In the left hand plots, Jupiter began the simulations at its true location in the Solar system, based on NASA's DE431 ephemeris (a = 5.203102 au). By contrast, the right hand plots show the scenario where Jupiter began the simulations at a = 3.203102 au. All other initial conditions were identical between these two simulations. It is immediately apparent that the amplitude of both the eccentricity and inclination excursions experienced by the Earth are broadly similar between the two runs. The maximum eccentricity in the scenario more closely resembling our own Solar system (left) was slightly higher than when Jupiter was closer to the Sun, but remained relatively small in both cases. However, the frequency of the cyclical behaviour in both eccentricity and inclination was much greater when Jupiter was placed closer to the Sun than when it began on its true orbit. Simply moving Jupiter inwards has had a significant effect on the Milankovitch cycles experienced by the Earth in these runs.

[3] A barycentric distance of 40 au was chosen for the 'ejection' distance as a reasonable compromise that allowed us to determine when the Solar system had destabilised. For any of the planets to reach 40 au would require a major re-arrangement of the system's architecture.

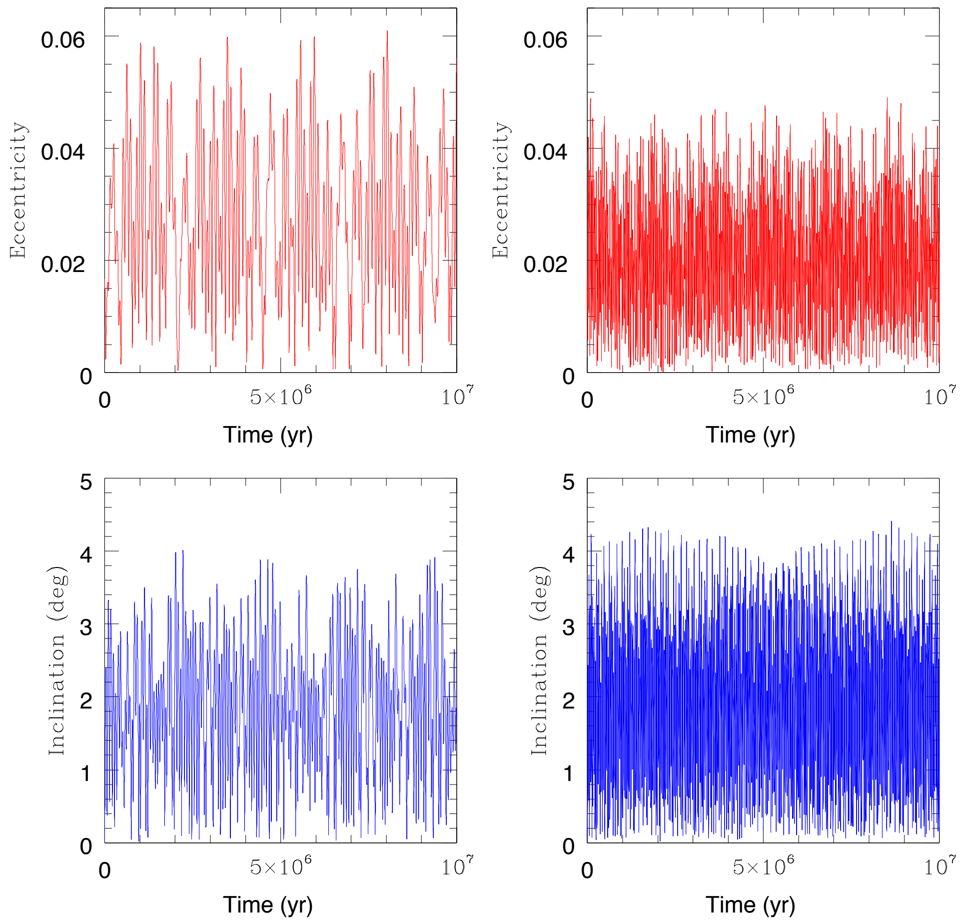


Fig. 1: The variation in the Earth's orbital eccentricity (top, red) and inclination (bottom, blue), for two of the versions of our Solar system studied in this work. The left hand data is from the system that most closely resembled our own, whilst the right is for the scenario where Jupiter was shifted inwards by a distance of 2 au. All other initial conditions were identical between the two runs. It is clear that, though the amplitude of the variations in eccentricity and inclination were broadly the same between these two runs, the speed at which the variations occurred was greater for the scenario where Jupiter was closer to the Sun.

The exemplar cases shown in Figure 1 clearly demonstrate how changing Jupiter's orbit can influence the Earth's Milankovitch cycles. However, changes to Jupiter's orbit can also have a much more dramatic effect – they can cause the Solar system to become unstable on very short timescales. As a result of the wide range of orbital architectures tested in this work, we found that a significant fraction of the Solar systems we created proved to be dynamically unstable, falling apart long before the end of our 10 Myr integrations. The stability of the Solar system as a function of Jupiter's initial semi-major axis and orbital eccentricity can be seen in Figure 2. Of the 159,021 versions of our Solar system we tested, almost 74% (117,549 systems) proved dynamically unstable within the ten million years of our simulations. The great majority of these featured Jupiters that were initially placed on orbits more eccentric than that displayed by our own Jupiter. However, there were values of initial Jovian semi-major axis where the Solar system was rendered unstable even for circular initial Jupiter orbits (such as those at around 4.25 and 4.95 au, where Jupiter and Saturn start in mutual 10:3 and 8:3 mean motion resonance, respectively). Similarly, there are two regions where the Solar system would remain stable even for moderately eccentric Jupiters – around 4.6 and 6 au. Once again, these are locations where Jupiter and Saturn would move on mutually resonant orbits – namely the 3:1 and 2:1 mean motion resonances, respectively.

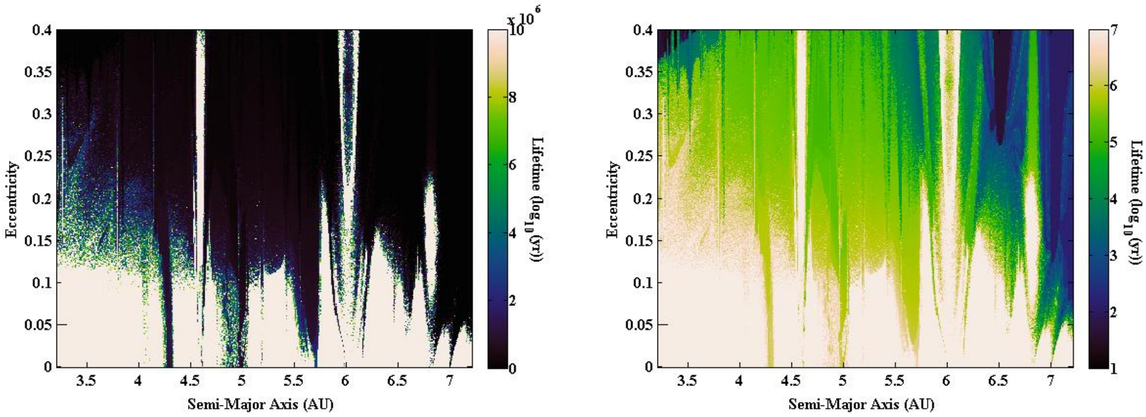


Fig. 2: The stability of the Solar system as a function of the initial semi-major axis, a, and eccentricity, e, of Jupiter's orbit. In these integrations, the initial orbits of the other planets were held at their DE431 ephemeris values, and their evolution was followed for 10 Myr under the influence of their mutual gravitation. The left plot shows the lifetime of each system on a linear scale, whilst the right shows the same information on a logarithmic scale. It is clear that the majority of systems were unstable, even on the short timescales considered.

The resonant features described above are clearly seen in the left panels of Figure 3, which shows the fraction of unstable simulations (top) and mean lifetime (bottom) of the Solar system as a function of Jupiter's initial semi-major axis and eccentricity. Both unstable and stable resonant features can be seen as troughs and peaks overlaid on a gradual trend to lower stability the closer Jupiter moves toward Saturn. The situation is more clear-cut when one considers the influence of eccentricity on the system's stability. Even scenarios featuring Jupiter on an initially circular orbit can prove unstable – but away from the destabilising influence of resonances (as seen in Fig. 2), the great majority of low-eccentricity solutions prove stable on the timescales considered in this work. Aside from those resonant regions, it is generally the case that scenarios where Jupiter has an initial orbital eccentricity less than ~0.125 prove dynamically stable – although this critical value is somewhat higher at low semi-major axes, and lower at greater semi-major axes. This transition region is evidenced in the right-hand panels of Figure 3 by the shoulder visible at $e \sim 0.1$, where the unstable fraction begins to climb more rapidly, and the mean lifetime begins to fall off more steeply.

For those regions where the system survived for the full 10 Myr of integration time, we have mapped the variability of the Earth's orbital elements as a function of time. In Figure 4, we show the variability of the Earth's orbital eccentricity over the integration period. In the left hand panel of that plot, we show the rms time variability of Earth's eccentricity, plotted on a logarithmic scale. In general, the closer to the Sun the initial orbit of Jupiter, the more rapidly the Earth's orbital eccentricity is driven to vary. As a second-order effect, the more eccentric the initial orbit of Jupiter, the more rapid are Earth's eccentricity excursions. This effect is particularly apparent in the strips of stability located at location of the 2:1 and 3:1 mean-motion resonances between Jupiter and Saturn.

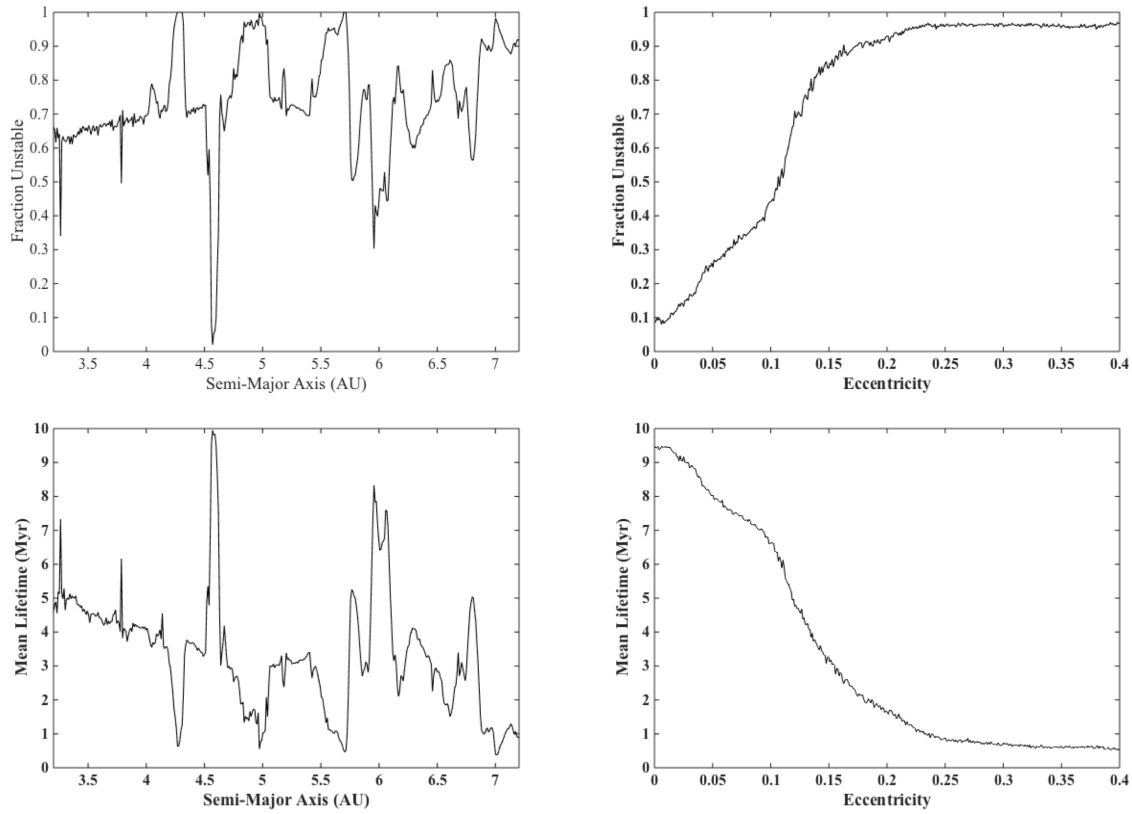


Fig. 3: The fraction of our simulations that proved dynamically unstable (top) and the mean lifetime of those integrations (bottom), as a function of the initial semi-major axis (left) and orbital eccentricity (right) of Jupiter. The increasing instability of the systems tested as a function of orbital eccentricity can be clearly seen in the right hand panels, whilst the influence of mean-motion resonances (such as the 3:1 and 2:1 MMRs between Jupiter and Saturn, at ~4.6 and 6.0 au respectively) can be clearly seen in the left hand panels.

The right-hand panel of Figure 4 shows the maximum eccentricity obtained by the Earth over the course of our integrations, again for those simulations where the Solar system survived intact for the full simulation time. In contrast to the rate at which Earth's eccentricity is driven to change, the scale of its maximum excursions seems unrelated to the initial orbital semi-major axis of Jupiter. Instead, it seems to be a function of the degree of stability of that planet's orbit. In general, towards the centre of the broad regions of stability, the maximum eccentricity obtained by the Earth's orbit is low, with it rising towards the edges of the stable regions – an indication that the Solar system is in the process of transitioning between a more stable and a less stable regime in those areas.

Figure 5 shows the variation of the Earth's orbital inclination over the course of our 10 Myr simulations. As with Figure 4, the left-hand panel shows the rms rate of change of Earth's orbital inclination, plotted on a logarithmic scale, whilst the right hand plot shows instead the maximum excursions seen in orbital inclination. As was the case with the rate at which Earth's orbital eccentricity was driven, it is clear from the left-hand panel of Figure 5 that the rate of inclination change is a strong function of Jupiter's initial semi-major axis. The closer that Jupiter orbits, the more rapidly the Earth's orbital inclination is driven. When taken in concert with the results shown in the left-hand panel of Figure 4, this is an indication that the frequency of the Milankovitch cycles might be a strong function of Jupiter's initial orbital location – the greater the Earth-Jupiter separation, the more slowly the cycles progress. Future work is necessary, however, before this conclusion can be confirmed.

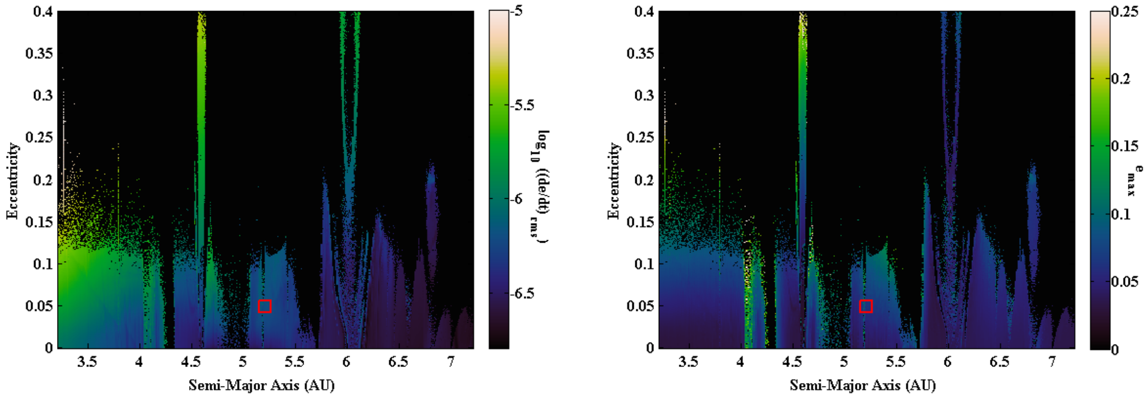


Fig. 4: The variation of the Earth's orbital eccentricity, as a function of Jupiter's initial orbital semi-major axis and eccentricity. The left hand panel shows the variation in the root-mean-squared value of the rate of change of Earth's orbital eccentricity with time, whilst the right shows the maximum value that eccentricity obtained over the 10 Myr of our integrations. The red square in the plots shows the location of Jupiter in our own Solar system.

The right hand panel of Figure 5 shows the maximum inclinations obtained by the Earth across the stable 10 Myr integrations we performed. The great majority of stable solutions show only very small excursions in inclination. There are, however, three bands (located at ~4.1, 4.7 and 6.3 au) where Earth's orbital inclination varies over a larger range. The band of values around ~15 degrees, at ~4.1, may well be the result of the 7:2 mean-motion resonance between Jupiter and Saturn, which would occur at ~4.14 AU. That narrow band at around 6.3 au is somewhat more mysterious, but given the apparent curved nature of the band (with an extension potentially visible at a ~6 au, e ~0.2) suggests that it may be the result of a secular resonance between the three planets in question (Jupiter, Saturn, and the Earth). Again, further work is needed to confirm or deny our suspicions in this case.

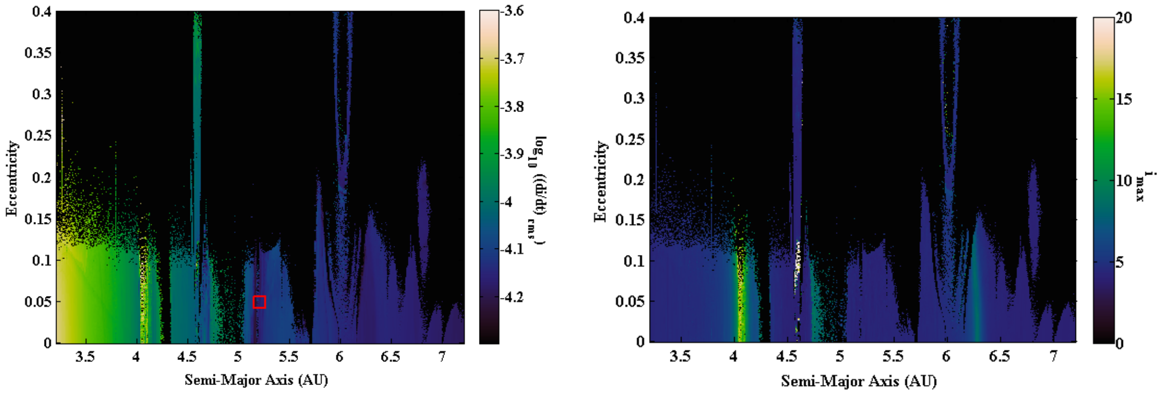


Fig. 5: The variation of the Earth's orbital inclination, as a function of Jupiter's initial semi-major axis and eccentricity. Left: The RMS value of the rate of change of the Earth's orbital inclination with time. Right: The maximum inclination attained by the orbit of the Earth's over the 10 Myr of our integrations. The red square in the left-hand plot shows the location of Jupiter's orbit within our own Solar system.

Future Work

We are currently in the process of taking the orbital elements output for the many Earths in our integrations and using them as the input for detailed climate modelling, to determine how the changes in Earth's orbital variability (as illustrated in Fig. 2 and Fig. 4) would influence the climate of our planet. It is not necessarily the case that large excursions in orbital eccentricity and inclination would render the climate less clement for the development of life – clearly the rate at which variations occur will play an important role in determining the climatic response to the Milankovitch cycles themselves (e.g. [34]).

Once we have a firm handle on the interaction between the Solar system's architecture and the Earth's climate variability, we will look to extend our modelling to consider planets in the habitable zone of known exoplanetary systems. Our long term goal is to construct a systematic approach by which the potential habitability of such planets can be quickly assessed, so that the best possible candidates can be selected for the search for life elsewhere. Such studies will form a critical component of the search for life beyond the Solar system, and will complement studies of the other factors that can influence planetary habitability – ranging from stellar activity (e.g. [35][36][37]) to the impact regimes (e.g. [38][39][40][41]) and delivery of volatiles (e.g. [42][43]) that might be experienced by potential targets, and even the presence (or lack) of giant satellites (e.g. [44][45]).

Taken in concert, these studies will eventually provide a vital resource that will help observers to target those exoplanets that offer the greatest likelihood of hosting life like that seen on the Earth (e.g. [14]). Since studying the Milankovitch cycles will provide important information about potential exoEarth climate variability, we anticipate that astronomers will want to obtain as much information as possible prior to committing the extremely complicated observations that will be necessary to look for biomarkers on newly discovered planets.

Acknowledgements

The work was supported by iVEC through the use of advanced computing resources located at the Murdoch University, in Western Australia.

References

1. Mayor, M. and Queloz, D., "A Jupiter-mass companion to a solar-type star", *Nature*, Vol. 378, 1995, pp. 355 - 359

2. Mary, G. W. and Butler, R. P., "A Planetary Companion to 70 Virginis", *Astrophysical Journal Letters*, Vol. 464, 1996, L147

3. Butler, R. P., Marcy, G. W., Williams, E., Hauser, H. and Shirts, P., "Three New '51 Pegasi-Type' Planets", *The Astrophysical Journal*, Vol. 474, 1997, L115 – L118

4. Cochran, W. D., Hatzes, A. P., Butler, R. P. and Marcy, G. W., "The Discovery of a Planetary Companion to 16 Cygni B", *The Astrophysical Journal*, Vol. 483, 1997, pp. 457-364

5. Howard, A. W., Johnson, J. A., Marcy, G. W., Fischer, D. A., Wright, J. T., Bernat, D., Henry, G. W., Peck, K. M. G., Isaacson, H., Apps, K., Endl, M., Cochran, W. D., Valenti, J. A., Anderson, J. and Piskunov, N. E., "The California Planet Survey. I. Four New Giant Exoplanets", *The Astrophysical Journal*, Vol. 731, 2010, pp. 1467 - 1481

6. Wittenmyer, R. A., Horner, J., Tuomi, M., Salter, G. S., Tinney, C. G., Butler, R. P., Jones, H. R. A., O'Toole, S. J., Bailey, J., Carter, B. D., Jenkins, J. S., Zhang, Z., Vogt, S. S. and Rivera, E. J., "The Anglo-Australian Planet Search. XXII. Two New Multi-planet Systems", *The Astrophysical Journal*, Vol. 453, 2012, id. 169

7. Robertson, P., Horner, J., Wittenmyer, R. A., Endl, M., Cochran, W. D., MacQueen, P. J., Brugamyer, E. J., Simon, A. E., Barnes, S. I. and Caldwell, C., "A Second Giant Planet in 3:2 Mean-motion Resonance in the HD 204313 System", *The Astrophysical Journal*, Vol. 754, 2012, id. 50

8. Borucki, W. J., Koch, D., Basri, G. et al., "Kepler Planet-Detection Mission: Introduction and First Results", *Science*, Vol. 327, 2010, pp. 977 - 980

9. Borucki, W. J., Koch, D., Basri, G. et al., "Characteristics of Planetary Candidates Observed by Kepler. II. Analysis of the First Four Months of Data", *The Astrophysical Journal*, Vol. 736, 2011, id. 19

10. Barclay, T., Rowe, J. F., Lissauer, J. J. et al., "A sub-Mercury-sized exoplanets", *Nature*, Vol. 494, 2013, pp. 452-454

11. Perryman, M., "The Exoplanet Handbook", by Michael Perryman, Cambridge, UK: Cambridge University Press, 2014

12. Fraine, J., Deming, D., Benneke, B., Knutson, H., Jordán, A., Espinoza, N., Madhusudhan, N., Wilkins, A. and Todorov, K., "Water Vapor in the Clear Atompshere of an exo-Neptune", *Nature*, Vol. 513, 2014, pp. 526 – 529

13. Beichman, C., Benneke, B., Knutson, H. et al., "Observations of Transiting Exoplanets with the James Webb Space Telescope (JWST)", *Publications of the Astronomical Society of the Pacific,* December 2014, arXiv:1411.1754

14. Horner, J. and Jones, B. W., "Determining habitability: which exoEarths should we search for life?", *International Journal of Astrobiology*, 9, 2010, pp. 273-291

15. Croll, J., 1875, "Climate and time in their geological relations, a theory of secular changes of the earth's climate", Published by Daldy, Tsbister & co., London

16. Köppen, W., Wegener, A., 1924, "Die Klimate der Geologischen Vorzeit", Gebrüder Borntraeger, Berlin.

17. Milankovitch, M., 1941, "Kanon der Erdbestrahlung und seine Andwendung auf das Eiszeiten-problem", R. Serbian Acad., Belgrade

18. Roe, G., 2006, "In defense of Milankovitch", Geophysical Review Letters, 33, L24703

19. Hays, J. D., Imbrie, J. and Shackleton, N. J., 1976, "Variations in the Earth's Orbit: Pacemaker of the Ice Ages", Science, 194, 1121-1132

20. Strom, R. G., Sprague, A. L., “Exploring Mercury: the iron planet”, eds. Strom, R. G., Sprague, A. L., Springer-Praxis Books in Astronomy and Space Sciences, London (UK): Springer, Chichester (UK): Praxis Publishing. ISBN 1-85233-731-1, 2003

21. Correia, A. C. M., Laskar, J., , “Mercury's capture into the 3/2 spin–orbit resonance including the effect of core–mantle friction”, Icarus, Vol. 201, 2009, 1-11

22. Horner, J., Waltham, D. and Koch, F. E., “The role of Jupiter in driving Earth’s orbital evolution”, *Proceedings of the 13th annual Australian Space Science Conference*, ISBN 13: 978-0-9775740-7-0, eds. Wayne Short and Iver Cairns, 2014, pp. 117 - 128

23. Chambers, J. E., “A hybrid symplectic integrator that permits close encounters between massive bodies”, *Monthly Notices of the Royal Astronomical Society*, Volume 304, 1999, pp. 793 – 799

24. Gilmore, J. B. and Ross, A., “Effective field theory calculation of second post-Newtonian binary dynamics”, *Physical Review D*, Vol. 78, 2008, 124021

25. Clemence, G. M., “The Relativity Effect in Planetary Motions”, *Reviews of Modern Physics*, Vol. 19, 1947, pp. 361-364

26. Nobili, A. M., Will, C. M., “The real value of Mercury’s perihelion advance”, *Nature*, Vol. 320, 1986, pp. 39 - 41

26. Folker, W. M., Williams, J. G., Boggs, D. H., Park, R. S. and Kuchynka, P., “The Planetary and Lunar Ephemerides DE430 and DE431”, *The Interplanetary Network Progress Report*, Volume 42, pp. 1-81

27. Horner, J., Marshall, J. P., Wittenmyer, R. A. and Tinney, C. G., “A dynamical analysis of the proposed HU Aquarii planetary system”, *Monthly Notices of the Royal Astronomical Society*, Volume 416, 2011, pp. 11 - 15

28. Horner, J., Hinse, T. C., Wittenmyer, R. A., Marshall, J. P. and Tinney, C. G., “A dynamical analysis of the proposed circumbinary HW Virginis planetary system”, *Monthly Notices of the Royal Astronomical Society*, Volume 427, 2012, pp. 2812 - 2823

29. Wittenmyer, R. A., Horner, J. and Marshall, J. P., “On the dynamical stability of the proposed planetary system orbiting NSVS 14256825”, *Monthly Notices of the Royal Astronomical Society*, Volume 431, 2013, pp. 2150 - 2154

30. Horner, J., Wittenmyer, R. A., Hinse, T. C., Marshall, J. P., Mustill, A. J. and Tinney, C. G., “A detailed dynamical investigation of the proposed QS Virginis planetary system”, *Monthly Notices of the Royal Astronomical Society*, 435, 2013, pp. 2033 - 2039

31. Horner, J., Lykawka, P. S., 2010, “2001 QR322: a dynamically unstable Neptune Trojan?”, *Monthly Notices of the Royal Astronomical Society*, 405, 49-56

32. Horner, J., Wittenmyer, R. A., Hinse, T. C., Tinney, C. G., “A detailed investigation of the proposed NN Serpentis planetary system”, *Monthly Notices of the Royal Astronomical Society*, 425, 2012, 749-756

33. Gilmore, J. B., "How reliable are latitudinal energy balance models for habitability calculations when using Earth's radiative properties?", *Monthly Notices of the Royal Astronomical Society*, Vol. 440, 2014, pp. 1435 – 1445

34. Waltham, D., "Testing Anthropic Selection: A Climate Change Example", *Astrobiology*, Vol. 11, 2011, pp. 105-114

35. Dehant, V., Lammer, H., Kulikov, Y. N., Grießmeier, J.-M., Breuer, D., Verhoeven, O., Karatekin, Ó., vanHoolst, T., Korablev, O. and Lognonné , P., "Planetary Magnetic Dynamo Effect on Atmospheric Protection of Early Earth and Mars," , vol. 129, pp. 279–300, Mar. 2007.

36. Lundin, R., Lammer, H. and Ribas, I., "Planetary Magnetic Fields and Solar Forcing: Implications for Atmospheric Evolution", *Space Science Reviews*, Vol. 129, 2007, pp. 245-278

37. Nicholson, B., Mengel, M., Brookshaw, L., Vidotto, A. A., Carter, B., Marsden, S., Soutter, J., Waite, I. and Horner, J., "Winds Of Planet Hosting Stars", *Proceedings of the 14th Australian Space Research Conference*, submitted, 2015

38. Horner, J. and Jones, B. W., "Jupiter friend or foe? I: The asteroids", *International Journal of Astrobiology*, 7, 2009, pp. 251-261

39. Horner, J. and Jones, B. W., "Jupiter - friend or foe? II: the Centaurs", *International Journal of Astrobiology*, 8, 2009, pp. 75-80

40. Horner, J., Jones, B. W. and Chambers, J., "Jupiter - friend or foe? III: the Oort cloud comets", *International Journal of Astrobiology*, 9, 2010, pp. 1-10

41. Horner, J. and Jones, B. W., "Jupiter - friend or foe? IV: the influence of orbital eccentricity and inclination", *International Journal of Astrobiology*, 11, 2012, pp.147-156

42. Horner, J., Mousis, O., Petit, J.-M. and Jones, B. W., "Differences between the impact regimes of the terrestrial planets: Implications for primordial D:H ratios", *Planetary and Space Science*, 57, 2009, pp. 1338-1345

43. Contro, B., Wittenmyer, R. A., Horner, J. and Marshall, J. P., "Towards a dynamics-based estimate of the extent of HR 8799's unresolved warm debris belt", *Proceedings of the 14th Australian Space Research Conference*, submitted, 2015

44. Waltham, D., "Anthropic Selection for the Moon's Mass", *Astrobiology*, Vol. 4, 2004, pp. 460-468

45. Waltham, D., "The large-moon hypothesis: can it be tested?", *International Journal of Astrobiology*, Vol. 5, 2006, pp. 327-331

46. Fressin, F., Torres, G., Charbonneau, D., Bryson, S. T., Christiansen, J., Dressing, C. D., Jenkins, J. M., Walkowicz, L. M. and Batalha, N. M., "The False Positive Rate of Kepler and the Occurrence of Planets", *The Astrophysical Journal*, Volume 766, 2013, article id. 81

47. Howard, A. W., Marcy, G. W., Bryson, S. T., Jenkins, J. M., Rowe, J. F., Batalha, N. M., Borucki, W. J., Koch, D. G., Dunham, E. W., Gautier, T. N., III, Van Cleve, J., Cochran, W. D., Latham, D. W., Lissauer, J. J., Torres, G., Brown, T. M., Gilliland, R. L., Buchhave, L. A., Caldwell, D. A., Christensen-Dalsgaard, J., Ciardi, D., Fressin, F., Haas, M. R., Howell, S. B., Kjeldsen, H., Seager, S., Rogers, L., Sasselov, D. D., Steffen, J. H., Basri, G. S., Charbonneau, D., Christiansen, J., Clarke, B., Dupree, A., Fabrycky, D. C., Fischer, D. A., Ford, E. B., Fortney, J. J., Tarter, J., Girouard, F. R., Holman, M. J., Johnson, J. A., Klaus, T. C., Machalek, P., Moorhead, A. V., Morehead, R. C., Ragozzine, D., Tenenbaum, P., Twicken, J. D., Quinn, S. N., Isaacson, H., Shporer, A., Lucas, P. W., Walkowicz, L. M., Welsh, W. F., Boss, A., Devore, E., Gould, A., Smith, J. C., Morris, R. L., Prsa, A., Morton, T. D., Still, M., Thompson, S. E., Mullally, F., Endl, M., and MacQueen, P. J., "Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler", *The Astrophysical Journal Supplement*, Volume 201, 2012, article id. 15

48. Howard A. W., Marcy, G. W., Johnson, J. A., Fischer, D. A., Wright, J. T., Isaacson, H., Valenti, J. A., Anderson, J., Lin, D. N. C. and Ida, S., "The Occurrence and Mass Distribution of Close-in Super-Earths, Neptunes, and Jupiters", *Science*, Volume 330, 2010, pp. 653

49. Wittenmyer, R. A., Tinney, C. G., Butler, R. P., O'Toole, S. J., Jones, H. R. A., Carter, B. D., Bailey, J. and Horner, J., "The Frequency of Low-mass Exoplanets. III. Toward $\eta_{\oplus}$ at Short Periods", *The Astrophysical Journal*, Volume 738, 2011, article id. 81

50. Swift, J. J., Bottom, M., Johnson, J. A., Wright, J. T., McCrady, N., Wittenmyer, R. A., Plavchan, P., Riddle, R., Muirhead, P. S., Herzig, E., Myles, J., Blake, C. H., Eastman, J., Beatty, T. G., Lin, B., Zhao, M., Gardner, P., Falco, E., Criswell, S., Nava, C., Robinson, C., Sliski, D. H., Hedrick, R., Ivarsen, K., Hjelstrom, A., de Vera, J. and Szentgyorgyi, A., "Miniature Exoplanet Radial Velocity Array (MINERVA) I. Design, Commissioning, and First Science Results", *SPIE, Astronomical Telescopes + Instrumentation,* in press, 2015, arXiv:1411.3724

51. Eastman, J. D., Brown, T. M., Hygelund, J., van Eyken, J., Tufts, J. R. and Barnes, S., "NRES: the network of robotic Echelle spectrographs", *Proceedings of the SPIE,* Volume 9147, 2014, id. 914716 pp. 15

52. Howell, S. B., Sobeck, C., Haas, M., Still, M., Barclay, T., Mullully, F., Troeltzsch, J., Aigrain, S., Bryson, S. T., Caldwell, D., Chaplin, W. J., Cochran, W. D., Huber, D., Marcy, G. W., Miglio, A., Najita, J. R., Smith, M., Twicken, J. D. and Fortney, J. J., "The K2 Mission: Characterization and Early Results", *Publications of the Astronomical Society of the Pacific*, Volume 126, Issue 938, 2014, pp. 398-408

53. Ricker, G. R., Winn, J. N., Vanderspek, R., Latham, D. W., Bakos, G. Á., Bean, J. L., Berta-Thompson, Z. K., Brown, T. M., Buchhave, L., Butler, N. R., Butler, R. P., Chaplin, W. J., Charbonneau, D., Christensen-Dalsgaard, J., Clampin, M., Deming, D., Doty, J., De Lee, N., Dressing, C., Dunham, E. W., Endl, M., Fressin, F., Ge, J., Henning, T., Holman, M. J., Howard, A. W., Ida, S., Jenkins, J., Jernigan, G., Johnson, J. A., Kaltenegger, L., Kawai, N., Kjeldsen, H., Laughlin, G., Levine, A. M., Lin, D., Lissauer, J. J., MacQueen, P., Marcy, G., McCullough, P. R., Morton, T. D., Narita, N., Paegert, M., Palle, E., Pepe, F., Pepper, J., Quirrenbach, A., Rinehart, S. A., Sasselov, D., Sato, B., Seager, S., Sozzetti, A., Stassun, K. G., Sullivan, P., Szentgyorgyi, A., Torres, G., Udry, S. and

Villasenor, J., "Transiting Exoplanet Survey Satellite (TESS)", *Proceedings of the SPIE*, Volume 9143, 2014, id. 914320

. Rauer, H., Catala, C., Aerts, C., Appourchaux, T., Benz, W., Brandeker, A., Christensen-Dalsgaard, J., Deleuil, M., Gizon, L., Goupil, M.-J., Güdel, M., Janot-Pacheco, E., Mas-Hesse, M., Pagano, I., Piotto, G., Pollacco, D., Santos, Ċ., Smith, A., Suárez, J.-C., Szabó, R., Udry, S., Adibekyan, V., Alibert, Y., Almenara, J.-M., Amaro-Seoane, P., Eiff, M. Ammler-von, Asplund, M., Antonello, E., Barnes, S., Baudin, F., Belkacem, K., Bergemann, M., Bihain, G., Birch, A. C., Bonfils, X., Boisse, I., Bonomo, A. S., Borsa, F., Brandão, I. M., Brocato, E., Brun, S., Burleigh, M., Burston, R., Cabrera, J., Cassisi, S., Chaplin, W., Charpinet, S., Chiappini, C., Church, R. P., Csizmadia, Sz., Cunha, M., Damasso, M., Davies, M. B., Deeg, H. J., Díaz, R. F., Dreizler, S., Dreyer, C., Eggenberger, P., Ehrenreich, D., Eigmüller, P., Erikson, A., Farmer, R., Feltzing, S., de Oliveira Fialho, F., Figueira, P., Forveille, T., Fridlund, M., García, R. A., Giommi, P., Giuffrida, G., Godolt, M., Gomes da Silva, J., Granzer, T., Grenfell, J. L., Grotsch-Noels, A., Günther, E., Haswell, C. A., Hatzes, A. P., Hébrard, G., Hekker, S., Helled, R., Heng, K., Jenkins, J. M., Johansen, A., Khodachenko, M. L., Kislyakova, K. G., Kley, W., Kolb, U., Krivova, N., Kupka, F., Lammer, H., Lanza, A. F., Lebreton, Y., Magrin, D., Marcos-Arenal, P., Marrese, P. M., Marques, J. P., Martins, J., Mathis, S., Mathur, S., Messina, S., Miglio, A., Montalban, J., Montalto, M., Monteiro, M. J. P. F. G., Moradi, H., Moravveji, E., Mordasini, C., Morel, T., Mortier, A., Nascimbeni, V., Nelson, R. P., Nielsen, M. B., Noack, L., Norton, A. J., Ofir, A., Oshagh, M., Ouazzani, R.-M., Pápics, P., Parro, V. C., Petit, P., Plez, B., Poretti, E., Quirrenbach, A., Ragazzoni, R., Raimondo, G., Rainer, M., Reese, D. R., Redmer, R., Reffert, S., Rojas-Ayala, B., Roxburgh, I. W., Salmon, S., Santerne, A., Schneider, J., Schou, J., Schuh, S., Schunker, H., Silva-Valio, A., Silvotti, R., Skillen, I., Snellen, I., Sohl, F., Sousa, S. G., Sozzetti, A., Stello, D., Strassmeier, K. G., Švanda, M., Szabó, Gy. M., Tkachenko, A., Valencia, D., Van Grootel, V., Vauclair, S. D., Ventura, P., Wagner, F. W., Walton, N. A., Weingrill, J., Werner, S. C., Wheatley, P. J. and Zwintz, K., "The PLATO 2.0 mission", *Experimental Astronomy*, Volume 38, Issue 1-2, 2014, pp. 249-330

The Dynamics of Centaurs in the Vicinity of the 2:1 Mean Motion Resonance of Neptune and Uranus Trojan Region

Jeremy Wood[1,2] Jonathan Horner[2,3]

[1] Hazard Community and Technical College 1 Community College Drive Hazard, Ky USA 41701
[2] University of Southern Queensland West St, Toowoomba QLD 4350, Australia
[3] Australian Centre for Astrobiology, UNSW Australia, Sydney, NSW 2052, Australia

Summary: In this work we present the results of a suite of dynamical simulations following the orbital evolution of 8,022 hypothetical "Centaur" objects. These Centaurs begin our integrations on orbits trapped within the 2:1 mean motion resonance with Neptune, and we follow their dynamical evolution for a period of 3 Myr under the gravitational influence of a motionless Sun and the four Jovian planets.

The great majority of the test particles studied rapidly escaped from Neptune's 2:1 mean motion resonance, and diffused throughout the outer Solar system. The average libration time of Centaurs in the 2:1 mean motion resonance of Neptune was found to be just 27 kyr – although two particles did remained trapped in that resonance for more than 1 Myr. Upon leaving the 2:1 resonance, the majority of test particles evolved by a process of random walk in semi-major axis, due to repeated close encounters with the giant planets.

Keywords: Centaurs, Celestial Mechanics: Mean Motion Resonances, methods: *n*-body simulations, Uranus Trojans, Comets

Introduction

Small bodies in the Solar system exist in orbits that extend from within that of the Earth to beyond Neptune's, and are collectively the detritus left behind from the formation of our planetary system. The hazard that these small bodies pose to Earth has been brought to public attention in recent years by the meteor strike in Chelyabinsk, Russia, in Feb. 2013 that injured over 1,100 people [1]. Numerous very close approaches between small bodies and the planets have also been observed, such as the near miss of 2012 DA14 [2][3] and 2014 HQ124 [4] with Earth and the close approach of comet C/2013 A1 (Siding Spring) [5][6][7] to Mars in Oct. 2014.

Two spectacular collisions between a planet and a small body occurred in the 20th century. In 1908 a small body exploded over Tunguska, Russia flattening an area of 2,150 km^2 [8][9][10] and in 1994, comet D/1993 F2 (Shoemaker-Levy 9) split into 21 pieces before colliding with Jupiter [11]. The collision left dark spots on the planet which lasted over a year [12]. The size and persistence of the scars left behind by the impacting cometary fragments illustrates the damage collisions between small bodies and planets can cause. Indeed, the energy contained in an impacting small body can be enough to cause an extinction level event. Such collisions have been proposed in order to explain a number of historical mass extinction events – most famously that which occurred some 65 Myr ago, resulting in the death of the dinosaurs [13]. Thus, one reason for studying small bodies of the Solar system is because of the potential hazard they pose.

Small bodies with initial orbits near Earth are not the only ones with the potential for an Earth strike. The immediate threat to Earth comes from three distinct populations of objects – the Near-Earth Objects (such as the Chelyabinsk impactor), the short-period comets (such as comet Shoemaker-Levy 9), and the long-period comets (such as comet Siding Spring). Each of these threatening populations are dynamically unstable, and are continually replenished from reservoirs of rocky or icy bodies further from the Sun [14][15][16][17]. The exact definition of each class of small body differs from one researcher to the next. Traditionally, Near Earth Objects (or NEOs) have been classified into three groups [18]. These are:

Atens – Objects with orbital semi-major axes less than that of the Earth whose orbits cross our own
Apollos – Objects with orbital semi-major axes greater than that of the Earth's, whose orbits cross the Earth's
Amors – Objects with orbits exterior to Earth's, and perihelia within 1.3 AU, that do not cross Earth's orbit

More recently, the discovery of Near-Earth Asteroids whose orbits are entirely within that of the Earth has resulted in the addition of a fourth sub-class – the Atira or Apohele asteroids, whose aphelia are closer to the Sun than the Earth's perihelion distance (e.g. [19]). In addition, those NEOs that can approach to within 0.05 AU of Earth are classified as Potentially Hazardous Astereoids or PHAs [20][21]. Currently 1,548 PHAs are known [20].

Objects that display noticeable activity (such as outgassing, or the presence of a tail) are classified as cometary. Amongst those bodies, the short-period comets are defined to be those having an orbital period of less than 200 years [22]. The Minor Planet Center refers to these as periodic comets [23]. The population of short period comets is replenished by the Centaurs [24]. The Centaurs are a group of small rock-ice bodies with semi-major axes between Jupiter and Neptune and perihelia beyond that of Jupiter's [21][22][25]. The frequency of the flux of Centaurs into a new Earth-crossing orbit is estimated to be 1 every 880 years [26]. Though the possibility of an eventual Earth collision with a Centaur is an order of magnitude lower than that of a Main Belt Asteroid, the threat is non-zero and more likely than a collision with a Long Period Comet [14][15][17].

Numerical studies show that small bodies in orbits beyond Earth's can be transported to the inner Solar system over time via the gravitational perturbation of their orbits [27]. Over their lifetime, small bodies may also enter into a mean motion orbital resonance with a planet. A mean motion orbital resonance occurs when the orbital periods of two bodies exist in a ratio of two small integers $r{:}\,s$. For example, Saturn orbits the Sun in about 29.5 years and Jupiter in about 11.86 years. Dividing these orbital periods yields almost 2.5 which is equal to the ratio 5:2. Thus, it can be said that Saturn and Jupiter are nearly in a 5 to 2 mean motion orbital resonance. In other words, Jupiter completes five complete orbits in the time it takes Saturn to complete two. For a body to be considered stuck in a resonance, the principal resonant angle associated with that resonance must also librate or change very slowly [40]. In the case of the 2:1 mean motion resonance of Neptune, the principal resonant angle is defined by

$$2\lambda_N - \lambda - \varpi \tag{2}$$

Where λ_N is the mean longitude of Neptune, λ is the mean longitude of the small body ($\lambda = M + \varpi$, where M is the mean anomaly) and ϖ is the longitude of perihelion of the small body.

Small bodies in mean motion resonances can either be in stable or unstable orbits. Some resonances are unstable due to the fact that they overlap with others. Over time, the eccentricity of these orbits increases leading to planetary orbit crossings or the objects colliding with the Sun [28][29]. Small bodies remain in these resonances for relatively short amounts of time. However, populations of small bodies also exist in stable mean motion resonances where they can remain for relatively longer periods of time. Examples include Jupiter Trojans (1:1 resonance with Jupiter; e.g. [30][31]) Neptune Trojans (1:1 resonance with Neptune; e.g. [32]), Hilda asteroids (3:2 resonance with Jupiter; e.g. [33]) and Plutinos (2:3 resonance with Neptune; e.g. [34]).

When a small body enters a stable mean motion orbital resonance with a planet, it becomes stuck in the resonance meaning that the semi-major axis of its orbit librates (oscillates in a quasi-periodic fashion) about the resonance location. This is known as resonance sticking [29][35].

Small bodies which orbit between Mars and Jupiter (in the asteroid belt) have received more attention in the literature than those that orbit between Jupiter and Neptune (the Centaurs) [36]. In this work, we present the initial results of a study of the dynamical behavior of bodies trapped in the 2:1 mean motion resonance of Neptune, and their orbital evolution after leaving it. First, in the following section, the theory of mean motion resonances, taxonomy of small bodies, and dynamical classification of Centaurs are discussed. We then present the method used to carry out the study, and describe how the resulting data were analyzed. Finally, we present and discuss our results, and detail our plans for future work.

Theory

The location of a given mean motion resonance for a planet, a_{MMR}, can be derived using the semi-major axis, a_p, of the planet, and the orbital period ratio of the resonance in question (s/r), using Kepler's 3rd Law. The result is:

$$a_{MMR} = a_p\left(\frac{s}{r}\right)^{2/3} \quad (1)$$

[37]. Here a_p is a planet's semi-major axis. The resonance is exterior to the planet's orbit if $r < s$ and interior if $r > s$. If $r = s$ then any small body in that resonance is called a Trojan asteroid [37][38]. In this paper, interior resonances are stated using the following format: "larger number:smaller number". For example the 2:1 mean motion resonance of Neptune is interior to the planet Neptune, but the 1:2 resonance is exterior to Neptune. Resonances actually exist in a region of a-e space, rather than at just one location, so it is possible that two resonances can overlap [39]. The location of the 2:1 mean motion resonance of Neptune is centered at 18.94 AU from the Sun. This is very close to the orbit of the planet Uranus, which lies at a = 19.19 AU. Neptune and Uranus are therefore very nearly in a 2:1 resonance with each other (with a period ratio of ~1.96), and as a result, the 2:1 resonance of Neptune and the Trojan region of Uranus overlap in a-e space. This potentially increases the instability of the orbits of any small bodies librating in this region [37][40].

The gravitational influence of the planets also perturbs the osculating orbital parameters of small bodies, even when the small body is not in a mean motion resonance. This is especially true during close planetary encounters. As bodies are often classified using osculating orbital parameters, the classification of a small body will often change during its dynamical lifetime [24]. Common orbital quantities used to classify small bodies include the semi-major axis, a, eccentricity, e, perihelion distance, q, and the Tisserand parameter, T_p, defined as:

$$T_p = \frac{1}{\frac{a}{a_p}} + 2\sqrt{\frac{a}{a_p}(1-e^2)}\cos\Delta i \quad (3)$$

Where Δi is the small body's orbital inclination with respect to the plane of the planet's orbit [41]. At any moment in time, the Tisserand parameter for a given small body will be different for the different planets. The Tisserand parameters of Jupiter, Saturn, Uranus and Neptune can be denoted as $\mathrm{T_J}$, $\mathrm{T_S}$, $\mathrm{T_U}$ and T_N respectively.

Over the years, a variety of different classification schemes have been put forward to identify different types of Solar system small bodies (e.g. [42][22][25]). In this work, we follow the classification scheme detailed in Table 1.

Table 1 One possible taxonomy for small bodies of the Solar system.

TYPE	$\boldsymbol{T_J}$	***a*** **(AU)**	***q*** **(AU)**	**REFERENCE**
Centaur	-	$a_{\mathrm{Jupiter}} < a < a_{\mathrm{Neptune}}$	$> a_{\mathrm{Jupiter}}$	[21]

Encke-Type Comet	> 3	< a_{Jupiter}	-	[43]
Jupiter Family Comet	$2 < T_J < 3$	-	-	[43]
Halley-Type Comet	< 2	< 40	-	[42]
Trans Neptunian Object (TNO)	-	> a_{Neptune}	-	[21]
Ambi-Neptunian Object		> a_{Neptune}	< a_{Neptune}	
Kuiper Belt Object (KBO)	-	$a_{\text{Neptune}} < a \leq 48$	-	[35]
Scattered Disk Object (SDO)	-	$48 < a \leq 1{,}000$	> a_{Neptune}	[35]
Oort Cloud Object	-	> 1,000	-	[44]

As for Centaurs, [41] and [36] state that Centaurs can be categorized in one of two dynamical classes referred to as 'random walk' and 'resonance hopping'. If the standard deviation, σ, of the semi-major axis values of a Centaur varies in time according to a power law given by:

$$\sigma = (2Dt)^H \tag{4}$$

then the Centaur is classified as a random-walk Centaur. Here t is time, D is the diffusion coefficient and H is the Hurst exponent [45]. In this case the semi-major axis is also said to undergo diffusion. From this equation it can also be shown that

$$log_{10}(\sigma) = Hlog_{10}(t) + constant \tag{5}$$

i.e. the log of the standard deviation varies linearly with the log of time. Centaurs whose semi-major axis values are not dominated by diffusion tend to jump from one resonance to another. These bodies are classified as resonance-hopping Centaurs. Resonance sticking dominates their dynamics. Centaurs may also display behavior of both types during their lifetimes [41].

Method

8,022 massless test particles were integrated for 3 Myr in the 6-Body problem (a motionless Sun; Jupiter, Saturn, Uranus, Neptune and test particle) using the regularized mixed-variable symplectic (RMVS) method in the SWIFT software package [46]. The time step used was 40 days, and data were output at intervals of 2 kyr. The initial values for a, e and i of the test particle orbits were randomly chosen from these ranges: 18.92 AU $\leq a \leq$ 19.16 AU, $0 < e \leq 0.7$, and $0° \leq i \leq 40°$. The initial a, e and i of the orbits of the giant planets were obtained directly from the standard SWIFT installation. The unusual range of a was because of the way in which the SWIFT integrator behaves. An a range symmetric about the resonance location was input, however this is the range that SWIFT used. Though odd, the range is totally usuable for this study. Both the planets and the test particles were initially placed at random locations in their orbits. The arguments of perihelion and longitudes of the ascending node were also randomised. Test particles were removed from the simulation by colliding with a planet, approaching too close to the Sun (~0.005 AU), obtaining a parabolic or hyperbolic orbit ($e \geq 1$), or by entering the Oort Cloud (defined as reaching a barycentric distance of 1,000 AU).

Once the simulations were complete, a random sample of 75 test particles was chosen, and their orbital elements plotted as a function of time, in order to determine approximate maximum allowed amplitudes of osculating and average semi-major axis within which a test particle is considered to display libration-like behavior about the 2:1 Neptunian mean motion resonance. Since the chosen data output period was too large for the true resonant angles to be determined (a compromise chosen to ensure the simulations could run in a reasonable amount of time), test particles are said to display only pseudo-librational behavior about the resonance, rather than true libration.

To determine the libration times, the dynamical lifetime of each test particle is divided into 10 kyr windows of time. If the last window has a duration of < 10 kyr then it is ignored. For example, if the dynamical lifetime of a test particle is 216 kyr then its lifetime is divided into 21 10 kyr windows, and the last 6 kyr are ignored.

To determine the longest consecutive libration time of a test particle, the average semi-major axis in each window is calculated and subtracted from the resonance location to calculate the amplitude. The maximum osculating semi-major axis value in the window is also determined. If both the amplitude of the average and osculating values of a are smaller than their respective maximum limits of 0.125 AU and 1.15 AU respectively from the resonance location, then the test particle is said to display libration-like behavior in the window. The largest number of consecutive windows displaying libration-like behavior is then determined, and multiplied by 10 kyr to find the total libration time.

The dynamical class of a sample of test particles is determined using the quantitative method of [41]. This method begins by finding the time interval within which 10 values of osculating semi-major axis occur. The dynamical lifetime is then divided into windows of time each equal to this interval, allowing consecutive windows to overlap by half a time interval. Any windows of duration longer than one-quarter the total time are discarded. The standard deviation of semi-major axis in each time window is calculated, and then the values are averaged over all windows. Following this, the process is repeated using up to 16 different window time intervals. To find the duration of other time intervals, we take the log base ten of the total number of data points during the dynamical lifetime of the test particle and subtract the log base ten of the fewest data points in any window (which is 10) and then dividing the difference by 16. Finally, find the slope of the graph of $\log_{10}$ of average standard deviation vs. $\log_{10}$ of window length. Following [41], we classify a test particle as "random-walk" if the slope of this graph lies between 0.22 and 0.95 and the correlation coefficient is > 0.85. If the correlation coefficient is between 0.7 and 0.85 the dynamical class is found qualitatively. We consider a test particle to be a "resonance-hopping"

Centaur if the slope is outside of this range and the correlation coefficient is < 0.9. Otherwise the dynamical class is found qualitatively.

Results and Discussion

Figure 1 shows the distribution of the test particles studied in this work in semi-major axis, at the start and end of our simulations (i.e. at t = 0 and 3 Myr). Of the 8,022 particles studied, 4,967 test particles survive after the full integration time, and move with semi-major axes in the range 5 AU – 35 AU at that time. Comparison of the two histograms shows that many of the test particles have diffused throughout the Solar system, escaping from the vicinity of the 2:1 resonance with Neptune in the process. The shape of the histogram at 3 Myr is roughly Gaussian with a peak within 19 AU – 19.5 AU. This overlaps the initial region and the orbit of Uranus.

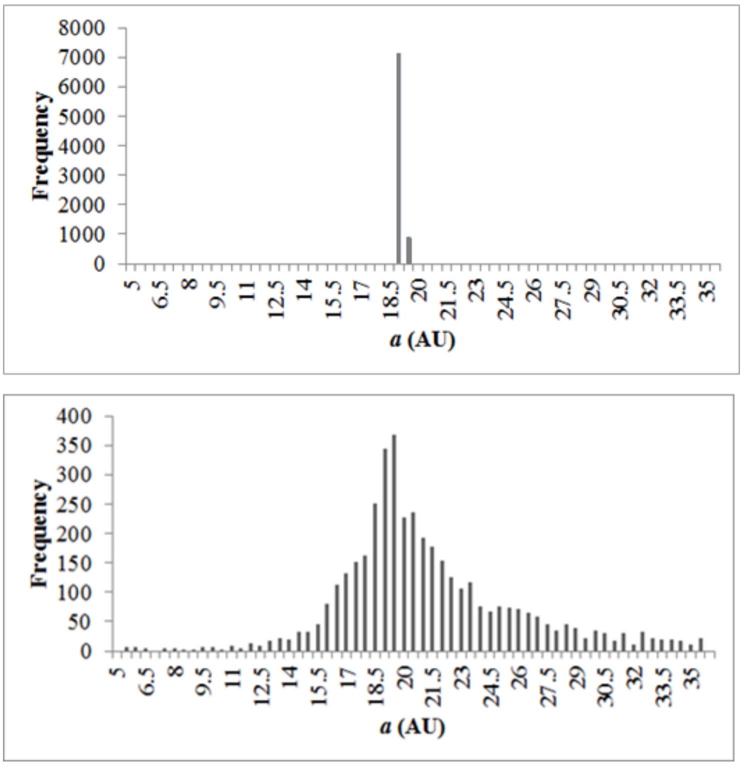


Figure 1. A histogram of the semi-major axis of test particles at the start (top; t = 0 Myr) and end (bottom; t = 3 Myr) of our simulations. The bin size is 0.5 AU for both histograms. At t = 0 Myr, 7,128 test particles are in the bin between 18.5 AU and 19 AU, and 894 test particles are in the bin between 19 AU and 19.5 AU. At the end of the simulations, 4,967 test particles survive in the range 5 AU < a < 35 AU. The diffusion of the test particles over time is clearly seen.

The number of test particles entering each class of small body at some point during their lifetime and the average time spent in each class is shown in Table 2. Of the classes inward from the Centaur region, the Jupiter Family Comet class (JFC) had the highest percentage of test particles entering the class, at 35.7%. Only 1.9% of our test particles became Encke-type comets. This result is not particularly unexpected – Encke-type comets are almost decoupled from the influence of Jupiter, and so are relatively dynamically stable. However, that stability in turn means that it is relatively difficult for objects to become captured on such orbits in the first place. 5.7% of the test particles studied became Mars crossers and a mere 3.1% became Earth crossers. Thus, the possibility of a Centaur coming from an orbit near Uranus and colliding with Earth within 3 Myr is small but nonzero. Of the classes outward from the Centaur region, KBO had the highest percentage of test particles entering the class at 53.7% and Ambi-Neptunian Object was 2nd at 47.8%. Only 0.11% became SDOs.

In regards to time spent in each class, the Centaur class had the highest average percentage of occupancy time of any class at 75% of a test particle's lifetime. Earth crossers and Mars crossers had the lowest average percentages of occupancy time overall at 3% and 3.7% respectively. JFCs spent on average 26% of their lives in the class. Encke-type comets had the highest average time of occupancy of the inward classes at 33%. Of the classes outward from the Centaur region KBO had the highest average occupancy time at 14.9% of a test particle's lifetime. The Ambi-Neptunian Object class had an average occupancy time of 27%.

About one-third of the test particles were removed from the integration by entering the Oort Cloud, obtaining a barycentric distance greater than 1,000 AU or by solar system ejection. 1.9% of test particles collided with the Sun.

Table 2 the number of test particles that entered each class of small body out of 8,022 and the average percentage of their lives spent in that class once there.

Class	**Oort Cloud**	**Encke-Type Comet**	**JFC**	**Centaur**	**Halley-Type Comet**	**KBO**	**SDO**	**Earth Crosser**
TOTAL	399	154	2,864	8,022	109	4,312	9	249
Avg %Time	-	33%	26%	75%	14%	14.9%	10%	3%

Class	**Ambi-Neptunian Object**	**Mars Crosser**
TOTAL	3,831	455
Avg %Time	27%	3.7%

Qualitative analysis showed that, whilst a given test particle displayed libration-like behavior about the 2:1 resonance, its osculating semi-major axis remained within 1.15 AU of the resonance location, and its mean semi-major axis remained within 0.125 AU of that location while a test particle displayed libration-like behavior about the resonance. Using these bounds, it was discovered that 2,036 test particles librated within the resonance for at least 10 kyr, 97 for 100 kyr or more and 2 for more than 1 Myr. The mean libration time for test particles that were trapped for at least 10 kyr was found to be 27 kyr, and the median time was 10 kyr. The longest libration time observed was 1,100 kyr, as shown in Figure 2.

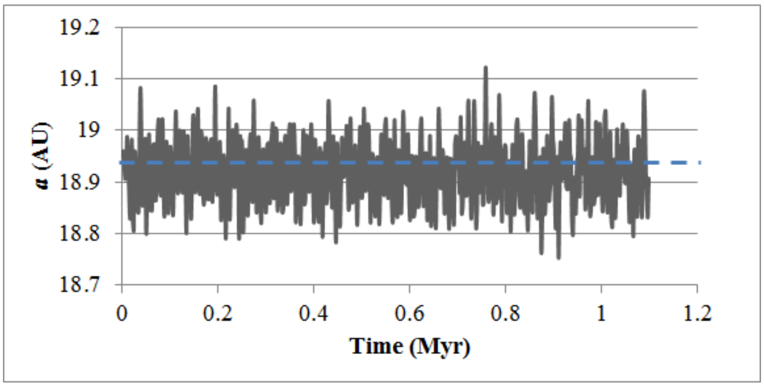


Figure 2. The libration of the most long lived 2:1 resonant object in our sample. The dashed line denotes the central location of the 2:1 resonance with Neptune. The initial orbital parameters for this test particle were a = 18.95 AU, e = 0.205 and i = 29.2°.

The dynamical classes of a random sample of 218 Earth-crossing test particles were determined. Given the recent interest in potentially hazardous objects, the particles that became Earth-crossing in this study are of particular interest. 35% of the test particles in the sample were identified as resonance-hopping Centaurs, whilst 65% were found to be random-walk Centaurs. Thus, the most likely method of dynamical transport of a Centaur to the inner Solar system is a random walk. Two case studies are shown in Figures 3 and 4.

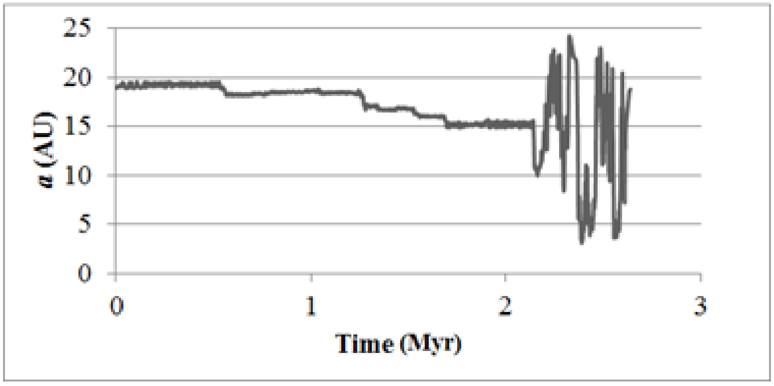


Figure 3. The first 2 Myr in the evolution of the semi-major axis of an exemplar test particle that eventually became an Earth-crossing object. Throughout this period, the particle is best classified as a resonance-hopping Centaur. Note the periods of libration in various resonances, which appear as nearly horizontal bands in the graph. This test particle later became an Earth crosser at 2.39 Myr. Data points beyond 25 AU are not shown, for clarity. Its initial orbital parameters were a = 18.99 AU, e = 0.063 and i = 37° respectively.

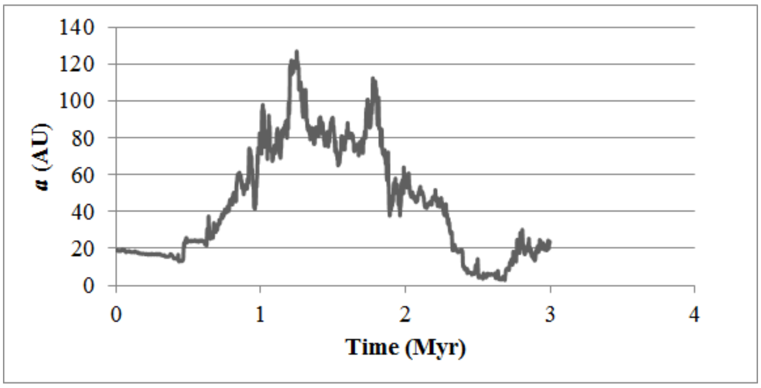


Figure 4. The evolution of the semi-major axis of an exemplar test particle that evolved to become an Earth-crossing object after a protracted period of random walking throughout the outer Solar system. It is classified as a random-walk Centaur. Note the lack of horizontal bands in the graph (aside from a short resonant capture at ~600 kyr). The test particle became an Earth crosser at about 2.65 Myr. Its initial orbital parameters were a = 19.02 AU, e = 0.44 and i = 23° respectively.

Conclusions and Future Work

Dynamical simulations were carried out of the evolution of 8,022 massless test particles, initially located in the vicinity of the 2:1 Neptunian mean-motion resonance, over a period of 3 Myr. The average libration time of test particles displaying libration-like behavior in the vicinity of the resonance for at least 10 kyr was found to be 27 kyr, and the median time was 10 kyr.

Just over one-third of test particles became JFCs, and just over half became KBOs. A negligible amount of test particles became SDOs which shows that most Centaurs do not evolve into orbits beyond the Kuiper belt with perihelia beyond Neptune's orbit. Less than 6% became Mars crossers and 3.1% Earth Crossers. This shows that the odds of a Centaur diffusing from an orbit near Uranus to an Earth-crossing orbit within 3 Myr is low but non-zero. Thus, this study supports the idea that Centaurs do represent a threat to Earth.

Of the taxonomical classes inward from the resonance, the Encke-type comet class had the highest average occupancy time at 33%. Of the outward classes, KBO had the highest average occupancy time at 14.9% of a test particle's lifetime.

65% of a sample of Earth crossers were random walkers and 35% resonance-hoppers. Graphs of semi-major axis vs. time for resonance-hopping Centaurs tend to have long horizontal bands, and those of the random-walk Centaurs tend to lack long horizontal bands.

Future work will be to determine the separatrices (boundaries) of the resonance in *a*-*e* space and the likelihood of capture as a Trojan of any Jovian planet.

References

1. Kohout, T., Gritsevich, M., Grokhovsky, V., Yakovlev, G., Haloda, J., Halodova, P., Michallik, R., Penttilä, A. and Muinonen, K., "Mineralogy, reflectance spectra, and physical properties of the Chelyabinsk LL5 chondrite - Insight into shock-induced changes in asteroid regoliths", *Icarus*, Vol. 228, 2014, pp. 78-85

2. Moskovitz, Nicholas, Endicott, T., Lister, T., Ryan, B., Ryan, E., Willman, M., Hergenrother, C., Binzel, R., Polishook, D., DeMeo, F., Benecchi, S., Sheppard, S., Marchis, F., Angusteijn, T., Birthwhistle, P., Verveer, A., Gulbis, A., Nagayama, T., Gilmore, A. and Kilmartin, P., "The Near-Earth Flyby of Asteroid 2012 DA14", *American Astronomical Society, DPS meeting #45*, 2013

3. de León, J., Ortiz, J. L., Pinilla-Alonso, N., Cabrera-Lavers, A., Alvarez-Candal, A., Morales, N., Duffard, R., Santos-Sanz, P., Licandro, J., Pérez-Romero, A., Lorenzi, V. and Cikota, S., "Visible and near-infrared observations of asteroid 2012 DA14 during its closest approach of February 15, 2013", *A&A*, Vol. 555, 2013, pp. L2

4. MPCHQ124web: MPCMPEC 2014-H67: 2014 HQ124, http://www.minorplanetcenter.net/mpec/K14/K14H67.html (accessed October 28, 2014)

5. MPCSidingSpringweb: Elements and Ephemeris for C/2013 A1 (Siding Spring), http://minorplanetcenter.net/cgi-bin/returnprepeph.cgi?d=c&o=CK13A010 (accessed October 28, 2014)

6. Farnocchia, D., Chesley, S. R., Chodas, P. W., Tricarico, P., Kelley, M. S. P. and Farnham, T. L., "Trajectory Analysis for the Nucleus and Dust of Comet C/2013 A1 (Siding Spring)", *AJ*, Vol. 790, No. 2, 2014, article id. 114

7. Ye, Q.-Z. and Hui, M.-T., "AN EARLY LOOK OF COMET C/2013 A1 (SIDING SPRING): BREATHTAKER OR NIGHTMARE?", *AJ*, Vol. 787, No. 2, 2014, article id. 115

8. Turco, R. P., Toon, O. B., Park, C., Whitten, R. C., Pollack, J. B. and Noerdlinger, P., "An analysis of the physical, chemical, optical, and historical impacts of the 1908 Tunguska meteor fall", *Icarus*, Vol. 50, 1982, pp. 1-52

9. Chyba, C. F., Thomas, P. J. and Zahnle, K. J., "The 1908 Tunguska explosion - Atmospheric disruption of a stony asteroid", *Nature*, Vol. 361, No. 6407, 1993, pp. 40-44

10. Farinella, P., Foschini, L., Froeschlé, Ch., Gonczi, R., Jopek, T. J. and Longo, P. G., "Probable asteroidal origin of the Tunguska Cosmic Body", *A&A*, Vol. 377, 2001, pp. 1081-1097

11. Hammel, H. B., Beebe, R. F., Ingersoll, A. P., Orton, G. S., Mills, J. R., Simon, A. A., Chodas, P., Clarke, J. T., de Jong, E., Dowling, T. E., Harrington, J., Huber, L. F.,

Karkoschka, E., Santori, C. M., Toigo, A., Yeomans, D. and West, R. A., "HST Imaging of Atmospheric Phenomena Created by the Impact of Comet Shoemaker-Levy 9", *Science*, Vol. 267, No. 5202, 1995, pp. 1288-1296

12. Jessup, K., Clarke, J., Ballester, G. and Hammel, H., "Ballistic Reconstruction of HST Observations of Ejecta Motion Following Shoemaker-Levy 9 Impacts into Jupiter", *Icarus*, Vol. 146, No. 1, 2000, pp. 19-42

13. Bottke, W. F., Vokrouhlicky, D. and Nesvorny, D., "An asteroid breakup 160Myr ago as the probable source of the K/T impactor", *Nature*, Vol. 449, No. 7158, 2007, pp. 48-53

14. Horner, J. and Jones, B. W., "Jupiter friend or foe? I: The asteroids", *International Journal of Astrobiology*, Vol. 7, 2008, pp. 251-261

15. Horner, J. and Jones, B. W., “Jupiter – friend or foe? II: the Centaurs”, *International Journal of Astrobiology*, Vol. 8, No. 2, 2009, pp. 75-80

16. Horner, J. and Jones, B. W., "Jupiter - friend or foe? IV: the influence of orbital eccentricity and inclination", *International Journal of Astrobiology*, Vol. 11, No. 3, 2012, pp. 147-156

17. Horner, J., Jones, B. W. and Chambers, J., "Jupiter - friend or foe? III: the Oort cloud comets", *International Journal of Astrobiology*, Vol. 9, No. 1, 2010, pp. 1-10

18. Morbidelli, A., Bottke, W. F., Froschele, Ch. and Michel, P., "Origin and Evolution of Near-Earth Objects", *Asteroids III*, W. F. Bottke Jr., A. Cellino, P. Paolicchi and R. P. Binzel, Eds, University of Arizona Press, Tucson, 2002, pp. 409-422

19. NEOsMPCweb: Near Earth Object Program, http://neo.jpl.nasa.gov/neo/groups.html (accessed 30th October 2014)

20. MPCPHAsweb: List Of The Potentially Hazardous Asteroids (PHAs), http://www.minorplanetcenter.net/iau/lists/t_dangerous.html (accessed October 28, 2014)

21. MPCUnusualweb: Unusual Minor Planets, http://www.minorplanetcenter.net/iau/lists/Unusual.html (accessed October 28, 2014)

22. Horner, J., Evans, N. W., Bailey, M. E. and Asher, D. J., "The populations of comet-like bodies in the Solar system", *MNRAS*, Vol. 343, No. 4, 2003, pp. 1057-1066

23. MPCCometaryDesWeb: IAU Minor Planet Center Cometary Designation, http://www.minorplanetcenter.net/iau/lists/CometResolution.html (accessed August 20, 2014)

24. Horner, J., Evans, N. W. and Bailey, M. E., "Simulations of the population of Centaurs - II. Individual objects", *MNRAS*, Vol. 355, No. 2, 2004b, pp. 321-329

25. Gladman, B., Marsden, B. G. and Vanlaearhoven, C., "Nomenclature in the Outer Solar System", *The Solar System Beyond Neptune*, M. A. Barucci, H. Boehnhardt, D. P. Cruikshank and A. Morbidelli, Eds, University of Arizona Press, Tucson, 2008, pp. 43-57

26. Horner, J., Evans, N. W. and Bailey, M. E., "Simulations of the population of Centaurs - I. The bulk statistics", *MNRAS*, Vol. 354, No. 3, 2004a, pp. 798-810

27. Wisdom, J., "The origin of the Kirkwood gaps - A mapping for asteroidal motion near the 3/1 commensurability", *AJ*, Vol. 87, 1982, pp. 577-593

28. Wisdom, J., "Chaotic behavior and the origin of the 3/1 Kirkwood gap", *Icarus*, Vol. 56, 1983, pp. 51-74

29. Morbidelli, A. and Moons, M., "Numerical evidence on the chaotic nature of the 3/1 mean motion commensurability", *Icarus*, Vol. 115, No. 1, 1995, pp. 60-65

30. Van Houten, C. J., van Houten-Groeneveld, A. and Gehrels, T., "Minor Planets and Related Objects. V. The Density of Trojans near the Preceding Lagrangian Point", *AJ*, Vol. 75, 1970, pp. 659

31. Levison, H. F., Shoemaker, E. M. and Shoemaker, C. S., "Dynamical evolution of Jupiter's Trojan asteroids", *Nature*, Vol. 385, No. 6611, 1997, pp. 42-44

32. Sheppard, S. S. and Trujillo, C. A., "A Thick Cloud of Neptune Trojans and Their Colors", *Science*, Vol. 313, No. 5786, 2006, pp. 511-514

33. Schubart, J., "Long-Period Effects in the Motion of Hilda-Type Planets", *AJ*, Vol. 73, 1968, pp. 99-103

34. Yu, Q. and Tremaine, S., "The Dynamics of Plutinos", *AJ*, Vol. 118, No. 4, 1999, pp. 1873-1881

35. Lykawka, P. S. and Mukai., T., "Dynamical classification of trans-neptunian objects: Probing their origin, evolution, and interrelation", *Icarus*, Vol. 189, No. 1, 2007, pp. 213-232

36. Tiscareno, M. and Malhotra, R., "The Dynamics of Known Centaurs", *AJ*, Vol. 126, No. 6, 2003, pp. 3122-3131

37. Murray, C.D. and Dermott, S. F., *Solar System Dynamics*, (Cambridge University Press), 1999

38. Horner, J. and Evans, N. W., "The capture of Centaurs as Trojans", *Monthly Notices Of The Royal Astronomical Society: Letters*, Vol. 367, No. 1, 2006, pp. L20-L23

39. Marzari, F. F. and Scholl, H. H., "The Role of Secular Resonances in the History of Trojans", *Icarus*, Vol. 146, No. 1, 2000, pp. 232-239

40. Smirnov, E. A. and Shevchenko, I.I., "Massive Identification of Asteroids in three-body Resonances", *Icarus*, Vol. 222, No. 1, 2013, pp. 220-228

41. Bailey, B. L. and Malhotra, R., "Two dynamical classes of Centaurs", *Icarus*, Vol. 203, No. 1, 2009, pp. 155-163

42. Levison, H. F., "Comet Taxonomy", *Completing the Inventory of the Solar System, Astronomical Society of the Pacific Conference Proceedings*, Vol. 107, T.W. Rettig and J.M. Hahn, Eds., 1996, pp. 173-191

43. Duncan, M., Levison, H. and Dones, L., "Dynamical Evolution of Ecliptic Comets", *Comets II*, M. C. Festou, H. U. Keller, and H. A. Weaver, Eds, University of Arizona Press, Tucson, 2004, pp. 193-204

44. Emel'yanenko, V. V., Asher, D. J. and Bailey, M. E., "A Model for the Common Origin of Jupiter Family and Halley Type Comets", *Earth, Moon & Planets*, Vol. 110, No. 1-2, 2013, pp. 105-130

45. Cordeiro, R.R., "Anomalous Diffusion in the Asteroid Belt", *AJ*, Vol. 132, No. 2114, 2006, pp. 2114-2126

46. Levison, H. F. and Duncan, M. J., "The long-term dynamical behavior of short-period comets", *Icarus*, Vol. 108, No. 1, 1994, pp. 18-36

Towards a dynamics-based estimate of the extent of HR 8799's unresolved warm debris belt

Bruna Contro[1], Rob Wittenmyer[1,2,3], Jonti Horner[2,3], Jonathan P. Marshall[1,3]

[1]*School of Physics, UNSW Australia, Sydney, New South Wales 2052, Australia*
[2]*Computational Engineering and Science Research Centre, University of Southern Queensland, Toowoomba, Queensland 4350, Australia*
[3]*Australian Centre for Astrobiology, UNSW Australia, Sydney, New South Wales 2052, Australia*

Summary: In many ways, the HR 8799 planetary system resembles our Solar system more closely than any other discovered to date – albeit significantly younger, and on a larger and more dramatic scale. The system features four giant planets and two debris belts. The first of these belts lies beyond the orbit of the outermost planet, and mirrors the location of our Solar system's Edgeworth-Kuiper belt. The second, which has yet to be fully observationally characterised, lies interior to the orbit of the innermost known planet, HR8799 e, and is an analogue to our Asteroid Belt. With such a similar architecture, the system is a valuable laboratory for examining exoplanetary dynamics, and the interaction between debris disks and giant planets.

In recent years, significant progress has been made in the characterisation of the outer of HR8799's debris disks, primarily using the Herschel Space Observatory. In contrast, the inner disk, which lies too close to its host star to be spatially resolved by that instrument, remains poorly understood. This, in turn, leaves significant questions over both the location of the planetesimals responsible for producing the observed dust, and the physical properties of those grains.

We have performed the most extensive simulations to date of the inner, unresolved debris belt around HR 8799, using UNSW Australia's supercomputing facility, *Katana*. In this work, we present the results of integrations following the evolution of a belt of dynamically hot debris interior to the innermost planet, HR8799 e, for a period of 60 million years, using an initial population of 500,000 massless test particles. These simulations have enabled the characterisation of the extent and structure of the inner belt, revealing that its outer edge must lie interior to the 3:1 mean-motion resonance with HR8799 e, at approximately 7.5 au, and highlighting the presence of fine structure analogous to the Solar system's Kirkwood gaps. In the future, our results will also allow us to calculate a first estimate of the small-body impact rate and water delivery prospects for any potential terrestrial planet(s) that might lurk, undetected, in the inner system.

Keywords:
Stars: individual: HR 8799, Stars: circumstellar matter, Planetary systems: minor bodies, Methods: *n*-body simulations, Astrobiology, Exoplanets, Habitability

Introduction

In the past few years, the search for planets around other stars has moved from focussing on the biggest and easiest planets to find (e.g. [1][2][3]), to looking for planets more like the Earth. As part of this process, the search for Solar system analogues has come to the fore, for a number of reasons. The first is purely practical – radial velocity observations of sun-like stars can only detect planets with orbital periods shorter than the time a given star has been targeted with observations. In order to discover a Jupiter-like planet, then, a given star must have been observed on a decadal timescale. For this reason, it is only in the last few years that the first true Jupiter-analogues have been found (e.g. [4][5][6]). Furthermore, the discovery of smaller, more Earth-like planets, has required the development of more precise tools, as well as larger observational datasets. Once again, that search has only recently begun to bear fruit (e.g. [7][8][9]).

In this push towards finding planetary systems that resemble our own, different techniques have brought complementary results to the table. The search for the smallest, most Earth-like planets, has been led by the *Kepler* spacecraft, using the transit technique (e.g. [7][10]). At the same time, the search for Jupiter-analogues has been led by direct imaging and radial velocity programs, which continue to find massive planets at ever increasing distances from their host stars.

Beyond the simple practical reasons for the slow onset of the search for Solar system analogues, an additional and key motivation for that search is that finding planetary systems that truly resemble our own is critically important, particularly since it enables us to gain a better understanding of our own Solar system and its place in the universe.

Despite the rapid progress made over the past two decades, technological limitations have, to date, resulted in very few exoplanetary systems being discovered that truly resemble our own. Possibly the prime example of such a system is that orbiting the star HR8799. The discovery of the first three planets in that system was announced in 2008 [11], using direct images of the system taken with the Keck and Gemini observatories. A fourth giant planet was found in 2010 [12] as a result of follow-up observations. With its four giant planets [12] and a circumstellar disk [13] comprised of two debris belts [14], this system stands as the most "Solar system like" of all that have been discovered to date, and as such is an important laboratory that will allow us to greatly improve our understanding of the formation, evolution, and scarcity of planetary systems like our own.

HR 8799 is a young A-type star, located ~ 40 pc from the Earth [15]. Whilst its age remains, to date, somewhat poorly determined, its relative youth is well accepted. Whilst [16] suggest that HR 8799 is no older than 50 million years (Myr), [17] suggest an age of 30 Myr, based on the stars likely membership of the Columba Association. When announcing the discovery of HR8799 e, [12] take a conservative perspective on the star's age. They consider two different ages for the planetary system – 30 Myr, based on the star's membership of the Columba Association, and 60 Myr, following previous work [11]. See Table 1 for a summary of the stellar parameters used in this work.

Table 1. Stellar Parameters for HR 8799.

Parameter	Value	Reference
Age (Myr)	30 60	[17] [11]
Teff (K)	7193 ± 87	[18]
L ($L_\odot$)	5.05 ± 0.29	[18]
M ($M_\odot$)	1.56	[12][19]

In Table 2, we present the best-fit orbital parameters of the planets, as presented in a recent dynamical study [19]. For this work we considered the primary values only, ignoring the errors. A schematic showing the orbits of the four planets in the HR 8799 system is shown in Figure 1. In this resonant architecture (planets trapped in a double Laplace resonance with orbital period ratios of 1:2:4:8, as described in [19]), the orbits of the planets remain dynamically stable on long timescales.

Table 2. The orbital elements of the four planets orbiting the star HR 8799, as used in our integrations. These elements are those of the double-Laplace resonant architecture proposed in Table 1 of [19]. Here, m is the mass of the planet, in Jupiter masses, a is the orbital semi-major axis, in au, and e is the eccentricity of the orbit. i is the inclination of the orbit to the plane of the sky, and Ω, ω and M are the longitude of the ascending node, the longitude of pericentre, and the mean anomaly at the epoch 1998.83, respectively.

	m [m_{jup}]	a [au]	e	i [deg]	Ω [deg]	ω [deg]	M [deg]
HR 8799 e	9±2	15.4±0.2	0.13±0.03	25±3	64±3	176±3	326±5
HR 8799 d	9±2	25.4±0.3	0.12±0.02			91±3	58±3
HR 8799 c	9±2	39.4±0.3	0.05±0.02			151±6	148±6
HR 8799 b	7±2	69.1±0.2	0.020±0.003			95±10	321±10

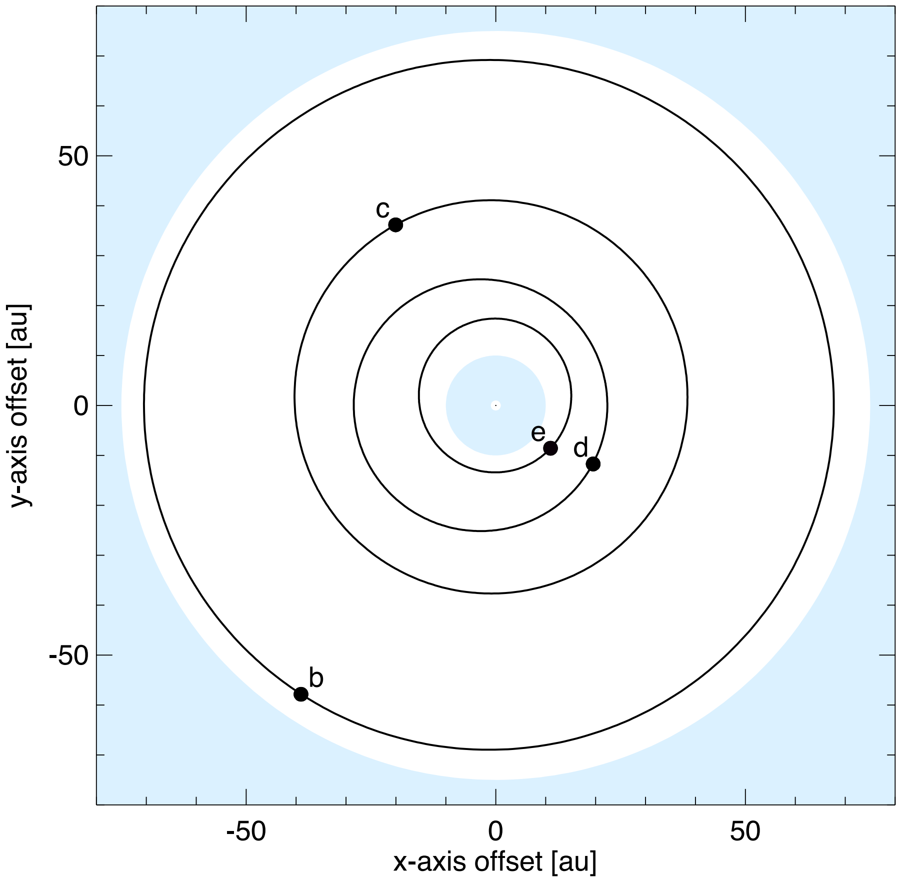


Figure 1. This is a schematic plot of the HR 8799 system. The four planets are at their correct orbital separations from the star, and location around the star, but are presented on circular rather than eccentric orbits. The light blue regions are those parts of the system where circumstellar debris is expected to reside in order to explain the observed infrared excess described in Matthews et al., 2014.

In addition to the four planets detected by [12], the HR 8799 system also hosts two debris belts and a halo of small grains and dust out to > 1000 au [14]. Whilst the outer belt has been spatially resolved by [13], [16] and [14], and shown to extend from 100 (±10 au) to 310 au, the inner, warm belt is, to date, poorly understood. The best estimate, based on a black body fit to the observed dust temperature ([20][21]), is that it must lie between 1 and 10 au. This information, coupled with our knowledge of the Keplerian orbital elements for HR 8799 b, c, d and e, gives us enough information to, for the first time, attempt to constrain its extent and interior structure using *n*-body simulations.

Material and Methods

We performed the most extensive simulations to date of the inner debris belt of HR 8799 using the numerical integrator package MERCURY [22], using UNSW's *Katana* supercomputing facility, and building on our earlier work studying the system [23]. The disk, initially composed of 500,000 massless test particles, was followed for 60 Myr (in order to ensure compatibility with the ages proposed by [11] and [12]) to trace its evolution, with an integration time-step of 7 days (approximately 1/2500th of the orbital period of HR8799 e).

Each test particle was followed (i) until it was ejected from the system (upon reaching a barycentric distance of 1000 au); (ii) impacted on one of the giant planets; or (iii) collided with the central body.

For this work, each of the six orbital elements for the test particles were randomly allocated to lie within a set range. The semi-major axes were chosen to lie between 1 and 10 au (with the inner and outer edges chosen in accordance with [24] and [14] descriptions of the inner disk). Once the semi-major axes were set, the eccentricity of the massless test particles was allocated to lie between 0.1 and 1.0[1]. Each of the test particles was then given an orbital inclination between 0 and 25 degrees (a range similar to that occupied by the great majority of objects in the Solar system's Asteroid and Edgeworth-Kuiper belts). The rotational orbital elements for each particle were each separately randomly dispersed between 0 and 360 degrees. The result of these processes was a dynamically excited disk of debris that filled the inner reaches of the HR 8799 planetary system, with orbital elements within the ranges detailed in Table 3. The orbital elements of the particles, as well as those of the four giant planets integrated, were output at 6 Myr intervals, allowing us to trace the evolution of the warm belt over the lifetime of the system.

Table 3: The range of values covered by the population of test particles considered in this work, at the start of the integrations. Within these values, the orbital elements of each test particle was randomly generated.

Orbital element	**Range of values considered**
Semi-major axis (au)	$1 < a < 10$
Eccentricity	$0.1 < e < 1.0$
Inclination (°)	$0 < i < 25$
Argument of Periastron (°)	$0 < \omega < 360$
Longitude of the Ascending Node (°)	$0 < \Omega < 360$
Mean Anomaly (°)	$0 < M < 360$
Subsidiary considerations	
Periastron distance (au)	$q > 0.1$

For this work, we followed the best-fit parameters for HR 8799 b, c, d and e and stellar mass of $1.56 M_\odot$ described by [19], as shown in Table 2. In addition, the location of the classical "Habitable Zone" (HZ; the region around the star in which an Earth like planet might reasonably be expected to be able to host liquid water on its surface, e.g. [25]) was estimated to be 1.974 (inner edge) to 3.407 au (outer edge), following [26] and [27], using the stellar parameters shown in Table 1.

We note that, in this paper, we consider the inner debris disk as a distribution of massless test particles. These particles individually represent a planetesimal inserted into the HR8799

[1] These simulations are constructed to complement those presented in [34], where we considered a dynamically cold, unexcited disk, and are intended primarily to determine the extremes at which the disk loses stability – hence the very broad range in orbital eccentricity considered. In practice, the highest eccentricity of any test particle in our initial sample was 0.992793 – a result of an additional constraint that no test particle move on an orbit with initial periastron less than 0.1 au.

planetary system on the given orbit. This approach neglects radiative forces and the perturbing influence of a massive asteroid belt on the orbits of the planets.

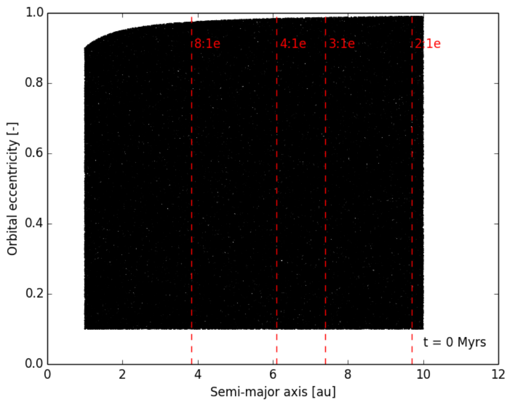
If we had used massive particles, the total mass of the debris disk would have become a limiting factor in the simulations, pulling on the planets and altering their orbits (as seen in simulations of a massive outer debris disk for HR8799 [33]). As such, these simulations are limited in relevance to the case where the inner disk is much less massive than the innermost planet. This is expected to be the case, since the collisional timescale for several tens of Earth masses of material at ~ 5-10 au would be << 60 Myr, causing a rapid depletion and removal of the inner disk.

By neglecting radiative forces we can only comment on the regions within which planetesimals would reside. Dust grains, which migrate under the influence of radiative forces, spread beyond their formation region both inward and outward through the system, sometimes becoming trapped in resonances with planets. As such the distribution of dust grains would not directly mirror that of their parent planetesimals.

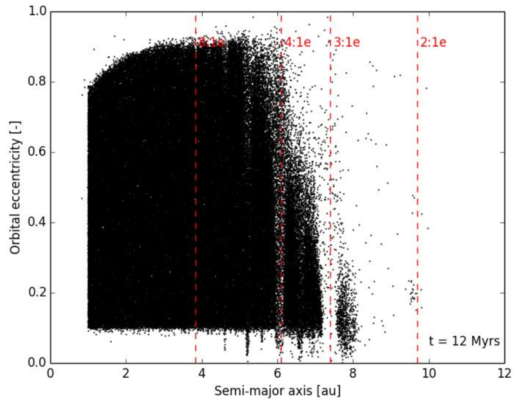
The results of these simulations, representing a simplification of the true system's dynamical interactions, still provide a first glimpse of the extent, structure and complexity of the inner system.

Results

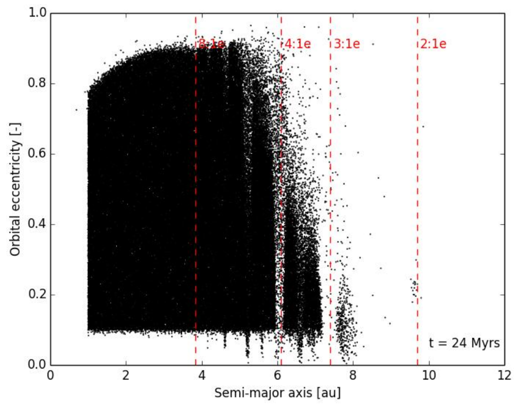
The instantaneous semi-major axes and eccentricities of the test particles in the inner debris disk of HR 8799 are shown as a function of time in Figure 2. Panel one shows the initial orbits of the 500,000 massless test particles, whilst the other panels show the elements of those particles surviving at each of the time-steps in question, in which each black dot represents a surviving test particle. Figure 3 shows in more detail the final distribution of the surviving test particles after the full 60 Myr have elapsed as well as the location of the HZ (green bar) and a few select mean-motion resonances with HR 8799 e (red lines), whilst Figure 4 shows the number of test particles that remain in the simulation as a function of time in our simulations.



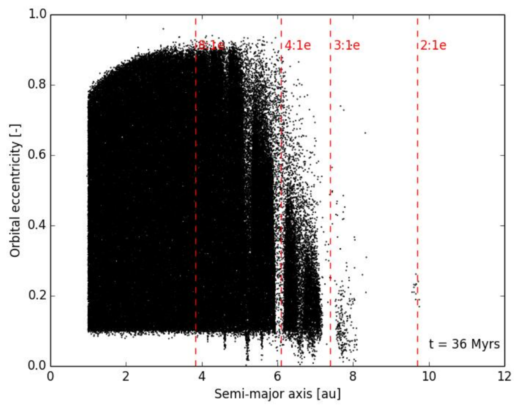
Panel 1: 0 Myr.



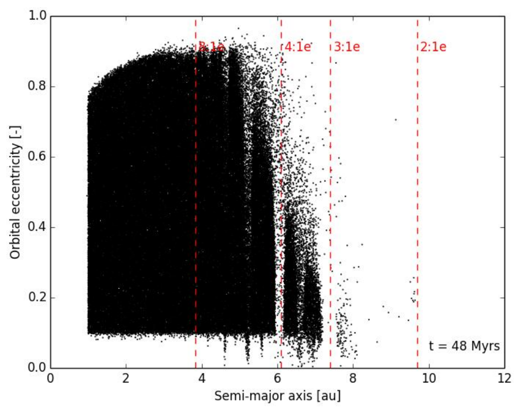
Panel 2: 12 Myr.



Panel 3: 24 Myr.



Panel 4: 36 Myr.



Panel 5: 48 Myr.

Fig 2. The distribution of surviving test particles orbiting HR 8799 in our simulations, as a function of time, from t = 0 to t = 48 Myr, in semi-major axis-eccentricity space. Note the rapid initial dispersal of debris beyond ~8 au, followed by a more sedentary sculpting of the disk over the remaining duration of the integration.

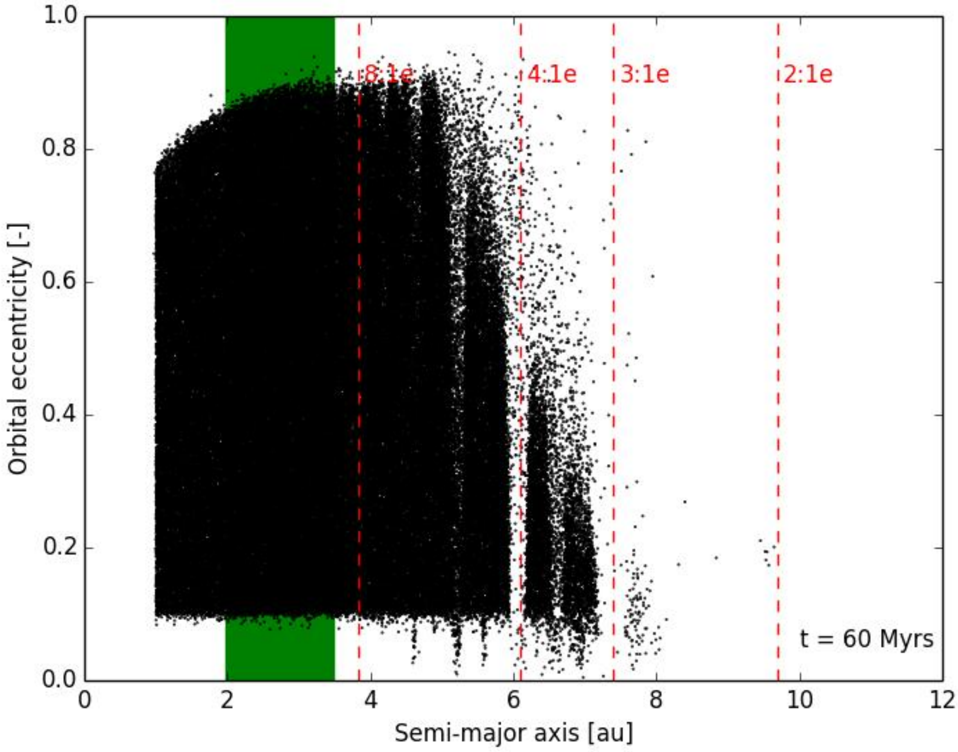


Fig 3. The final distribution of test particles orbiting HR 8799 in semi-major axis – eccentricity space after 60 Myr. The sculpting influence of mean-motion resonances between test particles and planets can be clearly seen both in the gaps introduced to the disk (e.g. just outside 6 au), and in those locations where test particles have been driven to stable orbits with lower eccentricities (the downward "spikes" in eccentricity seen between 4-8 AU. The region shown in green (1.974 to 3.407 au) is our best estimate of the location of the Habitable Zone in the system.

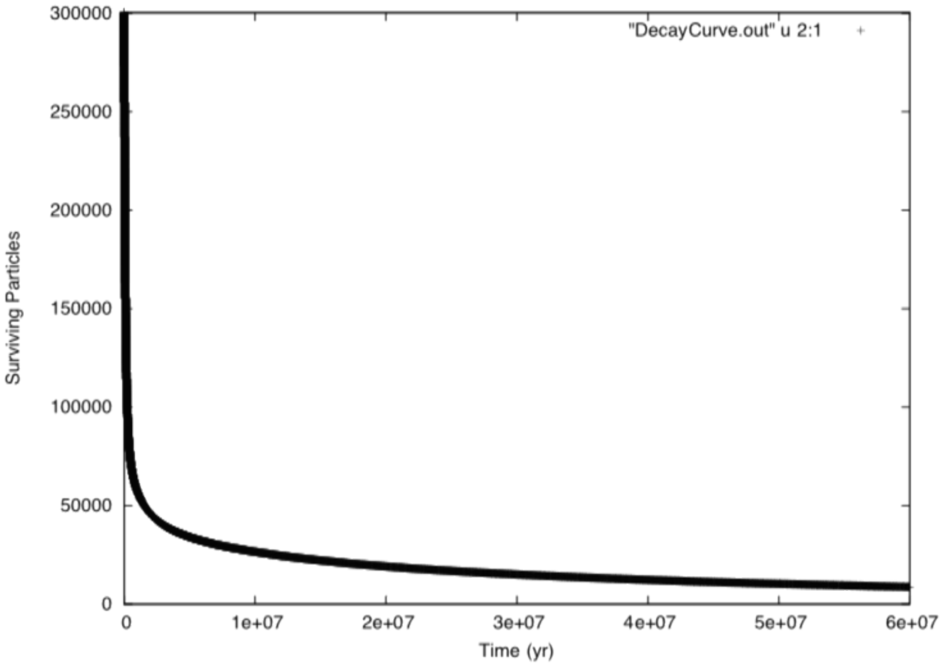


Figure 4. Plot showing the number of surviving test particles as a function of time. Unsurprisingly, given the extreme range of orbital eccentricities and semi-major axes tested in this work, almost 90% of the test particles are removed in the first 10 Myr of the disk's evolution.

Discussion

Over the course of 60 Myr of dynamical evolution, the warm belt orbiting HR 8799 undergoes drastic sculpting, particularly during the first 6 Myr of integration (as seen in Figure 4). This fast clearance is the result of strong dynamical stirring caused by HR 8799 e, the innermost planet, which quickly clears almost all objects located exterior to the 3:1 mean-motion resonance with that planet. Apart from few objects that remain at ~ 9.7 au, probably trapped in the 2:1 MMR and also at ~ 8 au, by the end of the simulations, the outer edge of the disk has become clearly defined at ~ 7.5 au, revealing that test particles cannot survive beyond this distance unless trapped in mean motion resonance with one of the giant planets.

Furthermore, at the final time-step, the internal structure of the disk can be also observed. Several resonant features are evident, observed as cleared gaps within the belt, such as that centred on the location of the 4:1 mean motion resonance with HR 8799e. Other high order resonances with HR 8799e can be seen between 8:1e and 3:1e, in which the massless test particles were either cleaned or damped to low eccentricity orbits. These structural features are strongly reminiscent of the Kirkwood gaps, which are observed within the Asteroid belt in our own Solar system (e.g [28][29]). Interestingly, our results agree nicely with [19], who examined whether a hypothetical fifth planet could exist in the HR 8799 system, interior to the orbit of HR 8799e, in an orbit resonant with that planet (and thereby with the others in the system). They found that a hypothetical HR 8799f would be dynamically stable if it were to orbit at either a ~ 9.7 au or a ~ 7.5 au. In addition, in their figure 22, they map the stable regions for objects interior to HR 8799e. Their results, which were obtained through million-year MEGNO ([35]) integrations of the system, reassuringly reveal the same stable and unstable regions as those resulting from our longer n-body dynamical simulations.

In addition, a clear sculpting can be seen at high eccentricities at the inner edge of the disk. These highly eccentric objects move on orbits with small periastra and orbital periods such that the 7-day time-step used for these integrations is insufficient. This feature is therefore a computational effect, rather than the result of a real dynamical process, although it is exacerbated by our initial setup decision that no test particle should move on an orbit with periastron distance less than 0.1 au.

We also note that very little sculpting can be seen within the classical HZ. This suggests that that region might well be able to host dynamically stable Earth-like planets. Unfortunately, the HR8799 system is certainly too young to be a sensible target for the future search for life beyond the Solar system (e.g. [30]). However, the possibility that such exo-Earths could exist in a system that so closely resembles our own remains intriguing. In future work, we will build on earlier studies of the influence of giant planets on the impact flux in our own Solar system (e.g. [31]; [32]) to examine the potential impact regimes that such exoEarths in the system might experience. This will allow us to build tools that will eventually be used to help determine which Earth-like planets, discovered in the years to come, are the most promising targets for the search for life beyond the Solar system [31], allowing astronomers to infer both the impact regime that would be experienced by those planets, and the potential level of hydration that they might have received from exogenous sources.

Conclusions

We have performed a detailed dynamical analysis of the inner planetesimal belt in the HR8799 planetary system. By simulating the orbital evolution of 500,000 massless test particles, initially distributed on dynamically hot orbits interior to the system's innermost planet, we have provided strong constraints on the structure and extent of the planetesimal belt.

Our results reveal that the outer edge of the inner belt can lie no further than ~7.5 au from its host star. The simulations also show that the belt must have a significant amount of internal structure, containing several gaps, cleared by the influence of mean-motion resonances with the giant planets in the system. These gaps, a direct analogue of the Solar system's Kirkwood gaps, are most clearly seen around the locations of the 4:1 and 3:1 mean-motion resonances with HR8799 e. Exterior to the location of that 3:1 mean-motion resonance, a small population of test particles can be seen trapped between ~ 7.5 to 8 au, at low eccentricities.

Complementary studies are required in order to investigate the evolution of a dynamically cold debris disk (the low eccentricity, low inclination counterpart to the disk studied in this work). Those simulations have recently been completed, and the results will be presented in a future paper. We will examine the impact rates that might be expected on any planets located in the HZ, as well as studying the possible routes by which water, and other volatiles, might be delivered to those planets. Further work will yield a better understanding of the dynamical structure of the inner debris belt, and offer clues as to the possibility and viability of a fifth, terrestrial planet in the HZ.

Acknowledgement

This work is supported by CAPES Foundation, Ministry of Education of Brazil, Brasilia – DF, Zip Code 70.040-020 and School of Physics - UNSW Australia, and made use of UNSW Australia's *Katana* supercomputing cluster.

References

1. Mayor, M. and Queloz, D., "A Jupiter-mass companion to a solar-type star", *Nature*, Vol.378, 1995, 335
2. Marcy, G. W. and Butler, R. P., "A Planetary Companion to 70 Virginis", *The Astrophysical Journal*, Vol. 464, 1996, L147
3. Butler, R. P. and Marcy, G. W., "A Planet Orbiting 47 Ursae Majoris", *The Astrophysical Journal*, Vol. 464, 1996, L153
4. Howard, A. W., Johnson, J. A., Marcy, G. W., Fischer, D. A., Wright, J. T., Bernat, D., Henry, G. W., Peek, K. M. G., Isaacson, H., Apps, K., Endl, M., Cochran, W. D., Valenti, J. A., Anderson, J. and Piskunov, N. E., "The California Planet Survey. I. Four New Giant Exoplanets", *The Astrophysical Journal*, Vol. 721, 2010, 1467
5. Wittenmyer, R. A., Horner, J., Tuomi, M., Salter, G. S., Tinney, C. G., Butler, R. P., Jones, H. R. A., O'Toole, S. J., Bailey, J., Carter, B. D., Jenkins, J. S., Zhang, Z., Vogt, S. S. and Rivera, E. J., "The Anglo-Australian Planet Search. XXII. Two New Multi-planet Systems", *The Astrophysical Journal*, Vol. 753, 2012, article id. 169
6. Robertson, P., Horner, J., Wittenmyer, R. A., Endl, M., Cochran, W. D., MacQueen, P. J., Brugamyer, E. J., Simon, A. E., Barnes, S. I. and Caldwell, C., "A Second Giant Planet in 3:2 Mean-motion Resonance in the HD 204313 System", *The Astrophysical Journal*, Vol. 754, 2012, article id. 50
7. Howard, A. W., Marcy, G. W., Bryson, S. T. et al., "Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler", *The Astrophysical Journal Supplement*, Vol. 201, 2012, article id. 15
8. Delfosse, X., Bonfils, X., Forveille, T., Udry, S., Mayor, M., Bouchy, F., Gillon, M., Lovis, C., Neves, V., Pepe, F., Perrier, C., Queloz, D., Santos, N. and Ségranson, D., "The HARPS search for southern extra-solar planets. XXXIII. Super-Earths around the M-dwarf neighbors Gl433 and Gl667C", *Astronomy & Astrophysics*, Vol. 553, 2013, article id. A8
9. Wittenmyer, R. A., Tuomi, M., Butler, R. P., Jones, H. R. A., Anglada-Escudé, G., Horner, J., Tinney, C. G., Marshall, J. P., Carter, B. D., Bailey, J., Salter, G. S., O'Toole, S. J., Wright, D., Crane, J. D., Schectman, S. A., Arriagada, P., Thompson, I., Minniti, D., Jenkins, J. S. and Diaz, M., "GJ 832c: A Super-Earth in the Habitable Zone", *The Astrophysical Journal*, Vol. 791, 2014, article id. 114
10. Borucki, W. J., Koch, D. G., Basri, G. et al., "Characteristics of Kepler Planetary Candidates Based on the First Data Set", *The Astrophysical Journal*, Vol. 728, 2011, article id. 117
11. Marois, C., Macintosh, B., Barman, T., et al., "Direct Imaging of Multiple Planets Orbiting the Star HR 8799", *Science,* Vol. 322, 2008, 1348

12. Marois, C., Zuckerman, B., Konopacky, Q. M., et al., "Images of a fourth planet orbiting HR 8799", *Nature*, Vol. 468, 2010, 1080
13. Rhee, J. H., Inseok, S., Zuckerman, B., et al., "Characterization of Dusty Debris Disks: The IRAS and Hipparcos Catalogs", *The Astrophysical Journal*, Vol. 660, 2007, 1556
14. Matthews, B., Kennedy, G., Sibthorpe, B., et al., "Resolved imaging of the HR 8799 debris disk with *Herschel", The Astrophysical Journal*, 2013, Vol. 780, 2014, article id. 97
15. van Leeuwen, F., "Hipparchos, The New Reduction of the Raw Data", *Astrophysics and Space Science Library*, Vol. 350, 2007, *by F. van Leeuwen*. Berlin: Springer.
16. Reidemeister, R., Krivov, A. V., Schimidt, T. O. B., et al., "A possible architecture of the planetary system HR 8799", *Astronomy & Astrophysics*, Vol. 503, 2009, 247
17. Zuckerman, B., Rhee, J. H., Song, I., Bessell, M. S., "The Tucana/Horologium, Columba, AB Doradus, and Argus Associations: New Members and Dusty Debris Disks", *The Astrophysical Journal*, Vol. 732, 2011, article id. 61
18. Baines, E. K., White, R. J., Huber, D., et al., "The CHARA array angular diameter of HR 8799 favors planetary masses for its imaged companions", *The Astrophysical Journal*, Vol. 761, 2012, article id. 57
19. Gozdziewski, K., Migaszewski, C., "Multiple mean motion resonances in the HR 8799 planetary system", *MNRAS*, Vol. 440, 2014, 3140
20. Booth, M., Kennedy, G., Sibthorpe, B., et al., "Resolved debris discs around A stars in the Herschel DEBRIS survey", *MNRAS*, Vol. 428, 2013, 1263
21. Pawellek, N., Krivov, A. V., Marshall, J. P., et al., "Disk Radii and Grain Sizes in Herschel-resolved Debris Disks", *The Astrophysical Journal*, Vol. 792, 2014, article id. 65
22. Chambers, J. E., "A hybrid symplectic integrator that permits close encounters between massive bodies", *MNRAS,* Vol. 304, 1999, 793
23. Marshall, J., Horner, J. and Carter, A., "Dynamical Simulations of the HR8799 planetary system", *International Journal of Astrobiology*, Vol. 9, 2010, 259
24. Su, K. Y. L., Rieke, G. H., Stapelfeldt, K. R., et al., "The Debris Disk Around HR 8799", *The Astrophysical Journal*, Vol. 705, 2009, 314
25. Kasting, J. F., Whitmire, D. P., Reynolds, R. T., "Habitable Zones around Main Sequence Stars", *Icarus*, Vol. 101, 1993, 108
26. Selsis, F., Kasting, J. F., Levrard, B., et al., "Habitable planets around the star Gliese 581?" *Astronomy & Astrophysics*, Vol. 476, 2007, 1373
27. Kopparapu, R. K., Ramirez, R. M., Schottel-Kote, J., et al., "Habitable Zones around Main-sequence Stars: Dependence on Planetary Mass", *The Astrophysical Journal*, Vol. 787, 2014, article id. 29
28. Kirkwood, D., "On the distribution of the minor planets", *The Observatory*, Vol. 5, 1882, 194
29. Wisdom, J., "The origin of the Kirkwood gaps – A mapping for asteroidal motion near the 3/1 commensurability", *Astronomical Journal*, Vol. 87, 1982, 577
30. Horner, J., Jones, B. W., "Determining habitability: which exoEarths should we search for life?", *International Journal of Astrobiology*, Vol. 9, 2010, 273
31. Horner, J., Jones, B. W., "Jupiter friend or foe? I: The asteroids", *International Journal of Astrobiology,* Vol. 7, 2008, 251
32. Horner, J., Mousis, O., Petit, J.-M., Jones, B. W., "Differences between the impact regimes of the terrestrial planets: Implications for primordial D:H ratios", *Planetary and Space Science*, Vol.57, 2009, 1338

33. Moore, A., Quillen, A. C., “Effects of a planetesimal debris disc on stability scenarios for the extrasolar planetary system HR8799”, *Monthly Notices of the Royal Astronomical Society*, Vol. 430, 2013, 320
34. Contro, B., Wittenmyer, R., Horner, J., Marshall, J. P., “The dynamical structure of HR8799’s inner debris disk”, *Proceedings of the Origins 2014 conference held at Nara, Japan, July 6 – July 11 2014, in press, arXiv:1502.02315*
35. Cincotta, P. M., Giordano, C. M., Simó, C., “Phase space structure of multi-dimensional systems by means of the mean exponential growth factor of nearby orbits”, *Physica D: Nonlinear Phenomena*, Vol. 182, Issue 3-4, 2003, 151

Winds of Planet Hosting Stars

B. A. Nicholson*, M. W. Mengel*, L. Brookshaw*, A. A. Vidotto†, B. D. Carter*, S. C. Marsden*, J. Soutter*, I. A. Waite* and J. Horner*‡

* *Computational Engineering and Science Research Centre, University of Southern Queensland, Toowoomba, QLD 4350, Australia*
† *Geneva Observatory, University of Geneva, Sauverny, Versoix CH-1290, Switzerland*
‡ *Australian Centre for Astrobiology, UNSW Australia, Sydney, NSW 2052, Australia*

Summary: The field of exoplanetary science is one of the most rapidly growing areas of astrophysical research. As more planets are discovered around other stars, new techniques have been developed that have allowed astronomers to begin to characterise them. Two of the most important factors in understanding the evolution of these planets, and potentially determining whether they are habitable, are the behaviour of the winds of the host star and the way in which they interact with the planet. The purpose of this project is to reconstruct the magnetic fields of planet hosting stars from spectropolarimetric observations, and to use these magnetic field maps to inform simulations of the stellar winds in those systems using the Block Adaptive Tree Solar-wind Roe Upwind Scheme (BATS-R-US) code. The BATS-R-US code was originally written to investigate the behaviour of the Solar wind, and so has been altered to be used in the context of other stellar systems. These simulations will give information about the velocity, pressure and density of the wind outward from the host star. They will also allow us to determine what influence the winds will have on the space weather environment of the planet. This paper presents the preliminary results of these simulations for the star τ Boötis, using a newly reconstructed magnetic field map based on previously published observations. These simulations show interesting structures in the wind velocity around the star, consistent with the complex topology of its magnetic field.

Keywords: Stellar winds, Stellar magnetic fields, Stellar evolution, Planetary evolution, Habitability, Astrobiology

Introduction

Exoplanetary science is one of the most dynamic and rapidly expanding fields in modern astrophysical research. Since the first discovery of a planet orbiting a sun-like star, less than 2 decades ago [1], astronomers have found over fifteen-hundred planets orbiting other stars[1]. The early discoveries were of planets of comparable mass to, or more massive than, Jupiter, and orbiting at distances much closer to their star than the Earth is from the Sun [3, 4, 5]. Advances in technology, and the increasing length in the temporal baseline of observations, have resulted in the discovery of systems that have large planets orbiting at similar distances to the gas giants of our Solar system [6, 7, 8]. Also, as a result of the exquisite precision with which the transits of planets across their host stars can now be measured, the size of the smallest discoverable planets has been decreasing, such that the first "super-Earths" have now been found [9, 10, 11].

[1]As of 6th November, 2014, the current tally stands at 1516 confirmed planets, with a further 3359 candidate planets awaiting confirmation, according to the Exoplanet Orbit Database at http://exoplanets.org [2].

As a result of the dramatic progress made in the discovery of exoplanets, the focus of exoplanetary science has shifted to include trying to characterise the planets that have been found. It will not be long before the first Earth-analogue planets will be discovered, and the hunt for life on those planets will begin. The characterisation of these Earth-like planets will be pivotal in deciding which ones to examine more closely. Determining the habitability of a planet based on a large range of variables is vital to choosing the most promising target for the search for life outside our Solar system [12].

A vital component in determining the habitability of a planet is the influence of the planet's host star. In particular, the activity and winds resulting from the magnetic field of an exoplanet's host-star will play a critical role in ensuring that the planet will be able to sustain a sufficient atmosphere to ensure that life can develop and thrive. In our own Solar system, Mars stands as an example of the destructive power winds can have on a planet's atmosphere. Over the four and a half billion year history of our Solar system, Mars's once thick atmosphere has been stripped by the wind of our Sun [13]. Earth has been able to maintain the atmosphere required to support life due to its protective magnetic field.

It is clear, then, that an understanding of the interaction between the stellar winds of planet host stars and the planets that orbit them will play an important role in the search for life outside our Solar system. However, this area of research is still in its infancy, due to the difficulties involved in either detecting and measuring stellar winds directly or in modelling them based on the information we can obtain about their magnetic field behaviour.

The first detections of Solar-like winds around other stars were made by examining the blue-shifted absorption of Lyman-α photons by heated atomic hydrogen gas in the Interstellar Medium (ISM) [14]. This is an indirect method of detecting winds, and attempts have been made to more directly detect winds through looking for x-ray emission from the interaction of the winds and the ISM [15], or radio emission direct from these winds [16]. Unfortunately, these methods have proven to be two to three orders of magnitude less sensitive than the Lyα absorption technique [14]. As a result of the lack of direct observation, the behaviour of stellar winds remains poorly understood. We are thus reliant on the modelling of these stars, constrained by the observations we can make, to investigate the behaviour of stellar winds.

In order to model the winds of other stars, we require information on the behaviour of their magnetic fields, and this can be obtained using Zeeman Doppler Imaging (ZDI) [17]. Making spectropolarimetric[2] observations at different phases of a star's rotation, ZDI can reconstruct the topology of the radial, azimuthal and meridional parts of its magnetic field (e.g. [19, 20]).

The purpose of this project is to take the reconstructed radial magnetic fields of planet hosting stars from spectropolarimetric observations, and use these magnetic field maps to inform a simulation of their winds using the BATS-R-US code [21]. In this paper, we will outline our chosen data for this project and target star for initial testing, and briefly discuss the construction of the magnetic field maps that will be used as input for the stellar wind modelling. We will then present some of the preliminary results of the project, and outline the work to be completed in the future.

[2]Spectropolarimetry is an observational technique that involves passing the light from a star through a polariser before sending it into a high-resolution spectrograph. Since magnetic fields will polarise light from atomic spectra, observing polarised spectra then detects the magnetic field of a star. A detailed explanation of this can be found in [18].

Data Sample

This project will make use of spectropolarimetric data available through our membership of the BCool[3] project [22], with a focus on investigating solar type, planet hosting stars. The first star to be examined for preliminary testing is τ Boötis (τ Boo), a solar type star with a ~ 5 Jupiter-mass planet that orbits at a distance of 0.05 AU and with a period of 3.31 days [4]. Table 1 gives a summary of the basic properties of this star taken from the Exoplanet Orbit Database[4] [2].

τ Boo is an active star, whose magnetic field and wind behaviour has been previously studied (e.g. [23, 24, 25, 26]), and so is an ideal test target. It displays a cycle in its magnetic field similar to that of our Sun's ~ 22 year magnetic cycle, but over a much shorter timescale of around 2 years [25]. Given the number of magnetic observations taken over many epochs, this makes it an ideal candidate for investigating the temporal evolution of the wind during its magnetic cycle, and the impact this may have on its planet.

Table 1: Summary of the basic properties of the star τ Boo.

Mass (Solar masses)	1.34 [27]
Radius (Solar radii)	1.46 [28]
Stellar $v \sin i$ (km s^{-1})	14.98 [29]
Stellar Inclination Angle (deg.)	40 [23]
Apparent Magnitude	4.5 [30]
Effective Temperature (K)	6387 [29]
Spectral Type	F6IV [31]

Stellar Magnetic Field Maps

The spectropolarimetric observations obtained of our target stars are used to map stellar magnetic fields using ZDI. The process for reducing the data is based on the latest version of the "Echelle Spectra Reduction: an Interactive Tool" software [32], called LIBRE-ESPRIT, developed by Jean-Francois Donati of the Observatoire Midi-Pyrenees. This software incorporates the optimal extraction of spectra from the CCD images, and uses least squares deconvolution (LSD) to co-add spectral lines, significantly increasing the sensitivity of magnetic detections in the spectra. This is needed, as the magnetic field signal is only $\sim 0.1\%$ of the total light from a star [32]. Since the magnetic field will effect the all spectral lines in a similar way, the thousands of spectral lines in each Echelle spectrum can be co-added in this way to improve the signal-to-noise of the magnetic field detection.

The method of ZDI then uses these LSD profiles from observations made at different phases in the star's rotation to reconstruct the map of the magnetic field for each of the radial, azimuthal and meridional field components[5]. Examples of such maps for young, solar

[3]The BCool project (http://bcool.ast.obs-mip.fr/) is an international collaboration that is carrying out ongoing observations to categorise and understand the magnetic fields of cool stars.

[4]http://exoplanets.org/. Data were obtained on the 17th October, 2014.

[5]More details on this can be found in [19].

type stars can be found in [33, 34, 35]. The stellar winds have been shown to be mainly governed by the radial component of the magnetic field [36], and so the radial field map alone is used as input to the simulations in our work.

Figure 1 shows the radial magnetic field configuration at the surface of τ Boo from observations made by Rim Fares and collaborators [25] in July 2008 using the Naval echelle spectropolarimeter on the Telescope Bernard Lyot (TBL) in France. This magnetic field map differs slightly from the those published of the same data set previously due to advances in the ZDI method. The map shows multiple regions of positive and negative magnetic field, with several large positive field spots and a large negative spot around the Northern pole. The field is less detailed at the Southern pole due to a lack of information resulting from the inclination angle of the star with respect to us. The black line indicates where the field is zero.

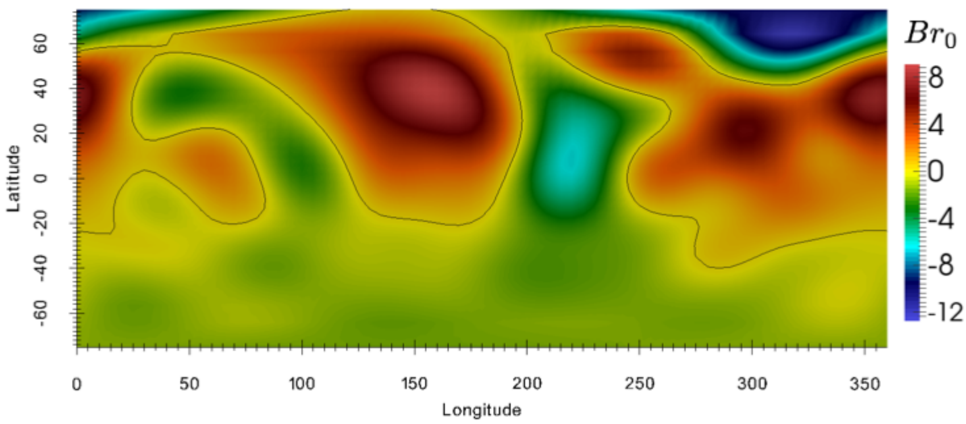


Fig. 1: A plot of the surface magnetic field of τ Boo from observation taken in July 2008 by Fares *et al.* [25]. The black line indicates where the magnetic field is zero. The most interesting feature of this plot are the multiple regions of positive and negative magnetic field, with a large negative spot around the northern pole, and the multiple, large positive regions in the mid-Northern latitudes. No information is available for the far southern region of the star as that is pointed away from the Earth.

Stellar Wind Modelling With BATS-R-US

The wind modelling for our target stars is done using the BATS-R-US code, originally developed to simulate the Solar wind [21]. BATS-R-US solves the three-dimensional, ideal magnetohydrodynamic equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{1}$$

$$\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot \left[\rho \mathbf{u} \otimes \mathbf{u} + \left(P + \frac{B^2}{8\pi} \right) I - \frac{\mathbf{B} \otimes \mathbf{B}}{4\pi} \right] = \rho g, \tag{2}$$

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot (\mathbf{u} \otimes \mathbf{B} - \mathbf{B}) = 0, \tag{3}$$

$$\frac{\partial \epsilon}{\partial t} + \nabla \cdot \left[\mathbf{u} \left(\epsilon + P + \frac{B^2}{8\pi} \right) - \frac{(\mathbf{u} \cdot \mathbf{B})\mathbf{B}}{4\pi} \right] = \rho \mathbf{g} \cdot \mathbf{u}, \tag{4}$$

which describe the conservation of mass, momentum, magnetic flux and energy, respectively. In these equations, ρ is the mass density, $\mathbf{u}$ is the plasma velocity, $\mathbf{B}$ is the magnetic field, P is the gas pressure, $\mathbf{g}$ is the gravitational acceleration of a star with mass M_* and radius R_*, I is the identity matrix, and ϵ is the total energy density, given by

$$\epsilon = \frac{\rho u^2}{2} + \frac{P}{\gamma - 1} + \frac{B^2}{8\pi}. \tag{5}$$

Here, γ is the polytropic index such that $P \propto \rho^\gamma$.

BATS-R-US has been chosen for this work as it has been used extensively [37, 38, 39], and reliably models the wind behaviour of the Sun. Additionally, the block refinement within the code allows the resolution to be increased in specific areas of interest, minimising the load on computational resources by not having to increase resolution everywhere. It is a modular code, designed to be used in different ways and with different physical parameters than the original solar parameters. In essence the code modifications involve the replacement of physical constants such as solar mass and radius sun by stellar parameters that can be varied as needed for each new stellar target. For more details see [40, 41].

Preliminary Results and Discussion

Some preliminary results from our implementation of the BATS-R-US code for the star τ Boo are detailed in Figure 2. The plot shows a slice in the equatorial plane of the star out to 20 stellar radii, with the colour map indicating the magnitude of the wind velocity in km s^{-1}. For context, contours of radial magnetic field from -0.2 to +0.2 Gauss are overlaid in black (negative field) and white (positive field). The complex magnetic field structure seen in the lobes of positive and negative field contours produces abrupt differences in wind velocities around the star. Further work is needed to confirm that the apparent sharp changes in wind velocity are completely physical, and not an artefact of the limited phase coverage of the magnetic mapping, or the limited resolution of the wind modelling used to produce these preliminary results. However, these 'streams' of higher velocity wind at the boundary between the areas of positive and negative magnetic field are particularly interesting, and could suggest the funnelling of plasma through these boundary regions.

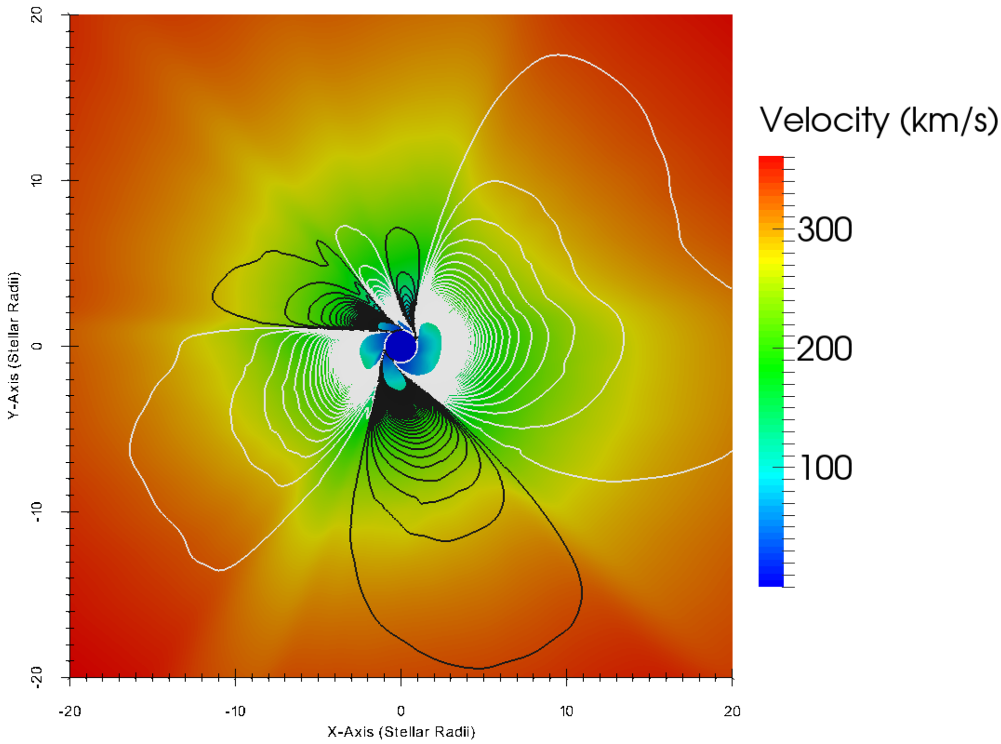


Fig. 2: A plot of the wind velocity of the star τ Boo from observations taken in July 2008 by Fares *et al.* [25], in the equatorial plane of the star out to 20 stellar radii. Overlaid are contours of radial magnetic field ranging from -0.2 to +0.2 Gauss, with white indicating positive field and black indicating negative field. Showing these contours highlights the interesting 'streams' of higher wind velocity at the boundaries between positive and negative magnetic field.

A future goal is to understand the temporal behaviour of the star τ Boo, as well as understanding the impact its planets will have on the behaviour of the wind. We will explore further targets, including others that have been mapped with ZDI, which will be used to investigate the time evolution of winds, and some that have not previously been explored with ZDI.

As well as producing wind models for the target stars, we will modify the BATS-R-US code to simulate the interaction of the stellar wind with a planet for each target. Beyond allowing us to better characterise these particular planetary systems, these simulations will also help us to better understand the formation and evolution of planetary atmospheres beyond our Solar system.

It is through that work that we aim to understand the impact of the winds on the planets that orbit these stars, and similarly how the presence of a planet, which may or may not have a magnetic field of its own, impacts on the behaviour of the wind.

Acknowledgments

This research is supported by the University if Southern Queensland's Strategic Research Fund StarWinds project. This research has made use of NASA's Astrophysics Data System. This research has made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at http://exoplanets.org.

References

[1] M. Mayor and D. Queloz, "A Jupiter-mass companion to a solar-type star," *Nature*, vol. 378, pp. 355–359, Nov. 1995.

[2] J. T. Wright, O. Fakhouri, G. W. Marcy, E. Han, Y. Feng, J. A. Johnson, A. W. Howard, D. A. Fischer, J. A. Valenti, J. Anderson, and N. Piskunov, "The Exoplanet Orbit Database," *Publications of the Astronomical Society of the Pacific*, vol. 123, no. 902, pp. 412–422, Apr. 2011.

[3] G. W. Marcy and R. P. Butler, "A Planetary Companion to 70 Virginis," *The Astrophysical Journal Letters*, vol. 464, p. L147, Jun. 1996.

[4] R. P. Butler, G. W. Marcy, E. Williams, H. Hauser, and P. Shirts, "Three New "51 Pegasi–Type" Planets," *The Astrophysical Journal*, vol. 474, no. 2, pp. L115–L118, Jan. 1997.

[5] R. P. Butler, G. W. Marcy, S. S. Vogt, and K. Apps, "A Planet with a 3.1 Day Period around a Solar Twin," *The Publications of the Astronomical Society of the Pacific*, vol. 110, pp. 1389–1393, Dec. 1998.

[6] A. W. Howard, J. A. Johnson, G. W. Marcy, D. A. Fischer, J. T. Wright, D. Bernat, G. W. Henry, K. M. G. Peek, H. Isaacson, K. Apps, M. Endl, W. D. Cochran, J. A. Valenti, J. Anderson, and N. E. Piskunov, "The California Planet Survey. I. Four New Giant Exoplanets," *The Astrophysical Journal*, vol. 721, pp. 1467–1481, Oct. 2010.

[7] R. A. Wittenmyer, J. Horner, M. Tuomi, G. S. Salter, C. G. Tinney, R. P. Butler, H. R. A. Jones, S. J. O'Toole, J. Bailey, B. D. Carter, J. S. Jenkins, Z. Zhang, S. S. Vogt, and E. J. Rivera, "The Anglo-Australian Planet Search. XXII. Two New Multi-planet Systems," *The Astrophysical Journal*, vol. 753, p. 169, Jul. 2012.

[8] P. Robertson, J. Horner, R. A. Wittenmyer, M. Endl, W. D. Cochran, P. J. MacQueen, E. J. Brugamyer, A. E. Simon, S. I. Barnes, and C. Caldwell, "A Second Giant Planet in 3:2 Mean-motion Resonance in the HD 204313 System," *The Astrophysical Journal*, vol. 754, p. 50, Jul. 2012.

[9] A. W. Howard, G. W. Marcy, S. T. Bryson, J. M. Jenkins, J. F. Rowe, N. M. Batalha, W. J. Borucki, D. G. Koch, E. W. Dunham, T. N. Gautier, III, J. Van Cleve, W. D. Cochran, D. W. Latham, J. J. Lissauer, G. Torres, T. M. Brown, R. L. Gilliland, L. A. Buchhave, D. A. Caldwell, J. Christensen-Dalsgaard, D. Ciardi, F. Fressin, M. R. Haas, S. B. Howell, H. Kjeldsen, S. Seager, L. Rogers, D. D. Sasselov, J. H. Steffen, G. S. Basri, D. Charbonneau, J. Christiansen, B. Clarke, A. Dupree, D. C. Fabrycky, D. A. Fischer, E. B. Ford, J. J. Fortney, J. Tarter, F. R. Girouard, M. J. Holman, J. A. Johnson, T. C. Klaus, P. Machalek, A. V. Moorhead, R. C. Morehead, D. Ragozzine, P. Tenenbaum, J. D. Twicken, S. N. Quinn, H. Isaacson, A. Shporer, P. W. Lucas, L. M. Walkowicz, W. F. Welsh, A. Boss, E. Devore, A. Gould, J. C. Smith, R. L. Morris, A. Prsa, T. D. Morton, M. Still, S. E. Thompson, F. Mullally, M. Endl, and P. J. MacQueen, "Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler," *The Astrophysical Journal Supplement Series*, vol. 201, p. 15, Aug. 2012.

[10] X. Delfosse, X. Bonfils, T. Forveille, S. Udry, M. Mayor, F. Bouchy, M. Gillon, C. Lovis, V. Neves, F. Pepe, C. Perrier, D. Queloz, N. C. Santos, and D. Ségransan, "The HARPS search for southern extra-solar planets. XXXIII. Super-Earths around the M-dwarf neighbors Gl 433 and Gl 667C," *Astronomy Astrophysics*, vol. 553, p. A8, May 2013.

[11] R. A. Wittenmyer, M. Tuomi, R. P. Butler, H. R. A. Jones, G. Anglada-Escudé, J. Horner, C. G. Tinney, J. P. Marshall, B. D. Carter, J. Bailey, G. S. Salter, S. J. O'Toole, D. Wright, J. D. Crane, S. A. Schectman, P. Arriagada, I. Thompson, D. Minniti, J. S. Jenkins, and M. Diaz, "GJ 832c: A Super-Earth in the Habitable Zone," *The Astrophysical Journal*, vol. 791, p. 114, Aug. 2014.

[12] J. Horner and B. W. Jones, "Determining habitability: which exoEarths should we search for life?" *International Journal of Astrobiology*, vol. 9, pp. 273–291, Oct. 2010.

[13] V. Dehant, H. Lammer, Y. N. Kulikov, J.-M. Grießmeier, D. Breuer, O. Verhoeven, Ö. Karatekin, T. van Hoolst, O. Korablev, and P. Lognonné, "Planetary Magnetic Dynamo Effect on Atmospheric Protection of Early Earth and Mars," *Space Science Reviews*, vol. 129, pp. 279–300, Mar. 2007.

[14] B. E. Wood, "Astrospheres and Solar-like Stellar Winds," *Living Reviews in Solar Physics*, vol. 1, 2004.

[15] B. J. Wargelin and J. J. Drake, "Stringent XRay Constraints on Mass Loss from Proxima Centauri," *The Astrophysical Journal*, vol. 578, no. 1, pp. 503–514, Oct. 2002.

[16] D. J. Mullan, J. G. Doyle, R. O. Redman, and M. Mathioudakis, "Limits on detectability of mass loss from cool dwarfs," *The Astrophysical Journal*, vol. 397, p. 225, Sep. 1992.

[17] M. Semel, "Zeeman-Doppler imaging of active stars. I - Basic principles," *Astronomy and Astrophysics*, vol. 225, pp. 456–466, Nov. 1989.

[18] B. Carter, S. Brown, J. F. Donati, D. Rees, and M. Semel, "Zeeman Doppler imaging of stars with the AAT," *Publications of the Astronomical Society of Australia*, vol. 13, pp. 150–155, May 1996.

[19] J. F. Donati and S. F. Brown, "Zeeman-Doppler imaging of active stars. V. Sensitivity of maximum entropy magnetic maps to field orientation." *Astronomy and Astrophysics*, vol. 326, pp. 1135–1142, Oct. 1997.

[20] J. F. Donati, "Magnetic cycles of HR 1099 and LQ Hydrae," *Monthly Notices of the Royal Astronomical Society*, vol. 302, no. 3, pp. 457–481, Jan. 1999.

[21] K. G. Powell, P. L. Roe, T. J. Linde, T. I. Gombosi, and D. L. De Zeeuw, "A Solution-Adaptive Upwind Scheme for Ideal Magnetohydrodynamics," *Journal of Computational Physics*, vol. 154, no. 2, pp. 284–309, Sep. 1999.

[22] S. C. Marsden, P. Petit, S. V. Jeffers, J. Morin, R. Fares, A. Reiners, J. D. do Nascimento, M. Aurière, J. Bouvier, B. D. Carter, C. Catala, B. Dintrans, J. F. Donati, T. Gastine, M. Jardine, R. Konstantinova-Antova, J. Lanoux, F. Lignières, A. Morgenthaler, J. C. Ramirez Velez, S. Théado, V. Van Grootel, and B. Collaboration, "A BCool magnetic snapshot survey of solar-type stars," *Monthly Notices of the Royal Astronomical Society*, vol. 444, no. 4, pp. 3517–3536, Aug. 2014.

[23] C. Catala, J. F. Donati, E. Shkolnik, D. Bohlender, and E. Alecian, "The magnetic field of the planet-hosting star τ Bootis," *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 374, no. 1, pp. L42–L46, Jan. 2007.

[24] J. F. Donati, C. Moutou, R. Fares, D. Bohlender, C. Catala, M. Deleuil, E. Shkolnik, A. C. Cameron, M. M. Jardine, and G. A. H. Walker, "Magnetic cycles of the planet-hosting star τ Bootis," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 3, pp. 1179–1185, Apr. 2008.

[25] R. Fares, J. F. Donati, C. Moutou, D. Bohlender, C. Catala, M. Deleuil, E. Shkolnik, A. Collier Cameron, M. M. Jardine, and G. A. H. Walker, "Magnetic cycles of the planet-hosting star τ Bootis - II. A second magnetic polarity reversal," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 3, pp. 1383–1391, Sep. 2009.

[26] A. A. Vidotto, R. Fares, M. Jardine, J. F. Donati, M. Opher, C. Moutou, C. Catala, and T. I. Gombosi, "The stellar wind cycles and planetary radio emission of the τ Boo system," *Monthly Notices of the Royal Astronomical Society*, vol. 423, no. 4, pp. 3285–3298, May 2012.

[27] G. Takeda, E. B. Ford, A. Sills, F. A. Rasio, D. A. Fischer, and J. A. Valenti, "Structure and Evolution of Nearby Stars with Planets. II. Physical Properties of ~1000 Cool Stars from the SPOCS Catalog," *The Astrophysical Journal Supplement Series*, vol. 168, pp. 297–318, Feb. 2007.

[28] G. Torres, J. Andersen, and A. Giménez, "Accurate masses and radii of normal stars: modern results and applications," *The Astronomy and Astrophysics Review*, vol. 18, pp. 67–126, Feb. 2010.

[29] J. A. Valenti and D. A. Fischer, "Spectroscopic Properties of Cool Stars (SPOCS). I. 1040 F, G, and K Dwarfs from Keck, Lick, and AAT Planet Search Programs," *The Astrophysical Journal Supplement Series*, vol. 159, pp. 141–166, Jul. 2005.

[30] G. T. van Belle and K. von Braun, "Directly Determined Linear Radii and Effective Temperatures of Exoplanet Host Stars," *The Astrophysical Journal*, vol. 694, pp. 1085–1098, Apr. 2009.

[31] B. D. Mason, G. L. Wycoff, W. I. Hartkopf, G. G. Douglass, and C. E. Worley, "The 2001 US Naval Observatory Double Star CD-ROM. I. The Washington Double Star Catalog," *The Astronomical Journal*, vol. 122, pp. 3466–3471, Dec. 2001.

[32] J. F. Donati, M. Semel, B. D. Carter, D. E. Rees, and A. Collier Cameron, "Spectropolarimetric observations of active stars," *Monthly Notices of the Royal Astronomical Society*, vol. 291, p. 658, Nov. 1997.

[33] S. C. Marsden, J. F. Donati, M. Semel, P. Petit, and B. D. Carter, "Surface differential rotation and photospheric magnetic field of the young solar-type star HD 171488 (V889 Her)," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 1, pp. 468–476, Jul. 2006.

[34] I. A. Waite, S. C. Marsden, B. D. Carter, R. Hart, J. F. Donati, J. C. Ramirez Velez, M. Semel, and N. Dunstone, "Magnetic fields and differential rotation on the pre-main sequence - III. The early-G star HD 106506," *Monthly Notices of the Royal Astronomical Society*, vol. 413, no. 3, pp. 1949–1960, Apr. 2011.

[35] S. C. Marsden, M. M. Jardine, J. C. Ramirez Velez, E. Alecian, C. J. Brown, B. D. Carter, J. F. Donati, N. Dunstone, R. Hart, M. Semel, and I. A. Waite, "Magnetic fields and differential rotation on the pre-main sequence -I. The early-G star HD 141943 - brightness and magnetic topologies," *Monthly Notices of the Royal Astronomical Society*, vol. 413, no. 3, pp. 1922–1938, Apr. 2011.

[36] M. Jardine, A. A. Vidotto, A. van Ballegooijen, J. F. Donati, J. Morin, R. Fares, and T. I. Gombosi, "Influence of surface stressing on stellar coronae and winds," *Monthly Notices of the Royal Astronomical Society*, vol. 431, no. 1, pp. 528–538, Apr. 2013.

[37] D. Ulusen, J. G. Luhmann, Y.-J. Ma, S. Ledvina, T. E. Cravens, K. Mandt, J. H. Waite, and J.-E. Wahlund, "Investigation of the force balance in the Titan ionosphere: Cassini T5 flyby model/data comparisons," *Icarus*, vol. 210, pp. 867–880, Dec. 2010.

[38] A. A. Vidotto, M. Jardine, M. Opher, J. F. Donati, and T. I. Gombosi, "Powerful winds from low-mass stars: V374 Peg," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 1, pp. 351–362, Mar. 2011.

[39] K. A. Kozarev, R. M. Evans, N. A. Schwadron, M. A. Dayeh, M. Opher, K. E. Korreck, and B. van der Holst, "Global Numerical Modeling of Energetic Proton Acceleration in a Coronal Mass Ejection Traveling through the Solar Corona," *The Astrophysical Journal*, vol. 778, p. 43, Nov. 2013.

[40] A. A. Vidotto, M. Opher, V. Jatenco-Pereira, and T. I. Gombosi, "Three-dimensional Numerical Simulations of Magnetized Winds of Solar-like Stars," *The Astrophysical Journal*, vol. 699, no. 1, pp. 441–452, Jul. 2009.

[41] A. A. Vidotto, M. Opher, V. Jatenco-Pereira, and T. I. Gombosi, "Simulations of Winds of Weak-lined T Tauri Stars. II. The Effects of a Tilted Magnetosphere and Planetary Interactions," *The Astrophysical Journal*, vol. 720, no. 2, pp. 1262–1280, Sep. 2010.

Cassini for Australian Stellar Science

Paul N. Stewart*, Peter G. Tuthill*, Phil D. Nicholson† and Matthew M. Hedman‡

* *Sydney Institute for Astronomy, School of Physics, H90, The University of Sydney, NSW 2006, Australia*
† *Department of Astronomy, Cornell University, Ithaca, NY 14853, USA*
‡ *Physics Department, University of Idaho, Moscow, ID 83844, USA*

Summary: Australia is significantly under-represented in space research globally. It is, however, possible for local researchers to use the public archives of the major space agencies to perform state-of-the-art research with large-scale multinational spacecraft, to which Australia is not a contributor. Taking advantage such public data we demonstrate the recovery of flux calibrated stellar spectra spanning the near infrared using observations from the Cassini spacecraft. In spite of the poor transmission of the atmosphere in the near infrared, there has yet to be a dedicated space telescope to provide a spectroscopic survey. Using these recovered spectra from Cassini observations we report on a pilot study to close this gap for many of the brightest stars. We demonstrate the use of stellar occultations by Saturn's rings to measure the size of molecular layers in the atmosphere of evolved stars. The use of this occultation technique to perform high resolution imaging using tomography is also shown.

Keywords: Cassini, Mira, Stellar Spectra, Stellar Imaging, Evolved Stars

Introduction

Cassini is a flagship-class spacecraft which was a joint mission between NASA, ESA and ASI (Agenzia Spaziale Italiana) to study Saturn and its complex environment. This includes, but is not limited to the planet's moons, rings, and magnetosphere. The $3.3 billion dollar cost of this mission was shared between these agencies in a 80/15/5 % split. As Australia is not a contributor to the mission, there is little room to be directly involved with the remarkable science produced. Nevertheless, it is possible to take advantage of NASA's open data policy to obtain data acquired with this spacecraft.

Amongst the most detrimental factors affecting the ability for astronomers to observe the stars is the Earth's atmosphere. It degrades starlight both with its physical motion, and molecular absorption and emission, which is dominated by H_2O and CO_2 in the near infrared. Therefore, as Cassini is far beyond any atmosphere, it is ideally situated for observing the stars. Stellar science is not within the intended mission envelope of the Cassini mission, so it has not been designed for this purpose. However, there exists an extensive archive of observations including stars, which were acquired for either calibration purposes or as part of studies of the Saturnian system which can be analysed for stellar science.

This first section is a discussion of Australia's global standing in space research and the flourishing open data movement. This is followed by the detailing of the use of Cassini's on-board near infrared spectrometer to observe stellar spectra. Stewart et al. [1] showed that Cassini can also be used for measuring spatial information of target stars. It was demonstrated that archival observations of a stellar occultations by the planet's rings yielded a measurement of the stellar angular diameter. The final section describes new enhancements made to this technique, enabling measurements of the spatial structure of circumstellar environments and in particular combining this technique with tomography to recover 2D images of stars.

Australia's Place in Global Space Research

Historically Australia was one of the early adopters in the space industry. The WRESAT satellite was designed, built, and launched by Australians from the Woomera Test Range. This 1967 launch made Australia the fourth country to launch their own satellite from their own territory, albeit with a modified US Redstone missile, donated for the launch. Unfortunately there have only been a small number of Australian designed and built satellites launched since. Most of these were for telecommunications use by Optus, and have been on-off devices.

Table 1 details the space research spending of the world's twenty largest national economies. Of these twenty richest countries, Australia is one of three which has no national space agency, and the only one which has no plans to join a multinational organisation. This complete lack of Australian commitment limits the active involvement in large scale space exploration for Australians. It also makes Australia entirely dependent on the grace of foreign powers for satellite access in support of critical factors such as weather, communications, and defence.

The Open Data Movement

Fortunately there is an opportunity for Australian researchers to take advantage of the remarkable work of other nations and still contribute to the field. This opportunity comes in the form of "Open Data": the idea that certain kinds of data should be publicly available for all to use without restriction or charge. In line with this principle, several of the major agencies endeavour to make their data publicly available. The ESA member countries mostly use ESA for their data archives, which can be found at the Research and Scientific Support Department website, rssd.esa.int. JAXA makes data from Japanese missions available at darts.isas.jaxa.jp. CNES publishes data for French only spacecraft, which can be found by searching smsc.cnes.fr for the specific mission (e.g. CoRoT).

NASA has the largest and most varied set of databases which are indexed at data.nasa.gov. The work in this paper takes advantage of the Planetary Data System's (PDS) Imaging Node Atlas found at pds-imaging.jpl.nasa.gov although there are many other databases for many different kinds of datasets. The PDS is the location for all public Cassini data from all twelve of its instruments. Data are released to the public through the PDS by JPL (NASA's Jet Propulsion Laboratory) and the individual instrument teams every few months. Usually the data in the release are around two years old at this time, allowing the instrument teams a secure priority on any science results to be obtained from the observations. The PDS imaging atlas enables fine detail searching of the extensive catalogue of observations. It is possible to refine a search by the expected instrument, observing modes, celestial coordinates, and other constrains, allowing for the identification and acquisition of specific datasets. These datasets do not contain any reference to a proposal, or request, so the specific purpose of the observations is usually not known. The PDS also provides some limited data reduction and processing tools, however these are not relevant to our project.

Table 1: Spending and major space activities of the world's largest 20 countries by GDP. All values are for 2014 where information is available, or for the nearest earlier year otherwise. Astronauts flown counts those from either national programs, intergovernmental programs (such as Interkosmos) or through commercial arrangements (e.g. UK and Saudi Arabia) (1): China does not publicly release details of their spending, so this number is at best an estimate, and the true value could be a factor of 2-3 either larger or smaller. (2): INTA states that their budget "exceeds a hundred million euros" [15]. A more accurate figure was not found. (3): Turkey is in the process of negotiation to join ESA. (4): Saudi Arabia is part of a group of Arab League nations who are planning to develop a "Pan-Arab Space Agency" (PASA) modelled on ESA.

Countries by GDP	GDP (M US$) [2]	Interplanetary Missions (Other Strengths)	Total Astronauts	National Agency	Budget (M US$)	% of GDP (with ESA)	ESA (M €) [3]
US	16,800,000	All Planets (Apollo, Shuttle, ISS)	334	NASA	17,715 [4]	0.105	
China	9,240,270	NEO, Moon (Own Space Station)	10	CNSA	~10,000 [5](1)	0.108	
Japan	4,901,529	Mercury, Venus, Mars, Asteroids	9	JAXA	2,030 [6]	0.041	
Germany	3,634,822	ESA and Galileo	11	DLR	2,000 [7]	0.055 (0.082)	766
France	2,734,949	ESA (Ariane launch systems)	9	CNES	2,410 [8]	0.088 (0.123)	755
UK	2,521,380	ESA, Mars	1	UKSA	517 [9]	0.021 (0.034)	270
Brazil	2,245,673	(ISS)	1	AEB	121 [10]	0.005	
Russia	2,096,777	Venus, Mars (Soyuz, Salut, Mir, ISS)	117	Roscosmos	5,500 [11]	0.262	
Italy	2,071,306	ESA and Cassini	6	ASI	1,300 [12]	0.063 (0.084)	350
India	1,876,797	Mars (Launch Systems)	1	ISRO	1,100 [13]	0.059	
Canada	1,826,768	(Shuttle, ISS, JWST)	9	CSA	438 [14]	0.024 (0.025)	20
Australia	1,560,597						
Spain	1,358,262	ESA	1	INTA	>126 [15](2)	0.009 (0.022)	140
South Korea	1,304,553	(Launch Systems)	1	KARI	443 [16]	0.034	
Mexico	1,260,914		1	AEM	8 [17]	0.001	
Indonesia	868,345			LAPAN	20 [17]	0.002	
Turkey	820,206			(3)			
Netherlands	800,173	ESA	2	NSO	139 [18]	0.017 (0.037)	125
Saudi Arabia	745,272		1	(4)			
Switzerland	650,376	ESA	1	SSO	9 [19]	0.001 (0.026)	127

Stellar Spectra with Cassini

Amongst the excellent suite of instruments on-board Cassini is a compound spectrometer spanning the spectrum from $350\,nm$ to $5.1\,\mu m$. This instrument is known as the Visible and Infrared Mapping Spectrometer (VIMS) and is comprised of two bore-sighted telescopes feeding independent spectrometers. The smaller telescope, with a $4.5\,cm$ diameter aperture observes 0.35-1.05 μm in 96 channels, whilst the larger, with a $23\,cm$ diameter aperture observes 0.8-5.1 μm in 256 channels providing broad and coarse spectral sampling [20]. VIMS is an imaging spectrometer, designed for hyperspectral imaging of surfaces and atmospheres within the Saturnian system.

This part of the spectrum is highly contaminated when observed from the ground mainly due to absorption and emission from H_2O, CO_2 and molecular oxygen. Other contributors include Rayleigh and Mie scattering which produce wavelength dependant partial opacity. For this reason stellar spectroscopy in these wavelengths tends to be limited to telluric windows where the atmosphere is relatively transparent. Alternatively it is possible to use a space platform to make such observations. ESA's Infrared Space Observatory (ISO) was a survey telescope capable of observing this spectral region redward of $2.4\,\mu m$. The International Ultraviolet Explorer (IUE) performed a similar function beyond the blue end of VIMS' range, without overlapping. Whilst the Hubble Space Telescope is capable of observing in these wavelengths, it is not a survey telescope so many stars have not been observed. DIRBE (Diffuse Infrared Background Experiment) on COBE (Cosmic Background Explorer) performed a broadband photometric survey to provide stellar magnitude at certain wavelengths within this region, rather than spectra for the brightest stars. As a consequence, VIMS has a unique opportunity to record spectra for stars in this spectral region, providing data unobtainable from any other source.

As part of an internal calibration and instrumental monitoring programme, VIMS has been used to observe many of the brightest stars in the sky since its launch over 17 years ago. These observations were never intended to be used for stellar science, but due to the unique opportunity VIMS potentially provided to the stellar astrophysics community the authors decided that the entire archive of such observations should be flux calibrated and published as a spectral atlas. Such an atlas is expected to have substantial scientific value far beyond the goals of the Cassini mission. Most of the stars which have been observed are evolved Carbon stars, Mira variables, and M and K giants, as well as some main sequence A stars, and nearby G stars.

Upon investigation, it was discovered that the visible data from VIMS-V rarely had sufficient signal-to-noise to be able to compete with existing ground based observations, so it was decided to focus only on the 256 spectral channels in VIMS-IR. The process for calibrating VIMS spectra is well established for the kinds of observations usually made, which is filled pixel imaging of atmospheres, rings or solid surfaces [21]. This calibration process needs modification for underfilled pixels, as occurs when observing a star, which significantly underfills VIMS' large 228 x 493 μrad pixel. In the case of an unresolved star, this size is the diffraction limit of the telescope, which for $3\,\mu m$ light in the infrared channel is $16\,\mu rad$. The centres of the pixels are separated on a 250 x 500 μrad grid, leaving gaps between. Without sub-pixel pointing control, the core of the PSF is going to fall partly into these gaps in approximately 10% of observations. Such an observational configuration prohibits accurate flux calibration and affected observations are disregarded.

For the majority of the data, unaffected by the pixel gaps, the calibration process was relatively straightforward. The raw counts are extracted from the spectral image cube, and the appropriate background subtraction applied dependent on the instrumental observing mode. These counts where then divided by the area of the effective aperture of the telescope and the exposure time producing a spectra in counts/m^2/s. This was converted to photons/m^2/s by multiplying by the sensitivity function, a combination of the detector sensitivity and electronic gain. We then divide by the instrumental bandpass $\Delta\lambda$ and multiply by the photon energy to produce a spectrum in J/m^2/s/μm. Finally this is converted to Janskys (10^{-26}W/m^2/Hz) in order to make comparisons to existing spectra in the literature, specifically ISO [22]. These recovered spectra were directly compared to overlapping DIRBE photometry and found to generally agree to within 1%.

Example Spectra

Figure 1 shows the recovered spectra for 10 evolved stars of varying types. The first 5 plots are of pulsating variable red giants with spectral classifications M4.5, M6, M6, M3 and M5.5. CW Leo (or IRC+10216) is a known carbon star with a spectrum very close to a black body due to its thick dusty envelope. L2 Pup is an M5 semi-regular pulsating star and the remaining plots are of different types of Mira variable. The dominant peaks and dips in the spectrum below 2.4 μm reveal water molecules in the stellar atmospheres, whilst the shape at longer wavelengths can reveal the abundance of other molecules, primarily CO (~4.2-4.8 μm), CO_2 (~4.2-4.5 μm) and OH (~3.0-3.6 μm) [23].

The spikes appearing in some of the spectra at the red end are due to low signal levels as both the stars get faint and the instrumental sensitivity drops. These occur due to a division by the instrumental sensitivity function on pixels where only a few detector counts are recorded. Hot pixels in the detector are the cause of the spikes that occur in some of the spectra at around 1.3 and 2.1 μm. The gaps visible in each spectra around 1.6, 3.0 and 3.9 μm are caused by partial obstructions due to filter edges within the instrument and are excluded. The spectra are too coarse to resolve individual lines, but do allow for the detection of spectral bands and broad photometry. In the 2 Cen plot there are four crosses indicating comparable photometry obtained with the DIRBE instrument on board the COBE spacecraft [24]. These crosses fall almost exactly onto the recovered spectra validating the recovery process.

Spatial Information from Cassini Stellar Occultations

One of the VIMS' observing modes involves staring at a background star as features of the Saturnian system intersect this line of sight. These events are known as stellar occultations and have been used extensively to study the planetary rings [25], [26]. From the Earth, the Moon has been used as an occulter for the study of the heavens since antiquity, with limited use of Lunar occultations for modern scientific stellar studies occurring over the last half century [29].

Stewart et al. [1] demonstrated the use of stellar occultations by Saturn's rings with VIMS for the study of the stellar source. The starlight is diffracted by sharp edges within the rings, such as the edges of gaps and ringlets, producing a Fresnel diffraction pattern. This pattern is recorded by Cassini as it moves along its trajectory producing a lightcurve, which is the

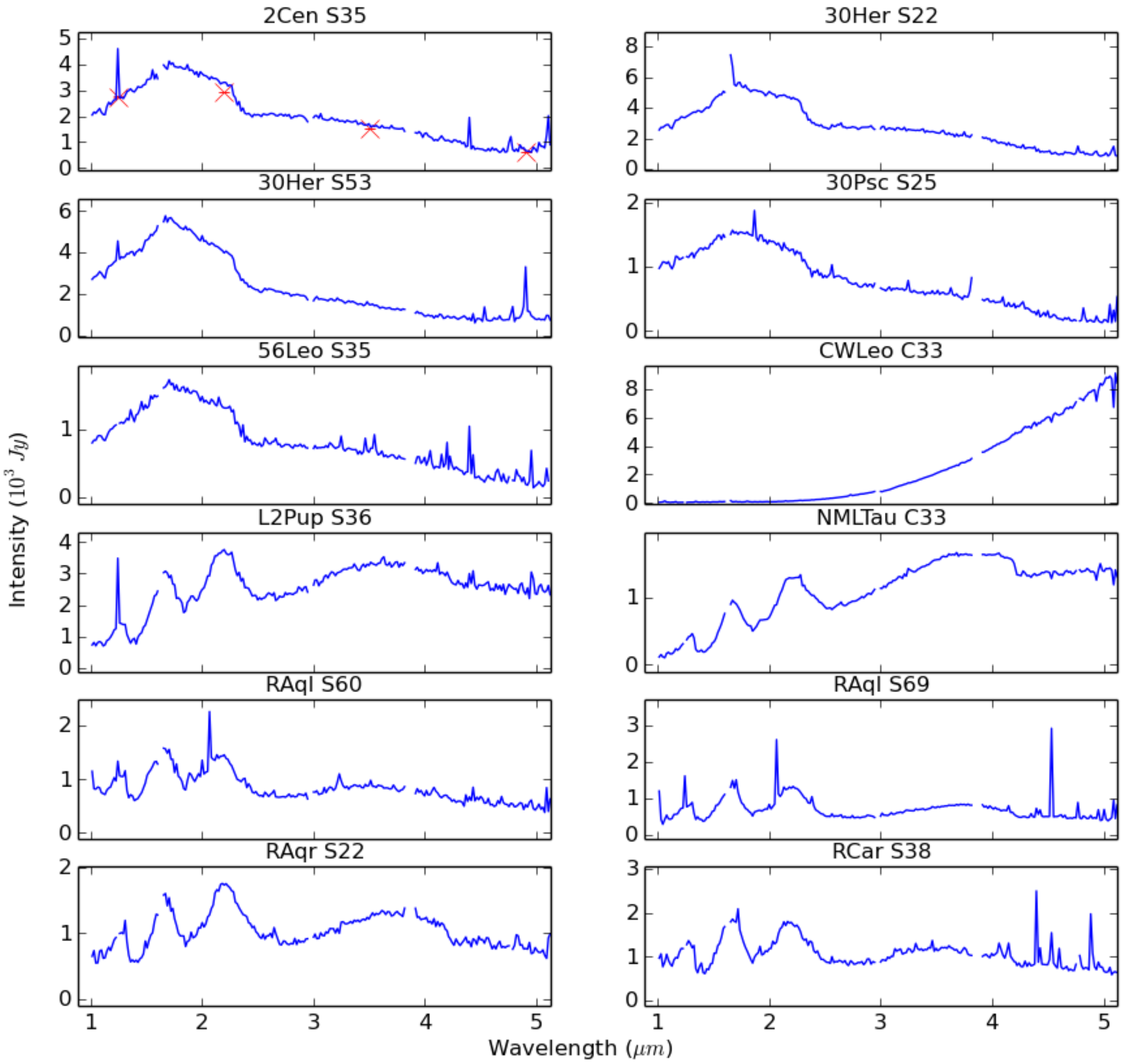


Fig. 1: Some example recovered spectra. The horizontal axis is wavelength in microns and the vertical axis is the calibrated intensity in units of 10^3 Janskys. The plot labels include the star name, followed by the 'sequence' they were observed in. 'C' and 'S' letters denote successive planning sequences for when the spacecraft was in cruise to the planet, or the time in orbit around it respectively. The red crosses in the 2Cen plot show corresponding DIRBE photometry.

convolution of the brightness profile of the source and diffraction pattern of a point source with the same occultation event orbital geometry. Determining this geometry makes it is possible to produce this point source lightcurve allowing the recovery of the stellar brightness profile. Using certain sharp edges within the ring system, angular diameter of the source star Mira were measured through the fitting of a simple model. This star is the archetype of the Mira variable class of stars, which are known to exhibit large variations in both size and luminosity with a period usually around one year. Such stars are known to be responsible for up to 75% of the chemical enrichment of the interstellar medium, and to eject dusty circumstellar shells.

Figure 2 shows a comparison between Cassini-derived size spectrum and measurements acquired using terrestrial interferometry. The model used to fit these data is a simple uniform

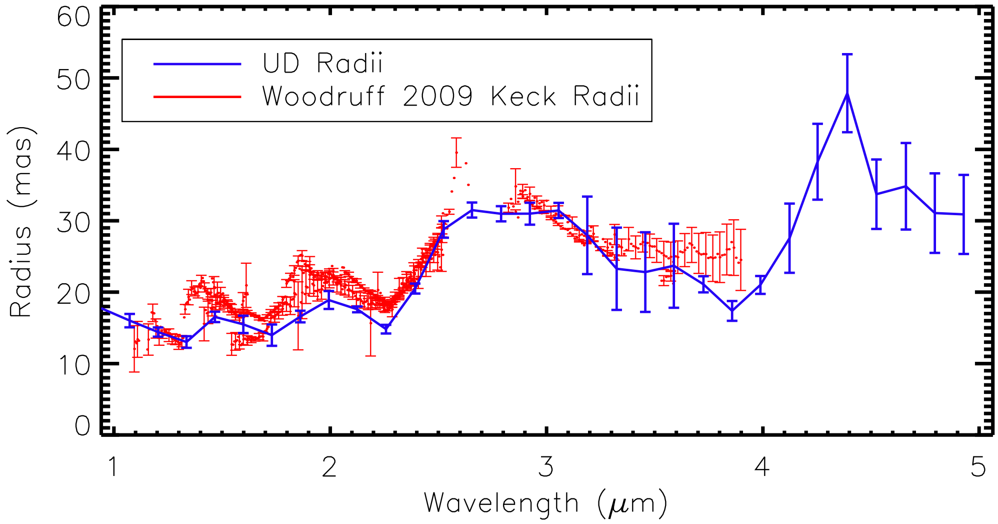


Fig. 2: from Fig.5, Stewart et al. [1]: Fitted uniform disc radii across all observed wavelengths for a single occultation event. The fitted uniform disc radii are shown in blue with one sigma uncertainties. The red data shows previously published interferometric measurements [30].

disc, where the star is assumed to be an evenly illuminated circle in the sky. This model was selected partly for its simplicity, having only a single free parameter. The uniform disc model is also used extensively for high angular resolution measurements in the literature, providing a direct comparison to other techniques in spite of not being physically realistic. Figure 2 illustrates the clear correlation between the measurements using VIMS occultations (in blue) and the literature values (in red). The slightly smaller value found with VIMS across all wavelengths is likely due to the inherent variable nature of the star.

Complex Model Fitting

As stars like Mira are known to exhibit departures from simple geometries and also to inhabit dusty environments, the uniform disc model may be an oversimplification. In order to better understand the circumstellar environment of the star more complex models were fitted. The first addition to this model attempted is a thin spherical concentric shell. This shell is allowed to vary in radius and flux ratio with respect to the star, which has a free radius, giving the model three free parameters. This model is expected to produce a more realistic description of the nature of the circumstellar environment of Mira.

Figure 3 shows the results of this more complex model fitted over all wavelengths. The first thing to note from this plot is the reduction in the size of the fitted uniform disc once the shell is added (orange versus blue). This is due to some of the flux previously counting as the outer parts of the disc now being belonging to the shell. It is also important to note that the radius of the single shell falls into two discrete wavelength-dependent levels. The inner of these detected shells is around twice the radius of the disc (~28 mas), whilst the outer is around five to six times the disc radius (~70 mas). The existence of different shells being

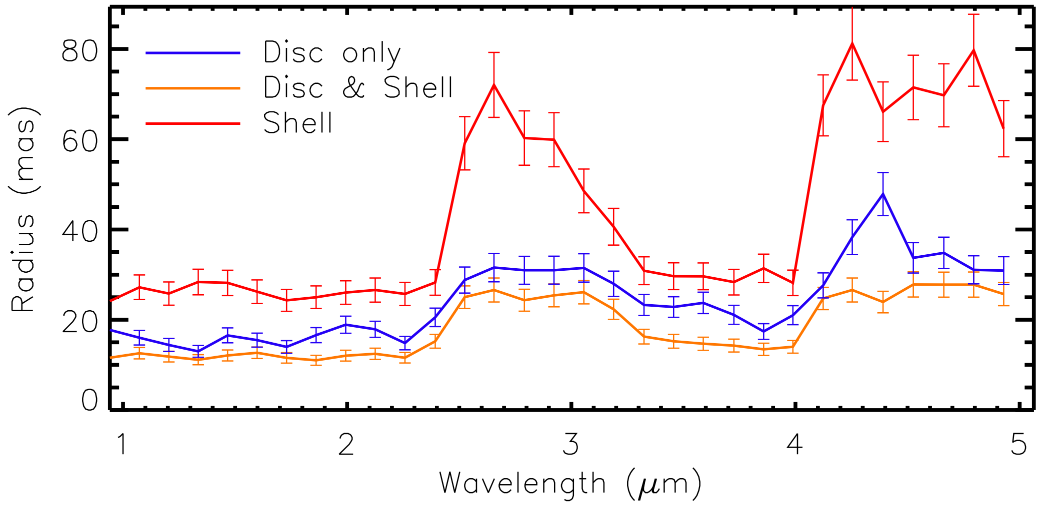


Fig. 3: The results of fitting a uniform disc model both with and without a circumstellar concentric shell. The blue line shows the fit to a uniform disc only, as per Figure 2. The yellow shows the radii of uniform disc after the addition of a shell. The radii of this fitted shell is shown in red. All error bars show at one sigma to fitted values.

dominant at different wavelengths reveals the thermal and molecular nature of these features, with both shells found to correspond to hot H_2O layers, the inner at ~2000 Kelvins, and the outer reaching only ~1200 Kelvins [23].

This relatively modest increase in model complexity has yielded a major enhancement in understanding the nature of the star. We have been able to measure the radius of circumstellar shells and determine their molecular composition and temperature due to the spectral observed dependence. Models allowing asymmetries and more complex morphologies are being investigated and will be presented in an forthcoming publication.

Tomographic Image Reconstruction

Tomography is a technique that uses overlapping projections in order to build a higher-dimensional image. With these observations, high resolution two dimensional images of the source star can be recovered using overlapping projections obtained from different ring edges. As the star is observed to pass behind the rings, it encounters many hard edges, and each of these yields a one dimensional projection of the star in a different direction. This direction is calculated to be an angle relative to the terrestrial celestial north, the direction towards the Earth's celestial north pole from any point in the sky. This angle of projection is in the direction normal to the ring edge as observed from Cassini's sky plane. Each of the one dimensional projections is a unique, but overlapping profile of the star. Figure 4 shows a comparison of projection angles available to two occultations of the rings with similar orbital geometry.

The models used above to determine the spatial structure of the circumstellar environment are very effective at parameterising the modelled features. However, they also enforce such structure in the brightness profiles, even in cases where it isn't realistic. By fitting a model-

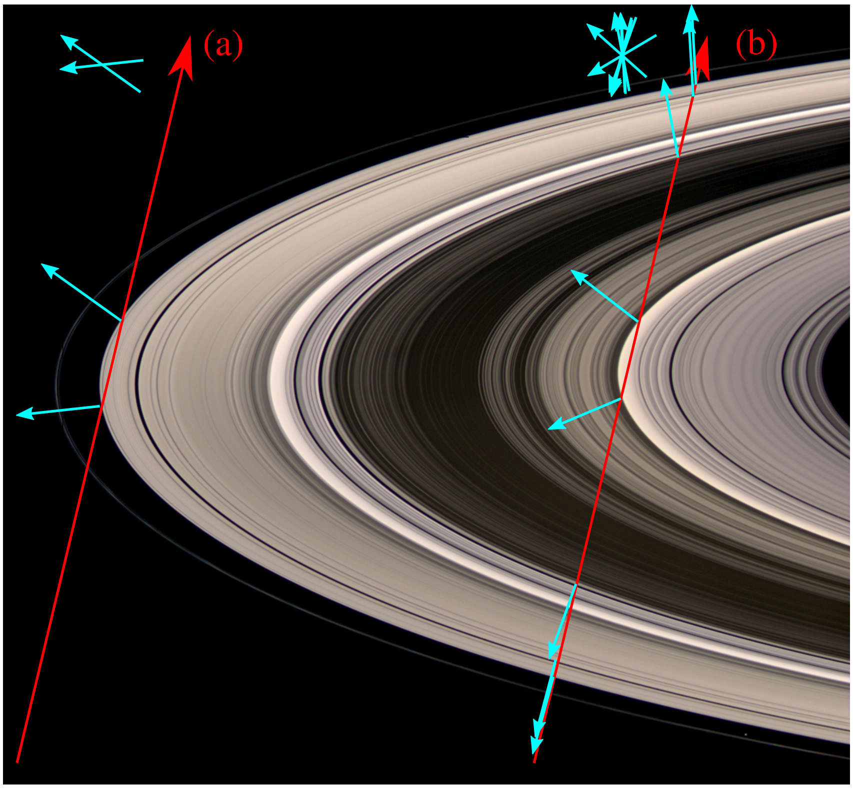


Fig. 4: from Fig. 3 Stewart et al. (2014 Submitted). A comparison of angular diversity for representative occultation geometries using indicative edges. The long red vectors indicate the apparent trajectory of the star behind the rings. The short cyan vectors identify the normal to some of the occulting edges the star would pass in these events. (a) shows a minimal occultation with only two edges that are parallel. (b) shows an almost identical orbital geometry, with vastly different angular diversity. The cyan vector clusters to the top left of each shows an angular diversity rose, a quick visual method to compare angular diversity between observations. Base image: PIA10446 [31]

independent brightness profile to each of these projections we have have a vision of what the star would look like, without any enforced structure. This is important for the imaging process in order to prevent the imposition of spurious structure.

These overlapping one dimensional projections and model-independent brightness profiles provide the requisite inputs to perform tomographic image reconstruction. The use of tomographic imaging within medicine is quite mature, although the use of most existing tools requires significantly higher density and uniformity of projection angles. For this work the *Scikit-learn* [32] project's machine learning regression algorithms, implemented in *Python* were found to be sufficiently versatile for the data available. The process can be repeated for each spectral channel enabling high angular resolution hyperspectral imaging of stellar targets. The full details of the image reconstruction process is presented in Stewart et al. (2014 Submitted).

Figure 5 shows the results of a tomographic image reconstruction revealing a high resolution image of Mira. The star and inner shell are seen to be surrounded by a much larger shell. This image exhibits the same shells measured with models above. The outer shell is seen to have a large degree of asymmetry, with significantly more flux in the north-east of the image.

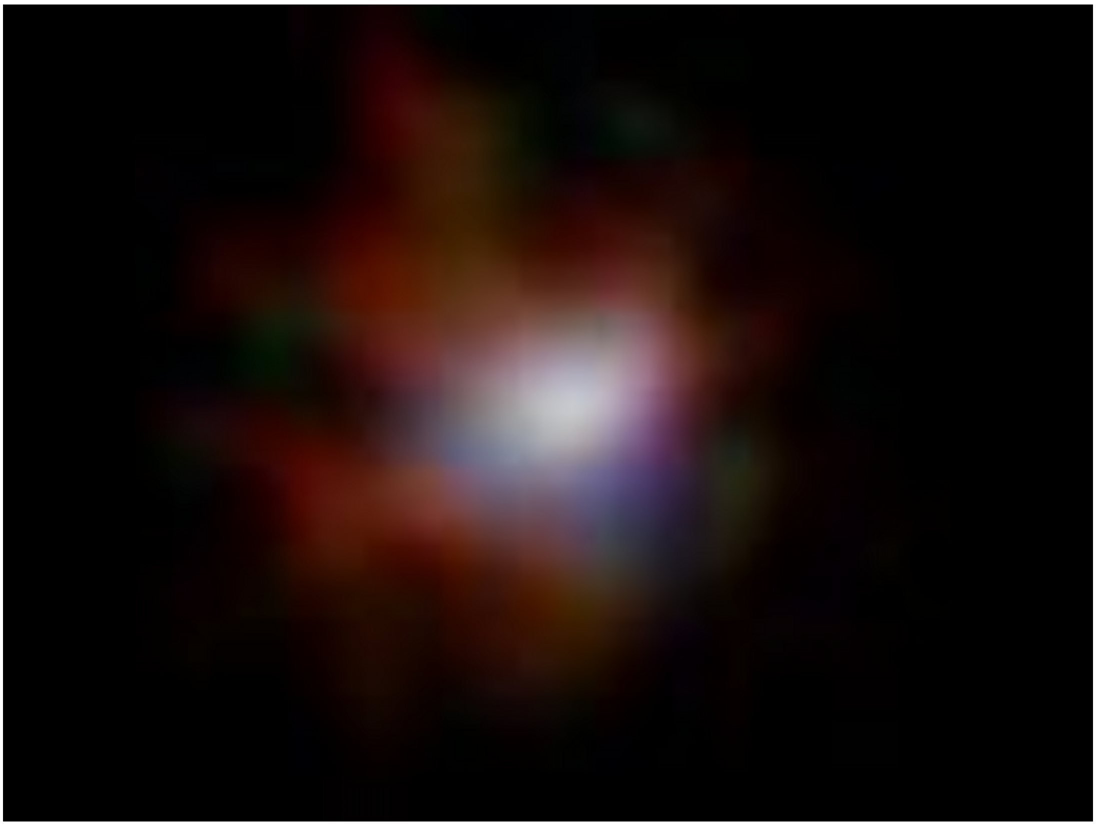


Fig. 5: A three colour tomographic reconstruction of Mira. The Red, Green, and Blue represent reconstructions at 2.0, 3.3 and 4.7 μm respectively. The bright core of the star (white) can be seen to be surrounded by cooler layers (red and green) which correspond to the shells detected in the one dimensional brightness profile model fitting.

Conclusion

Table 1 demonstrates quite clearly that Australia is lacking in the opportunities to be involved in space research. In spite of this unfortunate situation, it has been shown that taking advantage of the open data policies of the major global space agencies is one potential avenue in which Australians can be directly involved with cutting-edge space science. Using such data, we have demonstrated the recovery of flux-calibrated stellar spectra spanning the near infrared. We have also shown advances to the technique since Stewart et al. [1]. These include the expansion of the model being used to allow for more complex circumstellar environments, and the use of model-independent brightness profile fitting and tomographic image reconstruction to recover high resolution images of Mira and its circumstellar shells.

References

[1] Stewart, P.N., Tuthill, P.G., Hedman, M.M., Nicholson, P.D., and Lloyd, J.P. *High Angular Resolution Stellar Imaging with Occultations from the Cassini Spacecraft I: Observational Technique*, Monthly Notices of the Royal Astronomical Society, Oxford Journals, 2013. 433 (3), 2286-2293

[2] World Bank Open Data *The World Bank Group*, 2014 http://data.worldbank.org/indicator/NY.GDP.MKTP.CD
[3] ESA Budget 2014 *ESA*, 2014 http://www.esa.int/spaceinimages/Images/2014/01/ESA_budget_2014
[4] NASA *FY 2014 Budget Proposal*, 2014 http://www.nasa.gov/content/fy-2014-budget-proposal/
[5] Pike, J. *Budget Expenditures*, 2014 http://www.globalsecurity.org/space/world/china/budget.htm
[6] JAXA *Statements of Accounts*, 2014 http://www.jaxa.jp/about/finance/pdf/finance_25-04.pdf
[7] Drescher, J. *The Year Ahead in Space*, 2014 http://csis.org/files/attachments/101015_DLR%20Perspectives.pdf
[8] CNES *2012 CNES Budget*, 2012 http://www.cnes.fr/web/CNES-en/7403-2009-cnes-budget.php
[9] UKSA *UK Space Agency Annual Report and Accounts 2013-14*, 2014 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335126/Annual_Report_2013_to_2014.pdf
[10] AEB *National Program of Space Activities 2005-2014*, 2014 http://www.aeb.gov.br/wp-content/uploads/2013/03/PNAE_INGLES.pdf
[11] Russian Space Systems *News of Cosmonautics*, 2014 http://www.spacecorp.ru/press/branchnews/item5598.php
[12] ASI *Strategic Vision 2010-2020*, 2014 http://www.asi.it/files/ASI_DVS_2010_2020_ENG_0_0.pdf
[13] ISRO *OUTCOME BUDGET OF THE DEPARTMENT OF SPACE GOVERNMENT OF INDIA 2013-2014*, 2014 http://www.isro.gov.in/pdf/Outcome%20Budget2013-14.pdf
[14] CSA *THE CANADIAN SPACE AGENCY 2013-14 Estimates REPORT ON PLANS AND PRIORITIES* , 2014 http://publications.gc.ca/collections/collection_2013/asc-csa/ST96-7-2013-eng.pdf
[15] INTA *INTA in Figures*, 2014 http://www.inta.es/IntaEnCifras.aspx#A1
[16] KARI *Institutional management disclosures*, 2014 http://www.alio.go.kr/alio/public/p_ma_14_00_list.jsp?org_code=C0292
[17] Harding, R. *Space Policy in Developing Countries: The search for security and development on the final frontier*, 2013
[18] NSO *Programmes and Participations*, 2014 http://www.spaceoffice.nl/en/Space-Directory/Policy-and-Programmes/Programmes-and-Participations/
[19] Benz, W. *Switzerland and Space Sciences: an over 40 years old success story*, 2009 http://www.sps.ch/index.php?id=118&L=1
[20] Brown, Robert H., Baines, Kevin H., Bellucci, G., Bibring, J.-P., Buratti, Bonnie J., Capaccioni, F., Cerroni, P., Clark, Roger N., Coradini, A., Cruikshank, D. P., Drossart, P., Formisano, V., Jaumann, R., Langevin, Yves, Matson, D. L., Mccord, T. B., Mennella, V., Miller, Edward A., Nelson, R. M., Nicholson, Philip D., Sicardy, B., Sotin, C. *The Cassini Visual And Infrared Mapping Spectrometer (Vims) Investigation* Space Science Reviews, 115 (2004)
[21] McCord, T. B., Coradini, A., Hibbitts, C. A., Capaccioni, F., Hansen, G. B., Filacchione, G, Clark, Roger N., Cerroni, P., Brown, Robert H., Baines, Kevin H. *Cassini VIMS observations of the Galilean satellites including the VIMS calibration procedure* Icarus, 172 104-126, (2004)
[22] G.C. Sloan, K.E. Kraemer, S.D. Price, R.F. Shipman *A uniform database of 2.4-45.4 micron spectra from the ISO Short Wavelength Spectrometer* 2003, ApJS, 147, 379

[23] Yamamura, I., de Jong, T., Cami, J. *Infrared emission of hot water in the atmosphere of Mira* Astronomy and Astrophysics, v.348, p.L55-L58 (1999)

[24] Arendt, R. G., Berriman, G. B., Franz, B. A., Freudenreich, H. T., Lisse, C. M., Mitchell, K. J., Mitra, P. M., Moseley, S. H., Newmark, J. S., Odegard, N. P., Rawley, G. L., Reach, W. T., Silverberg, R. F., Skard, J. A. J., Stemwedel, S. W., Toller, G. N. *COBE Diffuse Infrared Background Experiment (DIRBE) Explanatory Supplement. COBE Ref. Pub. No98-A* 1998, 79-85

[25] Nicholson, P. D.; Hedman, M. M. *Self-gravity wake parameters in Saturns A and B rings* (2010), Icarus, Volume 206, Issue 2, p. 410-423.

[26] Harbison, Rebecca A.; Nicholson, Philip D.; Hedman, Matthew M. *The smallest particles in Saturns A and C Rings* (2013), Icarus, Volume 226, Issue 2, p. 1225-1240

[27] Hedman, M. M.; Gosmeyer, C. M.; Nicholson, P. D.; Sotin, C.; Brown, R. H.; Clark, R. N.; Baines, K. H.; Buratti, B. J.; Showalter, M. R. *An observed correlation between plume activity and tidal stresses on Enceladus* Nature, Volume 500, Issue 7461, pp. 182-184 (2013)

[28] Hedman, M. M.; Nicholson, P. D. *More Kronoseismology with Saturn's rings* (2014), Monthly Notices of the Royal Astronomical Society, Volume 444, Issue 2, p.1369-1388

[29] White, N *Lunar occultations: From conjecture to results* (1987), Vistas in Astronomy, 30, 13-25

[30] Woodruff, H. C.; Ireland, M. J.; Tuthill, P. G.; Monnier, J. D.; Bedding, T. R.; Danchi, W. C.; Scholz, M.; Townes, C. H.; Wood, P. R. *The Keck Aperture Masking Experiment: Spectro-Interferometry of Three Mira Variables from 1.1 to 3.8 m* The Astrophysical Journal, Volume 691, Issue 2, pp. 1328-1336 (2009)

[31] Cassini Imaging Team *Catalog Page for PIA12518* http://photojournal.jpl.nasa.gov/catalog/PIA12518 (2010)

[32] F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. *Scikit-learn: Machine Learning in Python* Journal of Machine Learning Research 12, 28252830 (2011)

Alteration minerals in the Great Artesian Basin: Comparison to weathering processes on Mars

Gemma E. Roberts [1], Patrice F. Rey [1], Helen E. A. Brand [2] and Balwant Singh [3]

[1] *EarthByte Group, School of Geosciences, The University of Sydney, Sydney, NSW, 2006, Australia*
[2] *Australian Synchrotron, Clayton, VIC, 2601, Australia*
[3] *Faculty of Agriculture and Environment, The University of Sydney, Sydney, NSW, 2006, Australia*

Summary: Terrestrial analogues are essential to study planetary surface processes as they can be examined in great detail and at low cost. Here we show that the regolith of the Great Artesian Basin (GAB) of central Australia has similar mineralogy to that documented at the surface of Mars. Our analyses of two drill cores (down to 300 m) from the GAB using visible to near infrared spectroscopy and synchrotron x-ray diffraction documents a four-layer weathering profile comparable to that documented at the surface on Mars. We conclude that the GAB could be one of the best terrestrial analogues for regional Martian weathering processes.

Keywords: Great Artesian Basin, terrestrial analogue, Mars, acidic weathering, VNIR spectroscopy, XRD

Introduction

High spectral resolution data of the Martian surface have been obtained by the CRISM (Compact Reconnaissance Imaging Spectrometer for Mars) and OMEGA (Observatoire pour la Mineralogie, l-Eau, la Glace et l'Activite) instruments. These have documented the spectral signatures for a range of hydrated alteration minerals including: Al-rich phyllosilicates (e.g. kaolinite and montmorillonite), Fe/Mg-rich phyllosilicates (e.g. nontronite and saponite), sulfates, iron oxides, and opaline silica [e.g. 1, 2].

On Mars, Al-rich phyllosilicates are commonly detected overlying Fe/Mg smectites in Nili Fossae [3], Valles Marineris [4], and Mawrth Vallis [1, 3, 5, 6]. Here we show that minerals in these Martian weathering profiles and their distribution within the profile are similar to those in the weathering profile of feldspar-rich volcaniclastic sedimentary rocks in the GAB of central Australia.

In the first section of this paper we present spectral data of the mineralogy across the weathering profile in the GAB. We then present an overview of the mineralogy observed on Mars, in particular in the regions of Nili Fossae, Valles Marineris, and Mawrth Vallis. Finally, we discuss the similarities between the weathering profiles in the GAB and those observed on Mars, and we suggest that central Australia could be one of the best terrestrial analogues for the evolution of the surface of Mars at the transition between the Noachian and Hesperian.

Alteration profiles observed in the Great Artesian Basin

From 130 to 97 Ma, volcaniclastic sediments were deposited in the cold, muddy, anoxic and shallow Eromanga Sea [7, 8]. Following sea regression from 97 to 93 Ma, a protracted period of uplift, erosion, and denudation lasting until *ca*. 60 Ma [9], resulting in the oxidation of biogenic pyrite releasing sulfuric acid into the environment [10, 11]. This long period of denudation under cold [12] and acidic oxidative conditions [13] produced a deep (~100 m) weathering profile, characterised by kaolinite, iron oxides, and amorphous silica, grading down through a mottled zone of kaolinite and montmorillonite, to montmorillonite at the base [14].

Materials and Methods

The mineralogy of the weathering profile in the GAB were documented along two cores using visible to near infrared (VNIR) spectroscopy, and synchrotron x-ray diffraction (XRD). Samples were collected from Tickalara 1 (142°09'25"E, 28°38'30"S) and Thargomindah 1 (143°27'00"E, 27°17'30"S) cores in southwest Queensland (Fig. 1).

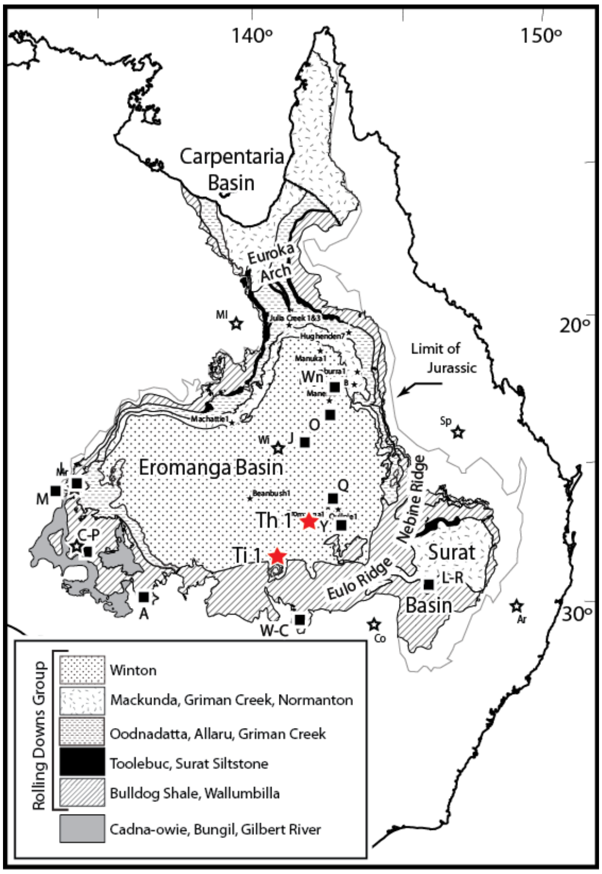


Fig. 1: Location of Tickalara 1 (Ti 1) and Thargomindah 1 (Th 1) on a simplified geological map of the Lower Cretaceous geology of the GAB (modified from [11]).

VNIR spectra were acquired through HyLogger^{TM} technology (a CSIRO developed non-destructive hyperspectral mapping technique) and an ASD FieldSpec Portable Spectroradiometer. The HyLogger^{TM} system uses an ASD VNIR spectrometer, scanning core and obtaining spectra at 8 mm intervals. Spectra is matched to laboratory spectra for mineral identification and corrected by a user, if required. VNIR spectroscopy was used to focus on a number of mineralogically representative intervals identified along the HyLogger™ scan of the top 300 m in both cores.

For higher resolution characterisation and sensitivity to minor phases, XRD analysis was collected at the powder diffraction beamline at the Australian Synchrotron using a Mythen II microstrip detector. Datasets were collected with a wavelength of 0.68934 Å between 3° and 83° two-theta for 5 minutes at two detector positions, and merged to remove gaps between the detector modules. An internal standard, 20% by weight NIST SRM 674 ZnO, was added to each sample to aid with quantification.

Tables 1 and 2 summarise the sampling dataset for Tickalara 1 and Thargomindah 1.

Table 1: Dataset for Tickalara 1

Core number	*Start depth (m)*	*End depth (m)*	*VNIR spectroscopy*	*XRD*
1-1	12.05	12.13	□	□
1-2	18.87	19.04	□	□
1-3	24.61	24.66	□	
1-4	25.00	25.03	□	
1-5	25.61	25.65	□	□
1-6	27.68	27.88	□	
1-7	30.56	30.58	□	
1-8	32.39	32.45	□	
1-9	38.96	39.00	□	□
1-10	53.61	53.73	□	
1-11	59.32	59.33	□	□
1-12	62.05	62.12	□	
1-13	67.59	67.66	□	
1-14	74.09	74.17	□	
1-15	77.68	77.80	□	
1-16	146.90	146.94	□	
1-17	154.05	154.10	□	□
1-18	174.83	174.88	□	
1-19	235.26	235.28	□	□
1-20	273.94	274.03	□	

Table 2: Dataset for Thargomindah 1

Core number	*Start depth (m)*	*End depth (m)*	*VNIR spectroscopy*	*XRD*
2-1	82.13	82.22	□	
2-2	82.32	82.36	□	
2-3	137.21	137.31	□	□
2-4	165.94	165.96	□	□

Mineralogy of the weathering profile

Spectral reflectance data

Twenty-four samples were collected down to a depth of 300 m in both wells for VNIR analyses. The spectra in Fig. 2 are from samples 2-1, 2-3, and 2-4, and were identified as being representative of the main mineralogical layers through the weathered section.

Kaolinite: Spectrum A (2-1) in Fig. 2 exhibits an asymmetrical Al-OH absorption doublet at 2.16 µm and 2.21 µm, a feature indicative of kaolin minerals [15].

Montmorillonite: Spectrum A in Fig. 2 exhibits hydration absorptions (H_2O and OH) at 1.41 µm and 1.91 µm, and an Al-OH absorption doublet (indicative of kaolinite) at 2.16 µm to 2.21 µm. Spectrum B (2-3) in Fig. 2 has a small singular absorption feature at 2.21 µm. These features resemble the reflectance spectrum of the Al smectite montmorillonite [15].

Goethite: Spectrum A in Fig. 2 has a weak and broad absorption band in the visible range at 0.92 µm. This feature is caused by electronic transitions involving ferric iron and is indicative of goethite [16].

Siderite: Spectrum B in Fig. 2 exhibits a broad absorption feature in the 1.0 µm to 1.3 µm region. This feature is commonly associated with ferrous iron [17]. The small absorption feature near 2.34 µm is likely due to CO_3 vibrations [15]. These two features closely resemble the reflectance spectrum of siderite.

Gypsum: Spectrum C (2-4) in Fig. 2 exhibits an H_2O absorption feature at 1.43 µm and an H_2O absorption doublet at 1.94 µm and 1.96 µm. Typically, these features are diagnostic of a minor amount of gypsum [18].

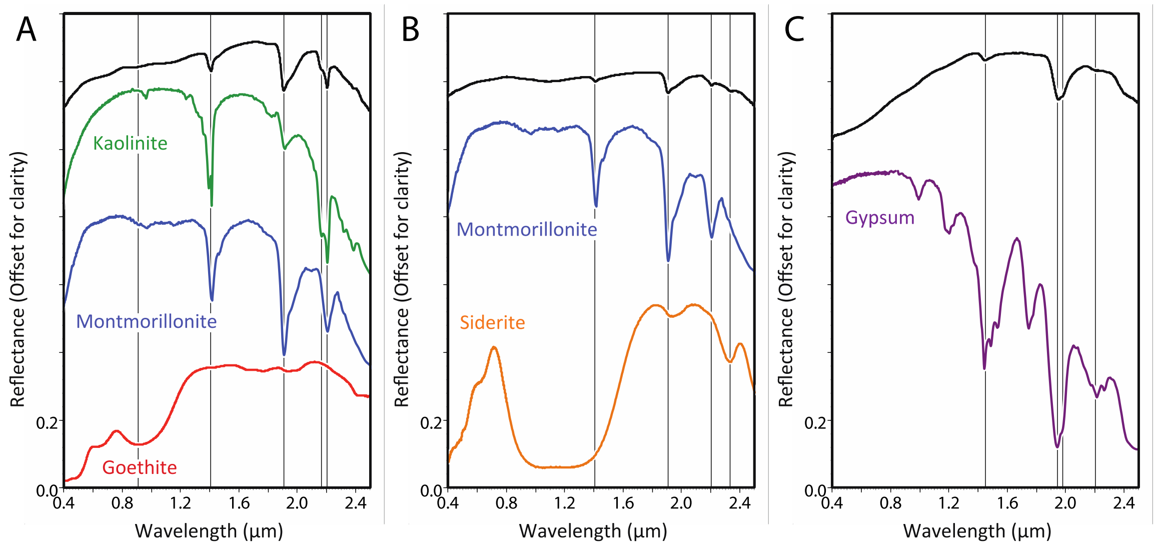


Fig. 2: Spectral reflectance spectra of major weathering minerals in Thargomindah 1. The top black spectrum in each plot was obtained using the ASD FieldSpec spectrometer, and the coloured lines shown below are USGS laboratory spectra of identified minerals. (A) Minerals present in sample 2-1 are goethite, kaolinite, and montmorillonite. Vertical lines mark absorption features at 1.41, 1.91, 2.16, and 2.21 µm, and the Fe^{3+} electronic transition feature at 0.92 µm; (B) Minerals present in sample 2-3 are montmorillonite and siderite. Vertical lines at 1.41, 1.91, 2.21, and 2.34 µm; (C) Minerals present in sample 2-4 are gypsum, possibly montmorillonite. Vertical lines at 1.43, 1.94, 1.96, and 2.21 µm.

Synchrotron XRD analyses

Nine samples were collected down to a depth of 300 m in both wells for synchrotron XRD analyses. Samples were identified as being representative of the main mineralogical layers through the weathered section. Seven samples from Tickalara 1 and two samples from Thargomindah 1 were analysed. These samples were chosen from those analysed using VNIR reflectance spectroscopy. Spectra shown in Fig. 3 are from samples taken from Tickalara 1. The minerals detected in these spectra were not detected by VNIR reflectance spectroscopy, most likely because these occur as minor phases and were unable to be identified by the lower resolution VNIR spectroscopy. These include nontronite (Fe smectite), detected in sample 1-11 (Fig. 3A), and pyrite, albite, and rozenite, detected in sample 1-19 (Fig. 3B).

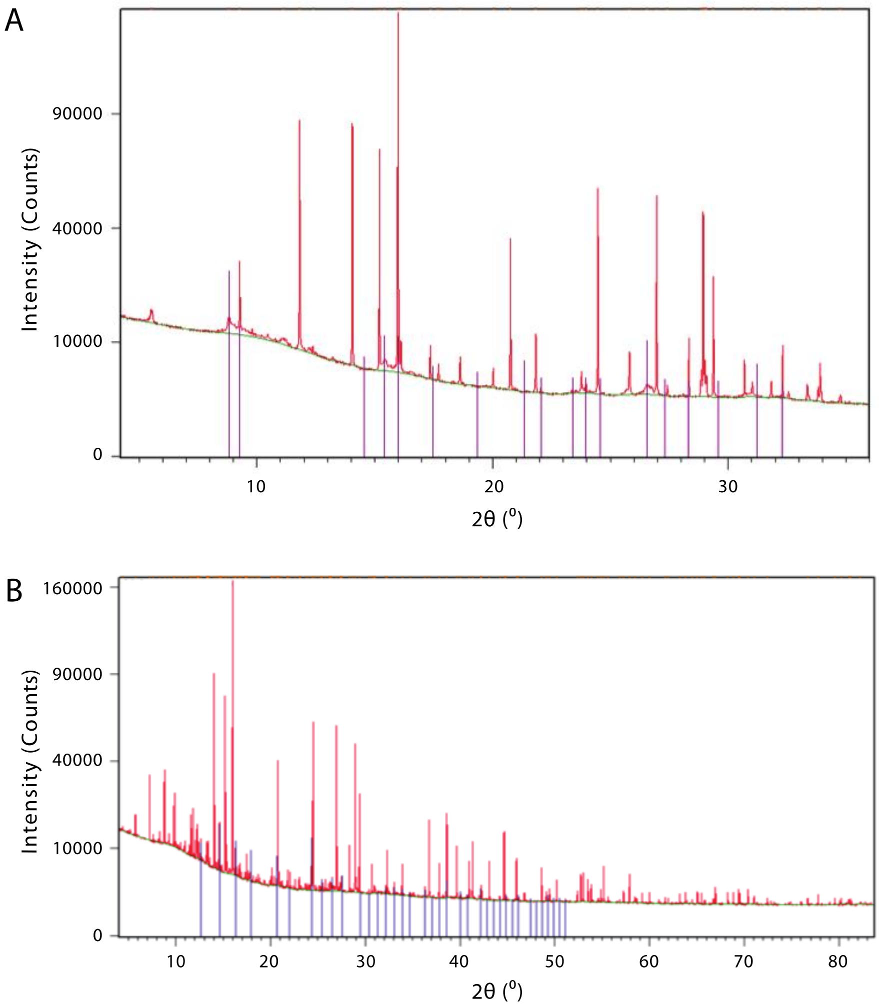


Fig. 3: Synchrotron XRD spectra of minor alteration minerals in Tickalara 1. (A) Spectrum of sample 1-11. Purple vertical lines below the red spectra indicate 2θ peaks corresponding to nontronite; (B) Spectrum of sample 1-19. Blue vertical lines indicate 2θ peaks corresponding to pyrite.

The identification of these minerals has led to the characterisation of a four–layer weathering profile (illustrated in Fig. 4):

1. The upper layer is principally composed of kaolinite, montmorillonite, and goethite. This layer is the host of precious opals in nearby Queensland opal fields, where thin (50 – 80 cm) opalised horizons extend laterally over many tens of kilometres, at a depth of 10 – 40 m [11].

2. Goethite is rare in the upper intermediate layer, in which kaolinite and montmorillonite remain the main minerals. Nontronite was detected via synchrotron XRD analysis.

3. The lower intermediate layer is dominated by montmorillonite with minor siderite and occasional gypsum.

4. The lowermost layer is montmorillonite and gypsum. Synchrotron XRD analysis revealed the presence of pyrite, rozenite, and albite.

A fifth layer of pore-lining and grain-coating nontronite has been documented in the eastern GAB a depth between 200 m and 300 m [19]. This nontronite is found within the Wallumbilla Formation (Fig. 1) (the top of which is at a depth of 420 m to 450 m in Tickalara 1 and Thargomindah 1), and is interpreted to be diagenetic in origin [19].

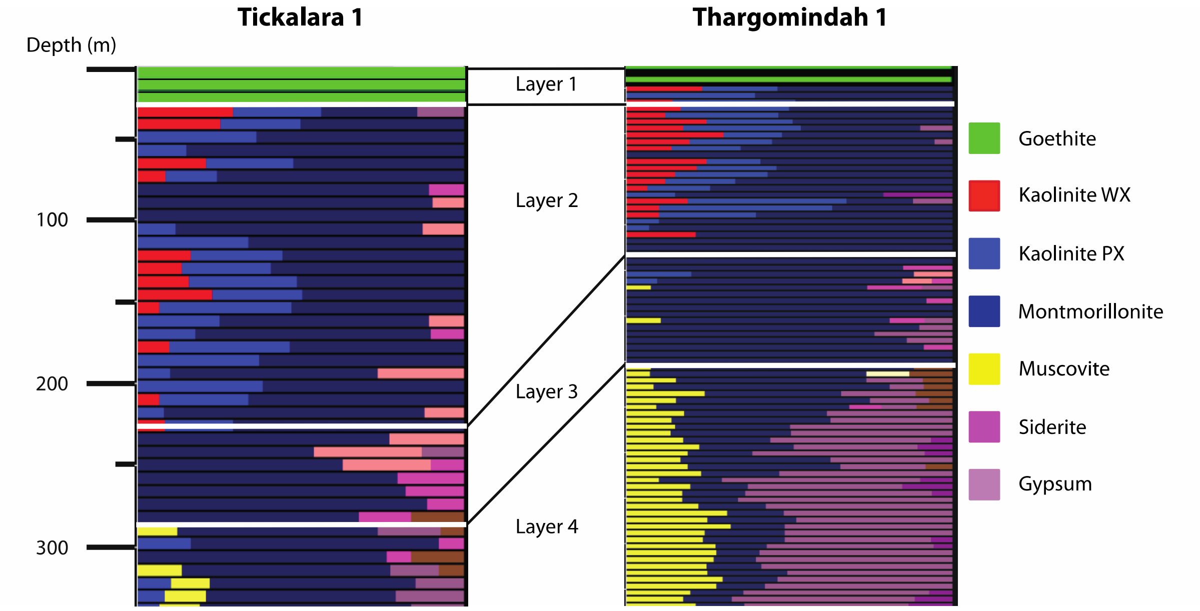


Figure 4: HyLogger^{TM} mineral distribution in Tickalara 1 and Thargomindah 1, down to a depth of 300 m. WX refers to well crystalline and PX refers to poorly crystalline.

HyLogger^{TM} data indicates that the contact between each of these zones is gradual, and occurs at varying depths in Tickalara 1 and Thargomindah 1. The thickness of the top goethite-bearing zone is approximately 25 m in both wells. The upper intermediate kaolinite zone is approximately 200 m in thickness in Tickalara 1, and thins to approximately 90 m in Thargomindah 1. The thickness of the lower intermediate montmorillonite zone in both wells is approximately 40 m to 50 m.

Alteration profiles documented on Mars

Al-rich phyllosilicates at the surface of Mars are commonly observed in Nili Fossae [3], Valles Marineris [4], and Mawrth Vallis [1, 3, 5, 6]. In these regions (Fig. 5) the Al-rich phyllosilicates overlie Fe/Mg smectites [20] and are often intermixed with hydrated silica [6].

No Fe/Mg smectites have been found interbedded with or overlying Al-rich phyllosilicates [2].

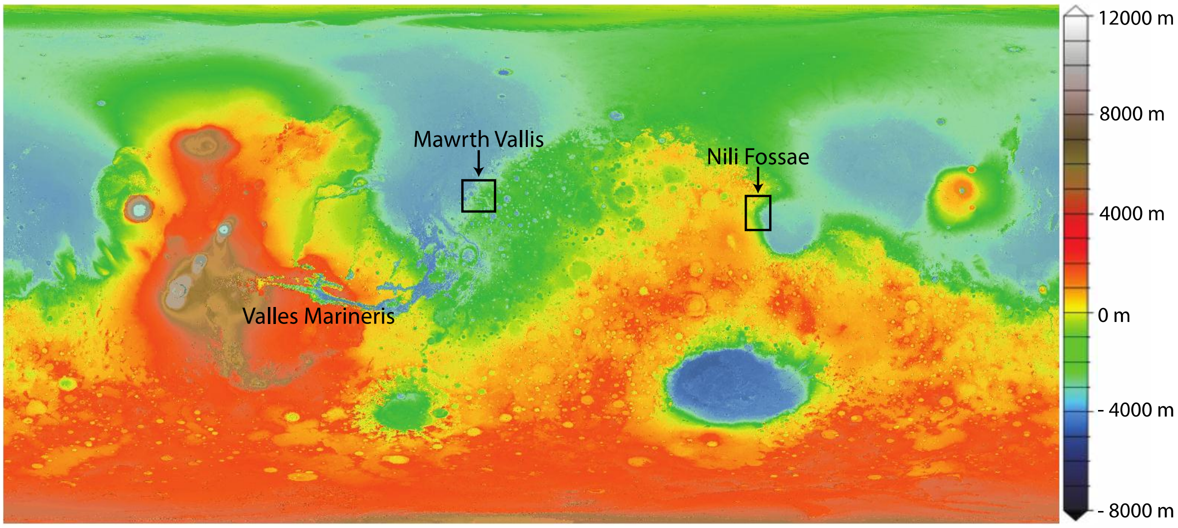


Figure 5: Location of major phyllosilicate outcrops on Mars overlying a MOLA (Mars Orbiter Laser Altimeter) elevation map.

Nili Fossae

Nili Fossae is the name given to a collection of grabens that have been heavily modified by erosion and infilling [21]. Fe/Mg smectites are the lowermost exposed unit in Nili Fossae, and are up to 600 m thick, as observed in the walls of a trough [21, 22]. A kaolinite-bearing unit always occurs above the Fe/Mg smectite unit in the area immediately adjacent to the Nili Fossae, and is tens of meters thick at the most [21]. Spectra clearly document a kaolinite group mineral [21], with evidence that this kaolinite-bearing unit may be intermixed with montmorillonite [3, 21]. Fe/Mg smectites and the kaolinite-bearing unit are observable where a caprock, without a distinct mineralogical signature, has been eroded [21, 22]. In the easternmost part of Nili Fossae olivine-bearing rocks replace kaolinite overlying Fe/Mg smectite [21]. In some places the olivine-bearing unit has been altered to magnesium carbonate [21]. Hydrated silica is typically found in the far west as a component of aeolian dunes [21].

Valles Marineris

The canyon system known as Valles Marineris is traceable over a length of 4000 km. The lowest layers (down to -4 km) in the walls are composed of low-Ca pyroxene- and olivine-bearing materials, overlain by a mineral spectrally consistent with dehydrated saponite or chlorite (Fe/Mg-rich phyllosilicates), with a high-Ca pyroxene-bearing layer above [21]. Overlying these horizons, in the uppermost portion of the wall and plateau (> 3 km), a two-layer sequence of Fe/Mg smectite overlain by Al smectite or kaolinite is observed [21]. Sulfates, including gypsum, occur in subhorizontal layers hundreds of meters thick in the plains surrounding Valles Marineris [23]. Finely stratified (< 1 m) silica-bearing deposits have been observed on the plains, and are less than 100 m thick [24].

Mawrth Vallis

The outflow channel and surrounding plains that encompass Mawrth Vallis present the region with the highest abundance in phyllosilicates on Mars [25], traceable laterally over hundreds to thousands of kilometres [26, 27, 28]. Fe/Mg smectites (typically nontronite) are overlain by

Al-rich phyllosilicates on both sides of the outflow channel, and on the channel floor [29]. Spectra indicate that the Al-rich phyllosilicates are composed of two horizons: 1) Al smectite, likely montmorillonite, intermixed with hydrated silica, sometimes overlain by 2) kaolinite intermixed with hydrated silica [4, 5, 22]. These Al-rich phyllosilicates are estimated to be between 30 m and 60 m thick [30].

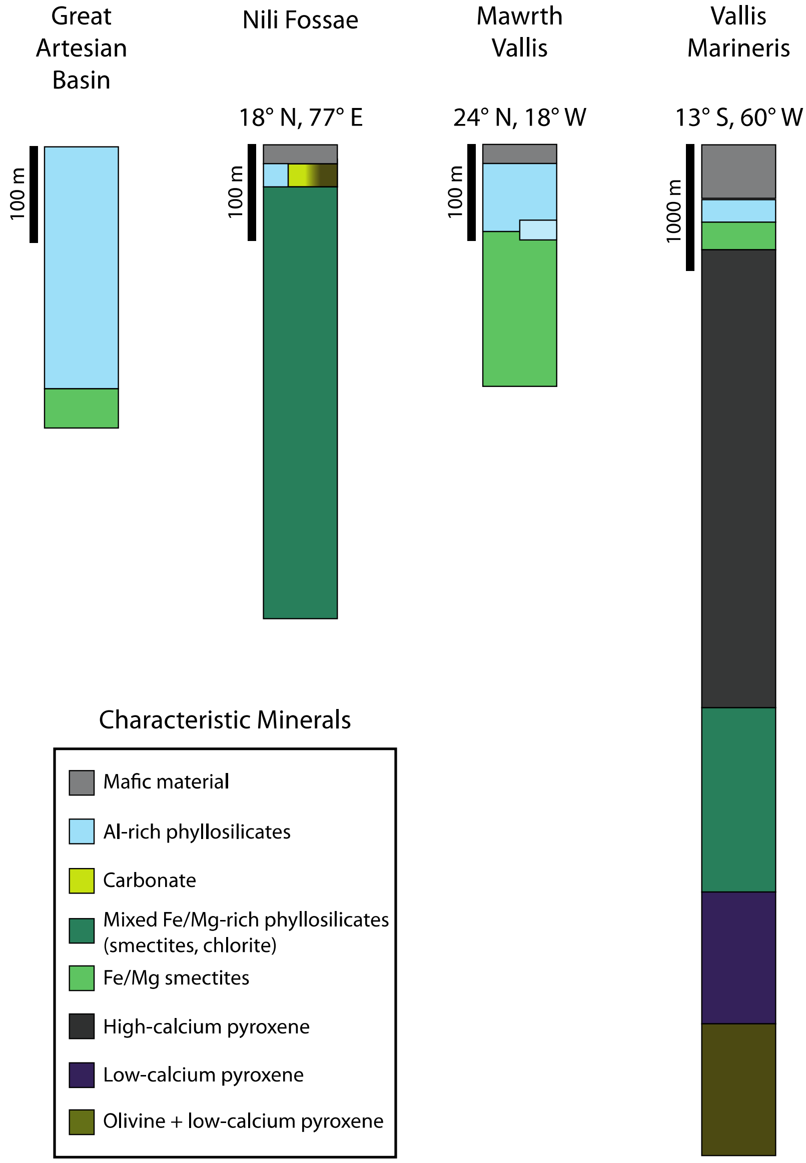


Figure 6: Simplified weathering profile of the GAB with comparison to the mineralogical profiles of Nili Fossae, Valles Marineris, and Mawrth Vallis (modified from [2]). Characteristic minerals or compositions that characterise each unit in VNIR spectra are shown. For ease of comparison, kaolinite-bearing and montmorillonite-bearing layers are illustrated as one Al-rich phyllosilicate layer. Colour gradation indicates that the corresponding minerals occupy the same unit, grading from one to the other

Discussion

The near-surface mineralogical profiles of the GAB and Nili Fossae, Valles Marineris, and Mawrth Vallis on Mars are characterised by a Fe/Mg smectite layer overlain by a layer comprised of Al-rich phyllosilicates, hydrated silica, and iron oxides (Fig. 6). In both the GAB and Mars, this mineralogical sequence has been interpreted as the result of regional-scale acidic oxidative weathering of feldspar-rich volcaniclastic sediments during the drying out of their respective surfaces [1, 11, 13, 31]. Very acidic conditions across the GAB were produced via the oxidation of biogenic pyrite into sulfuric acid during the drying out of the basin 97 myr ago [11], whereas on Mars, acidity derived from the photochemical conversion of volcanic sulfur dioxide into sulfuric acid in the presence of moisture [32]. The weathering profile described in the GAB is strikingly similar to mineralogical sequences observed on Mars, suggesting that the regional weathering processes may have involved similar ingredients (i.e. volcaniclastic sandstones and acidic conditions), despite the difference in proposed origin of acidity. The weathering profile described in the GAB has the potential to be a good terrestrial analogue for Martian regoliths formed via *in situ* surface weathering, hence providing a unique opportunity to explore in more detail weathering pathways, and the possible role of biology in the formation of Martian regoliths.

Acknowledgements

We would like to thank the staff at the Exploration Data Centre in Zillmere, QLD, for granting access to HyLogger^{TM} data for the two drill cores, Jonathan Clarke and Graziella Caprarelli for their helpful reviews, and the proceedings editor Wayne Short for his advice and support.

References

1. Poulet, F., Bibring, J.-P., Mustard, J.F., Gendrin, A., Mangold, N., Langevin, Y., Arvidson, R.E., Gondet, B., Gomez, C. and the Omega Team, "Phyllosilicates on Mars and implications for early martian climate", *Nature Articles*, Vol. 438, 2005, pp. 623-627.

2. Ehlmann, B.L., Mustard, J.F., Murchie, S.L., Bibring, J.-P., Meunier, A., Fraeman, A.A. and Langevin, Y., "Subsurface water and clay mineral formation during the early history of Mars", *Nature Review*, Vol. 479, 2011, pp. 53-60.

3. Cuadros, J. and Michalski, J.R., "Investigation of Al-rich clays on Mars: Evidence for kaolinite-smectite mixed-layer versus mixture of end-member phases", *Icarus*, Vol. 222, 2013, pp. 296-306.

4. Murchie, S.L., Mustard, J.F., Ehlmann, B.L., Milliken, R.E., Bishop, J.L., McKeown, N.K., Noe Dobrea, E.Z., Seelos, F.P., Buczkowski, D.L., Wiseman, S.M., Arvidson, R.E., Wray, J.J., Swayze, G., Clark, R.N., Des Marais, D.J., McEwen, A.S. and Bibring, J.-P., "A synthesis of Martian aqueous mineralogy after 1 Mars year of observations from the Mars Reconnaissance Orbiter", *Journal of Geophysical Research*, Vol. 144, 2009.

5. Bishop, J.L., Noe Dobrea, E.Z., McKeown, N.K., Parente, M., Ehlmann, B.L., Michalski, J.R., Milliken, R.E., Poulet, F., Swayze, G.A., Mustard, J.F., Murchie, S.L. and Bibring, J.-P., "Phyllosilicate Diversity and Past Aqueous Activity Revealed at Mawrth Vallis, Mars" *Science*, Vol. 321, No. 5890, 2008, pp. 830-833.

6. McKeown, N.K., Bishop, J.L., Noe Dobrea, E.Z., Ehlmann, B.L., Parente, M., Mustard, J.F., Murchie, S.L., Swayze, G.A., Bibring, J.-P. and Silver, E.A., "Characterization of

phyllosilicates observed in the central Mawrth Vallis region, Mars, their potential formational processes, and implications for past climate”, *Journal of Geophysical Research*, Vol. 114, 2009.

7. Exon, N.F. and Senior, B.R., “The Cretaceous of the Eromanga and Surat Basins”, *BMR Journal of Australian Geology & Geophysics*, Vol. 1, 1976, pp. 33-50.

8. Bryan, S.E., Constantine, A.E., Stephens, C.J., Ewart, A., Schön, R.W. and Parianos, J., “Early Cretaceous volcano-sedimentary succession along the eastern Australian continental margin: Implications for the break-up of eastern Gondwana”, *Earth and Planetary Science Letters*, Vol. 153, 1997, pp. 85-102.

9. Raza, A., Hill, K.C. and Korsch, R.J., “Mid-Cretaceous uplift and denudation of the Bowen and Surat Basins, eastern Australia: relationship to Tasman Sea rifting from apatite fission-track and vitrinite-reflectance data”, *Australian Journal of Earth Sciences*, Vol. 56, No. 3, 2009, pp. 501-531.

10. Ollier, C. and Pain, C., *Regolith, Soils and Landforms*, John Wiley & Sons Ltd., West Sussex, 1996.

11. Rey, P.F., “Opalisation of the Great Artesian Basin (central Australia): an Australian story with a Martian twist”, *Australian Journal of Earth Sciences: An International Geoscience Journal of the Geological Society of Australia*, Vol. 60, No. 3, 2013, pp. 291-314.

12. Bird, M.I. and Chivas, A.R., “Geomorphic and paleoclimatic implications of an oxygen-isotope chronology for Australian deeply weathered profiles”, *Australian Journal of Earth Sciences*, Vol. 40, No. 4, 1993, pp. 345-358.

13. Senior, B.R. and Mabbutt, J.A., “A proposed method of defining deeply weathered rock units based on regional geological mapping in southwest Queensland”, *Journal of the Geological Society of Australia*, Vol. 26, No. 5-6, 1979, pp. 237-254.

14. Senior, B.R., “Mineralogy and chemistry of weathered and parent sedimentary rocks in southwest Queensland”, *BMR Journal of Australian Geology & Geophysics*, Vol. 4, 1979, pp. 111-124.

15. Clark, R.N., *Chapter 1: Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy, Manual of Remote Sensing, Volume 3, Remote Sensing for the Earth Sciences*, (A.N. Rencz, ed.) John Wiley and Sons, New York, 1999, pp. 3-58.

16. Crowly, J.K., Williams, D.E., Hammarstrom, J.M., Piatak, N., Chou, I.-M. and Mars, J.C., “Spectral reflectance properties (0.3-2.5 μm) of secondary Fe-oxide, Fe-hydroxide, and Fe-sulphate-hydrate minerals associated with sulphide-bearing mine wastes”, *Geochemistry: Exploration, Environment, Analysis*, Vol. 3, 2003, pp. 219-228.

17. Hunt, G.R. and Salisbury, J.W., “Visible and near infrared spectra of minerals and rocks. II. Carbonates”, *Modern Geology*, Vol. 2, 1971, pp. 23-30.

18. King, S.J., Bishop, J.L., Garcia, G.C., Lafuente, B. and Fenton, L.K., “VNIR spectra of gypsum-rich field samples from White Sands New Mexico as an analog study for Olyumpus Undae, Mars”, *Abstract from the 45th Lunar and Planetary Science Conference*, The Woodlands, Texas, USA, March 17-21, 2014.

19. Hawlader, H.M., "Diagenesis and reservoir potential of volcanogenic sandstones – Cretaceous of the Surat Basin, Australia", *Sedimentary Geology*, Vol. 66, 1991, pp. 181-195.

20. Gaudin, A., Dehouck, E. and Mangold, N., "Evidence for weathering on early Mars from a comparison with terrestrial weathering profiles", *Icarus*, Vol. 216, 2011, pp. 257-268.

21. Ehlmann, B.L., Mustard, J.F., Swayze, G.A., Clark, R.N., Bishop, J.L., Poulet, F., Des Marais, D.J., Roach, L.H., Milliken, R.E., Wray, J.J., Bamouin-Jha, O. and Murchie, S.L., "Identification of hydrated silicate minerals on Mars using MRO-CRISM: Geologic context near Nili Fossae and implications for aqueous alteration", *Journal of Geophysical Research*, Vol. 114, 2009.

22. Mustard. J.F., Murchie, S.L., Pelkey, S.M., Ehlmann, B.L., Milliken, R.E., Grant, J.A., Bibring, J.-P., Poulet, F., Bishop, J., Noe Dobrea, E., Roach, L., Seelos, F., Arvidson, R.E., Wiseman, S., Green, R., Hash, C., Humm, D., Malaret, E., McGovern, J.A., Seelos, K., Clancy, T., Clark, R., Des Marais, D., Izenberg, N., Knudson, A., Langevin, Y., Martin, T., McGuire, P., Morris, R., Robinson, M., Roush, T., Smith, M., Swayze, G., Taylor, H., Titus, T. and Wolff, M., "Hydrated silicate minerals on Mars observed by the Mars Reconnaissance Orbiter CRISM instrument", *Letters to Nature*, Vol. 454, 2008, pp. 305-309.

23. Gendrin, A., Mangold, N., Bibring, J.-P., Langevin, Y., Gondet, B., Poulet, F., Bonello, G., Quantin, C., Mustard, J., Arvidson, R. and LeMouélic, S., "Sulfates in Martian Layered Terrains: The OMEGA/Mars Express View", *Science*, Vol. 307, 2005, pp. 1587-1591.

24. Milliken, R.E., Swayze, G.A., Arvidson, R.E., Bishop, J.L., Clark, R.N., Ehlmann, B.L., Green, R.O., Grotzinger, J.P., Morris, R.V., Murchie, S.L., Mustard, J.F. and Weitz, C., "Opaline silica in young deposits on Mars", *Geology*, Vol. 36, 2008, pp. 847-850.

25. Loizeau, D., Mangold, N., Poulet, F., Ansan, V., Hauber, E., Bibring, J.-P., Gondet, B., Langevin, Y., Masson, P. and Neukum, G., "Stratigraphy in the Mawrth Vallis region through OMEGA, HRSC color imagery and DTM", *Icarus*, Vol. 205, 2010, pp. 396-418.

26. Ehlmann, B., "Diverse aqueous environments during Mars' first billion years: the emerging view from orbital visible-near infrared spectroscopy", *Geochemical News*, Vol. 142, 2010.

27. Ehlmann, B.L., Berger, G., Mangold, N., Michalski, J.R., Catling, D.C., Ruff, S.W., Chassefière, E., Niles, P.B., Chevrier, V. and Poulet, F., "Geochemical Consequences of Widespread Clay Mineral Formation in Mars' Ancient Crust", *Space Science Review*, Vol. 174, 2013, pp. 329-364.

28. Loizeau, D., Mangold, N., Poulet, F., Bibring, J.-P., Gendrin, A., Ansan, V., Gomez, C., Gondet, B., Langevin, Y., Masson, P. and Neukum, G., "Phyllosilicates in the Mawrth Vallis region of Mars", *Journal of Geophysical Research*, Vol. 112, 2007.

29. Wray, J.J., Ehlmann, B.L., Squyres, S.W., Mustard, J.F. and Kirk, R.L., "Compositional stratigraphy of clay-bearing layered deposits at Mawrth Vallis, Mars", *Geophysical Research Letters*, Vol. 35, 2008.

30. Noe Dobrea, E.Z., Bishop, J.L., McKeown, N.K., Fu, R., Rossi, C.M., Michalski, J.R., Heinlein, C., Hanus, V., Poulet, F., Mustard, R.J.F., Murchie, S., McEwen, A.S., Swayze, G., Bibring, J.-P., Malaret, E. and Hash, C., “Mineralogy and stratigraphy of phyllosilicates-bearing and dark mantling units in the greater Mawrth Vallis/west Arabia Terra area: Constraints on geological origin”, *Journal of Geophysical Research*, Vol. 115, 2010.

31. Chevrier, V. and Mathé, P.E., “Mineralogy and evolution of the surface of Mars: A review”, *Planetary and Space Science*, Vol. 55, 2007, pp. 289-314.

32. Bullock, M.A. and Moore, J.M., “Atmospheric conditions on early Mars and the missing layered carbonates”, *Geophysical Research Letters*, Vol. 34, 2007.

Empirical Climatology of the Sporadic E Layer (ECSEL) from Radio Occultation Measurements

Lenard Pederick
Defence Science and Technology Organisation
Edinburgh, South Australia
Australia
Email: lenard.pederick@dsto.defence.gov.au

Summary: Sporadic E layers are thin, dense layers of ionization in the E region which appear intermittently and can have strong effects on the propagation of high frequency (HF) radio signals via the ionosphere. In order to be able to predict the effects of these layers, a global, statistical model of sporadic E layers is being developed, using radio occultation measurements taken by the satellite constellation COSMIC.

Introduction

Sporadic E layers are thin, dense layers of ionization in the E region of the ionosphere (about 90 to 120 km in altitude). These layers appear intermittently and are highly variable in time and space, unlike the regular E layer which is thicker and has consistent diurnal, seasonal and solar cycle behaviour. They can have a significant effect on the propagation of radio waves via the ionosphere, by reflecting radio waves up to frequencies of 100 MHz and by blocking propagation via the F region of the ionosphere [1].

Whilst sporadic E layers have been the subject of much research over several decades [2]–[5], their structure, behaviour and the processes that lead to their formation and destruction are still not completely understood.

The inclusion of the effects of sporadic E is important when modelling any system which involves the propagation of radio waves via the ionosphere. In order to account for the effects of sporadic E, a full climatological model of this layer was developed using historical measurements of its behaviour, modelling the geographic, diurnal and seasonal variations in the layer; this will provide a more complete model of the sporadic E layer than previous work done on mapping the sporadic E layer (e.g. [6]).

Data Sources

The data source chosen for use in this model was measurements made via GPS radio occultation. This is a technique which measures a GPS signal as it travels through a limb of the Earth's atmosphere. The retardation and bending due to the Earth's atmosphere results in a phase and Doppler shift, which can be measured very accurately by the GPS receiver aboard a satellite in low-Earth orbit. From these measurements the basic bending angle vs. impact parameter data can be calculated, and vertical profiles of refractivity as a function of tangent

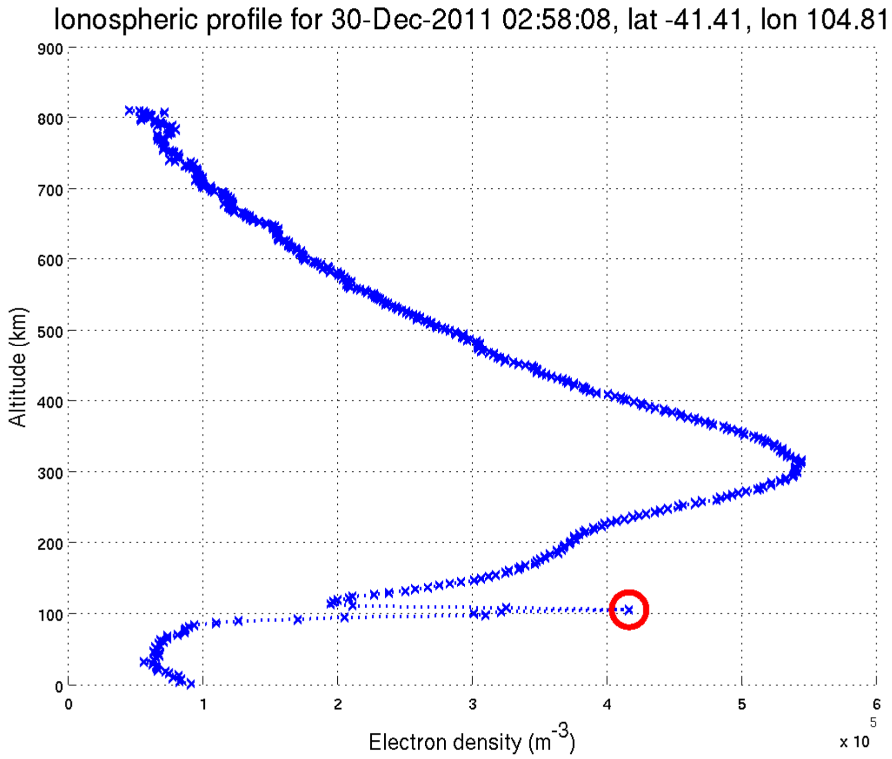


Fig. 1: An example ionospheric profile observed by a COSMIC satellite, with the sporadic E layer highlighted.

point radius can be derived. Further analysis converts the refractivity to electron density in the ionosphere, and temperature, pressure and water vapour in the lower atmosphere [7].

An example electron density profile is shown in Figure 1, which shows an observation with a prominent sporadic E layer, as well as the detected height and peak density of that layer. The layer appears as a narrow 'spike' in the electron density profile around 100 km in altitude. It should be noted that the thickness of the layer is very close to the vertical resolution of the radio occultation technique; thus the sensitivity of this technique to sporadic E layers will be reduced compared to other instruments such as ionospheric sounders.

A unique aspect of this technique is illustrated by Figure 2. This observation shows a regular E layer at a peak height of 111 km, and a sporadic E layer below it at 95 km. The sporadic E layer has a peak density below that of the regular E layer, thus it would be difficult to detect with an ionospheric sounder.

The effect of sporadic E on GPS signals is well known [8], [9], and GPS radio occultation has been previously used to look at the occurrence of sporadic E [10]–[15]. All of these studies analysed scintillations in the raw GPS signal, which under certain conditions can be caused by the small-scale spatial structure of sporadic E layers. However this technique does not provide information about the density of a sporadic E layer, which is a very important parameter for characterising its effect on radio propagation. In the model presented in this paper fully-processed ionospheric profiles were used, which provide density information. It should be noted that the presence of a sporadic E layer will generally violate the assumption of spherical symmetry used in the Abel transform when calculating an electron density profile

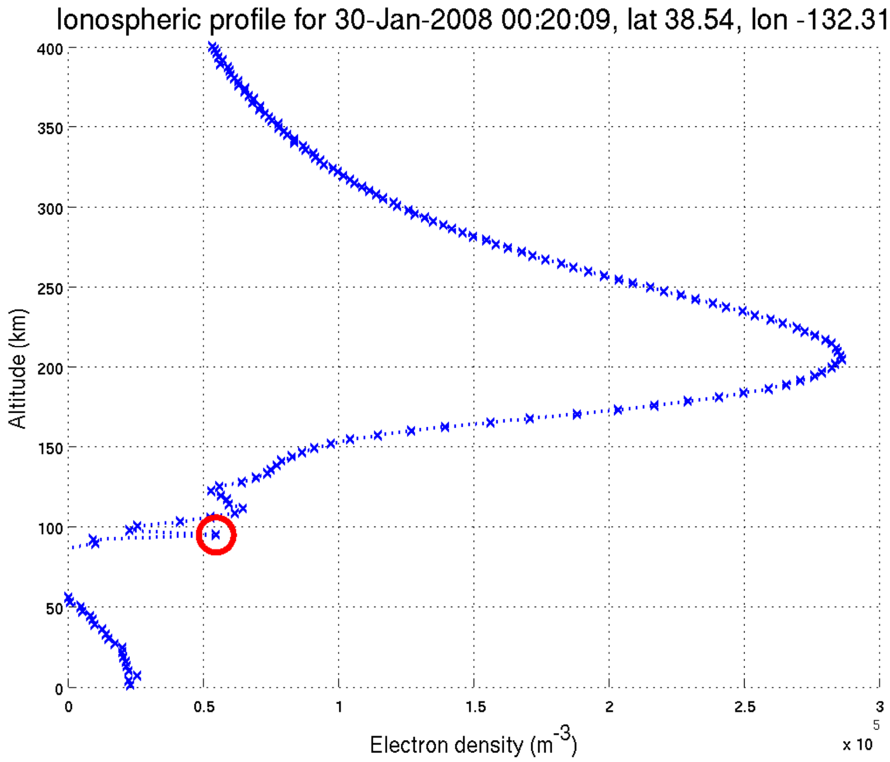


Fig. 2: An example ionospheric profile observed by a COSMIC satellite, showing a sporadic E layer with a peak electron density below that of the regular E layer.

from a radio occultation measurement [7].

The data used in this model came from three different satellite missions: CHAllenging Mini-satellite Payload (CHAMP) [16], Gravity Recovery and Climate Experiment (GRACE) [17] and Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) [18]. Most of the data (approximately 90%) came from COSMIC, which is also known as FORMOSAT-3. This is a collaborative project of the National Space Organization (NSPO) in Taiwan and the University Corporation for Atmospheric Research (UCAR) in the United States, to launch a constellation of satellites to provide radio-occultation soundings of the atmosphere and ionosphere, with the primary goal of demonstrating the value of near-real-time radio occultation observations in (tropospheric) numerical weather prediction [18]–[20].

The data from all 3 satellite missions were obtained from the COSMIC Data Analysis and Archival Center (CDAAC), hosted by UCAR. The CDAAC processes the data from the various missions into neutral atmospheric profiles and ionospheric products; both raw data and processed data products can be downloaded from CDAAC via their website at http://cdaac-www.cosmic.ucar.edu/cdaac/index.html.

Data Processing

For each radio occultation observation, the electron density profile data generated by CDAAC was processed to determine the presence of sporadic E, and if present, the height and peak density of the layer. A spatial high-pass filter is used to reject the regular E layer, which is much thicker than a sporadic E layer.

The observations were then separated into geographic cells and by time of day and day of the year. This allowed the measurement of the variation of sporadic E with geographic location, as well as its diurnal and seasonal variation. It is assumed that, although the sporadic E layer's behaviour varies seasonally, the layer's behaviour repeats with an annual cycle and thus aggregating data from different years (but from the same day of each year) is valid. The geographic cells were spaced 5 degrees apart at the equator and further apart towards the poles such that they had approximately equal area. The time resolution was 1 hour and the day resolution was 20 days. Measuring the solar cycle variation of sporadic E behaviour was considered but this was not done due to insufficient data.

For each geographic cell and time-of-day/day-of-year bin, two sets of statistics were generated: the median height of the sporadic E layer, and an inverse cumulative distribution function (CDF) of the critical frequency of the sporadic E layer, also known as f_oE_s, related to the peak electron density in the layer N by [1]

$$f_oE_s^2 = \frac{Ne^2}{\varepsilon_0 m}$$

where e is the elementary charge, m is the mass of the electron and ε_0 is the permittivity of free space. Note that this formula applies for the critical frequency of any ionospheric layer, not only sporadic E. The inverse CDF is the proportion of all f_oE_s observations with a value greater than a given plasma frequency; the model calculated the CDF for frequencies of 0 to 20 MHz. (It is labelled as an 'inverse' CDF as a standard CDF counts all the observations *below* a given value.) This allows the probability of a given radio signal being reflected from a sporadic E layer to be estimated directly.

Results

Figure 3 shows the height statistics generated for a particular grid cell and day of year. Both the median and mean height are shown; note that the median height is the quantity used by the model. Figure 4 shows the critical frequency statistics for the same grid cell and day of year. The corresponding (non-sporadic) E layer critical frequency (f_oE), from the International Reference Ionosphere [21], is also shown. Most of the f_oE_s measurements are very close to f_oE, indicating little or no sporadic E activity, however at certain times of the day, particularly around local noon (0340 UT), the distribution shows there is a small but significant probability of there being a more dense sporadic E layer. It can also be seen that there are sporadic E layers observed where $f_oE_s < f_oE$, (f_oE here referring to the value given by IRI, but in practice it is very close to the observed f_oE) where the sporadic E layer is at a different height to the peak of the regular E layer, as shown in Figure 2.

Figures 5 and 6 show the model's output over the entire Earth. For reference, both figures also show the solar terminator and three lines of magnetic latitude (the magnetic equator, 20 degrees north and 20 degrees south). Here the spatial variation of sporadic E can be seen, particularly in the probability distribution. There are three areas where sporadic E layers are likely to be seen: two areas at magnetic latitudes of approximately 30 degrees north and south, and a smaller area along the magnetic equator, all around local noon. These areas vary

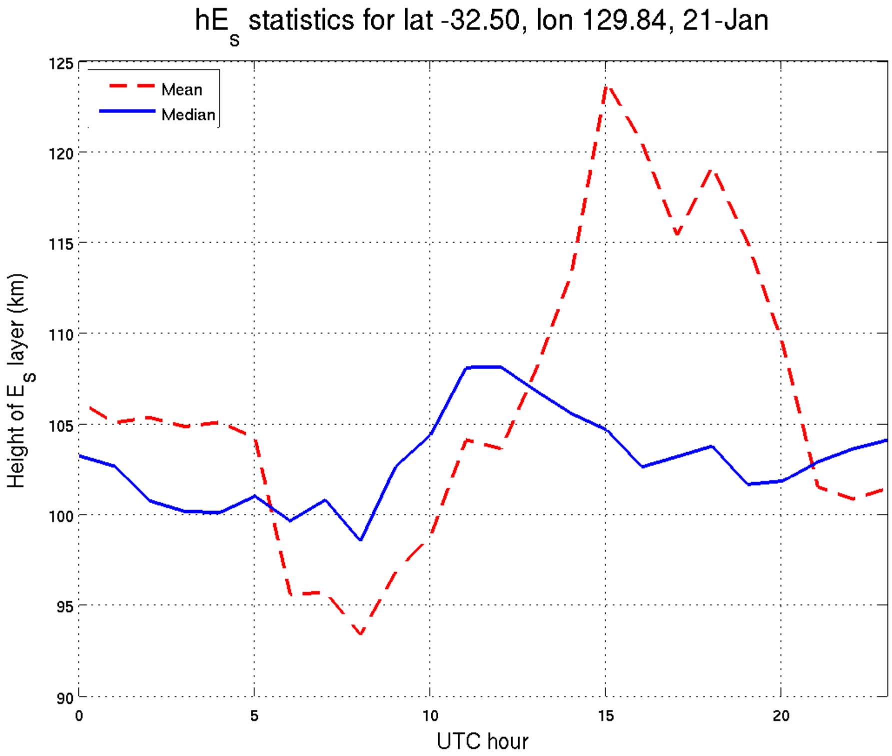


Fig. 3: Statistics on the height of the sporadic E layer observed at 32.50°S, 129.84°E, around January 21.

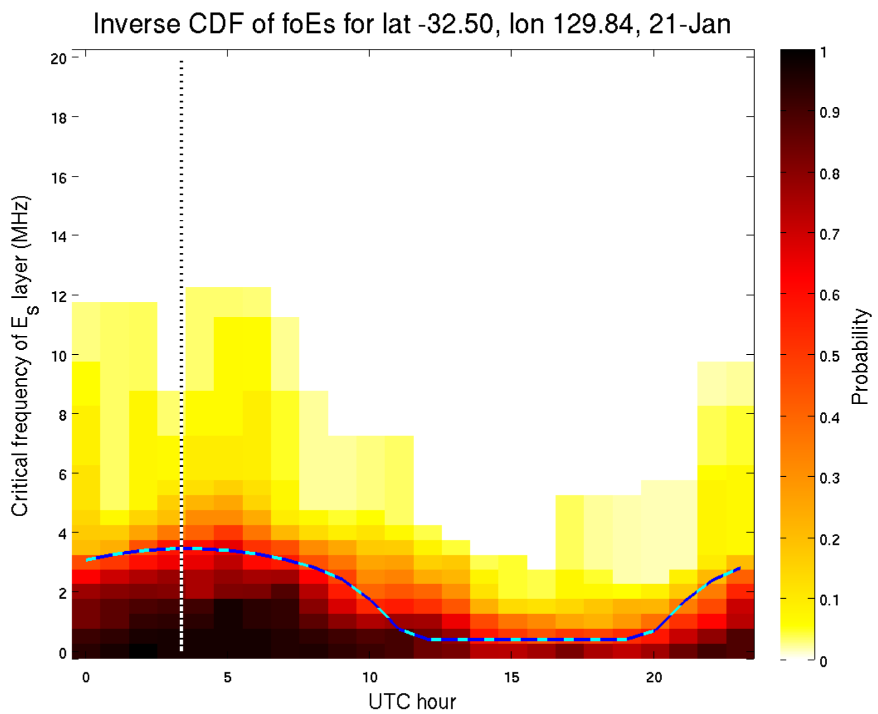


Fig. 4: Statistics on the critical frequency of the sporadic E layer observed at 32.50°S, 129.84°E, around January 21. The dashed line is the corresponding (non-sporadic) E layer critical frequency, from the International Reference Ionosphere. The vertical line corresponds to local noon.

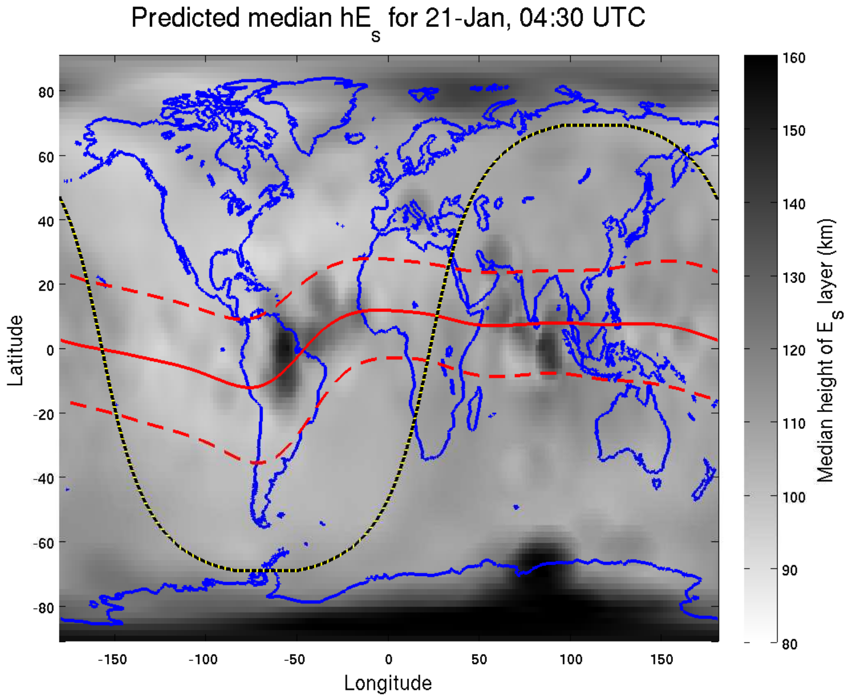


Fig. 5: Global sporadic E layer height, from the empirical model, for 0430 UT, January 21. The solar terminator, magnetic equator and magnetic latitudes 20 degrees north and south are also shown.

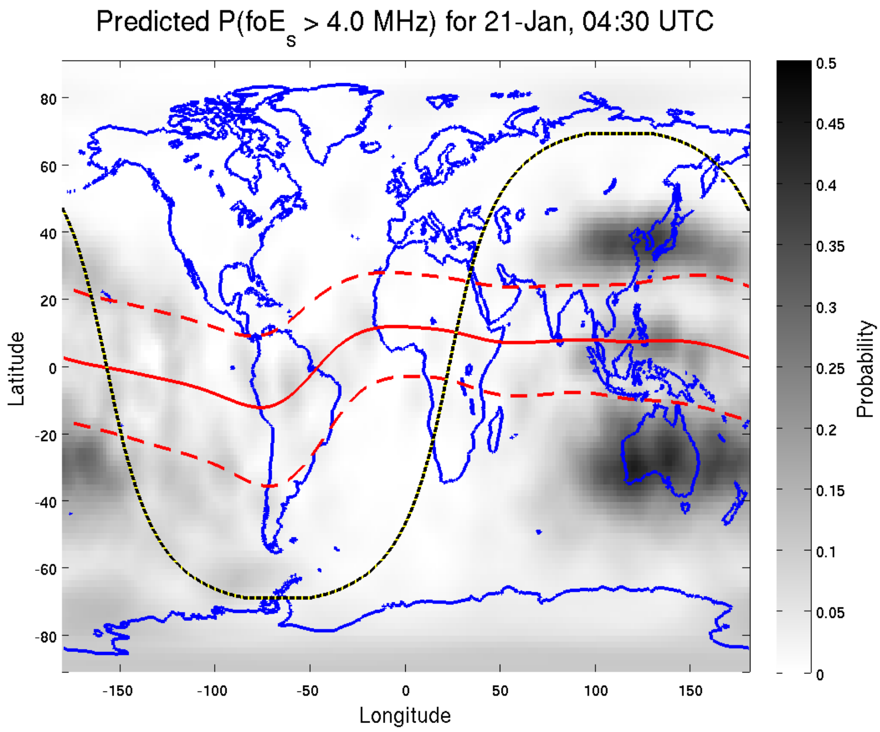


Fig. 6: Global sporadic E layer probability, from the empirical model, for 0430 UT, January 21. The solar terminator, magnetic equator and magnetic latitudes 20 degrees north and south are also shown.

somewhat with season; Figure 6 shows the result during southern hemisphere summer, with the southern hemisphere area showing a somewhat greater probability of sporadic E than the northern hemisphere area. This relation is reversed during the northern hemisphere summer, and is mostly caused by the variation in the underlying regular E layer which varies strongly with season.

Note that this spatial structure of the occurrence of sporadic E layers is similar to that observed by Arras et al [10], except that they did not detect the region of sporadic E along the magnetic equator. As they were using a different technique, looking at scintillation instead of the electron density profiles, this could be due to a number of different causes, such as a breakdown in the spherical symmetry assumed by the inversion technique, or a separate ionospheric effect (e.g. ionospheric disturbances) causing false signals in the electron density profiles. Further study, including data from other sensors, would be necessary to determine the validity of the sporadic E layers seen in the equatorial region.

Conclusion

An empirical, climatological model of sporadic E layers in the ionosphere has been developed, which can be used to predict the height of the layer and the probability of seeing a layer at any given location, time of day and day of year.

Further work is necessary in order to determine the sensitivity of GPS radio occultation to sporadic E layers, compared to other ionospheric instruments such as ionospheric sounders. This will allow the results from the model to be applied more accurately to radio propagation modelling.

References

[1] K. Davies, *Ionospheric Radio*. The Institution of Engineering and Technology, London, United Kingdom, 1990.

[2] J. Whitehead, "The formation of the sporadic-E layer in the temperate zones," *Journal of Atmospheric and Terrestrial Physics*, vol. 20, no. 1, pp. 49–58, 1961.

[3] ——, "Production and prediction of sporadic-E," *Reviews of Geophysics*, vol. 8, no. 1, pp. 65–144, 1970.

[4] ——, "Recent work on mid-latitude and equatorial sporadic-E," *Journal of Atmospheric and Terrestrial Physics*, vol. 51, no. 5, pp. 401–424, 1989.

[5] J. Mathews, "Sporadic E: current views and recent progress," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 60, no. 4, pp. 413–435, 1998.

[6] E. K. Smith, "Temperate zone sporadic-E maps (f_oE_s >7 MHz)," *Radio Science*, vol. 13, no. 3, pp. 571–575, 1978.

[7] G. A. Hajj, E. Kursinski, L. Romans, W. Bertiger, and S. Leroy, "A technical description of atmospheric sounding by GPS occultation," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 64, no. 4, pp. 451–469, 2002.

[8] Z. Zeng and S. Sokolovskiy, "Effect of sporadic E clouds on GPS radio occultation signals," *Geophysical Research Letters*, vol. 37, no. 18, 2010.

[9] W.-H. Yeh, C.-Y. Huang, T.-Y. Hsiao, T.-C. Chiu, C.-H. Lin, and Y.-A. Liou, "Amplitude morphology of GPS radio occultation data for sporadic-E layers," *Journal of Geophysical Research: Space Physics (1978–2012)*, vol. 117, no. A11, 2012.

[10] C. Arras, J. Wickert, G. Beyerle, S. Heise, T. Schmidt, and C. Jacobi, "A global climatology of ionospheric irregularities derived from GPS radio occultation," *Geophysical research letters*, vol. 35, no. 14, 2008.

[11] C. Arras, C. Jacobi, and J. Wickert, "Semidiurnal tidal signature in sporadic E occurrence rates derived from GPS radio occultation measurements at higher midlatitudes," in *Annales geophysicae*, vol. 27, no. 6. Copernicus GmbH, 2009, pp. 2555–2563.

[12] C. Arras, C. Jacobi, J. Wickert, S. Heise, and T. Schmidt, "Sporadic E signatures revealed from multi-satellite radio occultation measurements," *Advances in Radio Science*, vol. 8, no. 14, pp. 225–230, 2010.

[13] D. L. Wu, C. O. Ao, G. A. Hajj, M. de La Torre Juarez, and A. J. Mannucci, "Sporadic E morphology from GPS-CHAMP radio occultation," *Journal of Geophysical Research: Space Physics (1978–2012)*, vol. 110, no. A1, 2005.

[14] D. L. Wu, "Small-scale fluctuations and scintillations in high-resolution GPS/CHAMP SNR and phase data," *Journal of atmospheric and solar-terrestrial physics*, vol. 68, no. 9, pp. 999–1017, 2006.

[15] B. Carter, K. Zhang, R. Norman, V. Kumar, and S. Kumar, "On the occurrence of equatorial f-region irregularities during solar minimum using radio occultation measurements," *Journal of Geophysical Research: Space Physics*, vol. 118, no. 2, pp. 892–904, 2013.

[16] J. Wickert, C. Reigber, G. Beyerle, R. Knig, C. Marquardt, T. Schmidt, L. Grunwaldt, R. Galas, T. K. Meehan, W. G. Melbourne, and K. Hocke, "Atmosphere sounding by GPS radio occultation: First results from CHAMP," *Geophysical Research Letters*, vol. 28, no. 17, pp. 3263–3266, 2001. [Online]. Available: http://dx.doi.org/10.1029/2001GL013117

[17] B. D. Tapley, S. Bettadpur, M. Watkins, and C. Reigber, "The gravity recovery and climate experiment: Mission overview and early results," *Geophysical Research Letters*, vol. 31, no. 9, 2004.

[18] C. Rocken, K. Ying-Hwa, W. S. Schreiner, D. Hunt, S. Sokolovskiy, and C. McCormick, "COSMIC system description," *Terrestrial Atmospheric and Oceanic Sciences*, vol. 11, no. 1, pp. 21–52, 2000.

[19] R. A. Anthes, D. Ector, D. Hunt, Y. Kuo, C. Rocken, W. Schreiner, S. Sokolovskiy, S. Syndergaard, T. Wee, Z. Zeng *et al.*, "The COSMIC/FORMOSAT-3 mission: Early results," *Bulletin of the American Meteorological Society*, vol. 89, no. 3, pp. 313–333, 2008.

[20] J. Le Marshall, Y. Xiao, R. Norman, K. Zhang, A. Rea, L. Cucurull, R. Seecamp, P. Steinle, K. Puri, and T. Le, "The beneficial impact of radio occultation observations on australian region forecasts," *Aust. Meteorol. Oceanogr. J*, vol. 60, pp. 121–125, 2010.

[21] D. Bilitza, L. McKinnell, B. Reinisch, and T. Fuller-Rowell, "The International Reference Ionosphere today and in the future," *Journal of Geodesy*, vol. 85, no. 12, pp. 909–920, 2011.

Observations of a Mesospheric Bore over Edinburgh, Adelaide

Anne Unewisse, Andrew Cool, Manuel Cervera

Defence Science and Technology Organisation, PO Box 1500 Edinburgh SA, 5111

Summary:

A large mesospheric bore travelling Northwards over Edinburgh, South Australia, was imaged at 577.7 nm and 777.4 nm by the Defence Science and Technology Organisation's (DSTO) Thermospheric Radar Airglow Correlation Experiment (TRACE) on the 29th of March 2014 from 17 – 19 UT. Although several bores have been imaged by TRACE in the last few years, this particular event was notable for its 20% decrease in airglow intensity spanning the entire field of view from East to West and the resolution of its wavefronts. The background airglow intensity increased by nearly a factor of two in the six hours leading up to the bore, possibly associated with semi-diurnal tidal effects. There is evidence that the bore was detected by the light sensor of a locally located Boltwood Cloud Sensor, the first time such a measurement has been made.

Keywords: Airglow imaging, ionospheric studies, mesospheric bores

Introduction

Airglow is the visual or infrared emission seen in the Earth's atmosphere as a result of chemiluminescence. The emission wavelength is species dependent and commonly originates from oxygen, sodium or hydroxyl ions and molecules. The brightness of the airglow is proportional to the density of the species which makes it an ideal tool to explore atmospheric density variation.

Mesospheric bores are so named for their resemblance to tidal river bores. They appear as a sharp increase or decrease in airglow intensity spanning the entire field of view travelling at 20 – 100 ms^{-1} followed by either a series of phase-locked wave crests which appear as light and dark bands or a region of turbulence. The wavefronts may spread over time due to differing phase speed or slowly dissipate [1]. Light and dark bores often occur simultaneously at different altitudes.

Since the first published mesospheric bore images in 1995, imaged in the ALOHA-93 campaign [2], bores have been sighted in such diverse places as Brazil (1999) [3], the US (1999) [4], Argentina (2001) [5], Antarctica, (2001 [6] and 2007 [7]), Australia (2003) [8], Arecibo, Puerto Rico (2003) [9], India (2007) [10, 11] and China (2011) [12]. Occasionally they are visible to the naked-eye [4, 6, 12] and have been photographed [13] and videoed [14].

Models of mesospheric bores were generated almost immediately after the first sightings and continue to evolve as more observations are published [1, 15-18]. Terrestrial tidal bores occur when a shallow stable region of incompressible fluid bounded on three sides (a river bed and two banks) is impinged upon by tidal surge inducing a hydraulic jump. This travels down the river as a wave front followed by either a series of phase-locked wave crests or a region of turbulence, depending on the fractional increase in the river depth due to the bore [4]. Borrowing from this, it is postulated that mesospheric bores occur when a thin, bounded, stable region of the mesosphere, known as a 'duct', experiences a hydraulic jump which causes a series of waves to propagate down the duct. Lacking the 'river bed' region to confine the bottom side of the bore, waves are seen both top and bottom of the duct in opposite phases as an inverse pair.

The presence of a thermal duct is determined by a peak in the square of the Brunt-Väisälä frequency [12] , defined by

$$N^2 = \left(\frac{g}{T}\right)\left(\frac{dT}{dz}\frac{g}{c_p}\right) \tag{1}$$

where T is the temperature, c_p =1005 $Jkg^{-1}K^{-1}$ is the specific heat, dT/dz is the gradient of the temperature profile gradient and g=9.54 ms^{-2} is the acceleration due to gravity. An inversion in the temperature profile produces a peak in N^2 which defines the ducting region [4, 17]. Ducts can persist for hours and span up to 1200 km [4].

Mesospheric airglow tends to originate from narrow layers. If those layers lie close enough to the ducting region they will be affected by any bore waves propagating along them and the resulting density perturbations will appear as moving light and dark bands. Airglow originating from layers lying above the duct will be pushed upwards to less dense regions resulting in a dark bore with trailing light and dark waves. Conversely, airglow layers lying below the duct are pushed downwards into more dense regions forming bright bores. When two closely spaced airglow regions lie either side of a duct, simultaneous dark and light bores are seen [4, 6, 8, 15-19].

The actual cause of the hydraulic jump is much less well understood. In some cases, it is thought to be caused by weather fronts [4, 20, 21] and there is evidence that tropospheric lightning, known to generate concentric waves in airglow [22, 23], may play a role. Unlike auroral activity, there is no association of bores with Kp index [4]. Bores are sometimes seen during periods of large-scale gravity wave activity, semi-diurnal tides [3, 5, 24] and two day waves [19].

Bores are often preceded by several 'ripples'[3]. Reported as naked-eye phenomena as early as 1979 [25], the ripples are short-wavelength (15-20 km) transient 'wave packets' which appear suddenly in airglow images and typically last less than an hour. Ripples traverse the airglow image with group velocity directions that are unrelated to their phase velocity direction, sometimes even perpendicular [26]. Ripples are thought to arise from local convective or dynamic instabilities [11, 20] mainly due to their short wavelengths and lives.

This paper presents measurements and analysis of a significant bore observed over Edinburgh, Adelaide, Australia on the 29th of March 2014 by the Thermospheric Radar Airglow Correlation Experiment (TRACE).

The Edinburgh bore event

The Airglow Imager

The Defence Science and Technology Organisation (DSTO) has operated the Thermospheric Radar Airglow Correlation Experiment (TRACE) airglow imager from its Edinburgh, SA site since January 2012 in support of atmospheric studies at altitudes of interest to high frequency radar: specifically the ionospheric E and F layers.

TRACE consists of a Keo Sentry [27] camera with 180 degree field of view (FOV), 24mm / F4.0, achromatic Mamiya fisheye lens, a temperature controlled filter wheel and a 16 bit Princeton Instruments Acton back illuminated Pixis 1024B CCD camera with a pixel size of 13.3μm x 13.3μm. Details on the imager may be found in Unewisse et. al. [26]. The imager's FOV corresponds to image diameters of ~400 km at an altitude of 96 km [28, 29] and ~1000 km at an altitude of 250 km [30].

The imager has five available 2 nm wide filters centred on: 557.7 nm O(1S) oxygen emission originating from an altitude of 96 km with a small contribution from 250 km; 630.0 nm O(1D) oxygen emission (250 km); 777.4 nm O(5P) oxygen emission (350 km); 589.3 nm sodium doublet emission (92 km); and 572.4 nm used to measure local non-airglow conditions (the control image). In an urban environment, the sodium filter is generally too affected by light pollution to be scientifically useful. Images are obtained in a continuous cycle through selected filters with each exposure being typically 2 minutes.

Some of the filters image nearby lines which fall within the filter bandwidth [31]. In particular, the 777.4 nm filter designed to measure faint high altitude oxygen airglow will also include emission from two OH Meinel airglow features from 87 km: a Q1(3) doublet at 777.3 nm, and a P2(1) line emission at 778.4 nm. At times, this OH line dominates the images [5]. The 777.4 nm images are also sensitive to galactic plane emission and the back illuminated chip, while very sensitive, suffers from a fringing interference pattern known as etaloning which may be mitigated during image processing [32].

The TRACE is supported by two roof mounted environmental sensors: the Unihedron Sky Quality Meter (SQM) [33] which records calibrated light readings every minute and the Boltwood Cloud Sensor Mk II (BCS) [34, 35]which monitors weather and light conditions every 2.06 seconds.

Image Processing

The image processing method for this work was based on that detailed by the Mid-latitude All-sky-imager Network for Geophysical Observation (MANGO) project [36] and utilised routines in the Interactive Data Language (IDL) [37].

Firstly, a dark image is subtracted from the image. In this paper, the dark is the pixel by pixel median of a large number of closed shutter exposures taken over many nights with the same exposure time and temperature.

Secondly, the image is corrected for line-of-sight and van Rhijn effects [38] which produce opposite, radially proportional effects on the image brightness. Line-of-sight effects result in attenuation of signals received at the imager from a zenith angle θ (radians) due to different atmospheric thickness along the path. The received emission, I(θ), relative to the original emission, $I(\theta)_0$, is given by [39]

$$I(\theta) = I(\theta)_0 10^{-0.4aF(\theta)} \tag{2}$$

Where a is the atmospheric extinction co-efficient (assumed 0.2) and F(θ) is

$$F(\theta) = [\cos\theta + 1.5\,(93.885 - \theta)^{-1.253}]^{-1} \tag{3}$$

The van Rhijn effect results in the enhancement of airglow intensity received at the imager from a zenith angle θ due to path length through the airglow emission layer. It is defined by:

$$I(\theta) = \left[1 - \left(\frac{R_e}{R_e+h}\right)\sin^2\theta\right]^{-\frac{1}{2}} \tag{4}$$

where R_e is the Earth's radius in km and h is the airglow altitude in km [36, 39]. These two effects are modelled as a single mask and divided into each image.

Next, a star removal algorithm based on the publically available Airglowrsss IDL library [40] is used to reduce the impact of star trails.

The image is then converted from fisheye pixel units to linear km (unwarping). The unwarping method outlined in many sources [36, 39, 41] is based on a seminal paper [42] from the Weapons Research Laboratory (WRL), a precursor to the current Defence Science and Technology Organisation (DSTO). A lens function is created by identifying stars on a sample image and fitting a third order polynomial to a plot of radial distance from the image centre in pixels vs tabulated zenith angle at the image time. The zenith angle (θ) at the airglow imager is then converted to km using its relationship to the the Earth centered angle ϕ (radians):

$$\phi = \theta - \sin^{-1}\theta\left(\frac{R_e \sin\phi}{R_e+h}\right) \tag{5}$$

The distance from the image centre in km at the airglow height r_a is then approximately

$$r_a = (R_e + h)\phi \tag{6}$$

Using a method suggested by Theusner [43], a grid of 128 x 128 evenly spaced sample points on the image are transformed by the lens function and equation 6. The sample points and their resultant transformed co-ordinates are used in the IDL WARP_TRI function to warp all pixles in the image to km.

The final step is to rotate the image to conventional co-ordinates with North at the top and East to the left.

When higher contrast images are required, the airglow-free 'control' image is subtracted from the image to reduce background emission. Etaloning in the 777.4 nm images is minimized before co-ordinate conversion by creating a mask using the normalised median of a series of processed 777.4 nm images and then dividing the image by this mask. Other methods are currently being trialled.

Measurements

Figure 1 shows the progress of the bore from approximately 17-19 UT on the 29th of March 2014 in 557.7 nm airglow. Imaging halted automatically due to the impending dawn just after 19 UT. The images have 2 minute exposures and are approximately 8 minutes apart. At this altitude each image has a diameter of 400 km. Control image subtraction has been used to enhance contrast. The waves spread over time indicating wave ordering.

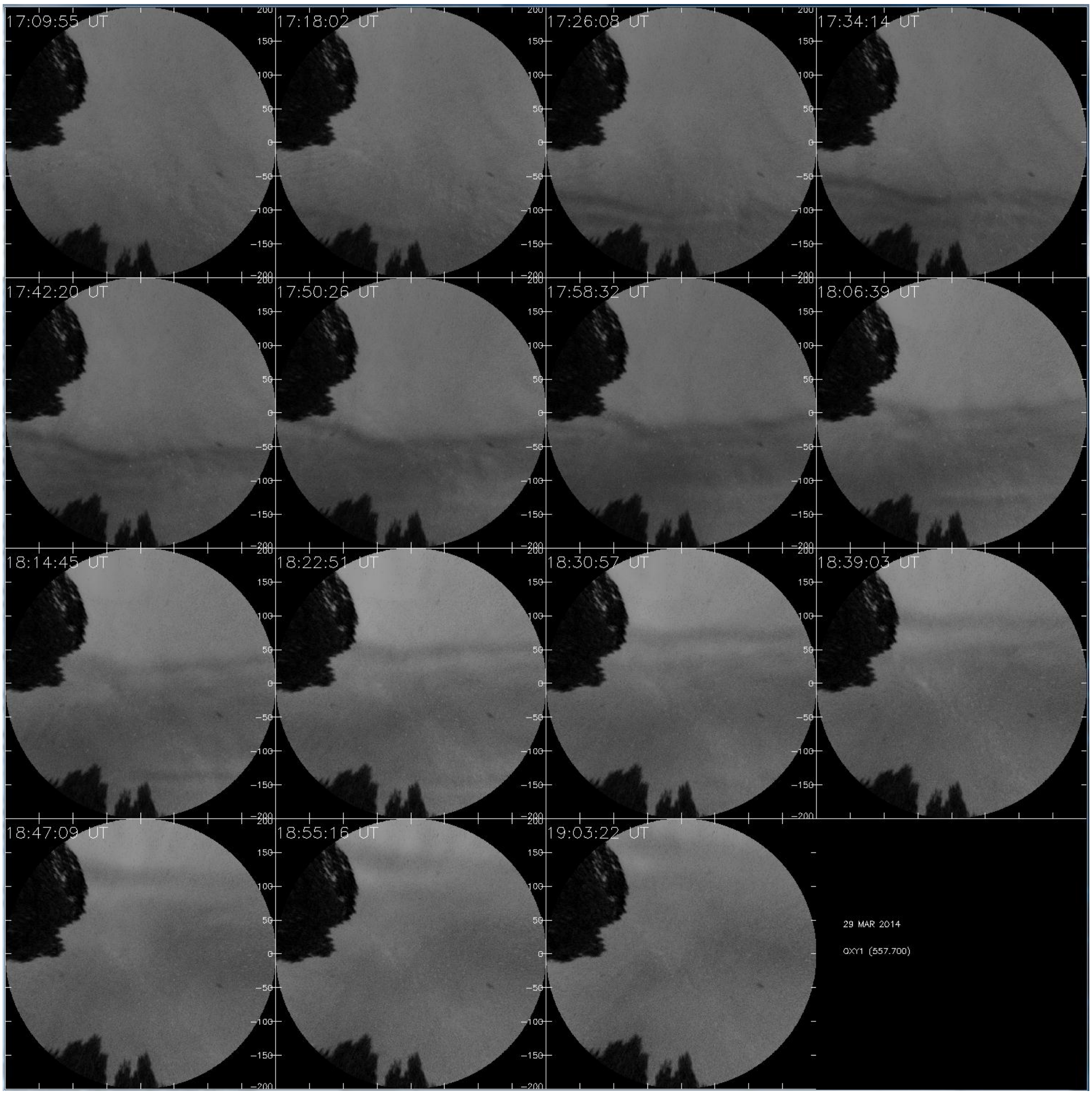


Figure 1 Mesospheric bore of the 29 March 2014 from 17-19UT.Each frame is 400 km x 400 km, North is up and East is to the left.

Smoothed North-South slices through 557.7 nm images of the bore without control subtraction are shown in Figure 2. A decrease in airglow intensity in arbitrary units from 500 to 400 is seen as the bore transits. The bore front has an average speed 46 ms^{-1} whereas the first trailing wave has a speed of ~34 ms^{-1}. The wavelength of the leading first wave varies from 12.5 km at 18:06 UT to 49 km at 18:47 UT. The wave period is approximately 20 minutes though the data is under sampled at a revisit rate of 8 minutes.

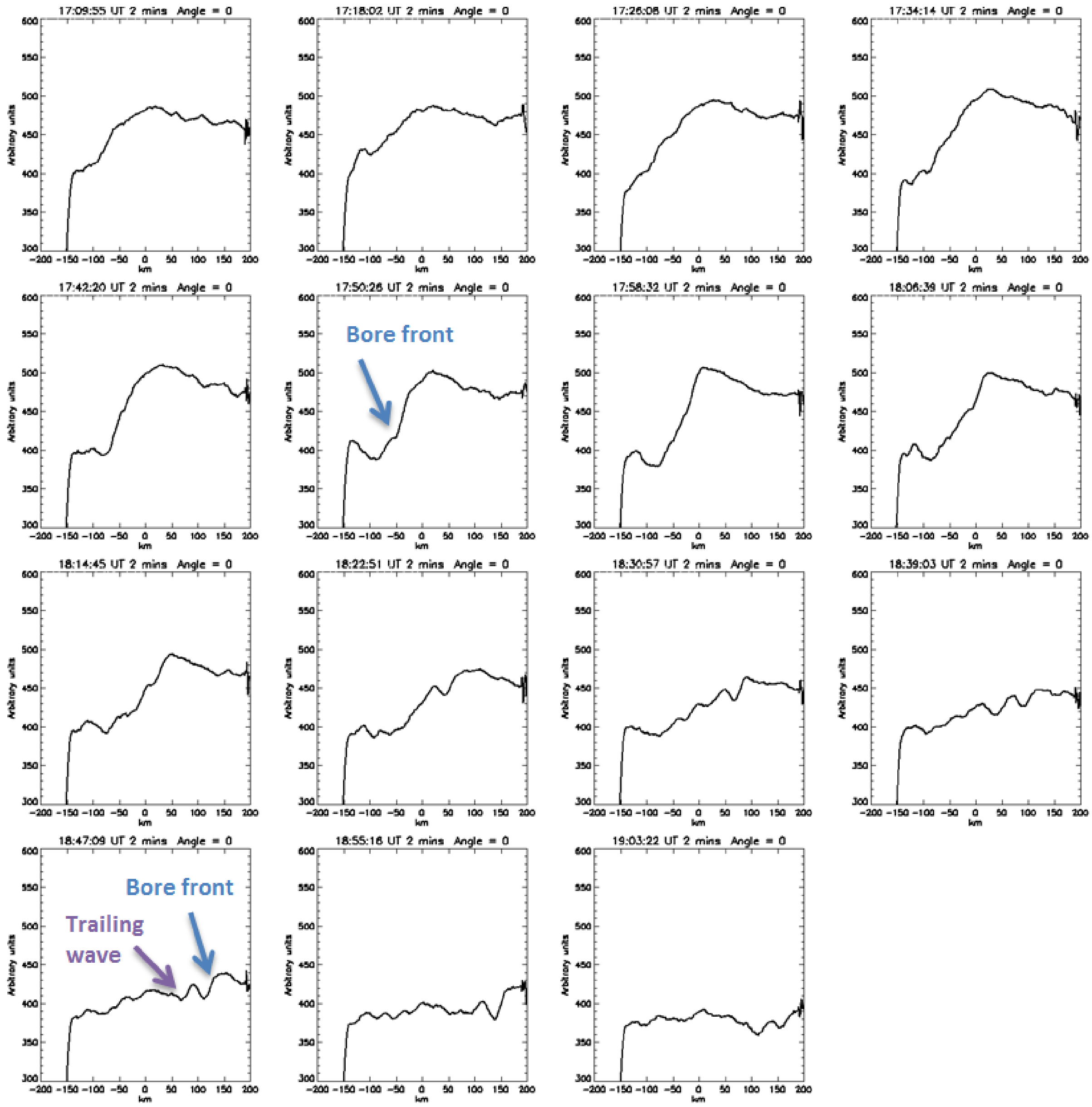


Figure 2 North-South slices through the bore images over time. The bore and first trailing wave appear as negative troughs.

The 777.4 nm airglow image revealed a faint inverse bore from contaminating OH Meinel bands in the filter bandwidth originating from 87 km. Figure 3 compares North-South slices of the last three OH images, where galactic plane interference is less, with the last three oxygen 557.7 nm images from Figure 2. Note the images are 4 minutes apart. The correspondence of OH peaks and 577.7 nm troughs can clearly be seen.

The bore event was preceded by a series of instability ‘ripples’ over 2 hours, one of which was visible in both the 557.7 nm and OH Meinel images simultaneously.

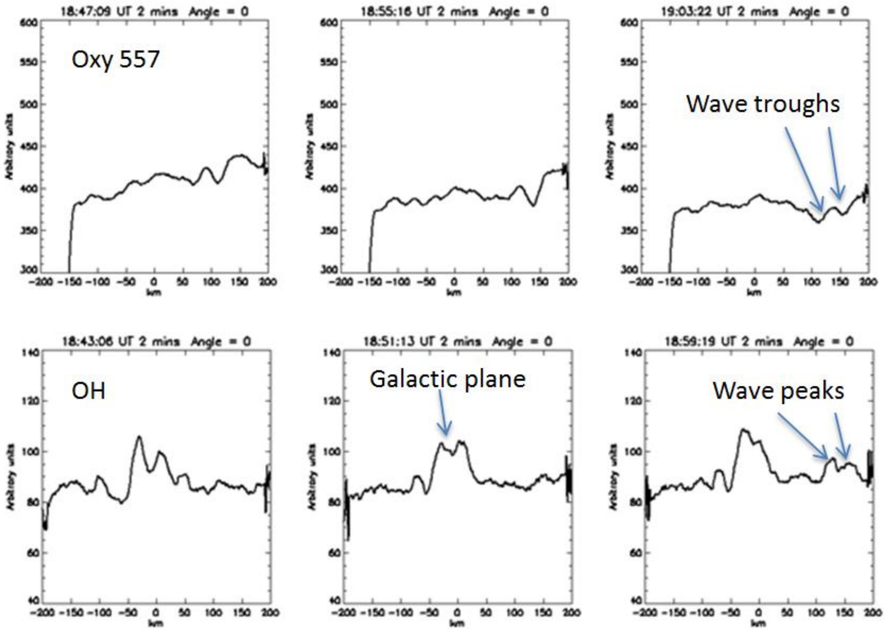


Figure 3 North South slices through the last three Oxygen 557.7 nm and OH Meinel images.

Discussion

Ducting

TIMED/SABER satellite [44, 45] Level 2A kinetic temperature profiles are shown in Figure 4a. Unfortunately, the geographically closest data was taken 4.5 hours before (12.45 UT and 12.47 UT) and 2 hours after (21.61 UT and 21.63 UT) the observed bore, the latter being after local sunrise (20.5 UT). The green line is the altitude of the oxygen 557.7 nm emissions (96 km) and the blue line is the altitude of the OH emission (87 km). There is a 30 K temperature inversion between 88 and 92 km in the 12.47 UT profile.

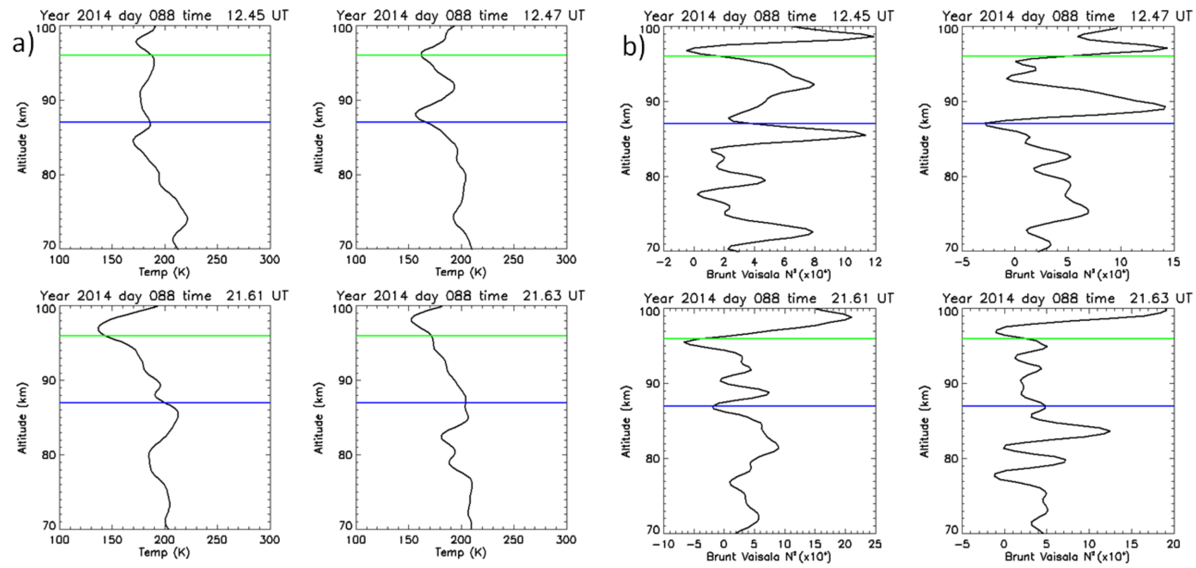


Figure 4 a) The TIMED/SABER temperature profiles for events geographically close to the observed bore and b) the corresponding Brunt-Vaisala squared. The altitudes of the 557.7 nm (green) and 777.3 nm (OH) airglow emission is also shown.

The corresponding Brunt-Väisälä squared profiles are shown Figure 4b. A duct is clearly seen between the two airglow layers in the hours leading up to the bore but not in the hours after the bore. The FWHM of the ducting regions is around 4 km.

SABER data is not able to confirm the presence of ducting at the time of the bore but it does indicate that there was one present between the airglow layers 4.5 hours beforehand. As ducts can persist for hours [4], it is likely that the duct was still present when the bore was observed.

Median airglow intensity.

Figure 5 shows the median of the central 512 x 512 image pixels for 557.7 nm (oxygen, 96 km), 572.4 nm (background) and 777.4 nm (oxygen, 350 km and OH, 87 km) airglow images. The figure clearly illustrates the problems associated with subtracting the 'background' control data from weak 777.4 nm data.

The 557.7 nm airglow intensity rose from 260 to 450 arbitrary units in the 6 hours leading up to the bore, starting from the appearance of ripples, and then dropped off as the bore passed. The OH airglow showed a corresponding but smaller decrease in intensity during this time which rose as the bore passed. Bores have previously been associated with large rises in the background airglow intensity [3, 5, 8, 19, 24] .

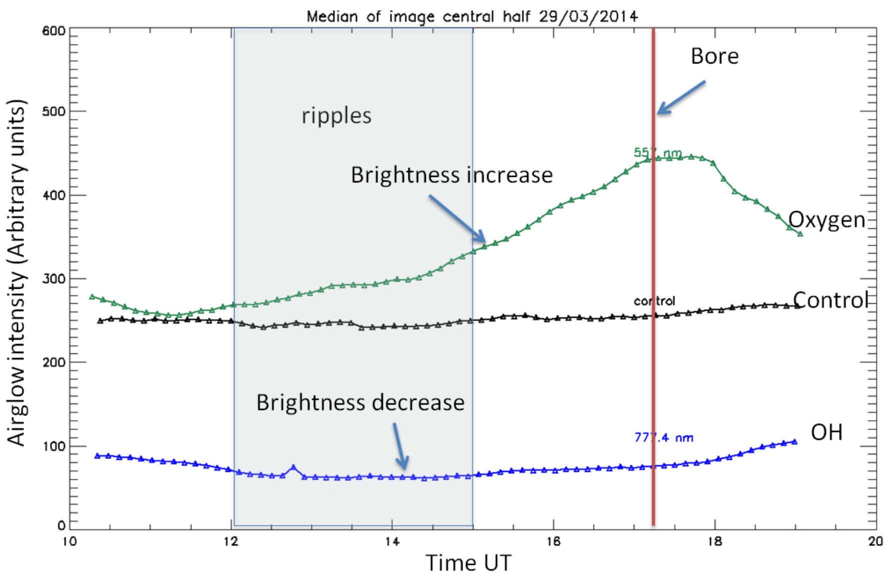


Figure 5 The median airglow intensity over the central 512x512 pixels of the image for three filters.

Light sensor bore detection

The Boltwood Cloud Sensor (BCS) daylight sensor [35] is an Osram SFH 213 photodiode. It has a 20 degree field of view and spectral range of 100 – 1100 nm (covering both IR and visual radiation) with a peak sensitivity at 850 nm [46]. The sensor records local weather and light conditions every 2.06 seconds. For practical reasons, the BCS is mounted over the container roof rather than over ground as recommended which does not significantly impact on its use in TRACE's context.

Unless there is strong artificial light or the sky is obscured by thick cloud, the average overnight light sensor reading at this site is 19 ± 9 arbitrary units (1σ). Figure 6 is a plot of smoothed light sensor output on the night of the bore showing that the light level rose to around 100 prior to the bore and then oscillated with a period of approximately 28 minutes as it passed. As stars were clearly visible in all images, it is assumed that this is not due to cloud cover. A similar but smaller rise in light reading was seen in a fainter bore on the 19th of August 2014 suggesting that this may be a method of bore detection.

The daylight sensor spectral range covers both visual and IR wavelengths. The BCS incorporates two 8 – 14 μm infrared temperature sensors, one ground facing and one sky facing. The difference between the two measurements is used to determine cloud cover. The ground temperature varies slowly but the sky temperature is sensitive to small changes. The overlayed arbitrarily scaled sky IR temperature in Figure 6 shows that at least some of the light readings are related to the variation in sky temperature.

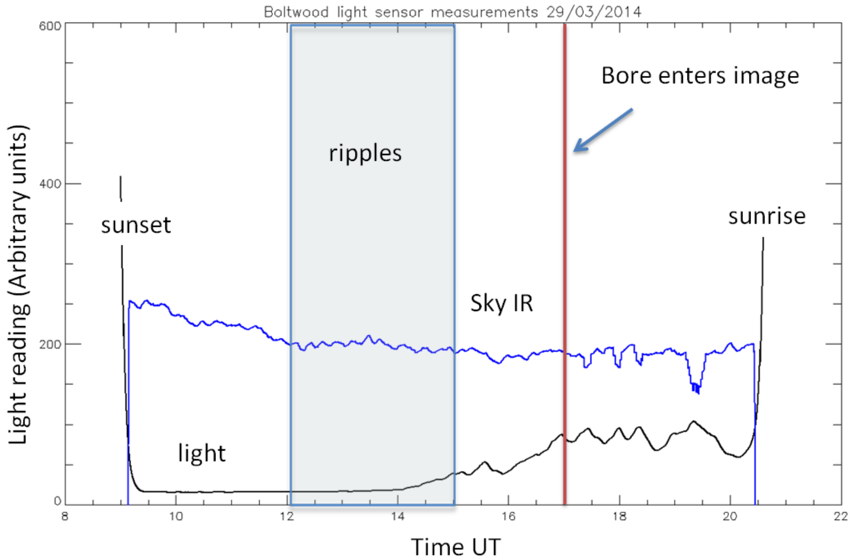


Figure 6 a) Light readings (black) and scaled SKY IR temperature (blue) from the BCS with ripple and bore times marked.

Although visual range airglow measurements were obtained at 557.7 nm, 630 nm, 777.4 nm and 572.4 nm, the 8 minute revisit time meant they were not sampled sufficiently to confirm the oscillations seen in the light measurements. There is a gradual increase in the 557.7 nm median intensity but it does not directly correlate with the increase in light measurements.

The Sky Quality Meter (SQM), designed to report calibrated light readings, experienced technical difficulties around 12 UT, just prior to the onset of the increase in light readings from the cloud sensor.

Conclusion

The Thermospheric Radar Airglow Correlation Experiment (TRACE) detected a large bore on the 29th of March 2014 over Edinburgh, South Australia. The bore exhibited wave ordering with the leading front significantly faster than trailing waves. In the hours leading up to the bore, the airglow intensity increased by almost a factor of 2 and there was evidence of local convective or dynamic instabilities. The bore was detected as a dark bore in 557.7 nm Oxygen from 96 km and a faint light bore in 87 km OH airglow. There is evidence that the bore was also detected by the light sensor of the onsite Boltwood Cloud Sensor and its sky IR sensor.

SABER data confirms the presence of ducting between the airglow layers 4.5 hours before the bore. As ducts can persist for hours, it is likely that the duct was still present when the bore was observed.

Acknowledgments

The conversion method for this work was influenced by code supplied by Michael Theusner. Thanks to Doug George on the Diffraction Limited forum for information on the BCS light sensor. We greatfully acknowledge the SABER Team for producing the temperature profiles used in the calculations.

References

1. R. G. Stockwell, *et al*., "The evolution of a breaking mesospheric bore wave packet," *Journal of Geophysical Research: Atmospheres,* vol. 116, p. D19102, 2011, doi:10.1029/2010JD015321.
2. M. J. Taylor, D. N. Turnbull, and R. P. Lowe, "Spectrometric and imaging measurements of a spectacular gravity wave event observed during the ALOHA-93 Campaign," *Geophysical Research Letters,* vol. 22, pp. 2849-2852, 1995, doi:10.1029/95GL02948.
3. A. F. Medeiros, *et al*., "An unusual airglow wave event observed at Cachoeira Paulista 23° S," *Advances in Space Research,* vol. 27, pp. 1749-1754, // 2001, doi:10.1016/S0273-1177(01)00317-9.
4. S. M. Smith, *et al*., "A multidiagnostic investigation of the mesospheric bore phenomenon," *Journal of Geophysical Research: Space Physics,* vol. 108, p. 1083, 2003, doi:10.1029/2002JA009500.
5. S. M. Smith, *et al*., "Characterization of exceptionally strong mesospheric wave events using all-sky and zenith airglow observations," *Journal of Geophysical Research: Space Physics,* vol. 111, p. A09309, 2006, doi:10.1029/2005JA011197.
6. K. Nielsen, *et al*., "An unusual mesospheric bore event observed at high latitudes over Antarctica," *Geophysical Research Letters,* vol. 33, p. L07803, 2006, doi:10.1029/2005GL025649.
7. J. V. Bageston, *et al*., "Case study of a mesospheric wall event over Ferraz station, Antarctica (62° S)," *Annales Geophysicae,* vol. 29, pp. 209-219, 01 / 01 / 2011, doi:10.5194/angeo-29-209-2011.
8. R. L. Walterscheid, *et al*., "An intense traveling airglow front in the upper mesosphere–lower thermosphere with characteristics of a bore observed over Alice Springs, Australia, during a strong 2 day wave episode," *Journal of Geophysical Research: Atmospheres,* vol. 117, p. D22105, 2012, doi:10.1029/2012JD017847.
9. S. M. Smith, *et al*., "Evidence of mesospheric bore formation from a breaking gravity wave event: simultaneous imaging and lidar measurements," *Journal of Atmospheric and Solar-Terrestrial Physics,* vol. 67, pp. 345-356, 3// 2005, doi:http://dx.doi.org/10.1016/j.jastp.2004.11.008.
10. V. Lakshmi Narayanan, S. Gurubaran, and K. Emperumal, "Nightglow imaging of different types of events, including a mesospheric bore observed on the night of February

15, 2007 from Tirunelveli (8.7°N)," *Journal of Atmospheric and Solar-Terrestrial Physics,* vol. 78–79, pp. 70-83, 4// 2012, doi:http://dx.doi.org/10.1016/j.jastp.2011.07.006.
11. V. L. Narayanan, S. Gurubaran, and K. Emperumal, "Imaging observations of upper mesospheric nightglow emissions from Tirunelveli (8.7oN)," *Indian J. Radio Space,* vol. 38, pp. 150-158, 2009.
12. Q. Li*, et al.*, "Investigation of a mesospheric bore event over northern China," *Annales Geophysicae (09927689),* vol. 31, pp. 409-418, 2013, doi:10.5194/angeo-31-409-2013.
13. T. Ashcroft and W. Lyons. (2014). *Photograph of a mesospheric bore*. Available: http://rammb.cira.colostate.edu/projects/npp/blog/index.php/uncategorized/severe-weather-in-the-mesosphere/attachment/mesospheric_bore/
14. A. Cherney. (2014). *Movie of bore over Heathcote, Victoria, Australia taken with Nikon D700*. Available: http://vimeo.com/52319627
15. E. M. Dewan and R. H. Picard, "Mesospheric bores," *Journal of Geophysical Research: Atmospheres (1984–2012),* vol. 103, pp. 6295-6305, 1998, doi:10.1029/97JD02498.
16. E. M. Dewan and R. H. Picard, "On the origin of mesospheric bores," *Journal of Geophysical Research: Atmospheres,* vol. 106, pp. 2921-2927, 2001, doi:10.1029/2000JD900697.
17. P. J. Loughmiller, "Observational and theorectocal studies of mesospheric dynamics and bores using airglow imagaers at mid-latitudes," Doctor of Philosophy Doctoral, Cornell University, 2008.
18. P. J. Loughmiller, M. C. Kelley, and M. P. Hickey, "Observational and modelling study of mesospheric bores," in *36th COSPAR Scientific Assembly*, Beijing, China, 2006, p. 1454.
19. J. H. Hecht and R. L. Walterscheid, "Observations of the OH Meinel (6,2) and O2 atmospheric (0,1) Nightglow emissions from Maui during the ALOHA-90 Campaign," *Geophysical Research Letters,* vol. 18, pp. 1341-1344, 1991, doi:10.1029/91GL01152.
20. M. J. Taylor and R. Edwards, "Observations of short period mesospheric wave patterns: In situ or tropospheric wave generation?," *Geophysical Research Letters,* vol. 18, pp. 1337-1340, 1991, doi:10.1029/91GL01516.
21. CIMSS. (2014, 22 May 2014). *Convectiviely-generated mesospheric airglow waves over Texas*. Available: http://cimss.ssec.wisc.edu/goes/blog/archives/15299
22. S. D. Miller*, et al.*, "Suomi satellite brings to light a unique frontier of nighttime environmental sensing capabilities," *Proceedings of the National Academy of Sciences,* vol. 109, pp. 15706-15711, September 25, 2012 2012, doi:10.1073/pnas.1207034109.
23. C. Seaman. (2014, 12 June). *Severe Weather in the Mesosphere*. Available: http://rammb.cira.colostate.edu/projects/npp/blog/index.php/tag/daynight-band/
24. J. H. Hecht*, et al.*, "Observations of variations in airglow emissions during ALOHA-93," *Geophysical Research Letters,* vol. 22, pp. 2817-2820, 1995, doi:10.1029/95GL03019.
25. A. W. Peterson, "Airglow events visible to the naked eye," *Applied Optics,* vol. 18, pp. 3390-3393, 1979/10/15 1979, doi:10.1364/AO.18.003390.
26. A. Unewisse*, et al.*, "TRACE: A New Relocatable Airglow Imager," in *Proceedings of the 13th Australian Space Science Conference* Sydney. 30 September - October 2, 2013, p. 149.
27. T. S. Trondsen. (2012, March 29). *Low Light Level Imaging Instrumentation for Airglow - Keo Scientific*. Available: http://www.keoscientific.com/aeronomy-imagers.php
28. S. Suzuki*, et al.*, "Statistical characteristics of gravity waves observed by an all-sky imager at Darwin, Australia," *J. Geophys. Res.-Atmos,* vol. 109, p. D20S07, 2004, doi:10.1029/2003jd004336.
29. K. Shiokawa, Y. Otsuka, and T. Ogawa, "Propagation characteristics of nighttime mesospheric and thermospheric waves observed by optical mesosphere thermosphere imagers at middle and low latitudes," *Earth Planets Space,* vol. 61, pp. 479-491, 2009.

30. T. Ogawa, *et al.*, "Simultaneous ground- and satellite-based airglow observations of geomagnetic conjugate plasma bubbles in the equatorial anomaly," *Earth Planets Space,* vol. 57, pp. 385-392, 2005.
31. D. E. Osterbrock, *et al.*, "Night-sky high-resolution spectral atlas of OH and O2 emission lines for echelle spectrograph wavelength calibration," *Publications of the Astronomical Society of the Pacific,* vol. 108, pp. 277-308, 1996.
32. M. Schirmer. (2010). *About superflatting and defringing*. Available: http://www.astro.uni-bonn.de/theli/gui/aboutsuperflatting.html
33. Unihedron, "Sky Quality Meter -LU Users manual," Optcorp, Ed., 1.1 ed, 2006.
34. J. Marchant, R. J. Smith, and I. A. Steele, "Calibration of the Boltwood Cloud Sensor," presented at the Proc. SPIE 7012, Ground-based and Airborne Telescopes II, 2008, doi:10.1117/12.789072.
35. DiffractionLimited, "Cloud Sensor II User's Manual," Boltwood Systems Corporation, V0026 ed, 2009.
36. F. F. Kenge, R. Mundra, and M. A. Rangel, "Mid-latitude All-sky-imager Network for Geophysical Observation," B.Sc. Major Qualifying Project, Faculty of Worcester Polytechnic Institute, 2013.
37. Exelis. (2014). *IDL - Interactive Data language*. Available: http://www.exelisvis.com/docs/using_idl_home.html
38. P. J. Van Rhijn, "On the brightness of the sky at night and the total amount of starlight," *Astrophys. J.,* vol. 50, 12/1919 1919, doi:10.1086/142513.
39. M. Kubota, H. Fukunishi, and S. Okano, "Characteristics of medium and large-scale TIDs over Japan derived from OI 630-nm nightglow observation," *Earth Planets Space,* vol. 53, pp. 741-751, 2001.
40. J. J. Makela, "Airglowrsss IDL library ", 9 June 2005 ed. University of Illinois, 2005, p. Google Airglowrss IDL code (http://code.google.com/p/airglowrsss/source/browse/trunk/IDL/Imaging/removestars.pro?spec=svn79&r=79).
41. F. J. Garcia, M. J. Taylor, and M. C. Kelley, "Two-dimensional spectral analysis of mesospheric airglow image data," *Appl. Optics,* vol. 36, pp. 7374-7385, 1997.
42. K. H. Lloyd, "Concise method for photogrammetry of objects in the sky," Weapons Research Establishment, Australia WRE-TECHNICAL NOTE-72 (WRB&D), 1971.
43. M. Theusner, "Image dewarping code," 23 June 2013 ed: private comunication, 2013, p. IDL code.
44. J. M. Russell III, *et al.*, "Overview of the SABER experiment and preliminary calibration results ", 1999, pp. 277-288.
45. NASA, HamptonUniversity, and GATS. (2014). *TIMED/SABER Browse Data Page*. Available: http://saber.gats-inc.com/browse_data.php
46. Osram, "Silicon PIN Photodiode: SFH 213, SFH 213 FA," 1.1 ed: OSRAM, 2014.

The Spectrum Difference Function technique for velocity cross-checking in TIGER application

Thanh Nguyen*, John Devlin*, Darel Elton*, Guang Deng*, Edhem Custovic* and Emma Bland †

* *Department of Electronic Engineering, LaTrobe University, Bundoora, Victoria, AUSTRALIA. E-mail: tt72nguyen@students.latrobe.edu.au*
† *Department of Physics, LaTrobe University, Bundoora, Victoria, AUSTRALIA*

Summary: Super Dual Auroral Radar Network (SuperDARN) is an international HF radar network employed to explore the impact of solar disturbances on Earth by monitoring the location of aurora and related phenomena occurring in the ionosphere. These radars utilise an Auto-Correlation Function (ACF) to estimate the target velocity. The ACF technique, however, has shown large and occasionally unrealistic errors. An independent cross-checking technique, therefore, is desirable for evaluating the velocity estimated by the ACF technique.

This paper proposes a novel Spectrum Difference Function (SDF) technique for the purpose of cross-checking the ACF results. In the proposed technique, the difference between the magnitude spectra of the transmitted and the received signals is defined as the SDF and employed to deduce target velocity. Mathematical derivation shows that the gradient of the SDF in the vicinity of the carrier frequency is directly proportional to the target Doppler frequency. Therefore, by calculating the gradient of the SDF, the Doppler frequency and consequently target velocity can be deduced.

Keywords: SuperDARN, Radar, Doppler Frequency Estimation, FFT

I. Introduction

Super Dual Auroral Radar Network (SuperDARN) is an international network of over 30 high frequency (HF) coherent scatter radars designed primarily for observing plasma convection in the high-latitude ionosphere, as well as a range of other ionospheric phenomena including magnetospheric substorms, gravity waves, magnetohydrodynamic waves, and the refraction of radio waves [1], [2]. SuperDARN radars use a multi-pulse technique to measure the complex autocorrelation function (ACF) of the received backscatter over 75–110 range gates in the radar's field of view. These ACFs are then used to determine the power (signal-to-noise ratio), line of sight Doppler velocity and Doppler spectral width associated with the scattering target by fitting model functions to the measured ACF power and phase [3], [20]. However, the ACF technique has been shown to produce large and occasionally unrealistic errors in the measured Doppler velocity [5]. Since all radars operate using common data analysis software, it is difficult to assess the accuracy of velocity measurements.

There have been several attempts to compare velocities measured by SuperDARN radars to other instruments. Drayton [16] pointed out that SuperDARN velocities were on average lower than F-region drift velocities measured by DMSP satellites. Gillies [17] also suggest that velocities derived from the SuperDARN ACFs were systematically lower than DMSP velocities. Similar results were found when SuperDARN derived velocities were compared with other incoherent scatter radars in [18] and in [19].

In this paper we present a new technique for measuring the Doppler velocity that is independent of, and parallel with, the ACF technique, and thus provides a means of cross-checking SuperDARN velocity measurements "on the fly" with the same radar. This technique takes advantage of the advanced capability of the TIGER 3 radar, a new SuperDARN radar, which was recently installed near Adelaide, South Australia [15].

The Tasman International Geospace Environment Radar (TIGER) is part of the SuperDARN expansion to the southern hemisphere, currently consisting of three HF radars [6], [7], [8]. TIGER 3 is an all-digital radar, integrated with the latest technology and Field Programmable Gate Arrays (FPGA) chips which could bring about greater performance and increased operational flexibility [9], [10], [11], [15]. Based on the TIGER 3 configuration, we propose a Spectrum Difference Function (SDF) technique, which runs parallel with the current ACF, and therefore, provides a means for cross-checking results of the ACF technique. In the proposed technique, the difference between the magnitude spectra of the transmitted and the received signals is defined as the SDF and employed to deduce target velocity. Mathematical derivation shows that the gradient of the SDF in the vicinity of the carrier frequency is directly proportional to the target Doppler frequency. Therefore, by calculating the gradient of the SDF, the Doppler frequency and hence the target velocity can be deduced.

This paper is organized as follows: section II presents the mathematical derivation of the proposed technique. Section III analyses Matlab simulation results for the application of the SDF technique on the TIGER 3 platform, in both noise-free and Gaussian noise scenarios. Section IV presents the new method for the velocity estimation in radar using the SDF technique. The computational consideration of the technique is given in Section V, followed by a discussion, future work and conclusion.

II. Mathematical derivation of the SDF technique

A. Motivation

Assume the magnitude spectrum function of the transmitted (Tx) signal is $G(f_c)$, where f_c is the carrier frequency, and G is the envelope function determined by the Fast Fourier Transform (FFT) of the Tx pulse envelope. Assume that the received (Rx) signal is at the same scale as the Tx signal and noise has little effect on the spectrum of Rx signal. The magnitude spectrum of the Rx signal, therefore, is similar but only a shifted version of the Tx by a Doppler frequency and can be approximate as $G(f_c+f_d)$. In most radar applications, f_d is negligible in comparison to f_c, therefore the difference between the Tx and Rx spectra can be calculated as:

$$G(f_c + f_d) - G(f_c) \approx f_d \times G^{'}(f_c) \tag{1}$$

Where $G^{'}(f_c)$ is the derivative function of $G(f_c)$.

Equation (1) shows the relationship between the Doppler frequency and spectra of the Tx and Rx signals. It suggests that the Doppler frequency can be calculated if the spectra of Tx and Rx signals are known.

B. Spectrum Difference Function

In this section, the Spectrum Difference Function (SDF) in the vicinity of carrier frequency is derived. Given that the radar Tx signal, $x(t)$, is a periodic rectangular pulse sequence with a carrier frequency f_c, pulse width T, and pulse repetition time τ:

$$x(t) = \begin{cases} \cos(2\pi f_c t) & \text{if } k\tau \le t \le k\tau + T \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Its Fourier transform is given by:

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi f t} dt \tag{3}$$

Assume there is only a single body moving target in a noise free environment where the received signal $y(t)$ is also periodic with a pulse width T and pulse repetition time τ. The received signal will exhibit a time shift (due to delay), Doppler frequency shift (due to the moving target) and attenuated magnitude A (due to loss in transmission path and reflection). As a result, the received signal can be expressed as:

$$y(t) = \begin{cases} A\cos\big(2\pi(f_c + f_d)(t + t_d)\big) & \text{if } k\tau \le t \le k\tau + T \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The Doppler frequency depends on the target movement. That is, if the target is moving away from radar, f_d is negative and if the target is approaching the radar, f_d is positive. As this technique compares spectra of the Tx and Rx signals to determine the Doppler frequency f_d, the signal magnitude and time delay do not play important roles. Upon time and amplitude standardization in the receiver, the Rx signal can be expressed as:

$$y(t) = \begin{cases} \cos\big(2\pi(f_c + f_d)t\big) & \text{if } k\tau \le t \le k\tau + T \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The Fourier transform of the Rx signal is given as:

$$Y(f) = \int_{-\infty}^{+\infty} y(t) e^{-j2\pi f t} dt \tag{6}$$

The SDF is defined as:

$$D(f) = |Y(f)| - |X(f)| \tag{7}$$

It is expected that the SDF is closely related to f_d. This relationship is proven in the following section.

C. Mathematical derivation of the SDF technique

To simplify the mathematical problem, a single pulse transmission signal will be examined first. That is:

$$x(t) = \begin{cases} \cos(2\pi f_c t) & \text{if } -\frac{T}{2} \le t \le \frac{T}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

$x(t)$ can be rewritten in the following form:

$$x(t) = g(t)\cos(2\pi f_c t)$$

where $g(t)$ is a rectangular window function in the form of:

$$g(t) = \begin{cases} 1 & \text{if } -\frac{T}{2} \le t \le \frac{T}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The Fourier transform of the rectangular window function $g(t)$ can be expressed as:

$$G(f) = \int_{-\infty}^{+\infty} g(t)e^{-j2\pi ft}dt \tag{10}$$

This results in the magnitude spectrum of the rectangular window exhibiting a sinc function form:

$$|G(f)| = \left|\frac{\sin(\pi fT)}{\pi f}\right| \tag{11}$$

The Fourier transform of the Tx signal then becomes:

$$\begin{aligned}
X(f) &= \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}dt \\
&= \int_{-\infty}^{+\infty} g(t)\cos(2\pi f_c t)e^{-j2\pi ft}dt \\
&= \int_{-\infty}^{+\infty} g(t)e^{-j2\pi ft}\frac{e^{-j2\pi f_c t} + e^{j2\pi f_c t}}{2}dt \\
X(f) &= \frac{1}{2}\Big[G(f + f_c) + G(f - f_c)\Big]
\end{aligned} \tag{12}$$

The above equation indicates the upper and lower side bands in the spectrum of the Tx signal. In SuperDARN radars, f_c ranges from 8 to 20 Mhz, the two bands can be considered non overlapping. Therefore, the magnitude spectrum of $x(t)$ can be approximated as:

$$|X(f)| = \frac{1}{2}\Big[|G(f + f_c)| + |G(f - f_c)|\Big] \tag{13}$$

The received signal upon standardization can be expressed as:

$$y(t) = g(t)\cos\big(2\pi(f_c + f_d)t\big)$$

Similarly, we have the magnitude spectrum of $y(t)$ is:

$$|Y(f)| = \frac{1}{2}\Big[|G(f + f_c + f_d)| + |G(f - f_c - f_d)|\Big] \tag{14}$$

From (13) and (14), the SDF can be calculated as:

$$\begin{aligned}
D(f) &\approx \frac{1}{2}\Big[\big(|G(f - f_c - f_d)| - |G(f - f_c)|\big) + \big(|G(f + f_c + f_d)| - |G(f + f_c)|\big)\Big] \\
&= \frac{1}{2}\big(D_1(f) + D_2(f)\big)
\end{aligned} \tag{15}$$

Let us consider the upper side band function:

$$\begin{aligned}
D_1(f) &= \frac{|G(f - f_c - f_d)| - |G(f - f_c)|}{2} \\
&= \frac{K(f - f_c - f_d) - K(f - f_c)}{2}
\end{aligned} \tag{16}$$

Where

$$K(f) = |G(f)| = \left|\frac{\sin(\pi fT)}{\pi f}\right|$$

In radar applications, f_d is significantly smaller in comparison to f_c. For the SuperDARN radars, the typical drift velocity is smaller than 1000 m/s [12]. Doppler frequency shift f_d, thus, is nominally smaller than 200 Hz. Therefore, the Taylor series of function $K(f - f_c - f_d)$

that is infinitely differentiable f_d in the neighbourhood $(f - f_c)$ can be applied:

$$\begin{aligned} K(f - f_c - f_d) &= K(f - f_c) + \frac{K^{(1)}(f - f_c)}{1!}(-f_d) + \frac{K^{(2)}(f - f_c)}{2!}(-f_d)^2 \\ &+ \frac{K^{(3)}(f - f_c)}{3!}(-f_d)^3 + ... \end{aligned} \quad (17)$$

Ignoring the high order terms in the above equation we then have the approximation:

$$D_1(f) \approx -\frac{K^{(1)}(f - f_c)f_d}{2} \quad (18)$$

It can be seen that the envelope of $K(f - f_c)$ is identical to that of $K(f)$, only being shifted in frequency. $K^{(1)}(f - f_c)$ in the neighbourhood of f_c in equation (18), therefore, is equal to $K^{(1)}(f)$ in the vicinity of f = 0 Hz. Let us examine the function $K^{(1)}(f)$ in the vicinity of f = 0Hz.

$$K(f) = \left|\frac{T \sin \Omega}{\Omega}\right| \quad \text{in that} \quad \Omega = \pi f T$$

In the vicinity of $f = 0Hz$, which is also the vicinity of $\Omega = 0$, we have:

$$\left|\frac{T \sin \Omega}{\Omega}\right| = \frac{T \sin \Omega}{\Omega}$$

Therefore:

$$K^{(1)}(f) = \pi T^2 \frac{\Omega \cos \Omega - \sin \Omega}{\Omega^2} \quad (19)$$

$$K^{(2)}(f) = \pi^2 T^3 \frac{(2 - \Omega^2) \sin \Omega - 2\Omega \cos \Omega}{\Omega^3} \quad (20)$$

$K^{(2)}(f)$ in the neighbourhood of $f = 0Hz$ can be found by:

$$\lim_{f \to 0} K^{(2)}(f) = \pi^2 T^3 \lim_{\Omega \to 0} \frac{(2 - \Omega^2) \sin \Omega - 2\Omega \cos \Omega}{\Omega^3} \quad (21)$$

In the above equation, the limits of the numerator and denominator, as Ω approaches 0, are 0. Therefore L'Hopital's rule can be applied to show:

$$\lim_{f \to 0} K^{(2)}(f) = -\frac{\pi^2 \tau^3}{3} \quad (22)$$

With the second derivative $K^{(2)}(f)$ being a constant in the neighbourhood of f = 0 Hz, the first derivative $K^{(1)}(f)$ will exhibit a linear function in that area. Similarly, $K^{(1)}(f - f_c)$ will show a linear function in the vicinity of f_c. From equation (18), this means the gradient of the SDF in the vicinity of f_c is a linear function of f_d.

From equations (18),(19),(22) the gradient, $Grad$, of the upper side band function $D_1(f)$ in the neighbourhood of f_c is determined as follows:

$$Grad = \frac{\pi^2 T^3}{6} f_d \quad (23)$$

Equation (23) shows that the gradient of the SDF in the neighbourhood of the carrier frequency is a linear function of the Doppler frequency. Therefore, if the value of $Grad$ is known, the Doppler frequency and hence the target velocity can be determined.

III. Simulation results

In this section, MATLAB is used to simulate Tx/Rx signals and their magnitude spectra. The SDF is then calculated, from which the gradient in the neighbourhood of carrier frequency is computed and the Doppler frequency is deduced. Variable Doppler frequency values are tested to confirm equation (23) from the previous section.

The current TIGER 3 transmits the Kat's 8-pulse sequence [20]. However, in this section, only a single pulse is simulated to demonstrate the validity of the new technique. For this simulation all other TIGER radar parameters are kept unchanged.

A. SDF in a noise-free scenario

The transmit signal is simulated in MATLAB with the following parameters: pulse length 300 μS and carrier frequency $f_c = 8$ MHz. The TIGER 3 radar operates at a sampling rate of $f_s = 125$ MHz, therefore, a single pulse contains 37,500 samples. A single period is 100 ms and contains 12.5 million samples.

The SDF technique is based on the linear region of the SDF in the vicinity of carrier frequency. This area proves narrow, in the order of a few thousand Hz. If the FFT of one Rx pulse is taken, the number of samples in the FFT is N = 37,500, the frequency bin width achieved is therefore:

$$f_w = \frac{f_s}{N} = 3333\,Hz$$

Therefore, only a few samples of the signal spectrum can be used for the SDF technique. This will not be enough for the technique to achieve a high level of accuracy. Zero-padding is, thus, necessary for the SDF technique to reduce frequency bin width and thereby, having more samples for processing. However, this comes at the cost of increased computation. In this paper, a zero-padding up to N = 12,500,000 points (equal to one pulse period) will be used to test the validity of the SDF technique.

With a sampling rate $f_s = 125$ MHz and the size of the FFT N = 12,500,000, the frequency bin width achieved is :

$$f_{res} = \frac{f_s}{N} = 10Hz$$

The carrier frequency is 8 MHz and, thus, located at bin 800,000. 4,000 samples around the carrier frequency (from bin 798,001 to bin 802,000) of the Tx magnitude spectrum are shown in Figure 1.

The Rx pulse is generated in the same manner as the Tx pulse, except the carrier frequency is marginally altered to 8,000,100 Hz, artificially inducing a Doppler shift, f_d, of 100 Hz. The SDF is formed by the difference in magnitude spectra of Tx and Rx signals. For this technique, only the SDF in the vicinity of the carrier frequency is of interest, located around bin 800,000 of the spectrum. 600 and 200 samples around bin 800,000 of the SDF are shown in Figure 2 and Figure 3, respectively.

As shown in Figure 2 and Figure 3, the SDF function in the vicinity of carrier frequency exhibits a linear function in a small region of 2,000 Hz, equivalent to 200 FFT samples. The

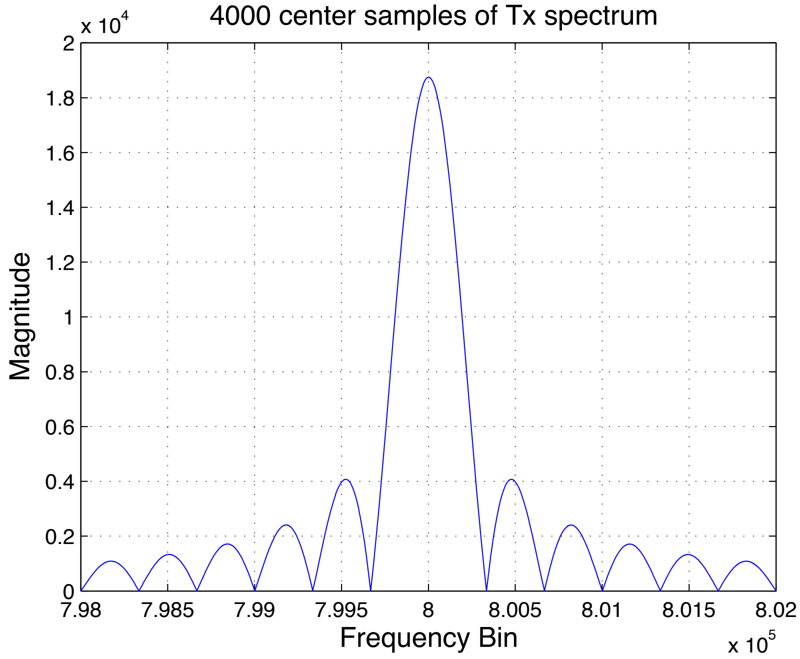


Fig. 1: Magnitude spectrum of Tx pulse in the neighbourhood of $f_c = 8MHz$. The spectrum is centered at bin 800,000

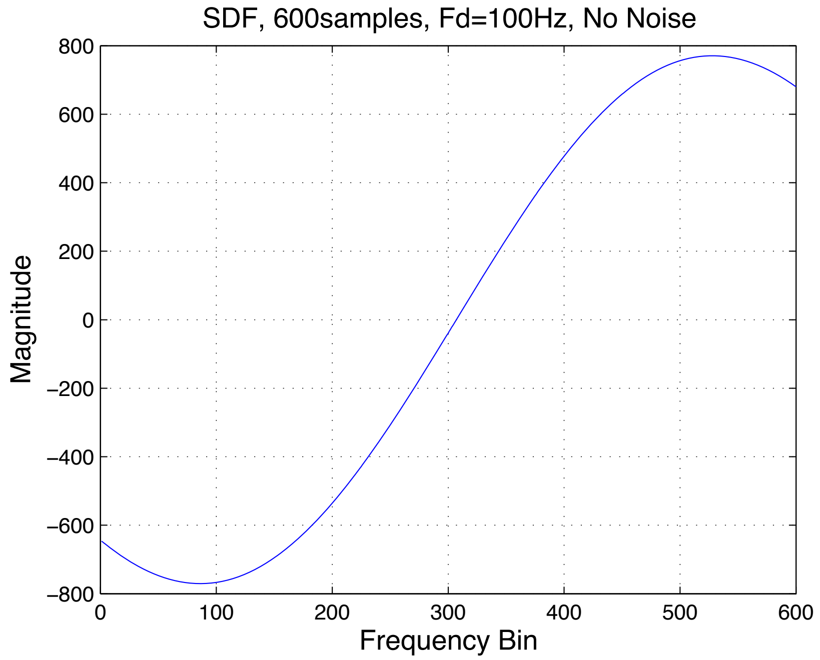


Fig. 2: 600 samples of SDF in the neighbourhood of f_c

more samples that are taken for SDF, the less linear it is. If more than 400 samples are taken, the SDF no longer exhibits linear but a cubic function. With a cubic fit, the computational complexity of the SDF technique will increase.

In this paper, the linear fit is implemented to reduce computational complexity. It is favourable to take less data samples because the error introduced by the linear fit will be smaller. However, under the presence of noise, more data samples are desirable for accurate

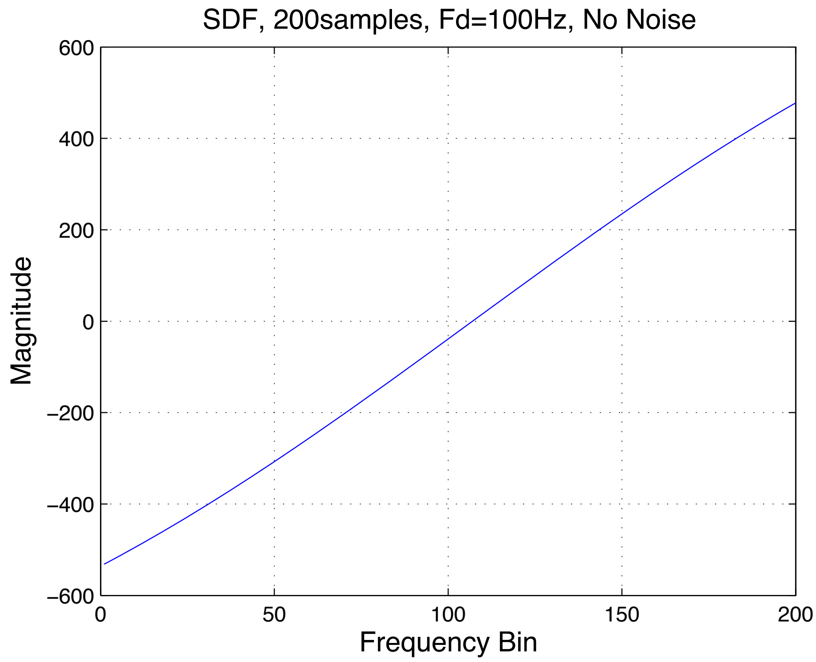


Fig. 3: 200 samples of SDF in the neighbourhood of f_c

interpolation. Both sides considered, a reasonable approach is to utilise the Least Square Error criteria (LSE) for a linear fit of the 200 samples of the SDF. Details of the LSE criteria are presented in [13].

Applying the LSE criteria to the 200 samples of the SDF when input $f_d = 100$ Hz equates to $Grad = 0.5253$. Varying the input Doppler frequencies from 0.1 to 200 Hz, the corresponding gradients are calculated, from which the estimated Doppler frequencies are deduced accordingly, using equation (23). The results are shown in Table 1 and Figure 4.

Table 1: Gradient and estimated f_d in noise free scenario

Input f_d (Hz)	Gradient	f_d estimate (Hz)
0.1	0.000526	0.095
1	0.005257	0.946
2	0.010517	1.892
5	0.026292	4.731
10	0.052585	9.462
20	0.105164	18.923
50	0.262844	47.308
100	0.525284	94.616
150	0.786974	141.924
200	1.047570	189.232

Table 1 and Figure 4 indicates the directly proportional relationship between f_d and the gradient of $SDF(f)$ in the vicinity of the carrier frequency. Consider, for example, the case where $f_d = 100\ Hz$,if the amplitude normalization is taken, the actual gradient of the SDF is thus given as:

$$Grad = \frac{0.52528}{125,000,000} = 42.022 \times 10^{-10}(unit/Hz) \qquad (24)$$

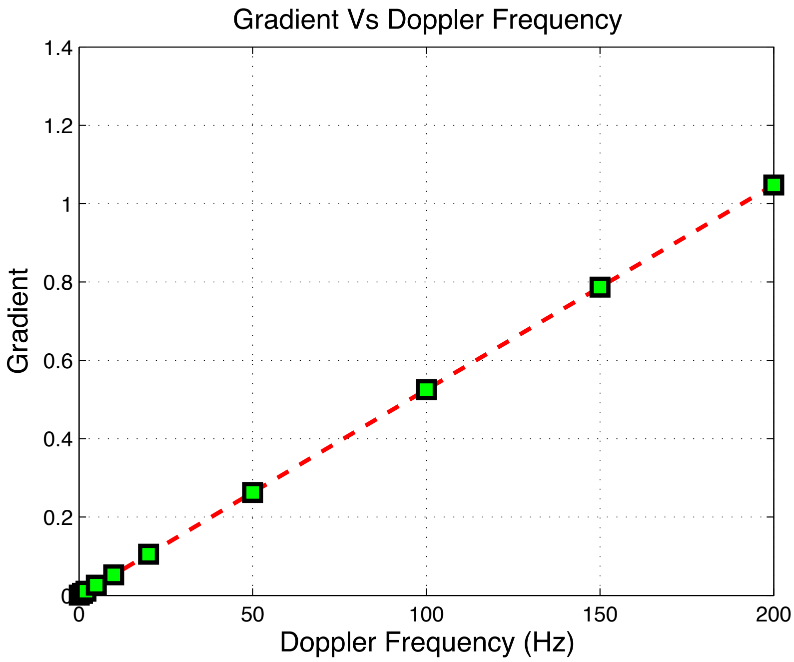


Fig. 4: Gradient versus f_d in noise free situation

Substituting the above value of $Grad$ and simulated data $\tau = 0.3$ ms into equation (23) to find f_d:

$$f_d = \frac{6 \times 42.022 \times 10^{-10}}{\pi^2(3 \times 10^{-4})^3} = 94.62(Hz) \tag{25}$$

With the simulated input value $f_d = 100$ Hz, equation (25) shows an estimation error of 5.38 %. This is considered the error of the technique and can be attributed to the following reasons:

1. Equation (23) is an approximation.
2. The SDF is not linear.
3. The LSE method itself introduces an error.

The 5.38 % error in a noise-free environment suggests that the result needs to be scaled up by a scale factor (SF) of $100/94.62 = 1.0569$ to compensate for the systematic error of the technique.

B. SDF in White Gaussian Noise scenario

In this section, let us assume that the noise at the radar receiver is additive white Gaussian noise $n(t)$:

$$r(t) = y(t) + n(t) \tag{26}$$

where $r(t)$ is the total signal at the radar receiver, and $y(t)$ is the reflected signal from a moving target.
The stepped procedure in this section is shown below:
Step 1: calculate magnitude spectrum of the Tx signal, take 200 central samples as reference spectrum.
Step 2: generate a single Doppler frequency f_d from 1 to 200 Hz.

Step 3: generate a single SNR value.
Step 4: generate the Rx signal with a random phase, given f_d and add white Gaussian noise with the given SNR, calculate magnitude spectrum of the Rx signal and SDF, apply LSE criteria to 200 centre samples of the SDF to find the gradient then substitute into equation (23) to find calculated f_d, scale up the calculated f_d by a SF to compensate for the estimated error, subtract this value from f_d generated in step 2 to get the absolute error.
Step 5: loop to step 4 for 10,000 times (which is considered a statistically sufficient number). For a single set of 10,000 random Rx signals, two variables are produced; the mean error (ME) and root mean squared error (RMSE). The two numbers are both taken as the quality benchmark of the SDF technique.
Step 6: loop to step 3 and change SNR.
Step 7: loop to step 2 and change Doppler frequency.

Simulation results where $f_d = 100$ Hz with $SNR = 20$ dB is shown in Figure 5. Simulation results for $f_d = 100$ Hz, 50 Hz, 10 Hz with SNR = 25 dB, 20 dB, 15 dB, respectively are shown in Table 2.

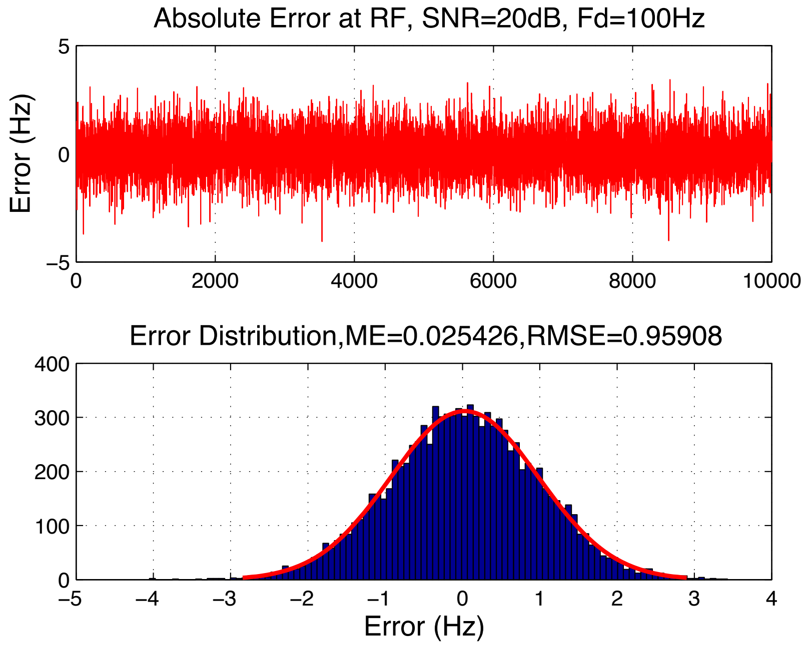


Fig. 5: Error distribution from 10,000 random Rx signals, SNR = 20dB, rectangular window, FFT size N=12,500,000, $f_d = 100$ Hz

Table 2: SDF technique error in a Gaussian noise scenario. All values are in Hz

	$SNR = 25dB$		$SNR = 20dB$		$SNR = 15dB$	
f_d	ME	RMSE	ME	RMSE	ME	RMSE
100	0.0393	0.5501	0.0254	0.9591	0.0547	1.7009
50	0.0591	0.5554	0.0528	0.9572	0.0358	1.6846
10	0.0734	0.5518	0.0634	0.9475	0.0501	1.6693

Table 2 shows that a high level of accuracy can be achieved with the new SDF technique using a rectangular pulse. At a typical level of SNR = 20 dB in TIGER 3 radar, the RMSE of the estimation is approximately 0.96Hz, irrespective of the Doppler frequency.

The ME of all cases in Table 2 are positive. Although small, they indicate a positively biased estimation. This could be resolved by reducing the SF, thereby pulling the ME back to zero. By doing this, the RMSE is also further reduced.

IV. New method for the velocity estimation

A new method for measuring target velocity based on the SDF technique can be summarized as follows:
Step 1: Calculate the magnitude spectrum of the Tx signal, take 200 centre samples and consider this as the reference spectrum
Step 2: Process data in one range gate at a time, including:

1) Standardize the Rx signal. All the calculations so far are based on the assumption that the magnitude of the Tx and Rx signals are the same. It is necessary to standardize the Rx signal magnitude to the Tx signal magnitude.
2) Calculate the SDF
3) Use the LSE criteria to calculate the gradient of the SDF in the vicinity of carrier frequency
4) Calculate f_d using equation (23)
5) Scale up the calculated f_d with a pre-determined SF. The value of the SF, however, depends on a number of variables.
6) Calculate the target's velocity

V. Computational consideration for the SDF technique

As discussed in the previous sections, zero padding is necessary because the technique utilises a small region in the SDF, in which there must be enough samples for processing. The more zero-padding to the signal, the better the representation of the spectrum and more samples for processing, which results in more accurate frequency estimation. On the other hand, further zero-padding increases the computation complexity of the system, mostly due to the size of the FFT.

In the previous sections, a zero-padding of Tx and Rx signal up to 12,500,000 points is implemented to simulate the technique's validity. With this arrangement, the frequency bin width achieved is 10 Hz. The linear region in the SDF is about 2000 Hz, therefore, consists of 200 samples. If SNR = 20 dB and a rectangular window is applied, the simulation results show that the estimation uncertainty of 0.95 Hz can be achieved with the SDF technique.

An FFT size of 12,500,000 points, however, is a huge computation and can pose difficulties in the implementation to any system. The current TIGER 3 configuration does not meet the real time processing requirement for an FFT of 12,500,000 points. Even if future configurations could meet this requirement, it is not considered efficient to do so. A reduction of the FFT size, therefore, is a must.

The first approach is to directly reduce the size of zero-padding. As a result, there are less samples in the area of interest for the SDF technique. The reduction of computation, therefore, comes at the cost of decreasing accuracy. However, the size of the FFT can be reduced to a level where the estimation accuracy still meets the system requirement.

If a zero-padding of Tx and Rx signal up to 1,250,000 points is implemented, for example, the frequency bin width achieved is 100 Hz and there are 20 samples in the linear region of the SDF for the linear fit. At the extreme, the minimum necessary zero-padding is at N=125,000. In this arrangement, the frequency bin width is 1000 Hz. Therefore, there are only 3 samples in the linear region for the linear fit. It is to be expected that the accuracy reduces significantly in this case.

Initial simulation results show that as the size of zero-padding reduces from 12,500,000 to 1,250,000 the accuracy of the SDF technique degrades insignificantly. From N = 625,000 downward, the RMSE and the ME of error both increase significantly, showing the estimation is highly biased and has large RMSE. At this point, there are insufficient samples for the linear fit which limits the feasibility of the SDF technique for smaller sizes of zero-padding. A zero-padding of around one million points could be a sensible solution as a trade-off between accuracy and computational complexity.

The second approach is to shift the processing domain from RF to an IF frequency. The SDF technique is based on the spectra of Tx and Rx signals, therefore the technique is feasible as long as the two spectra are well preserved. As the bandwidth of returned radar echoes is in the order of a few KHz, we could reduce the IF frequency significantly compared to the RF carrier signal. By moving the two spectra to a much lower IF frequency, the sampling rate can be reduced significantly and so the size of the FFT. This solution comes with the cost of extra mixers and filters which are necessary for shifting frequency and preserving spectra. Overall, the computation requirement can be reduced to a manageable level.

With a combination of the first and the second approaches, the size of the FFT could be reduced to a level of few thousand points. The current TIGER 3 configuration can meet both processing time and hardware requirement for this FFT size. The SDF technique, therefore, can be feasible in TIGER 3 platform.

VI. Discussion

The SDF technique is designed as a cross-checking mechanism for the standard ACF technique. The technique acts as an independent measurement of the velocity utilising the data available and as such is not a replacement of the ACF.

The technique uses only one Rx pulse. The magnitude spectrum of the Rx pulse in one range gate is compared to the magnitude spectrum of Tx pulse to find Doppler frequency in that range gate. The technique requires a “clean” Rx pulse from the range gate which is currently examined. The term “clean” in this instance refers to a signal from that range gate which only comes from one Tx pulse and is not “contaminated” by cross range interference caused by multiple pulse sequence.

Considering the 7-pulse sequence as an example: pulses are transmitted at lags [0, 9, 12, 20, 22, 26, 27] with one lag time being 2400 μS, and the pulse width is 300 μS [20]. There is, by design, a large gap between the first and the second pulse (9 lags which is equivalent to

72 range gates). For the time of the first 9 lags, the receiver captures only a "clean" Rx pulse from the first 72 range gates due to the first Tx pulse. After that time, the Rx signal from the receiver will be "contaminated" by the cross range interference caused by the multiple pulse sequence. These first 72 "clean" Rx pulses therefore could be utilised by the SDF for cross-checking velocity in their respective range gates.

Similarly with 8-pulse sequence, the pulses are transmitted at lags [0, 14, 22, 24, 27, 31, 42, 43] with a lag time of 1800 μS, and unchanged pulse width [20]. Therefore, the "clean" pulses could be used by the SDF technique for cross-checking velocity for the first 84 range gates.

Consequentially, the SDF technique is not capable of covering all TIGER radar range gates. However, for the purpose of velocity cross-checking of the ACF technique, 72 range gates (in the 7-pulse sequence) or 84 (in the 8-pulse sequence) are considered sufficient.

Finally, the technique does not take advantage of the signal processing gain that the multiple pulse sequence offers with the ACF. However, it uses multiple measurements from the oversampling of the received signal, which provides a signal processing gain in the data extraction. While the ACF technique can generate one result (average velocity) after a number of pulse sequences during the integration time, the SDF technique can be implemented in real time and produces one result every pulse sequence. At a common mode, the ACF technique correlates 70 pulse sequences (equivalent to an integration time of 7s) for one velocity estimate. This estimate is considered as the average velocity of the target within the integration time. At the same time, the SDF technique can generate 70 independent estimates of the same velocity. The temporal resolution of the SDF technique is, therefore, much higher than that of the ACF. Furthermore, because these 70 independent estimates all come from one slow changing velocity, the post-processing of these numbers can further reduce the uncertainty. In a simplistic example, the uncertainty due to noise in the SDF technique could be reduced by $\sqrt{N}$, where N is the number of pulses in the average.

VII. Future work

The simulation results in MATLAB are consistent with the mathematical derivation of the technique and show that the Doppler frequency can be estimated with a high level of accuracy. There are several technical issues of the technique which need to be further investigated:

1) The SF in this paper is calculated based on a given Doppler frequency f_d. However, f_d variation results in different values of the SF, although the differences are small. There are other factors, such as, the random phase of the Rx signal, pulse shape, fitting technique and the number of samples taken for the SDF, which all have an influence on the value of the SF. The value of the SF, therefore, varies within a certain range. It is necessary to decide on one value of the SF which introduces the smallest error.
2) The rectangular pulse has been used in this paper for simplicity and to simulate the technique. In reality, however, a rectangular pulse is not practical nor desirable. Currently, TIGER radars utilise a modified version of Gaussian pulse shaping for the Tx signal [14]. This pulse shape and several others, such as, Hanning, Hamming and Blackman need to be tested with the new technique to find the optimum solution for the SDF technique.

3) The simulation conducted thus far have ultilised Tx and Rx signals with the same magnitude. In reality, the Rx signal magnitude is very small and needs to be standardized to the Tx signal magnitude. A standardization algorithm needs to be developed for this purpose.
4) In the presence of noise, generally the more data samples are taken, the more advantageous it is for signal processing. However, if more data samples are taken, the SDF does not exhibit a linear relationship but a cubic one. It is, thus, necessary to test and compare linear and cubic fits. In each, different versions of the number of samples taken need to be tested to find the optimum solution for the implementation of the technique.

VIII. Conclusion

This paper has proposed a novel SDF technique for the Doppler frequency estimation in SuperDARN radars. This technique is expected to be useful for independently verifying the accuracy of SuperDARN velocity measurements, which may in turn lead to improvements to the standard ACF technique. Simulation results show that the new technique can achieve a high level of accuracy. However, further research needs to be conducted to study the implementation of the SDF technique on the TIGER 3 platform.

References

[1] R. A. Greenwald, K. B. Baker, J. R. Dudeney, M. Pinnock, T. B. Jones, E. C. Thomas, J. P. Villain, J. C. Cerisier, C. Senior, C. Hanuise, R. D. Hunsucker, G. Sofko, J. Koehler, E. Nielsen, P. Pellinen, A. D. M. Walker, N. Sato, and H. Yamagishi, "DARN/SUPERDARN a global view of the dynamics of high-latitude convection," *Space Science Rev*, vol. 71, no. 1-4, pp. 761-796, 1995.

[2] G. Chisham, M. Lester, S. Milan, M. Freeman, W. Bristow, A. Grocott, K. McWilliams, J. Ruohoniemi, T. Yeoman, P. Dyson, R. Greenwald, T. Kikuchi, M. Pinnock, J. Rash, N. Sato, G. Sofko, J.-P. Villain, and A. Walker,A decade of the super dual auroral radar network (superdarn): scientific achievements, new techniques and future directions," *Surveys in Geophysics*, vol. 28, no. 1, pp. 33-109, 2007. [Online]. Available:http://dx.doi.org/10.1007/s10712-007-9017-8

[3] P. V. Ponomarenko and C. L. Waters, "Spectral width of superdarn echoes: measurement, use and physical interpretation," *Annales Geophysicae*, vol. 24, no. 1, pp. 115-128, 2006. [Online]. Available: http://www.ann-geophys.net/24/115/2006/

[4] K.McWilliams, "Superdarn pulse sequences optimization and testing," in *2003 SuperDARN workshop*, Finland, 2003.

[5] P. Ponomarenko, "Superdarn velocity errors," in *2013 SuperDARN Workshop*, Moose Jaw, Saskatchewan, Canada, May 26 - May 31 2013.

[6] P. L. Dyson and J. C. Devlin, "The tasman international geospace environment radar," *The Physicist (The Australian Institute of Physics)*, vol. vol. 37, pp. 48-53, 2000.

[7] P. Dyson and J. C. Devlin, "The tiger radar - an extension of superdarn to sub-auroral latitudes," in *WARS'00 (Workshop on Applications of Radio Science) Proceedings*, National Committee for Radio Science, Australian Academy of Science, 2000, pp. 9-31.

[8] P. Dyson, J. Devlin, M. Parkinson, and J. Whittington, "The tasman international geospace environment radar (tiger) - current development and future plans," in *Radar Conference, 2003. Proceedings of the International*, Sept 2003, pp. 282-287.

[9] J. Whittington and J. C. Devlin, “Architecture for a fully digital implementation of a tiger radar,” in *SuperDARN Workshop Proceedings*, Rydal Hall, Cumbria, U.K, 16-20 May 2005.

[10] J. Whittington, M. L. Parkinson, and J. C. Devlin, “Advanced operational capabilities through implementation of a fully digital superdarn radar,” in *SuperDARN Workshop Proceedings*, Rydal Hall, Cumbria, U.K, 16-20 May 2005.

[11] W. J. and D. J., “A high performance digital radar for extended space weather investigations,” in *Australian Institute of Physics 17th National Congress 2006 Proceedings*. Australian Institute of Physics, 2006.

[12] R. A. Greenwald, K. B. Baker, R. A. Hutchins, and C. Hanuise, “An hf phased-array radar for studying small-scale structure in the high-latitude ionosphere,” *Radio Science*, vol. 20, no. 1, pp. 63-79, 1985. [Online]. Available: http://dx.doi.org/10.1029/RS020i001p00063

[13] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993, ch. 8, pp. 223-226.

[14] B. Bienvenu, N.-V. Vu, J. Whittington, and J. Devlin, “Variable-length quadratic envelope pulse shaping and automatic gain control for hf radars,” in *6th International Conference on Broadband and Biomedical Communications (IB2Com), 2011*, Nov 2011, pp. 259-263.

[15] E. Custovic, A. J. McDonald, T. A. Kane, J. Vu, N. V. Whittington, and J. C. Devlin, “Next generation of over the horizon HF radars and the determination of foF2 in real-time,” in *Proceedings of the 11th Australian Space Science Conference, Canberra*, vol. 1, September 2011, pp. 141–156.

[16] Drayton, R. A. and Koustov, A. V. and Hairston, M. R. and Villain, J.-P., “Comparison of DMSP cross-track ion drifts and SuperDARN line-of-sight velocities,” in *Annales Geophysicae*, vol. 23, no. 7, 2005, pp. 2479–2486. [Online]. Available: http://www.ann-geophys.net/23/2479/2005/

[17] Gillies, R. G. and Hussey, G. C. and Sofko, G. J. and McWilliams, K. A. and Fiori, R. A. D. and Ponomarenko, P. and St.-Maurice, J.-P., “Improvement of SuperDARN velocity measurements by estimating the index of refraction in the scattering region using interferometry,” in *Journal of Geophysical Research: Space Physics*, vol. 114, no. A7, 2009, pp. n/a–n/a. [Online]. Available: http://dx.doi.org/10.1029/2008JA013967

[18] Xu, L. and Koustov, A. V. and Thayer, J. and McCready, M. A., “SuperDARN convection and Sondrestrom plasma drift,” in *Annales Geophysicae*, vol. 19, no. 7, 2001, pp. 749–759. [Online]. Available: http://www.ann-geophys.net/19/749/2001/

[19] Davies, J. A. and Lester, M. and Milan, S. E. and Yeoman, T. K., “A comparison of velocity measurements from the CUTLASS Finland radar and the EISCAT UHF system,” in *Annales Geophysicae*, vol. 17, no. 7, 1999, pp. 892–902. [Online]. Available: http://www.ann-geophys.net/17/892/1999/

[20] McWilliams, K., “SuperDARN Pulse Sequences Optimization and Testing,” in *2003 SuperDARN workshop*, Finland, 2003.

Innovative Use of Additive Structures to Develop Future Rocket Engine Technology

Dean R. Mosseveld

RMIT Centre for Advanced Manufacturing, RMIT University, GPO Box 2476, Melbourne, VIC, 3001, Australia

Summary: Selective Laser Melting (SLM), a form of additive manufacturing is expanding the capabilities of 3D manufactured components, making it possible to build end use products directly from CAD geometry. Building upon current understanding, applications of SLM are being investigated in high temperature and high pressure areas as applied to a regeneratively cooled rocket engine. Through combining a conformal micro-lattice into the cooling chamber, it has been possible to reduce the structural mass while also increasing the heat transfer rate into the cooling fluid. The post processed surface roughness of SLM produced parts is usually a deterrent to this manufacturing process, however the application has found to increase the thermal transfer in the system, improving the effectiveness of the regenerative cooling. Through a combination of analytical, computational, and empirical analysis it has been possible to determine performance gains, and through testing prove the legitimacy of additive manufacturing in high stress and high temperature applications. The development of complex micro-lattice structures are beginning to push the computational performance of CAD software, requiring processes to be developed to reduce the computational downtime of SLM machines. Challenges still await this newly developed technology, however this novel design pushes the boundaries in the design of small and medium sized rocket motors.

Keywords: *Additive Manufacturing, Selective Laser Melting, Liquid Rocket Engine, Micro-Lattice*

Introduction

Liquid rocket engines (LRE) are a well-established form of propulsion capable of propelling scientific, commercial and military vehicles into earth orbits and even to the outer edges of our solar system. However, a new era of rocket engine design and manufacturing may soon immerge with the recent advances in Additive Manufacturing (AM) technologies. The development of increasingly larger additive manufacturing machines and the advances in technologies such as Selective Laser Melting (SLM) will allow metallic structures with complex geometries to be created on increasingly larger scales [1]. Although advancing at a fast pace, SLM can still have difficulties in obtaining 100% dense structures and the inconsistencies present in the SLM process are still viewed as barriers to the widespread adoption of such manufacturing techniques.

With AM there is a significant increase in the overall part complexity that is achievable over that of conventional machining. Such complexities including internal geometries and integrated assemblies have made way for the construction of highly complex structures, such as micro-lattices (shown in Fig.1), which have been proven to exceed strength-to-weight ratios of all engineering materials of densities below 1,000kg/m^3 [2].. The adoption of complex geometries into otherwise conventional designs have made way for significant mass reductions and increased optimisation, resulting in an overall increase in efficiencies [2].

(a) (b)

Fig. 1: Micro-lattice made using SLM (a) and aerospace bracket designed with 40-60% mass reduction (b) [3].

Given the increased design freedom available with SLM, a novel engine concept has been designed to take advantage the increased flexibility. The concept derived takes a standard regenerative LRE and applies a structural micro-lattice into the cooling jacket, reducing the overall dry mass. In addition traditionally assembled components have been integrated to reduce the overall number of parts to two, allowing further reductions in engine mass.

The design of liquid rocket engines are well understood [4], the difficulty is realised in the introduction of multifunctional heat exchangers derived from three-dimensional micro-lattice structures [5]. These heat exchangers form regenerative cooling channels in the engine walls that allow thermal energy to be absorbed by the coolant and returned to the injector [4]. The ability to create micro-architecture within the confines of these walls improves the thermal and structural capabilities of fluid cooling channels [6].

Ensuring the correct material is used in the AM process is crucial to producing models which avoid manufacturing flaws and are suited to the part application [1]. The ability to print H13 tool steel is advantageous to the manufacture of rocket engines due to the high resistance to thermal shock, thermal fatigue and high strength at elevated temperatures [7], however it also exhibits relatively low thermal conductivity and little corrosion resistance when compared to more favourable nickel based alloys.

With the availability of rocket design software [8], a new design of LRE has been created that will be able to be manufactured in a single piece with the available SLM machine. Currently, SLM machines have quite small build chambers limiting the size of engine that can be produced.

Designing a next generation Liquid Rocket Engine

Research was conducted into regenerative cooling channels to improve the life of rocket combustion chamber walls [9], the addition of multi-functional heat exchangers within the wall cavities is a new concept, however previous research has been conducted on similar structures for heat exchanges. Previously multi-objective optimisation has been used to compare the structural and thermal performance of multifunctional sandwich panels with micro-architectured cores [6], however these concentrate mainly on compressive loading scenarios which cannot be directly applied.

Optimisation of the thermal and structural lattice geometry is not the goal of this paper, rather this is a proof of concept based approach with optimisation to come in future stages. Initial models have been developed for use as a first iteration of design that can be compared to the

future experimental results to produce distinction between cooling channel and engine system designs.
A number of lattice chamber wall designs have been developed, only a small few of which have been manufactured and tested. Some designs have been used to allow a comparison between regenerative systems and non-lattice designs within these cavities.

The current opportunities have allowed for the printing of a range of sandwich lattice structures to compare micro-lattice layouts and help develop a scale engine with a complete regenerative system. The results of testing will go on to be used in the creation of a complete engine design, allowing for further refinements to be made in future.

Rocket System Design

The engine was designed first by analysing the combustion chamber profile. The key engine parameters are shown in table 1. The design of this rocket was restricted as it was required to be printed within RMIT's SLM 250 HL, which has a build volume of 250x250x350 mm. Similarly, the maximum operating combustion pressures were restricted to 2 MPa to ensure that the design was achievable within the funding restrictions and ensure achievable system pressures without the use of a turbo pump. A representation of the designed rocket engine with internal lattice structure can be seen in Fig. 2.

Table 1: LRE design parameters

Parameter	Value	Unit
Design thrust	2,000	N
Design exit pressure	101.325	kPa
Chamber pressure	2	MPa
Characteristic length, L*	1100	mm
Oxidiser	Liquid Oxygen	-
Fuel	Jet A (Kerosene)	-
Oxidiser to Fuel Ratio	2.3	-

The system utilises a kerosene regenerative cooling system to extract heat away from the combustion chamber wall, reducing its temperature, while effectively retaining most of the strength of the combustion chamber wall during operation. The kerosene, also acting as the engine fuel, will be pre-heated by this process, increasing the combustion efficiency of the system and utilising the waste system energy.

Fig. 2: Engine segment showing cooling channel and combustion chamber profile

The design of the engine combustion chamber only considers the aerodynamic properties of the combustion process to ensure stable and efficient burn. To complete the design of the engine; material, structural and fluid analysis needed to be conducted.

Selective Laser Melting

Selective Laser Melting (SLM) uses a layer based construction technique where a 30µm layer of metallic powder is deposited on which a high powered laser then passes, solidifying the molten metallic layer to the layer beneath within a tolerance of+/- 20µm [11]. This process results in a 99.9% dense component [12] from a range of powdered materials including: steel, aluminium, titanium and nickel alloys.
The opportunity to print models in H13 tool steel has been provided by the RMIT’s Advanced Manufacturing Precinct (AMP), and the resultant lattice fluid testing structures can be seen in Fig. 3 below.

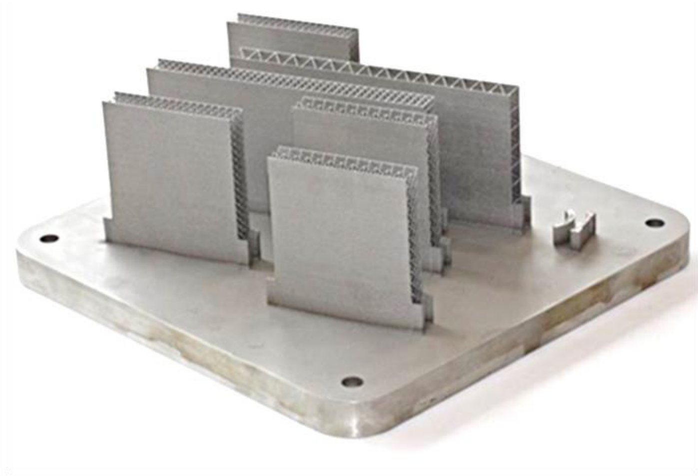


Fig. 3: Lattice fluid flow testing sections printed using SLM process.

The H13 tool steel has a few advantageous properties which make it idea for use in rocket motors, particularly its high strength when working at raised temperatures. The ideal material for this application would be a nickel based alloy, as in addition to the high strength at temperature it would also retain high thermal conductivity over the working range, increasing the efficiency of the regenerative system.

Micro lattice systems

The advantages of AM manufacturing and specifically the capabilities of the SLM process have enabled the manufacture of unique, high tolerance and micro-scale geometry which is unable to be manufactured using traditional techniques. One such geometry type is a micro-truss lattice, which can contain hundreds or even thousands of interlaced struts which form porous lattices, exhibiting high strength-to-mass ratio among other advantages.
The cell design can vary greatly between lattice types, however, due to limited resources, two conformal lattice types were selected for further investigation. The cell types face centred cubic (FCC) and body centred cubic (BCC) were selected due to the previous work conducted at RMIT’s Advanced Manufacturing Precinct and the favourable strut orientations which allow for ease of manufacturing in SLM processes.

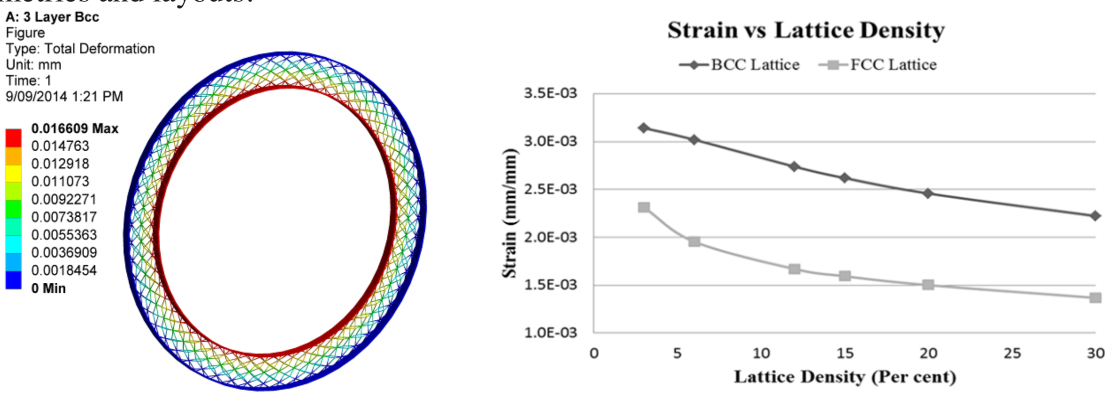
(a) (b) (c)

Fig. 4: Lattice cell Types (a), three layer BCC lattice (b) and three layer FCC lattice (c).

The body centred cubic cell as shown in Fig. 4(a) consists of eight individual struts, joining a node located in the centre of the cell to each of the eight corner nodes. The FCC cell consists of 24 struts, joining a node on the centre of each face of the cubic cell to each of the face corners. The FCC cell arrangement shares its boundaries with adjacent cells where the BCC cell will always have eight unique struts. The FCC cell arrangements have been modified in this project by removing four front and rear struts, to ensure the build orientation angle of any cell struts does not contain zero angle members. This design consideration is discussed in more detail in the proceeding sections.
Multi-layered lattice structures were also investigated to determine advantages of higher density lattice structures and different build orientations for future testing.
These lattice structures were analysed using a range of structural and fluid simulations to determine lattice suitability and to estimate system drag for future comparison.

Structural Analysis

When designing the rocket cooling channels, structural analysis was considered to ensure the channels were able to withstand the pressure of the coolant and heat produced in the combustion chamber. A computational approach was sought as it allowed for the analysis of a multitude of engine designs using a parametric geometry modification approach, reducing the need for long modelling times. This analysis was conducted on a series of radial engine segments with varying lattice geometries to determine how changes in geometry will affect the systems resistance to deflection under expected loads. Fig. 5(a) shows the deflection of a radial segment of the designed rocket engine with a three layer BCC lattice with an applied unit load. Unit loads were applied so a comparison could be made between the various lattice geometries and layouts.



(a) (b)

Fig. 5: Circular cross section of LRE under applied internal unit pressure load (a) and analysis results (b).

Fig. 5(b) shows the resulting decrease in lattice deformation with increasing lattice density. The analysis shows two key outcomes; BCC and FCC lattices with the same lattice densities behave differently under the same applied load and lattice structures can be tailored to a specific application depending on the internal load. The lattice densities of the two lattice cell configurations were achieved by varying the truss strut diameters.

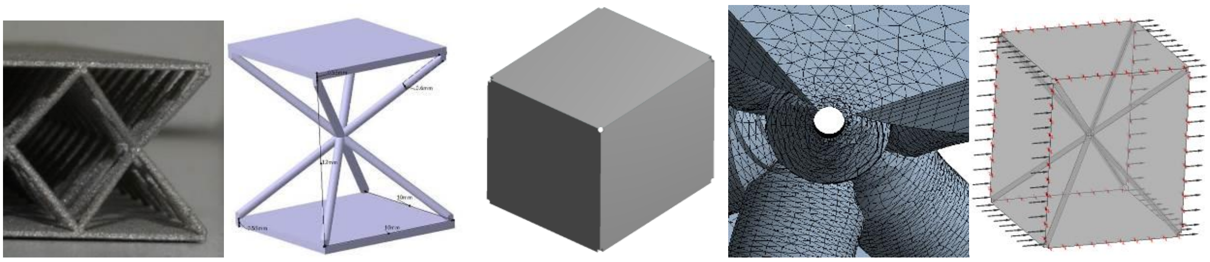
The first outcome is important to the design of the LRE as it shows the clear structural advantages of the FCC lattice configuration. Utilising the FCC lattice system over a BCC system would allow for a lighter engine dry mass and therefore better thrust to weight ratio.

The second outcome of this analysis was that by varying the density of the lattice structure, the cell slice mass could be optimised for a given internal load and combustion chamber diameter.

The structural analysis was primarily used to create iterations of the engine and lattice system under the expected loading conditions to ensure a working AM system could be developed for future testing.

Fluid Flow Analysis

A series of steady state CFD simulations were conducted to determine the fluid drag properties of a single representative lattice cell one layer in height, within a fluid channel replicating the LRE cooling jacket. The geometry encloses a fluid space of 12 x 10 x 10 mm and the struts vary in size from 0.4mm to 1mm. The analysis considered flow through single and multiple layer lattice structures and was used to determine a per-cell drag for comparison to future experimental results. Further analysis was conducted to determine how a change in the lattice strut diameter would affect the system and wall drag. The simulation modelling process can be seen in Fig. 6 below.

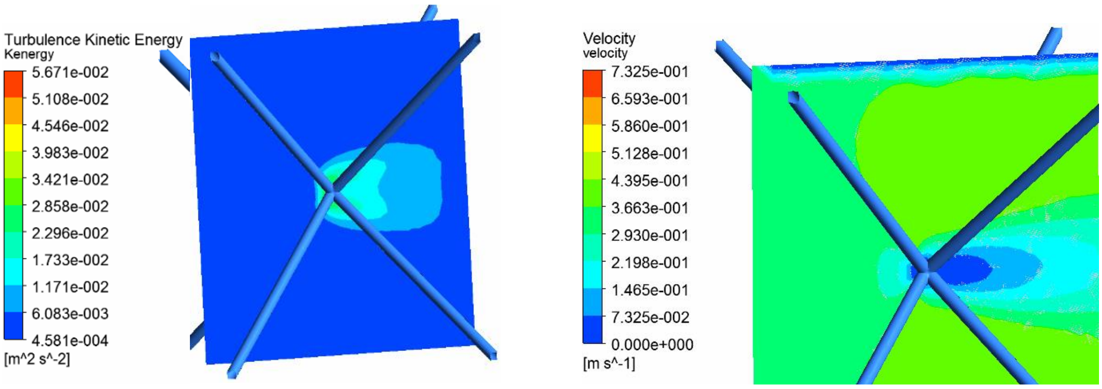


(a) (b) (c) (d) (e)

Fig. 6: CFD modelling process from model recreation (a,b) to fluid domain (c) boundary meshing (d) and boundary conditions (e).

The analysis was conducted with all geometric surfaces emulating a high surface roughness of 50 µm to replicate an as-manufactured SLM part. The contribution of the wall friction to the system drag was in the order of 4% for each of the systems with an increasing contribution as the strut size decreased.

The overall system drag increased by about 10% per 100µm strut diameter increase, caused primarily by an increase in the size of the system recirculation zone, as seen in Fig. 7(a).



(a) (b)

Fig. 7: Flow visualisation plots showing drag inducing phenomenon; Recirculation zones (bluff body drag) (a) and upper and lower wall friction drag (b).

The main contributors to the overall system drag were the high pressure flow stagnation points, zones of flow re-circulation downstream of the structure as seen in Fig. 7(a) and the low velocity boundary flows caused by the skin friction of the upper and lower walls of the system as seen in Fig. 7 (b).

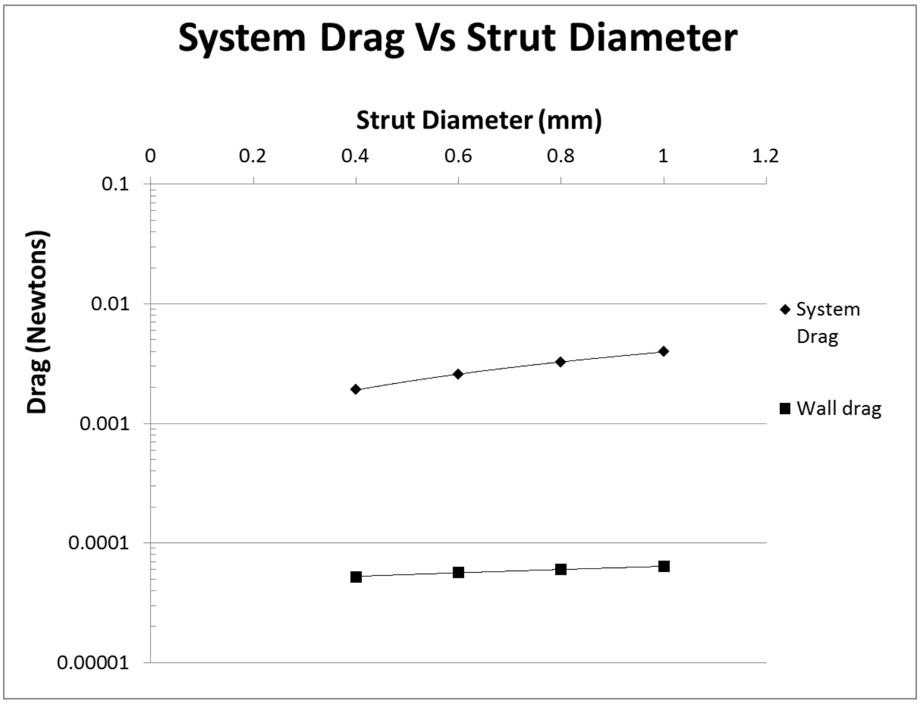


Fig. 8: System and wall drag plot with logarithmic y-axis scale

Fig. 8 shows the combined wall and strut drag (system drag) and the drag associated only with the walls of the system. The plot shows the gradual increase drag caused by an increasing strut diameter and difference in the effects of the strut induced drag compared to the drag caused by the wall friction. Equation 1 shows the relation between system drag to the element strut diameter for the given system:

$$Drag\ (mN) = 3.427 \ast D + 0.525 \qquad (1)$$

Where D is the diameter of the strut in mm

The equation was developed to determine the drag of a given system within the scope of the data and is not accurate when used to calculate the drag of the system outside of these values.

The contribution of the wall effects have been included in all analysis cases so that this data can be applied to multi-layer lattice systems. For example; a three layer lattice system will contain two cells with a single adjacent wall and one cell with no adjacent walls. This can be seen in Fig. 4(b) and (c)

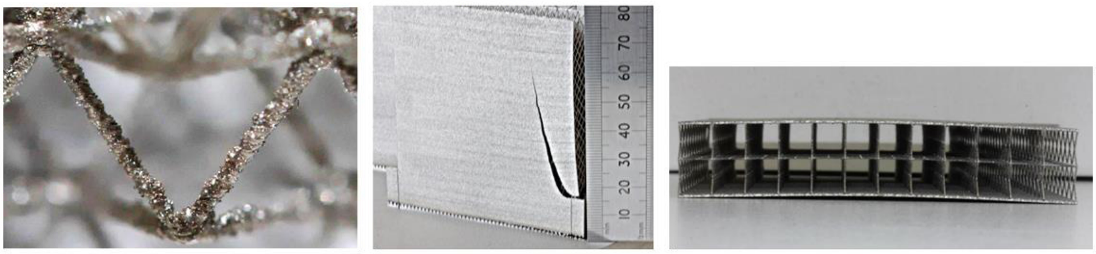
Micro Lattice Integration

There are many advantages to AM which make it very good manufacturing option, and many consider it to be a complete manufacturing solution, however, in reality AM still poses significant challenges. These challenges are different to that of traditional manufacturing, so care must be taken throughout the design and manufacture process.

A key design constraint encountered when developing the lattice systems is the inability of the SLM process to create low angle or overhanging geometries. This requires a part to have a minimum build angle of 30 degrees from horizontal and any build below this angle will likely contain geometric failures that will render the build process to fail. A 45 degree build angle is generally used to ensure a safe margin for most builds.

This is usually not a problem for solids structures as support structure can be added to strengthen part overhangs and removed during post processing of the part. This however, requires tool access to the support material which is not possible if the support material is located with a lattice structure. The unfused metallic powder will be removed using compressed air immediately after manufacture and a series of low pressure fluid tests will be used prior to high pressure flow tests to remove the remaining trapped powder.

As mentioned previously, two faces of the FCC lattice structures were removed to adhere to this build angle problem. These two faces would have a relative build angle of zero degrees and have no supporting structure underneath. As the surrounding unfused metallic material cannot supply any support, the struts along the two faces would fall from the build area and the next layer would have nothing to fuse onto, repeating the failed process and causing a non-uniform lattice structure.



(a) (b) (c)

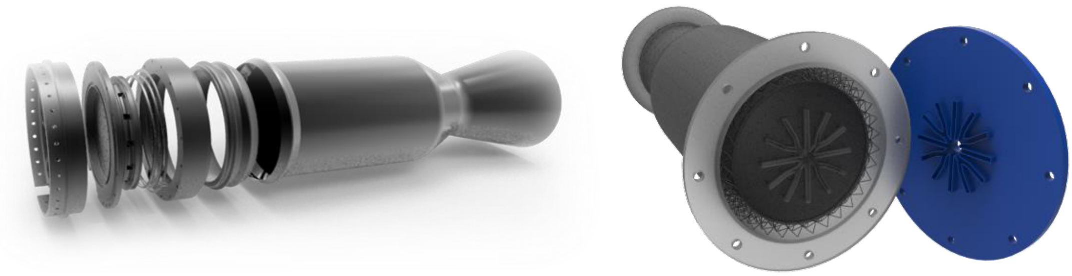
Fig. 9: Lattice surface roughness (a), Surface cracking (b) and warping (c).

Other challenges associated with the development of lattice parts using an SLM process include; the inherent 30 μm roughness as seen in Fig. 9(a), caused by adjacent metallic powder fusing with the molten layers, and surface cracking / deformation as seen in Fig. 9(b) and (c) where the residual system heat generated by the scanning laser has caused undesired geometry motion.

The surface roughness and geometric inaccuracies instigated a rethink in the integration of the injector plate design, specifically around manufacturing the injector orifices.

Injector Design

With the geometric complexity advantages made available by AM it was possible to re-think the way current rocket systems are manufactured. Previously, the rocket engine system would consist of a series of radially bolted, subtractive manufactured parts that each perform a separate function with O-rings used to seal any fluid joints, as seen in Fig. 10. This enabled a single complex system to be split into separate, simple and manufacturable parts.



(a) (b)

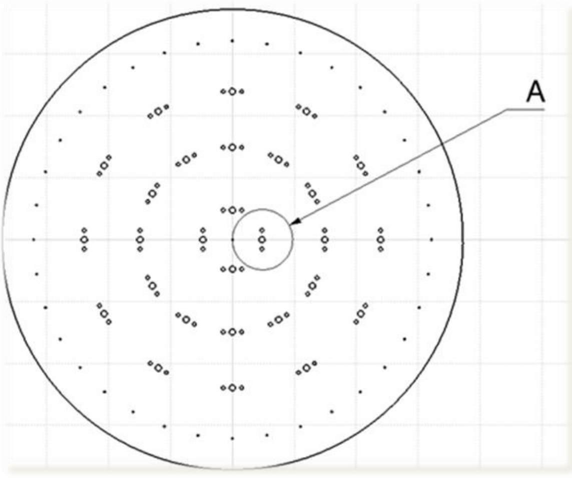
Fig. 10: Traditionally designed rocket system (a) and new, two part system design (b)

The simplification of a system is important in traditional manufacturing methods but cause large amounts of wasted material, and problems associated with high tolerance parts and the sealing of high pressure fluid channels. The new design concept allows for the manufacture of a single piece injector/ combustion chamber system which removes the need for radially bolted joints, fluid sealing and undesirable stress concentration in bolted connections or joints.

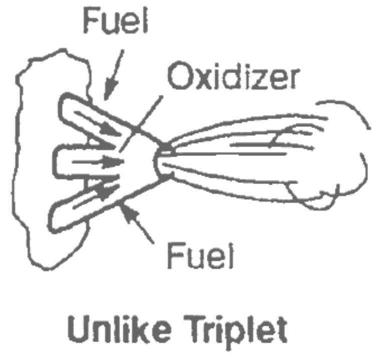
(a) (b)

Fig. 11: Engine segment with integrated injector plate (a) and matching backing plate (b)

An Injector is simply a plate which contains directed orifices which directs the LRE fuel and oxidiser into streams (similar to a shower head), which atomises the fuel when it impinges at a set distance into the combustion chamber, mixing the fuel and allowing for a highly efficient burn. The fluid separation channel has been integrated into the system to ensure there is no mixing of the fuel/ oxidiser prior to injection into the combustion chamber while still allowing each fluid to reach the injector orifices located throughout the injector face. The orifice pattern and fluid separation channel can be seen in Fig. 11(a).
The sizing of the injector orifices was critical to ensuring high combustion efficiency for the system as the type and ratio directly dictates the fuel mixing ratio. An unlike impinging triplet design as seen in Fig. 12(b) was used as it induces no fuel swirl, good fuel/oxidiser mixing and a simple orifice design.



(a)



(b)

Fig. 12: Injector plate orifice layout with channel to ensure separation of fuel and oxidiser (a) and orifice injection pattern (b) [4].

SLM is not able to produce injector orifices with sufficient accuracy [13] to ensure efficient fuel/oxidiser mixing. This is due to the 30 μm particle size that will cause the smallest orifice of 300 μm to vary by up to 20%. It was decided that these orifices were to be manufactured using traditional CNC milling methods after the completion of the printing process. This machining would require access to the back face of the injector requiring the integration of a removable backing plate as seen in Fig. 11(b). This backing plate could be manufactured using the same SLM process but would require machining around the mating fluid separation channel on both the backing plate and injector plate to ensure an adequate fluid seal can be created by the two mating parts.
This would still result in two part system with integrated combustion chamber and injector plate that has not yet been attempted to the author's knowledge.

Directions For Future Work

The work within this paper was limited to the progress of a final year thesis project and, as such has opened many interesting avenues of research but has also left many questions unanswered.
A thorough comparison of the material properties for additive metal creation compared to traditionally produced (and standards controlled) material would have helped the analysis in this paper immensely and is still considered a shortfall when mass production and repeatability is considered.
Optimisation of the lattice geometry is a complex but intriguing expansion to this research especially when the capabilities of the additive manufacturing can allow for some highly complex geometries that were not even considered during this project. The optimisation of the lattice geometry around the engine throat area could substantially increase the structural effectiveness of the system as this is the point where the largest pressure will occur and failure is likely to occur.
Traditional geometries such as simple cylindrical struts typically create a large amount of drag and are considered inefficient, especially in dense fluids. It would be advantageous to implement struts shaped like aircraft wings that would have a lower drag associated per unit area compared to the cylinders used in these analyses.
It was found that a large percentage of the system drag is contributed by the struts and only a small percentage (2%) is contributed by the wall friction, as we know that the amount of heat transfer is closely proportional to the amount of drag associated with the body, a large

improvement in the efficiency of the system can be achieved by optimising the structure to create better heat transfer from the inner wall surface (and adjacent strut geometry) and should be an immediate focus for further work in this subject.

Further optimisation is possible, such as increasing the surface area travelled by the fluid, which would increase the effective enthalpy of the fluid [10]. Introducing a helical flow path into the cooling fluid would increase the fluid travel path, and can readily be achieved with AM. This is however, outside of the scope of this paper due to the complexities that it would incur to the design.

The author acknowledges that the analysis conducted within this paper does not cover all of the possible fluid and thermal effects within the system and a more in depth analysis of each situation is needed to fully understand effects such as non-uniform wall temperatures, inner strut heat transfer and turbulent flow. A macro system based analysis is required to understand important effects in the system that cannot be encompassed by sectional analyses, and would be the subject of future work.

Conclusion

Various lattice geometries were printed using RMIT's SLM 250HL to determine the geometric accuracies of AM printed lattice sandwich structures. It was found that surface roughness, thermal distortion and cracking provided the greatest challenge.

It was discovered that BCC and FCC lattices with the same lattice densities behave differently under the same applied load and that a given lattice structure could be tailored to a specific application, loading conditions and safety requirements. Similarly individual cell slices could be optimised for a given internal load and combustion chamber diameter by varying lattice density, cell size and the strut angle relative to the loading direction.

The overall system drag was found to increase by about 10% per 0.1mm strut diameter increase and was caused primarily by an increase in the size of the system recirculation zone.

The main contributors to the overall system drag were discovered to be the high pressure flow stagnation points, zones of flow re-circulation downstream of the structure and the low velocity boundary flows caused by the skin friction of the upper and lower walls,

however it should be noted that the contribution of the wall friction to the system drag was only in the order of 4% of the total system drag. An equation was developed to relate the System drag to the element strut diameter for the given system.

It was decided that even though FCC lattice cell structures had favourable fluid and structural properties, the higher degree of thermal distortion seen by these structures, caused throughout the SLM printing process, encourages the use of BCC internal lattice structure in a regenerative cooling channel application due to their compliant and consistent deformation nature.

A unique single piece injector concept was developed that allows for a two piece LRE system, which minimises inter-part connections and post processing requirements while still maintaining high combustion and system efficiencies.

The analysis conducted has led to a first iteration design of an AM LRE, ready for initial system testing. There are many possibilities for future analysis and system optimisation for the use of AM in rocket engine design.

Acknowledgements

The Author would like to thank the team at RMIT's Advanced Manufacturing Precinct for their ongoing support and technical experience, Marten Jurg for his assistance in the project and Liam McQuellin (LPM Design) for donating hardware and software support for testing.

References

[1] Stucker, B., D. Rosen, and I. Gibson, *Additive Manufacturing Technologies*. 2010, New York: Springer Science+Business Media.

[2] Bauer, J., et al., *High-strength cellular ceramic composites with 3D microarchitecture*, 2014, PNAS: PNAS.org. p. 2453-2458.

[3] M. Brandt, S. Sun, M. Leary, S. Feih, J. Elambasseril, and Q. Liu, "High-Value SLM Aerospace Components: From Design to Manufacture," *Advanced Materials Research*, vol. 633, pp. 135-147, 2013.

[4] Huzel, D. and D. Huang, *Modern engineering for design of liquid-propellant rocket engines*, ed. R. SeeBass. Vol. 147. 1992, 370 L'Enfant Promenade, SW, Washington DC, 20024-2518: American institute of Aeronautics and Astronautics.

[5] Foli, K., et al., *Optimisation of micro heat exchanger: CFD, analytical approach and multi-objective evolutionary algorithms*. international journal of heat and mass transfer, 2006. **49**.

[6] Roper, C., *Multi-objective optimization for design of multifunctional sandwich panel heat pipes with micro-architecture truss cores*. International journal of heat and fluid flow, 2011. **32**.

[7] Maeda, K. and T. Childs, *Laser sintering (SLS) of hard metal powders for abrasion resistant coatings*. Journal of Materials Processing Technology, 2004. **149**: p. 609-615.

[8] A. Ponomarenko, "RPA – Tool for Rocket Propulsion Analysis," in *Space Propulsion Conference*, Cologne, Germany, 2013.

[9] Quentmeyer, R., *Rocket Combustion Chamber Life-Enhancing Design Concepts*, in AIAA/ASME/SAE/ASEE *Joint Propulsion Conference*1990, AIAA: Orlando,FL.

[10] Cengel, Y., R. Turner, and J. Cimbala, *Thermal-Fluid Sciences*. 2008, New York, NY: McGraw-Hill.

[11] SLM Solutions. (2012, 20 March). *Do you know everything about SLM®?* Available: http://stage.slm-solutions.com/download.php?f=ffe6c35abfc7f4968305f1234442c8ff.

[12] A. Gebhardt, *Understanding additive manufacturing: rapid prototyping, rapid tooling, rapid manufacturing*. Munich ; Cincinnati: Hanser Publishers, 2012.

[13] D. Thomas, "The Development of Design Rules for Selective Laser Melting," Doctor of Philosophy Thesis, University of Wales, 2010.

Robotic swarms as means to autonomously and rapidly characterise small celestial bodies

W. Crowe, N. Kinkaid, J. Olsen and J. Page

School of Mechanical and Manufacturing Engineering, UNSW Australia, UNSW Sydney, N.S.W. 2052, Australia

Summary: Spacecraft swarms have been suggested as an alternative design and operation strategy to improve asteroid characterisation mission costs, timeframes and resilience to failure. However, many of the benefits to swarming have been neglected. The problem of characterising asteroids has been presented, along with how single spacecraft have been used in the past to make measurements and infer properties. The studies on swarms that have already been proposed for asteroid characterisation missions have been outlined. The areas where these can be built upon, as well as further aspects of swarm behaviour that can be utilised are described. This leads on to the work currently being undertaken at the University of NSW.

Keywords: Spacecraft swarms, asteroid prospecting, asteroid mining, orbital mechanics

Introduction

Asteroids (and other small bodies in the Solar System) are of interest to humans because they could possibly destroy life on Earth [1], may hold evidence of the origin of the Solar System [2] and potentially hold resources that can be exploited for space travel and industry [3]. Study in these three areas requires information from a large number of the asteroid population. This will be possible in the near term only if rapid and cheap methods of characterisation are available.

For full asteroid characterisation, a spacecraft must be sent in physical proximity. The current approach is to send monetarily and computationally expensive spacecraft to perform these studies, with vast amounts of data then sent to Earth for human analysis on the ground. In addition to their expense and required intensity of human support, these spacecraft must approach an asteroid slowly in order to protect themselves from unforseen damage.

Spacecraft swarms have been proposed as a tool to characterise many asteroids in a short time [4, 5], with a lower capital cost, reduced size of required support team and with reduced risk as compared with single spacecraft. Spacecraft swarms are groups of simple robotic spacecraft, which interact with one another and the environment using simple rules. Other than in reducing mission risk, swarms can benefit missions through acting as a distributed sensing network and through using emergence of behaviour and swarm intelligence to make autonomous decisions.

This paper will present current methods used for determining the properties of asteroids and how robotic swarms are used in other applications. This will be followed by a review of past concepts and studies on swarming for asteroid characterisation, before synthesising where mission planners may benefit from using simple swarms. The paper then outlines work being undertaken by the School of Mechanical and Manufacturing Engineering at the University of New South Wales.

Characterising asteroids

No simple, concise definition of "asteroid" is available, as these celestial bodies exhibit a diversity of properties in their population as well as similarities with comets, Kuiper-Belt objects and some moons. The International Astronomical Union (IAU) does not have an explicit definition for an asteroid (or a minor planet), except by exclusion from other classes of celestial bodies – planets, dwarf planets, satellites and comets. It has been suggested [6] that only one physical assumption about asteroids is largely true; that the mass of an asteroid is tiny when compared with that of planetary or Solar masses, but very large when compared with visiting spacecraft. Other aspects vary greatly, including elemental compositions, the number of bodies constituting each one (and whether these are contacting, or freely orbiting one another), spin states and vast differences in size and shape.

Most known asteroids have been observed through telescopes, allowing for determination of their orbital elements. Indeed, in order to be designated as a minor planet, observations of orbital elements must be clearly noted and must be dissimilar to those of asteroids that have already been discovered. The majority of asteroids are at such a distance that their shape and many other characteristics cannot be observed from Earth, the exception being if an asteroid comes close (a "fly-by"). Listed in this section are some of the characteristics that *can* be observed. For the measurements from a visiting spacecraft (or swarm) to be beneficial, either the resolution or reliability of measurements of any one characteristic should be equal to, or better than, those that can be readily determined from Earth-based observation.

The determination of asteroidal mass was originally based upon loose estimates of density and size [7], depending upon the material assumed from observed spectography and assumed magnitude. These estimates could be refined after fly-bys of asteroids by spacecraft, where the change in trajectory and distance of closest approach of the spacecraft would be combined to determine mass [8]. More double and multi-bodied asteroid systems have been discovered recently, which can lead to an excellent mass estimate of the system from the orbital period of the smaller body and its distance from the major body [9]. Otherwise, without travelling to an asteroid, the mass parameter remains elusive.

Spectography [10] allows for the surface elements of asteroids to be determined with some accuracy. Many asteroids remain undifferentiated since their formation, so it can be inferred that the properties on the surface are similar to those of the asteroid in bulk, although it is still unclear whether surface coverings of dust are able to form preferentially over time and so obscure the true composition of an asteroid. There is some difficulty with matching the mineralogy of asteroids to meteorites that have fallen to Earth. As such, only loose classification systems [11] have been produced, with their efficacies currently unknown. Sharper measurements have been made with visiting spacecraft. The NEAR-Shoemaker mission used several high-energy sensors [12] to determine chemical abundances, but was unable to provide insight into the exact form of the mineralogical content. The Hayabusa mission was planned to take a sample of an asteroid [13] and return it to Earth. The sample was not taken as designed, but some dust was captured and returned to the Earth's surface for analysis [14].

The shape and size of an asteroid provides clues to its density and gravity asymmetry [15]. This can be attempted from Earth by using the asteroid light-curve, which provides a magnitudinal estimate. Another method of determining these properties can be undertaken if the asteroid flies close to the Earth, where a radar reflection is made off the asteroid surface [16]. The beauty of this method is that it does not matter how the light from the Sun is oriented and many observations can resolve the majority of the asteroid surface. Visiting spacecraft have led to well-determined shapes from intensive analysis of high-quality photos [17] and from LiDAR measurements.

The spin of an asteroid can be determined from Earth using its observed light curve [18]. This method is prone to inaccuracies, especially where applied to a rounded asteroid. If an asteroid flies close to the Earth, a radar reflection can reveal rotation through Doppler shift [16]. Visiting spacecraft can determine how fast an asteroid is rotating by terrain matching [19].

Determining the geomorphology of an asteroid has been limited to scanning surface features in the past, then inferring internal structure [20]. Imaging from rendezvous spacecraft is used to determine where regolith has infilled surface features and covered impact craters [21]. Analytic and simulation techniques are being used currently to investigate how the asteroid gravity, spin [22] and static charge [23] can lead to regolithic dust migration on asteroids, but photographs showing infilling of craters are the main method to determine regolithic depth at different points.

Robotic swarms

Swarms are often observed in nature, where biological “agents” (swarm members) cooperate using simple rules and stimuli to create complex emergent behaviour (that is, behaviour that is not obvious from noting interactions at a smaller scale). Typical examples are the swarming behaviours of creatures such as birds, bees and ants. Robotic swarms have been inspired largely by these biological swarms (biomimetics [24]). An example is in route optimisation, where many agents search for a target, with a pheromone algorithm then making it more attractive for other agents to follow the fastest agent to return (in much the same way as ants find and then reinforce paths to food). More recently, physical laws have been used as inspiration for swarming (physicomimetics [25]). These can take the role of artificial attractive and repelling forces of strengths that vary with distance from the agents. An example is where an obstacle is programmed to repel but a goal to attract and there exists a stable distance between agents that is neither attracting nor repelling. Information is then sent through the swarm when the agents approach obstacles, so that they move around it as a group. However, the agents are still ultimately attracted to the goal (much as liquid in a river will flow towards the lowest point of a slope, but flow around an island).

Swarm intelligence is a way of thinking about machines, not only leading to swarms being homogenous, where every agent is identical, but also heterogeneous, where some agents may have different functions or abilities. One could even extend the logic of swarms to single robots, where subsystems cooperate to produce different functions of the whole. Thinking of a swarm as components of a single unit is intuitive to designers, although development of distributed systems could produce disruptive advancement in some areas [26]. Swarms in the distributed sense are still new, with many possibilities yet untried. There are several differences between single robots and distributed swarm robots where swarms can produce value in mission scenarios. Some of the main examples are described in this section.

Swarms can be more robust to mission failure, as the functions of a swarm as a whole are resilient to partial or complete function loss in single units. Replication of individual functions exists over many units in a fully homogenous swarm or partially exists in a heterogeneous swarm.

Swarm spacecraft can engage in group problem solving, where information from multiple units can be collected centrally, or in a distributed manner and then used to make decisions or process data. This means that agents with varying histories can contribute together to provide a solution that accounts for obstacles and opportunities that may not be obvious to a single decision maker.

Another great advantage of swarms is using their emergent behaviour to solve problems or provide a more resilient structure for accommodating information [25, 27]. Emergent behaviour is that which is not expected to occur, but does and is dependent on configuration and environment. Emergent behaviour can be changed by changing the environment or the rule-based interaction amongst the individual swarm agents. Emergence can assist in the resilience or problem-solving abilities of swarms. Patterns in the way that emergent behaviour evolves in a swarm can lead an observer to make qualitative judgments about the environment that it passes through.

A final way to differentiate between single spacecraft and swarms is through their use as a system of distributed sensors. Distributed sensing has been studied in the past as a way of obtaining many data points simultaneously within an environment, which could provide the network with rich data [28]. This system type also creates robustness for remote sensors, as the data can help account for noise, failure, drift and loss of calibration. When creating this capability between individual agents, there must be a way to address the requirement to process and store more data, as well as which data to save and delete over time.

Proposed swarms for asteroid characterisation

Studies have been carried out by both the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA) in the past in relation to sending spacecraft swarms to asteroids and other celestial bodies [29]. Many papers have been submitted for both studies, although references [30] and [5, 31] are representative for the ESA and NASA efforts respectively. These research efforts have focussed on how to characterise tens or hundreds of asteroids in a short amount of time. They acknowledge the resilience of swarms to destruction or failure of single agents and improved costs per robotic unit when they are homogenous.

The limiting roles of long-range communications and propulsion have been addressed in both studies. These two aspects of spaceflight can be accommodated on large and complex spacecraft, but integration to swarm spacecraft could jeopardise their main benefit of simplicity. The ESA study suggests using a large spacecraft to augment both communications and propulsion, by acting to propel the swarm spacecraft most of the way and then relaying their communications back to Earth. The NASA study suggests having the swarm spacecraft launch from a Lagrange point in order to take advantage of the low launch costs from there and then use the Solar wind and photonic propulsion to propel the spacecraft to their goal. The Solar sails could then be used for orbital station-keeping whilst in proximity to an asteroid, taking advantage of the low gravity environment. Communications would be through using specialised “messenger” agents to ship data physically to Earth orbit and transmit to Earth from there. The messenger method requires that the swarm is fully automated when outside of

Earth orbit. It adds risk by increasing the successful journeys required to two, although it is noted that this is already required for sample return missions.

The ability of swarms to carry many sensors in a highly heterogeneous configuration has been studied by NASA. The ability of swarms to characterise many asteroids due to a distributed search area has been investigated by the ESA team.

There are currently two major asteroid mining companies [32, 33], both of which are investigating swarms for the purpose of characterising asteroids. No complete mission or spacecraft designs have been made public as yet. These have been complemented by researchers from universities, some of whom have publicised the winning of research grants from governments for investigation into the subject.

Several studies are being undertaken to identify possible payloads for NASA's proposed Asteroid Redirect Mission (ARM) [34, 35], including several on nanosat swarms. This means that the swarm will be delivered to the asteroid by the ARM spacecraft and will be able to use the ARM's communication system to correspond with Earth. The ARM will then return to an orbit about the Moon, with the payload potentially accessible to visiting astronauts. Some of the swarms developed for ARM could be used as technology demonstrators for asteroid mining companies. They are envisioned to use distributed sensing systems to determine compositional and geologic features of the visited asteroid.

The gaps in swarming about asteroids

Past studies of swarm missions have identified them as a way of reducing risk through use of multiple units or as distributed sensor networks. The other two areas of value addition, group decision-making and emergent behaviour, have been neglected or under-realised. The four areas have been detailed below.

Robustness to mission failure is an often-identified positive for swarm based missions. This is especially significant when the intention is to move in proximity to an asteroid, where the environment is largely unknown prior to rendezvous and where moonlets, charged dust and highly distended shapes are common and hold the potential for interference with spacecraft. Avoidance of these dangers has been a source of complexity, time and, subsequently, cost in previous missions. This is because a single spacecraft must slowly descend to acquire iterative knowledge about the environment surrounding its target. Further time can be added to avoid areas of specific danger regions of the asteroid visited once they are identified. This can involve moving through places of outgassing (in the case of Rosetta mission) over the asteroid surface in proximity (when desired, this is performed last in a mission, to prevent risks to other tasks, for example NEAR/Shoemaker mission), and avoiding debris from impacts (in the case of Hayabusa 2 mission).

There may be a secondary advantage to the redundancy that leads to robustness; in that purposefully crashing one of the agents into an asteroid has the potential to present data that may not be possible to obtain otherwise. For example, NASA's Lunar Crater Observation and Sensing Satellite (LCROSS) mission used an almost spent rocket stage to impact one of the moon's polar craters and release dust and vapour, proving that water ice does in fact exist at the lunar poles [36]. A similar impact on an asteroid could expose the sub-regolithic portion of the asteroid or throw up dust and vapour to be analysed by other spacecraft in the swarm. Such a process may also be able to obtain seismic data if a detector had previously been deployed on the surface of the asteroid. The Hayabusa-2 spacecraft mission is set to launch an impactor into the asteroid that it visits. This will achieve the same purpose, but with a specialised impacting pin [37]. Although designed specifically for this function, this system increases mission complexity and has added to the cost of the Hayabusa 2 spacecraft when compared with its predecessor, Hayabusa 1. The spacecraft can be made to act as penetrators, useful in the case of swarms, as several instruments could be deployed over many points of interest. Past and future penetration strategies to date can be found in Lorenz' work [38]It is noted that due to the unreliability of penetrators, especially in unexplored ground, swarms provide the opportunity for many attempts to be made of penetrating the asteroid in a number of dispersed locations. The unknown nature of asteroid formation and the strength of adhering forces, means that this type of research could provide significant progress in asteroid geology and planetary formation sciences. The paper cited suggests that the VESTA mission has been crafted in the past for two penetrators to be launched to an asteroid, whilst the mothership flew by en route to Mars, an early and small-scale example of swarm missions.

Collective problem solving has been investigated in asteroid swarm studies to the extent that data from each agent of the swarm is sent to a central decision-maker. This method, while taking advantage of a greater range of information, creates an information bottle-neck and weak point, where mission success can be reliant on a single spacecraft and requires significant data transport. One area where this can be improved is through collective decision-making, or localised decision-making, where agents can come to their own decisions based upon information from neighbouring agents. Distributing information throughout the swarm can allow it to be retained in a robust manner, important in maintaining a robust mission function.

At this point in time, emergent behaviours of swarms about asteroids have received little attention. The emergence of behaviour in a simple system, with well-defined variables, is difficult to measure. However, when the system becomes more complex, emergence can be especially useful when the desired amount of data may be small and use of more data will not necessarily provide greater insight to the problem. The gravity about asteroids is commonly subject to shape and mass distribution asymmetries. This produces an environment that can lead to chaos in three spatial dimensions. The full gravity field can be resolved by a single spacecraft, but this is for the sake of using numerical techniques to infer other details about the dynamics properties of orbital vehicles. Using the swarm to find dynamics of interest is a way of reducing the requirement of information transfer by transmitting only the data desired. As an example, a set of spacecraft orbiting an asteroid, with an arbitrary distribution as the initial conditions, there is a tendency for those spacecraft about unstable orbits to move to orbits of stability or escape, whilst those about stable orbits will remain there. To speed the creation of this emergent behaviour, spacecraft can have rules applied where they fire thrusters to act to destabilise orbits. Use of emergent behaviour in the near-asteroid environment is one of the main investigations being carried out currently at the University of NSW, with figure 1 showing a simulation tool that illustrates the abovementioned example.

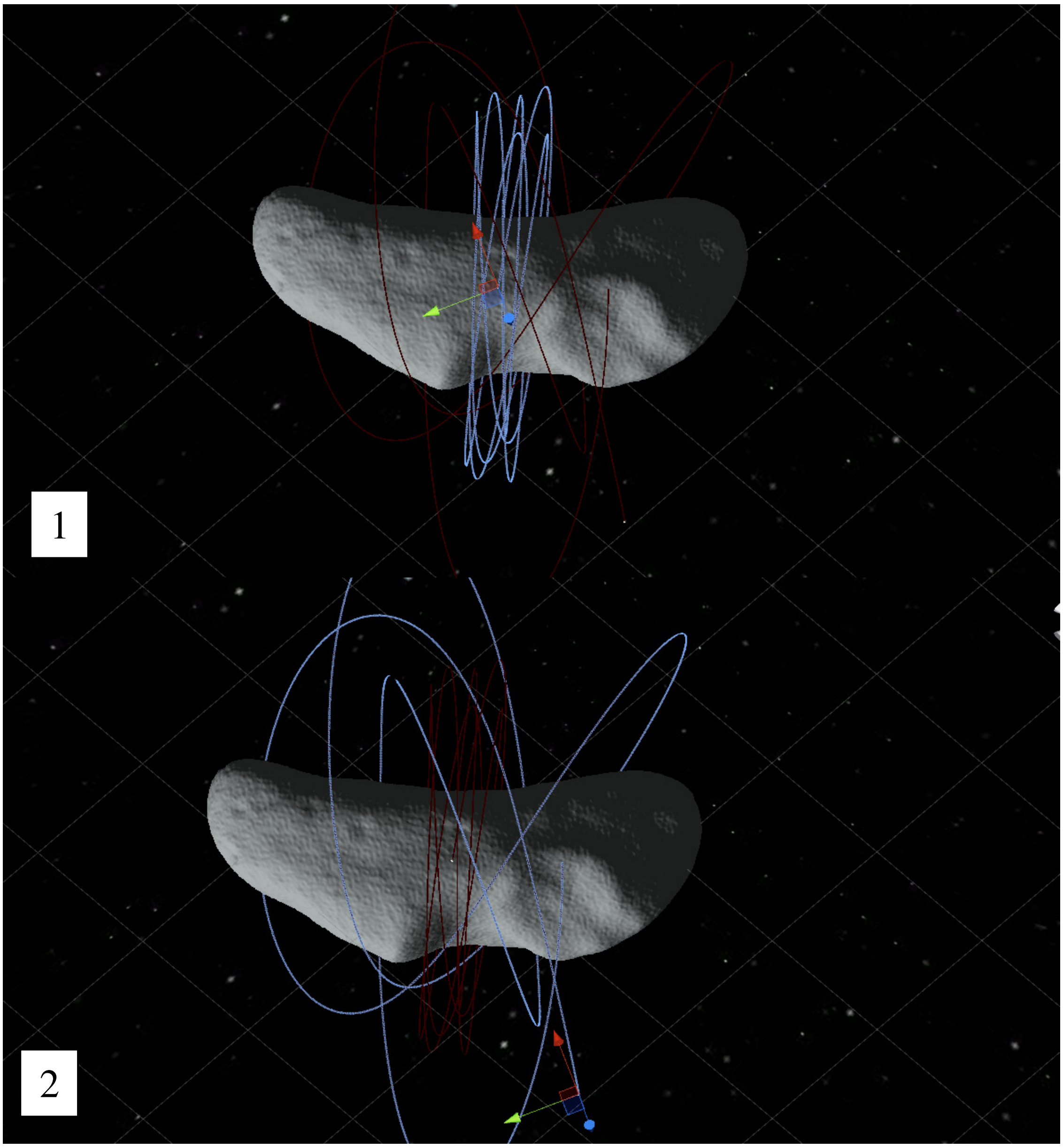


Figure 1: Two spacecraft are shown here orbiting a model of the asteroid 433 Eros. The first spacecraft orbits with some stability, but the second, offset from the first by several kilometres, an unstable orbit is exhibited which slowly moves to orbit the most stable axis, orthogonal to the initial axis.

The environment about asteroids is highly variable, that is, environmental interactions can be very different at close intervals. This means that information from single spacecraft must be collected rigorously and globally. Swarms of spacecraft have the capacity to maintain constant global coverage of an asteroid, allowing for a full picture of the state of an asteroid to be picked up at any time. A sensor network like this can act to self-calibrate and help distinguish between noise and actual data, meaning there can be increased simplicity of sensors and thus increased simplicity of spacecraft. The authors envision that this would be performed in many local networks as a self-organised swarm, to reduce the requirement to centralise data transfer.

Designing swarms to accommodate all of these areas allows for a more complete leveraging of the possibilities, and thus to a reduction of the complexity of the individual swarming robotic spacecraft. Although simplicity conflicts with normal deep-space spacecraft development, it provides great opportunity for reduction in costs, human input and time of development.

UNSW's School of Mechanical and Manufacturing Engineering has worked on swarms for aviation [39], robotics [40] and spaceflight [41] in the past. Progress has been made into visualising swarms in simulations via virtual reality. A set of tools for simulating swarms in the vicinity of asteroids has been produced and tested, using both Matlab™ and Unity™ which have proven physics capabilities. Some preliminary simulations have been carried out for proof of concept, with more rigorous simulations to model specific swarming scenarios now set to be run [42]. Through simulation, a set of results will help show when it is most beneficial to characterise asteroids using simple swarms. As secondary objectives, it will be found how best to swarm to collect desired data and what decisions must be made during swarming.

Conclusion

When characterising asteroids for the purpose of mining, swarms of robotic spacecraft may be used, and in many ways are more beneficial to use, over conventional methods. The data required in characterisation is still unknown for this future application, so the limits to simplicity cannot yet be fully explored, although it is useful to find what data sets can be produced with current technology and to what quality.

An emphasis should be placed on taking advantage of all of the capabilities that swarms have to offer. This does not mean neglecting the capabilities that have already been suggested through reports and papers in the past, but taking account of the activities that can only be determined through simulation.

Work is ongoing at the University of NSW to identify how best to swarm about asteroids, using simple spacecraft to infer data.

References

1. Lu, E. *The project b612 concept*. in *Collection of Technical Papers—2004 Planetary Defense Conference: Protecting Earth from Asteroids, American Institute of Aeronautics and Astronautics Inc., Reston, VA, USA*. 2004.
2. Glassmeier, K.-H., et al., *The Rosetta Mission: Flying Towards the Origin of the Solar System.* Space Science Reviews, 2007. **128**(1-4): p. 1-21.
3. Sonter, M.J., *The technical and economic feasibility of mining the near-earth asteroids.* Acta Astronautica, 1997. **41**(4): p. 637-647.
4. D'Arrigo, P. and S. Santandrea, *APIES: A mission for the exploration of the main asteroid belt using a swarm of microsatellites.* Acta Astronautica, 2006. **59**(8-11): p. 689-699.
5. Curtis, S.A., et al. *ANTS for Human Exploration and Development of Space*. in *Aerospace Conference, 2003. Proceedings. 2003 IEEE*. 2003.
6. Scheeres, D.J., *Orbital Motion in Strongly Perturbed Environments.* Orbital Motion in Strongly Perturbed Environments, by Scheeres, Daniel J. ISBN: 978-3-642-03255-4. Berlin: Springer, 2012, 2012. **1**.
7. Carry, B., *Density of asteroids.* Planetary and Space Science, 2012. **73**(1): p. 98-118.
8. Yeomans, D., et al., *Estimating the mass of asteroid 253 Mathilde from tracking data during the NEAR flyby.* Science, 1997. **278**(5346): p. 2106-2109.
9. Margot, J.L., et al., *Binary Asteroids in the Near-Earth Object Population.* Science, 2002. **296**(5572): p. 1445-1448.
10. Abell, P.A., et al., *Mineralogical composition of (25143) Itokawa 1998 SF36 from visible and near-infrared reflectance spectroscopy: Evidence for partial melting.* Meteoritics & Planetary Science, 2007. **42**(12): p. 2165-2177.
11. Bus, S.J. and R.P. Binzel, *Phase II of the Small Main-Belt Asteroid Spectroscopic Survey: A Feature-Based Taxonomy.* Icarus, 2002. **158**(1): p. 146-177.

12. Cheng, A.F., et al., *Near-Earth Asteroid Rendezvous: Mission overview.* Journal of Geophysical Research: Planets (1991–2012), 1997. **102**(E10): p. 23695-23708.
13. Yano, H., et al., *Touchdown of the Hayabusa spacecraft at the Muses Sea on Itokawa.* Science, 2006. **312**(5778): p. 1350-3.
14. Nakamura, T., et al., *Itokawa Dust Particles: A Direct Link Between S-Type Asteroids and Ordinary Chondrites.* Science, 2011. **333**(6046): p. 1113-1116.
15. Scheeres, D., *Orbital mechanics about small bodies.* Acta Astronautica, 2012. **72**: p. 1-14.
16. Ostro, S.J., et al., *Asteroid 4179 Toutatis: 1996 Radar Observations.* Icarus, 1999. **137**(1): p. 122-139.
17. Miller, J.K., et al., *Determination of Shape, Gravity, and Rotational State of Asteroid 433 Eros.* Icarus, 2002. **155**(1): p. 3-17.
18. Masiero, J., et al., *The Thousand Asteroid Light Curve Survey.* Icarus, 2009. **204**(1): p. 145-171.
19. Veverka, J., et al., *The landing of the NEAR-Shoemaker spacecraft on asteroid 433 Eros.* Nature, 2001. **413**(6854): p. 390-393.
20. Prockter, L., et al., *Surface Expressions of Structural Features on Eros.* Icarus, 2002. **155**(1): p. 75-93.
21. Veverka, J., et al., *Imaging of Small-Scale Features on 433 Eros from NEAR: Evidence for a Complex Regolith.* Science, 2001. **292**(5516): p. 484-488.
22. Liu, X., H. Baoyin, and X. Ma, *Dynamics of surface motion on a rotating massive homogeneous body.* Science China Physics, Mechanics and Astronomy, 2013. **56**(4): p. 818-829.
23. Hartzell, C.M. and D.J. Scheeres, *The role of cohesive forces in particle launching on the Moon and asteroids.* Planetary and Space Science, 2011. **59**(14): p. 1758-1768.
24. Fisher, L., *The perfect swarm: the science of complexity in everyday life*. 2009: Basic Books.
25. Spears, W.M., et al., *An overview of physicomimetics*. 2005: Springer.
26. Barnhart, D.J., T. Vladimirova, and M.N. Sweeting, *Very-Small-Satellite Design for Distributed Space Missions.* Journal of Spacecraft and Rockets, 2007. **44**(6): p. 1294-1306.
27. Hinchey, M.G., R. Sterritt, and C. Rouff, *Swarms and Swarm Intelligence.* IEEE Computer, 2007. **40**(4): p. 111-113.
28. Brooks, R.R., C. Griffin, and D.S. Friedlander, *Self-Organized Distributed Sensor Network Entity Tracking.* International Journal of High Performance Computing Applications, 2002. **16**(3): p. 207-219.
29. Truszkowski, W., et al., *Swarms in Space Missions*, in *Autonomous and Autonomic Systems: With Applications to NASA Intelligent Spacecraft Operations and Exploration Systems*. 2010, Springer London. p. 207-221.
30. D'Arrigo, P. and S. Santandrea, *The APIES mission to explore the asteroid belt.* Advances in Space Research, 2006. **38**(9): p. 2060-2067.
31. Clark, P., M. Rilee, and S. Curtis. *PAM: Biologically inspired engineering and exploration mission concept, components, and requirements for asteroid population survey*. in *55 th International Astronautical Congress*. 2004.
32. Industries, D.S. *Deep Space Industries*. 2014 [cited 2014 2014-11-4]; Available from: http://deepspaceindustries.com/.
33. Resources, P. *Planetary Resources - The asteroid mining company*. 2014 [cited 2014 2014-11-4]; Available from: http://www.planetaryresources.com/#home-team.
34. Brophy, J.R. and B. Muirhead. *Near-Earth Asteroid Retrieval Mission (ARM) Study*. in *33rd International Electric Propulsion Conference, Washington, DC, October 6–10, 2013*. 2013. Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration, 2013.
35. Moore, C., *Technology development for NASA's Asteroid Redirect Mission.* 2014.
36. Schultz, P.H., et al., *The LCROSS Cratering Experiment.* Science, 2010. **330**(6003): p. 468-472.
37. Tsuda, Y., et al., *System design of the Hayabusa 2—Asteroid sample return mission to 1999 JU3.* Acta Astronautica, 2013. **91**(0): p. 356-362.
38. Lorenz, R.D., *Planetary penetrators: Their origins, history and future.* Advances in Space Research, 2011. **48**(3): p. 403-431.
39. Chi, T., et al., *Evolving swarm of UAVs.* 2014.
40. Chi, M.T.-C., M.J. Page, and J. Olsen, *Simulation of Search and Track Vehicles.* Message from the SimTecT 2013 Papers Committee, 2013: p. 174.
41. Cheng, H., *Behaviour-Based Control Law for Spacecraft Swarm Operation Haoyang Cheng (, John R. Page (2), Nathan M. Kinkaid (3) and John F. Olsen (4).*
42. Crowe, W., et al. *Simulation of spacecraft about collinear Lagrange points of small bodies in the Solar System*. in *SimTect 2014*. 2014.

Assessing the viability of a Harpoon as an Active Space Debris Removal Device

Roger Dudziak [a], Sean Tuttle [a], Simon Barraclough [b]

[a] *School of Engineering and Information Technology, University of New South Wales, Australian Defence Force Academy Campus, Canberra, Australia*
[b] *Airbus Defence and Space Ltd, Stevenage, UK*

This paper presents the results of empirical testing and numerical modeling carried out to demonstrate the effectiveness of a harpoon as an Active Debris Removal (ADR) device. The parameters tested include: the relationship between tip shape and both ballistic limit and the creation of secondary debris; the ability to penetrate targets at oblique impact angles, low temperature and with heat pipe obstructions; the ability to lock onto targets post-penetration and withstand the loads expected during de-orbiting maneuvers; and the effect of tip shape on the penetration of CFRP targets. Testing involving the impact of blunt and conical shaped steel tips into 3mm aluminium (Al) plate showed that the ballistic limit varies in proportion to the tip circumference, with conical shapes having a higher relative ballistic limit due to the additional energy required for petalling. In regards to secondary debris, it was found that blunt shapes created a plug during penetration as a result of shearing around the periphery of the projectile, whilst conical tips resulted in minor spalling and fragmentation. Preliminary oblique impact testing up to 40° showed that the ballistic limit increases with obliquity at a greater rate for blunt tips than conical ones. This was supported by simulations up to a 60° impact angle. Impact testing of 3mm Al plate with conical projectiles at low temperatures showed a more brittle fracture mode when compared with targets impacted at room temperature. This resulted in a cleaner fracture surface and an increased ballistic limit. Impact testing of Al panels obstructed with fixed heat pipes showed that the harpoon could successfully penetrate a target panel with such an obstruction due to shearing of the pipe flange. Testing of two lock-on mechanisms showed that both a spring activated and integrated toggle could reliably open on target penetration. Tensile load testing was also conducted and showed that both designs could withstand the loads expected during de-orbiting maneuvers, with the integrated toggle being more robust. Simulation was used to evaluate the effect of varying the diameter of a conical tip on the ballistic limit. The results showed that the ballistic limit increased with diameter. A Smooth Particle Hydrodynamics (SPH) solver, which is better suited to modeling impact into brittle materials, was used to model impact into CFRP targets. This showed that, in comparison with blunt tips, conical tips had a higher ballistic limit. In addition, the debris formed was less coherent and had a higher terminal velocity. Further simulation showed that numerical modeling can provide an accurate prediction of the ballistic limit for Al plate impacted by conical projectiles, excellent prediction of impact failure modes, good predictions of debris likely to be created during the impact process, and an efficient means of testing prototype tip shapes. Finally, preliminary studies of harpoon impact into low mass targets with up to 3 DOF showed that the harpoon was still able to penetrate such targets under simulated μ-gravity conditions.

I. Introduction

Current estimates suggest that approximately 5,500 metric tons of space debris, comprised of spent rocket bodies, decommissioned satellites, and other disused material is orbiting the Earth [1]. Apart from the obvious threat that this quantity of debris presents to both manned and unmanned equipment in orbit, an additional threat of uncontrolled, cascading collisions, commonly referred to as the Kessler syndrome [2], also exists. The harpoon is one of several Active Debris Removal (ADR) systems currently being investigated by the ESA to interrupt this potential cascade of collisions by removing large, uncontrolled debris [1]. The research undertaken as part of this project aims to contribute to studies being conducted by Airbus Defence & Space on the viability of using a harpoon in an ADR role.

In order for a harpoon to be used successfully in an ADR role it must meet several requirements. The first of these includes being able to penetrate a target with minimal energy and creation of secondary debris. In regards to the former, minimising required impact energy reduces the propellant needed to launch the harpoon. This reduces the overall mass of the ADR system and would not be insignificant in the likely event that multiple harpoons are launched on a single spacecraft. In addition, as will be identified in this paper, reducing impact energy also reduces the creation of secondary debris. This not only reduces the risk of adding to the space debris problem, but also damaging vulnerable equipment within the target. This could include fuel and heat pipes, as well as pressure vessels and batteries. Penetration of any of this equipment could lead to fragmentation or explosions [3]. Performance against this requirement was assessed by undertaking experimental tests to establish a relationship between tip shape, ballistic limit and secondary debris creation. In addition, tip parameterisation was undertaken using simulation methods to reduce the quantity of experiments required.

The second requirement identified was that the harpoon must be capable of penetrating a range of target materials under a variety of conditions and in different configurations. The materials include those that are typically found on solar arrays (CFRP, and 20mm Al core with 0.25mm CFRP skin), and equipment panels and payload adaptors (3-5mm Al panels or Al Sandwich panels with 0.5mm thick skin) [1]. Whilst penetration of Al and Al honeycomb (Al H/C) panels was tested experimentally, due to the limited availability of CFRP, the effectiveness of different tip shapes against this material was compared using simulation. In regards to target conditions, its temperature could vary significantly depending on its orientation in relation to the sun at the time of impact. This could have a significant effect on the properties of the target material. As a result, impact testing on Al targets cooled to very low temperatures was conducted to determine the effect of extreme cold on the ballistic limit of Al plate as well as the creation of secondary debris.[1] In regards to target configuration, this relates to panels which may be obstructed by equipment boxes and/or attached heat pipes. Heat pipes in particular are a significant consideration as they typically have a 3mm wall thickness (Astrium, 2012). As such, impact tests were conducted on target panels fitted with representative heat pipes to determine the effect of such obstructions on harpoon penetration. Target configuration also relates to its orientation at the time of impact. As potential targets could be spinning or tumbling, impact is likely to occur at an oblique angle. Limited impact testing was therefore conducted on targets up to an angle of obliquity of 40°. This limited experimental testing was supplemented by simulating impacts up to an angle of 60°.

The final requirement identified relates to the lock-on mechanism used to securely attach the harpoon to the target to facilitate stabilisation and de-orbiting maneuvers. The forces generated during this phase of the mission are expected to average approximately 500N, with a peak as high as 2-3kN in the event of a high thrust de-orbit manoeuvre [1]. Testing of two different lock-on mechanisms to confirm reliability of deployment as well as load capacity was therefore conducted.

Of note, the empirical impact testing discussed in this paper was designed to simulate impact into debris representative of that expected to be targeted as part of a harpoon ADR mission. These targets, such as large satellites and rocket bodies, are likely to have high inertia properties relative to the harpoon. As such, the majority of testing used fixed targets and the effect of impacting a target along an axis not coincident with its CoG was not considered. Empirical testing of harpoon impact into low mass targets in simulated μ-gravity was undertaken as a preliminary study only, the results of which are discussed in section H.

II. Experimental Set-Up

The UNSW medium velocity gas gun was used to conduct the experimental component of this study. This gas gun was chosen as it was able to impart the estimated energy required to allow the projectile to penetrate the thickest target configuration tested. Projectile speed was initially measured using two sheets of Al foil spaced a given distance apart and connected to an Arduino microcontroller. The timer was triggered as the projectile broke through each piece of foil. This method proved to be highly inaccurate however due to the inconsistent flexing of the foil pieces prior to projectile penetration. As a result, a modified speed sensor using two infrared

[1] Temperature of orbiting objects such as the ISS varies between -128°C in eclipse to 93°C in direct solar radiation, http://www.nasa.gov/content/cooling-system-keeps-space-station-safe-productive/#.U3BVvvmSyuI, viewed on 10 May 14

sensor/transmitter pairs was developed. This enabled the time of flight between the two sensor pairs to be measured to a resolution of 1×10^{-6} sec. This set-up was positioned at the bore exit and is detailed in Fig. 1.

All target plates used during the impact experiments were 250mm^2 in dimension and varied in thickness. These were secured to a target box which was positioned on a steel frame. In order to decrease experimental uncertainty due to target movement upon impact, the plates were attached to the target box with two 6mm thick Al holders which were bolted to a 4mm thick steel adapter plate. This adapter plate was secured to the box with M10 bolts. The box itself was filled with clay and bolted to the stand. In order to further reduce the chance of movement, solid steel plates were placed behind and on the target stand. This set-up can be seen in Fig. 2. This configuration was later modified for impact testing of obstructed panels. The target box was removed and the target adapter plate fixed directly to a steel frame. This configuration allowed high speed video footage to be taken of the rear of the target during impact and therefore the nature of heat pipe failure to be recorded.

The harpoon used for experimental testing was manufactured from 10mm diameter mild steel rod. Several iterations were developed to improve reliability against failure. Whilst the initial version was modular to allow replacement of harpoon segments in the event of failure, or to change its length, it was found that this design actually contributed to failure on impact. The final version was therefore manufactured from a single threaded rod. The thread enabled the fitting of different attachments such as non-discarding sabots as well as damper systems. The end of the harpoon was threaded to allow for the fitting of different tips. These tips were also manufactured from mild steel and were drilled and tapped with a 3/16 thread. This thread size was chosen as it corresponded to that used in the *Ramset* toggle trialed as a lock-on device. The harpoon and tips are shown in Fig. 3.

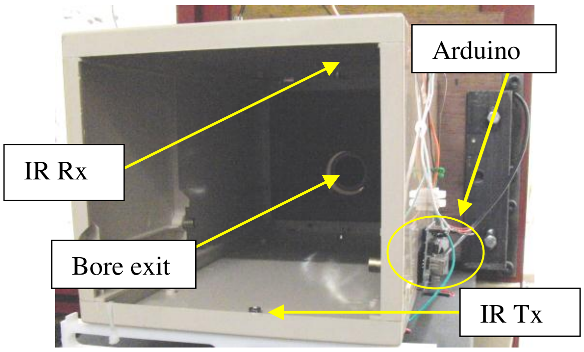


Figure 1. Sensor used to measure harpoon speed.

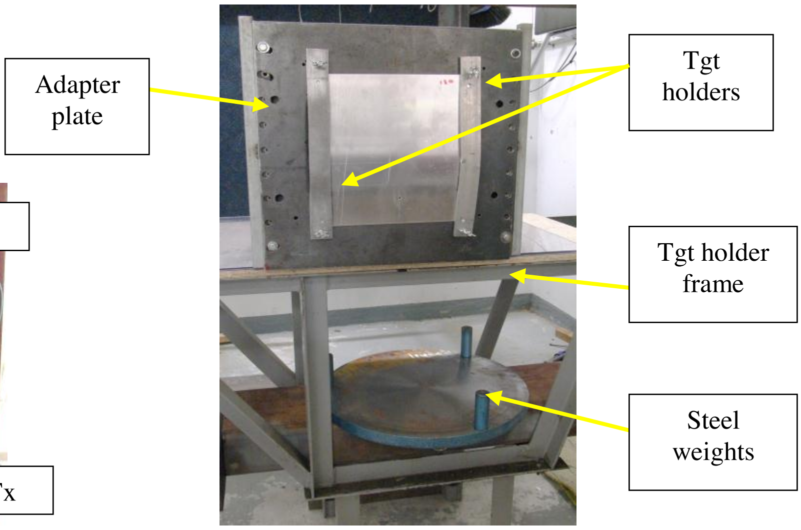


Figure 2. Target holder set-up used for ballistic limit and oblique impact testing.

Figure 3. Harpoon used for experimental testing (left) and tips used for ballistic limit comparison (right).

III. Results and Discussion

A. Ballistic Limit Testing

The ballistic limit of 3mm, Al5005H34 plate for both conical and blunt tip shapes was determined experimentally. Specifically, the purpose of this testing was to identify which tip shape had the lowest limit as, in addition to reducing the propellant required, minimising the ballistic limit will reduce the shock-induced stresses in the target material. This is important as it reduces the risk of spalling and therefore further debris creation. Whilst spalling is commonly associated with the tensile stresses resulting from the reflection of shock waves from the free surface of a target [4], it can also occur through a process known as ductile spall damage.

This is particularly relevant for Al. This process leads to the development of spall as a result of void nucleation and growth [4]. As the likelihood of damage is a function of the magnitude of the shock-induced stress in the target material, reducing the impact energy is likely to reduce the onset of ductile spall damage.

Potential errors associated with the testing regime include deflection of the target box and adapter plate as a result of the high impact energy, and the anisotropic nature of the target material. In regards to the latter, yield strength will vary depending on the orientation of the grains in a given sample. In regards to deflection, as discussed in the introduction, this was minimised by mounting the target securely to the target box. Resultant deflection was recorded using a high speed camera at 1200 frames per second. From the footage it could be seen that deflection occurred post impact rather than during penetration as a result of shock wave propagation through the target and box, as well as the inertia of the target set-up. As a result, the effect on recorded ballistic limits is evaluated as minor.

The results of the experimental ballistic limit testing, including uncertainty ranges, are detailed in Table 1. The upper limit corresponds to the velocity at which complete penetration occurred, while the lower limit represents the closest recorded velocity below this value at which full penetration did not occur. From these results it can be seen that, with the exception of the cone, the ballistic limit is proportional to the circumference of each tip.

Table 1: Ballistic Limits for 3mm Al5005H34 Plate Impacted by Different Harpoon Tip Shapes

Tip Shape	**Dimensions (mm, deg)**				**Ballistic Limit (m/s)**	**Kinetic Energy (J)**	$\left(\frac{E_{i,ave}}{E_{cone,ave}}\right)$
	Dia.	Length	Angle	Circ.			
Cone	16	18	24	49	42.9 - 44.2	206.8 - 220.3	1.0
Cube	N/A	16 (side length)	N/A	64	40.9 - 42.2	204.1 - 217.1	0.99
Hex	N/A	16 (across flats)	N/A	54	38.6 - 40.3	190.41 - 207.8	0.93
Cylinder	16	11.5	N/A	50	33.3 - 35.5	128.4 - 145.8	0.64

The fact that the ballistic limit of the cone deviates from this trend can be explained by the difference in failure modes of blunt versus conical tips. Perforation of plates by blunt or rounded projectiles 'usually involves punching out a plug resulting from shear failures that initiate along or near the boundary of the impact surface' [5], however, perforation by conical or ogive tips results in piercing, crack propagation and petaling [6]. As the latter involves high radial and circumferential tensile stresses [7], more energy is required.

B. Secondary Debris Creation

As the intent of the harpoon system is to remove debris from low earth orbit, this study also included an assessment of the relative amount of debris created by each tip shape. From this testing, it was confirmed that impact of thin Al targets with conical tips can produce ductile spalling and minor fragmentation. Impact by blunt tips can produce plugs, and occasionally minor fragmentation. Of note, any debris produced during the experimental impact tests was opposite to the side impacted, which would typically be internal to the debris object. Examples of the debris produced by blunt and conical tip shapes are given in Fig. 4.

Figure 4. Plugs produced from impact of 3mm Al5005H34 plate by blunt projectiles (left) noting that dimensions match those of respective tips given in table 1. Spall (centre) and fragmentation (right) produced from impact of 3mm Al5005H34 plate by conical and blunt projectiles respectively at 8x magnification. Maximum spall diameter (centre) is 2.6mm. Fragmentation dimensions (right) are 15.2mm long, 0.6mm wide.

Of note, in regards to impact into spacecraft, unless the impact site is a radiator panel or solar array most of the structure will be covered with MLI and this has been shown to help trap inside the satellite any small particles created during the harpoon impact [8].

C. Low Temperature Impact Testing

Although published data was not available for Al5005H34, the tensile strength of Al2024, commonly used in space applications, has been shown to increase by approximately 276MPa over a temperature range of room temperature to -269°C. In addition, the % elongation and ductility of some Al alloys decreases with a significant reduction in temperature [9]. The reason for this testing therefore was to investigate the effect of low temperature on the ballistic limit and, more importantly, to investigate whether the increased brittleness of Al at these temperatures would increase the risk of secondary debris creation.

The low temperature testing was conducted on 3mm Al5005H34 plate using a conical tip. The target material was cooled to -171°C by immersion in liquid nitrogen before being impacted by the harpoon with the temperature of the target being monitored with a K-type thermocouple. Due to the time delay between removing the target from the liquid nitrogen and firing the gun, the target temperature increased prior to impact. It is estimated that an uncertainty of approximately 5°C exists in the impact temperature readings. This is due to the fact that only one thermocouple was placed at the base of the target and uniform temperature across the target plate cannot be assumed. The recorded impact temperatures as well as the results of this experiment are detailed in Table 2.

Table 2: Results From Impact of a Conical Projectile Into 3mm Al5005H34 Plate at Low Temperatures

Test	Impact Energy(J)	Impact Temperature(°C) +/-2°C	Perforation diameter(mm) +/-0.5mm
1	247	-104	14.5
2	255	-83	16
3	255	-119	13.5
4	265	-102	16
5	265	-86	Complete penetration
6	148	-107	12.6
7	241	-130	13.6
8	217	-138	12.2
9	202	-136	12.7
10	202	-132	13.1
11	224	-80	14

From these results it can be seen that for the same impact velocity e.g., data points four versus five and nine versus ten, the penetration depth decreases with decreasing temperature. In addition, although the impact velocity for test four was 20% higher than the ballistic limit for the Al impacted at room temperature, complete penetration did not occur. This provides good evidence that the ballistic limit of Al plate increases with low temperature. This correlates well with the increase in strength at low temperatures noted in the aforementioned research.

As well as increasing the ballistic limit, cooling Al plate to low temperatures resulted in reduced spall and fragmentation of the target plate. This difference is likely to be the result of the reduction in ductility with temperature of the sample. This reduction in ductility reduces the likelihood of ductile spalling mentioned earlier. Photographs of two impacted samples, taken at 8 x magnification, show a clear difference in the fracture surface of Al plate impacted at room temperature in comparison to that impacted at low temperatures. The sample shown in Fig. 5 was impacted at room temperature at a velocity of 42.2m/s resulting in a perforation diameter of 14.5mm. The sample shown in Fig. 6 was impacted at -119°C at 47.3m/s resulting in a perforation diameter of 13.5mm. Fig. 5 shows the onset of spalling (arrow) as well as the formation of fragmentation (circled), however, Fig. 6 shows relatively clean fracture surfaces. From this testing it can therefore be concluded that, whilst impacting Al alloy targets at low temperatures may require a slight increase in impact energy, the potential for additional debris formation may actually be reduced.

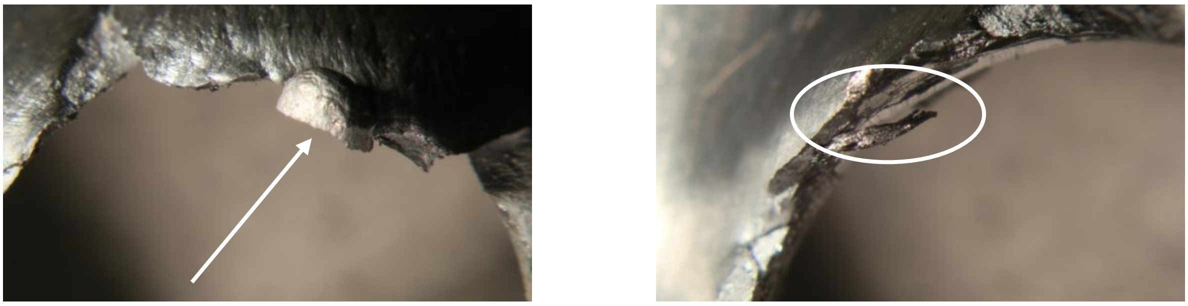


Figure 5. Fracture surface of 3mm Al5005H34 plate impacted at room temperature at 8 x magnification (left). The onset of spall formation is evident in centre right of image. Fracture surface of 3mm Al5005H34 plate impacted at room temperature at 8 x magnification (right). Formation of fragmentation is evident in centre left of image.

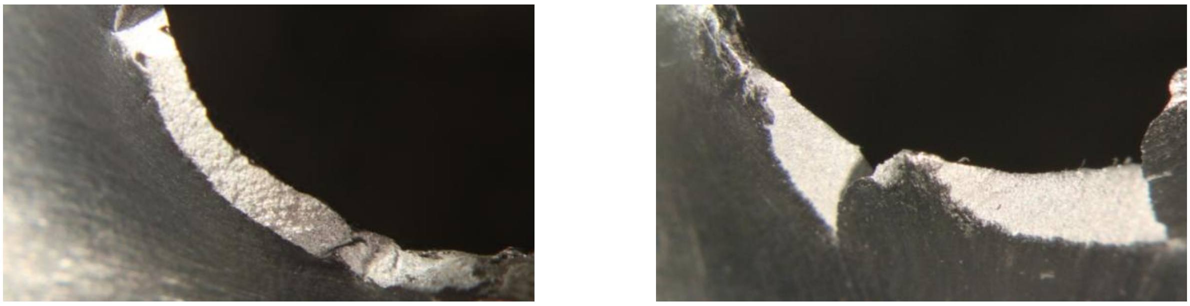


Figure 6. Fracture surface of 3mm Al5005H34 plate impacted at -119°C at 8 x magnification.

D. Heat Pipe Obstruction

In the event that the harpoon is used as an ADR system, it is unlikely that it will penetrate a completely unobstructed panel. Potential targets such as satellites will commonly have heat pipes as well as equipment boxes fitted. As such, it was deemed important to obtain empirical results for impact into panels fitted with such obstructions. Impact testing was therefore conducted to determine the ability of the harpoon to penetrate 2mm thick Al6061-T6 plate with a representative heat pipe secured to the back. The heat pipe was manufactured from 6061-T6 billet Al in accordance with industry specifications. The pipe was secured to the plate with M4 bolts and a representative adhesive, also in accordance with these specifications, to ensure an accurately representative configuration. The M4 bolts were tightened to a torque value of approximately 2.3Nm. The particular grade of Al was chosen as it was more readily available than the 6063-T6 and 7020-T6/2024-T6 grades of Al used for the actual heat pipes and panels respectively, and had similar yield properties. The target configuration is shown in Fig. 7.

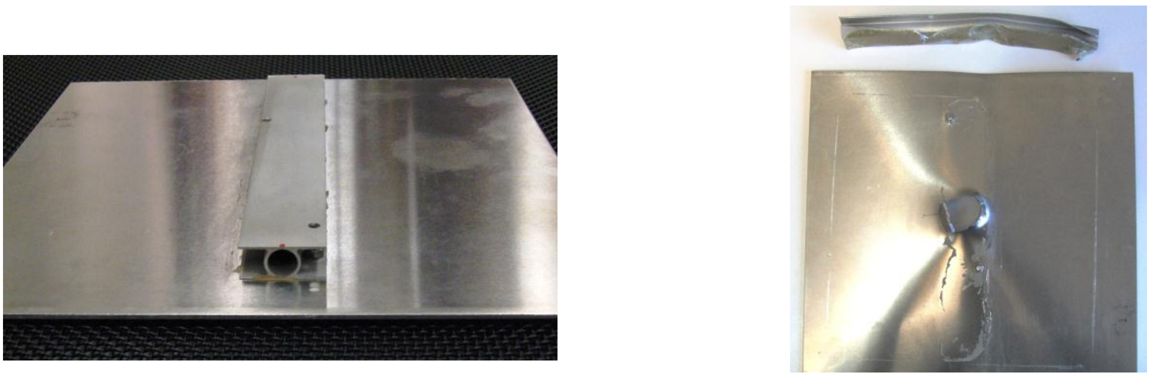


Figure 7. Rear surface of Al6061-T6 fitted with representative heat pipe manufactured to industry specifications (left). Result of impact with blunt projectile (hex) at 66m/s into 2mm Al6061-T6 plate fitted with a representative heat pipe showing complete separation of heat pipe. Impact was along centre-line of pipe (right).

A series of three tests were conducted. These involved impacting the target panel with different tips as well as at different velocities and alignments. The primary purpose of this series of tests was to determine the effect that an obstruction such as a heat pipe would have on the harpoon's ability to penetrate the target panel. The results of this series of tests are detailed in table 3.

Table 3: Results from impact of harpoon into 2mm Al6061-T6 panel fitted with representative heat pipe

Test	Impact Velocity (m/s)	Tip	Impact Location	Result
1	65.85	Blunt (Hex)	Heat pipe centreline	Complete separation of heat pipe due to shear of flanges at mounting bolts.
2	49.97	Cone	Left of centre	Partial separation of heat pipe due to shear of left flange at mounting bolt. Harpoon remained attached to panel post penetration.
3	42.18	Integrated toggle	Heat pipe centreline	Partial penetration of panel and separation of heat pipe at bolt location closest to point of impact.

From these results it can be seen that a heat pipe does not prevent the harpoon from successfully penetrating a target panel. In all cases, the common failure mode was shear through the pipe flange at the mounting bolt(s) resulting in separation of the pipe from the target panel. High speed footage of test 1 showed bending of the pipe with initial harpoon penetration followed by clear separation from the panel. As a result of penetrating the panel at different locations with respect to the pipe centre line it could be seen that this separation occurred whether the harpoon impacted directly along the centre line of the pipe or along its edge. This is of significant benefit for an actual mission scenario where the precise locations of heat pipes, as well as harpoon impact location, may be unknown. In addition, the results of test 3 showed that even if the projectile impact energy is too low for full penetration of the target panel, the pipe will still separate. From these tests it can therefore be concluded that the flange is the weakest point for this target configuration and will reliably fail on impact, allowing successful harpoon penetration.

E. Oblique Impact Testing

In regards to impact angle, although normal impact and perforation of thin and moderately thick plates has been extensively investigated both analytically and experimentally by authors such as Goldsmith, Backman, Gupta, and Finnegan, data related to oblique impacts are not readily available. Oblique impact needs to be considered for an ADR harpoon system due to the possibility that the target may be spinning and even tumbling. The limited existing research shows that the ballistic limit for a given projectile changes with varying degrees of obliquity [10]. This is largely a result of increasing path length with increased obliquity as well as increased bending stresses due to the asymmetry of the loading [7].

In order to determine the effect of obliquity on harpoon penetration, a series of preliminary tests were conducted with conical and blunt tips. These experiments were conducted at an angle of obliquity up to 40° by angling the target box. As the angles were determined with a standard protractor, a conservative uncertainty margin of +/-1.0° has been included. Testing was concluded at 40° due to damage which was incurred to the harpoon as a result of the projectile sliding along the face of the target plate and impacting the target holders. The results of the empirical tests are detailed in Table 3.

Table 3: Results of Oblique Impact Testing for 3mm Al5005H34 Target Plate

Tip Shape	Angle of obliquity +/-1.0°	Impact velocity (m/s)	Result
Cone (16mm, 24°)	20	43.5	Through
	30	44.3	Through
	40	44.3	Ricochet
	40	46.6	Embedded in target
	40	48.2	Through
Cube (16mm side length)	20	43.5	Through
	30	43	Through
	40	48	Skidded along target face

From these preliminary results it can be seen that the blunt tip appears to require a greater relative increase in energy in order to penetrate targets at increasing angles of obliquity. This can be concluded from the fact that, although it had a greater relative increase in energy between 30° and 40° in comparison to the conical tip, it failed to even embed in the target. These limited results were supplemented with simulation. These results are discussed in section G.

F. Effectiveness of Harpoon Lock-on Mechanisms

In order to allow de-orbiting of the debris, an effective lock-on mechanism needs to be developed. Preliminary studies already completed by Airbus Defence & Space have identified a number of options for this purpose including iterations of fixed and deployable barbs. Due to the increase in impact energy required for devices with fixed barbs, as well as a higher likelihood of debris creation, this project focused on testing two different deployable toggle mechanisms. The first used a simple spring activated toggle mechanism, whilst the second was an integrated tip/toggle design (Fig. 8). The second mechanism deployed under the influence of gravity during testing. As such, it would require a simple strain-energy device, such as a tape spring, to facilitate deployment in zero-g. This tip was designed to improve reliability in impacts where the target panel may be obstructed by heat pipes or obstructions. In both cases, premature deployment of the toggle mechanisms was prevented by fitting collars which were forced to release upon impact with the target panel. Both toggle mechanisms were tested during impact with 23mm thick Al H/C panels with 1mm Al face sheets. As shown in Fig. 8, both mechanisms operated reliably and effectively.

Figure 8. Harpoon with spring activated toggle (left) and integrated toggle (right) in deployed configurations.

In addition to testing the deployment of these mechanisms, their ability to withstand the expected forces associated with de-orbiting manoeuvres was also established. A study conducted by Airbus Defence & Space suggests that an average force of approximately 500N, with a peak as high as 2-3kN in the event of a high thrust de-orbit manoeuvre, could be experienced. This testing was conducted on Al H/C panels. These samples were chosen instead of Al plate as they have a lower resistance to shear and therefore represent the worst case scenario. They are also typical of spacecraft structures. Of note, the CFRP panel also used in satellite applications was not available for testing. The testing involved clamping the target panel with embedded harpoon in a tensile testing machine and recording the peak force achieved before failure of either the mechanism or target material. The results of the testing are detailed in Fig. 9.

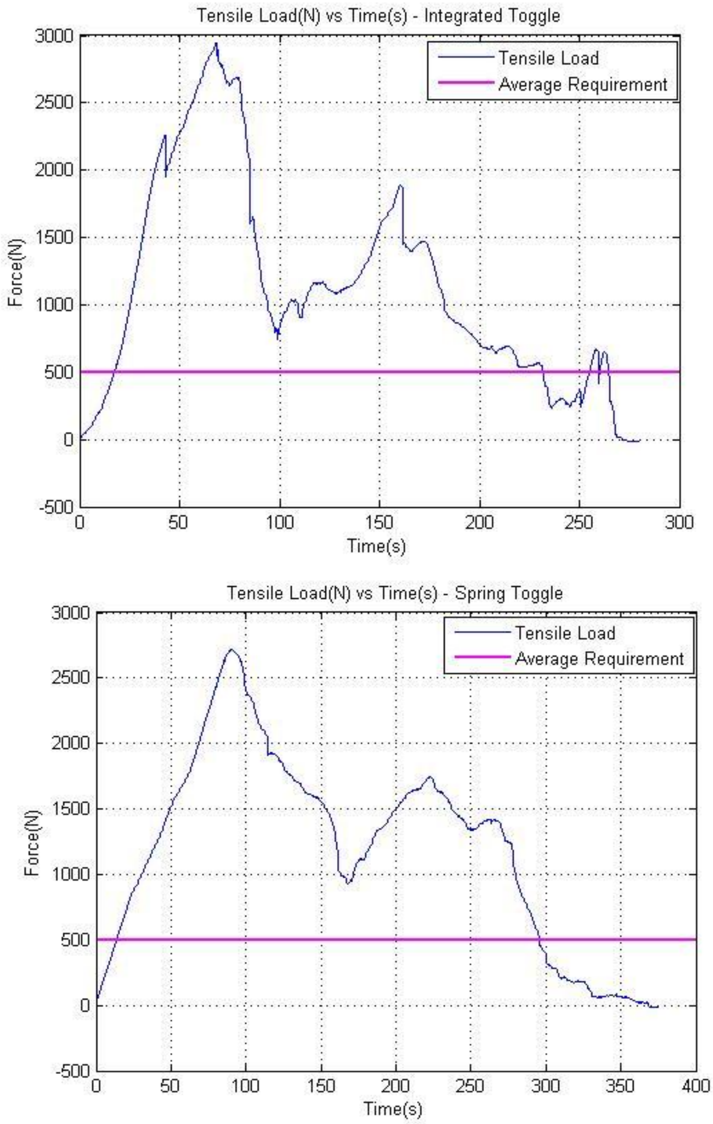


Figure 9. Pull-out resistance of integrated toggle locked onto Al H/C panel (top), and pull-out resistance of spring activated toggle locked onto Al H/C panel (bottom).

As a result of conducting this test it was found that the peak load experienced by the integrated toggle was 2,950N, while the spring activated toggle experienced a peak load of 2,720N. As a result, both were capable of withstanding the minimum load requirement of 500N. Pull out occurred during the spring toggle test as a result of the failure of the toggle. The test of the integrated toggle resulted in failure of the honeycomb panel with the toggle remaining entirely undamaged. This suggests that the integrated toggle would be capable of withstanding much higher tensile forces and is better suited to applications where high peak forces are likely. This enhanced capability is due to its more robust design and larger surface area over which the tensile force can act. In addition, although the integrated toggle closed up as the honeycomb failed, it was noted that it was still capable of supporting the minimum load requirement of 500N in a partially deployed configuration. This suggests that this type of toggle can be relied upon to provide sufficient lock-on to the target even if it does not fully deploy on impact.

G. Impact simulation using ANSYS/Autodyn

Part of this study involved numerical modeling to simulate impact into Al and CFRP targets. The reasons for conducting simulation were: to identify the suitability of using simulation as a harpoon development tool; to obtain further results for the study into oblique impact; to parameterise tip dimensions and therefore reduce the amount of experimental testing required; and to simulate impact into CFRP targets. In regards to the models simulating impact into Al targets, a linear equation of state was used. This was chosen due to the fact that the relatively low velocity of the impacts being simulated meant that only the initial (linear) portion of the associated material Hugoniot was relevant. The strength model chosen was the Johnson-Cook model, whilst the failure model was based on effective plastic strain. The Johnson-Cook strength model was chosen as it incorporates a strain hardening component which is an important consideration for low velocity impact into Al targets. The model also includes strain rate and thermal softening effects. The results of quasi-static testing of 3mm Al5005H34 samples were used to determine the primary constants required for the Johnson-Cook strength model. Extensive convergence testing was also conducted to improve the accuracy of the simulations. This involved mesh refinement, adjustments of hourglass damping controls, determining the sensitivity of results to variations in the Johnson-Cook parameters, and optimising target panel size to minimise edge effects whilst maintaining realistic computation times. In addition, the hourglass energy was monitored and maintained at below 10% for the majority of the simulations and the energy error was maintained below 5%.[2]

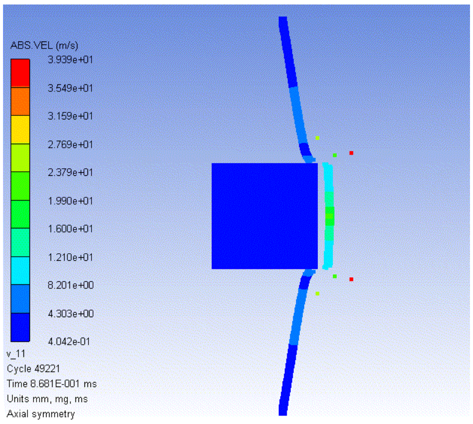
In regards to mesh size, convergence testing was conducted where the target mesh size was reduced from 10mm to 1mm in 1mm increments. This showed that the variation in energy errors reduced significantly once the mesh size reached 2mm. As such, this size was used for target meshes in all of the simulations undertaken. In regards to mesh type, initially the automatically generated configuration was maintained which included a hexahedral mesh for the target and tetrahedral mesh for the harpoon. As the default mesh size was usually quite large however, it was reduced in order to decrease the final energy error values. Although this configuration was suitable for the ballistic limit simulations, it resulted in excessive energy errors once oblique impact simulations were conducted with conical and complex tips. As a result, the target mesh was changed to a tetrahedral structure. This significantly reduced the hourglass energy values for these remaining simulations.

In order to assess the suitability of using simulation to support the harpoon development process, ballistic limit simulations were run and convergence with empirical results assessed. This showed that varying degrees of agreement existed between experimental and simulated ballistic limits. Whilst the simulated ballistic limit for the conical tip exactly matched that obtained experimentally, the model over-predicted the ballistic limit for the blunt projectiles. The relationship between circumference and ballistic limit identified experimentally however was confirmed with simulation. In addition, the petaling and plugging failure modes shown by the simulations were representative of those achieved experimentally.

The oblique impact simulations conducted confirmed that a blunt tip requires more energy to perforate a 3mm Al plate at high angles of obliquity in comparison to a conical tip, particularly at angles in excess of 45°. Custom tips were also designed and tested for their performance at increasing angles of obliquity however, their performance was worse than that achieved by a conical tip shape. Simulations involving the parameterisation of conical tips were also run where the tip diameter was varied between 16-20mm. This showed that the ballistic limit increases with tip diameter.

In regards to harpoon impact into CFRP, this simulation used an SPH solver in Autodyn-2D. This solver was chosen due to its meshless nature and therefore greater applicability to impact simulations of brittle materials in comparison to the Lagrange solver. A data set developed in a previous study on testing, material modeling, and numerical simulation of impact on CFRP was used in this solver [11]. Simulations of impact into 1mm CFRP panel by both blunt and conical projectiles at their ballistic limits of 11m/s and 35m/s respectively, detailed in Fig. 10, showed significant differences in debris formation, debris velocity, and projectile deceleration. Specifically, the debris formed during impact by a blunt projectile is more coherent and has a lower velocity than that produced by a conical tip. The maximum velocity of these debris particles is approximately 40m/s whilst debris created by a conical projectile reaches approximately 590m/s. The significant difference in ballistic limits can be explained by the fact that CFRP has low impact strength but high tensile yield strength [12]. As penetration by the conical tip results in piercing and therefore a high tensile load, it experiences greater deceleration than the blunt tip which causes the CFRP to fail as a result of high impact-loading. It is also worth noting that debris formation is initiated on the back-face of the target. This is typical of impact damage incurred by thin composite targets [7] and means that any debris created by the harpoon will remain internal to the target.

[2] Hourglass energy is defined on the LS-Dyna Support site, http://www.dynasupport.com/howtos/element/hourglass, as '....nonphysical, zero-energy modes of deformation that produce zero strain and no stress.' They therefore need to be minimised in simulations as they can result in unphysical results.



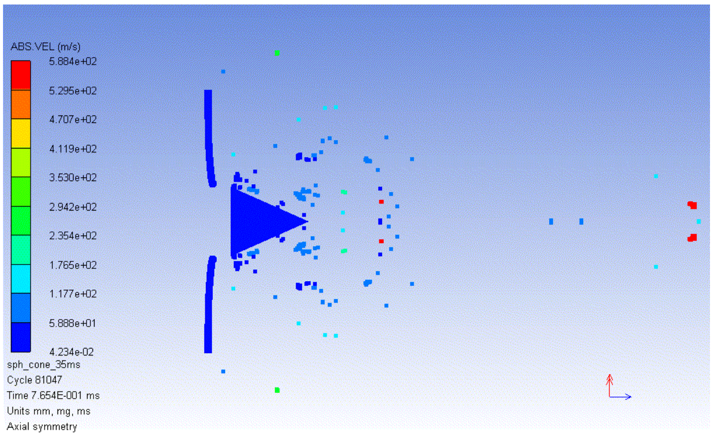


Figure 10. Simulation of blunt projectile impacting 1mm CFRP at ballistic limit (11m/s) (left), and simulation of conical projectile impacting 1mm CFRP at ballistic limit (35m/s) (right).

H. Harpoon Effectiveness in μ-Gravity

The experimental results presented in this paper have been for impact into fixed targets. This test regime was designed to be representative of impact into high mass targets only. These include rocket bodies, and large, inoperable satellites in the 1 to 8 tonne class with high inertia properties. In addition to these results, however, preliminary studies have been conducted to assess the effectiveness of a harpoon ADR system against low mass targets. In order to conduct such a test regime, a testing apparatus which allowed unconstrained movement of the target was required. Although the construction of a rig allowing a full 6 DOF would be desirable, the initial rig developed was limited to 2 DOF. This allowed the rotation and translation of an approximately 1kg target, designed to represent of a Cubesat. This 'Cubesat' was fitted with Al H/C panel and can be seen in Fig.11. The harpoon used for this test had a mass of approximately 0.045kg and the impact speed was approximately 18ms-1. Translation was achieved with low friction linear bearings and rotation achieved with roller bearings. Of note, the friction from the linear bearings was quantified by comparing the theoretical ΔV of the target, which was obtained from a simple momentum balance calculation, with the actual ΔV obtained from high speed footage of the impact. This calculation showed that the friction was negligible. As a result of conducting this test, it was shown that a low mass harpoon was capable of penetrating an Al H/C target with 2 DOF. A snapshot from the high speed footage showing successful penetration of the 'Cubesat' is shown in Fig.12. In addition, this test rig was further modified to allow 3 DOF. Testing with this rig showed that the harpoon was still able to achieve successful penetration.

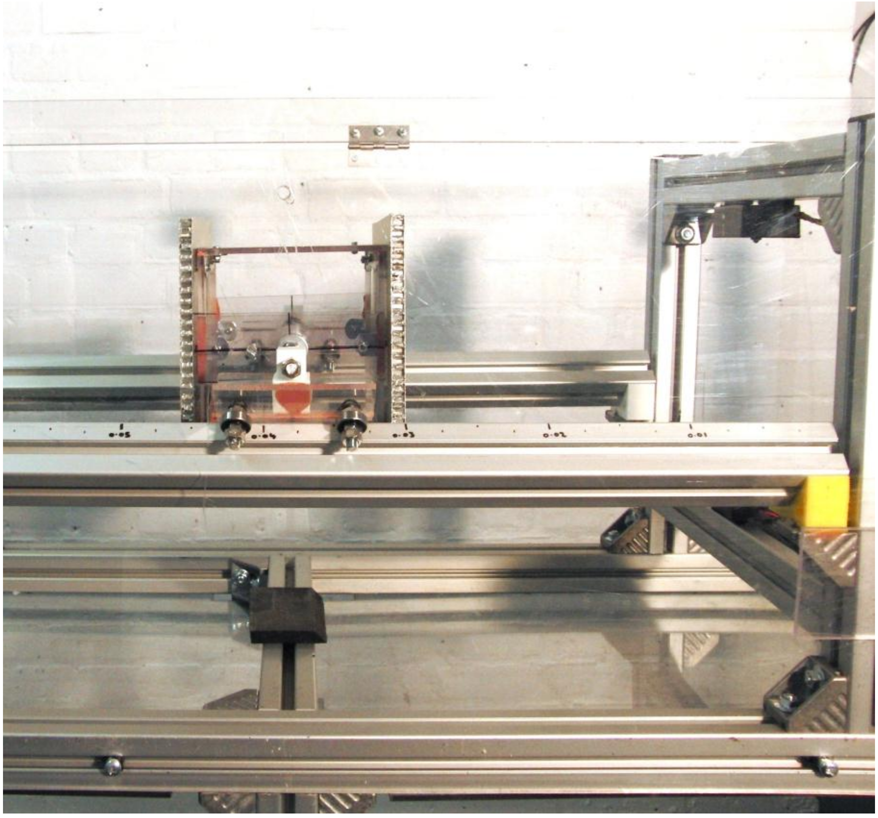


Figure 11. Test-rig used to assess harpoon impact of low mass target with 2 DOF. Test rig is comprised of a mock 'Cubesat' fitted with AL H/C panels and mounted on linear bearings. These bearings allow translation on rails with very low friction. 'Cubesat' is also mounted on a single axle supported by bearings to allow rotation of target.

Figure 12. Image taken from high speed footage showing successful penetration of 'Cubesat' by low mass harpoon with impact speed of approximately 18ms^{-1}. Harpoon is white Delron rod. Rod is fitted with Kapton tape covered damper seen on immediate right of AL H/C panel to ensure triggering of IR speed sensors.

IV. Conclusion

In this paper the results of a preliminary study to determine the effectiveness of using a harpoon in an ADR application were presented. These results have demonstrated the feasibility of the harpoon as a method of active debris removal. Specifically, the work performed showed that:

Outcome	Result	Implication
A	The ballistic limit for penetration is less for a blunt nosed projectile than a conical one, however, this only applies to normal impact. The conical tip requires a less significant increase in energy for oblique impacts.	A conical tip would be more versatile in an ADR mission as targets are likely to be rotating or tumbling therefore resulting in an oblique impact.
B	Secondary debris is created by both blunt and conical projectiles. Blunt projectiles create a plug, while conical projectiles can create minor fragmentation and spalling. Secondary debris for both cases is on the side opposite to impact.	Neither blunt nor conical tips prevent secondary debris formation, therefore, this criteria should not be used in the selection of tip shape.
C	Penetration of 3mm Al plate targets at low temperatures reduces secondary debris.[3] This results from increased brittleness of the target material and therefore a reduction in ductile spalling. The energy required to penetrate a low temperature Al target is also increased.	Targets in eclipse may be successfully penetrated with an increase in impact energy and a reduction in the likelihood of secondary debris formation, if a conical tip is used.
D	A harpoon can successfully penetrate Al targets, including those with heat pipe obstructions with the creation of internal debris only.	A harpoon can successfully penetrate a target panel with heat pipe obstructions. This applies whether the impact site coincides with the heat pipe centreline or is partially offset.
E	A conical projectile was shown to be capable of penetrating a 3mm Al target up to an angle of obliquity of 40°.	A harpoon with a conical tip can penetrate targets that are rotating or tumbling where the resulting impact angle is 40° or less. Follow on work is required to assess the limits of obliquity.
F	Results of toggle testing showed that the harpoon could successfully lock on to an Al target panel and withstand the forces expected during deorbiting manoeuvres.	Current toggle designs are sufficient to facilitate deorbiting of Al H/C and Al panel targets.
F	An integrated toggle design was proven to be more robust than a spring loaded toggle.	An integrated toggle may provide a more reliable solution than a spring toggle for actual mission scenarios.
G	Simulation undertaken with ANSYS/Autodyn showed that numerical modeling can be used as an effective tool to predict failure modes, likely debris creation for both Al plate and CFRP, and relative (but not absolute) impact speeds required for different tip shapes during normal and oblique impacts.	Simulation can be used to reduce the quantity of experimental testing required for further harpoon characterisation studies, such as oblique impact modeling.
H	Preliminary testing of harpoon impact into targets with low inertia and multiple DOF showed that successful penetration was still possible under such conditions.	The harpoon represents a viable ADR system for capture of low mass targets in μ-gravity.

[3] This applies to the 5005H34 grade of Al used in this study. The applicability to other grades of Al will be verified in further work.

V. Further Work

Additional parametric studies are proposed to further validate the use of a harpoon in an ADR application. These include additional empirical and simulation studies of oblique impacts, testing toggle deployment during impact with obstructed and oblique panels, and conducting a comparison between Al5005H34 and higher strength Al to allow extrapolation of the results to satellite structures made with different Al alloys. Tip optimisation which develops on the investigation of ballistic limit and tip shape should also be undertaken. In addition, the design and testing of an effective damper system is required. Such a damper will prevent the harpoon from passing through a target panel due to excess impact energy. Finally, a preliminary study of tether deployment in vacuum and micro-gravity is proposed to improve confidence in the ability of the tether to unravel freely during firing of the harpoon.

References

1. Astrium UK, *Active Debris Removal Harpoon Design Report*, Oct 12, Stevenage, UK

2. Nikolaev, S., 'Modeling Kessler syndrome using brute force approach', AIP Conference, 1464, 456 (2012), New Mexico, 2012

3. Schaefer, F.K., Ryan, S., Lamber, M., Putzar, R., 'Ballistic Limit Equation for Equipment Placed Behind Satellite Structure Walls', International Journal of Impact Engineering, Vol 35, 2008

4. Meyers, M.A., Dynamic Behaviour of Materials, John Wiley & Sons, USA, 1994

5. Goldsmith, W., & Finnegan, S.A., 'Normal and Oblique Impact of Cylindro-Conical and Cylindrical Projectiles on Metallic Plates', International Journal of Impact Engineering, Vol 4, No. 2, 1986

6. Levy, N., & Goldsmith, W., 'Normal Impact and Perforation of Thin Plates by Hemispherically Tipped Projectiles – II. Experimental Results, International Journal of Impact Engineering, Vol 2, No. 4, 1984

7. Zukas, J.A., Impact Dynamics, John Wiley & Sons, USA, 1982

8. Reed, J., & Barraclough, S., 'Development of Harpoon System for Capturing Space Debris', 6th European Conference on Space Debris, Darmstadt, Germany, 2013

9. Hickey, C., Mechanical Properties of Titanium and Aluminium Alloys at Cryogenic Temperatures, Watertown Arsenal Laboratories, 1962

10. Backman, M.E., & Goldsmith, W., 'The Mechanics of Penetration of Projectiles into Targets', International Journal of Engineering Science, Vol 16, 1978

11. Wickleim, M., Ryan, S., White, D., Clegg, R., 'Hypervelocity Impact on CFRP: Testing, Material Modelling, and Numerical Simulation,' International Journal of Impact Engineering, 35, 12 (2008) 1861

12. Greszczuk, L., et al., Investigation of Brittle Fractures in Graphite-Epoxy Composites Subjected to Impact, McDonnell Douglas Astronautics Company, California, USA, 1975, viewed on-line at http://primoa.library.unsw.edu.au, on 31 Aug 14

Fast Cold-Start Acquisition of GPS Signals Using the Delay Doppler Map Accelerator

Eamonn P. Glennon [1], Joon Wayn Cheong [1], Kevin Parkinson [3] and Andrew G. Dempster [1,2]

[1] *Australian Centre for Space Engineering Research (ACSER), UNSW*
[2] *School of Electrical Engineering and Telecommunications (EET), UNSW*
University of New South Wales, Sydney, NSW, 2052 Australia

[3] *General Dynamics Corporation, New Zealand*

Summary:
The Namuru V3.2R3A and Kea V4.1SB Global Positioning System (GPS) receivers are Field Programmable Gate Array (FPGA) based receivers designed to operate in low earth orbit (LEO). Being FPGA based, each has the flexibility to incorporate additional hardware blocks, such as the Delay Doppler Map Accelerator (DDMA), which was specifically designed for GPS Reflectometry applications.

In this paper, we examine how this same hardware block can also be employed to improve receiver's cold-start time by virtue of the wide-bandwidth afforded by the DDMA. This includes a brief review of the hardware and software signal-processing performed by the receiver, followed by some test results performed using ACSER's GPS simulator. We also discuss how improvements can be made to future products using some of the techniques described here.

Keywords: space-borne GPS receiver, signal acquisition, Namuru GPS, Kea GPS, Delay Doppler Map (DDM), search engine, detection

Introduction

The University of New South Wales (UNSW) is well known for its pioneering work in Field Programmable Gate Array (FPGA) based Global Positioning System (GPS) receiver design, which was undertaken by the Satellite Navigation And Positioning Laboratory (SNAP Lab)[1] and privately held General Dynamics Corporation of New Zealand. The outcome of this work was the development of the Namuru FPGA based GPS receiver [1], which in turn resulted in the opportunity to further develop the GPS receiver technology for operation in space as part of the Garada and Biarri projects, carried out by the Australian Centre for Space Engineering Research (ACSER). The culmination of this work has been the opportunity to fly the GPS receivers on board several CubeSat missions. This includes the three-satellite Biarri constellation, which has been done through the Defence Science and Technology Organisation (DSTO), the UNSW QB50 Cubesat EC0 and the DSTO Biarri follow-on mission 'Buccaneer'[2].

An important consequence of having a real world, high-cost mission depending on a technology is a focus to ensure that the technology is reliable and fit for purpose. For these

[1] SNAP Lab is a joint collaboration between the Surveying and Geospatial Engineering (SAGE) group of the School of Civil and Environmental Engineering (CIVEN) and the School of Electrical Engineering and Telecommunications (EET) at UNSW.

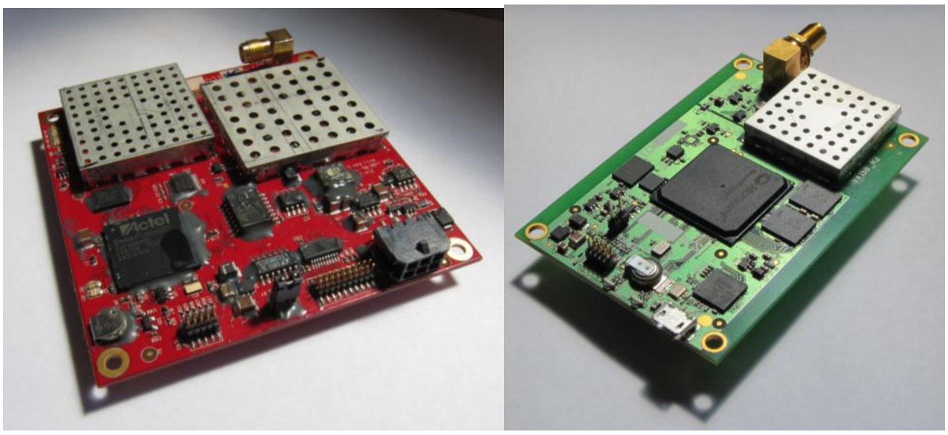


Fig. 1: Namuru V32R3A flight model (left) and Kea V41SBR2 prototype (right)

projects, the Namuru and Kea GPS receivers are no exception and significant effort has therefore been put into ensuring that the best possible performance and reliability is achieved from the receiver in order to ensure mission success. This includes the development of bespoke hardware and firmware to meet the specific requirements associated with each program.

A fundamental requirement of a space-borne GPS receiver is to navigate in orbit with a sufficiently short time-to-first-fix (TTFF). However, such a receiver often faces significant difficulty with acquisition of the satellites when the receiver is first powered up. Receivers typically rely on the use of prior information, including knowledge of the satellite orbits, the current time and the receiver position and velocity in order to reduce the TTFF, but such information is typically not available following a receiver factory-reset. A factory-reset refers to the process of clearing the receiver's non-volatile memories and real-time-clock, thereby selecting a default state.

One of the more unusual requirements for the UNSW EC0 mission is to perform GPS remote sensing through the addition of a custom LHCP antenna array, combined with an ability to extract and store digitized intermediate-frequency (IF) GPS signals. This IF data is then processed with another custom hardware block, the Delay Doppler Map Accelerator (DDMA) in order to produce a DDM image from which sea surface wind-speed and direction can be derived. Other groups throughout the world have also designed such hardware, with examples including [3], [4] and [5].

In this paper, we first review the Namuru V32R3 [2] and Kea V41SBR2 FPGA based GPS receiver, the DDMA hardware block [6] and the problem of acquiring GPS signals in low earth orbit following a factory reset. We then explore how the DDMA hardware as implemented on the Namuru V32R3 receiver can be used to speed up the cold-start acquisition process, as well as describing the firmware changes required to implement such a process. Experimental results obtained using UNSW's GPS simulator are then presented and discussed. Finally, we conclude with recommendations for the Kea series of GPS receiver to achieve improved future performance.

Namuru and Kea GPS Receivers

The Namuru V32Rx and Kea V41SBRx are the FPGA based GPS receivers that have been developed specifically for the Biarri and UNSW EC0/DSTO Buccaneer spacecraft, respectively.

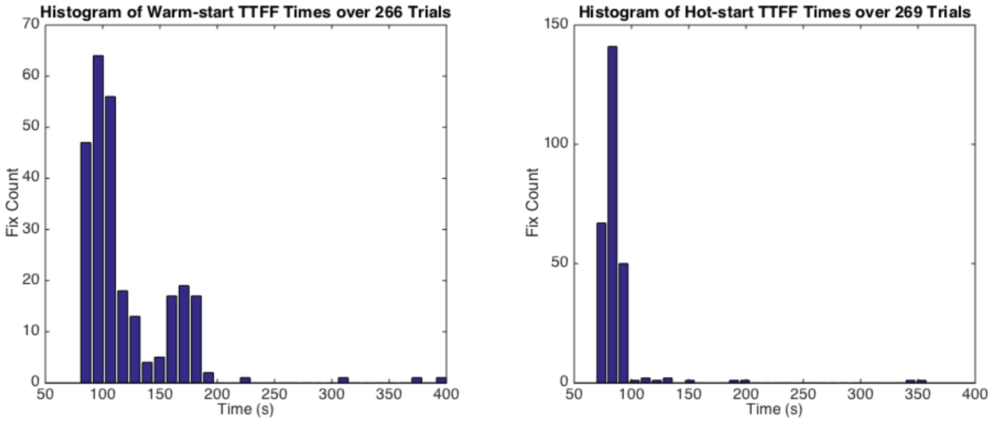


Fig. 2: Warm & Hot start times for a Namuru V32 receiver using TLEs in LEO

In the case of the earlier Namuru V32Rx receiver, the design was customized for a Boeing Colony II Cubesat platform. It employed a Microsemi SmartFusion System on Chip (SoC) and ProASIC3E FPGA for the digital signal-processing and a Zarlink GP2015 RF front end for the GPS analog down-conversion and quantisation. When the Namuru V32Rx receivers were first designed, the SmartFusion SoC was selected because it contained a built-in FPGA, which it was hoped would be sufficiently large to accept the GPS correlator baseband. However, the GPS correlator was subsequently found to require almost three times the available resource and it was therefore necessary to include the additional ProASIC3E FPGA, which was sufficiently large to accept the GPS correlator design from previous versions of Namuru, albeit with little margin for future expansion. The Namuru V32Rx receivers implement 12 conventional GPS correlators each containing in-phase and quadrature-phase channels and 3 code-phase taps separated by 0.5 chips. Each correlator dumps its accumulate and dump outputs every 1 ms and therefore allows 1 kHz of bandwidth to be observed. The FPGA resource within the SmartFusion was therefore surplus to requirements and was subsequently used for the Delay Doppler Map Accelerator hardware block.

The Kea receiver design represents a privately funded evolution of the Namuru receiver and represents a significant improvement to the receiver design. Use of a MicroSemi SmartFusion2 SOC with 50,000 logic elements has allowed the ProASIC3E to be eliminated, while use of a General Dynamics Corporation BL2627 CMOS L1 RF front-end has allowed major reductions in the power consumption and size of the RF front end. CalPoly CubeSats are supported via a custom interface board that is used to ensure not only mechanical and electrical compatibility with the CalPoly cubesat standard, but also provides features such as antenna-switching, arbitrary reference frequency generation with excellent phase noise from a GPS disciplined source, as well as RS232 and RS422 interfaces. Photographs of the Namuru V32R3A flight model receiver and the second revision of the Kea V41 are shown in Figure 1.

Acquisition of GPS Signals in Low Earth Orbit

Warm Start

Operating a GPS receiver in low earth orbit presents many challenges to the GPS receiver designer. One of the first problems concerns acquisition of the satellites when the receiver is first switched on, with several scenarios being applicable.

The most frequent scenario, typically referred to as a *warm-start*, involves the case where the receiver has prior knowledge of the GPS constellation by virtue of a previously downloaded almanac, the current time obtained from a built-in real-time-clock, and knowledge of the current position and velocity. A *hot-start* is similarly to a *warm-start,* except that previously stored ephemeris is employed thereby saving at least 30 seconds of data extraction for each satellite. For a terrestrial receiver, the current position can usually be taken as the previous position and the velocity can be safely set to zero. However, such a strategy cannot be used for the LEO scenario because the receiver circles the earth roughly every 90 minutes and experiences Doppler frequencies of up to ±40 kHz, rather than the ±4 kHz experienced by the terrestrial receiver.

One solution to this dilemma is to calculate a set of very coarse Two Line Elements (TLE) for the receiver similar to those calculated by NORAD. In this way, the Namuru or Kea receiver is able to calculate the satellite visibility and Doppler frequencies when the receiver is powered up and a rapid TTFF can be achieved. Experimental results obtained using a Namuru V32 in a low earth orbit scenario are shown in Figure 2, where it can be seen that typical warm start times are less than 200 seconds, although some outliers are occasionally observed.

Cold Start

In some cases it may be necessary to completely reset the GPS receiver, in which case the prior knowledge required for a *warm-start* will not be available. When the receiver is powered up under such conditions, it has to perform what is commonly referred to as a *cold-start*, which amounts to blindly searching for all the satellites until sufficient satellites to perform a fix are acquired.

The time to perform a cold start with the conventional correlators of the Namuru or Kea receivers can be calculated easily.

The number of iterations $N_{1,1}$ that a single hardware block must perform in order to search the entire search space for a single space-vehicle is given by:

$$N_{1,1} = \frac{BW_{sv}\, C_{sv}}{BW_c C_c}$$

BW_{sv} is the total search bandwidth and C_{sv} is the total search code-phase space to be searched. The bandwidth and code-phases visible to a channel during each search-iteration are given by BW_c and C_c, respectively. Each search-iteration is typically made up of several rounds of non-coherent integration N_{NCR} of some coherent integration period T_{IP}. Here, T_{IP} is defined as the time taken to perform a 1 ms coherent integration, which for a real time correlator is usually 1 ms, but for a postprocessing DDMA might be longer, especially if more code phase taps are calculated over time. The time to perform the above search is therefore given by:

$$T_{1,1} = N_{\ 1,1} T_{IP} N_{NCR} = \frac{BW_{sv}\, C_{sv}}{BW_c\, C_c} T_{IP} N_{NCR}$$

When there are N_{sv} satellites to find using N_c channels, the above result scales accordingly.

$$T_{N_{sv},N_c} = T_{1,1} \frac{N_{sv}}{N_c}$$

Table 1: Conventional correlator search time parameters

Parameter	Value	Parameter	Value
BW_{sv}	80,000 Hz	T_{IP}	0.001 s
C_{sv}	1023 chips	N_{NCR}	4
BW_c	750 Hz	$T_{1,1}$	290.9866 s
C_c	1.5 chips	N_{sv}	32
$N_{1,1}$	72,746.6	N_c	12
		$T_{32,12}$	775.9644

The cold start search time can therefore be calculated to be 776 seconds or 12 minutes 57 seconds using the values formula values given in Table 1. This process reduces the bandwidth of each correlator search from 1 kHz to 750 Hz order to reduce the scalloping losses at the frequency boundaries.

The problems of performing a cold-start are therefore plain to see, with almost 13 minutes required to search the entire constellation when operating in LEO. The process is further complicated by the fact that over such a long interval the satellite visibility is almost certain to change, meaning that some satellites may no longer be visible by the time that portion of the search is reached. Clearly, cold-starts in LEO are to be avoided if possible.

Space represents one of the harshest environments to which electronic systems can be subjected and the probability of needing to perform a factory reset and therefore a cold-start is expected to be not insignificant. Given this situation, it is essential that a receiver be capable of cold-starting and recovering as fast as possible. That said, one way of avoiding the need to cold-start following a factory reset would be for the satellite on-board-computer (OBC) to restore essential parameters deleted during a factory reset. These parameters would include up to date almanac, recent TLE parameters, the most recent oscillator frequency offset and up to date time. However, this requires additional procedures to be performed in such an event, resulting in a more complicated concept of operations (CONOPS).

Delay Doppler Map Accelerator (DDMA)

Purpose

The purpose of the DDMA is to help generate DDMs, which are simply GPS correlations calculated across a range of code phase taps and frequency bins and usually non-coherently integrated in order to improve the output signal-to-noise ratio. DDMs are typically generated from IF signals reflected from the earth's surface and downlinked to earth for post-processing using software-defined-radio techniques. However, this is not feasible for many cubesat platforms that have insufficient communications bandwidth to downlink large datasets. The DDMA permits the post-processing to be performed on board the spacecraft and the result down-linked instead. This hardware has been optimised for space based DDMs, for which the coherence interval is limited to no more than 1 ms [7], although in general, the coherence interval is dependent on the properties of the reflecting surface.

Hardware Implementation

The DDMA is a hardware correlator block that can be used to perform GPS correlations on previously captured IF data with minimal additional processor load being placed on the GPS receiver microprocessor. The DDMA hardware block was first described in [6] and is briefly summarized here.

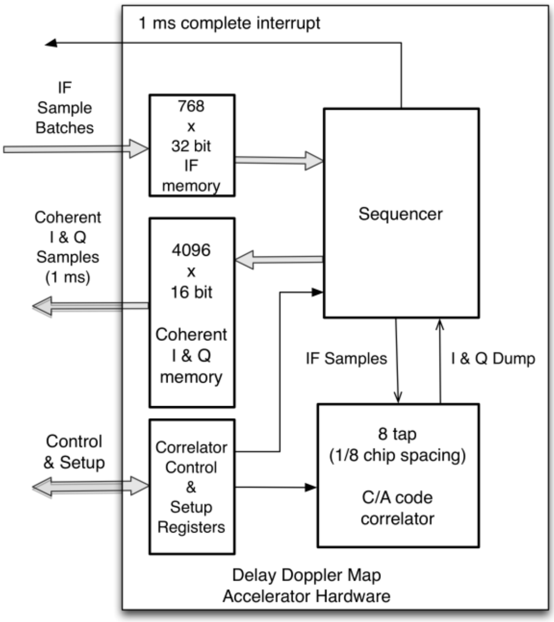


Fig. 3: DDMA Hardware Block Diagram

The DDMA hardware can be thought of as a normal hardware channel that re-processes the same 1 ms worth of IF samples with different initial code-phase settings and carrier frequency settings, storing the outputs in a memory that can be further processed by the firmware. A block diagram of the hardware block is shown in Figure 3.

It is implemented as a channel single correlator with 8 code phase taps that takes its input from one of two 384 × 32 bit ping-pong buffers loaded by the firmware from a store of previously captured IF samples. Because DDMs are required to show the correlation curve in detail, the spacing for the code phase taps can be widened or narrowed, with selectable chip spacing of 1/2, 1/4 or 1/8 of a chip. Similarly, because a bandwidth greater than 1 kHz is often required, the dump interval from the correlator can be set to 1 ms, 0.5 ms, 0.25 ms or 0.125 ms. thereby setting the effective bandwidth to 1 kHz, 2 kHz, 4 kHz or 8 kHz, respectively. Irrespective of the selected dump interval, up to 1 ms of IF data is processed during each round of the DDMA. The results of each round are stored in a 4096 × 16 bit output memory that holds 1 ms worth of I and Q dump samples for each code phase tap (equivalent to 8 sets of dump samples when a dump interval of 0.125 ms has been selected) and a user specified number of code phase taps, up to the limit of the available memory. The types of DDMs that can be generated using this hardware in a single pass, which is to say without re-processing the dataset with a completely new code-phase or Doppler offset, are

Table 2: DDMA Configuration Options, with fast-acquisition option highlighted (Note that the Taps column can be calculated as the Chip Range / Chip Spacing)

Chip Spacing Chip Resolution (chips)	Dump Interval (ms)	Taps	Dumps Per ms	Chip Range (chips)	Bandwith (kHz)	"1ms Look" Compute Time (ms)
1/8	1/8	80	8	10	8	10
1/8	1/8	256	8	32	8	32
1/4	1/8	256	8	64	8	32
1/2	1/8	256	8	128	8	32
1/2	1/4	512	4	256	4	64
1/2	1/2	1024	2	512	2	128
1/2	1	2048	1	1024	1	256

Table 3: DDMA hardware search time parameters (Equivalent to the conventional correlator scenario in Table 1)

Parameter	Value	Parameter	Value
BW_{sv}	80,000 Hz	T_{IP}	0.032 s
C_{sv}	1023 chips	N_{NCR}	4
BW_c	8,000 Hz	$T_{1,1}$	10.23 s
C_c	128 chips	N_{sv}	32
$N_{1,1}$	79.92	N_c	1
		$T_{32,1}$	327.36

given by Table 2.

It should be recognized by the reader that the sizes of the memories and configuration of the block have been determined by the need to calculate DDMs and the available resources (memory and FPGA fabric logic) of the SmartFusion SoC. As will later become apparent, these are not necessarily optimized for the purpose of performing fast-acquisition

Additional Firmware Processing

In order to generate a complete DDM, or indeed to use the DDMA to search for satellites, it is necessary to perform additional processing of the DDMA output, as well as replenishing the input ping-pong buffer and output I&Q buffers. There are several reasons for this, but foremost among them was the need to limit the complexity and resources consumed by the hardware block.

The first task that needs to be performed is to convert a sequence of two 0.5 ms dump samples, four 0.25 ms dump samples or eight 0.125 ms dump samples into two, four or eight 1 ms dump samples. This is done by taking the input sequences and performing an FFT, which can be done efficiently because of the small size of the associated FFT. This process needs to be performed using all of the samples from a particular code-phase tap, and therefore needs to be completed across all of the code phase taps. The output of this process is a set of I&Q dump samples for each of the code-phase taps and produces frequency bins that are separated by 1 kHz. If frequency interpolation of the above results is required then the FFT should be zero padded and a larger FFT performed instead. However, if an interpolated FFT is carried out, it should be recognized that this is an interpolation and does not represent additional frequency resolution. This is generally acceptable for DDMs from a space-borne platform because the coherence interval of the signal is of the order of 1 ms [7], so additional coherent integration would not be beneficial.

The second process is to perform non-coherent integration of the 1 ms results produced in the previous step. Any results from the first stage can be non-coherently integrated until such time as the required signal-to-noise ratio is reached or the input IF signal stream has been fully processed.

Both of the above steps can be performed in parallel with the operation of the DDMA hardware block. This means that 1 ms worth of output can be processed by the DDMA while the previous 1 ms outputs get FFT'ed and non-coherently integrated. Recognising this fact can make a significant difference to the total processing time because even with hardware acceleration, calculating DDMs is still computationally intensive.

Cold-Start Using the DDMA

Hardware Configuration and Search Time

In order to use the DDMA to search for visible satellites, it is necessary to configure the hardware to search the maximum amount of frequency space and code-space in the minimum amount of time. For the DDMA as implemented on the Namuru V32R3 and described above, this amounts to setting the hardware dump interval to 0.125 ms/dump, which is the widest bandwidth, and selecting the widest possible code-phase separation, namely 0.5 chips. This configuration is highlighted in Table 2. It can be seen that this allows 128 chips and 8 kHz of bandwidth can be searched in 32 ms. Using these parameters, the hardware search time can be calculated as 328 seconds, or 5 minutes 28 seconds.

Firmware Processing

Generation of DDMs or the detection of visible satellites also requires firmware processing, as shown in Figure 4.

The first step, after reading the outputs from the DDMA hardware (and then starting the next batch of hardware processing) is to perform an 8-point complex FFT on each code phase tap. This transforms a sequence of 8 samples with a sample period of 0.125 ms into what would have been measured using 8 correlators each setup for a 1 ms integrate & dump interval. This needs to be performed 256 times every 32 ms and it is therefore important that this processing be fully optimised. In fact, such was the concern for this portion of the operation that a custom 8 point complex FFT routine was created, resulting in the FFT processing run time on an 80 MHz Cortex M3 processor being reduced from 24 us to 6 us.

The subsequent steps involve taking the magnitudes of each complex output and non-coherently integrating those outputs, as well as keeping track of where the peak output of the non-coherently integrations is to be found. At the end of 32 ms, a single round of non-coherent integration will have been performed and after 4 × 32 ms, a 128 chip × 8 kHz DDM will have been calculated with roughly similar sensitivity to the conventional correlator

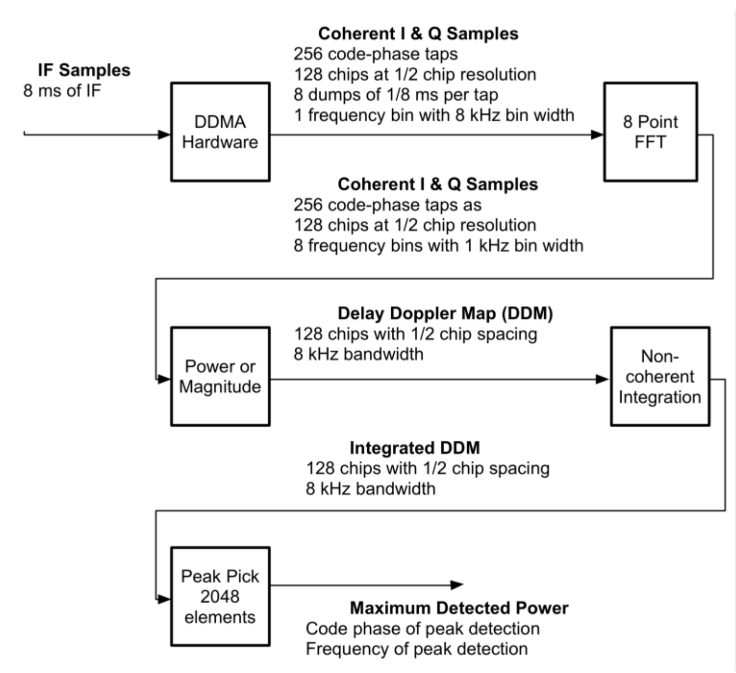


Fig. 4: Firmware signal processing

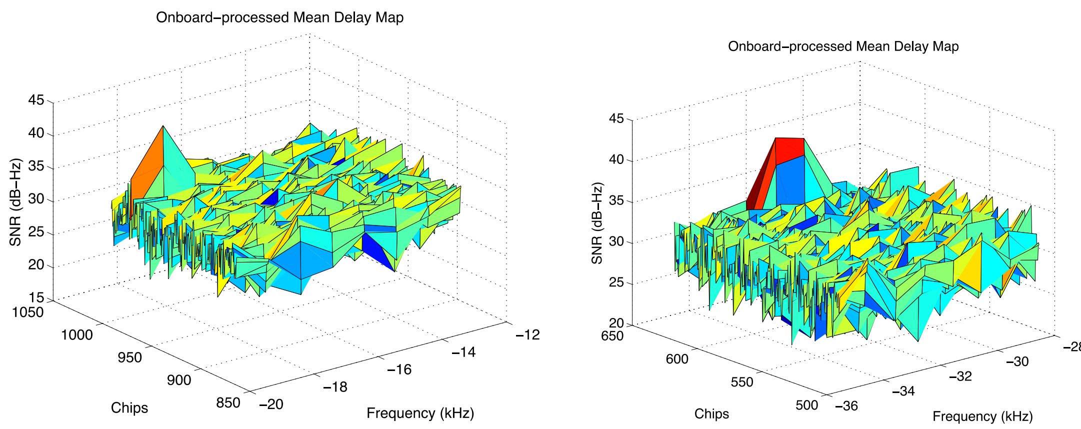


Fig. 5: DDMA detection correlations for SV1 (left) and SV11 (right)

scenario described earlier. A record of the peak value can then be compared with any previously detected peaks and updated should the new maximum be larger. This entire process is repeated for the same satellite until the total area of 1023 chips × 80 kHz has been covered by all of the smaller DDMs. It is not necessary to store all of the DDMs, only to record whether the peak power exceeds any previously detected peak power. At the end of the entire process, a decision can be made as to whether the largest detection is sufficiently large to represent a real signal.

Experimental Results

In order to validate these concepts, the firmware signal processing associated with this process was added to the Namuru V32R3 Aquarius firmware. This involved adding an additional task to the processing in order to sequence through the process of calculating each of the individual DDMs, performing the FFTs, accumulating the results until the full search space had been covered and recording the maximum detection for each satellite. The task automatically applied this process for every satellite in the constellation and returned the maximum signal for each satellite.

Following the firmware update, a GPS signal detection experiment for a LEO vehicle was created for an STR6560 GPS simulator and some preliminary testing performed with the Namuru V32R2. The results of this experiment are shown in Table 4 and Figure 5, where upon close examination it can be seen that although most of the signals were detected using the hardware, some were not. Part of the reason for these failures is due to the scalloping losses of up to 3 dB that occur when the signal to be detected lies midway between two FFT frequency bins. These losses are reduced for the conventional search because the frequency spacing between successive searches is set to 750 Hz rather than the maximum allowable 1 kHz. To eliminate this effect for the DDMA search, it would be necessary to zero pad the FFT input and to then perform an FFT, although this would come at the expense of additional processor loading.

Correlation loss is another cause of losses during detection, where this takes place when the peak signal correlation occurs midway between two code phase taps. This can result in an additional 2.5 dB of loss. The final cause of losses occurs within the non-coherent integration process. Normally when the coherent integration period is increased, there is 3 dB worth of signal to noise ratio (SNR) with each doubling in the integration period. However, this rule of thumb does not apply when non-coherently integrating because in such cases, small gains are observed when the input SNR is smaller. This is known as squaring loss [8].

Table 4: LEO GPS signal detection using DDMA

SV	Doppler (Hz)	C/N0 (dBHz)	DDMA SNR (dBHz)	DDMA Doppler (Hz)	Detected
1	-18,992	48	42.3	-18,000	Y
11	-31,856	46	40.7	-32,000	Y
13	6,918	48	41.4	7,000	Y
20	20,891	47	42.1	21,000	Y
23	17,267	48	41.9	18,000	Y
24	-20,775	46	38.1	-20,000	N
32	11,545	45	38.2	12,000	N

Actual correlations generated by the receiver are shown in Figure 5. The time to perform the entire search was longer than expected. Rather than the 327 seconds expected from the estimate of Table 3, the actual time was measured as 501 seconds or 8 minutes 21 seconds. This amounts to an increase by a factor of 1.53 and was caused by the firmware processing taking longer than the 128 ms for each DDM. Part of the reason is that the processor is also used to perform other GPS functions such as track satellites, extract data from the GPS navigation message and perform navigation solutions, so the full processor is not available for the processing.

Future Enhancements

Although the utility of the method as implemented was less than that envisaged from the original processing time estimates, the process of going through this exercise has resulted in some useful recommendations for improving the acquisition performance of future receivers, such as the new Kea receivers.

One way of improving performance is by modification of the DDMA, which if changed to allow 1/16 ms dump intervals would allow a further doubling of the bandwidth of each search. Even were the output memory to be held constant at 4096 × 16 bits, this would still allow 64 chips × 16 kHz to be calculated every 16 ms, thereby halving the time of the hardware portion of the search. This should come close to halving the time for the complete search, assuming that the processing capacity to perform the FFTs, non-coherent integrations and peak picking was available. This should be feasible given that the ARM Cortex M3 on the Kea board runs at 100 MHz compared to 80 MHz on the Namuru V32R3.

The second change concerns operation of the correlators themselves. Based on the optimised FFT, it is clear that the software overhead of adding additional frequency bins to the correlators is very low. This means that were the correlators modified to store 4 dump samples every millisecond rather than a single dump sample, this would be equivalent to increasing the number of effective correlators by a factor of 4. The hardware cost of such a feature would be the additional storage required to store each of the dumps, although with the availability of large FPGAs, this could be an acceptable tradeoff. The problems caused by scalloping losses could also be easily dealt with by zero padding the 4 complex dump samples with 4 zeroes and then performing an 8 point complex FFT, at a cost of 6 us per code phase tap. Indeed, this processing need not always be performed, but could be limited to those periods when a wider bandwidth was particularly useful, such as during a cold start or during very high dynamics.

Conclusions

In this paper, we first described the Namuru & Kea FPGA based GPS receivers and reviewed the difficulties associated with the acquisition of GPS signals in LEO. We also reviewed the DDMA hardware block originally designed for GPS reflectometry experiments and showed how the block could also be used for detection of GPS signals in LEO.

Experiments carried out on the Namuru V32Rx showed that the time for detection of signals during a cold-start could be reduced from a worst case of 776 seconds when using the conventional tracking channel correlators to approximately 501 seconds when using the DDMA. The time taken when using the DDMA was approximately 53% longer than what was expected given the original analysis due to higher firmware processing overheads than originally envisaged. Nonetheless, the method was still shown to be useful.

This analysis and work also resulted in some recommendations for future products. One of these recommendations is to allow the bandwidth of the DDMA to be wider than currently supported. The second recommendation is to use the same techniques that allow wider bandwidths for the DDMA to also be applied to the tracking correlators. This would allow faster cold-start even with the use of standard correlators. A fourfold reduction in cold-start TTFF should be achievable if these changes are implemented.

References

1. P. J. Mumford, K. J. Parkinson, and A. G. Dempster, "The Namuru Open GNSS Research Receiver," in *19'th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation* Fort Worth, TX, 2006, pp. 2847-2855.
2. E. Glennon, J. P. Gauthier, M. Choudhury, K. Parkinson, and A. G. Dempster, "Project Biarri and the Namuru V3.2 Spaceborne GPS Receiver," in *IGNSS 2013 Symposium*, Outrigger Gold Coast, Australia, 2013.
3. A. C. E. Valencia, J.F. Marchan-Hernandez, X. Bosch-Lluis, N. Rodriguez-Alvarez, I. Ramos-Perez, "Advanced architectures for real-time Delay-Doppler Map GNSS-reflectometers: The GPS reflectometer instrument for PAU (griPAU)," *Advances in Space Research,* vol. 46, pp. 196 - 207, 2010.
4. W. Li, D. Yang, M. Zhang, and Q. Zhang, "Design of a delay/Doppler mapping receiver for GPS remote sensing," 2009, pp. 765128-765128-6.
5. R. de Vos van Steenwijk, M. Unwin, and P. Jales, "Introducing the SGR-ReSI: A next generation spaceborne GNSS receiver for navigation and remote-sensing," in *Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC), 2010 5th ESA Workshop on*, 2010, pp. 1-7.
6. E. P. Glennon, K. Parkinson, J. W. Cheong, A. G. Dempster, C. Rizos, and K. Yu, "CubeSat-based GPS Reflectometry Using the Namuru V3.2 GPS Receiver," in *13'th Australian Space Science Conference*, Sydney, 2013.
7. S. Gleason, "Remote Sensing of Ocean, Ice and Land Surfaces Using Bistatically Scattered GNSS Signals From Low Earth Orbit," PhD University of Surrey, December 2006
8. F. van Diggelen, *A-GPS: Assisted GPS, GNSS, and SBAS*: Artech House, 2009.

Multi-Antenna Switching for Spaceborne GPS Receivers

Scott O'Brien [2], Kevin Parkinson [2,3], Eamonn Glennon [1], Joon Wayn Cheong [1], and Andrew Dempster [1,2]

[1] *Australian Centre for Space Engineering Research (ACSER)*
[2] *School of Electrical Engineering and Telecommunications (EET)*
University of New South Wales (UNSW), Kensington, NSW 2052, Australia
[3] *General Dynamics Ltd, New Zealand*

Summary: This paper gives an overview of the requirements, development and testing of an antenna switching daughterboard for the UNSW Educational CubeSat Zero project. This board has been designed to support multiple GNSS antennas, pointing in various directions, for a number of GNSS-based experiments including single frequency radio occultation using the GNSS satellites, and earth remote sensing by way of reflections of the GNSS satellites off the Earth's surface. An assessment of the actual antenna switching circuitry, by way of in-circuit measurement of the switch isolation, and signal attenuation is reported. Finally an overview of the next stages in the development of the EC0 CubeSat are given.

Keywords: EC0, UNSW, ACSER, QB50, GNSS-R, reflectometry, GNSS-RO, radio occultation

QB50 Project Overview

The QB50 project is part of the European Commission Seventh Framework Programme effort to, amongst a variety of other goals, facilitate access to space. The goal is for universities around the world to develop 50 CubeSats that will contain a supplied sensor payload for atmospheric monitoring of the thermosphere (a choice of three different sensors is offered), in addition to any other payloads developed by the individual universities [1].

The project is being managed by the Von Karmen Institute for Fluid Dynamics (VKI), a non-profit educational and research organization based in Brussels. VKI are responsible for supplying the science sensor payload, and arranging the launch and deployment of the satellites, currently targeted for January 2016 using a Cyclone-4 rocket launched from Brazil.

The three VKI-supplied sensor payloads that participants may choose from are:

- Ion-Neutral Mass Spectrometer (INMS)
- Flux-ф-Probe Experiment (FIPEX)
- Multi-Needle Langmuir Probe (m-NLP)

These sensors are designed to measure the Earths' thermosphere – the region in the atmosphere between 200 and 380 km in altitude. This region has historically been little understood or examined due to the difficult nature of obtaining sufficient useful measurements. Current methods are restricted to sensors on vehicles in highly elliptical orbits, or sounding rockets, but both these provide only highly temporally limited and spatially restricted results. The QB50 project on the other hand will provide many measurements at multiple locations simultaneously over, hopefully, a number of months.

EC0 Mission Overview

The UNSW QB50 entry is known as Educational CubeSat Zero (EC0). In addition to the chosen VKI-supplied INMS atmospheric sensor payload, a number of other scientific payloads are currently under development at UNSW for inclusion in EC0.

These EC0 science payloads include a fully 3D-printed structure, a micro-kernel operating system evaluation experiment, an FPGA single event upset recovery monitoring experiment, and a number of Global Navigation Satellite Systems (GNSS)-based experiments [2]. These GNSS-based experiments are the focus of this paper.

The GNSS-based experiments include the space heritage-proving of the UNSW-developed Namuru / Kea GNSS receiver board (shown in Fig. 1) [3], a GNSS-based radio occultation (GNSS-RO) experiment, and a GNSS-based reflectometry (GNSS-R) experiment.

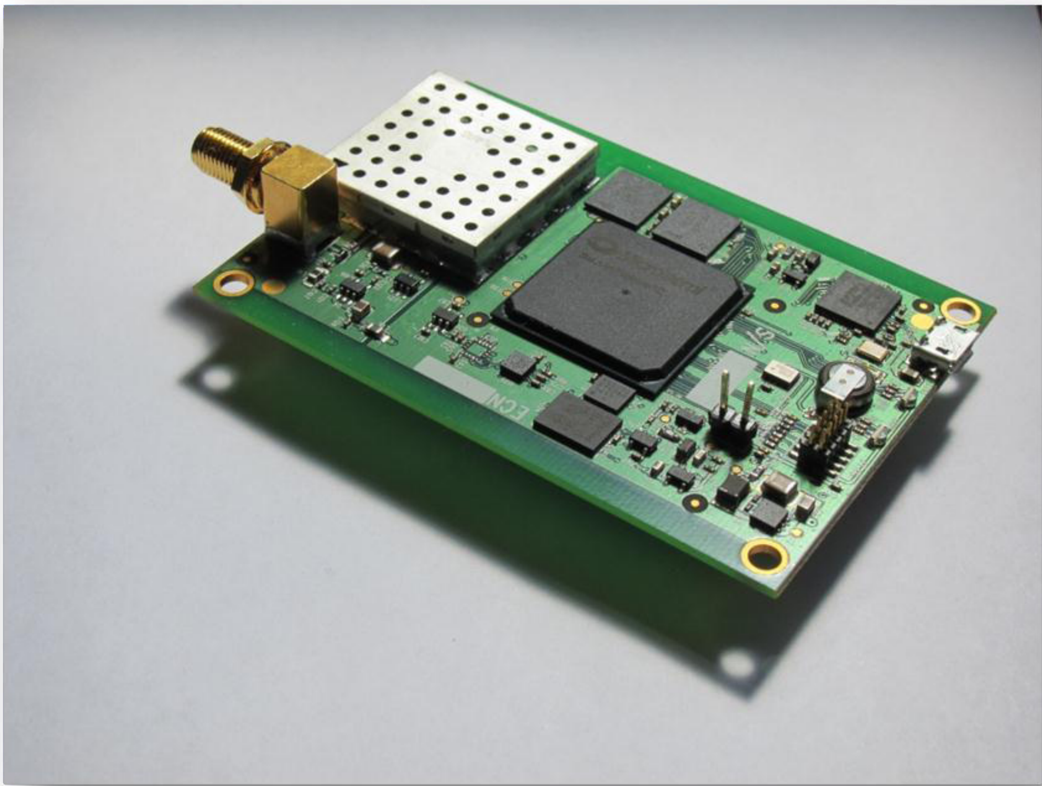


Fig. 1: The UNSW-developed Kea GNSS receiver

A complication arises when considering the antenna pointing requirements for each experiment. The position, velocity, and timing (PVT) operation of the receiver requires a zenith-pointing antenna to receive the direct signals from the GNSS satellites. The GNNS-RO experiment requires a horizontally pointing antenna to observe GNSS signal occultations. Finally, the GNSS-R experiment requires a nadir-pointing antenna to observe the GNSS signals reflected off the surface of the Earth [2].

One option is to employ a single receiver antenna and reorient the spacecraft as required. This option has been discounted for four reasons. Firstly reorientation would involve expending power, power that is simply not available in the space- and weight-limited 3U CubeSat form-factor. Secondly any reorientation would necessarily mean a period of loss of communication with ground stations due to communications transmitting antenna placement. Thirdly, the antenna for the GNSS-R experiment is ideally left-hand circularly polarised (as opposed to right-hand circularly polarised for normal PVT operation and GNSS-RO operation). Finally, it would mean that only one GNSS-based experiment could be undertaken at any point in time. In other words, PVT would not be available whilst performing GNSS-R or GNSS-RO; having PVT information available during those experiments is critical for the science.

The second, and chosen option, is to have multiple antennas pointing as required, and switching between them as required. The spacecraft has therefore been designed with a zenith antenna for PVT, a horizontal antenna for GNSS-RO, and a suitably-polarised nadir antenna for GNSS-R.

To enable switching between these 3 antennas, an antenna switching daughter-board (ASDB) has been developed. This ASDB, the blue PCB pictured in Fig. 2 during testing, is designed to plug directly into both the Namuru / Kea receiver, and the spacecraft bus.

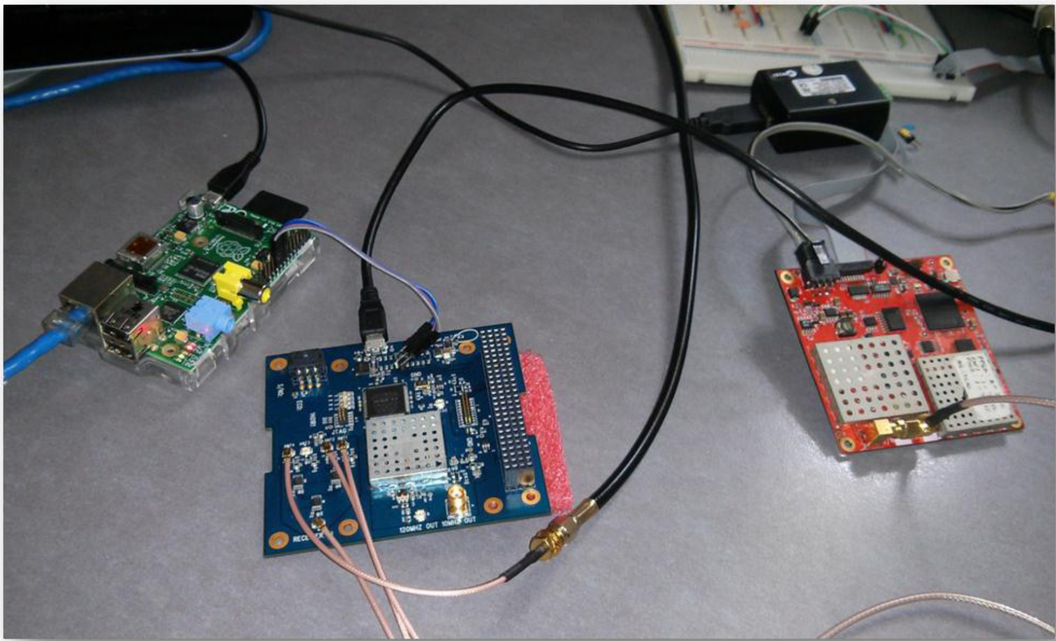


Fig. 2: Testing the ASDB switches

ASDB Design

The requirements for the ASDB included the ability to switch between a number of antenna inputs, under control of the flight computer and the Namuru / Kea receiver, as determined by mission planning requirements. Additional requirements for the EC0 mission, and other subsequent missions, means the board also has the ability to output additional clock signals at varying frequencies, hence the other circuitry visible on the board. This additional circuitry is not discussed in this paper.

The Peregrine Semiconductor PE4250 SPDT RF Switch chip was chosen to perform the required switching. According to the product data sheet [4], this switch has a low insertion loss, approximately 0.70 dB @ 1.5 GHz @ 25°C, and a high isolation of approximately 48 dB @ 1.5 GHz @ 25°C. Both these factors are of particular importance to GNSS-R due to the very low signal reception levels likely to be experienced. Both are also expected to have slightly improved performance at the lower temperatures likely to be experienced in space. The switching circuit design is shown in Fig. 3.

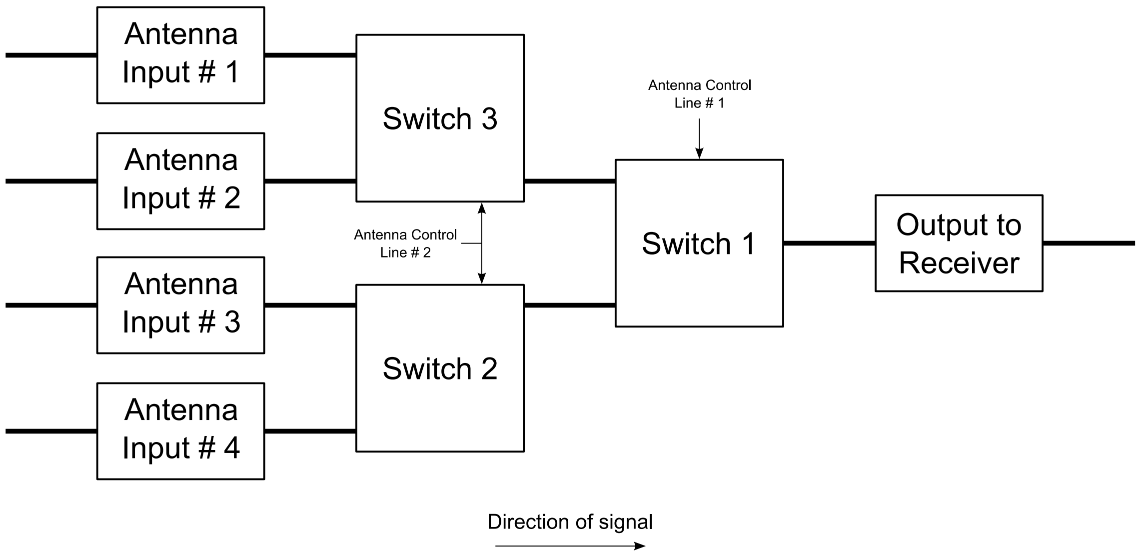


Fig. 3: ASDB switching circuit design

A further consideration for the antenna switching is the supply of power to the active receiver antennas. It is common to supply power to the amplifiers built-in to the antennas via a DC bias in the RF signal line. This is commonly called a bias-T. A detailed section of the circuit schematic showing the bias-T arrangement is shown in Fig. 4.

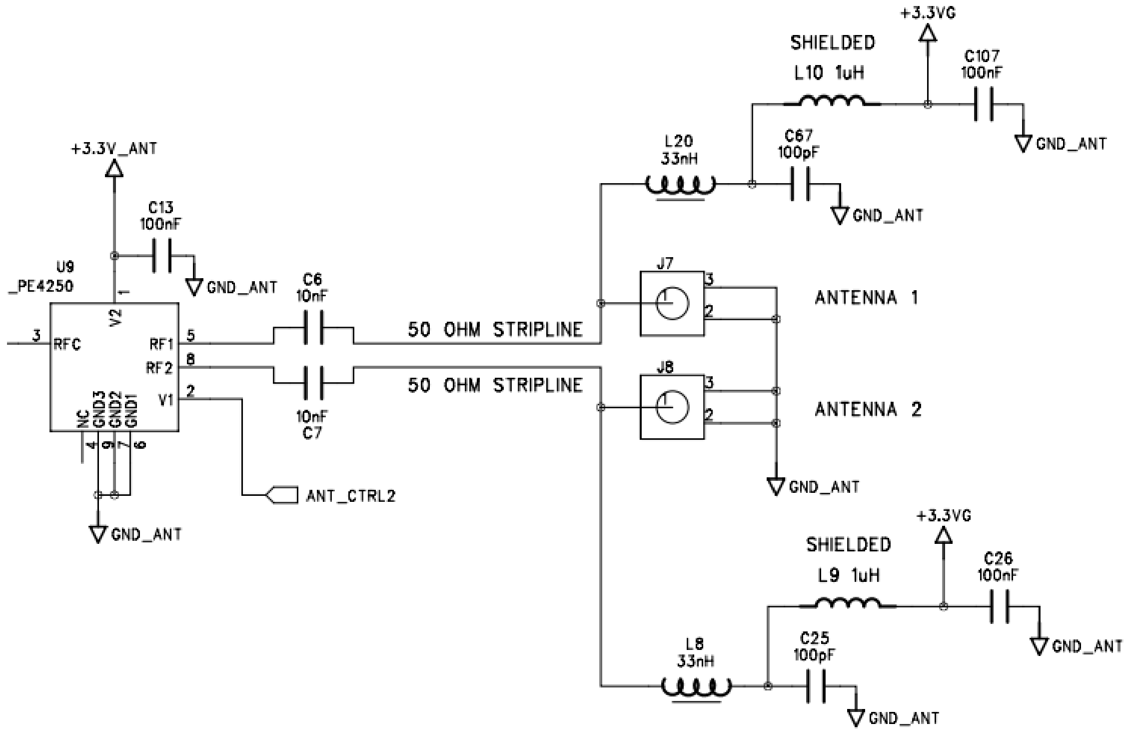


Fig. 4: ASDB bias-T detail from the circuit schematic

ADSB Testing

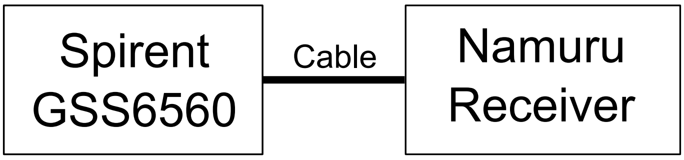
To ensure the ASDB switching performed adequately, bench tests were performed on the switching mechanism, the signal attenuation incurred, and the switch isolation.

Switch Operational Testing

To ensure the switches were actually working, a Namuru v3.2 receiver was connected to a GPS antenna via the ASDB switches. Control of the switches was via Python code running on a Raspberry Pi single board computer. Each switch output was tested and as a PVT solution was obtained in each case, the switches were determined to be working as expected.

Signal Attenuation Testing

To test signal attenuation, a Namuru v3.2 receiver was connected to a GPS signal simulator, a Spirent GSS6560, with and without the ASDB in the signal path. These configurations are shown in simplified form in Fig. 5.

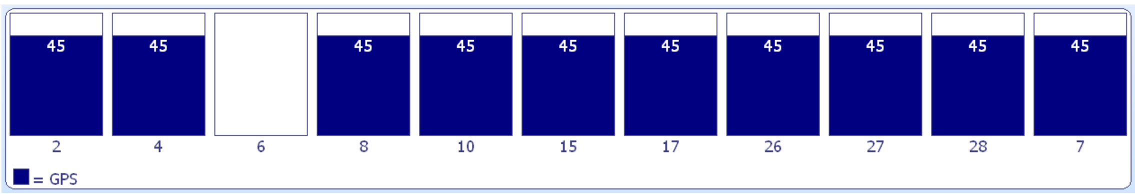


(a)

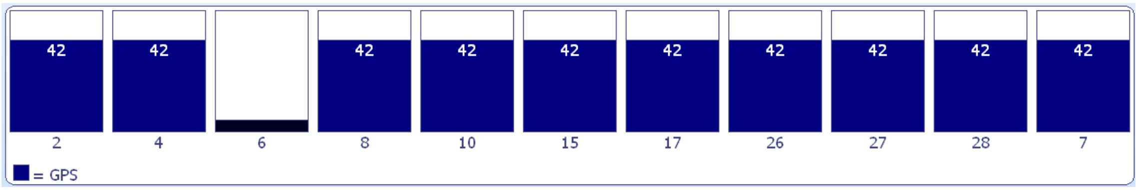
(b)

Fig. 5: Attenuation test configuration a). without ASDB, and b). with ASDB

By noting the C/N_0 values output by the Namuru receiver in its NMEA messages, any signal attenuation variations can be seen. As can be seen in Figure 6, the results with and without the ASDB for the visible GPS satellites show a 3 dB signal level reduction.



(a)



(b)

Fig. 6: C/N_0 values for each visible satellite as output by the Namuru receiver a). without ASDB, and b). with ASDB

From the PE4250 datasheets, a minimum of 1.4 dB (2 x 0.7 dB) was expected. The difference between the observed 3 dB and 1.4 dB can be attributed to NMEA message value rounding, and additional attenuation due to long cables and additional connectors converting one connector type to another (uFL to SMA in this case). Once installed on the spacecraft the attenuation should be lower. A 3 dB attenuation level is not expected to be a problem for any of the experiments.

Switch Isolation Testing

To test each individual switch isolation, an Agilent N9310A RF Signal Generator was connected to an Anritsu MS2711D Spectrum Master spectrum analyser, with the ASDB in the signal path. This configuration is shown in simplified form in Fig. 6.

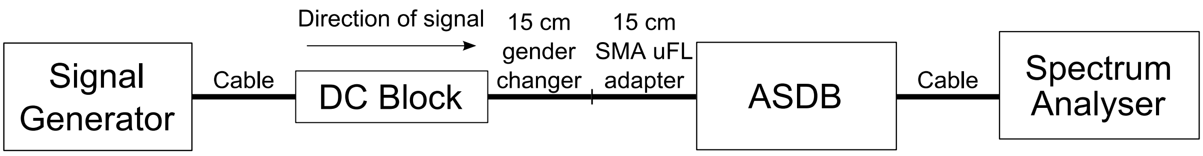


Fig. 6: Switch isolation test configuration

A signal level of -30 dBm at 1.57542 GHz was output by the signal generator. The signal was connected to each antenna input in turn, and the output signal level was then measured by the spectrum analyser on the output port to the receiver in all four switch positions. The measured results are shown in Table 1.

Table 1: Switch isolation measurements

Signal into antenna #	Antenna switch #	Measured value (dBm)
1	1	-35.5
	2	-88
	3	-76
	4	-81
2	1	-90
	2	-35.4
	3	-86
	4	-75
3	1	**-71**
	2	-93
	3	**-35.4**
	4	-78
4	1	-92
	2	-73
	3	-76
	4	-35.4

As can be seen in Table 1, the minimum isolation level is when a signal is input into antenna input 3, and antenna 1 is selected by the switch, giving an isolation value of -35.4 dBM - (-71 dBm) = 35.60 dB. This figure is lower than expected and can most likely be explained by crosstalk from the wiring configuration during the testing, as shown in Fig. 7. Coupling on the PCB may also have an effect. Once installed on the spacecraft the isolation should be higher, but this level is not expected to be a problem for any of the experiments.

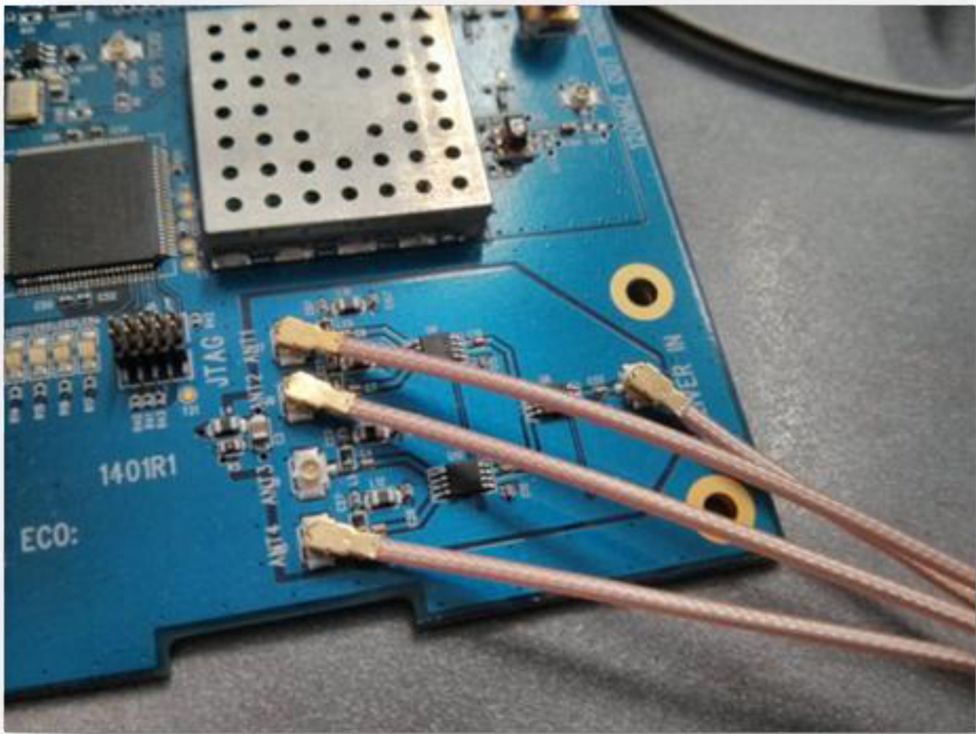


Fig. 7: Switching circuitry and test wiring detail

Future Work

Going forward, development continues on the software and procedures to control the antenna switching. In addition, software development continues on the Namuru / Kea receiver to facilitate extraction of raw science data for the GNSS-R and GNSS-RO data.

Hardware testing is also planned by means of high altitude balloon flights. After assembly and integration, thermal vacuum testing of the satellite is planned using the new test facilities at the Advanced Instrumentation and Technology Centre at Mt Stromlo, Canberra.

Conclusion

This paper has given an overview of the requirements, development and testing of the ASDB for the UNSW EC0 CubeSat project. This ASDB has been designed to support multiple GNSS antennas, pointing in various directions, for PVT, GNSS-R, and GNSS-RO experiments. An assessment of the actual antenna switching circuitry, by way of in-circuit measurement of the switch isolation, and signal attenuation was performed and the switching was found to be working largely as expected.

References

1. von Karman Institute for Fluid Dynamics. (2013, 8 August 2013). *QB50 Home Page*. Available: https://www.qb50.eu/
2. Li Qiao, Eamonn Glennon, Andrew G Dempster, and Sebastian Chaoui, "Using CubeSats as platforms for remote sensing with satellite navigation signals," in *IEEE International Geoscience and Remote Sensing Symposium IGARSS 2013*, 2013, pp. 852-855.
3. E. Glennon, J. Gauthier, M. Choudhury, K. Parkinson, and A. Dempster, "Project Biarri and the Namuru V3.2 Spaceborne GPS Receiver," in *International Global Navigation Satellite Systems Society Symposium IGNSS 2013*, Gold Coast, Australia, 2013.
4. Peregrine Semiconductor. (2009, 3 November 2014). PE4250 SPDT UltraCMOS RF Swtch, Product Specification. [Product data sheet]. Available: http://www.psemi.com/pdf/datasheets/pe4250ds.pdf

The Kea GNSS Receiver for Spacecraft Operations

K J Parkinson, E P Glennon, and A G Dempster

University of New South Wales, Sydney, NSW, Australia

Summary: Field Programmable Gate Array (FPGA) GNSS receivers have continued to develop at UNSW as research has progressed. As a result, a low cost receiver, the new V4 Kea GNSS platform has been developed for space and terrestrial operations. Although not fully space qualified in terms of radiation tolerance, the space version of the Kea receiver is designed to recover gracefully using redundancy where possible to maintain operational status in the harsh environment of space. The operational failure risks associated with single event upsets (SEU) in volatile memory are reduced through the use of careful design and selection of components. Factors affecting power consumption are carefully minimised to avoid overheating in zero atmosphere. The receiver is built on a custom designed printed circuit board, uses custom base-band logic in the FPGA portion of a System On Chip (SOC) component with supporting application firmware. The combined techniques presented deliver a more affordable solution where full space qualification is not required. This paper describes the architecture, design challenges and development path of the Kea receiver with features aimed at delivering robust performance and flexibility for space operations.

Keywords: Kea, Space, Receiver, FPGA, Base-band

Introduction

Small satellite and sounding rocket missions have usually been based on commercially available GNSS receiver products. These often prevent the implementation of anything further than tracking the GPS L1 signal within the limits of the receiver's capability. The receiver internal architecture is usually proprietary with the software and base-band not accessible to the end user.

In 2004 the Namuru GNSS receiver platform was developed at the University of New South Wales (UNSW) as reference design to support ongoing research resulting in many successful applications and research projects . One of the main features of the Namuru receiver is that it is fully open source which allows access to the previously unavailable receiver internals. An FPGA is used for both the base-band and the soft core processor which allows extreme flexibility. Although not specifically designed for space applications the Namuru V2 receiver was tested to determine its suitability for sounding rocket experiments at the German Space Centre (DLR) (A. Grillenberger, et al., 2008). After finding that the receiver was able to meet their expectations with minor modifications, it was successfully launched on a sounding rocket mission in April 2010 by DLR (A. Grillenberger and M. Markgraf, 2011). Because this mission was by design quite short, there was no prolonged exposure to environmental factors that may impact operation. However, the mission was entirely successful and the Namuru V2 receiver performed well.

In late 2010 the Garada project (GARADA, 2010), funded by the Australian Space Research Program (ASRP), was launched at the Australian Centre for Space Engineering Research (ACSER) located at UNSW. This collaborative project developed the design, the technologies and the business case for a constellation of low cost and light-weight Synthetic Aperture Radar (SAR) satellites that can acquire images of the earth at night and in any weather. Part of this research further developed the UNSW GNSS platform for deployment on satellites with prolonged exposure to the space environment. Two receivers were developed: one for GPS L1, and a more sophisticated L1/E1/L5/E5 receiver.

In the following sections we discuss the issues relating to the development of new versions of GNSS receivers, some that have been designed to operate reliably in space.

1 The changing approach to space electronics

Electronic parts that are capable of withstanding space environments are known as radiation hardened or “rad-hard” parts. While they can operate very reliably in harsh space environments they are more expensive and often less advanced than today’s commercially available components. Radiation tolerant components of this calibre were originally designed for long range deep space missions where the environment is more challenging and the mission costs greater. Only the well funded military and aerospace programs have the budgets to use these parts.

More recently, there has been a move away from using traditional radiation hardened parts generated by the increasing commercialisation of space. Less emphasis is being placed on radiation hardening for some Low Earth Orbit (LEO) missions because the environment is less challenging in this area of space. There is a desire to keep costs down and missions are less critical, thus commercially available components can be used coupled with design strategies to allow failure detection and system recovery. As a result it has become more affordable to build commercial LEO satellites.

2 Risks and strategies

There are a number of risks that need to be considered when designing and building any electronic equipment intended to operate reliably in space.

2.1 Mechanical construction

As expected, the greatest mechanical vibration and stress is present at the time of launch. The mechanical design must take this into account in order to survive beyond the launch. The selection of bonding materials and printed circuit board construction plays an important role in this.

The modern lead-free component soldering materials require precise and elevated temperatures during manufacture to make reliable electrical connections. To ensure quality levels X-ray techniques should be used to confirm connection bonding which drives up cost of manufacture.

Multilayer printed circuit boards use small holes known as "vias" that are plated with copper to create connections between electrical layers. The plating is very thin and often breaks down under mechanical and thermal stress. The thin plating does very little to conduct heat between layers. The solution is to fill the via hole with solid copper by either over plating or inserting copper plug.

2.2 Temperature range

The space environment can present a wide range of temperatures, often passing through cyclic variations caused by spacecraft rotation and intermittent solar exposure. Under these conditions continuous thermal stress is applied to components as they pass through the temperature range leading to joint failures and sometimes mechanical failure.

2.3 High vacuum

Many electronic components typically give off a small amount of vapour known as "out gassing" long after manufacture as the bonding materials continue to settle down and cure. This increases in the vacuum of space sometimes leading to added stress and early failure through circuit breakdown. Careful component selection is essential using only solid state passive devices such as multi-layer ceramic capacitors and inductors.

A more critical problem is the thermal design in a high vacuum where the absence of an atmosphere limits the ability of components to dissipate heat in the normal way. In general this can be solved by using active components rated at three or four times the required capacity for very little extra cost. However, in most cases careful management of the surface area and understanding thermal resistance issues will eliminate the risks as found in the tests on the earlier Namuru V2 receiver at DLR (A. Grillenberger, et al., 2008).

2.4 Radiation exposure

The radioactive environment of space is generally destructive to most commercial semiconductor devices due to the presence of trapped electrons, trapped protons, solar protons and cosmic rays as described in (M. Dowd, 2010). These conditions cause random operational failures in semiconductor devices in three main categories: Total Ionisation Dose (TID), Single Event Latch-up (SEL) and Single Event Upset (SEU).
TID is caused by energised particle build up over time in semiconductor devices which leads to total device failure by eliminating their ability to operate as a semiconductor.

SEL is when highly charged particles cause a semiconductor device to be driven out of specification and to draw excessive continuous current leading to device destruction through overheating as described in (D. Layton, et al., 2010).

SEU is caused by a highly charged particle disturbing the contents of a data storage element, such as a memory cell, leaving it with incorrect data.

The effect of each of these conditions is shown in (Table 1).

Table 1. Space radiation effects.

Fault Type	Digital Circuit	Analog Circuit	Recovery Possible	Recovery Method
TID	Failure	Failure	No	None
SEL	Failure	Failure	Yes	Power cycle
SEU	Data loss	Spike	Yes	Data error correction

Flash memory provides some immunity to data corruption from SEU because the semiconductor architecture of each memory cell is non-volatile and constructed using larger geometry capable of dissipating energy more quickly. While not completely immune, this type of memory is able to withstand greater levels of radiation.

3 The Kea V4 GNSS Space Receiver

The Kea V4 receiver (Figure 1) adopts the approach of using less costly commercially available components and applying reliability enhancing techniques to reduce the operational risks when operating in space.

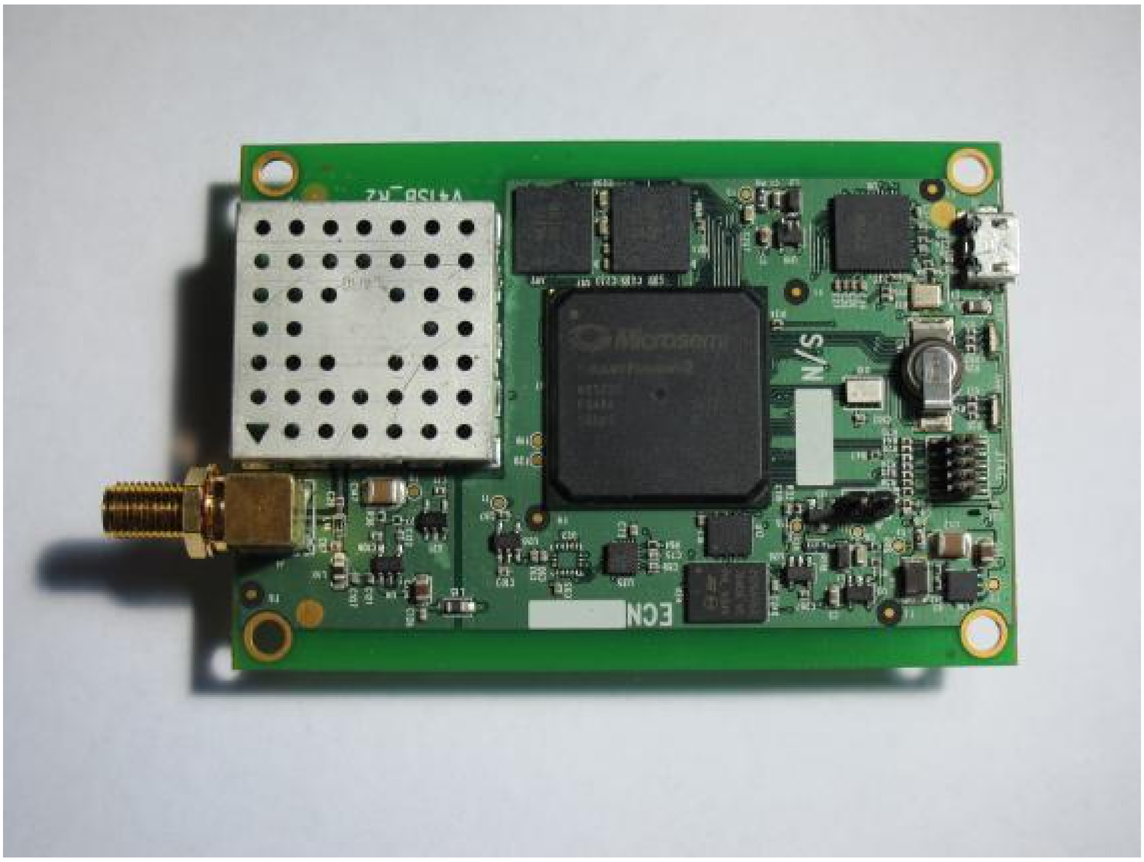


Figure 1. Kea V4.1SB Receiver.

While not intended for mission critical or safety of life applications, the receiver is suitable for operation on LEO space missions. It is built on a custom Printed Circuit Board (PCB) using the major components shown in (Figure 2).

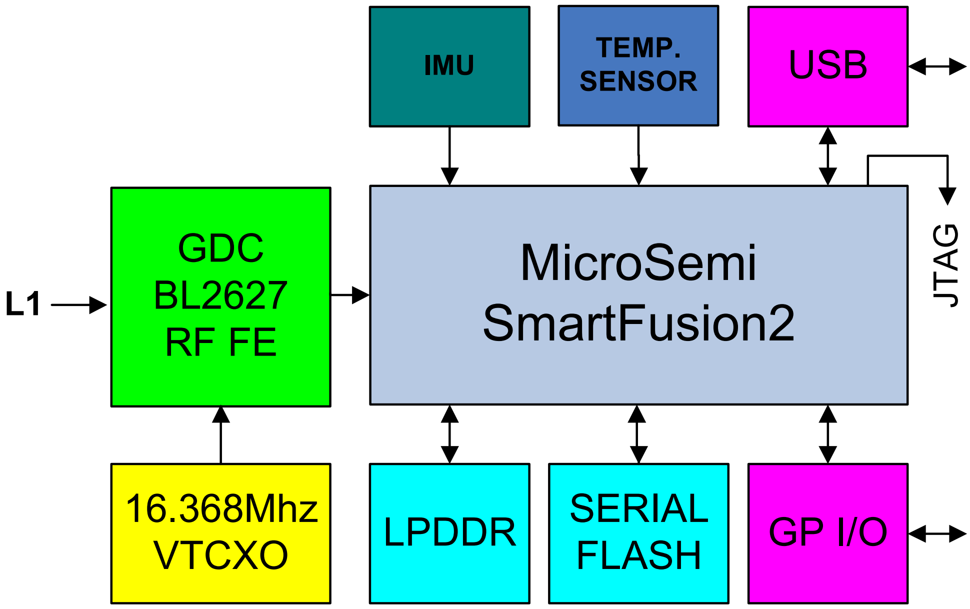


Figure 2. Kea V4.1SB GNSS Receiver.

The receiver design uses non "rad-hard" components throughout with some important features that have been added to improve reliability and provide fault tolerance as follows:

3.1 CPU and FPGA

On the digital side of the receiver the main component is a SmartFusion2 SOC device made by MicroSemi with an embedded 166MHz ARM Cortex M3 processor (MicroSemi Inc, 2013). This flash based device provides all CPU functions combined with a range of digital inputs and outputs and separate section of programmable logic. The base-band signal processing functions are contained in the programmable logic that is internally connected to the bus of the ARM processor. The advantage of using these flash based SmartFusion2 parts is that the ARM software and the FPGA logic image is stored in the on-chip flash which greatly reduces the risk of errors due to SEU problems compared with using Static Random Access Memory (SRAM) as on previous Namuru receiver generations. Power consumption is also reduced by not using on-chip SRAM. MicroSemi also offers more costly versions of pin compatible parts with a full radiation hardened specification. These can be fitted to the Kea V4 receivers in place of the commercial parts where required as an upgrade.

3.2 Oscillator

The main oscillator for a GNSS receiver needs careful attention, often requiring short term stability of better than 2.0ppm. As discussed, in space the oscillator will most likely need to maintain stability over a greater range of temperatures. The Kea V4 receivers are equipped with a Voltage Controlled Temperature Compensated Oscillator (VCTCXO) allowing the option to pull the oscillator up to 5ppm either side of the nominal frequency. A software algorithm is used to drive the VCTCXO with corrections while monitoring a precise temperature sensor. As a further option, an Oven Controlled Crystal Oscilator (OCXO) with the same frequency adjusting capability can be fitted.

3.3 Radio Frequency Front end

A custom Application Specific Integrated Circuit (ASIC) Radio Frequency (RF) down converter is used on all Kea receivers. This chip has been developed for robustness and the space environment in mind. It uses 130nm silicon which is not small geometry by today's standards and as a result allows greater channel width in the silicon to deliver more robust SEU mitigation. Low Voltage Differential Signal (LVDS) interface components have been used to minimise voltage gradients between the RF and digital sections. This significantly reduces the likelihood of SEL problems in the digital sections because of the low voltages and low impedances. The differential nature of the digital signals also helps minimise noise contamination in the RF section from the digital section. Other techniques drawn from the earlier Namuru V2 receivers have been applied to also control digital noise in the receiver (K. J. Parkinson, et al., 2006).

3.4 Construction

Traditional tin-lead paste alloy is used on the Kea V4 space receiver because it is well understood and found to be reliable for component connections, as opposed to some of the recent lead free products which have not yet been proven for long term reliability (D. R. Frear, et al., 1994).

The multilayer printed circuit board is constructed with solid copper plating or plugs through via holes for electrical connection between layers. This gives greater reliability under variable temperature conditions and improved heat transfer to the inner layers where larger copper areas can dissipate heat.

3.5 Random Access Memory (RAM)

The ARM processor uses a small amount of on-chip SRAM and 64Mbyte of external, fast, low power, double data rate Dynamic RAM (DRAM). Both of these areas are vulnerable to SEU problems. However, the Smartfusion2 SOC has been designed with SEU mitigation in mind to the extent that most of the on board silicon has been built to withstand SEU attacks and to recover dynamically. The internal memory and the external DRAM controller contain silicon features to deliver Single Error Correction with Double Error Detection (SECDED). Recovery from double error detection and beyond can then be handled in software. With the processing speed available, and the large SRAM capacity, a software redundancy mechanism can be employed in the Kea receivers to create double redundancy in multiple RAM locations to overcome data corruption.

3.6 Power supplies

All power rail voltages are derived from high frequency (1MHz and above) high efficiency switch mode regulators to minimise heat dissipation and reduce power consumption. In general each switch mode supply is capable of at least twice the required current to ensure low thermal dissipation especially while recovering from latch-up conditions. Each supply also has the ability to switch over to cycle skip mode on light loads to economise on power consumption.

3.7 Shielding

The RF front end section is shielded using a metal enclosure. Additional shielding between circuit sections is provided using careful printed circuit board layout by dividing large copper plane areas into segments. This ensures less coupling between low and high signal level areas. Some consideration is being given to the new PolyRAD shielding material being developed for NASA that is predicted to substantially reduce TID (NASA, 2002).

4 Conclusions

A low cost space capable GNSS receiver development has been described using careful selection of commercial components. The design exploits the flexibility of more complex modern components without the need for radiation hardening. Fault tolerance and recovery is delivered which is suitable for missions that can tolerate a lesser degree of assurance.

5 Acknowledgments

This work is supported by the Australian Centre for Space Engineering Research (ACSER) and the Satellite Navigation, Positioning Laboratory (SNAP) at the University of New South Wales (UNSW), and General Dynamics Ltd.

References

1. Dowd, M.(2010) How Rad Hard Do You Need? The Changing Approach To Space Parts Selection?, Maxwell Technologies, San Diego, CA

2. Frear, D.R., Burchett, S. and Morgan, H.S. (1994) *The Mechanics of solder alloy interconnects*. Springer.

3. www.garada.unsw.edu.au, Garada, 2010

4. Grillenberger, A. and Markgraf, M. (2011) Flight test results of a novel integrated GPS receiver for sounding rockets. in *20th ESA Symposium on European Rocket & Balloon* (Hyeres, France),

5. Grillenberger, A., Rivas, R., Markgraf, M., Mumford, P.J. and Parkinson, K.J. (2008) THE NAMURU RECEIVER AS DEVELOPMENT PLATFORM FOR SPACEBORNE GNSS APPLICATIONS. in *Navitec 2008*, Noorwijk, The Netherlands

6. Layton, D., Czajkowski, D., Marchall, J., Anthony, H. and Boss, R.(2010) Single Event Latchup Protection Of integrated Circuits, Maxwell Technologies, San Diego, CA

7. Microsemi Inc.(2013) SmartFusion2 SOC Product Brief

8. www.sbir.gov/sbirsearch/detail/218094, NASA, 2002

9. Parkinson, K.J., Dempster, A.G., Prof, Mumford, P.J. and Rizos, C., Prof. (2006) Improving signal quality in FPGA based GPS receiver designs. in *IGNSS Symposium 2006* (Surfers Paradise, Australia), International Global Navigation Satellite Systems Society

Space-borne GNSS Based Orbit Determination using a SPIRENT GNSS Simulator

Sanat Biswas*, Li Qiao†, Andrew Dempster*

* *Australian Centre for Space Engineering Research, The University of New South Wales, NSW, Australia, 2052*
† *School of Engineering and Information Technology, The University of New South Wales, Canberra, ACT, Australia, 2600*

Summary: Ground simulation is essential for assessing the performance of space-borne Global Navigation Satellite System (GNSS) based positioning algorithms and hardware. SPIRENT GNSS simulator is widely used for establishing the simulation environment. This paper presents a new approach to simulate a space-borne GNSS based positioning scenario in a SPIRENT simulator using Precise Orbit Determination (POD) data. A methodology is developed to convert the available POD data to the SPIRENT-compatible format. The POD data of TANDEM-X satellite was used to define the user satellite motion. The recent GPS navigation data is used to define the motion of the GPS satellites and hypothetical Galileo constellation orbits are used to define the motion of the Galileo satellites. This setup is used to simulate a multi-GNSS based navigation scenario. The simulated code and carrier phase measurements are recorded and then used in a MATLAB based Kalman Filter routine to estimate the position of the satellite.

Keywords: GNSS, Kalman Filter, orbit determination, simulation

Introduction

Use of multiple Global Navigation Satellite Systems (GNSS) for real-time on-board navigation of LEO satellites is an intriguing and cost effective technique. Landsat-4 was the first satellite to carry a space-borne GPS receiver [1] . For precise positioning by post processing, the use of GPS flight data has been a well accepted technique since then. The success of TOPEX/Poseidon, GRACE and CHAMP missions has proven the use of the GPS in LEO missions to be a low cost and simple solution. The possibility of multi-GNSS receiver applications in space missions is being assessed [2]. Numerous simulations have to be executed to assess performance of the space-borne GNSS receivers, hardware modules and the associated algorithms. SPIRENT GNSS simulator is widely used to simulate the GNSS constellations, signals transmitted by the navigation satellites, signal received by the user vehicle and the motion of the user vehicle. For characterization and testing of UNSW Namuru space-borne multi-GNSS receiver, a SPIRENT simulator was used to simulate the user satellite motion using the inbuilt orbit model and the provided initial conditions [3, 4]. The accuracy of the considered true position of the satellite in the simulation experiment depends on the accuracy of the inbuilt orbit model of the simulator. In this paper, a new method is proposed to design a more realistic simulation scenario for space-borne GNSS based orbit determination. Precise Orbit Determination (POD) data is used in this method to define the exact motion of the user satellite instead of using the SPIRENT generated user satellite motion. These data are highly accurate and provided in Earth Centred Earth Fixed Frame (ECEF). POD data of several satellites are publicly available. Most of the time the available data format is not compatible with SPIRENT. A methodology is developed in MATLAB to convert the available POD data to a SPIRENT-compatible format. Use of POD data enables a high fidelity Space-borne GNSS based orbit determination simulation using a SPIRENT GNSS simulator and facilitates efficient evaluation of position estimation algorithms. From the simulated scenario, the code and carrier phase measurements are recorded and then used in a MATLAB based Kalman Filter routine to estimate the position of the user satellite.

Simulation Procedure and Assumptions

The SPIRENT GSS8000 GNSS simulator is capable of simulating the GPS, GLONASS and Galileo signals. This simulator takes the vehicle motion as input, generates the GNSS satellites' positions from the ephemeris and simulates the signals to be received by the user GNSS-receiver from the Visible satellites. The simulator can generate tropospheric and ionospheric errors in the measurements and also can incorporate receiver clock bias, if the actual receiver is not available in the simulation. Fig. 1 shows the basic operation of the SPIRENT simulator. For the LEO satellite position estimation simulation without any GNSS

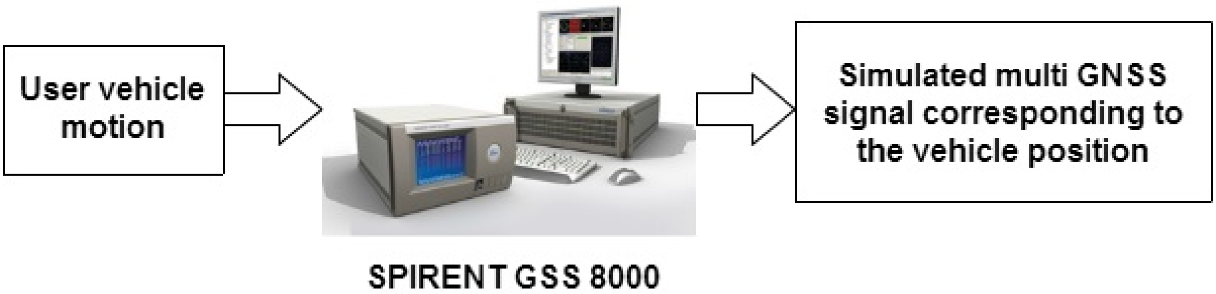


Fig. 1: SPIRENT simulator operation

receiver in the process, a reference orbit of the user satellite was defined in the SPIRENT GNSS simulator. Based on the reference orbit, the simulator generates GNSS signals and measurements. These simulated measurements were recorded and used in the MATLAB environment to estimate the position of the user satellite.

Use of Reference Orbit in SPIRENT

The SPIRENT GNSS simulator has a provision for defining the Keplerian orbit parameters to describe the orbit under consideration. These parameters are used in a pre-defined orbit model to obtain the reference position of the satellite. Another way of providing the reference orbit to the simulation scenario is directly defining the position of the satellite in the Cartesian coordinate system with respect to time. If the precise position of the satellite is known, then the second approach is more accurate than generating the reference orbit in the simulation. GFZ Potsdam provides the POD solution for TanDEM-X satellite in the Earth Centred Earth Fixed (ECEF) frame which is accurate to sub-decimetre level. These data are available in Consolidated Product Format (CPF). The position data are collected and converted into a series of user motion commands (UMT format) which are recognised by the SPIRENT simulator. The POD solution provides the position of the satellite every 120s. The SPIRENT simulator

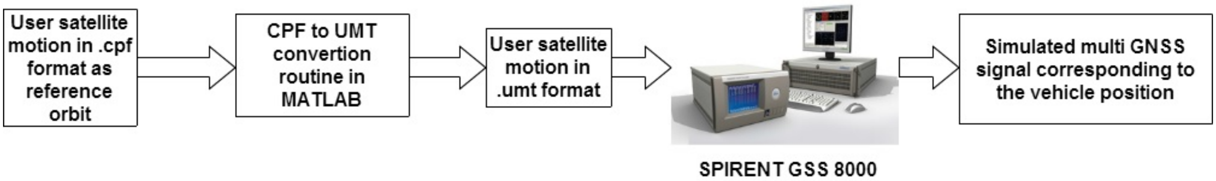


Fig. 2: Using POD data in SPIRENT

interpolates the position of the satellite during this time interval. In fig. 2 the sequential process of providing reference orbit to SPIRENT is shown. This method was utilized to provide the best possible reference orbit to the simulation scenario. This reference orbit was considered as the truth and used to calculate the estimation error in the simulations. The starting epoch was selected arbitrarily as 17th December 2013.

GNSS measurement simulation

For the GNSS signal and measurement simulations, all the 32 satellites of the GPS are considered and the full constellation of Galileo is assumed. The SPIRENT GNSS simulator provides both pseudo-range and carrier-range data directly, when a GNSS receiver is not connected to the simulator. The receiver clock bias is simulated in the scenario. The carrier-range does not contain the integer ambiguity but, the ionospheric delay is considered and simulated in the scenario, using the default SPIRENT ionospheric delay model. The measurements are generated at 1Hz rate.

Model for state propagation

In Extended Kalman Filter based estimation technique the estimable state vector is propagated to calculate *apriori* state. In the present simulation experiment a simple orbit model is used for propagation. Following the conventional approach, the state vector associated with the satellite motion is selected as:

$$\mathbf{X_{sat}} = \begin{bmatrix} \mathbf{r} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} x & y & z & v_x & v_y & v_z \end{bmatrix}^T \tag{1}$$

where $\mathbf{r} = [x \quad y \quad z]^T$ is the position vector of the satellite in the Earth Centred Inertial (ECI) frame and $\mathbf{v} = [v_x \quad v_y \quad v_z]^T$ is the velocity vector of the satellite in the ECI frame. The dynamical equation of the state vector includes the acceleration of the satellite. As the scenario is a LEO mission, the effect of the Sun and the Moon on the acceleration are neglected. Considering J_2, J_3 and J_4 zonal harmonics, the acceleration of the satellite due to the Earth's gravity field in ECI frame is expressed as [5] :

$$\ddot{\mathbf{r}} = \begin{bmatrix} -\frac{\mu_e x}{r^3}(1 + J_2C_{1x} + J_3C_{2x} + J_4C_{3x}) \\ -\frac{\mu_e y}{r^3}(1 + J_2C_{1y} + J_3C_{2y} + J_4C_{3y}) \\ -\frac{\mu_e z}{r^3}(1 + J_2C_{1z} + J_3C_{2z} + \frac{mu_e}{r^2}J_3C_{3z} + J_4C_{4z}) \end{bmatrix} \tag{2}$$

where

$$C_{1x} = \left(\frac{R_e}{r}\right)^2 \frac{3}{2}\left(1 - 5\frac{z^2}{r^2}\right)$$
$$C_{2x} = \left(\frac{R_e}{r}\right)^3 \frac{5}{2}\left(3 - 7\frac{z^2}{r^2}\right)\frac{z}{r}$$
$$C_{3x} = -\left(\frac{R_e}{r}\right)^4 \frac{5}{8}\left(3 - 42\frac{z^2}{r^2} + 63\frac{z^4}{r^4}\right)$$

$$C_{1y} = \left(\frac{R_e}{r}\right)^2 \frac{3}{2}\left(1 - 5\frac{z^2}{r^2}\right)$$
$$C_{2y} = \left(\frac{R_e}{r}\right)^3 \frac{5}{2}\left(3 - 7\frac{z^2}{r^2}\right)\frac{z}{r}$$
$$C_{3y} = -\left(\frac{R_e}{r}\right)^4 \frac{5}{8}\left(3 - 42\frac{z^2}{r^2} + 63\frac{z^4}{r^4}\right)$$
$$C_{1z} = \left(\frac{R_e}{r}\right)^2 \frac{3}{2}\left(3 - 5\frac{z^2}{r^2}\right)$$
$$C_{2z} = \left(\frac{R_e}{r}\right)^3 \frac{5}{2}\left(6 - 7\frac{z^2}{r^2}\right)\frac{z}{r}$$
$$C_{3z} = \left(\frac{R_e}{r}\right)^2 \frac{3}{2}$$
$$C_{4z} = -\left(\frac{R_e}{r}\right)^4 \frac{5}{8}\left(15 - 70\frac{z^2}{r^2} + 63\frac{z^4}{r^4}\right)$$

and $r = \sqrt{x^2 + y^2 + z^2}$, R_e and μ_e are the mean radius of the Earth and the gravitational parameter of the Earth respectively. The differential equation for the state vector can be represented as:

$$\dot{\mathbf{X}}_{\mathbf{sat}} = \begin{bmatrix} \dot{\mathbf{r}} \\ \dot{\mathbf{v}} \end{bmatrix} + \mathbf{W}_{\mathbf{sat}}(\mathbf{t}) = \begin{bmatrix} \mathbf{v} \\ \ddot{\mathbf{r}} \end{bmatrix} + \mathbf{W}_{\mathbf{sat}}(\mathbf{t}) \tag{3}$$

Here the $\mathbf{W}_{\mathbf{sat}}$ vector is the process noise vector which represents the unmodeled dynamics of the motion. Apart from the states associated with the satellite motion, the GNSS receiver clock bias and the bias rate are also considered as estimable states. The Namuru V3.3 receiver uses a single clock for both the GPS and Galileo constellations. For this reason, two states associated with the receiver clock bias are included. A receiver clock bias model can be represented as a first order Markov process. The receiver clock bias dynamics can be represented as [6] :

$$\dot{\mathbf{X}}_{\mathbf{clk}} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{1}{\nu} \end{bmatrix} \mathbf{X}_{\mathbf{clk}} + \begin{bmatrix} w_u \\ w_{ru} \end{bmatrix} \tag{4}$$

where, $\mathbf{X}_{\mathbf{clk}} = [\delta t_u \quad \delta t_{ru}]^T$, ν is the clock model correlation time, $\mathbf{W}_{\mathbf{clk}} = [w_u \quad w_{ru}]^T$ is random noise. δt_u is the receiver clock bias and δt_{ru} is the receiver clock bias rate. The complete state vector and the process noise for the estimation are:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{\mathbf{sat}} \\ \mathbf{X}_{\mathbf{clk}} \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} \mathbf{W}_{\mathbf{sat}} \\ \mathbf{W}_{\mathbf{clk}} \end{bmatrix}$$

Measurement model

Pseudo-range and carrier range measurements of the GNSS is modelled as [7, 8] :

$$\rho_i(t) = r_i(t) + c[\delta t_u(t) - \delta t_i(t-\tau)] + \mathbf{I}(t) + \epsilon_\rho(t) \tag{5}$$

$$\Phi_i(t) = r_i(t) + c[\delta t_u(t) - \delta t_i(t-\tau)] + \mathbf{I}(t) + \lambda N + \epsilon_\Phi(t) \tag{6}$$

where

i	is GNSS satellite index
ρ_i	is pseudo-range from the LEO satellite to the navigation satellite i
Φ_i	is carrier-range from the LEO satellite to the navigation satellite i
r_i	is geometric distance from LEO satellite to the navigation satellite i
δt	is receiver clock bias
δt_i	is clock bias of the navigation satellite
τ	is signal transmission time
c	is velocity of light
$\mathbf{I}(t)$	is ionospheric error
λ	is the wavelength of the carrier signal
N	is intiger ambiguity
$\epsilon_\rho(t)$	is random noise in pseudo-range measurement
$\epsilon_\Phi(t)$	is random noise in carrier-range measurement

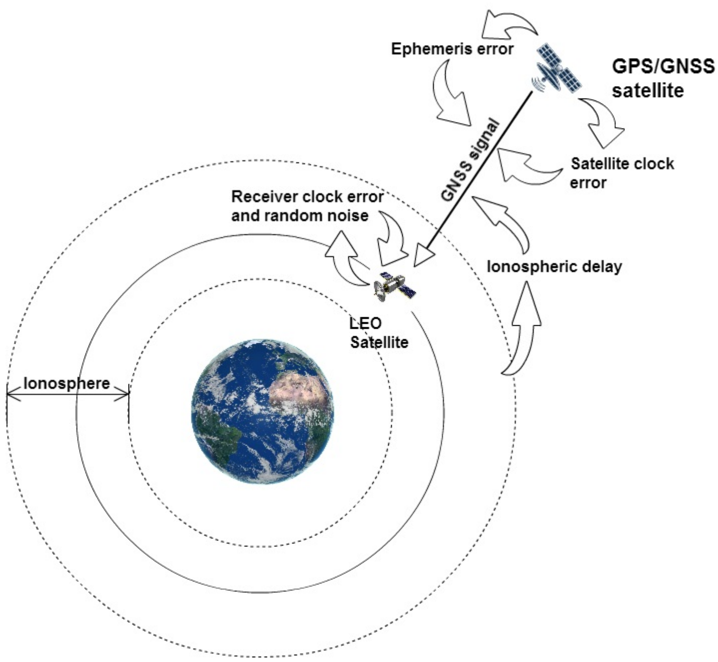


Fig. 3: Measurement errors

In fig. 3 the GNSS signal path and the associated errors are shown. The clock biases and ionospheric error introduces bias in the estimation of position. Navigation satellite clock biases are broadcast to the user receiver and the receiver clock bias is estimated. Using the GRAPHIC technique [9] the ionospheric error can be eliminated, when the integer ambiguity is resolved. This method is utilized in the simulation experiments without any receiver, because the simulated carrier phase measurements in these cases do not contain the integer ambiguity.

Estimation

A conventional Extended Kalman Filter (EKF) algorithm [10] is used to estimate the state vector. The filter tuning parameters are the process noise covariance matrix $\mathbf{Q}$, the measurement noise covariance matrix $\mathbf{R}$ and the initial value of the state covariance matrix $\mathbf{P}$. For the initialization of the Extended Kalman Filter an initial guess of the state vector is required. The least squares method is utilized to make an initial guess of the state vector. The $\mathbf{P}$ matrix is initialized using the uncertainties calculated from least square estimation. $\mathbf{Q}$ and $\mathbf{R}$ are selected as $E[\mathbf{W}\mathbf{W}^{\mathbf{T}}]$ and $E[\mathbf{V}\mathbf{V}^{\mathbf{T}}]$. $\mathbf{V}$ is the measurement noise vector which contains all the measurement noises associated with the measurements at a particular time. Both the process noise and the measurement noise are considered as auto-correlated and not cross-correlated [11]. This assumption results in diagonal $\mathbf{Q}$ and $\mathbf{R}$ matrices. For utilizing both the GPS and Galileo measurements a single Extended Kalman Filter is used. All the measurements from the GPS and Galileo at the same time are considered in a single measurement vector and subsequently processed through the EKF. In fig. 4 the estimation process is shown in a blok diagram.

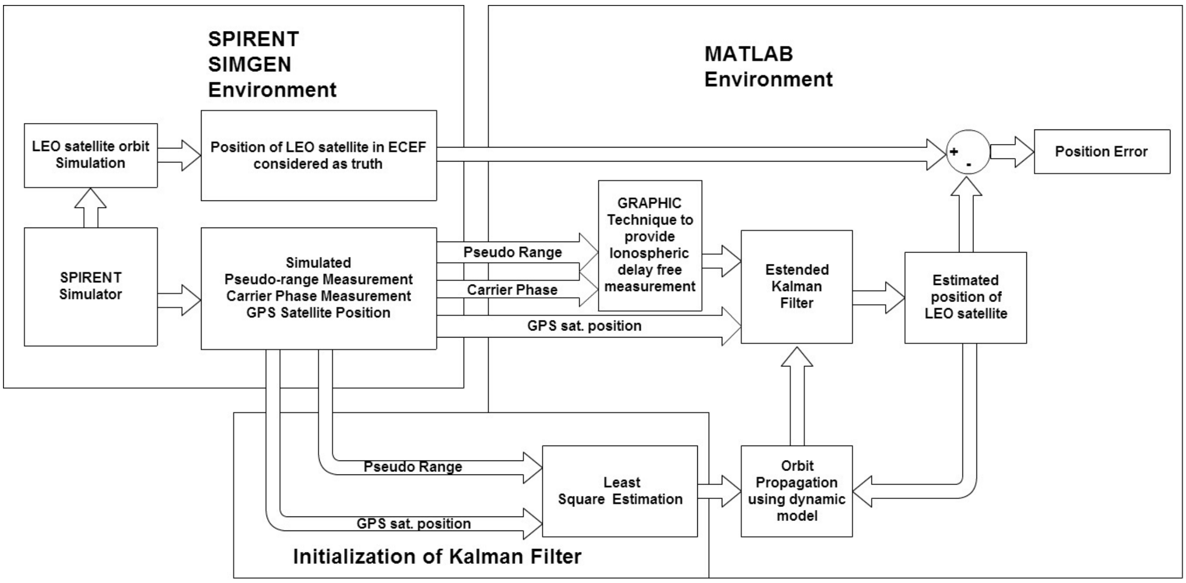


Fig. 4: Block diagram of the estimation process

Simulation result and discussion

Using the described method, a Lower Earth Orbit satellite mission is simulated. The POD data of TanDEM-X satellite was obtained from GFZ Potsdam and used as the reference orbit in simulation. Both the GPS and Galileo constellations were simulated and the measurements were used in an Extended Kalman Filter based algorithm to determine the satellite position. Figure 5 shows the position estimation error in each axis of the ECI frame. The root mean

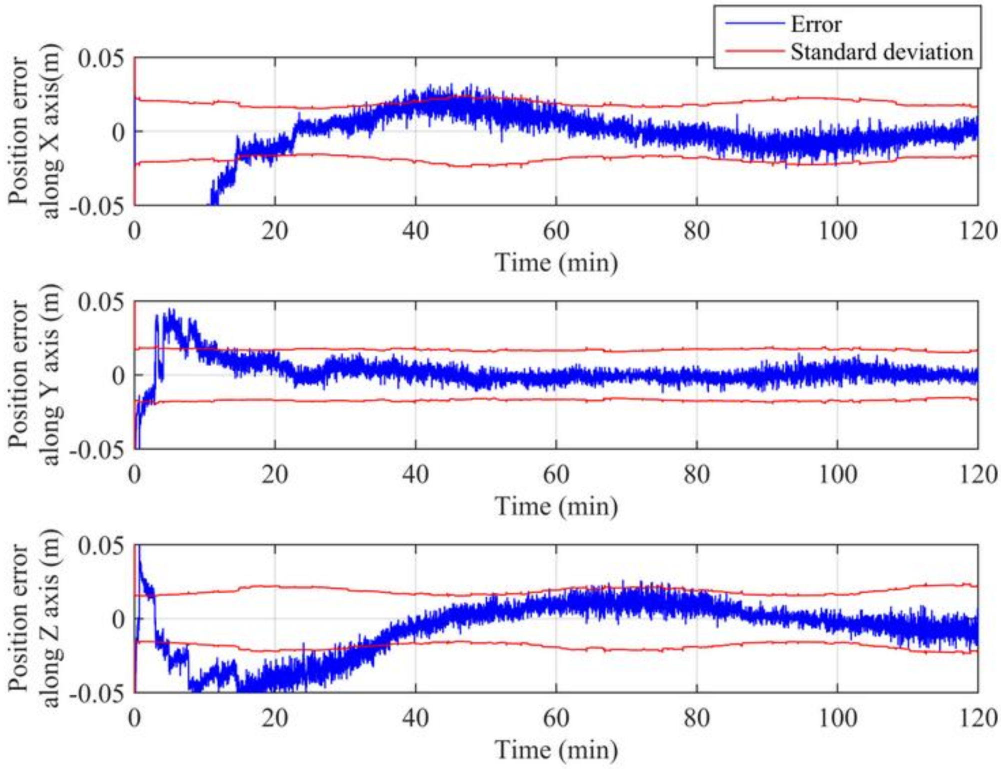


Fig. 5: Position estimation error using GPS and Galileo measurements

square error is 3.19 cm and the measurement rate is 1 Hz. It is observed that, the simulations using only SPIRENT measurements result in satellite position estimation error within the centimetre level. These simulations use the precise position of the TanDEM-X as a reference orbit to generate the GPS as well as Galileo signals. The receiver clock bias and the ionospheric delay were simulated and random noises were incorporated in the measurements artificially. In the EKF a simplified orbit and clock model were used. The receiver clock bias was constant and both the pseudo-range and the carrier-range measurements were used for the estimation. The constant clock bias resulted in very accurate receiver clock bias estimation using a simplified receiver clock model. Using both code and carrier phase measurements the ionospheric error was eliminated. Also the positions of the GNSS satellites were known accurately. These assumptions resulted in centimetre level accuracy in the LEO satellite position estimation using the EKF.

Conclusion

A method is developed to simulate a GNSS based LEO satellite orbit determination scenario using the best possible reference orbit as the input to the SPIRENT GSS 8000 simulator. This procedure is useful for evaluating the performance of a position estimation algorithm in a LEO mission scenario. The presented procedure makes the GNSS based orbit determination simulation highly modular because the EKF based estimation algorithm is easily replaceable with other estimation technique, which enables the performance comparison of different estimation algorithms and the performance of different receivers with different estimation algorithms can also be demonstrated by connecting the GNSS receivers with the SPIRENT simulator.

References

[1] Montenbruck, O., Garcia-Fernandez, M., Helm, A., and Markgraf, M., "GPS FOR MICROSATELLITES STATUS AND PERSPECTIVES," *In proceedings of 6th IAA Symposium on Small Satellites for Earth Observation, April 23-26, Berlin*, 2007.

[2] Capuano, V., Botteron, C., and Farine, P.-A., "Gnss Performances for Meo, Geo and Heo," in *64th International Astronautical Congress, Beijing, China*, 2013.

[3] Choudhury, M., Glennon, E., Dempster, A. G., and Mumford, P., "Characterization of the namuru v3. 2 spaceborne gps receiver," in *12th Australian Space Science Conference*, pp. 24–27, 2012.

[4] Choudhury, M., Cheong, J. W., Wu, J., Shivaramaiah, N. C., and Dempster, A. G., "Initial Test Results of Namuru Dual-GNSS Space-borne Receiver," in *IGNSS Symposium 2013*, (Outrigger Gold Coast, Qld Australia), pp. 1–8, 2013.

[5] Vallado, D. A., *Fundamentals of astrodynamics and applications*, vol. 12. Springer Science & Business Media, 2001.

[6] Qiao, L., Lim, S., Rizos, C., and Liu, J., "A Multiple GNSS-based Orbit Determination Algorithm for Geostationary Satellites," in *IGNSS Symposium 2009*, no. December, 2009.

[7] Kaplan, E. D. and Hegarty, C. J., "Understanding GPS: principles and applications," pp. 237–260, Artech house, 2005.

[8] Misra, P. and Enge, P., *Global Positioning System : Signals , Measurements and Performance*. Massachusetts: Ganga-Jamuna Press, 2006.

[9] Yunck, T. P., "Coping with the atmosphere and ionosphere in precise satellite and ground positioning," in *Environmental Effects on Spacecraft Positioning and Trajectories* (Jones, A. V., ed.), Geophysical Monograph Series, Wiley, 1993.

[10] Montenbruck, O. and Gill, E., *Satellite orbits: models, methods and applications*. Springer, 2012.

[11] Montenbruck, O. and Ramos-Bosch, P., "Precision real-time navigation of LEO satellites using global positioning system measurements," *GPS Solutions*, vol. 12, pp. 187–198, Oct. 2008.

Attitude Determination of the UNSW EC0 CubeSat

Li Qiao[1], Andrew G Dempster [2], Barnaby Osborne [2]

[1]*School of Engineering and Information Technology, University of New South Wales, Canberra, 2610, ACT, Australia*
[2]*Australian Centre for Space Engineering Research, University of New South Wales, 2052, NSW, Australia*

Summary: The UNSW QB50 project satellite "UNSW EC0", is being designed and constructed at the Australian Centre for Space Engineering Research (ACSER). This paper describes the design and simulation of the attitude estimation method. The attitude estimation uses an Extended Kalman Filter to blend the multiple vectored measurements from the Sun, the Earth horizon, the Earth's magnetic field and angular velocity data from Gyroscopes. The sensor configuration consists of a three-axis magnetometer, three-axis MEMS gyroscope and Cubesense, which is a combination of a Sun Sensor and an Earth sensor. The results from the simulation show that the developed Extended Kalman Filter (EKF) can achieve the expected performance with demonstrated utility in fusing the various sensors at various update rates.

Keywords: Attitude estimation, Extended Kalman filtering, CubeSats, Magnetometer, Gyroscope, CubeSense

Introduction

The University of New South Wales (UNSW) established itself as a leading centre for satellite system engineering research and education in Australia following receipt of a major grant under the Australian Space Research Program (ASRP). One of the major projects under ACSER is the Warrawal project, a Comprehensive Tertiary Education Program in Satellite Systems Engineering. As part of the space engineering related research and education program, UNSW proposed and was awarded the design and construction of one of the QB50 constellation satellites. UNSW's QB50 2U CubeSat is called UNSW QB50 Educational CubeSat 0 (UNSW EC0).

UNSW EC0 will be part of a larger constellation of satellites launched through the QB50 project run by the Von Karman Institute (VKI). The primary goal of the QB50 mission is to carry out a series of atmospheric sounding experiments in the thermosphere. To this end, all participating CubeSats will carry a VKI payload in addition to the host institute payloads. The network of 50 CubeSats will be launched into a 98 degree inclined circular orbit at an altitude of 350-380 km in a string-of-peals configuration [1]. The constellation of CubeSats are made up of a mix of 2U and 3U CubeSats, with the majority carrying a set of standardised sensors for multi-point, in-situ, long-duration measurements of key parameters and constituents in the largely unexplored lower thermosphere and ionosphere.

Satellite Configuration

The UNSW EC0 is a 2U CubeSat, with physical dimensions of approximately 101mm ×101mm× 227mm and a maximum allowed weight of 2 kg.

The ADCS has two components consisting of the attitude control (ACS) system and the attitude determination system (ADS). The work regarding the ACS component design and simulation was presented in Reference [2]. This paper will show the design of the attitude determination system through on-board measurement. ADCS is one of housekeeping subsystems for a satellite bus. The task for ADS component is to determine the satellite's orientation relatively to the Earth, Sun or other object. Once the orientation is determined, the ACS could control the attitude actuators to point the satellite as desired.

A satellite is often composed of many subsystems, including experiment payloads that impose requirements on attitude determination and control. The design requirements imposed by the payloads necessitate precise 3-axis attitude determination of the spacecraft when it is in orbit. In our case, it is important to know the orientation of EC0 relatively to the Earth so it can receive the GPS signals. For large satellites, subsystems are redundant for reliability purposes. However, for size and weight constraint reasons, there is no redundant design for ADCS in most CubeSats. Miniaturization is a key approach in order to meet the mass and volume budget. The body size of the CubeSat is $10.1\mathrm{cm}\times 10.1\mathrm{cm}\times 22.7\mathrm{cm}$ and the mass is around $1.92\mathrm{kg}$.

Basics in ADS

This section describes the brief definitions of the different reference frames used throughout the paper. An assumption in this paper is that the satellite can be treated as a rigid body with a constant inertia matrix and a non-moving centre-of-mass.

Earth Centred Inertial (ECI) Frame: An inertial frame used for navigation. The origin of the frame is located at the centre of the earth with the z-axis pointing towards the North Pole. The x-axis points towards vernal equinox, which is the point where the plane of the Earth's orbit about the Sun crosses the Equator going from south to north, and the y-axis completes the right hand Cartesian coordinate system.

Orbit Reference Frame: Orbit frame is a kind of Vehicle Velocity Local Horizontal (VVLH) coordinate frame. The origin of the orbit frame lies with the satellite centre of mass and it has the x-axis pointing along the positive velocity vector direction tangentially to the orbit, the z-axis pointing toward the centre of the Earth, and the y-axis completes the right-handed Cartesian coordinate system. The satellite attitude is described by roll, pitch and yaw, which is the rotation around the x-, y-, z- axis respectively.

Body Reference Frame: The body frame is fixed to the geometry of the satellite and moves and rotates with the satellite. Its origin lies with the centre of satellite mass, the x-axis points in the forward direction, the z-axis points down side, and the y-axis completes the right-handed orthogonal system.

Euler Angles: Roll, pitch and yaw. Euler angles shown in Fig. 1 are used to represent attitude. There are twelve possible sets of Euler angles. The first, second and third rotation is about the $\mathrm{XYZ}-$ axis of the body frame, respectively. Roll is the angle about the $Xb-$ axis. Pitch is the angle about $\mathrm{Yb}-$ axis. Yaw is the angle about the $\mathrm{Zb}-$ axis. Euler angle errors are used to evaluate the attitude determination performance.

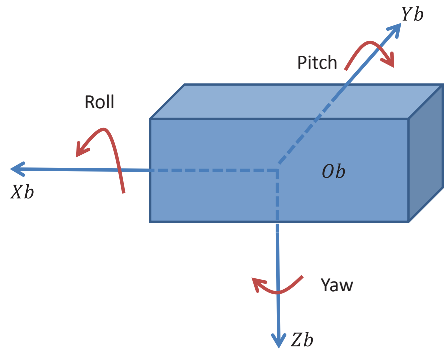


Fig. 1. Euler angles: Roll, pitch and yaw

Attitude quaternion: As an alternative to the Euler angle representation, a quaternion is a four element-vector commonly used to represent attitude. Note that both quaternions and Euler angles are representatives of attitude rotations and conversions between the two can be performed. The attitude quaternion is used to present the satellite kinematics in Kalman filter, which will be discussed later in this paper.

ADS Components

ADS cannot be dependent on only one sensor to provide acceptable accuracy; therefore it needs several different types of sensors and applies a Kalman Filter for the data fusion.

In this mission, the attitude sensors include fine sun sensing and earth horizon sensing via the CubeSense board, rotation rates from a Tri-axial Gyroscope, Tri-axial Earth magnetic field vector sensing via an externally mounted 3-axis magnetometer, as well as coarse sun sensing via the solar panels. The considerations during selection are:

1) The key selection criteria would be power, size and cost.
2) These selected sensors are not high-end sensors as the CubeSat uses only magnetorquers to provide a low-cost and low accuracy control, with the pointing requirement being limited to within$\pm 2°$.
3) Having four different types of sensors is to ensure continuous solutions and allows for sensor failure.
4) The coarse sun sensors integrated into solar cells are used only to indicate the Sun as they are too inaccurate.

In addition, a GPS receiver will also be incorporated, as part of the orbit determination system (ODS). The GPS receiver embedded software contains a Kalman Filter and can continuously provide the position, velocity and time information for the ADS.

All the sensors, except the gyroscopes, are vector-based and require the corresponding reference vector to determine the attitude. The sensor types selected are mostly inertially referenced, except the magnetometer where the referenced magnetic field is in the ECEF(Earth Centred Earth Fixed) frame. Table 1 list the measurements and the required reference values of each attitude sensor. Where $\sim$ indicated the measured values, and the superscripts indicate the reference frame with respect to.

Table 1 Attitude sensor

Number	Device	Measurements	Reference
1	Sun sensor	$\tilde{\mathbf{S}}^{B}$	$\boldsymbol{S}^{I}_{ref}$
2	Earth sensor/Nadir sensor	$\tilde{\mathbf{E}}^{\mathbf{B}}$	$\boldsymbol{E}^{I}_{ref}$

3	Magnetometer	$\tilde{\boldsymbol{B}}^B$	$\boldsymbol{B}^E$
4	MEMS Angular rate Gyroscope	$\tilde{\boldsymbol{\omega}}_{\mathrm{IB}}^{\mathrm{B}}$	

1. CubeSense

An integrated sun and nadir sensor, a CubeSense module from Innovus [3], will be used as primary attitude sensor. Two 640×480 pixel CMOS cameras are dedicated to sun sensing and Earth horizon detection, respectively. Both cameras have wide field-of-view optics (180°) based on a fisheye lens for increased operating range. The primary outputs of the sensor are the measured sun vector and nadir vector in the sensor's coordinate frame. The measured vectors are output as azimuth/elevation angles relative to the camera bore-sight. Mapping the azimuth and elevation from spherical to three-dimensional Cartesian coordinates, the measured vector can be expressed with the two angles. For instance, the measured sun vector in body reference frame $\tilde{\mathrm{S}}^{\mathrm{B}}$can be found as

$$\tilde{\boldsymbol{S}}^{\boldsymbol{B}} = \begin{bmatrix} \cos(Elevation) \cdot \cos(Azimuth) \\ \cos(Elevation) \cdot \sin(Azimuth) \\ \sin(Elevation) \end{bmatrix} \tag{1}$$

Similarly, the measured Earth vector $\tilde{\mathrm{E}}^{\mathrm{B}}$ can be calculated with the elevation and azimuth.

The Sun and Earth reference vector $\mathrm{S}_{\mathrm{ref}}^{\mathrm{I}}$ and $\mathrm{E}_{\mathrm{ref}}^{\mathrm{I}}$ in Table 1 are computed in the on-board computer as the CubeSense hardware does not contain reference vector generator. For the Sun reference vector, the on-board computer will install the astronomical Almanac model, which contains ephemeris data for the Sun and also provides formulas for computing the Sun's position. The following formulas outline the method provides by the Astronomical Almanac[4].

$$\begin{cases} \mathrm{n} = \mathrm{JD} - 2454545.0 \\ \mathrm{L} = 280.460° + 0.9856474°\mathrm{n} \\ \mathrm{g} = 357.528° + 0.9856003°\mathrm{n} \\ \lambda = \mathrm{L} + 1.915° \sin(\mathrm{g}) + 0.020° \sin(2\mathrm{g}) \\ \beta = 0° \\ \epsilon = 23.439° - 0.0000004°\mathrm{n} \\ \boldsymbol{S}_{\mathrm{ref}}^{\mathrm{I}} = \begin{bmatrix} \mathrm{X}_{\mathrm{S}}^{\mathrm{I}} \\ \mathrm{Y}_{\mathrm{S}}^{\mathrm{I}} \\ \mathrm{Z}_{\mathrm{S}}^{\mathrm{I}} \end{bmatrix} = \begin{bmatrix} \cos(\lambda) \\ \cos(\epsilon)\sin(\lambda) \\ \sin(\epsilon)\sin(\lambda) \end{bmatrix} \end{cases} \tag{2}$$

where n is the number of days since J2000; JD is the Julian Date; L is the mean longitude of the Sun L, g is the mean anomaly; L and g are used to find the ecliptic longitude λ; the ecliptic latitude is zero; ϵ is the obliquity of the ecliptic. Finally, the direction in the Sun $\mathrm{S}_{\mathrm{ref}}^{\mathrm{I}}$ can be calculated in ECI. JD can be transformed from GPS time or UTC time.

For the Earth reference vector $\mathrm{E}_{\mathrm{ref}}^{\mathrm{I}}$ is calculated from the positioning subsystem using Equations(3).

$$\begin{cases} \boldsymbol{r}^I = \boldsymbol{R}_E^I \boldsymbol{r}^E \\ \boldsymbol{R}_E^I = \boldsymbol{R}_I^{E^T} = \begin{bmatrix} \cos(GAST) & -\sin(GAST) & 0 \\ \sin(GAST) & \cos(GAST) & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ GAST = \omega_{ie} t \\ \boldsymbol{E}_{ref}^I = -\dfrac{\boldsymbol{r}^I}{\|\boldsymbol{r}^I\|} \end{cases} \tag{3}$$

Where, r^{E} is the position of the satellite in the ECEF frame the GPS receiver gives; the $\mathrm{r}^{\mathrm{E}}$is transformed to r^{I} with the Greenwich Apparent Sideral Time(GAST); ω_{ie} is the Earth rotation; and t is the time since the ECEF and ECI frames were aligned. GAST can be transformed from JD.

2. Tri-axial MEMS Gyroscope/IRU

The Tri-axial gyroscope suite, also called the Inertial Reference Unit (IRU), measures the angular rate of the satellite in three axes $\tilde{\omega}_{\mathrm{IB}}^{\mathrm{B}}$. The gyroscope is modelled as

$$\begin{cases} \tilde{\boldsymbol{\omega}}_{rg} = \boldsymbol{\omega}_{IB}^B + \boldsymbol{b} + \boldsymbol{\eta}_v \\ \dot{\boldsymbol{b}} = \boldsymbol{\eta}_u \end{cases} \tag{4}$$

Where $\boldsymbol{\omega}_{IB}^{B}$ is the realistic measurement of the gyroscope, b is the bias, η_{v} is the ARW (Angular Random Walk) unit$°/\sqrt{\mathrm{s}}$, and η_{u} is the RRW (Rate Random Walk) unit$°/\sqrt{\mathrm{s}}$. This is a simplified model as the gyro scale factor errors and misalignment errors are neglected. The noise has to be as small as possible because it is mitigated with filtering. However the noise of MEMS gyroscopes is relatively large, therefore the gyroscope selection should be well considered. The final actual hardware selection for the IRU hasn't decided yet.

3. Tri-axial magnetometer

A magnetometer senses magnetic field strength and, when used in a three-axis triad, magnetic field direction. The model of the magnetometer is

$$\widetilde{\boldsymbol{B}}^B = \boldsymbol{B}^B + \boldsymbol{\varepsilon} \tag{5}$$

Where $\boldsymbol{B}^{\boldsymbol{B}}$ is the realistic measurement; ε is the noise.

As a spacecraft navigational aid, sensed field strength and direction is compared to a map of the Earth's magnetic field $\boldsymbol{B}^{\boldsymbol{E}}$ stored in the memory of an on-board or ground-based guidance computer. If satellite position is known then attitude can be inferred. Due to the very limited storage of the on-board computer, storing large amount data for expected magnetic field values is not feasible. Instead, using the satellite position provided by the orbit model, the reference magnetic field value is estimated by using the international Geomagnetic Reference Field (IGRF) 2011. This model is an approximation near and above the Earth's surface, working with an orbit estimator [5]. In order to reduce the computing load, this model can be implemented with lower order such as 8. $\boldsymbol{B}^{\boldsymbol{E}}$ can be presented in Inertial frame within Equation (6)

$$\boldsymbol{B}^{\boldsymbol{I}} = \boldsymbol{R}_E^I \boldsymbol{B}^{\boldsymbol{E}} \tag{6}$$

Algorithm Design

Specifically the Discrete EKF will form the bases of the EKF developed as all the sensors are modelled as discrete. Fig. 2 presents the design of the ADS to process the multiple and various data rate data in an efficient way.

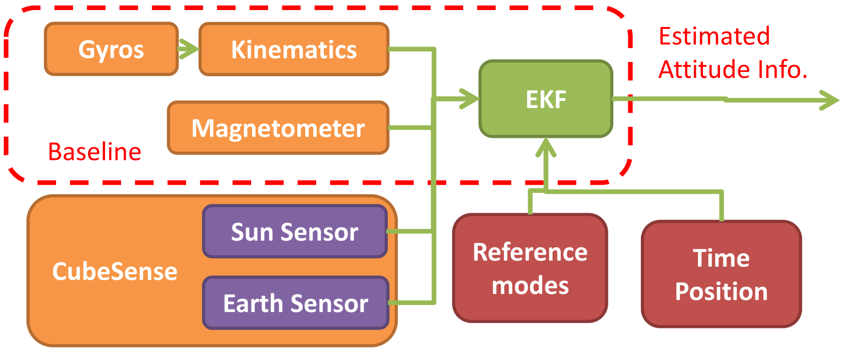


Fig. 2: The block diagram of ADS

All the above mentioned sensors except gyroscopes will general vector measures. Generally there are two ways to handle the vector measurements: 1) use the TRIAD or QUEST algorithm to calculate the original attitude information in quaternion format from two vectors, and then use the quaternion to blend in the EKF[6]; 2) put the available vector measurements to EKF directly. The first method requires at least two vector measurements are available at the same time. To avoid this constraint and for the efficiency, the second method is chosen in this paper. Different types of sensor have different update rates and the availability of some measures is not guaranteed. This may result the different combinations, and then cause the formulation of the observation matrix H_k and the calculation of the residual ε for each of the sensor types. The solution is to use the superposition of the measures from Reference [7]. The detailed EKF design is presented below.

1. Attitude kinematical model

The state vectors of the EKF are composed of four quaternion elements and three gyro bias elements, expressed in

$$\boldsymbol{X} = [\epsilon_1 \quad \epsilon_2 \quad \epsilon_3 \quad \eta \quad \beta_1 \quad \beta_2 \quad \beta_3] \tag{7}$$

The first four elements are the unit quaternion, written as

$$\boldsymbol{q} = [\boldsymbol{\epsilon} \quad \eta]^T = [\varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3 \quad \eta]^T \tag{8}$$

Where $\boldsymbol{q}$ is normalized $\boldsymbol{q}^T\boldsymbol{q} = \eta^2 + \boldsymbol{\epsilon}^T\boldsymbol{\epsilon} = \mathbf{1}$. η and $\varepsilon = [\varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3]^T$ are the real part and the three imaginary parts, respectively. According to the gyroscope modelling presented in Equation (4), the bias vector $\boldsymbol{b} = [\beta_1 \quad \beta_2 \quad \beta_3]^T$ is estimated as part of the states. To implement the EKF, the differential value $\boldsymbol{\delta q}$ can be driven from calculating multiplicative error quaternion.

$$\delta\boldsymbol{q} = \boldsymbol{q}\otimes\hat{\boldsymbol{q}}^{-1} \tag{9}$$

$$\boldsymbol{\delta q} = [\delta\epsilon \quad \delta\eta]^T \tag{10}$$

Then a well-known quaternion kinematics can be expressed as

$$\dot{\mathbf{q}} = \frac{\mathbf{1}}{\mathbf{2}}\,\boldsymbol{\Omega}(\omega)\mathbf{q} \tag{11}$$

Where

$$\boldsymbol{\Omega}(\boldsymbol{\omega}) = \begin{bmatrix} 0 & +\omega_z & -\omega_y & +\omega_x \\ -\omega_z & 0 & +\omega_x & +\omega_y \\ +\omega_y & -\omega_x & 0 & +\omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix} = \begin{bmatrix} -[\boldsymbol{\omega}\times] & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^{\boldsymbol{T}} & 0 \end{bmatrix} \tag{12}$$

$[\boldsymbol{\omega}\times]$ is the skew-symmetric matrix, indicting the cross product of $\boldsymbol{\omega}$. Equation (11) needs to be transformed in discrete format in the EKF as all the sensors are modelled as discrete. The quaternion propagation can be written as

$$\hat{\mathbf{q}}_{\mathrm{k+1}}^{-} = \bar{\boldsymbol{\Omega}}(\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+})\hat{\mathbf{q}}_{\mathrm{k}}^{+} \tag{13}$$

The $\bar{\boldsymbol{\Omega}}(\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+})$ is expressed in

$$\bar{\boldsymbol{\Omega}}(\hat{\boldsymbol{\omega}}_{\mathbf{k}}^{+}) = \begin{bmatrix} I_{3\times3}\cos\left(\frac{1}{2}\|\hat{\omega}_{\mathrm{k}}^{+}\|\Delta t\right) - [\hat{\Psi}_k^{+}\times] & \hat{\Psi}_k^{+} \\ -[\hat{\Psi}_k^{+}]^T & \cos\left(\frac{1}{2}\|\hat{\omega}_{\mathrm{k}}^{+}\|\Delta t\right) \end{bmatrix} \tag{14}$$

$$\hat{\Psi}_k^{+} = \frac{\sin\left(\frac{1}{2}\|\hat{\omega}_{\mathrm{k}}^{+}\|\Delta t\right)\hat{\omega}_{\mathrm{k}}^{+}}{\|\hat{\omega}_{\mathrm{k}}^{+}\|} \tag{15}$$

$\bar{\Omega}(\hat{\omega}_{\mathrm{k}}^{+})$ is a 4×4 matrix, and $\hat{\Psi}_k^{+}$ is a 3×1 matrix. $\Delta \mathrm{t}$ is the discrete sample time. $\hat{\omega}_{\mathrm{k}}^{+}$ is the post-update estimate of ω, and $\hat{\mathrm{q}}_{\mathrm{k}}^{+}$ is the post-update estimate of $\boldsymbol{q}$**.** Equation (11) or its discrete form Equation (13) relates the gyroscopes angular velocity outputs and the current quaternion to produce the derivative of the quaternion as the quaternion propagation. I.e. the gyroscopes provide attitude information in a near continuous fashion.

The propagation for the last three state elements, which are three gyroscopes bias values in Equation (7), is defined as

$$\begin{cases} \hat{\omega}_{\mathrm{k}}^{+} = \tilde{\omega}_k - \hat{\beta}_k^{+} \\ \hat{\beta}_k^{-} = \hat{\beta}_k^{+} \end{cases} \tag{16}$$

In the EKF, the predicted covariance estimate matrix $\mathrm{P}_{\mathrm{k+1}}^{-}$ is written as

$$P_{k+1}^{-} = \Phi_k P_k^{+} \Phi_k^T + G_k Q_k G_k^T$$

Where

$$\boldsymbol{\Phi}_k = \begin{bmatrix} \boldsymbol{\Phi}_{11} & \boldsymbol{\Phi}_{12} \\ \mathbf{0}_{3\times3} & \boldsymbol{I}_{3\times3} \end{bmatrix}$$

$$\boldsymbol{\Phi}_{11} = \boldsymbol{I}_{3\times3} - [\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\times]\frac{\sin(\|\hat{\omega}_{\mathrm{k}}^{+}\|\Delta t)}{\|\hat{\omega}_{\mathrm{k}}^{+}\|} + [\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\times]^2\frac{1 - cos(\|\hat{\omega}_{\mathrm{k}}^{+}\|\Delta t)}{\|\hat{\omega}_{\mathrm{k}}^{+}\|^2}$$

$$\boldsymbol{\Phi}_{12} = [\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\times]\frac{1 - \cos(\|\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\|\Delta t)}{\|\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\|^2} - \boldsymbol{I}_{3\times3}\Delta t - [\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\times]^2\frac{\|\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\|\Delta t - \sin(\|\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\|\Delta t)}{\|\hat{\boldsymbol{\omega}}_{\mathrm{k}}^{+}\|^3}$$

$$G_k = \begin{bmatrix} -\boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \boldsymbol{I}_{3\times3} \end{bmatrix}$$

$$Q_k = \begin{bmatrix} (\sigma_v^2\Delta t + \frac{1}{3}\sigma_u^2\Delta t^3)\boldsymbol{I}_{3\times3} & -(\frac{1}{2}\sigma_u^2\Delta t^2)\boldsymbol{I}_{3\times3} \\ -(\frac{1}{2}\sigma_u^2\Delta t^2)\boldsymbol{I}_{3\times3} & (\sigma_u^2\Delta t)\boldsymbol{I}_{3\times3} \end{bmatrix}$$

Where σ_{v}^2 and σ_{u}^2 are the variances of the gyroscope noises. Note that $\mathrm{P}_{\mathrm{k+1}}^{-}$ is a 6×6 matrix, as the estimated error state $\Delta\hat{\tilde{\mathrm{X}}}_{\mathrm{k}}^{+}$ is a $[6\times1]$ vector rather than a $[7\times1]$ state vector defined in

Equation (7). In the update phase, the states to be updated is $\Delta\widehat{\mathrm{X}}_{\mathrm{k}}^{+} = [\Delta\widehat{\alpha}_{\mathrm{k}}^{+} \quad \Delta\widehat{\beta}_{\mathrm{k}}^{+}]^{\mathrm{T}}$. Using the small angle approximation $\delta\epsilon \approx \delta\alpha/2$ and $\eta \approx 1$, the four state quaternion δq defined in Equation (10) has been replaced by the three state Euler error angle vector.

2. Measurement update

Due to ARW and RRW, the gyroscope propagation results will diverge without the measurement update. The observation matrix for the Sun sensor, Earth sensor and magnetic field sensor are $\boldsymbol{H^S}$, $\boldsymbol{H^E}$ and $\boldsymbol{H^M}$, respectively. They are expressed in Equation (17) to (19)

$$\boldsymbol{H}_k^S(\hat{X}_k^-) = \left[\left[A(\widehat{\boldsymbol{q}}_k^-)\boldsymbol{S}_{\boldsymbol{ref}}^{\boldsymbol{I}}\times\right] \quad \boldsymbol{0}_{3\times3}\right] \tag{17}$$

$$\boldsymbol{H}_k^E(\hat{X}_k^-) = \left[\left[A(\widehat{\boldsymbol{q}}_k^-)\boldsymbol{E}_{\boldsymbol{ref}}^{\boldsymbol{I}}\times\right] \quad \boldsymbol{0}_{3\times3}\right] \tag{18}$$

$$\boldsymbol{H}_{\boldsymbol{k}}^{\boldsymbol{M}}(\widehat{X}_k^-) = \left[\left[A(\widehat{\boldsymbol{q}}_{\boldsymbol{k}}^-)\boldsymbol{B}_{\boldsymbol{ref}}^{\boldsymbol{I}}\times\right] \quad \boldsymbol{0}_{3\times3}\right] \tag{19}$$

The standard EKF form for the error state vector update is

$$\Delta\widehat{\widehat{\boldsymbol{X}}}_{\boldsymbol{k}}^{+} = \boldsymbol{K}_{\boldsymbol{k}}\boldsymbol{res}_{\boldsymbol{k}} \tag{20}$$

$\boldsymbol{res_k}$ is the measurement residual. The residual of the Sun, Earth and Magnetometer are expressed in Equation (21).

$$\begin{cases} \boldsymbol{res}^S = \widetilde{\boldsymbol{S}}^{\boldsymbol{B}} - \boldsymbol{A}(\widehat{\boldsymbol{q}}_{\boldsymbol{k}}^-)\boldsymbol{S}_{\boldsymbol{ref}}^{\boldsymbol{I}} \\ \boldsymbol{res}^E = \widetilde{\boldsymbol{E}}^{\boldsymbol{B}} - \boldsymbol{A}(\widehat{\boldsymbol{q}}_{\boldsymbol{k}}^-)\boldsymbol{E}_{\boldsymbol{ref}}^{\boldsymbol{I}} \\ \boldsymbol{res}^M = \widetilde{\boldsymbol{B}}^B - \boldsymbol{A}(\widehat{\boldsymbol{q}}_{\boldsymbol{k}}^-)\boldsymbol{B}_{\boldsymbol{ref}}^{\boldsymbol{I}} \end{cases} \tag{21}$$

where

$$\boldsymbol{A}(\widehat{\boldsymbol{q}}_{\boldsymbol{k}}^-) = (\eta^2 - \|\epsilon\|^2)\boldsymbol{I}_{\boldsymbol{3\times3}} + 2\widehat{\boldsymbol{q}}_{\boldsymbol{k}}^-\widehat{\boldsymbol{q}}_{\boldsymbol{k}}^{-\boldsymbol{T}} - 2\eta[\epsilon\times] \tag{22}$$

The quaternion update and the bias update equations are expressed in Equations (23) to (25).

$$\widehat{\boldsymbol{q}}_k^+ = \widehat{\boldsymbol{q}}_k^- + \frac{1}{2}\Xi(\widehat{\boldsymbol{q}}_k^-)\delta\widehat{\boldsymbol{\alpha}}_k^+ \tag{23}$$

$$\widehat{\boldsymbol{\beta}}_k^+ = \widehat{\boldsymbol{\beta}}_k^- + \Delta\widehat{\boldsymbol{\beta}}_k^+ \tag{24}$$

$$\Xi(\boldsymbol{q}) \triangleq \begin{bmatrix} \eta\boldsymbol{I}_{3\times3} + [\boldsymbol{\epsilon}\times] \\ -\boldsymbol{\epsilon}^T \end{bmatrix} = \begin{bmatrix} \eta & -\epsilon_3 & \epsilon_2 \\ \epsilon_3 & \eta & -\epsilon_1 \\ -\epsilon_2 & \epsilon_1 & \eta \\ -\epsilon_3 & -\epsilon_2 & -\epsilon_1 \end{bmatrix} \tag{25}$$

The error covariance matrix update and the gain calculation equations are

$$\boldsymbol{P}_{\boldsymbol{k}}^+ = \left[\boldsymbol{I}_{\boldsymbol{6\times6}} - \boldsymbol{K}_k\boldsymbol{H}_k(\widehat{\boldsymbol{X}}_{\boldsymbol{k}}^-)\right]\boldsymbol{P}_{\boldsymbol{k}}^- \tag{26}$$

Where the gain matrix $\boldsymbol{K_k}$ is written as

$$\boldsymbol{K}_k = \boldsymbol{P}_k^- \boldsymbol{H}_k^T(\widehat{\boldsymbol{X}}_k^-)[\boldsymbol{H}_k(\widehat{\boldsymbol{X}}_k^-)\boldsymbol{P}_k^- \boldsymbol{H}_k^T(\widehat{\boldsymbol{X}}_k^-) + \boldsymbol{R}_k]^{-1} \tag{27}$$

The detailed steps of the discrete Extended Kalman Filter are implemented accruing to Reference [8].

Simulation Environment

In this section, the attitude determination algorithm has been implemented in Malab and STK (System Tool Kit). The Kalman filter algorithm is coded in Matlab and the positions of CubeSat, Sun and Earth and timing in various frames are generated from STK.

The simulation parameters are summarized in Table. 2. Simulation is carried out for a duration of three orbits. EKF estimations are done at 1Hz. This interval ensures all the sensor data are reliably available all the time. The sampling frequency of Gyro output is set to 10 Hz, thus estimations are done every 10 times Gyro integration. The orbit used in the simulation is a Circular Polar Low Earth Orbit (LEO) with an inclination of approximately 98° and an orbit height of approximately 350 km. This orbit is considered typical for QB50.

Table. 2 Simulation Parameters

Orbit altitude	350 km	EKF filtering rate	1Hz
Gyroscope sampling rate	10 Hz	Gyro initial bias	$\pm 3°$
Gyro ARW	$0.029°/s$	Gyro RRW	$0.0002°/s$
Orbit period	5492.4s	Simulation carried out duration	3 orbits 16477s

The initial state is set as $\widehat{\mathrm{X}}_0^- = [\mathrm{q}_0 \quad 0_{3\times1}]^{\mathrm{T}} = [0_{3\times1} \quad 1 \quad 0_{3\times1}]^{\mathrm{T}}$, and the initial error covariance is set as $\boldsymbol{P}_0 = \sqrt[4]{\Delta t((\sigma^S)^2 + (\sigma^E)^2 + (\sigma^M)^2)(\sigma_v^2 + 2\sigma_u\sigma_v)}$. Where σ_v^2 and σ_v^2 are the variance associated with η_v and η_u respectively in Gyroscope modelling Equation(4). The measurement noise R matrix in Equation (27) is calculated as.

$$\boldsymbol{R}_k = \sigma^2 \boldsymbol{I}_{3*3}$$

The value of noise variance of Sun, Earth and magnetometer used are chosen based on datasheets since no real measurements have been obtained from the hardware. The values are listed below. Noted the accuracy of sun varies with the incident angle range.

$$\sigma^S = \begin{cases} 0.3° & (\pm 40°) \\ 0.5° & (\pm 60°) \\ 1° & (\pm 90°) \end{cases}$$
$$\sigma^E = 0.5°$$
$$\sigma^M = 1.25\times10^{-7}\ Tesla$$

Back to the inherent problem: The different types of sensors produce solutions at different rates. For instance, the sun sensor is not available in eclipse as it uses the Sun as a reference vector even when the Sun sensor is switched on. To solve this problem, a signal matrix is designed to control the formulation of the observation matrix $\boldsymbol{H}_k$ and the calculation of the residual $\boldsymbol{\varepsilon}$ for each of the sensor types. Use 1 or 0 to indicate the sensor is switched on or off, the number of combination of the three sensors is eight theoretically. *Fig. 4* lists the observation matrix and measurement residual formulations.

Table 3 The formulation of the observation matrix and residual

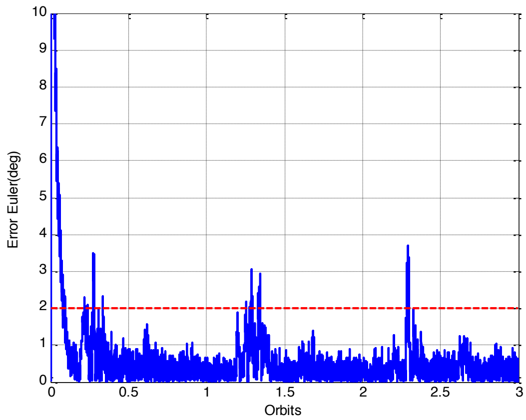
Case No.	$[M \quad S \quad E]$	Observation matrix	Measurement residual
1.	$[1 \quad 1 \quad 1]$	$diag[\boldsymbol{H}^M \quad \boldsymbol{H}^S \quad \boldsymbol{H}^E]^T$	$[\boldsymbol{res}^M \quad \boldsymbol{res}^S \quad \boldsymbol{res}^E]^T$
2.	$[1 \quad 1 \quad 0]$	$diag[\boldsymbol{H}^M \quad \boldsymbol{H}^S]^T$	$[\boldsymbol{res}^M \quad \boldsymbol{res}^S]^T$
3.	$[1 \quad 0 \quad 1]$	$diag[\boldsymbol{H}^M \quad \boldsymbol{H}^E]^T$	$[\boldsymbol{res}^M \quad \boldsymbol{res}^E]^T$
4.	$[0 \quad 1 \quad 1]$	$diag[\boldsymbol{H}^S \quad \boldsymbol{H}^E]^T$	$[\boldsymbol{res}^S \quad \boldsymbol{res}^E]^T$
5.	$[1 \quad 0 \quad 0]$	$\boldsymbol{H}^M$	$\boldsymbol{res}^M$
6.	$[0 \quad 1 \quad 0]$	$\boldsymbol{H}^S$	$\boldsymbol{res}^S$
7.	$[0 \quad 0 \quad 1]$	$\boldsymbol{H}^E$	$\boldsymbol{res}^E$
8.	$[0 \quad 0 \quad 0]$	No filtering, only propagation.	

Results

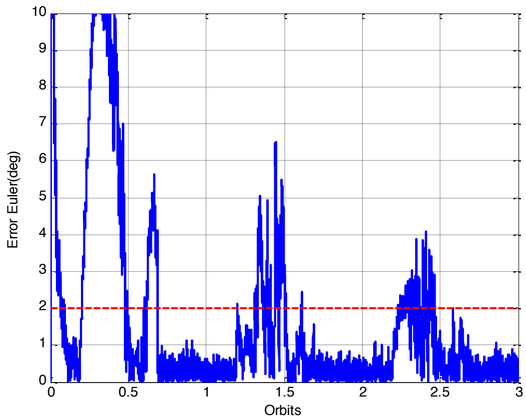
A number of simulations were performed on the cases listed in Table 3. Although there are eight possible combinations, not all of them could meet the $\pm 2°$ pointing accuracy requirement after convergence. The four plots in Fig. 3 show the attitude determination performance of the cases that met the pointing accuracy. The performance is evaluated using the Euler angle error$\Delta\emptyset$, defined in Equation (28)

$$\Delta\emptyset = \sqrt{\Delta Roll^2 + \Delta Pitch^2 + \Delta Yaw^2} \tag{28}$$

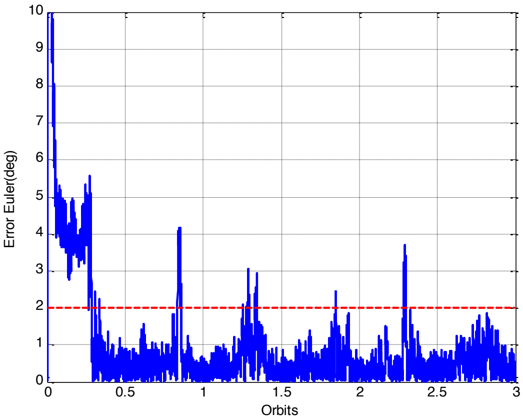
where ΔRoll, $\Delta Pitch$ and ΔYaw are the individual Euler angle errors.



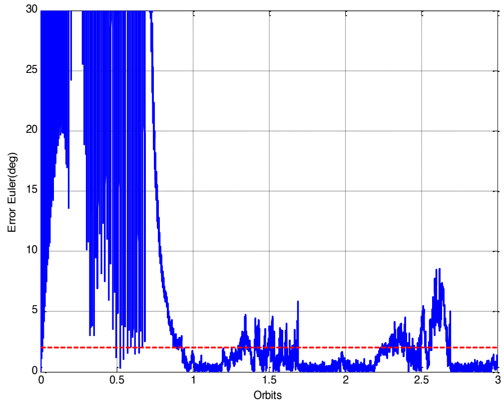
a) Case 1: $[M \quad S \quad E] = [1 \quad 1 \quad 1]$



b) Case 2:$[M \quad S \quad E] = [1 \quad 1 \quad 0]$



c) Case 3:$[M \quad S \quad E] = [1 \quad 0 \quad 1]$



d) Case 4:$[M \quad S \quad E] = [0 \quad 1 \quad 1]$

Fig. 3: The Estimated Euler Angle Error

The four plots in Fig. 3 show the estimated Euler angle error of the first four cases listed in Table 3. The first region within one orbit is defined as the convergence phase. With all sensors switched on, the Case 1 shown in Fig. 3 a) takes only half orbit before the error has converged. *Fig. 3* b) shows the result without Earth vector. It can be seen that the convergence is longer than Fig. 3 a) due to the absence of Earth. The common point in Fig. 4 a) and b) is that better pointing accuracy can be achieved when the sun vector is measured. Fig. 3 c) presents the result without Sun observation. The Earth vector ensures the convergence but the accuracy is lower. To further explain the importance of Sun vector,
Fig. *4* shows the norm of the error covariance matrix versus the availability of the Sun vector. Obviously the error covariance increases during eclipse (Sun Signal =0) and reduces quickly once the Sun is measured (Sun Signal =1).

The cases 1, 2 and 3 have set the magnetometer on all the time. Though the magnetometer is the lowest accurate sensors, it ensures the convergence of the results. From Fig. 3 d) which plots the results without magnetometers, it can be seen that the convergence phase is longer and the estimation crosses the boundary after the convergence.

The results of the case 5 to 8 are not presented here as these cases are failed to meet the pointing accuracy requirement. The main reason is that only one vector makes the attitude unobservable. In the occasions when these cases (case 5, 6, 7, 8) happened, though the designed filter can still work and output the simultaneous attitude results, the results will be untrustworthy. This means the onboard control system should be capable to restart the failed sensor after it was detected off.

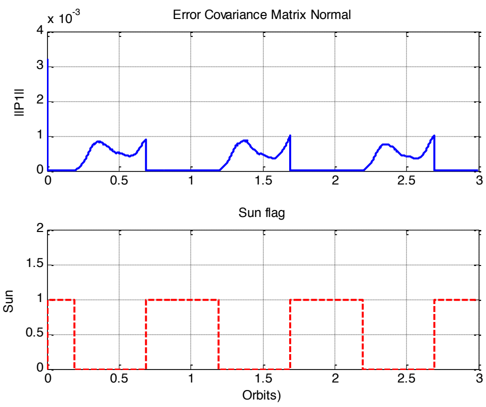


Fig. 4: The effect of Sun to the filter in case 2

Conclusion and Further Work

The paper presents the algorithm and simulation results of the designed attitude determination system in UNSW EC0 project. The results from the simulation show that the developed EKF can achieve the expected performance with demonstrated utility in fusing the various sensors at various update rates. Gyroscopic inputs are used to establish the attitude kinematic system and the magnetic data is the measurement used to overcome the divergence of the attitude propagation. It is considered as the baseline design for the ADS, i.e. ADS has full access to the gyroscopes and magnetometer. An integrated ADS consisting of a Sun Sensor and an Earth Sensor will provide a better solution, As the Sun sensor provides the most accurate measurement of all of the sensors, it is suggested that the Sun Sensor should be switched on when satellite flies under the Sun. There are some uncertainties in the attitude determination algorithm. For example, the question as to whether to use the IGRF model or a simpler periodic model for estimating the earth's magnetic field will be determined according to the actual hardware performance.

Future work is to establish the experimental characterizations of the selected sensors that are required to put the algorithm into practise. For instance, the specifications of the gyroscope on its data sheet are needed to verity the gyro model with the real experimental data. In addition, the placement of all the sensors might be problematic due to the magnetic interference. Thus it is recommended to do the interference test following assembly of the CubeSat.

Further work also includes integration of the attitude control, as well as adding orbit determination part to ADCS. A full Monte Carlo verification of the EKF performance should be performed to provide detailed performance parameters for the EKF and the simulation. Furthermore, running EKF in hardware in loop simulation will provide more realistic performance information which should be tested in various scenarios before the flight.

Acknowledgements

The authors gratefully acknowledge the support of the following partners. The UNSW EC0 is part of the education programme Space Engineering supported by UNSW and Australian Space Research Programme (ASRP).

References

[1]. QB50 Project Description, <https://www.qb50.eu/index.php/project-description> (Retrieved November 2013)

[2]. Qiao, L., Osborne B, Dempster A.G., Attitude Control System (ACS) of UNSW QB50 project "UNSW EC0". Australian Space Science Conference. UNSW2013

[3]. Innovus. CubeSense <http://www.innovus.co.za/media/component-sales-pdfs/CubeSense_web.pdf> (Retrieved November 2013)

[4]. Bowen, J.A., On-board Orbit Determination and 3 axis attitude determination for Pico satellite applications. Master thesis. California Polytechnic State University, July 2009

[5]. Kane, P.C., and Carlisle, P., “Spacecraft attitude dynamics”. J. Wiley, 1986.

[6]. Shuster, M. D. and OH, S. D., "Three-axis attitude determination from vector observations", Journal of Guidance, Control, and Dynamics, Vol. 4, No. 1 (1981), pp. 70-77.

[7]. Murrel, J.W., Precision Attitude Determination For Multimission Spacecraft. In Proceeding of the AIAA Guidance, Navigation and Control Conference, Palo Alto.CA. 1978 p70-87

[8]. Crassidis J. and Lunkins J., “Optimal Estimation of Dynamic Systems,” New York, USA Chapman & Hall/CRC, 2004

Field Testing Marsobot, A Mars Society Australia Robotics Project.

S.W. Hobbs [1], J.D.A. Clarke [2], G.A. Mann[3] [3]

[1] *School of Physical, Environmental and Mathematical Sciences, University of New South Wales, Australian Defence Force Academy, Northcott Drive, Canberra, Australian Capital Territory 2600, Australia*

[2] *Mars Society Australia. P.O. Box 327, Clifton Hill, VIC 3068, Australia.*

[3] *School of Engineering and Information Technology, Murdoch University, South Street, Murdoch WA 6150, Australia*

Summary: This paper introduces Mars Society Australia's Marsobot project, and describes the performance of robots during field testing at Arkaroola, South Australia. Two teleoperated rovers, one four-wheeled (Little Blue) and a larger eight wheeled machine (Miner), have been built using low-cost off the shelf components. Both rovers underwent standardised DHS-NIST-ASTM tests over the period 5 – 12 July, 2014. The tests were conducted in controlled conditions and were designed to provide useful engineering data on the rovers' range and mobility, as well as highlight potential flaws and limitations in design. Both Little Blue and the Miner performed well in the tests, though specific limitations in design robustness and endurance were observed. Lessons learned from these tests will be incorporated into future improvements of the rovers, and refining of the Marsobot project overall.

Keywords: Mars, analogue research, rover

Introduction

No manned mission has ever been to Mars and the only methods currently available for exploration are remote sensing and the use of robotic lander missions. To date eight successful landing mission have visited Mars, including static stations and rover missions [1]. Although the static landers provided a wealth of data regarding their immediate environment, the use of mobile robotic platforms vastly extended the amount of terrain that could be explored [2]. Thus the use of mobile platforms continues to be a preferred method of Martian exploration. Earth-based concept and trainer hardware has been and remains an integral component for developing space missions. Examples of this include the Rocky series of rover trainers that were used as concept testers for the Sojourner mission [3], as well as the FIDO precursor to the Mars Exploration Rovers (MER) Spirit and Opportunity [4]. In addition, Earth hardware is ideal for training students in the applications of remote exploration [5]. Similarly, a range of trainers have been used to study the role of human-centred robotics in exploration [6, 7]

We have developed Marsobots, wheeled training platforms based on commercial off the shelf (COTS) hardware and open source software within a limited budget. In this contribution we describe the testing of two Marsobot rovers against engineering criteria developed by the US National Institute of Standards and Technology (DHS-NIST-ASTM for response robots [8, 9]) during MSA's Arkaroola Robot Challenge in July 2014. We will also compare these designs with some existing machines. We will discuss their performance in the tests and suitability for planetary science test platforms as well as lessons that were learned from the experience.

2.0 Marsobot Platform Description

The Martian environment has proven extremely harsh for machines, with over half of all mission sent to the Red Planet failing [1]. Successful surface based operations have been hampered by the large temperature differential between day and night, the abrasive dust and the overall harshness of the Martian terrain. Although the development of space-qualified equipment is far beyond the budget of the Marsobot project, engineering and mission testing of Earth based vehicles in analogous environments has proven invaluable for preparing for actual Martian operations [4].

Arkaroola is located in the Northern Flinders Ranges, South Australia, and was chosen as the Mars Society Australia's (MSA) primary Mars analogue research site [10, 11]. The diverse geology, semi-arid conditions and temperature extremes have made it an ideal site for testing various aspects of Mars research [12].

The Marsobot project fielded two rovers of different scales. We considered purchasing existing robotics platforms such as the Pioneer 3-AT, Coroware Corobot robot platform, or the Seekur Jr for the Marsobot project [13]. These platforms are four wheel drive machines designed for smaller scale robotics applications, education and research and are approximately the desired size for the Marsobot project. Table 1 lists key attributes of each considered platform as per manufacturer specifications [13, 14].

Table 1: Comparison of established rover platforms.

Name	Pioneer 3-AT	Corobot	Seekur Jr
Weight [kg]	12	8	77
Dimensions (LxWxH) [mm]	500x490x260	350(L)x340(W)x220(H)	1198x835x494
Drive	4WD skid steer	4WD skid steer	4WD skid steer
Max speed [m/s]	0.8	0.36*	1.2
Endurance [hr]	8 hrs	2.5 hr	3.5 hr
Max Slope [degrees]	20	~35*	35
Sensors	Ultrasonic range finders, laser scanner, GPS	IR range detectors, camera, GPS	LIDAR, laser range finders, stereo vision, IMU
Cost (USD)	10500	5200	8000

* Determined by NIST testing

Although the manufacturer specifications would have allowed any of the above platforms to be used in a simulated Martian environment the cost of these platforms are orders of magnitude greater than our budget. Both Little Blue and the Miner were built for approximately $1600. Additionally, purchase of these platforms would negate the flexibility and training value of designing and constructing custom made hardware from readily available, open source materials.

The engineering and mission constraints and chosen hardware we used for both rovers are described in greater detail below. Little Blue (Fig. 1) was built to approximate the size of Sojourner, a microrover deployed to Mars in 1997 [15]. The Little Blue chassis is based on a heavily modified Toyabi Monster truck. The original motors have been replaced with four Dagu 1:131 12V motors coupled to the hub of each wheel and driven by a Sabretooth 2 X 24A motor controller to provide four wheel drive.

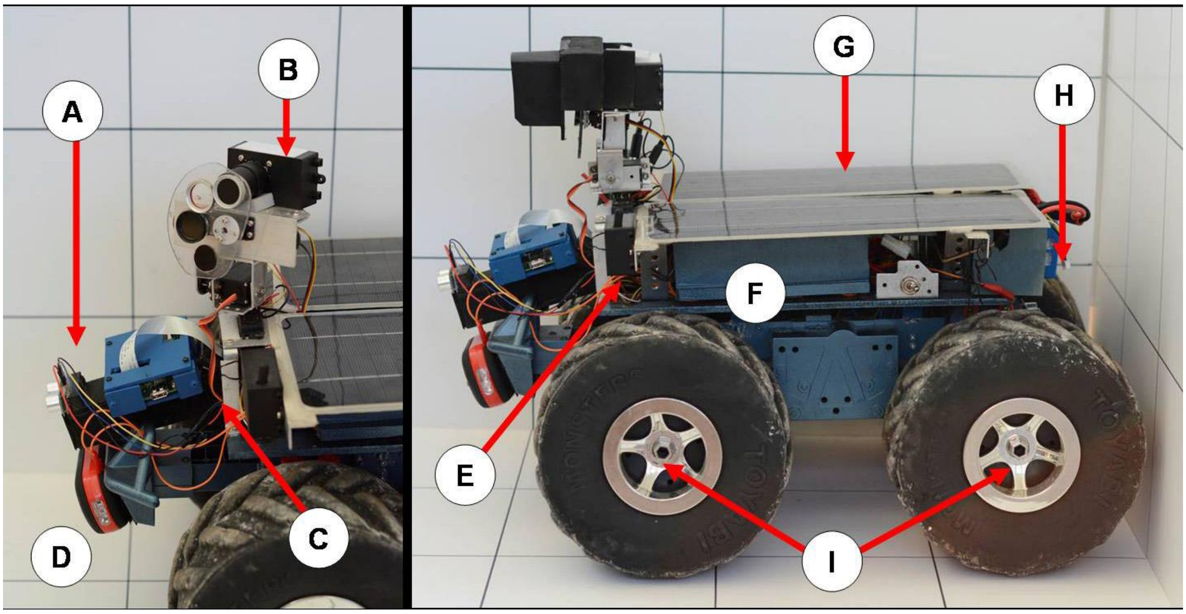


Fig 1. Overview of Little Blue rover with key components marked. (A) Forward distance sensor. (B) Multispectral imaging system. (C) Reflectance spectrometer. (E) Weather sensors. (F) Electronics box. (G) Solar panels. (H) Rear distance sensor. (I) Four wheel drive skid steer mobility.

Little Blue is equipped with range finding ultrasonic sensors [16] and a first person video (FPV) camera fitted with a filter wheel. The camera captures multispectral images by sequentially photographing through one visible and three near infrared (NIR) filters. These separate images are transmitted to a ground station where they are exploited further. This type of camera was also used on the Mars Pathfinder mission [15]. Little Blue also carries a visible light reflectance spectrometer, controlled by a Raspberry Pi Model A on a WiFi link. This instrument collects spectra from materials of interest in order to identify mineralogical composition [17]. A BMP085 temperature and pressure sensor [18] and a DHT 11 temperature and humidity sensor [19] provide Little Blue with the ability to sense weather information. Fig. 1 shows the location of the power and controller systems for Little Blue. We have based control of the sensor and science suite of both rovers on the Arduino 8-bit microcontroller. This microcontroller is well supported and has been used in many robotics applications [20, 13]. The Arduino Uno is programmed using the open source Arduino environment and possesses 14 digital input/output pins and six analog input pins. We use these pins on Little Blue to operate a relay to power the camera, interface with the forward and rear ultrasonic sensor, control the camera filter wheel servo and read data from the two weather sensors.

The rover employs independent power systems for the operation of the mobility system and the science system to provide redundancy allowing for useful science to be conducted in case of mobility failure, and vice versa. The mobility system is powered with a 14.6V 5Ah Lithium Polymer (LiPo) battery, while an 11.2V 2.4Ah LiPo battery powers the wireless camera and Arduino. Power is supplied through two solar panels to a maximum of 16 W. The solar panels are connected via a LiPo charger wired into the 11.2V battery. Aft of the solar panels is a rear facing ultrasonic sensor, providing some degree of hazard detection while reversing the rover.

Little Blue is achieved using skid steered in a similar manner to the Soviet Lunokhod rover [21]. Skid steering (tank steering) operates by varying the speed of motors on either side of the vehicle. We note skid steering, due to its simplicity and robustness has proved to be the mechanism of choice for many established education and military robotics platforms [13, 22, 23].

Design and overall layout for the Miner rover (Fig. 2) is similar to that of Little Blue, though the Miner has twice as many wheels and is much larger. The Miner is fitted with a visible light FPV camera and a two-band multispectral camera atop its imaging mast. This instrument uses an NIR sensitive camera filtered with a #25 red filter, and controlled by a Raspberry Pi. This camera system allows for the capture of red/NIR ratios and has been used for vegetation analysis [24]. The 12V 12Ah battery of the Miner is supplemented by a 10W solar panel. An additional 11.1V battery for its camera system provides redundancy.

The Miner has sufficient torque to carry additional payloads of over 6 kg, such as an environmental logger that operates independently and passively records UV light intensity, temperature and humidity onto an SD card for post mission download and analysis (Fig. 2). Additional instruments could include a BMP085 temperature and pressure sensor identical to Little Blue and a Ublox GPS module. Suspension is designed around a skid-steered, eight wheeled rocker bogie system with each wheel driven by a 6V electric motor drawing up to 2A at full load. The Miner rover also uses a similar Arduino Uno architecture, and interfaces with GPS, sensors and camera relay. Communication between the Arduino of both rovers is accomplished by using AP220 2.4 Ghz serial wireless modems communicating at 9600 baud.

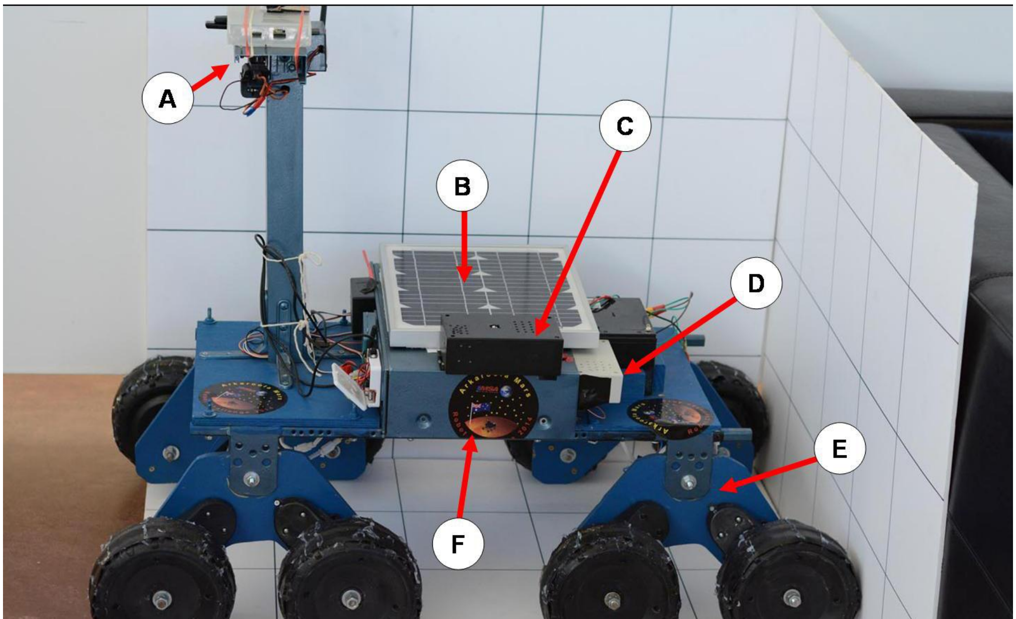


Fig 2. Overview of the Miner rover with key components marked. (A) Multispectral imaging system. (B) Solar panel. (C) Weather logger. (D) GPS and weather sensors. (E) Bogie suspension. (F) Electronics box.

Chassis and mobility

Control and Programming

All of the rovers so far sent to Mars have possessed a degree of autonomous operation (high-level commanding) necessitated by the 8 - 48 minute time delay for communication signals to make the round trip between Earth and Mars [1]. In order to simplify the overall design and remain within budget we opted for a teleoperated mode of rover control. This mode is frequently used for the operation of search and rescue and military support robots where simplicity and reliability are critical requirements [8]. It is appropriate for robots that would assist astronauts in operational tasks [25], where real-time or near real-time communication is feasible, such as at a manned Mars base or in low Mars orbit. Studies postulate future astronaut missions to Mars would benefit from the assistance of robotic devices deployed near the astronauts base of operation and teleoperated directly [26]. Both Little Blue and the Miner rovers ~~was achieved~~ are controlled by a combination of a 2.4 MHz six channel R/C system. This was chosen for its low cost, simplicity, reliability and range, (approximately 200m). Similar R/C systems are routinely used in response robots [8].

DHS-NIST-ASTM Testing

Although we designed and built the Marsobot rovers to broadly conform to existing planetary machines we aimed to quantify their engineering performance based on standardised, widely accepted tests. Over 60 standardised robotic testing methods were developed by NIST in order to characterise the performance and applicability of robots to emergency response situations [9]. The test methods were developed to standardise critical robot capabilities including endurance, mobility, communications, remote sensing, logistics, safety and human/machine interaction [8, 9]. The NIST's extensive usage of these tests on a range of machines have provided emergency responders, engineers and developers useful and repeatable measurements in simulated emergency environments [9], though they do not release details publically. This study consisted of six engineering tests in standard, reproducible test apparatuses and two operational tests i.e. more closely approximating a real operational scenario for the machine as in [27]. The target machine is to be teleoperated in real time by an operator who is out of direct visual and audible contact and are designed to characterise the strengths and limitations of competing designs. Tests performed by Little Blue and the Miner included the following:

- Logistics: Robot Test Config and Cache Packing. The process required the completion of forms for every participating machine to capture details of the physical properties, equipment specifications, configurations, toolkit, packing and transport logistics.

- Energy/Power: Endurance: Terrains: Pitch/Roll Ramps. A test rig consisting of 15° wooden ramps measuring 1200 x 600mm was laid out in a specified alternating sawtooth pattern to repeatably measure the robots' performance on discontinuous terrain. Operators guided the robots around a 15m figure-eight path on the ramps around two vertical pylons. Distance and time from full battery charge to inoperability are measured.

- Mobility: Terrains: Flat/Paved Surfaces (100m). Two pylons were placed 50m apart on a flat surface. The ground around each was marked with a circle 2m in diameter. The robots were to make 10 timed figure-of-eight laps around this course, without deviating from the circumscribed path. Thus both speed and control are important.

- Mobility: Towing: Grasped Sleds (100m). The robots dragged an aluminum sled, carrying an operator-designated payload, around 10 figure-of-eight laps on the 100m course specified in the third test.

- Radio Comms:Line-Of-Sight Environments. The robots were tested for navigation control and video feed on a straight course at 50m, then stations every 100 m thereafter. The last station at which both navigation control and video of 100 x 100 mm hazardous targets were perfectly reliable (complete circle and all four visual tests correct) was reported.

- Sensors:Video:Acuity Charts and Field of View Measures. The robots were placed on a 15° ramp 6 m and then 40 cm from a far-field Landolt-C vision chart. The operator viewed the chart at their control station via the robot's camera and read down the chart to the smallest line at which the orientations of the C shapes were discernible. No more than two errors were permitted on a line. This is reported as a percentage of the 6-6 (20:20) vision standard.

The operational tests were conducted in an old road base quarry at Wooltana station, chosen as it presented a Mars-like variety of smooth, rocky, eroded and rough terrain within a small area. The operational tests included:

- Irregular Terrain Traversal. A 106m course consisting of four gates (1.2m pylons spaced 2 m apart) was arranged over rough, natural, Mars-like terrain. It included slopes of between approximately 20° - 40°, loose sand, and large irregular stones.

- Context Imaging. A small, brightly painted 100g target object was 43 to 76m from the starting point. The operator was to locate the object as quickly as possible, then photograph it in context. Time to locate the target and distance to target were recorded. Each of four best images were examined by three expert field geologists who scored each. The mean rating over all images and criteria was then calculated. The Miner was unable to participate in this test due to time constraints.

An additional, qualitative field test that was not part of the NIST suite was conducted on Little Blue's imaging system. Arkaroola contains an abundance of fossil stromatalites – some of the oldest forms of life on Earth [28]. As discovery of similar evidence of fossil life on Mars would be of great interest to exobiologists, a robotic method of identification of these fossils would be of great interest to the planetary science community [29]. Our trial consisted of Little Blue photographing stromatolite bearing rocks with non-stromatolitic control rocks at set distances of 1m, 50cm and 25cm respectively. The photographs were then sequentially presented to an experienced palaeontologist who was then asked to positively identify the stromatolite bearing rocks from the control group.

Results and Lessons Learned

Fig. 3A-C shows the series of true-colour images returned from the Little Blue vision system of the stromatolite identification test site. The maximum resolution of the images was 640 X 480 pixels and progressively shows greater detail as the rover was driven closer to the target. Textures became visible in the target rocks as Little Blue reached 50 cm (Fig. 3B), with striations and banding discernable at 25 cm, the closest Little Blue was able to approach the target (Fig. 3C). We note that although these details were observed in Little Blue imagery, as well as the overall colours of the rocks, the paleontologist was unable to positively distinguish between stromatolites and naturally occurring features. The results of Little Blue's qualitative image acuity test has implications on the type and resolution of sensors required for positively identifying potential fossil life using a remote sensing platform. The difficulty of identifying past or present forms of life strictly by surface remote sensing and tradeoffs between camera resolution and power requirements has been a concern for planetary missions [4, 29, 30, 31, 32].

Possible methods for improving the capability of Little Blue in discerning between biological and non-biological forms would be an upgrade to a higher resolution camera. Field measurements of the rock samples used for the camera test indicated an average spacing of stromatolite layering, a diagnostic feature used to discern this life form [28], to be on the order of ~1 layer per pixel. Little Blue's camera was capable of resolving 2 mm/pixel at a distance of 25 cm from the target rocks. We thus estimate that the resolution of the camera would need to be 1024x768 pixels or greater to achieve the required resolving power, equivalent to the Pancams carried on board the MER machines [31].

Microscopic analysis has also been an established process for investigating existing or relict life forms [32, 33, 34] and provides a level of detail not available to mast-mounted imaging systems, such as those aboard MER [31]. Past microscopic imagers carried aboard Fido, MER and Curiosity were designed to simulate a geologist's hand lens and were capable of resolutions of ~30um/pixel [35, 36]. Data returned from these instruments have provided critical information on fine scale geology and would have greatly assisted in distinguishing the fossil stromatolites from the control samples. The addition of a microscopic imager, and possibly an IR spectrometer to determine mineralogy, similar to those carried aboard MER [31] would also provide detailed closeup resolving power, plus the added advantage of identifying structure at the microscopic scale [36]. The inclusion of such instruments would come at the expense of greater power and bandwidth requirements, and higher cost. Additionally a microscopic imager would require some kind of postioning system, such as an articulated arm, in order to place the instrument. These would equate to a larger vehicle size in order to accommodate the increased weight from larger solar panels and batteries, as well as requiring a larger, more powerful drive system.

Positive imaging results from a future Mars missions may not be sufficient to resolve the question of extraterrestrial life on Mars – many independent observations may be needed. Nevertheless, Little Blue's qualitative test has highlighted the importance of some of the factors, such as image resolution that are required to make meaningful assessments on the identification of extraterrestrial fossils.

Fig. 4A-D shows an overview of testing of the Marsobot rovers. The mobility and endurance testing of the Miner rover and Little Blue is shown in Fig. 4A-B, and the preliminary results of context imaging tests from Little Blue are shown in Fig. 4C-D. Table 2 lists the NIST test performance of both of the Marsobot rovers and Corobot in more detail. Since the NIST tests are comparatively new and because the NIST does not publish individual results for specific machines, few comparative data are yet available. A comparison is made here with the only comparable COTS machine to undertake the same trials at Arkaroola, namely a Corware Corobot. As shown in Table 2 Little Blue's overall performance was good considering the cost. Little Blue's small dimensions and large wheels with respect to the rover size assisted the rover's maneuverability. The small size and high power torque of the motors assisted its negotiation of the 15^{o} maneuvering ramps (Fig. 4B), as well as the off-road test (Table 2). The solar panels delivered enough energy to the science systems to keep them operating for approximately 2 hrs.

Table 2: Test results of Little Blue, Miner and Corobot rovers.

Rover	**Little Blue**	
NIST Test Method (metrics)	**Result**	**Comments**
Energy/Power:Endurance:Terrains:Pitch/Role Ramps (distance, time - full charge to inoperability)	105 m 2363s	Completed 7 repetitions (mean speed 0.04m/sec)
Mobility:Maneuvering Tasks:Sustained Speed 100m (mean time per repetition)	214.8 s/rep.	Completed 10 repetitions (mean speed 0.47m/sec)
Radio Comms: Line-of-Sight Environments (Max distance for functioning control and vision)	200m	Note: range measured at 100m stations
Visual Acuity (Far Field lowest line Landolt C chart, Near Field lowest line)	15% 25%,	
Operational Test Method (metrics)	**Result**	**Comments**
Operational Task: Irregular Terrain Traversal (time to complete, average speed)	531s 0.2m/s	Performed well despite hard wheels
Operational Task: Context Imaging (time to locate target, % mean rated quality of 4 context images)	417 s 65%	Distance to target 43.12m Quick to locate target object!
Rover	**The Miner**	
NIST Test Method (metrics)	**Result**	**Comments**
Energy/Power:Endurance:Terrains:Pitch/Role Ramps (distance, time - full charge to inoperability)	Abstained	Vehicle dimensions do not agree with ramp dimensions
Mobility:Maneuvering Tasks:Sustained Speed 100m (mean time per repetition)	297.75s/rep.	Completed 8 repetitions (mean speed 0.34m/sec)
Mobility:Maneuvering Tasks:Grasped/ Hitched Sleds 100m (weight, mean time per repetition)	6.47kg 299s/rep	One repetition only – test aborted due to battery failure
Visual Acuity (Far Field lowest line Landolt C chart, Near Field lowest line)	15% 20%,	
Operational Test Method (metrics)	**Result**	**Comments**
Operational Task: Irregular Terrain Traversal (time to complete, average speed)	941s 0.11m/s	Lost some time on slopes
Rover	**Corobot**	
NIST Test Method (metrics)	**Result**	**Comments**
Energy/Power:Endurance:Terrains:Pitch/Role Ramps (distance, time - full charge to inoperability)	150m 2032s	V_{ini}=13.8v V_{final}=12.77v 8 topples (top heavy)
Mobility:Maneuvering Tasks:Sustained Speed 100m (mean time per repetition)	280s/rep.	Completed 10 repetitions mean speed 0.36m/sec
Mobility:Maneuvering Tasks:Grasped/ Hitched Sleds 100m (weight, mean time per repetition)	3.47kg 369.1s/rep	Completed 10 reps, no fails (mean speed 0.27m/sec)
Visual Acuity (Far Field lowest line Landolt C chart, Near Field lowest line)	5% 120°	Genius Widecam 1050
Operational Test Method (metrics)	**Result**	**Comments**
Operational Task: Irregular Terrain Traversal (time to complete, average speed)	1282s 0.08m/s	Forfeited Gate 1 – slope too steep – lost time here

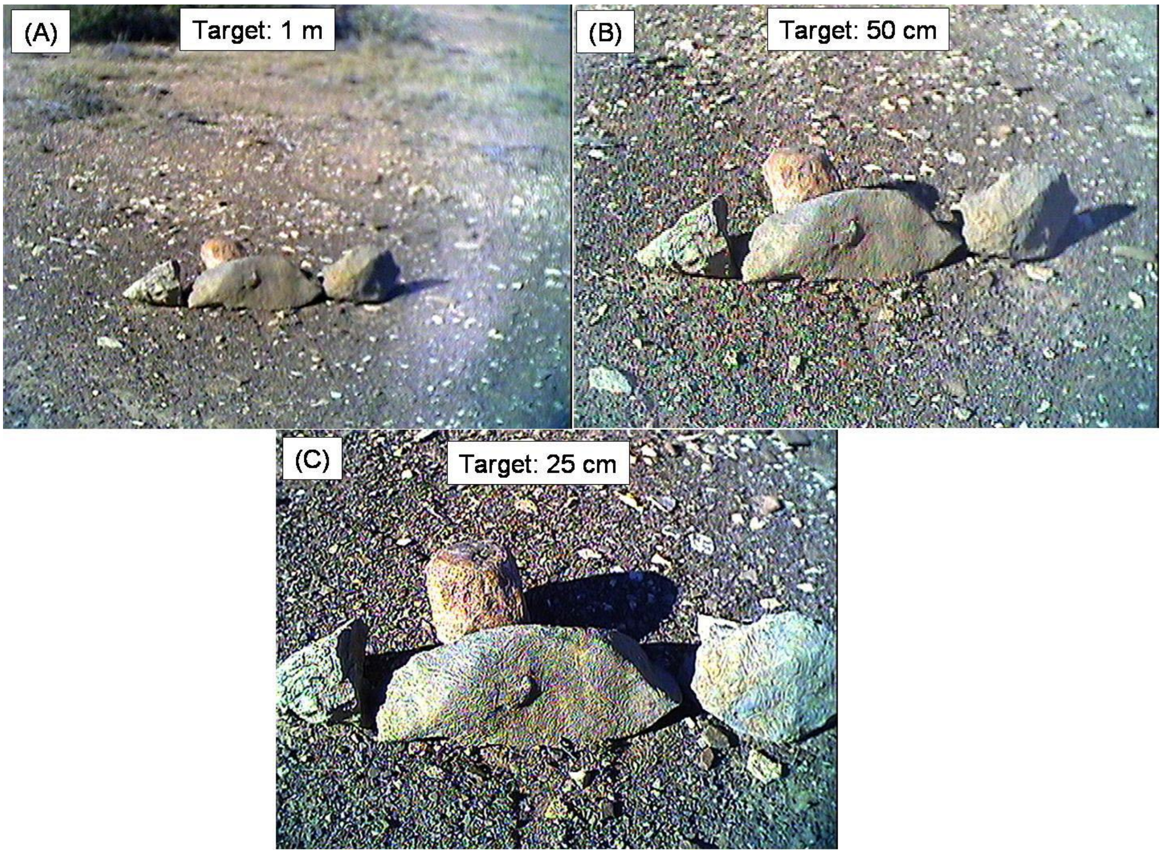


Fig 3.Results from Little Blue qualitative fossil identification trial. Target consists of a mixture of stromatolite and non stromatolite bearing rocks. (A) Image from 1 m distance. (B) Image from 50 cm distance. (C) Image from 25 cm distance.

Little Blue was able to complete 10 laps of the sustained speed test, as well as seven complete laps of the energy/endurance test (Table 2). In addition the R/C controller, FPV camera and wireless modem ranges (~200 m, Table 2) exceeded the initial engineering design for Little Blue.

Visual acuity of Little Blue's camera system was 15-20%, reflecting the limited resolution of the wireless camera. The visual acuity of the system was sufficient to allow for finding of an imaging target in the search and find field test in under seven minutes (Fig 4C, Table 2). In addition a false colour NIR image was able to be obtained (Fig. 4D), from which the absorbance characteristics of the target were illustrated.

Despite the overall performance of Little Blue, problems were observed. Little Blue shed its front left wheel during the ramp test. Although field repairs were quickly made, this catastrophic failure highlighted the criticality of a robust coupling of wheels to motor. Further modifications for future designs may include additional batteries in order to provide greater endurance, though this would be at the cost of increased weight and greater strain on the motors.

The Miner Rover also performed well on the DHS-NIST-ASTM tests [8] (Table2). The Miner was able to negotiate the outdoor irregular terrain traversal test at a speed of 0.11m/s, despite suffering a failure of a microprocessor power regulator in one irregular terrain traversal. Visual acuity was lower than for Little Blue, (15-20% cf 15-25%, Table 2), probably due to the wider field of view of the imaging system. In addition, the Miner was able to pull more than 6 kg over flat ground during the sled dragging attempt.

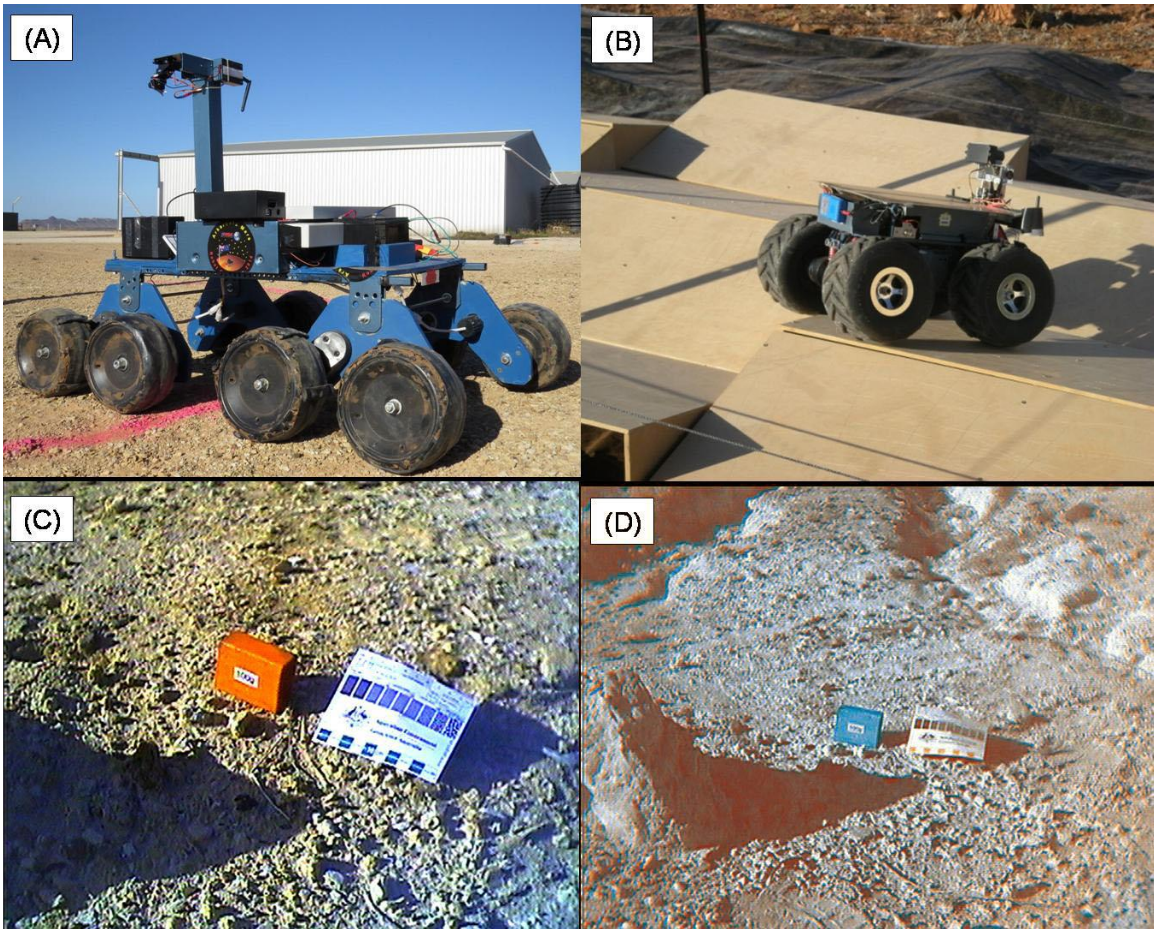


Fig 4. Overview of NIST tests for Marsobot rovers. (A) The Miner undertaking the endurance tests. (B) Little Blue rover on mobility ramps. (C) Context image of target imaged by Little Blue rover. (D) False colour image of context target using NIR. Blue tone of target denotes NIR absorbance of target.

As with Little Blue, the testing revealed some design flaws and limitations of the Miner rover. It was found that the power supply for the Miner was inadequate, with the rover unable to complete 10 laps of the sustained speed test (Table 2). This was due to complete discharge of the 12V battery supply, for which the solar array was unable to compensate. On flat ground similar of the type used in the NIST test the Miner draws power of up to 200W, the motors together drawing 16A. Replacing the existing 12 Ahr battery with a higher capacity accumulator of at least 16 Ahr would allow the Miner to complete a future NIST endurance test of this type. Additionally the replacement of the existing 10 watt panel with a 60 watt version, similar to that used in the Solar Rover [22], would increase the power accumulation capacity of the Miner by 600%. The implementation of a 1:3 drive/rest cycle would improve power management of the vehicle and ensure adequate power generation for the vehicle. The Miner has capacity for a heavier duty power supply which will be included in future upgrades and be the subject of future research.

The plastic wheels of the Miner provided insufficient grip on the surface to climb steep slopes. This issue was partially mitigated on site by addition of rubber strip treads across the wheel width of four wheels to increase traction. Further traction losses arose from rocker bogie pivots being too stiff and suspension lift off occurring during high energy turns [37], caused by the wheel base being too narrow. Extending the width of the Miner rover and providing better freedom of movement for the bogies would provide for better all-wheel ground contact over irregular terrain and thus greater mobility. Integration of the ground station controls would also be of benefit the operation of both rovers. The current system requires the rover operator to switch attention between three independent systems (R/C control, sending and receiving commands, capturing imagery). Combining these interfaces would be likely to significantly reduce the workload of the rover operator, reducing the need to coordinate between systems for mobility control, image acquisition and receiving telemetry data.

Part of the study involved evaluation of the robots in the astronaut support role. Both robots accompanied pairs of people wearing simulated space suits and provided support such as video coverage of their activities, imagery of rocks and, in the case of Little Blue, spectral data (Fig 5). While both robots were able to perform these tasks, they found it difficult to keep up with the simulated EVA team over smooth ground and were not able to deal at all with rougher terrain. While further analysis is needed, these preliminary observations suggest that small rovers are better suited to deploy-and-leave operations, where they are carried to a site and then left to carry out further studies independently or to explore area unsuited to astronaut teams, such as soft ground, or lower priority scientific areas. The EVA support role might be better carried out by suitably equipped unpressurised support vehicles. This is exemplified in the R-Gator tested by the United States Military, which is able to act as a light transport for personnel, autonomously follow them while they are on foot, or operate under remote control in a support role [38].

Fig 5. Marsobot rovers in EVA support role. (A) Little Blue. (B) Miner.

Conclusion

The Marsobot project represents the Mars Society's work in developing analogue robotic vehicles that are capable of trialling future experiments for actual planetary missions. Readily available, off the shelf components were used, and an open source development system employed to allow for ease of development and collaboration. The Arkaroola NIST tests provided an opportunity to quantify the capabilities and limitations of both rover vehicles in an outdoor environment. Both the Little Blue and Miner rovers performed well in most of the tests, proving the ability of low cost systems to operate in a Mars analog environment. Lessons learned from the results of these tests will be incorporated into future improvements of the rovers, and refining of the Marsobot project.

Acknowledgements

We gratefully acknowledge the tireless efforts of Dr Graham Mann to design, set up and employ the NIST testing and the help of expedition volunteers. We also acknowledge the support of the staff at Arkaroola Village during the field expedition. The expedition was partly funded by VSIRO and the Australia-India Council

References

1. Barlow, N.G., *Mars: An Introduction to its Interior, Surface and Atmosphere*, Cambridge University Press, Cambridge, 264 pp, 2008.

2. Soderblom, L.A., Bell III, J.F., "Exploration of the Martian surface 1992-2007", in: *The Martian Surface: Composition, Mineralogy and Physical Properties*, ed. J.F. Bell III, Cambridge University Press, New York, 2008.

3. Hayati, S., Volpe, R., Backes, P., Balaram, J., Welch, R., et al., "The Rocky 7 Rover: a Mars sciencecraft prototype". Proceedings of the 1997 IEEE International Conference on Robotics and Automation, 2458-2464, 1997.

4. Anderson, R.C., Haldermann, A.F.C., Dohm, J. and Huntsberger, T., "A Dress Rehersal for the 2003 Mars Exploration Rovers", *Mars Analog Research*, San Diego, California, 2006.

5. Sipos, A., Vizi, P.G., "Simulated Mars model competition". *Proceedings of the 42nd Annual Lunar and Planetary Science Conference*, The Woodlands, Texas. Lunar and Planetary Institute, Houston, TX, abst 2014, 2011.

6. Akin, D. L., Bowden, M. L., Saripalli, S., and Hodges, K. "Developing technologies and techniques for robot-augmented human surface science", in: *AIAA Space 2010 Conference and Exhibition*, Anaheim, California, 2010.

7. Fong, T., Bualat, M., Edwards, L., Flueckiger, L., Kunz, C., Lee, S. Y., Park, E., To, V.,Utz, H., Ackner, N., Armstrong-Crews, N., and Gannon, J. "Human-Robot Site Survey and Sampling for Space Exploration," *AIAA-2006-7425. Proc. AIAA Space 2006,* San Jose, CA, 2006.

8. Jacob, A., Downs, A., Huang, H., Messina, E., Saidi, K., Sheh, R., Virts, A., "Standard test methods for response robot". *ASTM International Committee on Homeland Security Applications: Operational Equipment*; Robots (E54.08.01), 2014.

9. Jacoff, A., Huang, H., Virts, A., Downs, A., Sheh, R., "Emergency response robot evaluation exercise". *Association for Computing Machinery,* 145-154, 2012.

10. Mann, G.A., Clarke, J.D.A. & Gostin, V.A., "Surveying for Mars analogue research sites in the Central Australian deserts". *Australian Geographical Studies*, 42(1), 116-124, doi: 10.1111/j.1467-8470.2004.00247.x, 2004.

11. Heldmann, J., Brown, A., Clarke, J. D. A., Rupert, S. M., and Thomas, M. “Follow the water: applying a Mars exploration strategy to the Arkaroola analog region, South Australia”. *Mars Analog Research*, San Diego, California, 2006.

12. Willson, D., J.D.A. Clarke and Murphy, G.M., “MARS-OZ: a Design for a Simulated Mars Base in the Australian Outback”. *Journal of the British Interplanetary Society* 58, 292-293, 2003.

13. Gonullu, M.K., “Development of a Mobile Robot Platform to be used in Mobile Robot Research”. *Masters thesis*, Middle East technical University, 2013.

14. Corobot User’s Guide, Accessed February 2015 (Online). http://www.coroware.com/products/corobot/userguide.pdf.

15. Farrand, W.H., Bell III, J.F., Johnson, J.R., Bishop, J.L., Morris, R.V., “Multispectral imaging from Mars Pathfinder”, in: *The Martian Surface: Composition, Mineralogy and Physical Properties*, ed. J.F. Bell III, Cambridge University Press, New York, 2008.

16. HC-SR04 Ultrasonic Distance Measuring Module Product Page, Accessed January 2015 (Online). http://dx.com/p/hc-sr04-ultrasonic-sensor-distance-measuring-module-133696.

17. Wang, S.X., Zhou, X.J., “Spectroscopic Sensor on Mobile Phone”, Patent 7,420,663 (US, 2008).

18. Bosch Sensortec, "BMP085 digital pressure sensor datasheet", Accessed January 2015 (Online) http://dlnmh9ip6v2uc.cloudfront.net/datasheets/Sensors/Pressure/BST- BMP085-DS000-06.pdf.

19. Nguyen, T., Slonaker, J., Kadous, M., “Semi-autonomous wireless controlled robot.” *Capstone Senior Design Project*, Indiana University, 2012.

20. Shue, S., Hargrove, C., Conrad, J., “Low Cost Semi-Autonomous Sentry Robot”. *IEEE Proceedings*, 2012.

21. "NASA National Space Science Data Center. Luna 17/lunokhod, 2005. Accessed November 2014 (Online) http://nssdc.gsfc.nasa.gov/database/MasterCatalog?sc=1973-001A.

22. Roman, M.J., “Design and Analysis of a Four Wheeled Planetary Rover”. *Master’s thesis*, University of Oklahoma, 2005.

23. Jacoff, A., Wavering, A., Messina, E., Virts, A., Huang, H., Downs, A., “Response Robots – DHS/NIST sponsored evaluation exercises”. *NIST Technology Administration, US Department of Defence* 90, pp231, 2010.

24. Sugiura, R., Noguchi, N., Ishii, K., “Remote-sensing technology for Vegetation Monitoring using an Unmanned Helicopter”. *Biosystems Engineering* 90, pp369-379, 2005.

25. Culbert, C., et al. “Activities of the NASA Exploration Team Human Robotics Working Group”. *Proceedings of the American Institute of Aeronautics and Astronautics Space 2003 Conference,* Long Beach, Calif., 2003.

26. Sylvan, R.L., “Increased cost effectiveness of Mars exploration using human-robotic synergy”. in Crossman, F., Zubrin, R. (eds) *On to Mars*, 265 pp, 2005.

27. Jacoff, A., Messina, E., Evans, J., “Experiences in Deploying Test Arenas for Autonomous Mobile Robots”. *Intelligent Robots and Systems Proceedings*, 2003.

28. Bengtson, S., “Early Cambrian fossils from South Australia”. *Australasian Association of Palaeontologists Memoir 9*, 364p, 1990.

29. Westalla, F., and 13 colleagues, “An ESA study for the search for life on Mars”, *Planetary and Space Science* 48, 181-202, 2000.

30. Mutch, T.A., Jones, K. L., “*The Martian Landscape*”, NASA SP-425, 160pp, 1978.

31. Bell III, J.F., and 24 colleagues, “Mars Exploration Rover Athena Panoramic Camera (Pancam) Investigation”, *Journal of Geophysical Research*, 2003.

32. McKay, C.P., “The Search for Life on Mars”. In: Whitter, D.C.B. (Ed). *Planetary and Interstellar Processes Relevant to the Origins of Life*, Springer Netherlands, 289pp, 1997.

33. Schinteie, R., Campbell, K.A., Browne, P.R.L., “Microfacies of stromatalitic sinter from acid-sulphate-chloride springs at Pararaki Stream, Rotorua Geothermal Field, New Zealand.” *Palaeontologia Electronigia* 10, 2007.

34. Lynne, B.Y., Campbell, K.A., “Diagenetic transformations (opal-A to quartz) of low- and mid-temperature microbial textures in siliceous hot spring deposits, Taupo Volcanic Zone, New Zealand." *Can. J. Earth Sci*. 40, 2003.

35. Grotzinger, S J., and 25 colleagues, “Mars Science Laboratory mission and science investigation”. *Space Sci. Rev*. 170, 2012.

36. Herkhenhoff, K.E., and 22 colleagues, “Textures of the soils and rocks at Gusev Crater from Spirit’s microscopic imager. *Science* 305, 2004.

37. Carlone, T.J., “Investigating Suspension Lift-Off of Skid-Steer Rovers with Passive Differential Suspension”. Master’s Thesis, Carnegie mellon University, Pennsylvania, 2013.

38. Moorehead, S J., Wellington, C K., Paulino, H., Reid, J.F. “R-Gator: An Unmanned Utility Vehicle”. *SPIE Defense and Security Conference, Proceeding of the SPIE* 7692, 2010.

Field Robotics, Astrobiology and Mars Analogue Research on the Arkaroola Mars Robot Challenge Expedition

Jonathan D. A. Clarke†, Jason M. Held‡, Astrid Dahl‡, Nicolette Wheaton† and the Arkaroola Mars Robot Challenge Expeditioners[1]

†Mars Society Australia, c/o 43 Michell St. Monash, ACT 2904, Australia
Australian Centre for Astrobiology
‡Saber Astronautics Australia Pty. Ltd. (Sydney) Australia

Summary: Mars Society Australia's Arkaroola Mars Rover Challenge Expedition was an initiative of Saber Astronautics and Mars Society Australia. The expedition was based round a program of field trials of robots developed by students from Murdoch University, the University of New South Wales, Mars Society India and Mars Society Australia. This was the first time such an expedition had been attempted in Australia. Robot performance, remote operations, simulated extra-vehicular activities and simulated space suit evaluations were accompanied by geobiological research and training in field geology and media relations. The successful conclusion of the expedition paves the way for further expeditions of this type for Australian and international students and researchers.

Keywords: field robotics, planetary exploration, astrobiology, analogue research, Mars, Arkaroola.

Introduction

The Arkaroola Mars Rover Challenge Expedition, a collaboration with Saber Astronautics and Mars Society India (MSI), is the latest in a series of expeditions run by Mars Society Australia (MSA) since 2001. Destinations of previous expeditions have included the Pilbara [1], the Mars Desert Research Station [2] in Utah as well as Arkaroola [3].

The goal of the expedition was to conduct research across several fields relevant to Mars exploration in a Mars-like environment side by side with educational and public engagement programs. A Mars analogue is an environment or a region that has analogous characteristics to Mars [4].

The Arkaroola region (Fig. 1) was selected by MSA on the basis of 1) its many scientific features of interest to planetary geologists, geomorphologists, and astrobiologists; 2) a diversity of different terrains, materials and surfaces ideal for engineering tests; 3) the education and outreach

[1] K. Abetz, W. Andrews, M. Cooper, A. Dahl, P. Gatkine, S. George, G. Hobbs, S. Hobbs, S. Houlahan, Y. Kabaran, A. Kotwal, A. Kshirsagar, J. Lai, G. Mann, S. McArthur, G. Murphy, R. Prajapat, A. Rajagopal, K. Silburn, N. Small, T. Smith, P. Spinks, V. Sudarsanan, A. Tandi, M. Tanner, N. Wheaton, and T. Wheaton.

potential supplied by this diversity; 4) the logistic support available through the Arkaroola Resort [5], and 5) a heritage of Mars analogue and astrobiological research by MSA, the Australian Centre for Astrobiology and others [6, 7].

Fig. 1: Location of Arkaroola in the Flinders Ranges (Google Earth Image).

The expedition was conducted over a two week period in July 2014 and involved 28 participants. They included students from four universities in Australia and India (Murdoch, UNSW, Macquarie and IIT Bombay), researchers from two Australian universities (Murdoch and Macquarie), MSA volunteers, two researchers from Saber Astronautics and a representative from Fairfax Media. Several researchers and volunteers located in Sydney, Canberra and Melbourne were also tasked by MSA and Saber Astronautics to provide expedition support.

This paper summarises the expedition rationale and activities, providing a context for other papers, including some those elsewhere in these proceedings, which will detail expedition results (e.g. [8, 9. 10]).

Rationale

Purpose of Mars analogue research

Analogue research is an invaluable step in space research and development between the laboratory and an actual mission. Mars (and other Solar System) analogues are also very useful in education and outreach as they place educators and students in environments that resemble aspects of those found on Mars and elsewhere in the Solar System.

Three types of analogue research were outlined by Persaud [4]. These were programs of discovery, opportunity, and investigation. Programs of discovery are those that consist of holistic and realistic simulations (which may none the less be of low fidelity in some aspects) that set up a situation and make qualitative observations of the problems and challenges that arise, both anticipated and unanticipated. Programs of opportunity are those where engineering and scientific research (e.g. geology, biology) is carried out of the analogue region or environment. Programs of investigation carry out experiments of specific parameters and make qualitative observations on operational problems, usually connected to the design and operation of engineering items and instrumentation. Under this classification, the research undertaken on the Arkaroola Mars Rover Challenge Expedition fell into the category of programs of opportunity and investigation with respect to field robotics and operations research, and of opportunity with respect to geobiology.

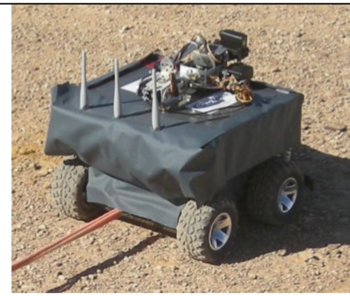
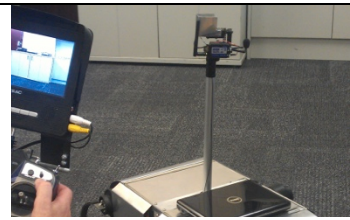
Expedition aims

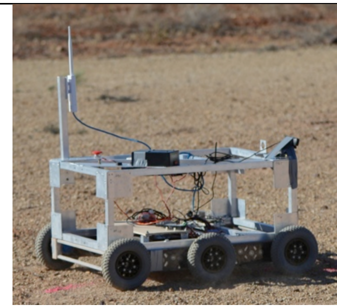
The Arkaroola Mars Rover Challenge Expedition had the following objectives (terminology of Persaud [4]):

- To investigatively test a range of field robotic systems under standardised conditions with goal of developing concepts for planetary operations, especially Mars.
- To opportunistically test robots in the field against conditions and features that approximated in one or more aspects those that might be found on Mars and other Solar System bodies.
- To opportunistically test control of remote robotic and human field operations at Arkaroola from Saber Astronautics' control centre in Sydney.
- To opportunistically conduct a series of simulated extra-vehicular-activities (EVAs), utilising observational data of human-robot systems to model mission performance
- To opportunistically explore effectiveness of field science while wearing simulated space suits.
- To opportunistically explore geobiological features relevant to astrobiology, testing hypotheses regarding the presence of fossil bacteria in veins and presence of sponge biomarkers in sediments.
- To provide instruction to participants into astrobiology, astrogeology, and science communications, and encourage interaction and cross-stimulation between students, researchers and volunteers.
- Enable public outreach via lectures and media contacts as to the importance of planetary science, field robotics, and astrobiology, and
- Inspire and equip STEM teachers via the Spaceward Bound program through exposure to science & engineering related to planetary exploration.

These goals were all met during the expedition, as detailed in the following sections.

Table 1: Arkaroola Mars Robot Challenge Expedition robots.

Robot	Operator	Description	Investigations	Other features	Illustration
MSI rover	MSI	Custom 6X6 rover bogie rover with skid and 4-wheel steering	Remote operations tests	Power: 11.1V LiPo 56Ah Batteries. Sensors: 3 FPV colour cameras 180 degrees FoV, 1 forward, 1 navigation (rear mount) and 1 on the arm; compass and GPS. Comms: 250mW 900MHz RS-232 serial command and control link (ruuning at 19200 bps); 500mW 2.4GHz wireless data link for camera feeds. Other features: 3 DoF (degrees of freedom) arm, with soil sampling scoop attachment or optional gripper attachment.	
Corobot	Murdoch	Modified OTS 4X4 rover	Rover trials	Power: 2 x 6V 5Ah NiMH batteries. Sensors: fixed camera, 2 IR detectors, GPS, force sensors, bump switches. Comms: 802.11n WiFi to laptop. Other features: 5 DoF manipulator with joint angle sensors, power screwdriver	
Mascot	Murdoch	Custom hexapod rover with drive revolute spring legs	Rover trials	Power 2 x 18v 5Ah motors, 1x 11 Ah LiPo batteries. Sensors: Pan-tilt 380 line PAL camera with 1.4 Ghz 1 W video transmitter, microphone. Comms: 900 MHz analog RC (control). Other features: 5 DoF Still camera mount	

Robot	Operator	Description	Investigations	Other features	Illustration
UNSW	UNSW	Modified OTS 6X6 skid steer rover	Rover trials	Power: 2 x 18Ah SLA batteries. Sensors: 1 pan tilt camera, 1 fixed camera, GPS, pitch-roll inclinometers. Comms: 0.63W 2.4Ghz 802.11n WiFi to laptop. Other features: custom 4Dof manipulator, lidar	
Miner	MSA	Custom 8X4 in rocker bogie pairs, skid steer rover	Rover trials & remote operations tests	Power: 1 x 12V 12Ah SLA batteries, 1 x 12.2V 2.4Ah LiPo. Sensors: pan-tilt camera with 750mW 5.8Ghz video transmitter, 2 x sonar detectors. Comms: 0.75W 2.4Ghz analogue RC. Other features: 20W solar panel	
Little Blue	MSA	Custom 4X4 skid steering rover	Rover trials & remote operations tests	Power: 1 x 14.8V 5Ah batteries, 1 x 12.2 Ah LiPo battery. Sensors: pan-tilt camera with 750mW 5.8Ghz video transmitter, 2 x sonar detectors. Comms: 0.75W 2.4Ghz analogue RC. Other features: custom spectrometer using filer wheel, 40W solar panel, LED lighting option	
Phantom 2	MSA	OTS quadcopter	Rover trials & remote operations tests	Power: 5.2 Ah LiPo batter/balancer. Sensors: pan-tiltcamera with 600mW 5.8GHz transmitter to 177.8 mm HDMI video receiver, GPS. Comms: 2.4 GHz analogue RC control. Other features: Helical and omni-directional antennae for video, PC on-screen display	

Expedition Activities

Field robotics

Expedition robots

Seven robots (Fig. 2) were tested on the expedition, their descriptions and operations are summarised in Table 1.

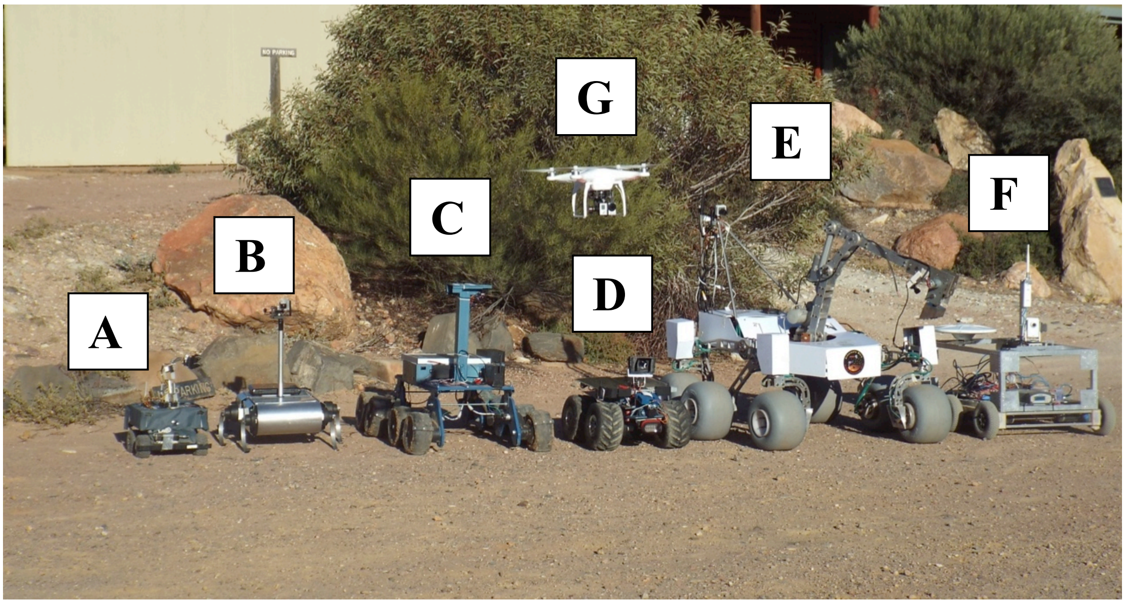


Fig. 2: The seven robots that took part in the Arkaroola Mars Robot Challenge Expedition. A – Corobot, B -Mascot, C-Miner, D- Little Blue, E – MSI Rover, F – UNSW rover, G – Phantom 2.

Robot field trials

Planetary robots can be considered to be a specialised sub-type of field robots. Controlled testing using standard tasks within standard environments allow meaningful comparison of the performance of different designs and guide both improvement of particular designs and design selection for specific tasks. The participating robots in the Arkaroola trials were not envisaged as autonomous explorers such as the current generation of planetary rovers like Yutu on the Moon and Curiosity on Mars. Rather they were conceived as supporting human activity on the surface. Therefore they can be controlled locally, rather than over interplanetary distances. This means that issues of latency and autonomy were not considered in these trials. Such robots could be used for remote activities, required to perform routine maintenance tasks or support astronaut EVAs.

The trials (Fig 3) were devised by Graham Mann of Murdoch University to evaluate the performance of six of the seven robots on the expedition in these areas. Six different standardised tasks, based on those developed by the National Institute of Standards (DHS-NIST-ASTM) in the United States, were performed. Three operational tasks were also developed to evaluate the robots in tasks approximating their intended purpose [8]. The standardised tests were:

1. **Logistics: Robot Test Configuration & Cache Packing.** The process required the completion of forms for every participating machine to capture details of the physical properties, equipment specifications, configurations, toolkit, packing and transport logistics. The information includes specific photographs of a robot, in different poses and from various angles, against a calibrated background. The information is particularly important for managing the configuration of robots from one test to another.

2. **Energy/Power: Endurance: Terrains: Pitch/Roll Ramps.** A test rig consisting of 15° wooden ramps measuring 1200 x 600mm was laid out in a specified alternating sawtooth pattern to repeatably measure the robots' performance on discontinuous terrain. Operators guided the robots around a 15m figure-eight path on the ramps around two vertical pylons. Distance and time from full battery charge to inoperability are measured. It turned out to be impractical to bench test sets of batteries through multiple charge-recharge cycles in the field.

3. **Mobility: Terrains: Flat/Paved Surfaces (100m).** Two pylons were placed 50m apart on a flat surface. The ground around each was marked with a circle 2m in diameter. The robots were to make 10 timed figure-of-eight laps around this course, without deviating from the circumscribed path. Thus both speed and control are important. Average speeds in metres per second were recorded. `

4. **Mobility: Towing: Grasped Sleds (100m).** The robots dragged an aluminum sled, carrying an operator-designated payload, around 10 figure-of-eight laps on the 100m course specified in the third test. Average velocities and maximum achieved weights were recorded.

5. **Radio Communications: Line-Of-Sight Environments.** The robots were tested for navigation control and video feed on a straight course with pylons at 50m and then every 100m thereafter. The robot circumnavigated each station at a radius of 2m, imaging a 35 x 35mm bold letter and a 100 x 100 mm figure on the four faces box atop the pylons. The distance of last station at which both navigation control and video were perfectly reliable (complete circle and all four visual tests correct) was recorded.

6. **Sensors: Video: Acuity Charts and Field of View Measures.** The robots were placed on a 15° ramp 6m from a far-field Landolt-C vision chart. The operator viewed the chart at their control station via the robot's camera and read down the chart to the smallest line at which the orientations of the C shapes were discernible. No more than two errors were permitted on a line. This is reported as a percentage of the 6-6 (20:20) vision standard. The same procedure was used for the near-field Landolt-C chart, except that the distance was then 40cm. The horizontal field of view was calculated by measuring the distance between the chart and the camera.

The three operational tests were carried out at an abandoned road metal quarry. Terrain included slopes of 20 to 40 degrees, loose sand, and large irregular rocks. The operational tests evaluated:

1. **Irregular Terrain Traversal.** A 106m course consisting of four gates (1.2m pylons spaced 2m apart) was arranged over rough, natural, Mars-like terrain. It included a slopes of between approximately 20° - 40°, loose sand, and large irregular stones. Operators

were allowed to walk the course before robot testing. The robots were video recorded and timed during their traversal of the course.

2. **Context Imaging.** A small, brightly painted 100g target object was placed at a random locations on roughly level ground at distances of between 43 and 76m from the starting point. The operator was given the object's GPS coordinates. The operator was to locate the object as quickly as possible, then photograph it in context. Time to locate the target and distance to target were recorded. Each operator chose his best four images to be rated for quality. Each image was made anonymous, then examined by three expert field geologists who rated each according to five criteria: object in context (shows surrounding structure), image composition, brightness and contrast, sharpness of focus and image resolution. The mean rating over all images, experts and criteria was then computed from the ratings and expressed as a percentage of the perfect score.

3. **Sample Return**: Robots equipped with a manipulator had the option to use it in a variation of the Context Imaging task. The robots had to carry a small geologist's scale, place it alongside the located target object, photograph the object in context, collect the object then return it to the starting location. Time to return was reported.

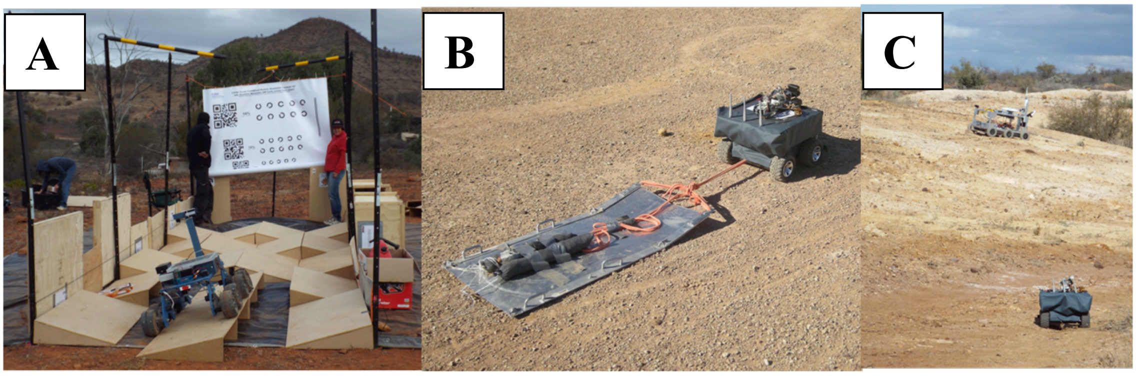


Fig. 3: Robot trials. A – Miner rover in pitch and roll ramps test area undergoing visual acuity test. D – Corbot pulling sled. C- UNSW robot and Corbot in quarry test area.

Remote operations

Space operations centre trials

Control of space assets is a major challenge to space operations. The Saber Astronautics has developed the "Responsive Space Operations Centre" (RSOC) which implements new techniques in diagnosing faults in space systems. These techniques are novel in that they consider space systems and space environments together in the same *global* probabilistic model, allowing operators the ability to rapidly respond to problems that can occur. This is an advantage over current methods which are limited by frequent retraining and a high degree of domain knowledge.

The same principles can be applied to surface operations on other Solar System bodies and to both manned and un-manned missions. They can also include mission analysis and performance optimisation. For example, optimising a human-robotic team, or tracking consumption logistics and workload, can increase the science output of the team. The Responsive Space Operations

Centre provides the tools to tackle the important research question of how the performance of the Mars base as a whole, EVA teams, or indivuaal astronaut or robotic systems can be measured during actual or simulated missions: during analogue missions such quantification is a way to better assess what impacts success or failure in field science.

Throughout the Arkaroola Mars Rover Challenge Expedition, Saber Astronautics tested the performance of the Responsive Space Operations Centre in supporting a simulated Mars mission. A communications link was established between Arkaroola and the Responsive Space Operations Centre at the Saber laboratory at Chippendale NSW. This enabled transmission of rover telemetry and GPS tracking data from the field and Mission support actions such as weather monitoring and remote scientist support. All communications with the expedition were subject to a 20 minute time delay representative of the latency of Mars operations. Live data streams from space weather protection services were also included into the service as part of the operational modelling.

Saber Astronautics also conducted a series of simulated missions, collecting observational data toward further research into global, probabilistic mission performance models. Simulated missions included robot-only scenarios utilising the MSI rover, human-only EVA teams, and human-robot EVA teams utilising the Phantom 2 quadcopter, Miner and Little Blue rovers. Several missions were conducted with both a surface rover and the quadcopter. Robot components were operated both directly and remotely from mission control. EVA personnel (in addition to mission control) consisted of pairs of Expeditioners, who wore simulated space suits loaned from the Victorian Space Science Education Centre (VSSEC) but originally manufactured by MSA to its own design.

Table 2: Key simulated missions conducted at Arkaroola by Saber Astronautics.

# Experiments	Assets	Mission Type
Two	2 astronauts, Little Blue, Phantom 2	Geological sortie
Two	2 astronauts	Maintenance (weather station; rover)
One	2 astronauts, Phantom 2	Geological sortie
Six	2 astronauts, Phantom 2	Exploration
Five	MSI Rover	Exploration

Missions were conducted over a wide range of terrains from smooth dust to extremely rocky creek beds with typical deployment duration of around 60 minutes. Mission goals fell into three main categories: exploration, geological ‘sorties’ or sample collection, and equipment maintenance. Rover-only mission goals included both acquisition of visual data via a video link and surface sampling with a robot arm. The arm was able to collect a range of materials, for example aeolian dust, fine alluvial sand, and alluvial gravel.

A range of specific tasks (Table 2) were attempted in astronaut-only, robot-only and mixed astronaut-robot EVA missions. These included a geologist monitoring non-geologist EVA teams via a video link and directing them over UHF radio to carry out specific geological tasks such as observation and sample collection. Other tasks in exploration EVAs involved multiple deployments and recoveries of the quadcopter. Perhaps the most notable EVA was the servicing of a weather station at one of the observatories behind Arkaroola (Fig. 4).

Data collected during each mission included biometric, GPS, motion and other data from sensors deployed on robots and personnel. In addition, human-to-human communications via UHF radio were collected, as well as local weather conditions via a weather station deployed on site.

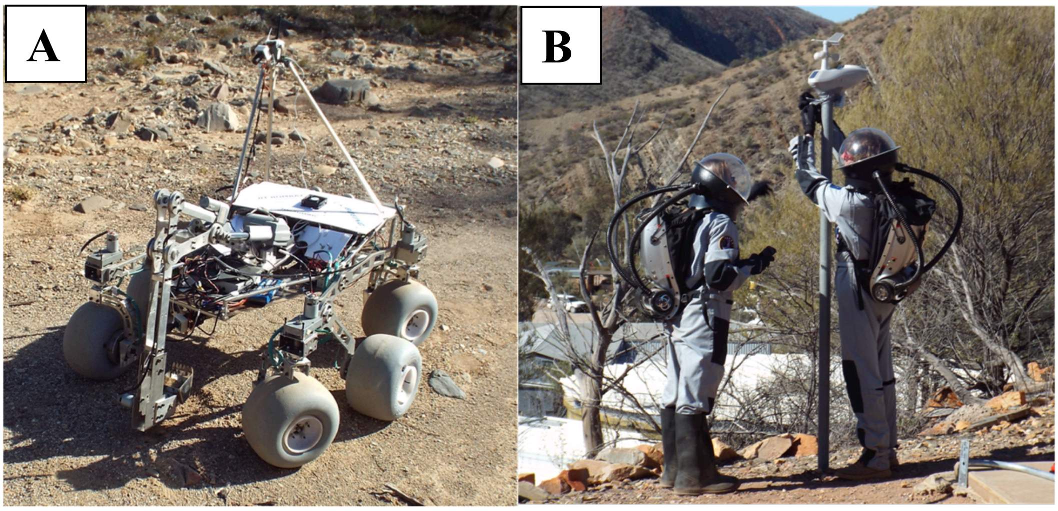


Fig. 4: Remote operations trials. A –MSI rover sampling alluvial gravel under remote control. B –Simulated EVA installing automatic weather station at Arkaroola Old Observatory.

Suit trials

Understanding the constraints of field work while in a space suit is critical when considering the capabilities and limits of crewed exploration on the surface of the Moon, Mars and other accessible Solar System bodies. The impact of a pressure suit was studied during a previous MSA expedition to the Pilbara. The Pilbara study used the University of North Dakota's NDX-1 suit and investigated the ability of geologists and non-geologists to recognize stromatolites despite the physical and temporal limitations of wearing the suit [1].

A follow up investigation using the same methodology was carried out on the Arkaroola Mars Robot Challenge Expedition (Fig. 5). A VSSEC simulated space suit was used. While not pressurized, the suit impaired sensory awareness and mobility to some extent analogous to an actual suit. A number of volunteers, both geologists and non-geologists, were asked to walk along a limestone outcrop and identity features that they considered as possible and probable stromatolites over a period of 20 minutes. Analysis is in progress and it is expected to present results at the 2015 ASRC.

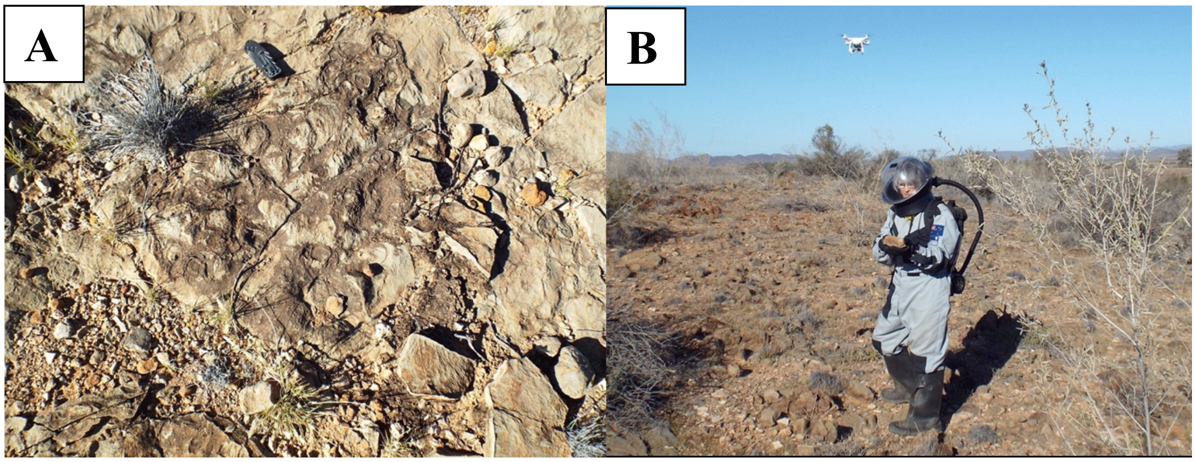


Fig. 5: Suit trials. A – bedding plane exposure of stromatolites of the Trenzona Limestone. B – geologist in VSSEC suit assesses whether a hand specimen contains a stromatolites. Experiment is being recorded by the Phantom 2 quadcopter.

Geobiology

Geobiological research, led by Simon George and Sarah Houlahan on the expedition focused on the testing hypotheses regarding the evidence for ancient biospheres found in Cryogenian (850 to 635 Mya) rocks the Arkaroola area (Fig. 6). The first hypothesis was regarding the biogenicity of possible microfossil structures found in veins within the Tapley Hill Formation which have been postulated to represent evidence of a deep hot biosphere living in basin fluids of Neoproterozoic age [11]. The second hypothesis was whether or not the carbonates of Balacanoona Formation and the contemporaneous siltstones of the Tapley Hill Formation contain organic compounds or biomarkers, specifically 24-*n*-isopropylcholestane [12], that may be indicative of the fossil sponges that some researchers have identified in these rocks [13].

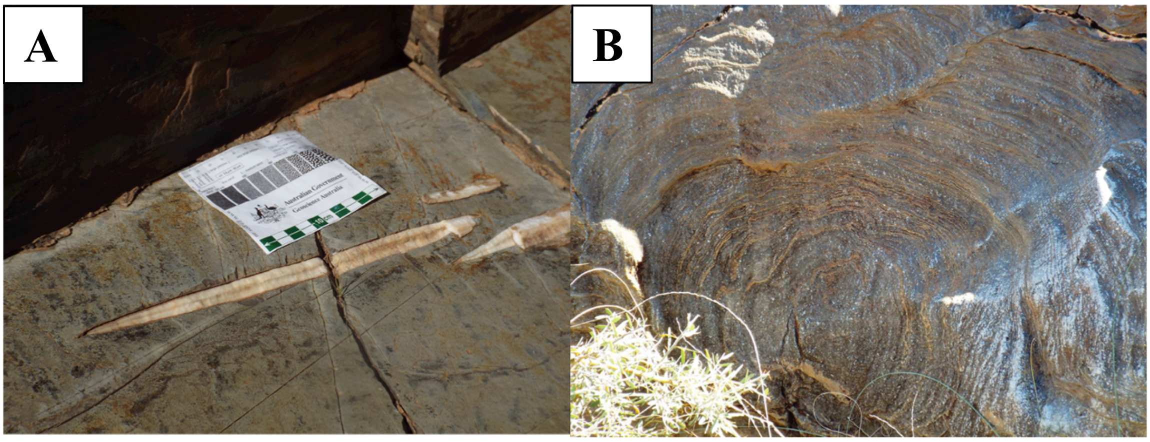


Fig. 6: A -Fibrous calcite veins in the Tapley Hill Formation that host possible microfossils of a deep hot biosphere. B - Stromatolite of the Balcanoona Formation, ~1.5 m high.

In addition to the importance of terrestrial geological of understanding the antiquity of the deep biosphere and Metazoa (complex animals) this research is relevant to astrobiology. The putative microfossils resemble structures found in some martian meteorites e.g. ALH84001 [14], testing their biogenicity refines tools for the interpretation of these martian structures. Likewise biomarkers are potentially the best tool to search for past or present life [15] in martian meteorites, samples being analysed on the surface of Mars or material returned by space missions.

Teacher experience

Three school teachers were among the Expeditioners. They contributed resources, including radios, vehicles, and the Phantom 2 quadcopter, took part in the experiments, gave talks and collected information for their classes. Their participation was part of their professional development as defined by the Australian Professional Standards for Teachers [16], specifically Standard 2, “Know the Content and How to Teach It” and Standard 6, “Engage in Professional Learning”. As a result of their participation the teachers were able to demonstrate that they were able to select and organise geology, Mars and space exploration content in a rich and authentic context following the expedition to Arkaroola (Standard 2) and had taken part in a high quality learning opportunity to improve our teaching practice by enriching our knowledge and understanding of the connection between geology, Mars and space exploration (Standard 6).

Training and outreach

Workshops and field trips

The geological framework of the Arkaroola area is an important part of its analogue value. The diverse rocks and landscapes provide diversity of surfaces for robot tests, counterparts for the types of landscapes and targets studied by planetary robots, and numerous astrobiologically interesting niches. As most of the expeditioners had not been to Arkaroola before and had no background in geology, the first full day was spent on a geological tour of the area, introducing them to the principles of field geology, the geological and geobiological highlights, the challenges facing designers of planetary rovers, and providing spatial orientation.

Part of the conditions of the Australia-India Council grant was the provision of media training for the expeditions. Peter Spinks, a science journalist with Fairfax Media provided a two-day workshop that gave an introduction to the different new media, how it works and its priorities, and how to write releases and articles for it. Expeditioners wrote test releases and received feedback. This training proved most helpful when post-expedition releases were made.

Public engagement

Educating the public about Mars and spacde exploration generally is a core goal of MSA. Consequently public engagement occurred throughout the Expedition. They included planned public lectures at Arkaroola, media coverage from arising from previous alerts, media visits, casual engagement with the public, and post-expedition reporting. A series of evening public lectures were delivered at Arkaroola by several expeditioners, covering topics such as Space Camp (K. Silburn), life on Mars and the history of Life on Earth (S. George), planetary robotics (G. Mann) and human missions to Mars (the lead author). The talks were attended by between 20 and 60 people. Melbourne University students who happened to be on a field camp in the area, attended two of the lectures. A group of students from Mitcham Girl’s High School in

Adelaide visited with the MSI and UNSW students and Saber Astronautics, watching the robots being put through their paces and talking about science and engineering opportunities for women.

There was extensive coverage over the expedition in the Indian media, including a report in the Times of India. Eloise Fuss of ABC Regional South Australia visited Arkaroola during the Expedition and several stories appeared subsequently on the ABC Regional SA web site. Peter Spinks of Fairfax Media wrote a number of excellent online articles and webcasts on TheAge.com.au blog "Science Matters" about the expedition.

Lastly, the presence of the expeditions and their numerous robots, people in simulated spacesuits the advertised public lectures generated a considerable number of casual interest and questions from other people visiting Arkaroola at the time.

Conclusion

The Arkaroola Mars Robot Challenge was a very successful expedition, with all the pre-expedition goals being met. The Expedition showed the value of multi-goal expeditions for both research and training in an analogue setting. Preliminary results are being published [8, 9, 10] and more publications are expected. Mars Society Australia plans future events to follow up on the success of this expedition, including possibly a field robotics student competition similar to the University Rover Challenge run by the US Mars Society in Utah [17].

Acknowledgments

The authors would like to The Arkaroola Mars Robot Challenge expeditioners for their enthusiasm and hard work which made this expedition possible. The expedition was partially funded by CSIRO, the Australia India Council and Saber Astronautics. In-kind support was provided by Arkaroola Wilderness Sanctuary, Graham Mann of Murdoch University, and the Victorian Space Science Education Centre. Lastly I thank the hard work of the editor and the reviewers who made this paper fit for publication.

References

[1] Willson, D., Rask, J.C., George, S.C., deLeon, P., Bonaccorsi, R., Blank, J., Slocombe, J., Silburn, K., Steele, H., Gargarno, M., and McKay, C.P. "The performance of field scientists undertaking observations of early life fossils while in simulated spacesuit", Acta Astronautica Vol. 93, 2014, pp 193–206.

[2] Persaud, R., Rupert-Robles, S., Clarke, J. D. A., Dawson, S, Mann, G., Waldie, J., Piechocinski, S., and Roesch, J. "Expedition One: a Mars Analog Research Station 30-day mission," In Cockell, C. S. (ed.). Martian Expedition Planning. American Astronautical Society Science and Technology Series Vol. 107, 2004, pp 53-88.

[3] Clarke, J. D. A., Persaud, R., Rupert, S., and the Expedition Two crew. "A Multi-Goal Mars

Analogue Expedition (Expedition Two) to the Arkaroola Region, Australia", In Clarke, J. D. A. (ed.), Mars Analog Research, American Astronautical Society Science and Technology Series Vol. 111, 2006, pp 3-16.

[4] Persaud, R. "A systematic approach to investigations at Mars analog research stations." In Cockell, C. S. (ed.). Martian Expedition Planning. American Astronautical Society Science and Technology Series 107, 104-121, 2004.

[5] Mann, G. A., Clarke, J. D. A., Gostin, V. A. "Surveying for Mars Analogue Research Sites in the Central Australian Deserts", Australian Geographical Studies, Vol. 30, No. 1, 2004 116-124.

[6] Clarke, J. D. A., Thomas, M., Norman, M. "The Arkaroola Mars analog region, South Australia." Abstracts of the 34th Lunar and Planetary Science Conference Abstract #1029, 2004.

[7] Thomas, M., Clarke, J. D. A., Gostin, V. A., and Williams, G. E. "The Flinders ranges and surrounds, South Australia: a window into astrobiology and planetary geology", Episodes Vol. 35, No. 1, 2012, pp 226-235.

[8] Mann, G. A., Small, N., Clarke, J. D. A. and Sheh, R. "Standardized Engineering Assessment of Field Robots in a Mars-like Environment", IEEE Robotics & Automation Magazine. Submitted.

[9] Hobbs, S.W., Clarke, J.D.A., Mann, G.M. Field Testing Marsobot, A Mars Society Australia Robotics Project. Proceedings of the 14th Australian Space research Conference, 2014, (this volume).

[10] Kabaran, Y., McArthur, S., Bartlett, I.S., and Kinkaid, N. BLUEsat robotics to Mars: lesson on real world testing for student projects. Abstracts of the 14th Australian Space research Conference, 2014, p 110.

[11] Bons, P. D., Montenari, M., Bakker, R. J., and Elbur, M. "Potential evidence of fossilised Neoproterozoic deep life: SEM observations on calcite veins from Oppaminda Creek, Arkaroola, South Australia", International Journal of Earth Sciences, Vol. 98, No. 2, 2009, pp 327-343.

[12] Love G.D., Grosjean, E., Stalvies, C., Fike, D.A., Grotzinger, J.P., Bradley, A.S., Kelly, A.E., Bhatia, M., Meredith, W., Snape, C.E., Bowring, S.A., Condon, D.J. and Summons, R.E. "Fossil steroids record the appearance of Demospongiae during the Cryogenian period", Nature Vol. 457, 2009, pp 718-722.

[13] Wallace, M.W., Hood, A.v.S., Woon, E.M.S., Hoffmann, H.-H., and Reed, C.P., "Enigmatic chambered structures in Cryogenian reefs: The oldest sponge-grade organisms?" Precambrian Research Vol. 255, 2014, pp 109–123.

[14] McKay, D.S., Gibson, E.K., Thomas-Keprta, K.L., Vali, H., Romanek, C.S., Clemett, S.J., Chillier, X.D.F., Maechling, C.R., Zare, R.N. "Search for Past Life on Mars: Possible Relic Biogenic Activity in Martian Meteorite ALH84001", Science Vol. 273, No. 5277, 1996, pp 924–930.

[15] Ehrenfreund, P., Röling, W.F.M., Thiel, C.S., Quinn, R., Septhon, M.A., Stoker, C., Direito, S., Kotler, J.M., Martins, Z., Orzechowska, G.E., Kidd, R.D., van Sluis, K. and Foing, B.H., "Astrobiology and habitability studies in preparation for future Mars missions: trends from investigating minerals, organics and biota", Int. J. Astrobiology, Vol. 10, 2011, pp 239-253.

[16] NSW Institute of Teachers, 2011 "Australian Professional Standards for Teachers". Address when accessed http://www.nswteachers.nsw.edu.au/DownloadDocument.ashx?DocumentID=1122

[17] Post, M.A., Lee, R. Lee, "Lessons learned from the York University Rover Team (YURT) at the University Rover Challenge 2008–2009', Acta Astronautica Vol. 68, No. 7–8, 2011, pp 1343–1352.

Coseismic Deformation Inferred from DInSAR and Model Results for 2008 Sichuan and 2009 L'Aquila Earthquakes

Youtian Liu, Linlin Ge, and Xiaojing Li

Geoscience and Earth Observing Systems Group (GEOS)
School of Civil and Environmental Engineering, UNSW Australia, Sydney NSW 2052
Australia

Summary: Using DInSAR technique to detect earth surface deformation at sub-centimetre level has promised its application in measuring co-seismic deformation along the fault for the needs of earthquake disaster mitigation. The 2008 Sichuan earthquake and 2009 L'Aquila earthquake are two good samples for such applications. In this paper, multiple pairs of L-band ALOS PALSAR images focusing on the Sichuan earthquake and two pairs of C-band ENVISAT ASAR images covering the L'Aquila earthquake are used to map the co-seismic deformation for each event. The measured deformation caused by the Sichuan earthquake is up to 112 cm in the radar line of sight direction, and about 28 cm for the L'Aquila earthquake. The interferograms generated for these two different events are then utilised to constrain the elastic fault model. Modelling results are compared with the profiles of unwrapped displacements from DInSAR technique. The authors focus on validating the superior ability of DInSAR in mapping co-seismic deformation over large area, and in finding the best-fit model result for the earthquakes.

Keywords: DInSAR, Coseismic Deformation, Okada Model Inversion

Introduction

Ground surveying, seismological measurement, and GPS network are the most common techniques for studying earthquake activities. But all these techniques are time consuming and highly costly for measurements covering large area and it may even be impossible to use some of them at locations difficult to access. To overcome these problems with improved efficiency and accuracy on deformation measurement over large area, geospatial mapping technologies have been developed. In particular, the Interferometric Synthetic Aperture Radar (InSAR) technique has been increasingly used due to its large coverage and the cloud penetration ability of radar wave.

Following an M_w 7.9 devastating earthquake occurred on 12th May 2008 in Sichuan, China, another M_w 6.3 earthquake struck the region of Abruzzo in central Italy on 6th April 2009. The study of ground surface deformation, especially co-seismic deformation, related to the two earthquakes can not only help our understanding of the two events but also assist us further on disaster mitigation for future earthquakes.

Co-seismic deformation is the change of ground surface shortly before and after the earthquake. Comparing to conventional surveying methods, Differential InSAR (DInSAR) technique has several advantages for co-seismic deformation measurement, such as large coverage, dense observations, low cost, and the ability to access remote areas. Comparing with the costly dense GPS network, DInSAR can deliver accurate measurements with much higher spatial density, e.g. 100 or more measurements per square kilometre [1]. These dense DInSAR measurements enable the elastic dislocation model in half-space to achieve better modelling results.

The 2008 Sichuan earthquake occurred along the Longmen Shan (LMS) thrust system which bounds the eastern margin of the Tibetan plateau, and consists of three major faults (namely, Wenchuan Fault, Yinxiu-Beichuan Fault and Guanxian-Anxian Fault) that strike SW-NE. The main shock took place on 12th May 2008 deformed a very large area of about 300 km E-W and 250 km N-S. Previously published studies about this quake using multi-frames ALOS PALSAR (Phase Array L-band SAR of the Japan Aerospace Exploration Agency) imagery can be found by Chini et al., Ge et al., and Hong et al. [2-5].

As for the 2009 L'Aquila earthquake, it is located at a depth of about 9 km, near the 5 km-long Mt. Stabiata fault along Monticchio-Fossa fault, Mt. Bazzano and Paganica faults [6]. These zones depict a surface subsidence of 13 km in length striking at 130° – 140° [7]. The main shock of this quake occurred on 6th April 2009. ENVISAT ASAR (Advanced SAR of the European Space Agency) datasets were utilised by many researchers to study it, e.g. [8-10].

To measure the co-seismic deformation generated by an earthquake, a pair of SAR images, one acquired before and another after the quake, is processed to produce the interferogram. The final displacement map is obtained with the two-pass DInSAR technique using SRTM DEM [11]. In this paper, the ALOS PALSAR L-band data is used for mapping Sichuan earthquake co-seismic deformation, while ENVISAT ASAR C-band data is utilised for measuring L'Aquila earthquake deformation. With the L-band and C-band DInSAR results being analysed for each earthquake, the authors are able to demonstrate the ability of DInSAR technology in co-seismic research.

There are successful applications of using DInSAR in monitoring earthquakes. A very early one was done in 1993 by Massonnet et al. [11] for the 1992 Landers earthquake. Later researchers also utilised this technique in other cases, such as measuring the deformation of a 1998 earthquake near Mt Iwate by Fujiwara et al. [12], estimating the Iranian Bam earthquake in 2004 by Talebian et al. [13] and the discovery of a blind strike-slip fault, and 3-D co-seismic displacement inversion of the 2005 Kashmir earthquake by Wang et al. [14]. The near real-time preliminary mapping result of co-seismic deformation for the 2008 Sichuan earthquake was done by Ge et al. [4]. These DInSAR results all can be used to model a co-seismic deformation respectively and invert its fault parameters to describe an earthquake.

In this paper, the authors present the DInSAR measurements of 2008 Sichuan earthquake and 2009 L'Aquila earthquake with the modelling results inverted from the Okada dislocation elastic model [15]. The ability of two-pass DInSAR with phase information for co-seismic mapping and monitoring thus has been verified. The advantages and limitations of deformation monitoring using phase information from radar imagery also discussed with more details.

DInSAR Data Analysis

To use DInSAR technique to map the co-seismic deformation of a selected seismic zone, we need to process a pair of radar satellite images acquired before and after the earthquake event. The pairs of ALOS PALSAR images for the Sichuan earthquake were collected from the ascending orbit listed in Table 1, which covers the Beichuan-Yinxiu fault (YB fault). In contrast, the L'Aquila earthquake area was mapped with ENVISAT SAR images taken from both ascending and descending orbits. Table 1 also lists the details for these pairs of images.

Table 1 Images for Interferometric Pairs Used

Earthquake	Satellite	Orbit	Interferometric Pair Date (dd/mm/yyyy)	Perpendicular Baseline (m)
Sichuan	ALOS	Ascending (Track 473 R620-630 Track 474 R610-620 Track 475 R613)	17/02/2008 – 19/05/2008 05/03/2008 – 05/06/2008 20/06/2007 – 22/06/2008	245/249 286/292 -42
L'Aquila	ENVISAT	Ascending (Track 401)	23/02/2009 – 04/05/2009	-141
L'Aquila	ENVISAT	Descending (Track 79)	18/03/2009 – 12/04/2009	-534

Both ALOS PALSAR datasets and ENVISAT ASAR datasets are processed with the two-pass DInSAR technique using SARscape software. However, since DInSAR provides deformation in the LOS (Line of Sight) direction, it is a one-dimensional measurement in the three dimensional coordinate system. The orthogonal surface displacement components, U_E, U_N and U_Z, can be related to the measured LOS displacement d_{los} with the Equation (1):

$$[U_E \quad U_N \quad U_Z]\begin{bmatrix}-\sin\theta\cos\alpha \\ \sin\theta\sin\alpha \\ \cos\theta\end{bmatrix} + \delta_{los} = d_{los} \tag{1}$$

Where α is the azimuth angle of the satellite, θ is the radar incidence angle at the scatterer point, and δ_{los} is the measurement error which consists of atmospheric delays of the signal, low coherence of phase, coarse DEM, imprecise satellite orbit and so on [16]. Figure 1 depicts the Equation (1) in ascending situation, and similar works on 3-D deduction can be referred to [3, 5, 9, 17, 18].

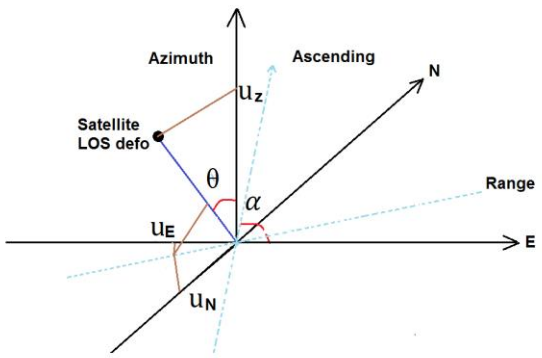


Figure 1: 1-D LOS deformation to 3-D projection at Ascending Measurement

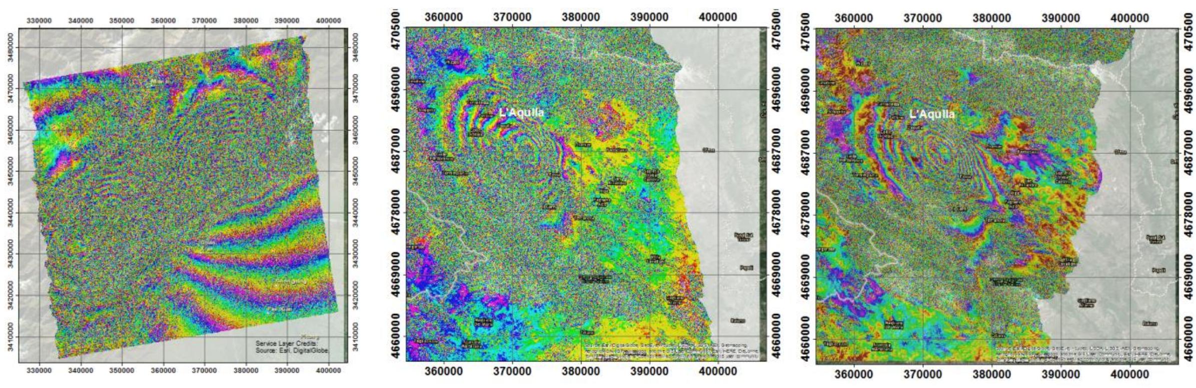
For a raw interferogram, the phase difference Δϕ consists of many factors, and can be described by Equation (2) [19]:

$$\Delta\phi = \phi_{Topo} + \phi_{Defo} + \phi_{Atmo} + \phi_{Orbit} + \phi_{Noise} \tag{2}$$

Where the phase (ϕ_{Topo}) is due to topography contribution, the phase (ϕ_{Defo}) is due to geometric displacement contribution, the phase (ϕ_{Atmo}) is caused by atmospheric disturbances such as ionosphere and humidity, the phase (ϕ_{Orbit}) is caused from orbit estimation error, and the phase (ϕ_{Noise}) is for system noise.

Theoretically, to obtain the deformation of the earth surface, all the other phase information should be removed except the phase caused by geometric displacement. To process these SAR images listed in Table 1, firstly, the SNR (Signal Noise Ratio) of an interferogram shall be improved with multilook. An azimuth to range multilook factor of 5 to 1 for ENVISAT images and 4 to 2 or 1 for ALOS images are utilised. Then, the 90 m SRTM was used as the external DEM for subtracting the topography phase. For removing the orbital error of ENVISAT datasets, the precise orbits files which are distributed by Delft Institute for Earth-Oriented Space Research (DEOS) are utilised. On the other hand, ALOS PALSAR dataset contains its precise orbit data. A further step to eliminate the orbit error was that ground control points were selected at the locations of no deformation or evidence of residual topographic fringes. The atmospheric error can be reduced with GPS measurement or other corrections, while the phase due to noise was reduced by using the Goldstein adaptive filter [20].

Since the phase information obtained from raw interferogram is wrapped, which ranges from (-π, π), the step to rebuild the relative phase data is phase unwrapping. In the case of the 2008 Sichuan earthquake, the seismic event ruptured a wide zone approximately of 300 km E-W by 250 km N-S. To cover the whole coseismic area, multi-frame pairs of ALOS PALSAR imagery could be used [2-4]. However, the pair taken on 20 June 2007 and 22 June 2008 from ascending track 475, frame 613, only covers the epicentre of the Mw 7.9 earthquake, with looking angle around 34.3°. Due to the L-band's ability of penetrating vegetation, using ALOS images to compute the deformation can achieve better coherence result. Fig. 2 (a) shows the differential interferogram of track 475, from which we can see the fringes along YB fault and use the information to visually define the fault area. To compute the displacement, the minimum cost flow algorithm [21] for phase unwrapping has been used to convert phase into displacement. However MCF underestimates the unwrapped phase value close to the fault due to high displacement gradient, mask of the incoherent areas is suggested.



(a) (b) (c)

Figure 2: (a) Interferogram of ALOS PALSAR for Sichuan earthquake (Track 475)

(b) Interferogram from ENVISAT ASAR for 2009 L'Aquila (Track 79)

(c) Interferogram from ENVISAT ASAR for 2009 L'Aquila (Track 401)

For the 2009 L'Aquila quake area, the LOS deformation is from -41 cm to 38 cm in track 79 and -24 to 28 cm in track 401. The difference between the two tracks measurements can be caused by temporal difference and atmospheric impact and so on due to the image pairs acquired at different time. The interferograms of the L'Aquila seismic area are shown in Figure 2 (b) and (c).

Modelling Method

To further analyse the fringe pattern of phase obtained from the interferogram, the unwrapped phase result of displacement map is modelled with a finite dislocation model for earth surface deformation in an elastic and homogeneous half-space [15] to estimate the earthquake focal mechanism.

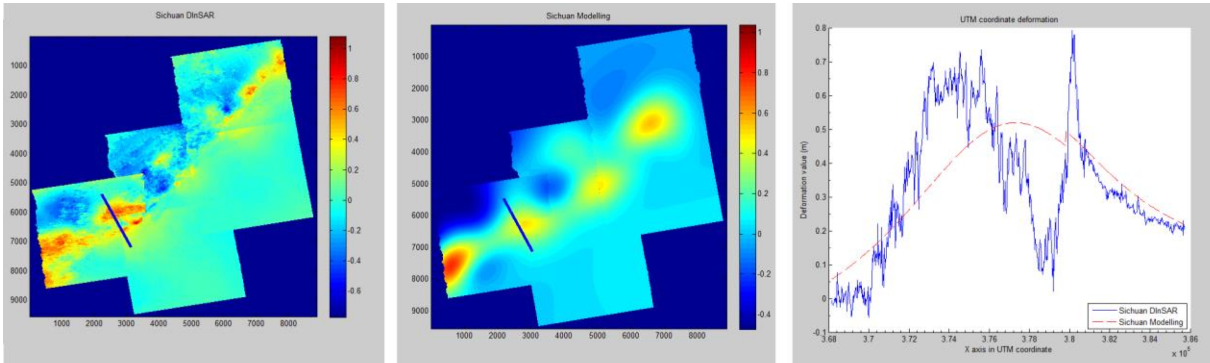
As we can see from the interferogram results, it offers us millions of observations points. To improve the computational efficiency, we subsampled the L'Aquila displacement map at one point for every 500 m × 500 m around the deformed area and one point for every 2000 m × 2000 m for the rest area. In total 9958 points and 7634 points for the track 401 pair and track 79 pair were obtained, respectively. On the other hand, the Sichuan earthquake ruptured a very large area, and also the seismic zone had low coherence for further analysis. Therefore, we only selected one point for every 500 m × 500 m near the fault and one point for every 1000 m × 1000 m to the rest, which yields 43234 points in total.

The subsampled data was then utilised for estimating the parameters to describe the earthquakes. The parameters consist of length and width of the fault, depth of the hypocentre, strike angle and dip angle about the fault, rake angle for slip vector, the amount of slip and dilation of the earthquake. The mentioned eight parameters are then inverted by comparing the forward modelling of the displacement for east, north and up direction. This is a nonlinear inversion problem, thus a nonlinear inversion algorithm [22] assuming uniform slip on the fault plane has been used. The inverted fault parameters are obtained by minimising the misfit between the modelling result and DInSAR result and then finding the best-fit solution. The inversion problem is initialised with global CMT (Central Moment Tensor) result to set the bounds and save computational time. After the nonlinear inversion, the estimation of the slip distribution and fault plane in patches are inverted with linear algorithm.

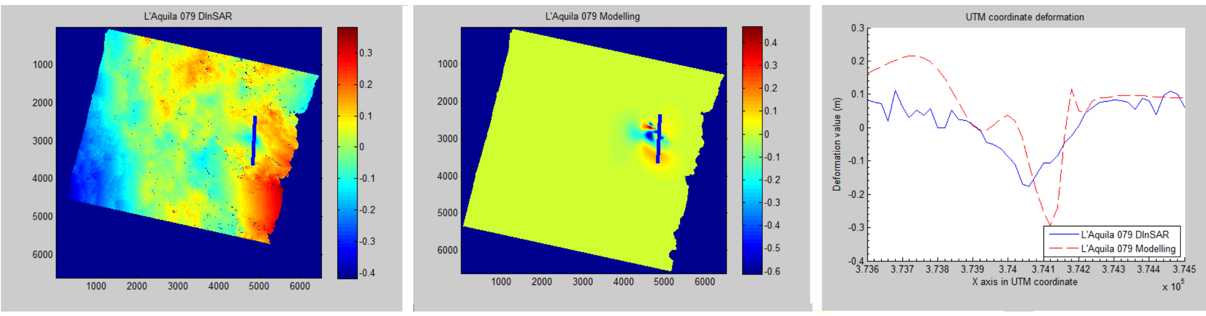
Discussion

Figure 2 shows the interferometric phase change caused by the earthquakes and aftershocks during the period of images collection. After phase unwrapping and geocoding steps, the final displacement maps are generated for both earthquakes (Figure 3). The deformation for the 2008 Sichuan earthquake is measured with the range from 0.9 m to -1.12 m in the LOS (Line of Sight) direction, which is about 0.4 m difference compared with result in [2]. For the 2009 L'Aquila Earthquake generated deformation, the measurement is from -0.24 m to 0.28 m in ascending pair and -0.41 m to 0.38 m in descending pair. The difference between the two pairs' results could be due to the difference of temporal separation and looking geometry. The positive sign for the LOS deformation is defined as the distance between scatterer and satellite is decreased and negative vice versa. Nonlinear and linear inversions would have offset between the DInSAR and the model results. In this work, the offset values are 9.9 cm, 7.5 cm and 2.2 cm for the best fit model for Sichuan case, L'Aquila descending case, and L'Aquila ascending case, respectively. The ramp value has the relationship shown in Equation (3), where the offset is estimated by the nonlinear results. X and Y are the coordinates for each point with the coefficients a and b for the points.

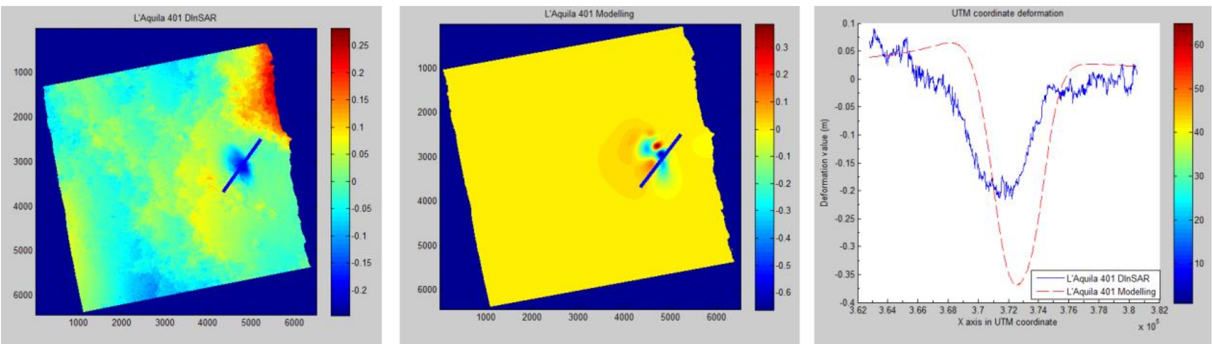
$$\mathrm{Ramp} = \mathrm{offset} + X * a + Y * b \tag{3}$$



(a) 2008 Sichuan earthquake DInSAR and Modelling comparison



(b) 2009 L'Aquila earthquake DInSAR and Modelling comparison (track 79)



(c) 2009 L'Aquila earthquake DInSAR and Modelling comparison (track 401)

Figure 3: DInSAR measurement and Modelling result comparison

The authors took a profile across the co-seismic area observed in each DInSAR displacement image and modelling image to compare the results. Figure 3 shows the subsidence in the seismic zone for the two earthquakes and modelling results correspondingly. Figure 3 (a) shows the mosaicked displacement map and modelling result for Sichuan earthquake, which covers most part of YB fault. The DInSAR result is different from [2], which could due to the use of a different pair of track 475 with a longer temporal baseline, but more clear interferogram. In the L'Aquila case, descending track 79 pair shows more complicated deformation from modelling results, due to a result of the different looking geometry of sensitivity. However, with the DInSAR results we can define the most likely strike angle and check with the modelling results. The coseismic parameters are modelled by minimising the misfit between modelling result and DInSAR result. Compared to the Sichuan earthquake, the L'Aquila earthquake impacted a much smaller area so that one ENVISAT ASAR image can cover the whole seismic area. The modelling results for the two seismic events are listed in Table 2.

Table 2 Nonlinear Results for 2008 Sichuan earthquake and 2009 L'Aquila earthquake

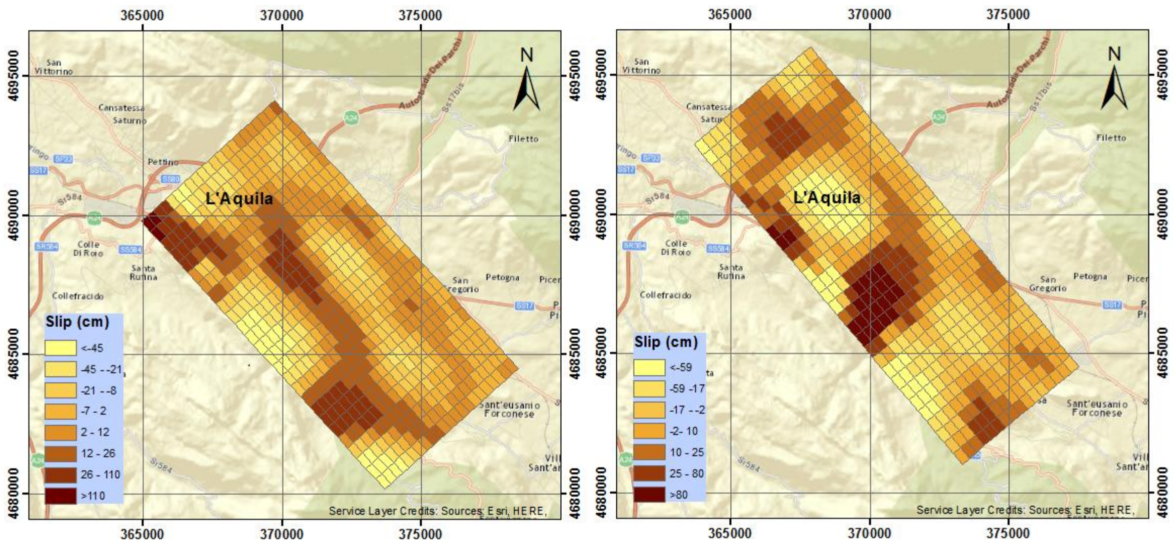
Event	Width (km)	Length (km)	Top Depth (km)	Strike (deg)	Dip (deg)	Rake (deg)	Slip (m)	Open (m)
Sichuan (YB fault)	36	159	1.9	236	38	139	3.16	0
CMT*			12.8	231	35	138		
L'Aquila (Track 79)	10.2	12.1	0.9	137	45	-100	0.42	0
L'Aquila (Track 401)	12.9	13.2	2.6	138	45	-100	0.54	0
CMT*			12.0	120	54	-113		

* *Stands for global Centroid Moment Tensor results [23] and the depth is epicentre*

As for Sichuan earthquake we selected the interferogram covering the 12th May 2008 event epicentral area, a single fault model along the YB fault is utilised for the study. From global CMT results we know there were 9 aftershocks larger than Mw 5.0 occurred before the slave image acquired for the last pair to form the DInSAR. Different temporal baseline DInSAR pairs were mosaicked but had been trying to get close time line for the second image. Taking track 475 as an example, from the wrapped phase interferogram (Figure 2 (a)), we can set the strike angle with a boundary of 200° to 260° with respect to the initial strike angle from global CMT result of 231°. Then, we utilise the interferograms from ALOS PALSAR Fine Beam mode as it covers the YB fault area, choosing HH polarization for better backscatter and coherence. We set lame constant λ and shear modulus μ for 30 GPa assuming the uniform layer under deep earth. The modelling result shows right-lateral trend and a gentler dip-slip angle (15°) with reverse fault. Since the length are shorter as the interferogram covers smaller area compared to the whole YB fault, the moment magnitude estimated is Mw 7.9, which is close to global CMT result. Aftershock investigation and multi-frame interferograms could produce closer result to the CMT result. However, the LOS deformation of Sichuan we measured is much less than the field survey [24]. This is because the L band signal coherence is lost when the surface displacement is too high (i.e. in the near fault field). In other words, the radar phase measurement reaches its saturation when the displacement gradient is higher that half wavelength per pixel, which is the case in the near fault field area. These phenomenon leads to error propagation and underestimation of the actual deformation measurement in the area. We achieved the moment magnitude of the earthquake 5.82×10^{20} N·m (Mw 7.8), which is a little less than USGS moment tensor result of 9×10^{20} N·m (Mw 7.9). This is because the fault did not include the Beichuan-Qiangchuan (BQ fault). A two faults modelling for the whole fault system can be referred to Chini et al. [2].

The cases for the L'Aquila event are different. The fringes are clear for the whole seismic area and it has less area losing coherence than the Sichuan earthquake. Figure 2 (b) shows area covered by both track 79 and track 401 and the relevant fringes. In other words, the ENVISAT ASAR interferograms captured the seismic event with short temporal baseline. The nonlinear inversion results show similar patterns from ascending and descending view for the earthquake, indicating it is a normal fault with slightly right lateral component. The geodetic moments for track 79 and track 401 inversions are 2.15×10^{18} N·m (Mw 6.2) and 4.03×10^{18} N·m (Mw 6.3), respectively, a little smaller than USGS result of 10×10^{18} N·m (Mw 6.3). The rigidity modulus is assumed to be 30 GPa. The inversion results with damping value can be referred to Atzori and Antonioli [25] for detail. Track 401 and track 79 achieved different results because the temporal baseline is different and the looking geometry is different. The deformation shown for the two L'Aquila displacement maps depict reasonable surface change. The modelling results are close to previous studies [8, 9].

The slip distributions for both tracks mapping the 2009 L'Aquila earthquake are inverted by linear inversion. The inverted slip distributions for each track are shown in Figure 4.



(a) (b)

Figure 4: (a) Slip distribution estimated for L'Aquila earthquake in 2009 of track 79. (b) Slip distribution estimated from track 401.

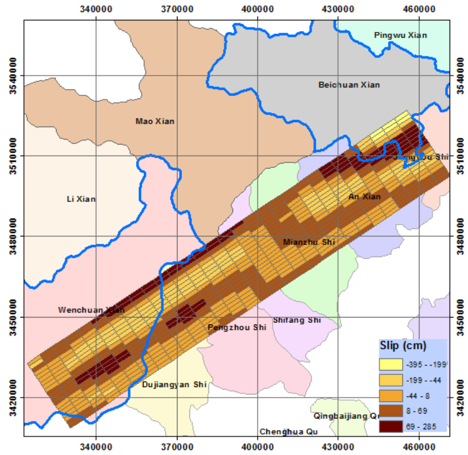


Figure 5: Slip distribution of the Sichuan YB fault

Figure 5 shows the modelling result of Sichuan earthquake, which models part of the faults. The slip distributions are less as the mosaicked displacement map has an area with low coherence area for consideration and not fully cover the second fault along Beichuan to Qingchuan. However, comparing to the 2008 Sichuan earthquake using L-band interferogram, it looks like that L'Aquila earthquake with C-band interferogram gives more clear fringes. This is because C-band has shorter wavelength and more sensitive to deformation. In the case of the Sichuan earthquake, due to the rugged terrain of the region, dense vegetation and rainfall, C-band could be more susceptible to be decorrelation than L-band. This is the reason we choose L-band interferogram for studying the Sichuan earthquake.

Conclusion

In this paper, we have studied both the 12th May 2008 Sichuan earthquake and 6th April 2009 L'Aquila earthquake using DInSAR technique to measure the surface deformation. The ability of using L-band and C-band SAR data to measure deformation is also shown and analysed. The saturation in radar phase measurement should be taken into consideration to constraint the error propagation. An elastic model named Okada in half-space for dislocation is used and the modelling results are in good agreement with DInSAR results, global CMT results and USGS CMT results. Both Interferometry and seismology results show the 2008 Sichuan earthquake is striking along South-West and North-East of inverted fault with right lateral strike slip, and 2009 L'Aquila earthquake is striking along South-East with SW-dipping and right lateral strike slip.

As a further analysis, we subsampled the interferograms into tens of thousands points, which is almost impossible for other conventional methods to achieve in a short period of time. Moreover, phase change is very sensitive to LOS surface change with centimetric accuracy, and these advantages make elastic half-space dislocation modelling based on DInSAR more reliable. We generated DInSAR results and modelling results using different bands (L-band and C-band) for the two seismic events and obtained satisfactory results. This shows the capability of DInSAR to map earthquakes and deliver comprehensive geological information to assess local seismic hazard. There are some differences between DInSAR results and modelling results, and the atmospheric delay might be one of the reasons causing uncertainty in DInSAR measurements. Furthermore, it is possible to deliver quick response and produce near real-time assessment for future earthquakes with new generation of radar satellites.

Acknowledgement

We thank the JAXA and ESA for providing the valuable datasets for Sichuan and L'Aquila area. We are very thankful to the two reviewers for their detailed and constructive suggestions, which helped to improve the quality of this manuscript. This research was supported under Australian Research Council's Discovery funding scheme (project number DP130101694).

References

1. A. Ferretti, A. Monti-Guarnieri, C. Prati, F. Rocca, and D. Massonet, "InSAR Principles-Guidelines for SAR Interferometry Processing and Interpretation," *ESA Training Manual,* 2007, vol. 19.
2. M. Chini, S. Atzori, E. Trasatti, C. Bignami, C. Kyriakopoulos, C. Tolomei*, et al.*, "The May 12, 2008, (Mw 7.9) Sichuan Earthquake (China): Multiframe ALOS-PALSAR DInSAR Analysis of Coseismic Deformation," *Geoscience and Remote Sensing Letters, IEEE,* 2010, vol. 7, pp. 266-270.
3. S.-y. Hong, M. Liu, Z.-r. Liu, Y.-q. Dai, X.-j. Shan, X.-h. Shen*, et al.*, "Precision evaluation and characteristics analysis of the coseismic deformation field of the 12th may 2008 Wenchuan Ms8.0 earthquake," in *Synthetic Aperture Radar, 2009. APSAR 2009. 2nd Asian-Pacific Conference*, 2009, pp. 148-154.
4. L. Ge, K. Zhang, A. Ng, Y. Dong, H.-c. Chang, and C. Rizos, "Preliminary Results of Satellite Radar Differential Interferometry for the Co-seismic Deformation of the 12 May 2008 Ms8.0 Wenchuan Earthquake," *Geographic Information Sciences,* 2008, vol. 14, pp. 12-19.
5. M. De Michele, D. Raucoules, J. De Sigoyer, M. Pubellier, and N. Chamot-Rooke, "Three-dimensional surface displacement of the 2008 May 12 Sichuan earthquake (China) derived from Synthetic Aperture Radar: evidence for rupture on a blind thrust," *Geophysical Journal International,* 2010, vol. 183, pp. 1097-1103.
6. E. W. Group, "Evidence for surface rupture associated with the Mw 6.3 L'Aquila earthquake sequence of April 2009 (central Italy)," *Terra Nova,* 2010, vol. 22, pp. 43-51.
7. P. Boncio, A. Pizzi, F. Brozzetti, G. Pomposo, G. Lavecchia, D. Di Naccio*, et al.*, "Coseismic ground deformation of the 6 April 2009 L'Aquila earthquake (central Italy, Mw6.3)," *Geophysical Research Letters,* 2010, vol. 37, p. L06308.
8. R. J. Walters, J. R. Elliott, N. D'Agostino, P. C. England, I. Hunstad, J. A. Jackson*, et al.*, "The 2009 L'Aquila earthquake (central Italy): A source mechanism and implications for seismic hazard," *Geophysical Research Letters,* 2009, vol. 36, p. L17312.
9. S. Atzori, I. Hunstad, M. Chini, S. Salvi, C. Tolomei, C. Bignami*, et al.*, "Finite fault inversion of DInSAR coseismic displacement of the 2009 L'Aquila earthquake (central Italy)," *Geophysical Research Letters,* 2009, vol. 36, p. L15305.

10. Z. Wang, J. Zhang, G. Huang, and T. e. Chen, "Monitoring the coseismic deformation of L'Aquila earthquake using D-InSAR technique," in *Advanced Computer Control (ICACC), 2010 2nd International Conference on*, 2010, pp. 521-525.
11. D. Massonnet, M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigl*, et al.*, "The displacement field of the Landers earthquake mapped by radar interferometry," *Nature,* 1993, vol. 364, pp. 138-142.
12. S. Fujiwara, T. Nishimura, M. Murakami, H. Nakagawa, M. Tobita, and P. A. Rosen, "2.5-D surface deformation of M6.1 earthquake near Mt Iwate detected by SAR interferometry," *Geophysical Research Letters,* 2000, vol. 27, pp. 2049-2052.
13. M. Talebian, E. J. Fielding, G. J. Funning, M. Ghorashi, J. Jackson, H. Nazari*, et al.*, "The 2003 Bam (Iran) earthquake: Rupture of a blind strike-slip fault," *Geophysical Research Letters,* 2004, vol. 31, p. L11611.
14. H. Wang, L. Ge, C. Xu, and Z. Du, "3-D coseismic displacement field of the 2005 Kashmir earthquake inferred from satellite radar imagery," *Earth Planets and Space,* 2007, vol. 59, p. 343.
15. Y. Okada, "Surface deformation due to shear and tensile faults in a half-space," *Bulletin of the seismological society of America,* 1985, vol. 75, pp. 1135-1154.
16. Y. Fialko, M. Simons, and D. Agnew, "The complete (3-D) surface displacement field in the epicentral area of the 1999 MW7.1 Hector Mine Earthquake, California, from space geodetic observations," *Geophysical Research Letters,* 2001, vol. 28, pp. 3063-3066.
17. A. H.-M. Ng, L. Ge, K. Zhang, H.-C. Chang, X. Li, C. Rizos*, et al.*, "Deformation mapping in three dimensions for underground mining using InSAR – Southern highland coalfield in New South Wales, Australia," *International Journal of Remote Sensing,* 2011, vol. 32, pp. 7227-7256.
18. T. J. Wright, Z. Lu, and C. Wicks, "Source model for the Mw 6.7, 23 October 2002, Nenana Mountain Earthquake (Alaska) from InSAR," *Geophysical Research Letters,* 2003, vol. 30, p. 1974.
19. B. M. Kampes, *Radar interferometry: persistent scatterer technique* vol. 12: Springer, 2006.
20. R. M. Goldstein and C. L. Werner, "Radar interferogram filtering for geophysical applications," *Geophysical Research Letters,* 1998, vol. 25, pp. 4035-4038.
21. M. Costantini, "A novel phase unwrapping method based on network programming," *Geoscience and Remote Sensing, IEEE Transactions on,* 1998, vol. 36, pp. 813-821.
22. D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial & Applied Mathematics,* 1963, vol. 11, pp. 431-441.
23. G. Ekström, M. Nettles, and A. M. Dziewoński, "The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes," *Physics of the Earth and Planetary Interiors,* 2012, vol. 200–201, pp. 1-9.
24. Y. Li, R. Zhou, A. Densmore, L. Yan, N. Richardson, S. Dong*, et al.*, "Surface rupture and deformation of the Yingxiu-Beichuan fault by the Wenchuan earthquake," *Acta Geologica Sinica,* 2008, vol. 82, pp. 1688-1706.
25. S. Atzori and A. Antonioli, "Optimal fault resolution in geodetic inversion of coseismic data," *Geophysical Journal International,* 2011, vol. 185, pp. 529-538.

Building Footprint Extraction from Airborne LiDAR unite Landsat Data

Xiaofeng. Chang [1], Linlin. Ge [2] and Xiaojing. Li[3]

[1] *Geoscience and Earth Observing Systems Group,*
School of Civil and Environmental Engineering,
UNSW Australia, Kensington, NSW 2052, Australia

Summary:
This paper aims to present an approach for a fast, efficient and low cost algorithm for extracting building footprints from airborne Light Detection and Ranging (LiDAR) and Landsat data. LiDAR provides an accurate and efficient remote sensing method to measure ground features, which can generate a high resolution irregularly-spaced point clouds map in a 3-dimensional space. To identify building segments from LiDAR data, the common method is to separate the ground firstly and then classify buildings from the non-ground features, which is based on an accurate digital elevation model (DEM). The non-ground features mainly represent buildings, vegetation and cars. It is hard to distinguish buildings and dense vegetation among the non-ground features. It is possible to determine the density of green on a path of land by using Normalised Difference Vegetation Index (NDVI). The key objective of the research is to develop a cost efficient method to classify dense vegetation from buildings by introducing the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper plus (ETM+) data. NDVI images can be thus generated from multiple Landsat images a couple of month before and after (if applicable) the LiDAR data. Using NDVI map improves the Landsat spectral unmixing map to generate a more accurate abundance map. The classification map, which is in around 2-3 meter spatial resolution, is generated by using the sub-pixel mapping technique to help LiDAR vegetation classification.

Keywords: LiDAR, Landset, footprint, NDVI, statistic, spectral umixing, sub-pixel mapping, Classification

Introduction

Airborne light detection and ranging (LiDAR) provides an accurate and efficient remote sensing method to measure ground objects, which can generate a high resolution irregularly-spaced point clouds map in a 3-dimentional space. To identify building segments from LiDAR data, the common method is to separate the ground first and then classify building from the non-ground features such as buildings, cars and vegetation. Those features are identified based on the LiDAR generated Digital Terrain Model (DTM)[1]. The classified and extracted data from LiDAR can be applied in many applications. For example the digital terrain model can be used to estimate the cost of road construction, or to build a hydro-geological model to estimate flooding areas. The extracted vegetation can be used for environmental protection and monitoring purpose [2]. Building footprints, which outline the buildings, play an important role in wide range of applications in civil and environmental engineering. These can be used to estimate energy demands, urban development and road traffic designs [3].

There are two ways to extract buildings from LiDAR data. One is separating the ground, trees, buildings and other ground features individually. The more popular way approach is separating the ground first and then identifying other features [4]. There are many scenarios have been proposed for the building footprint extraction. Alharthy and Bethel present a heuristic filter to separate buildings and trees [5]. Trees can be identified by comparing the elevation difference between the digital surface model and the digital terrain model, which are generated from the LiDAR first and last return data respectively. It is because that the laser beam can penetrate vegetation not buildings. However, the large and dense trees cannot be classified as the tree edge has large elevation difference as well. Zhang presented a region-grow algorithm based on the plane-fitting technique to identify non-ground features. The buildings and vegetation are separated based on the surface roughness [4]. However, it is still difficult to distinguish large dense vegetation and small isolated buildings.

The critical step of building footprint extraction is to distinguish building and dense vegetation among the non-ground features [4]. It is possible to determine the density of green on a path of land by using Normalized Difference Vegetation Index (NDVI). The key objective of the research is to develop a cost efficient method to classify dense vegetation from buildings by introducing the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) data. Landsat images, which are frequently updated and even allow extracting historical information on urban vegetation [6], could be used to do vegetation statistical analysis in the area of interest.

The primary objective of this research is to extract building foot from LiDAR data with auxiliary optical satellite images from Landsat. This research aims to improve current methods of extraction building footprints from exclusive LiDAR cloud point data source. To overcome the challenge of distinguishing large and dense vegetation from buildings, a free and easily accessible data source, Landsat satellite image, is brought in. Abundance maps for end-members such as bare soil, buildings, vegetation and roads can be generated, by using spectral unmixing technique on Landsat images. The NDVI values of the area of interest calculated from Landsat images, can be used to refine the vegetation proportion value in the abundance map, thus the vegetation classification accuracy will be improved. Sub-pixel mapping technique can be used on the abundance maps to generate a finer scale classification map, thus building will be better extracted.

Current Progress

By doing segmentation on the slope raster image, most of the building outlines can be identified, however the result is not very well. The algorithm assumes that the site topography is relatively flat, which results in the object misclassification. Objects that with low slopes and pitches will correspond to low-value pixels in a slope image; other source data need to be introduced to help segmentation and classification; a DEM can be added instead of using the ground value assumption.

Pre-processing

The original point cloud data need to be filtered and interpolated before processing, as the original LiDAR raw data are irregularly-spaced points, each point is assigned to a cell in x and y coordinates. If more than one point falls into the same cell, the lowest elevation value is selected as the value in the cell. If there are no data in to cell, the nearest neighbor algorithm is used to interpolate the cell value. If there is a big area without value then the area is considered as water. The DSM raster image and slope image shown below is generated based on the original data. The slope image distinguishes building and tree edges clearly, which also contains the unique classification information.

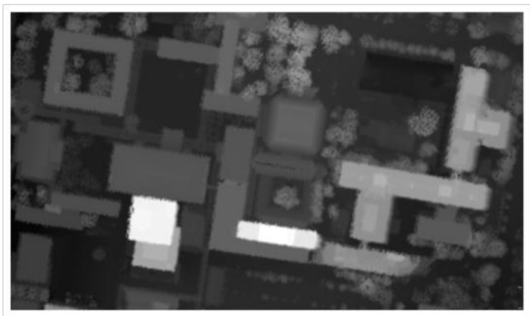


Figure 1 Digital Surface model

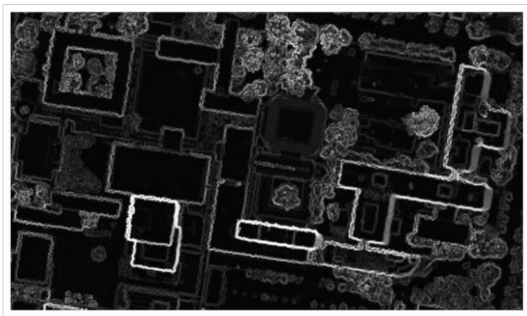


Figure 2 The slope image

Segmentation

The purpose of contrast split segmentation is to create target objects on the slope layer. The objects are classified into 'dark' and 'bright' groups based on the slope value. The purpose of using multi-resolution segmentation around buildings on the aerial optical image layer is trying to generate small segments to apply region grow algorithm to have a better building outline.

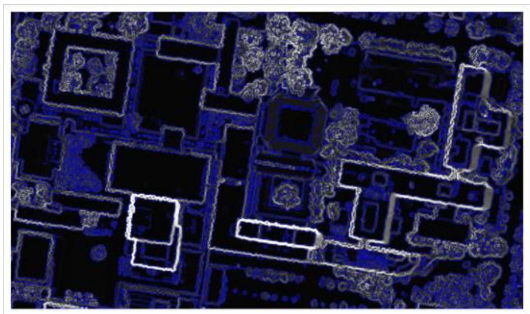


Figure 3: Contrast -split segmentation

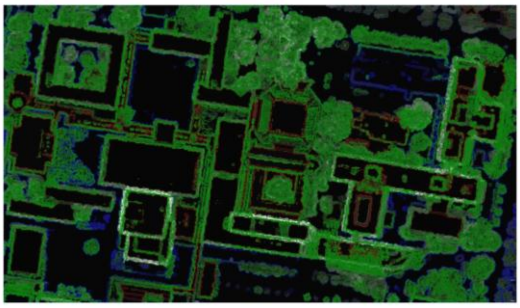


Figure 4: Multi-resolution segmentation

Classification

Firstly, we classify the ground by matching minimum pixel values in the object to the ground, which is calculated by using the quartile statistic method. This scenario classifies building as ground where the slope is not relatively small by mistakes. Then, we further classify and refine buildings by analyzing elevation differences from its neighborhood, taking into account area features. The building outline is modified by observing the spectral similarity and implementing the building region growing. The last step is filling small holes on roofs, which are surrounded and enclosed by buildings.

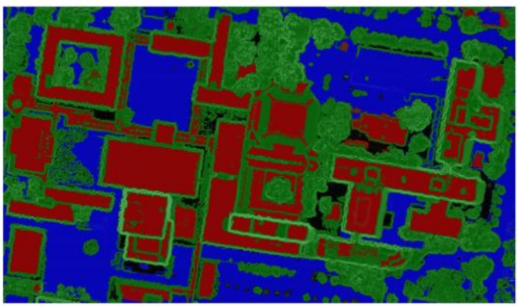


Figure 5: Assigned buildings

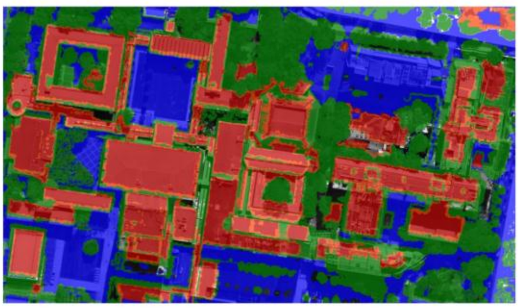


Figure 6: Building region growing

Result

The current result in Figure 7 shows that most of the buildings are well recognized. However, there are blemish on two buildings. The first one is due to glass roofs, which allow most of the LiDAR signal penetrating down to the building floor, so this building is not clearly distinguished. For the second one, vegetation areas surrounded by buildings are misclassified, when using the scenario to fill the unclassified small holes on the roof.

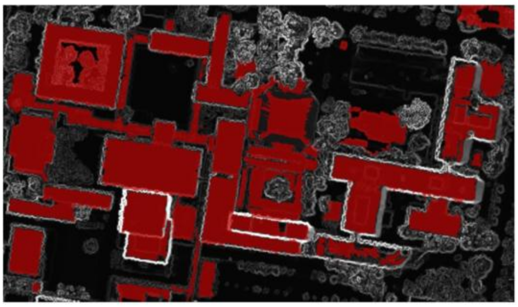


Figure 7: Current experiment result

The ground features, which have low slopes and pitches, will correspond to low-value pixels in a slope image. Figure 8 shows that a low slope building which most area of roof (red) is classified as vegetation (green). To improve this scenario, other data source could be introduced to help segmentation and classification. Using a Digital Elevation Model (DEM) instead of using the ground value assumption.

Figure 8: Side view, top view and classification of low slop building

Further works

As the misclassification occurs by using the statistic value to classify ground, an accurate Digital Terrain Model (DTM) by using a progressive morphological filter is necessary and helpful for the classification. By comparing each pixel elevation value on the Digital Surface Model (DSM) with a reasonable threshold, it is easy to separate bare ground from other features. Moreover it is also a good reference for filtering objects with a relatively low height, such as low bush, cars and etc.

Normalized Difference Vegetation Index (NDVI) could also be introduced to classify vegetation areas. However, this index number depends on the season and species of trees. Landsat program has been collecting images since 1972 and its extensive archive provides multiple images for the area of interest which were captured similar time period of the LiDAR image. By doing the statistical analysis on multiple NDVI maps and comparing with the high resolution aerial photo, a weighted map could be generated from the NDVI maps. Then the weighted map could be used to adjust the abundance map, which shows the proportion of vegetation in each pixel. The number of returns in each region is used to classify objects on their penetration features. The non-penetrate objects such as bare ground, buildings only have one signal return to the sensor. By contrast, the vegetation areas generally produce multiple returns from 2 to 5, dependent on the density of the vegetation.

To overcome the current methodology weakness on distinguishing dense vegetation and buildings, the following method is developed.

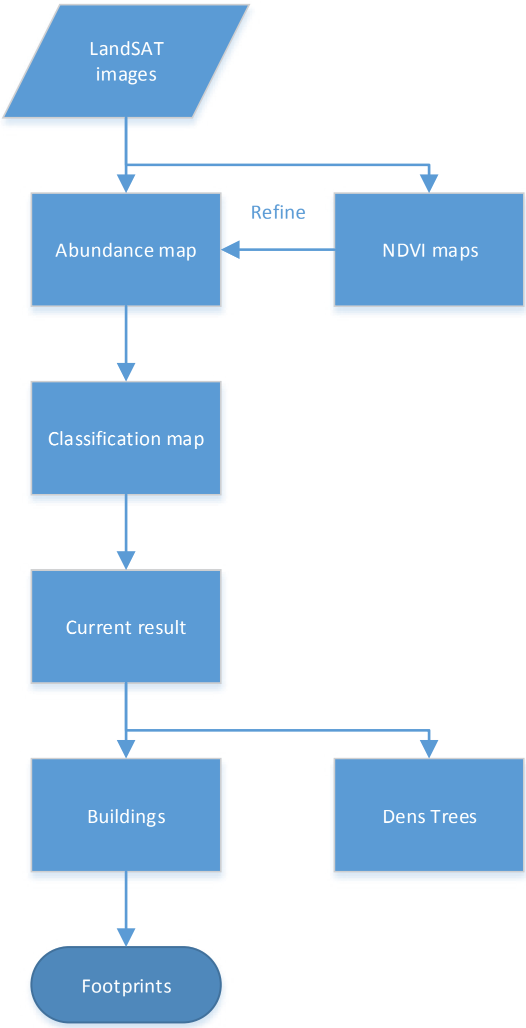


Figure 9: The flowchart of developed method to overcome the current weakness

Spectral unmixing

Unlike the active LiDAR remote sensing, the Landsat images were captured passively. The TM and ETM+ sensors collect solar radiation reflection from the ground features in different bands. Different ground features have different reflectance respect to the wavelength. Figure 10 illustrates the linear mixing where the incident solar radiation is reflected from surface through a single bounce and the surface consists of distinct endmembers [7]

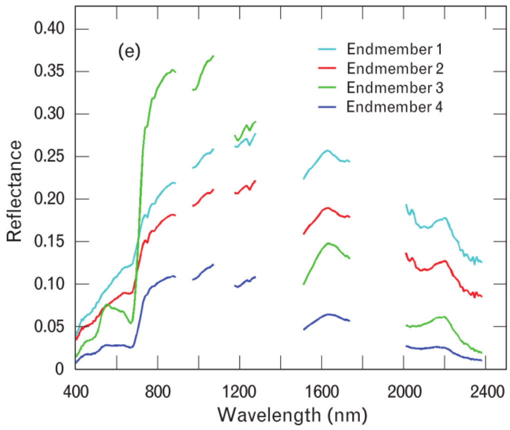


Figure 10: Endmember reflectance spectra derived using geometric endmember determination

The spectral unmixing technique can be used to generate the abundance map for endmembers[8]. Figure 11 shows the abundance map for four different ground features, where endmember 1 is bare soil and road, endmember 2 is associated with grass and low bush and endmember 3 and 4 correspond to two varieties of trees. This technique can be used on the area of interest in Landsat images.

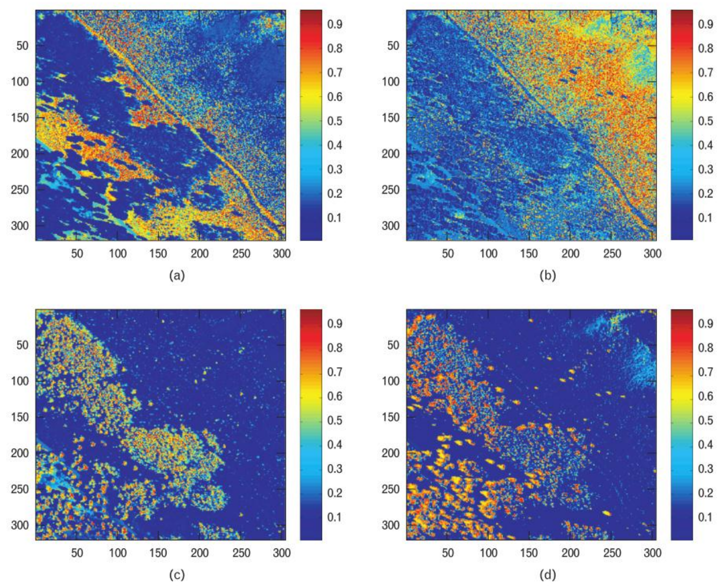


Figure 11: Abundance maps for four endmembers

The result abundance map gives the proportion of each ground feature in the fraction of a pixel size, which is 30 meters. The Landsat images are introduced to improve the classification of dense vegetation and buildings, so the area of interest only concerns dense vegetation, bare soil, and buildings. The rest ground features such as small bushes, cars and rivers are all identified or filtered out in the previous steps. As the first step is to filter out bare ground by using progressive morphology filter, the ground abundance map can be refined by calculating the LiDAR digital terrain model.

Abundance map refinement by NDVI

To classify vegetation from buildings, a Normalized Difference Vegetation Index map is calculated by using NIR and Red Band from Landsat image. Landsat images which are around one month before or after the LiDAR acquire time are selected. One image is selected as master image and co-registration while the rest are slave images. Then the calibration is used to remove the radiometric reflectance. Meanwhile the Landsat master image and LiDAR image co-registration in same coordinate system need to be done as well. Multiple Landsat images before and after LiDAR capture time can be used to generate vegetation weight based on statistics, which is used to refine the abundance map. The refined abundance map is at the medium resolution, which is 30 meters. The abundance map can be used to generate a classification map.

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

Classification map

The refined abundance map can be used to generate a classification map which is spatial resolution improved. Figure 12 shows the procedure of using the sub-pixel mapping technique to generate the hard classification fine scale map.[9]

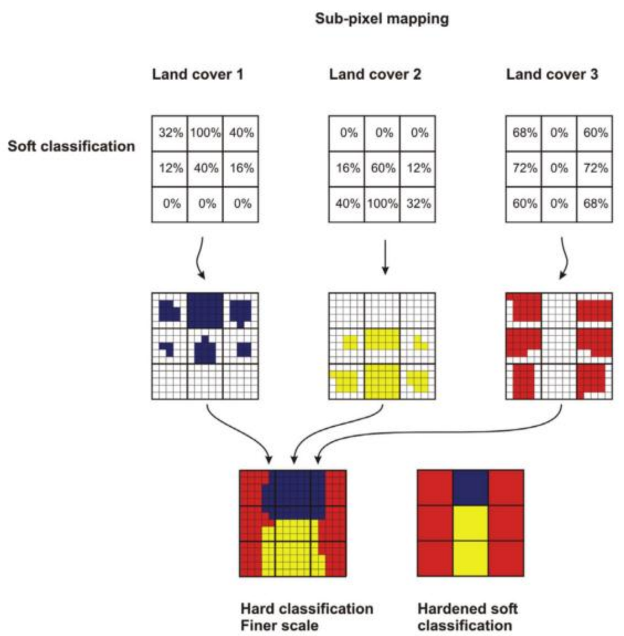


Figure 12: Hard classification with finer scale

Conclusions and perspective

From the current experiment result, it can be concluded that objects that with low slopes and pitches will correspond to low-value pixels in a slope image. Therefore, other source data need to be introduced to help segmentation and classification. To overcome this, this research implements the spectral unmixing and the subpixel mapping technique on multiple Landsate images for a finer resolution classification map, which helps LiDAR distinguish dense vegetation and building. Better results on extracting buildings from LiDAR data can be achieved to meet the general purpose of building footprints extraction, as the building footprint, which outlines buildings, plays an important role in a wide range of applications in the civil and environmental engineering.

References

1. Rottensteiner, F. and J. Jansa, *Automatic extraction of buildings from LIDAR data and aerial images*. International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 2002. **34**(4): p. 569-574.
2. Forlani, G., et al., *Complete classification of raw LIDAR data and 3D reconstruction of buildings*. Pattern Analysis and Applications, 2006. **8**(4): p. 357-374.
3. Jensen, J.R., *Remote Sensing of the Environment: An Earth Resource Perspective 2/e*. 2009: Pearson Education India.
4. Keqi, Z., et al., *A progressive morphological filter for removing nonground measurements from airborne LIDAR data*. Geoscience and Remote Sensing, IEEE Transactions on, 2003. **41**(4): p. 872-882.
5. Alharthy, A. and J. Bethel, *Heuristic filtering and 3D feature extraction from LIDAR data*. International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 2002. **34**(3/A): p. 29-34.
6. Van de Voorde, T., J. Vlaeminck, and F. Canters, *Comparing different approaches for mapping urban vegetation cover from Landsat ETM+ data: a case study on Brussels*. Sensors, 2008. **8**(6): p. 3880-3902.
7. Keshava, N. and J.F. Mustard, *Spectral unmixing*. Signal Processing Magazine, IEEE, 2002. **19**(1): p. 44-57.
8. Cuartero, A., A.M. Felicisimo, and F.J. Ariza, *Accuracy, reliability, and depuration of SPOT HRV and Terra ASTER digital elevation models*. Geoscience and Remote Sensing, IEEE Transactions on, 2005. **43**(2): p. 404-407.
9. Mertens, K., *Sub-pixel mapping: A comparison of techniques*. 2008, Ghent University.

Remote Sensing Image Classification of Environmental Hazards Using Information from Multiple Perspectives

Liyuan Li, Linlin Ge, and Xiaojing Li

Geoscience and Earth Observing Systems Group (GEOS), School of Civil and Environmental Engineering, UNSW Australia, Sydney NSW 2052, Australia

Summary: Environmental hazards cause tremendous damage around the world every year. As one of the typical approaches to analyse environmental hazards, automatic image classification using remote sensing data has been broadly applied. Nonetheless, the classification problem can be complicated and the quality of classification results can be affected. This research introduces the complexity of environmental hazard images, and develops a new image classification method for further discussion. Experiments on three different types of hazards (flood, bushfire and oil spill) are demonstrated. The experiments show that the proposed method is able to generate high quality mapping results for the three cases. To be more specific, noise-like misclassification spots are reduced and the boundaries of regions are better preserved. Hence these results could be preferable when they are converted into polygons. While the study on various cases are still necessary, the proposed method shows a potential to provide better support for the emergency response to some environmental hazards.

Keywords: Image classification, environmental hazards, remote sensing, emergency response

I. Introduction

The 2009 Black Saturday Bushfires in Victoria resulted in AU$4.4 billion damage and 173 fatalities. In total 7,562 people were displaced and 1,100,000 acres of land were burnt. Apart from bushfires, however, environmental hazards such as flood and oil spill also cause tremendous damage in Australia. Fig. 1 illustrates some major disasters in Australia in recent years.

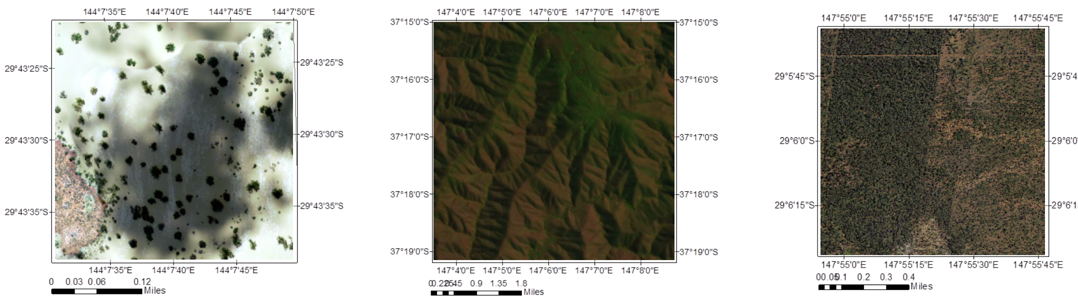
a) b) c)

Fig. 1 Some major disasters in Australia. a) The 2009 Black Saturday Bushfires[1], b) the 2010-2011 Queensland floods[2], c) the 2009 Montara oil spill[3].

Since it is difficult to make accurate prediction on environmental hazards, quick

[1] Available at: http://www.blacksaturdaybushfires.com.au/
[2] Available at: http://en.wikipedia.org/wiki/2010%E2%80%9311_Queensland_floods
[3] Available at: http://www.news.com.au/national/photos-e6frfkp9-1225960174416?page=2

assessment of the hazards is essential to enable prompt response and mitigate their effects. The extent of hazards can be assessed with ground survey. This approach often takes a considerable period of time to collect data. An alternative solution is to collect remote sensing data and run automatic analysis process on computers. Comparing with ground survey, the remote sensing approach usually costs less, requires less human involvement and takes less time to acquire intelligence [1]. Moreover, remote sensing data allow indirect access to remote and dangerous areas, cover vast areas, and provide dense samples. The accuracy of image classification results, however, is restricted by the resolution of images to some degree. When it is compared with the accuracy of total stations (could be millimetres or higher [2]), for instance, it could be lower. The classification quality can be affected by some characteristics of the images (such as the variation of illumination conditions, the variation of terrain, and the non-homogeneity within each class, refer to Fig. 2) as well. The resulting classification maps usually have some noticeable artefact.



) b) c)

Fig. 2 Some characteristics of the images that affect automatic classification results. a) Possible cloud shadow in the aerial image makes the illumination of flood area different; b) the variation of terrain: due to position of the sun, left sides of the mountains appear darker; c) the non-homogeneity within each class: trees in the left side show different patterns from the right side.

In order to deliver better results, some researches have proposed to use environmental indices for classification. In [3,4], the authors used Normalised Difference Vegetation Index (NDVI) and Normalised Burn Ratio (NBR) to map the extent of bushfire. Although the indices can depict bushfire clearer than individual bands, the classification results still contain some noise-like misclassification spots and can be further improved. Another example is in [5], Normalised Difference Water Index (NDWI) had been applied for the mapping of flood. The index was shown to be more effective than density-slicing approach and tasselled cap transformation. In general, however, environmental indices are effective for limited types of objects. Their effectiveness varies from case to case, and sometimes may not be good enough to render the objects (e.g., [6] showed that NBR can perform poorly in rendering bushfires).

These approaches are pixel-wise image classification based on spectral features. They are limited in two aspects. First, it is sometimes difficult for them to distinguish different classes when only a few bands are provided. An example is in the mapping of oil spill, where single polarisation radar images (typically C band) are usually applied. In order to distinguish oil slicks and look-alikes (which can be complicated in oil spill mapping [7]) using a few radar bands, methods have been developed using texture (e.g., statistics features [8,9], and wavelet / Gabor filtered bands [10]). Second, colour (or broadly speaking, spectrum) is only part of the useful information contains in a remote sensing dataset. Texture, region [11,12], shape [13] and even phase [14] could also be the important features that define the difference between classes. Some multiple features based image classification methods taking into account these features have been developed [15,16]. They usually performed better (than methods based on single type of features) in complicated cases, and have been applied to the classification of urban areas, vegetation, etc.

This research develops a new image classification method for the mapping of environmental hazard, which integrate spectral, textural and regional features to improve the classification quality. And the applications on different environmental hazard cases are studied. This article is structured as following. In Section I, research background, related studies on the mapping of environmental hazard and the aim of this research are introduced. Then the features we applied for classification and the details of the new classification method are revealed in Sections II and III. In order to validate the new method, hazards and the corresponding experiments (i.e., mapping of flood, bushfire and oil spill extent) are discussed in Sections IV and V. Finally, Section VI concludes this research effort and discussed about the future work.

III. Integrating Features from Different Perspectives

Image classification consists of two major steps: feature extraction and classification. Owing to the variety of remote sensing data (visible imagery, infrared imagery, radar imagery, etc.) and the variety of earth surface, the useful information for image classification can be very diverse. In order to maintain as much useful information as possible for analysis, we proposed to make use of the information from three different perspectives, namely, spectrum, texture and region.

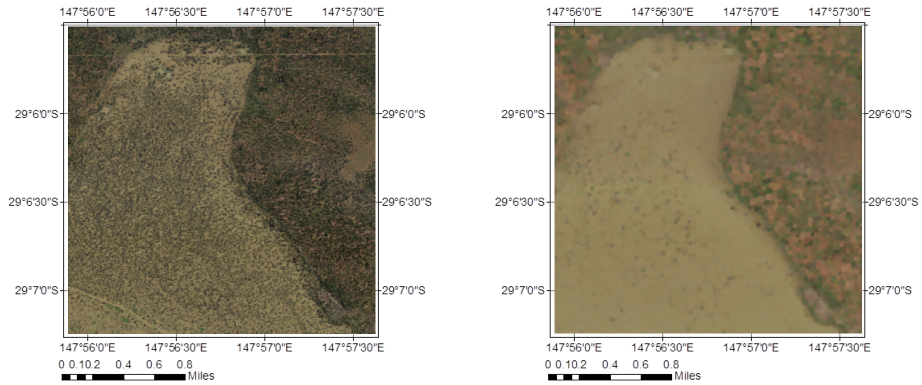
1) Spectral features

In many cases, data from multiple bands and sensors are available. Since it is usually ineffective to use all the bands for classification (In [17], the authors showed that it is not the more input features the better – classification accuracy can be affected by Hughes phenomenon), a proper scheme need to be developed to reduce the volume of input features while maintaining most of the useful spectral information. To meet the requirements, different environmental indices are used to depict the most significant features (e.g., for bushfire, NDVI and NBR will be calculated). Then, all input bands are normalized by their variances, before Principal Component Analysis (PCA) is applied to compress the majority of spectral information into a few principal bands. That is, principal components and environmental indices are used as spectral features. In some cases, however, the situation can be more complicated. Fig. 3 shows a true colour image of flood area. Trees that are half submerged by flood make the image classification complicated (i.e., the spectral information of flood are partially "hidden" by the trees). If spectral features were extracted directly from the images, the half submerged trees could be misclassified as non-flooded. To avoid this problem, these trees are regarded as impulsive noise and a pair of mathematical morphology operators has been used for the noise suppression (this filtering process is done in MATLAB). The operators are opening and closing, which are illustrated below.

$$(f \ominus b) = \min\{f(x+x', y+y') - b(x', y') \mid (x', y') \in D_b\}, \tag{1}$$

$$(f \oplus b) = \max\{f(x - x', y - y') + b(x', y') \mid (x', y') \in D_b\}, \tag{2}$$
$$f \circ b = (f \ominus b) \oplus b, \tag{3}$$
$$f \cdot b = (f \oplus b) \ominus b, \tag{4}$$

where $\ominus$ and $\oplus$ are morphological erosion and dilation resp.,
f is the input image,
x, y are the coordinates of the pixel,
b is structural element,
D_b is the set of relative locations that are included in a structural element,
$\circ$ and $\cdot$ are morphological opening and closing, respectively.



a) b)

Fig. 3 True colour image of flood area. a) Original image (the dark spots in the left side are trees that are half submerged by flood), b) result of morphological filtering.

2) Textural features

Although spectral features are considered to be crucial in image classification, the support of textural features is sometimes necessary to differentiate different classes (especially when only one or a few bands are available, and the patterns of different classes are detectable). Textural features that are commonly applied in image classification include variance, entropy, Gray-Level Co-occurrence Matrix (GLCM) features, wavelet bands, Gabor filtered bands, etc. Most of these features are extracted by sliding a window across the whole image and running the corresponding calculation within each window. The window is typically slid pixel by pixel. Running this processing on a remote sensing image, however, can be slow and may not satisfy the requirement of quick analysis – the volume of remote sensing data are sometimes huge (e.g., when the full extent of a satellite image is considered) and the pixel-wise sliding scheme increases the amount of computation significantly. In order to avoid this problem while maintaining the accuracy of feature extraction, a set of overlapped windows are developed to extract the textural features (the feature extraction process is done in MATLAB in our experiment). The windows are designed to have some degree of overlapping, and the statistic results of each window will be assigned only to the central pixels. The feature extraction process can be expressed as

$$f'(x, y) = S(f(W_i)), \forall\ (x, y) \in C_i,$$

where f is the input image,
f' is the textural features,

x, y are the coordinates of the pixel,
S is the operator that specifies the feature being extracted (e.g., variance);
C_i is the set of pixel coordinates corresponding to the ith window centre ;
W_i is the set of pixel coordinates corresponding to the ith window;
C_i is in the centre of W_i;
The size of window centre should be a factor of image size and window size.

An example is showed in Fig. 4. There are three 4×4 windows with 50% overlapping rate. The statistic result of each window is only assigned to the central pixels to maintain accuracy. Comparing with pixel-wise sliding scheme, this setting increases the computing speed by 4 times.

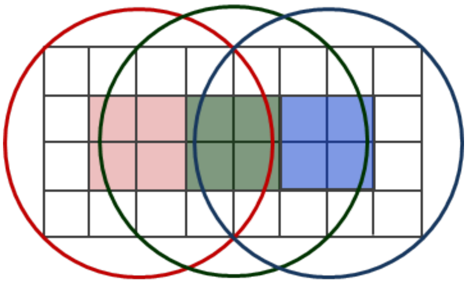


Fig. 4 Overlapped windows for textural feature extraction. There are three 4×4 windows with 50% overlapping rate. The statistic result of each window is only assigned to the central pixels to maintain accuracy.

3) Regional information

Since remote sensing imagery is the representation of earth surface, adjacent pixels are more or less related. In [11], the concept of spectral and spatial information is discussed, and the authors proposed to use segmentation results to support image classification. By dividing image into small segments, pixels in the same region can be considered as similar, despite their difference from a global point of view. This can help to reduce the noise-like misclassification spots significantly. To support image classification, the segmentation method applied to generate segments should be "conservative". That is, it is usually better to over-segment than to under-segment, since large segments have a high risk of involving pixels from different classes. In order to generate slightly over-segmented maps (i.e., the sizes of segments are slightly small than the objects that we are interested in), avoid extremely large segments and too fragmental results, conventional methods such as watershed and region growing have not been used. The methods tend to generate segments in various sizes. Some are too small while some could be too large. That is, when the settings (e.g., filtering method, edge detection method and connectivity in watershed) are "conservative" enough, the segmentation results are usually too fragmental. Instead, full λ-schedule segmentation [18] is applied to generate segments (the segmentation is done on ERDAS in our experiments). And once the image is segmented using full λ-schedule, the mean values of each region, etc. are extracted as regional features. Full λ-schedule segmentation merges adjacent regions iteratively depending on a merging cost. For adjacent regions *i* and *j*, the merging cost t_{ij} (ranges from 0.0 to 100.0) is

$$t_{ij} = \frac{\frac{|O_i|\times|O_j|}{|O_i| + |O_j|}\left\|\boldsymbol{u}_i - \boldsymbol{u}_j\right\|^2}{\text{length}\big(\partial(O_i, O_j)\big)}, \tag{5}$$

where O_i, O_j are regions i and j of the image,

$|O_i|$, $|O_j|$ are the areas of regions i and j,

$\boldsymbol{u}_i$, $\boldsymbol{u}_j$ are the mean values of regions i and j,

length($\partial(O_i, O_j)$) is the length of the common boundary of O_i and O_j.

IV. Mapping Environmental Hazards

There is a common concern in the analysis of environmental hazards, namely, the extents of the hazards. This can be estimated by image classification. In a two classes classification, extent of hazard is depicted by the class of abnormal area (in opposite to the class of normal area). Combing with the features discussed in Section II, a new image classification method for the mapping of environmental hazards is developed. The overall structure is available in Fig. 5. The method consists of three major steps: image pre-processing, feature extraction and classification. This pre-processing is shown in the flow chart. Once it is done, three types of features will be extracted from the integrated dataset.

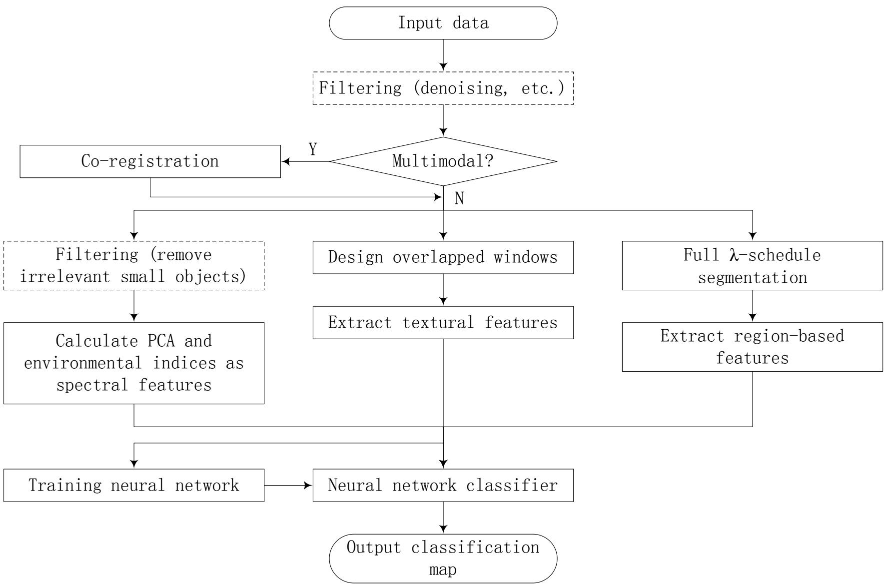


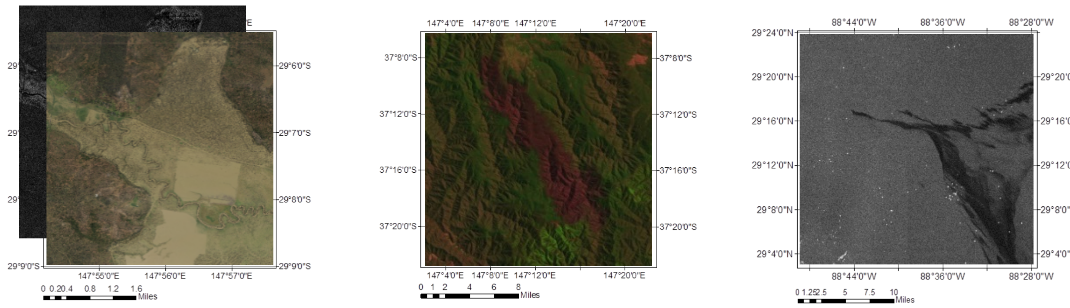
Fig. 5 Overall structure of the proposed method.

1) Spectral features: the first 2~3 bands of PCA, related environmental indices (e.g., NDVI, NDWI and NBR). The original data will be filtered before feature extraction, if necessary;
2) Textural features: statistic features such as mean, variance and entropy. A set of overlapping windows are applied to extract the features. These windows can help to improve computing efficiency while maintain accuracy;
3) Region-based features: the features of each region (e.g., mean value of the region). The regions are acquired by segmenting the original image using full λ-schedule.

Given the features from different perspectives, a multi-dimensional feature space is built. Each dimension represents an individual feature and each point in the feature space represents the features related to an individual pixel. In our method, feedforward neural network (NN) with 2 hidden layers is trained for the classification task in feature space (the process is done in MATLAB in our experiment). The output of neural network classifier is used to plot the final classification map.

IV. Cases of Environmental Hazard

Three remote sensing datasets on different environmental hazards are provided to validate the proposed method. Details of the three hazards are shown below.



a) b) c)

Fig. 6 Remote sensing data of environmental hazards. a) The 2009 New Angledool flood: ADS40 airborne optical data and COSMO-SkyMed radar data; b) the 2009 Victorian bushfires: Landsat-5 optical data; c) the 2010 Deepwater Horizon oil spill: COSMO-SkyMed radar data.

1) The 2009 New Angledool flood. In the early months of 2009, a vast area of Queensland was affected by flooding, due to heavy rainfall. According to the news (March 18, 2009) of State Emergency Service (SES), NSW, New Angledool was expected to be "isolated (by flood) later this week"[19]. Two sets of remote sensing images and a helicopter survey data provided by SES (polygon data collected by helicopter GPS surveying) are available for the cases. They are selected since the datasets are acquired a few days after the flooding and cloud free. The details are available in Table 1.

Table 1 Details of the datasets

Dataset	Size	Resolution (m)	Pre-processing	Acquire date
COSMOS-SkyMed	27304×29566	2.5	SLC	19&20/03/2009
ADS40	51950×41202	0.2	Calibrated & Geocoded	March, after flooding
Helicopter survey	NA	NA	NA	March, after flooding

2) The 2009 White Timber Spur fire in Dargo, Victoria. Due to the extreme weather, a series of bushfires were burnt across Victoria in 2009. Bushfire was reported near Dargo, Victoria on February 07, 2009 [20]. Cloud free Landsat 5 TM data (Level 1, resolution: 30m, size: 8121×7191) acquired on February 25, 2009 are available for

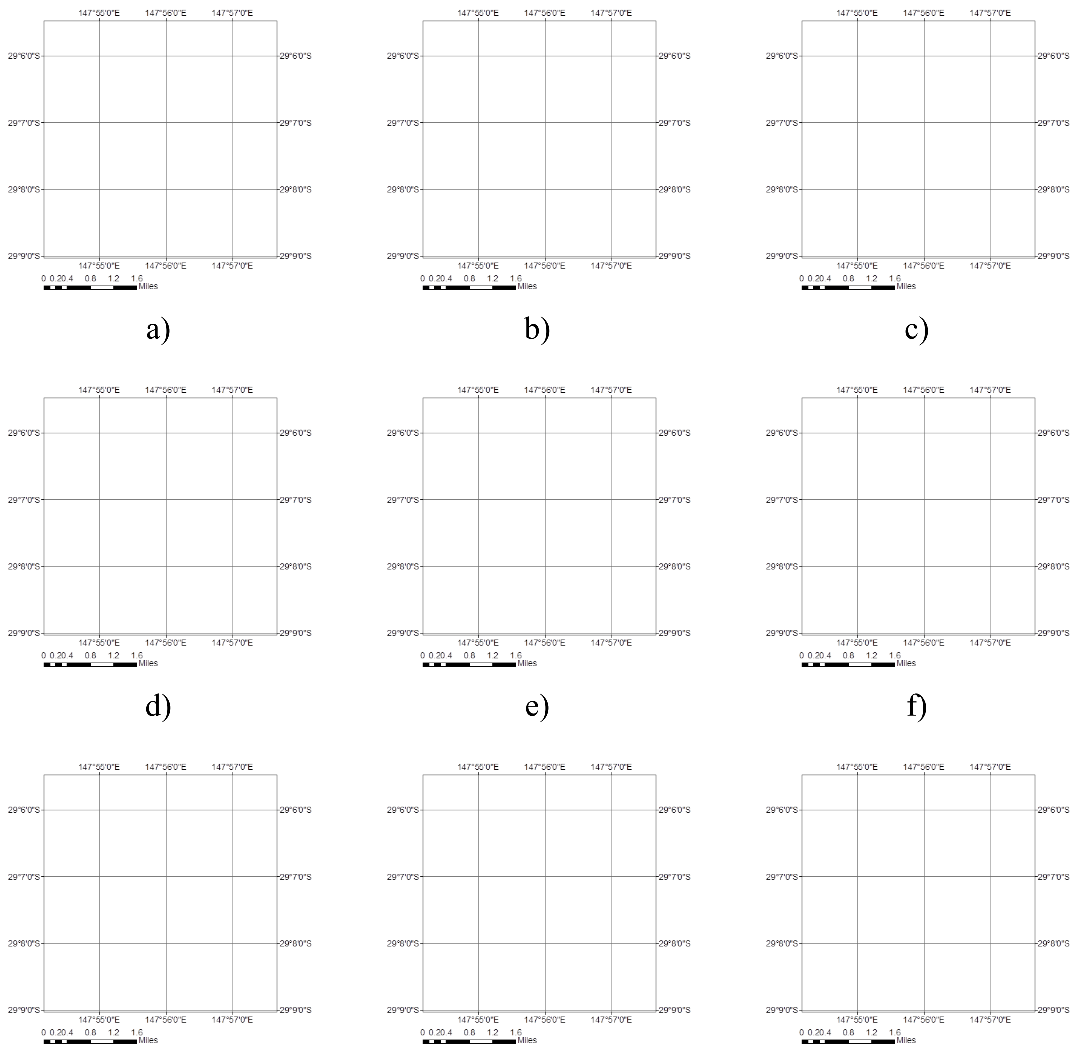
the case. As it takes months for the area to recover [20], the Landsat data are considered to be suitable.

3) The 2010 Deepwater Horizon oil spill. The oil spill began on 20 April 2010 in the Gulf of Mexico due to explosion. By the end of April, the oil slicks were "reaching the vulnerable Louisiana coastal wetlands"[21]. COSMOS-SkyMed (calibrated and geocoded from SLC data, size: 4103×2980, resolution: 4 m) acquired on May 2, 2010 are available. Since the data acquired soon after the spreading of oil slicks, it is able to depict the effect of this oil spill.

V. Experiments

Experiments are conducted for the three environmental hazards introduced above. In the first case, both the ADS40 and COSMOS-SkyMed data are resampled to 4m resolution for registration.

1) The 2009 New Angledool flood. Flood extent measured by helicopter surveying is illustrated in Fig. 7. And in



a) b) c)

d) e) f)

g) h) i)

Fig. 8, the classification results using different methods, the overlapped views of the classification maps and original optical image, and the comparison between classification maps and survey data are shown. By comparing the results with helicopter survey data, classification accuracy is assessed and illustrated in Table 2.

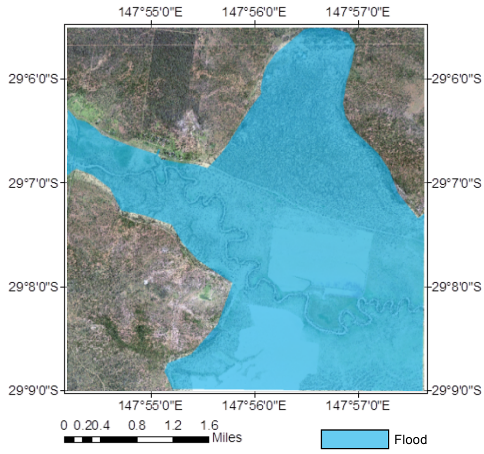
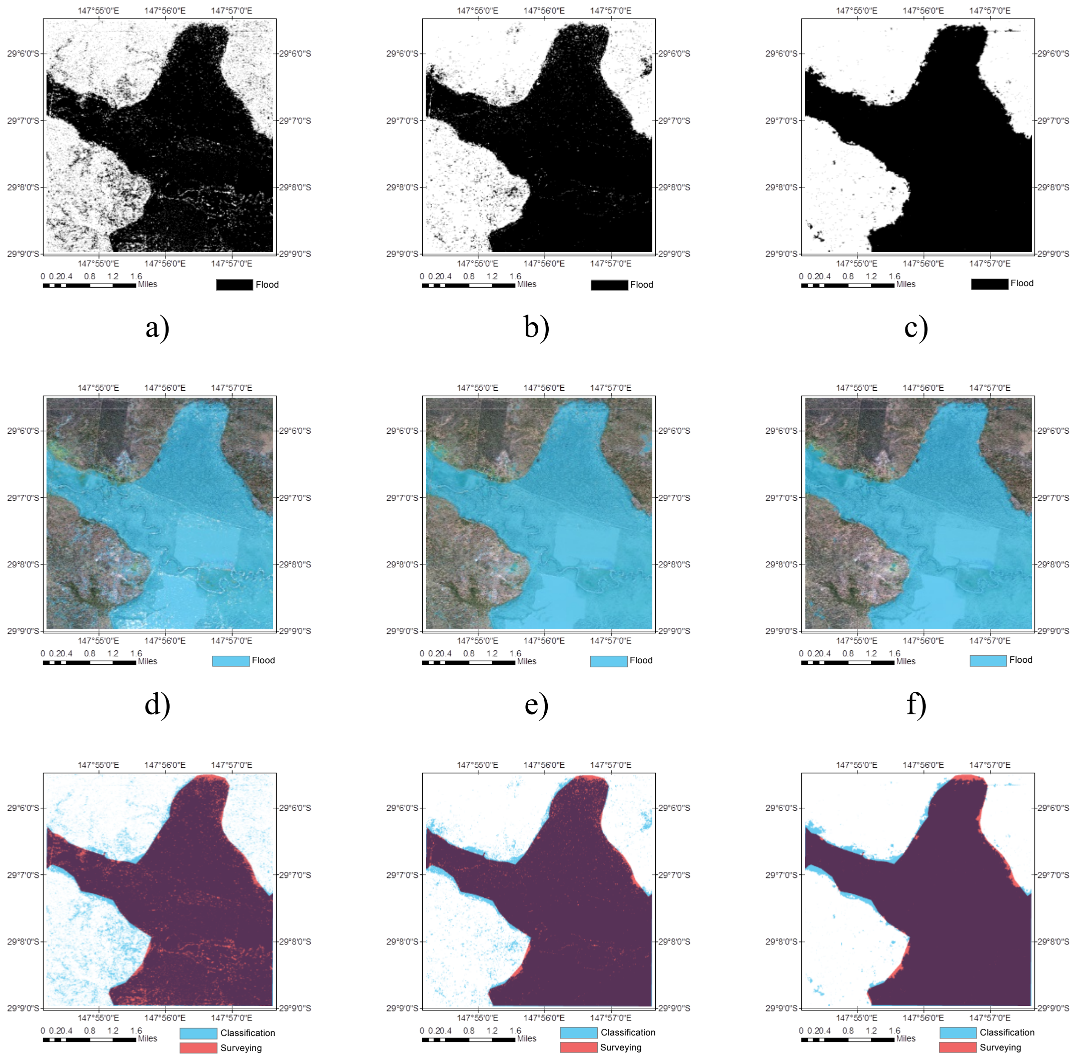


Fig. 7 Flood extent estimated by helicopter survey.



a) b) c)

d) e) f)

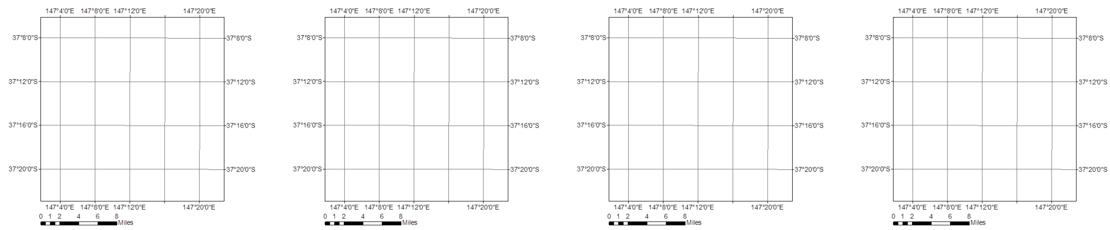
g) h) i)

Fig. 8 Classification results using different methods. From a) to c): classification results using spectral features + NN, textural features + NN, and the proposed method; from d) to f): overlapped views of the classification maps and original optical image; from g) to i): the comparison between classification maps (blue) and survey data (red).

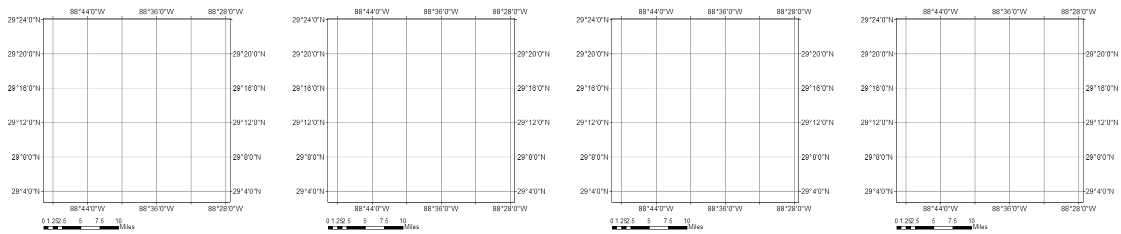
Table 2 The similarity between classification maps (denoted as “C”) and helicopter survey data (denoted as “S”).

Spectral + NN			Textural + NN			The proposed method		
S \ C	Flood	Others	S \ C	Flood	Others	S \ C	Flood	Others
Flood	94.2%	5.8%	Flood	97.3%	2.7%	Flood	98.3%	1.7%
Others	13.5%	86.5%	Others	7.5%	92.5%	Others	5.0%	95.0%

2) The 2009 White Timber Spur fire in Dargo, Victoria, and the 2010 Deepwater Horizon oil spill.

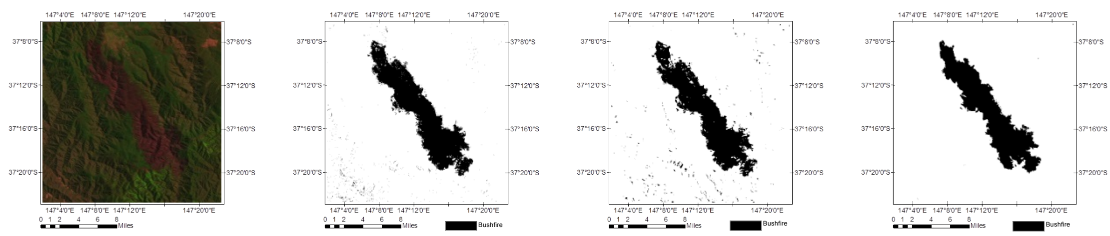


a) b) c) d)

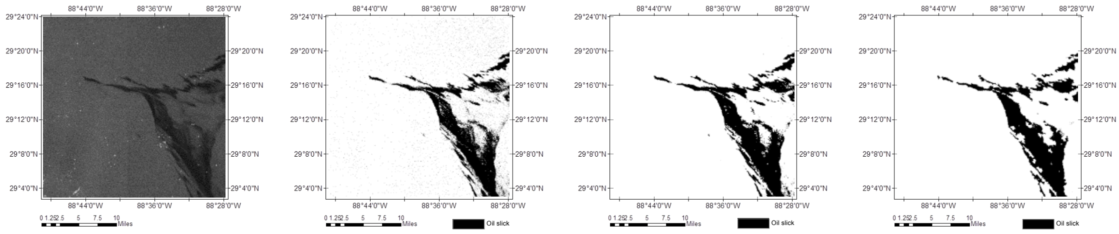


e) f) g) h)

Fig. 9 shows the classification results using different methods. The overlapping views of the classification maps and original images are also shown in Fig. 10.

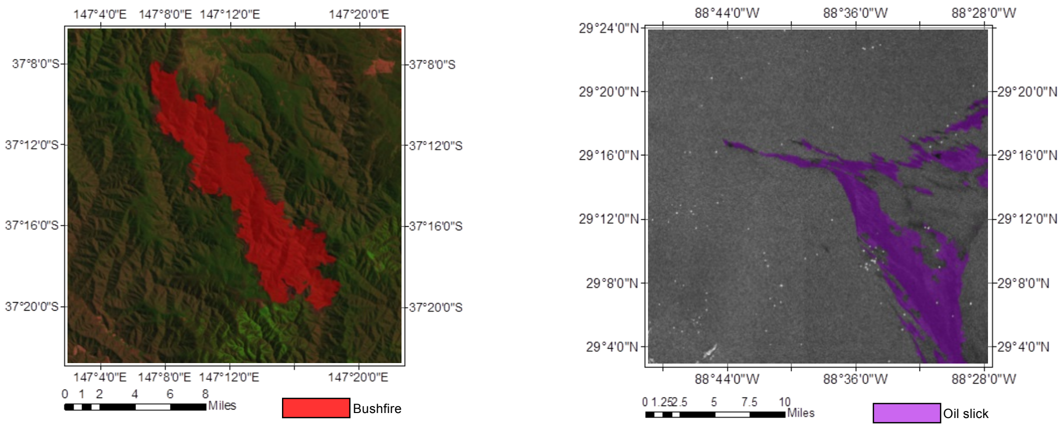


a) b) c) d)



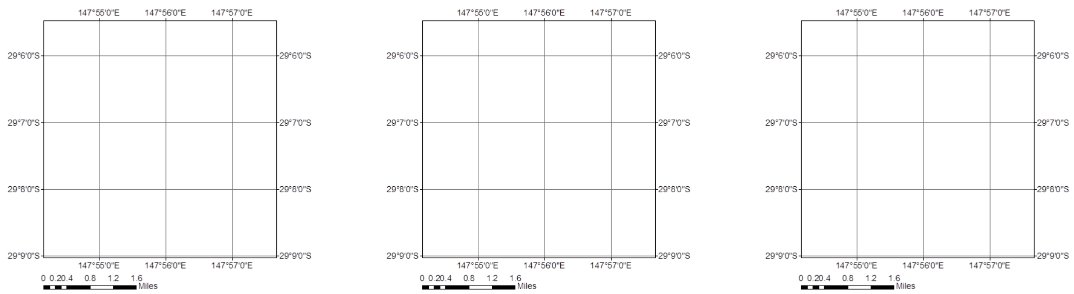
e) f) g) h)

Fig. 9 Classification results using different methods. From left to right: original images, classification results using spectral features + NN, textural features + NN, and the proposed method.

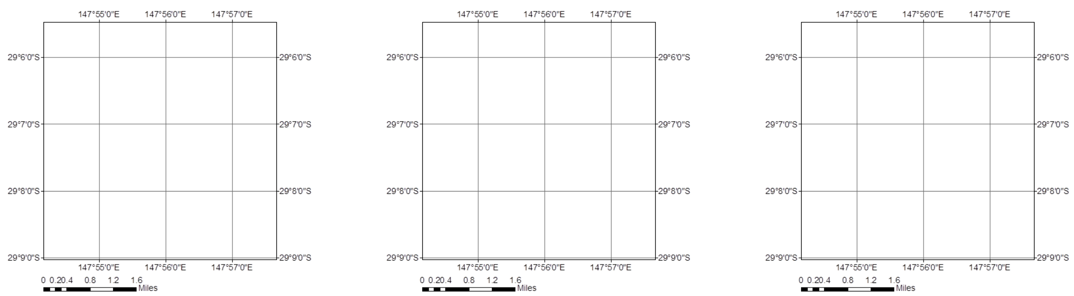


a) b)

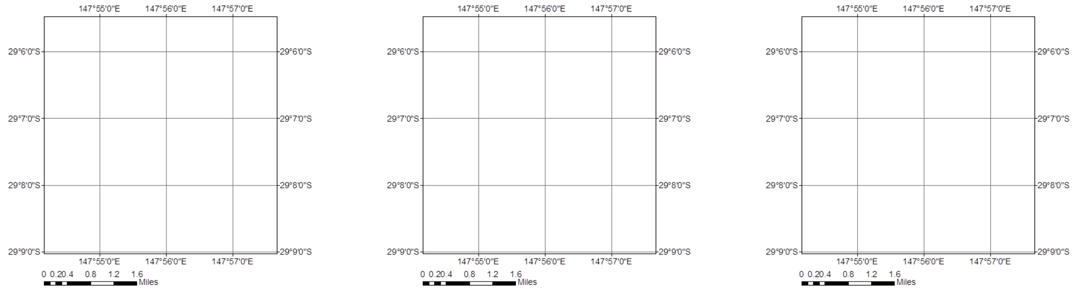
Fig. 10 Overlapping views of the classification maps and original images: a) bushfire, b) oil spill.



a) b) c)

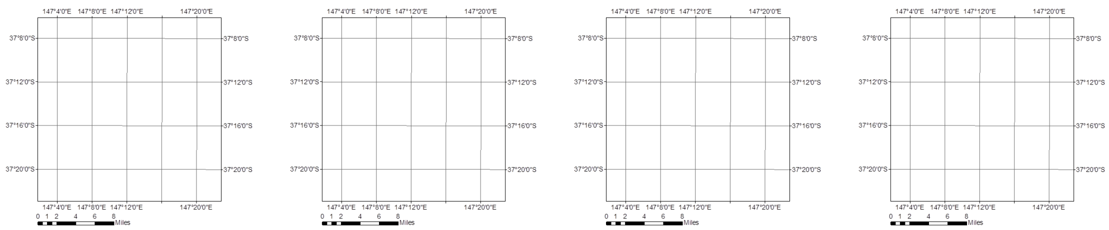


d) e) f)

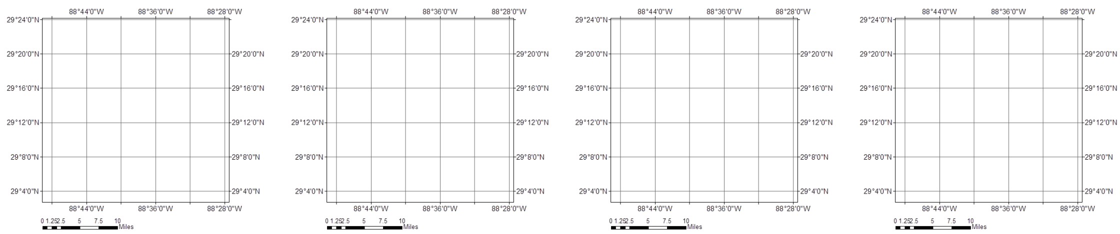


g) h) i)

Fig. 8 and Table 2 show that the result of the proposed method has the highest similarity with the helicopter survey data, in both flooded areas and non-flooded areas. Specifically, the mapping similarity of flood extent is higher than 98%. Moreover, noise-like misclassification spots appear in the first two results are reduced significantly in the third one. This improves the visual quality of classification map and makes the boundary of flood more explicit. Meanwhile,



a) b) c) d)



e) f) g) h)

Fig. 9 indicates that the mapping results of the proposed method contain less misclassification spots and have clear boundaries as well. In short, for the three cases of environmental hazards, the proposed method showed the ability to generate mapping results with less misclassification spots and more explicit boundaries. And a very high similarity with helicopter survey data is shown in the first case. The performance could be explained as below. 1) PC bands and environmental indexes were used for spectral features after morphologic filtering, this can alleviate the influence of irrelevant objects (e.g., half-submerged trees in flood) and preserve significant information for classification; 2) by combing different types of features (spectral, textural and regional) in neural network classifier, more comprehensive interpretation on the image are provided. Besides, instead of classifying images segment by segment using segmentation results, the segmentation results are used to support the extraction of regional information. This can alleviates the problem caused by segmentation error.

VI. Conclusion

Owing to some advantages that ground survey is difficult to achieve, remote sensing data is used for the near real-time analysis of environmental hazards. The complexity of the scenes, however, could sometimes make the classification difficult. In the paper, we discussed the way to improve the quality of classification results, and a new image classification method using features from multiple perspectives was developed. The method integrates features from multiple perspectives in an NN classifier to classify images. Experimental results indicated that, for some of the environmental hazard cases, the method performs well in reducing classification artefact and preserving boundaries. And in the case of the 2009 New Angledool flood, a 98% classification similarity with helicopter survey data was shown. Besides, the design on extracting textural features (which usually take a relatively long time) helps to improve computing efficiency.

Since the mapping results contain less noise and have explicit boundary, it is preferable when converted into polygons. The results can be used to support volume calculation, damage assessment, the planning of evacuation and rescue, etc. In the future study, we will look into the design of classifier and post-classification process, thus to further improve the method.

References

1. Lindsey, R., "Satellites Do It Faster, Cheaper", webpage, 23 October 2013, assessed 19 February, 2015, available at: http://earthobservatory.nasa.gov/Features/BAER/baer2.php
2. Leica Geosystems, "Total Stations (TPS) ", webpage, 2015, assessed February 15, 2015, available at: http://www.leica-geosystems.com/en/Total-Stations-TPS_4207.htm
3. Guo, L., "Land cover change detection using Landsat TM imagery of the 2009 Victorian bushfires", Master Thesis, The University of New South Wales, Australia, 2011.
4. Brumby, S. P., Harvey, N. R., Bloch, J. J., et al., "Evolving forest fire burn severity classification algorithms for multispectral imagery", *Proc. SPIE 4381, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII*, 2001, pp. 236-245.
5. Jain, S. K., Singh, R. D., Jain, M. K., et al., "Delineation of flood-prone areas using remote sensing techniques", *Water Resources Management*, Vol. 19, No. 4, 2005, pp. 333-347.
6. Roy, D. P., Boschetti, L., and Trigg, S.N., "Remote sensing of fire severity: assessing the performance of the normalized burn ratio", *Geoscience and Remote Sensing Letters*, IEEE Vol. 3, No. 1, 2006, pp. 112-116.
7. Brekke, C, Solberg, A. H. S., "Oil spill detection by satellite remote sensing", Remote Sensing of Environment, Vol. 95, No.1, 2005, 1-13.
8. Nyoungui, A. N., Tonye, E., and Akono, A., "Evaluation of speckle filtering and texture analysis methods for land cover classification from SAR images", *International Journal of Remote Sensing*, Vol. 23, No. 9, 2002, pp. 1895-1925.
9. Agüera, F., Aguilar, F. J., and Aguilar, M. A., "Using texture analysis to improve

per-pixel classification of very high resolution images for mapping plastic greenhouses", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 63, No. 6, 2008, pp. 635-646.
10. Clausi, D. A., and Deng, H., "Design-based texture feature fusion using Gabor filters and co-occurrence probabilities", *Image Processing, IEEE Transactions on*, Vol.14, No.7, 2005, pp. 925-936.
11. Tarabalka, Y., Benediktsson, J. A., Chanussot, J., et al., "Multiple spectral–spatial classification approach for hyperspectral data", *Geoscience and Remote Sensing, IEEE Transactions on*, Vol. 48, No. 11, 2010, pp. 4122-4132.
12. Bruzzone, L., and Carlin, L., "A multilevel context-based system for classification of very high spatial resolution images", *Geoscience and Remote Sensing, IEEE Transactions on*, Vol. 44, No. 9, 2006, pp. 2587-2600.
13. Zhang, L., Huang, X., Huang, B., et al., "A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery", *Geoscience and Remote Sensing, IEEE Transactions on*, Vol. 44, No. 10, 2006, pp. 2950-2961.
14. Ma, W. Y., & Manjunath, B. S., "EdgeFlow: a technique for boundary detection and image segmentation", *Image Processing, IEEE Transactions on*, Vol. 9, No. 8, 2000, pp. 1375-1388.
15. Zhang, L., Zhang, L., Tao, D., et al., "On combining multiple features for hyperspectral remote sensing image classification", *Geoscience and Remote Sensing, IEEE Transactions on*, Vol. 50, No. 3, 2012, pp. 879-893.
16. Fauvel, M., Chanussot, J., and Benediktsson, J. A., "A spatial-spectral kernel-based approach for the classification of remote-sensing images", *Pattern Recognition*, Vol. 45, No. 1, 2012, pp. 381-392.
17. Chu, H. T., "Integration of Multisource Remote Sensing Data for Improvement of Land Cover Classification", Doctoral dissertation, The University of New South Wales, 2013, pp. 99-101.
18. Robinson, D. J., Redding, N. J., and Crisp, D. J., "Implementation of a fast algorithm for segmenting SAR imagery", DSTO Electronics and Surveillance Research Laboratory, 2002.
19. State Emergency Service NSW, "SES Flood Update 18 March", webpage, 2009, assessed February 17, 2015, available at: http://www.ses.nsw.gov.au/news/2010/ses-flood-update-18-march
20. Raadik. T., and Nicol, M., "Assessment of the post-fire status and distribution of the Dargo Galaxias (Galaxias sp. 6), affected by the White Timber Spur fire, upper Dargo River system", Department of Sustainability and Environment, 2012.
21. BBC News, "Timeline: BP oil spill", webpage, 2010, assessed February 17, 2015, available at: http://www.bbc.com/news/world-us-canada-10656239

Coastal Erosion Detection Using Multispectral Images with Tidal Model

Qingxiang LIU, Linlin GE and Xiaojing LI

Geoscience and Earth Observing Systems Group (GEOS), School of Civil and Environmental Engineering, UNSW Australia, Sydney NSW 2052, Australia

Summary: Coastal erosion has long been a threat to the safety and property of coastal residents in many countries including Australia. When coastal erosion occurs, coastal roads and infrastructures will be destroyed, which can increase the flooding potential and lead to transportation issues. Satellite imagery has the advantage of large temporal coverage, which enables long-term measurements with data from many years till now. In this paper, an integrated methodology using satellite imagery to detect the long-term coastal erosion is proposed. Linear spectral unmixing and sub-pixel classification are the key techniques to attain the water-land boundaries in a sub-pixel level and the introducing of a simplified tidal model enables the identification of eroded areas. In the experiment, Landsat TM 5 imagery is utilised to detect the coastal erosion at East Gippsland, Victoria, Australia. The results prove that Landsat multi-spectral images with medium resolution have the potential to detect long-term coastal erosions in a sub-pixel level.

Keywords: Coastal erosion, spectral unmixing, sub-pixel classification, tidal model, multi-spectral images

1. Introduction

Coastline describes the land-water interface (LWI) segregating a standing body of water from land [1]. Coastal erosion can decrease the protection from coastal infrastructures, which increased flooding potential. Since the beach-front roads and buildings can be eroded, it can lead to transportation issues as well as coastal property loss. Moreover, the change of coastlines might affect the plants and animals and cause eco-system shifts.

Traditional beach surveying method [2] using total stations and levels can provide accurate results. However, it is labour-intensive and time-consuming. A more recent in-situ method using kinematic differential GPS is accurate and relatively rapid [3], while it still involves field work. Aerial photography [4] and airborne LiDAR [5] are limited in their temporal and spatial availability [6]. Video imaging [7] method can provide temporally dense data while it is spatially limited [6]. Compared with those approaches, satellite remote sensing can provide data with a very large coverage as well as long period of time.

So far, a large amount of research on coastline extraction with remote sensing imagery has been conducted. Puissant et al. [8] used mathematical morphology method to precisely locate the coastlines in Very High Resolution (VHR) multispectral images. CHEN et al. [9] developed a software package to extract the coastline from Terra satellite ASTER images using MATLAB as the programming tool. However, many papers were focused more on the extraction of the instant water-land boundary without further analysis of results over a long period.

This paper aims to propose an integrated methodology to detect the long-term coastal erosion using multi-spectral images. The most significant advantage of Landsat imagery is its large temporal coverage, which is appropriate for detecting slow changes like coastal erosion. Although the imagery has a medium resolution of 30m (and band 6 is 120m) which is not high enough to detect coastal erosion, pixels can be separated into several sub-pixels by spectral unmixing and sub-pixel classification; thus the resolution of extracted boundaries can be significantly improved. The expected results will not only benefit coastal residents as disaster prediction, but also provide useful information for urban planning and research of climate changes.

2. Experimental Methodology

The eastern beach of Gippsland Basin, Victoria, Australia was chosen as the study area as shown in Fig. 1. The latitude and longitude of the centre are 38°13'58"S and 147°22'27"E.

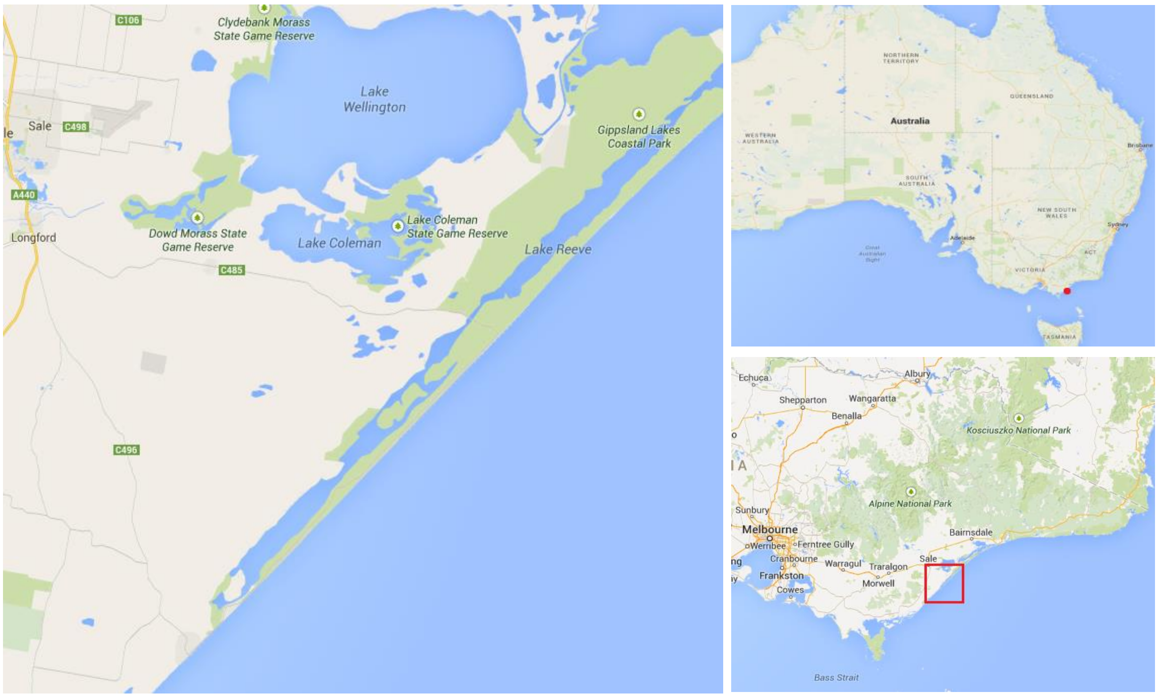


Fig. 1: Study area: the eastern coast of Gippsland, Victoria, Australia

20 Landsat TM 5 level 1T products from year 2007 to year 2011 were downloaded from the United States Geological Survey (USGS) website http://earthexplorer.usgs.gov/. The band 6 (thermal band) was not utilised in the experiment as a result of its lower spatial resolution of 120m. Table 1 shows the specifications of the data.

Table 1 Specifications of experiment Landsat TM data

Number	Acquisition date	Path number	Row number
1	2rd January 2007	91	86
2	2rd January 2007	91	87
3	4th February 2007	91	86
4	4th February 2007	91	87
5	31st August 2007	91	86
6	31st August 2007	91	87
7	4th October 2008	91	86
8	4th October 2008	91	87
9	8th January 2009	91	86
10	8th January 2009	91	87
11	24th January 2009	91	86
12	24th January 2009	91	87
13	8th November 2009	91	86
14	8th November 2009	91	87
15	10th October 2010	91	86
16	10th October 2010	91	87
17	30th January 2011	91	86
18	30th January 2011	91	87
19	19th March 2011	91	86
20	19th March 2011	91	87

Fig. 2 shows the general flow chart of the proposed methodology, where the parallelograms represent data or results while the rectangles represent the processing or operation methods.

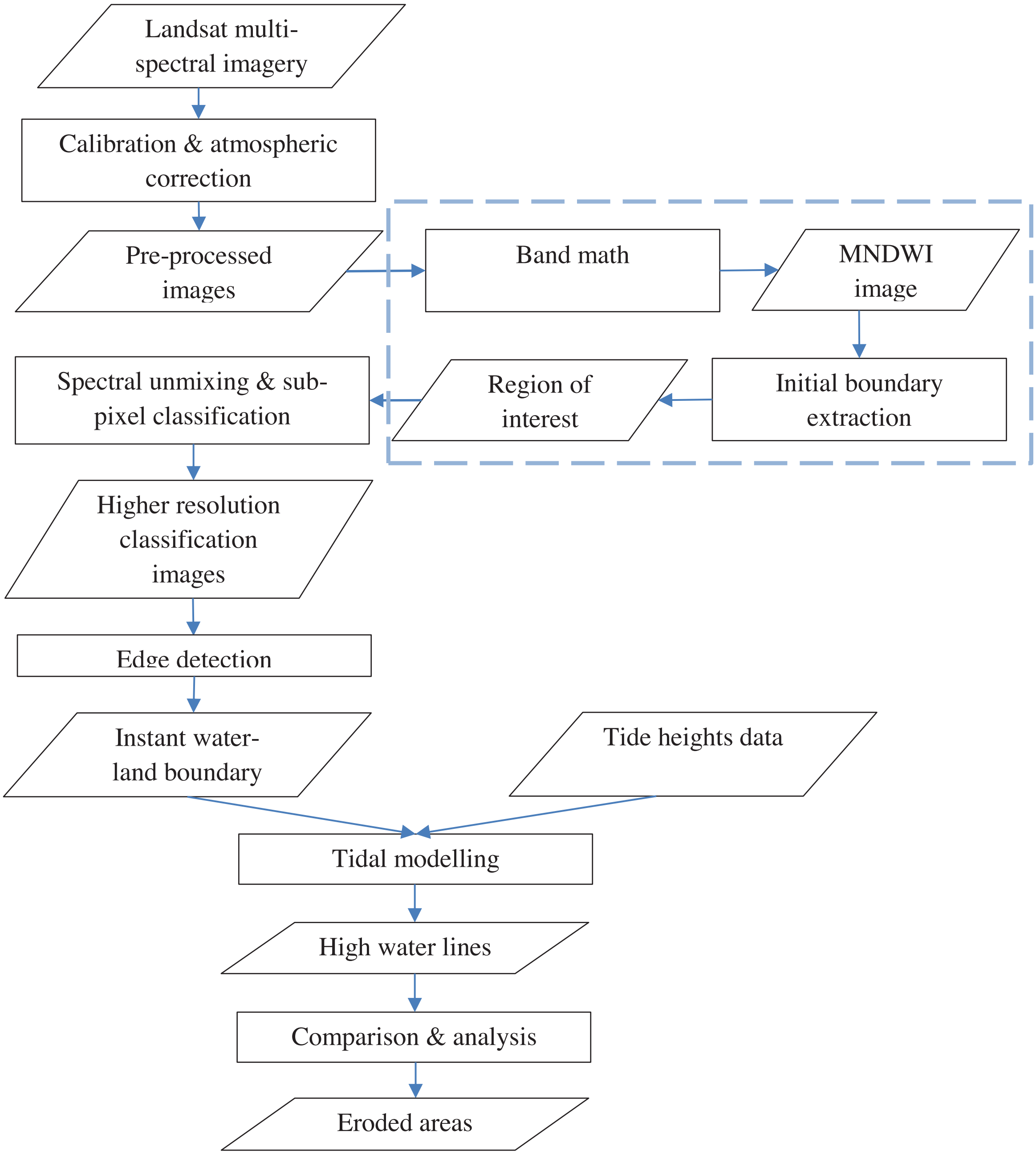


Fig. 2: Flow chart of coastal erosion detection

It should be noted that all the multispectral images need to be radiometric calibrated and atmospherically corrected before further processing, which was conducted by using ENVI software. However, the whole procedure is mainly realised by MATLAB programming, which includes the key steps: initial boundary extraction, linear spectral unmixing, sub-pixel classification and integration with a tidal model.

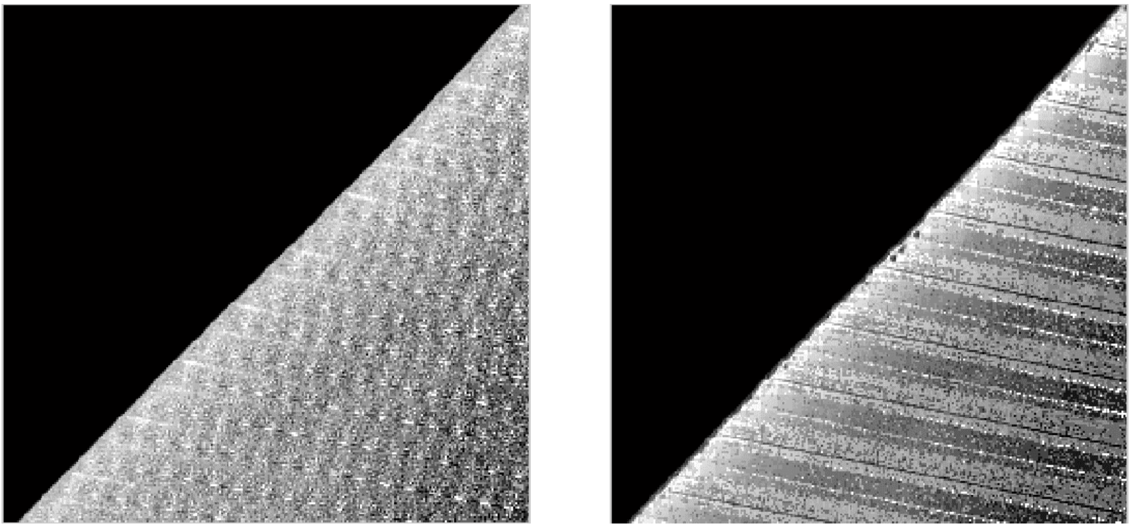
2.1 Initial Boundary Extraction

In order to improve the accuracy of endmember extraction as well as to reduce the computation time, the initial water-land boundary was extracted, as shown in the dash line rectangle area. Based on the initial boundary in a pixel level, a buffer area was created as the region of interest (ROI) for the further processing steps. In this case, the modified normalised difference water index (MNDWI) [10] was calculated by the calculation between bands as

Eqn 1. This index can increase the contrast between water and soil thus the boundary can be clearer and easier to be extracted.

$$MNDWI = \frac{Green - MIR}{Green + MIR} \quad (1)$$

Where *Green* represents green band and *MIR* represents middle infrared band, which correspond to band 2 and band 5 for Landsat TM 5 images respectively. As shown in Fig. 3, compared with NDWI (normalised difference water index) [11], the boundary is clearer since the impacts of non-water features can be better supressed.



(a) *(b)*

Fig. 3: MNDWI image (a) and NDWI image (b)

Once the MNDWI image is attained, the boundary can be easily extracted by the built-in edge detection function *edge* in MATLAB after binarisation. The extracted boundary (the red line) overlapped with the clipped Landsat image is shown as Fig. 4 (a). After dilating the line by several numbers of pixels (e.g. 5 pixels), a ROI where can be created as Fig. 4 (b).

(a) *(b)*

Fig. 4: Extracted initial boundary (a) and expanded ROI (b)

2.2 Linear spectral Unmixing

Since the spatial resolution of Landsat TM 5 images is not high enough for detecting the coastal erosion. To tackle this issue, the combination of spectral unmixing and sub-pixel classification is applied to improve the resolution of the extracted boundary.

For remote sensing images, every pixel might not contain only one ground feature but is generated as a mixed combination of different ground features, which is caused by the resolution limitation of the sensor or that an object covers across the pixel boundaries. Spectral unmixing theory was built on the fact that each image pixel is presenting the joint combination of spectral reflectance from several ground features (endmembers) [12]. Since the indirect illumination scattered from adjacent features to the water can be neglected, linear unmixing [13] which assumes the combination is linear is applicable. The linear mixing model for a single pixel is as [14]

$$x_i = \sum_{j=1}^{N} \rho_{ij}\alpha_j + w_i \tag{2}$$

Where x_i is the reflectance of the pixel at band i, N is the number of endmembers, ρ_{ij} represents the reflectance of endmember j at band i, α_j is the fractional abundance of endmember j and w_i is additive observation noise.

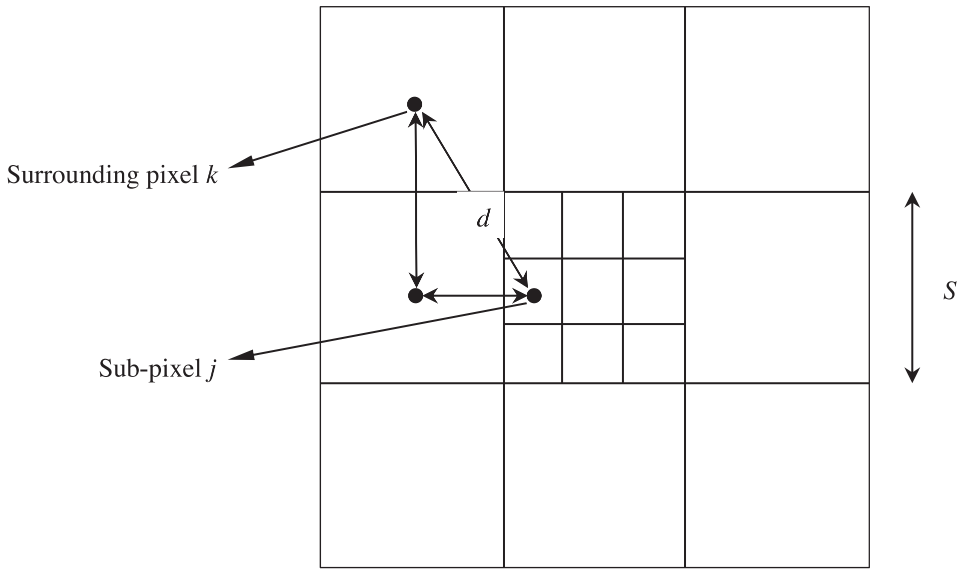
In this paper, three endmembers: vegetation, soil and water which are the major components of the coastal pixels were considered. There are many algorithms to identify the endmembers for multi/hyperspectral images based on the notion of spectral mixture modelling including N-FINDR and pixel purity index (PPI)[14]. A simpler method which combines the indices of normalised difference vegetation index (NDVI), NDWI and band 5 was introduced. Firstly, thresholds for the indices were applied to identify three groups of candidate pixels for those three endmembers. For example, pixels with NDWI above 0.5 were chosen as candidates for water body; pixels with NDVI above 0.6 and band 5 above 0.1 for vegetation; pixels with NDVI below 0.3, NDWI below -0.35 and band 5 reflectance above 0.1 for soil. Then the average reflectance of the three pixel groups at band i was considered as the reflectance of the three endmembers respectively, which means ρ_{ij} in Eqn 2 was attained. By using the linear least square method to solve Eqn 2, the proportions (fraction abundances) of the three ground features can be solved.

2.3 Sub-pixel Classification

With the fraction image attained by spectral unmixing, sub-pixel classification is to separate every pixel into several sub-pixels and to assign them with their corresponding ground features. Sub-pixel classification was firstly raised by Atkinson[15] and there are several algorithms for sub-pixel mapping including genetic-based algorithm, neural networks algorithm, artificial neural networks (ANN) predicted wavelet coefficients and spatial attraction sub-pixel mapping [16]. Considering the computation time, the sub-pixel classification method based on spatial attraction model is applied. In this algorithm, assuming S_{ij} is the spatial dependence on ground feature i for sub-pixel j, it can be expressed as

$$S_{ij} = \sum_{k=1}^{N} w_k * Fraction_k \tag{3}$$

In Eqn 3, for sub-pixel j, k is one of the pixels surrounding the centre pixel which sub-pixel j belongs to; N is the number of its surrounding pixels; w_k is the weight of surrounding pixel k and $Fraction_k$ is its fractional abundance. In this case, the weight w_k is the inverse distance d between the sub-pixel j and its surrounding pixel k. One example of the calculation of d is shown as Fig. 5, where S represents the decomposition factor, where S^2 means how many sub-pixels each pixel can be divided into.



$$d^2 = S^2 + (0.5S + 0.5)^2$$

Fig. 5: Distance between sub-pixel j and its surrounding pixel k

For sub-pixel j, if the calculated spatial dependence of ground feature i is larger than that of any other ground feature, then sub-pixel j will be assigned with ground feature i [16]. It was found that the decomposition factor S can be as large as 9 for Landsat TM 5 images to be saturated [17], which means every pixel can be separated into 81 sub-pixels. After the sub-pixel classification, the water-land boundary can be extracted from the classified higher resolution image by simply using an edge detection function in MATLAB.

2.4 Integration with a Tidal Model

Since the extracted curves are based on the instant images when the satellites passed over the imaging area, they only represent the instant water-land boundaries. To get the real eroded areas, the instant boundaries need to be corrected to the coastline. There are many indicators of coastline while the most commonly applied one is the high water line (HWL) [6]. To correct the extracted boundary to HWL, a tidal model which describes how the tide heights change with the position of boundaries is required. For sandy beach like the eastern coasts of Gippsland, a conventional and commonly applied tidal model was proposed by Shen et al. [18]. The geometric model can be described as Fig. 6:

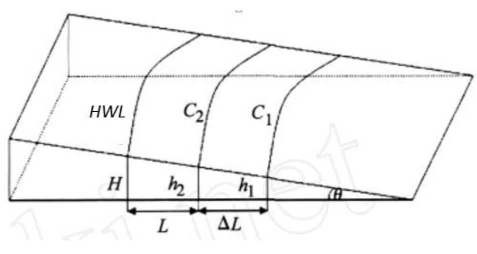


Fig. 6: A simple tidal model [18]

Where C_1 and C_2 represent the extracted instant water-land boundaries from two images, h_1 and h_2 are the tidal heights at the two imaging times, H is the tide height for HWL and θ is the slope angle. The intertidal distance between C_2 and HWL can be described by L where

$$L = \frac{H - h_2}{\tan\theta} \tag{4}$$

Similarly, the largest intertidal distance L_{max} can be calculated using the tide height of the HWL and the accumulated low tide line h

$$L_{max} = \frac{H - h}{\tan\theta} \tag{5}$$

The tide heights data were collected by the Australian Bureau of Meteorology (BOM), from which the tide height at the imaging time can be acquired. The tide heights at imaging time (10:00 am of local time) of the five years were calculated. Fig. 7 shows the variations of tide heights at 10:00am of year 2007 and 2008 together:

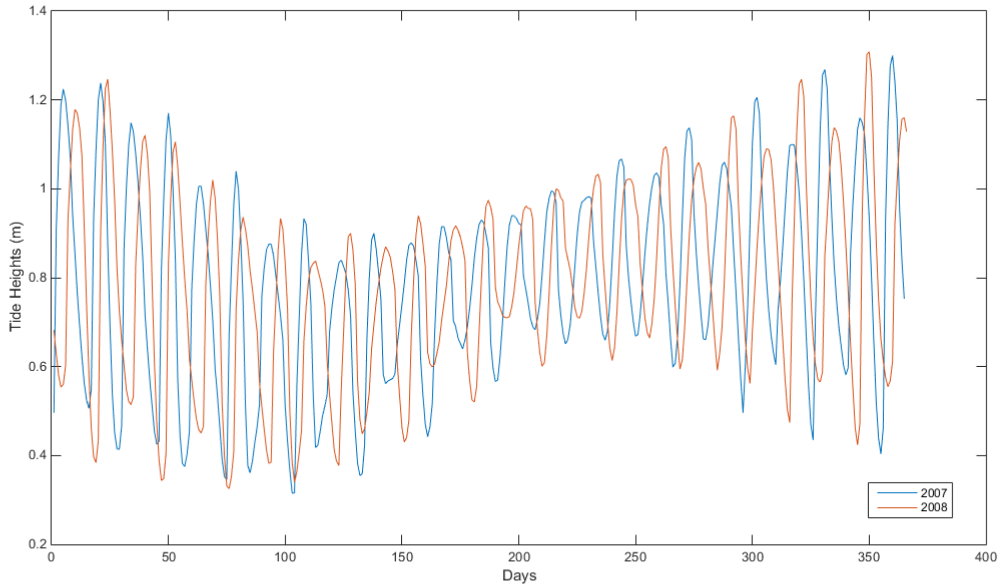


Fig. 7: Tide heights variations at 10:00am of year 2007 and 2008

From the plotted data, we can identify the tidal ranges for the five years are 0.98m, 0.98m, 1.00m, 1.02m and 0.96m respectively and the mean value is 0.99m. As Eqn 5, the largest intertidal distance can be calculated with the mean value 0.99m and a known slope angle. For Gippsland beach, since the slope angle ranges from 1° to 1.5° [19]. Therefore, the largest intertidal distance is between 37.9m and 56.8m.

The comparison and analysis were realised using ArcMap 10.2. The seaward boundary of the lines extracted from those images was extracted as the accumulated low tide line. Fig. 8 shows the extracted water-land boundaries (numbered according to Table 1) and the low tide line, overlapped with the Landsat TM image acquired on 4th February 2007. By applying a buffer with the buffer width of the largest intertidal distance, the intertidal areas exceeding the buffer zone (risk zone) were identified as water-eroded areas. Fig. 9 shows the pessimistic risk zone (37.9m) and the optimistic risk zone (56.8m) when slope angle is 1.5° and 1° respectively.

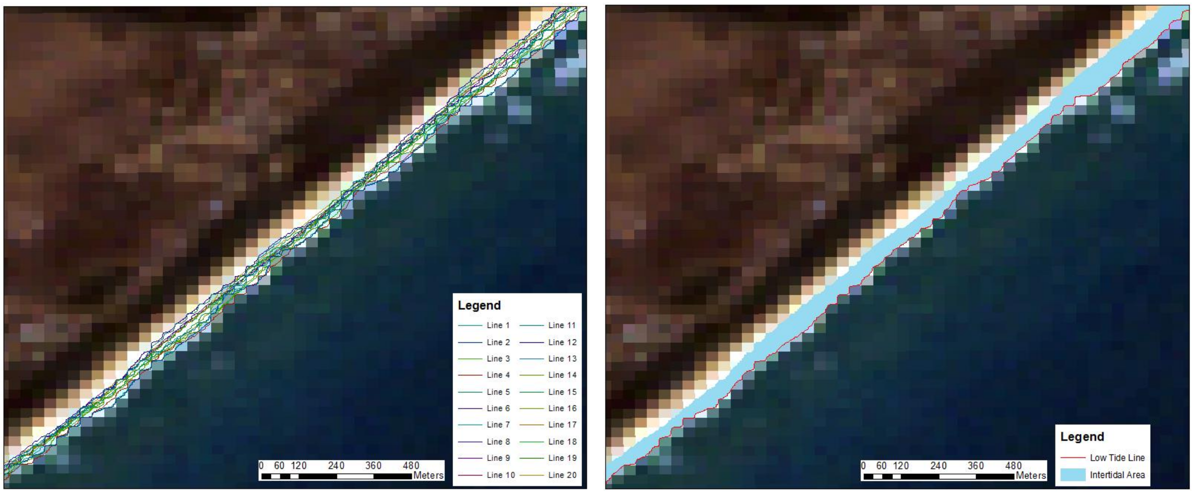


Fig. 8: Extracted water-land boundaries (left figure) and the low tide line (the red curve in the right figure)

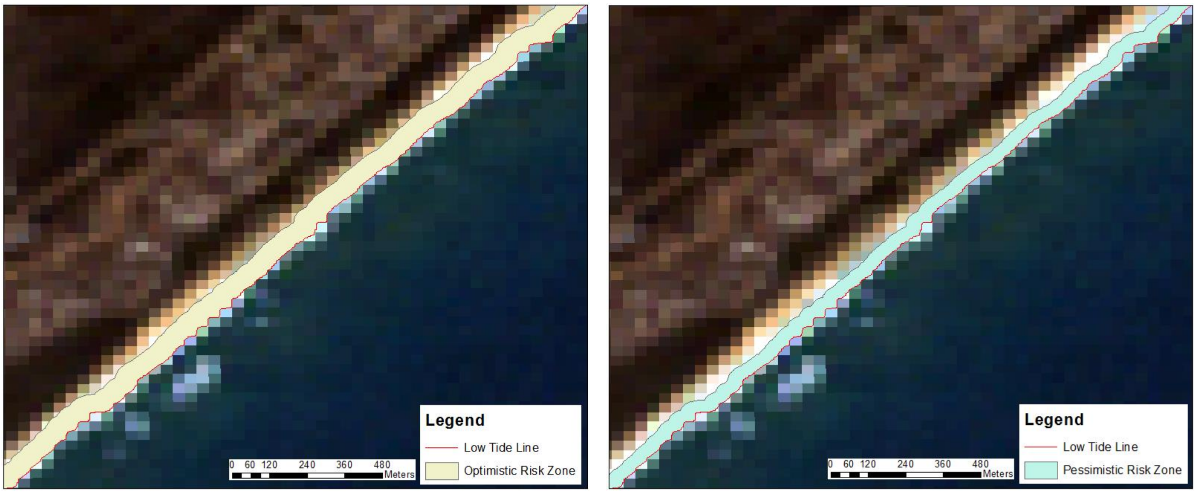


Fig. 9: Optimistic risk zone (left figure) and pessimistic risk zone (right figure)

3. Results and Discussion

Several areas were identified as water-eroded areas with the pessimistic risk zone. However, only two sites were identified as relatively obvious erosion (over 10 metres) with both optimistic risk zone and pessimistic risk zone. Fig. 10 shows those two eroded areas (site A and site B).

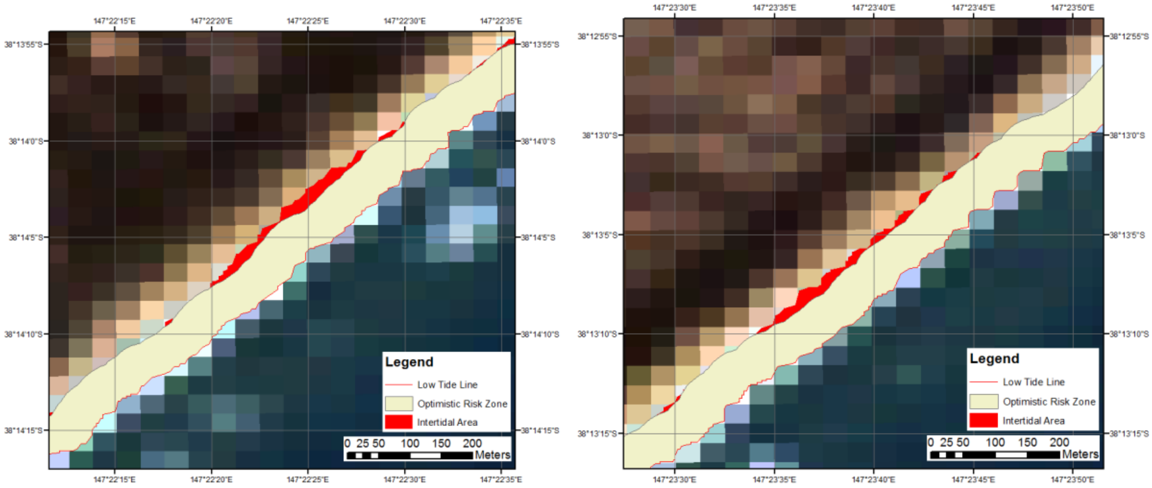


Fig. 10: Identified obviously eroded areas: site A (left) and site B (right)

Fig. 11 shows the change of shoreline over ten years observed from high resolution images for those two sites. Those images were collected from Google Earth software.

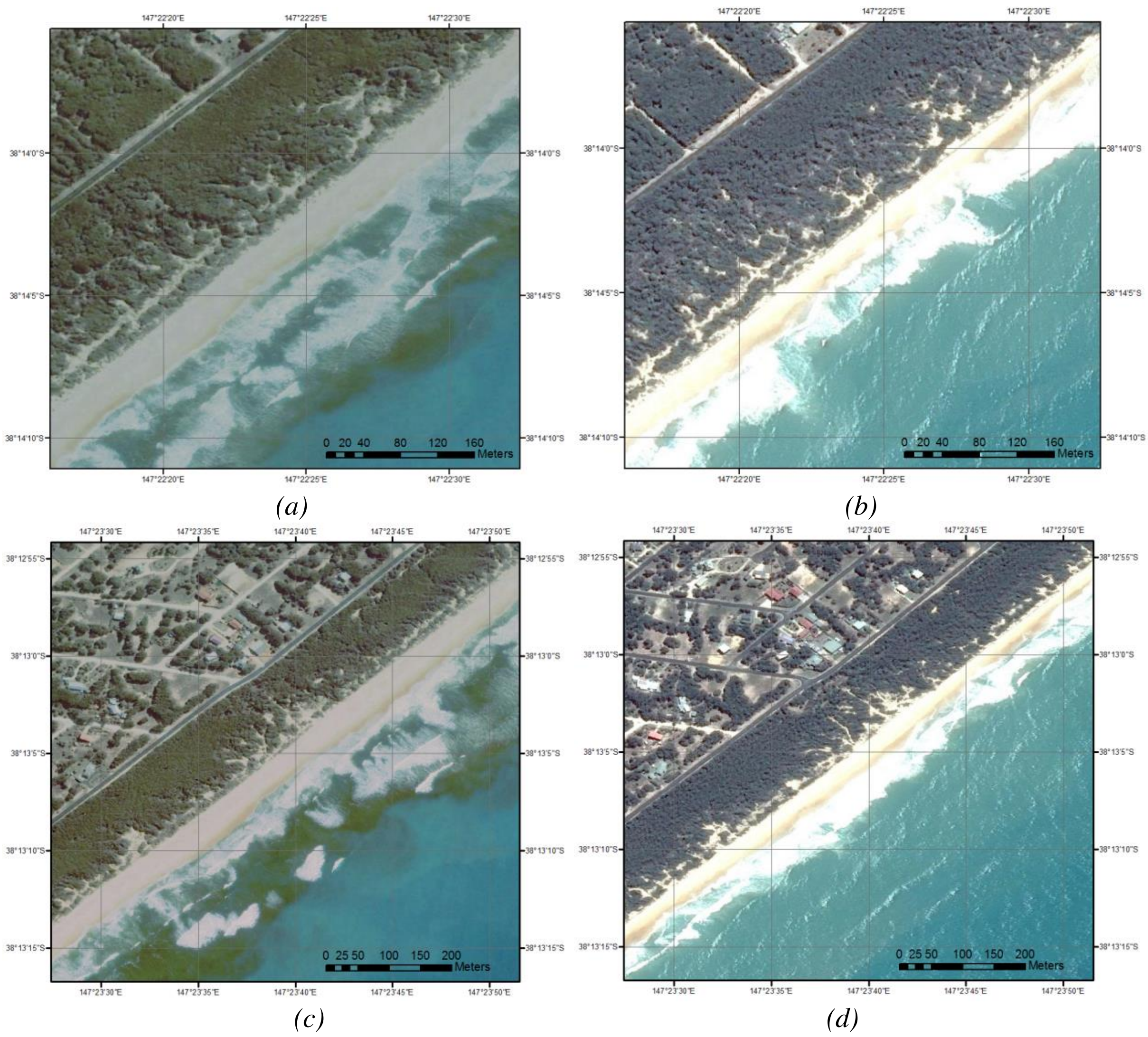


Fig. 11: Shoreline change over ten years: (a) site A on 7th April 2003 (b) site A on 6th October 2013 (c) site B on 7th April 2003 (d) site B on 6th October 2013

By visual comparison, it is not hard to find that the beach zones of the two sites became narrower and the shoreline shows a landward trend. Although the time span is longer than that of our experiment images, it shows the same trend with the detected results. However, even if we use Landsat images from year 2003 to 2013, the verification work will still be challenging. This is because we have no idea the exact imaging time (and the tide height at the imaging time) for the images from Google Earth, which means the instant boundaries for the high resolution images cannot be corrected to HWL using the tidal model. The daily position ambiguity of the shoreline can be much larger than its long-term change over years, for example, the largest intertidal distance on 7th April 2003 was 19m to 28m and on 6th October 2013 was 30m to 46m when the slope angle was between 1° to 1.5°, thus it is unknown whether the water-land boundary change is caused by daily tidal change or by long-term erosion.

Therefore, in the future work the results will firstly need to be verified with other high resolution reference data (e.g. GPS ground surveying data, high resolution satellite data) with more detailed information including the exact imaging time.

Besides, the endmember extraction is a key step in the whole processing since it determines the fraction of each selected ground feature and largely affected the sub-pixel classification result. The multiple indices method applied in the paper only considered three ground features, which is only an approximate method and remains to be improved. Moreover, this method requires the exploration of optimal thresholds. The authors are trying other methods such as learning from hyperspectral image unmixing methods to improve this step.

Lastly, it is worth trying to add other data source including radar images in the dataset to reduce the date interval as well as to improve the spatial resolution by image fusion.

4. Conclusion

In this paper, an integrated methodology is proposed, in order to detect the long-term coastal erosion in an efficient and economic method. Although the resolution of Landsat TM 5 images is 30m, the pixel size can be reduced to about 3m by spectral unmixing and sub-pixel mapping with the maximum decomposition factor 9. Even though the accuracy of the result was not quantitatively verified, the result shows a consistent trend with that of high resolution images from Google Earth. The proposed methodology shows the possibility and potential of using medium resolution images to detect long-term coastal erosion with improved resolution. Linear spectral unmixing, sub-pixel classification are the key techniques using spectral information to extract the water-land boundary in a sub-pixel level.

Acknowledgements

The authors would like to thank the U.S. Geological Survey for providing free Landsat TM5 images, which is the experiment dataset in the experiment.

References

[1] J. B. Shaw, M. A. Wolinsky, C. Paola, and V. R. Voller, "An image-based method for shoreline mapping on complex coasts," *Geophysical Research Letters,* vol. 35, p. L12405, 2008.

[2] S. K. May, R. Dolan, and B. P. Hayden, "Erosion of U.S. shorelines," *Eos, Transactions American Geophysical Union,* vol. 64, pp. 521-523, 1983.

[3] R. A. Morton and F. M. Speed, "Evaluation of shorelines and legal boundaries controlled by water levels on sandy beaches," *Journal of Coastal Research,* pp. 1373-1384, 1998.

[4] G. L. Smith and G. A. Zarillo, "Calculating Long-Term Shoreline Recession Rates Using Aerial Photographic and Beach Profiling Techniques," *Journal of Coastal Research,* vol. 6, pp. 111-120, 1990.

[5] H. F. Stockdonf, A. H. Sallenger Jr, J. H. List, and R. A. Holman, "Estimation of shoreline position and change using airborne topographic lidar data," *Journal of Coastal Research,* pp. 502-513, 2002.

[6] E. H. Boak and I. L. Turner, "Shoreline definition and detection: a review," *Journal of Coastal Research,* pp. 688-703, 2005.

[7] R. A. Holman, A. Sallenger, T. C. Lippmann, and J. W. Haines, "The application of video image processing to the study of nearshore processes," *Oceanography,* vol. 6, pp. 78-85, 1993.

[8] A. Puissant, S. Lefèvre, and J. Weber, "Coastline extraction in VHR imagery using mathematical morphology with spatial and spectral knowledge," in *Proceedings of The XXI Congress of International Society of Photogrammetry and Remote Sensing*, 2008, pp. 1305-1310.

[9] M.-q. CHEN, H.-l. SUN, X.-y. LI, and Z. WANG, "Coastline Detection and Extraction for Remote Sensing Image by Using Matlab," *Journal of Hubei University of Technology,* vol. 5, p. 020, 2011.

[10] H. Xu, "Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery," *International Journal of Remote Sensing,* vol. 27, pp. 3025-3033, 2006.

[11] S. McFeeters, "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features," *International journal of remote sensing,* vol. 17, pp. 1425-1432, 1996.

[12] F. Van Der Meer and S. De Jong, "Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of end-members," *International Journal of Remote Sensing,* vol. 21, pp. 2781-2797, 2000.

[13] N. Keshava, "A Survey of Spectral Unmixing Algorithms," *Lincoln Laboraroty Journal,* vol. 14, p. 24, 2003.

[14] J. Plaza, E. M. Hendrix, I. García, G. Martín, and A. Plaza, "On endmember identification in hyperspectral images without pure pixels: A comparison of algorithms," *Journal of Mathematical Imaging and Vision,* vol. 42, pp. 163-175, 2012.

[15] P. Atkinson, M. Cutler, and H. Lewis, "Mapping sub-pixel proportional land cover with AVHRR imagery," *International Journal of Remote Sensing,* vol. 18, pp. 917-935, 1997.

[16] K. Mertens, "Sub-pixel mapping: A comparison of techniques," Ghent University, 2008.

[17] G. Linlin, L. Xiaojing, W. Fan, and I. L. Turner, "Coastal erosion mapping through intergration of SAR and Landsat TM imagery," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, 2013, pp. 2266-2269.

[18] J. Shen, J. Zhai, and H. Guo, "Study on coastline extraction technology," *Hydrographic Surveying and Charting,* vol. 29, pp. 72-77, 2009.

[19] R. Freij-Ayoub, J. Underschultz, F. Li, C. Trefry, A. Hennig, C. Otto*, et al.*, *Simulation of coastal subsidence and storm wave inundation risk in the Gippsland Basin*: CSIRO Petroleum, 2007.

APPENDIX A

14ASRC
List of Presentations & Posters

List of Presentations

Abstracts are listed in alphabetical order of first author

Name of Presenter	Name of all authors	Title
Ruken Alac Barut	Ruken Alac Barut, Gulcin Ozurlan Agacgozgu, Linlin Ge, Jean Xiaojing Li	Investigation of Ancient Water Supply Systems with Geophysical VLF-EM Modelling and Remote Sensing
Simon Alexander	Simon Alexander, Kefei Zhang, Robert Norman	GPS Radio Occultation remote sensing for Antarctic atmospheric research
Jeremy Bailey	Jeremy Bailey, Lucyna Kedziora-Chudczer	Using astrobiology as the focus for a large online university course
Ian Bartlett	Ian S. Bartlett, Thomas Dixon, Amy J. Geddes, Dr. Elias Aboutanios	UNSW BLUEsat's Heavy Balloon Platform: A Versatile Bus for Student Near-Space Flights
James Bennett	J.C. Bennett, C. Smith, J. Sang, K. Zhang	Manoeuvring space debris for collision avoidance using ground-based lasers
Francis Bennet	Francis Bennet, Celine D'Orgeville, Yue Gao, William Gardhouse, Nicolas Paulin, Ian Price, Francois Rigaut, Ian Ritchie, Craig Smith, Kristina Uhlendorf, Yanjie Wang	Space environment management with ground based adaptive optic enhanced LIDAR
Philip Bland	P. A. Bland, M. C. Towner, J. P. Paxman, R. M. Howie, E. K. Sansom, M. Cupak, G. K. Benedix, S. J. Tingay, J. A. Harrison, K. A. Dyl, A. W. R. Bevan, J. A. Kennewell and M. J. Galloway	The Desert Fireball Network: A continent-scale observational facility serving the planetary science, space debris, and astronomy research communities
Kimberly Bott	Kimberly Bott, Lucyna Kedziora-Chudczer, Daniel Cotton, Jeremy Bailey	Deuterium on Uranus: a diagnostic for the planet's formation
Zahra Bouya	Z.Bouya, M. Terkildsen, M. Francis	Total Electron Content forecast model over Australia
Melrose Brown	Russell Boyce, Melrose Brown	The UNSW Canberra Space Initiative - Towards Routine Affordable In-Orbit Space Science
Joshua Brandt	Joshua Brandt, Barnaby Osborne, Naomi Tsafnat, Tracie Barber	The Breakup of Liquid Jets in Reduced Gravity Conditions
Andrew Brawley	Andrew Brawley	Silanna: Manufacturing in Australia High-Reliability Integrated Circuits for Space Applications
Iver Cairns	Iver Cairns, Vasily Lobzin, Alina Donea, Steven Tingay, Divya Oberoi, Michael Reiner, and Don Melrose	Solar magnetic Reconnection at Low Altitudes and Associated Type III Solar Radio Bursts and X-Ray Emission
Graziella Caprarelli	Graziella Caprarelli	Revisiting Mars planetary data sets: New insights
Brett Carter	Brett Carter, Robert Norman, John Retterer, Endawoke Yizengaw, Kyle Wiens, Keith Groves, Ronald Caton, Leo McNamara, Christopher Bridgwood, Matthew Francis, Michael Terkildsen and Kefei Zhang	Recent advances in understanding the drivers of day-to-day variability in the generation of scintillation-causing Equatorial Plasma Bubbles

Name of Presenter	Name of all authors	Title
Brett Carter	Brett Carter	Severe space weather events and their impact on our technology-dependent society
Sarah Chamberlain	S. Chamberlain , V. Wilquet, A. Mahieux, S. Robert, I. Thomas, A. C. Vandaele	SOIR/VEX observations of minor species in the Venus Mesosphere
Joon Wayn Cheong	Joon Wayn Cheong, Jinghui Wu, Nagaraj Shivaramaiah, Mazher Choudhury, Eamonn Glennon, Kevin Parkinson and Andrew Dempster	Real-Time Acquisition and Tracking of Galileo E1 Signals in Low Earth Orbit Scenarios
Jonathan Clarke	Jonathan Clarke	Arkaroola Mars Robot Challenge Expedition
Philip A. Clarke	Philip A. Clarke	Australian Aboriginal perceptions of space
Kimberley Clayfield	Kimberley Clayfield	CSIRO Astronomy and Space Science
Gavin Conibeer	Gavin Conibeer, Simon Chung	Challenges for manufacturing photovoltaic cells in-situ from minerals mined in space
David Cooper	P.D. Cooper	A Satellite Survey of Terrestrial Impact Features Across Outback Australia
William Crowe	William Crowe, Dr Nathan Kinkaid, Dr John Olsen, Mr John Page	Robotic swarms as means to autonomously and rapidly characterise small celestial bodies
Paul Curnow	Paul Curnow	Aboririginal Skies
Julie Currie	Julie Currie, Colin Waters, Murray Sciffer, Fred Menk	Pi 2 signatures in Australian Radar data
Luke Daly	L. Daly, P. A. Bland, K. A. Dyl, L. V. Forman, P. W. Trimby, S. Moody, and S. P. Ringer	High resolution analysis of tiny nuggets of platinum group elements in carbonaceous chondrites
Linda Davis	Linda M. Davis, Ying Chen	Double the Data for Satellite Relay Links
Linda Davis	Linda Davis	6S Success: Satellite Services, Systems, Spectrum, Software and Signals
Scott Dorrington	Scott Dorrington	Trajectory design for asteroid retrieval missions
Kerrie Dougherty	Kerrie Dougherty	Lift-off from Down Under: early proposals for equatorial launch facilities in Australia
Roger Dudziak	Roger Dudziak, Sean Tuttle	Harpoon Technology Development for the Active Removal of Space Debris
Kathryn A. Dyl	Kathryn A. Dyl, Philip A. Bland, James S. Cleverley, Chris G. Ryan	High-Resolution Mapping of Trace Elements in Primitive Meteorites: Unravelling the Record of the Early Solar System
Lewis Freeland	Lewis Freeland, Fred Menk	Ionospheric signatures of dayside reconnection at polar latitudes
Eamonn Glennon	Eamonn Glennon and Kevin Parkinson	The Kea V4.1 Spaceborne GPS Receiver
Ali Haydar Göktoğan	Ali Haydar Göktoğan, Steven Potiris, Anthony Tompkins	Navigating the Mars Yard: An Autonomous Path Planning and Tracking System for an Experimental Mars Rover
Ali Haydar Göktoğan	Ali Haydar Göktoğan	Design, Development and Operation of the Experimental Mars Rover Mawson
Mark Graham	Mark Graham	Operations Analysis of Australian Sensor Contribution to Space Situational Awareness
Mary Griffiths	Mary Griffiths, Sumen Rai	Getting serious about 'hacking' space-derived data: designing sustainable open data hackathons for maximum impact
Andrew Heitmann	Andrew Heitmann, David Holdsworth, Robert Gardiner-Garden & Michael Turley	Extraction of ionospheric heights using wide area land-sea maps from HF sky-wave radar

Name of Presenter	Name of all authors	Title
Teck Seng Ho	T. S. Ho, C. Charles, and R. Boswell	A plasma-generated electron beam as a neutraliser for ion thrusters
Steven Hobbs	J. D. A. Clarke and S. W. Hobbs	Mars'O Bot: A Mars Society Australia Robotics project
David Holdsworth	David Holdsworth, Michael Turley, Andrew Heitmann & Robert Gardiner-Garden	Wide-area land-sea mapping using high-frequency sky-wave radar
Jonathan Horner	Jonti Horner, James Gilmore & Dave Waltham	The role of Jupiter in driving Earth's orbital evolution
Robert Howie	R. M. Howie, J. P. Paxman, P. A. Bland, M. C. Towner, M. Cupák, E. K. Sansom	Development of the Automated Digital Fireball Observatory for the Desert Fireball Network
Garland Hu	Garland Hu, Leon Stepan	Magnetic Attitude Control of a 3U Cubesat with Large Deployables
Yilser Kabaran	Yilser Kabaran, Stephanie R. McArthur, Ian S. Bartlett, Dr. Nathan Kinkaid	BLUEsat Robotics to Mars: Lessons on real world testing for student projects
Lucyna Kedziora-Chudczer	Lucyna Kedziora-Chudczer	Cloudiness and composition of the atmospheres of the hottest "hot Jupiters"
Lucyna Kedziora-Chudczer	Lucyna Kedziora-Chudczer	Diversity of Planetary Atmospheres: New results and Open Questions
John Kennewell	John Kennewell	The asteriod hazard to Geosats
Champlian Kenyi	Champlain Kenyi, Jeremy Bailey and Daniel V. Cotton	Analysis of Greenhouse Gas Concentrations Retrieved from Observations with a Ground-based Spectrometer in the Near-infrared
Penny King	King, P.L., Berger, J.A., Izawa, M.R., Loiselle, L.M., Henley, R.W., Wykes, J.L., Gellert, R., Troitzsch, U., Moore, C.L., Hyde, B.C. Renggli, C.J.1 and Semmler, N.M.	Informing space science developments through simulating "volatile" materials and processes on planetary bodies
Isabelle Kingsley	Isabelle Kingsley, Craig Browne, James Oliver, Peter Mahony, Jennifer Fergusson, Carol Oliver	Using the Mars Lab and Project Based Learning to deliver authentic science experiences
Alexey Kondyurin	Alexey Kondyurin	Large space constructions directly cured in free space environment
Matthew Lau	John Page, Matthew Lau	Immersive virtual reality for space engineering applications
J. Henry Ledyard	J. Henry Ledyard	Designing for Space: Notes from a repair technician.
Seungho Lee	Seungho Lee, Linlin GE, Xiaojing Li	Natural colour composition with regression analysis using Landsat satellite image simulation
Liyuan Li	Liyuan Li, Linlin Ge, Xiaojing Li	Remote Sensing Image Classification of Environmental Hazards using Information from Multiple Perspectives
Youtian Liu	Youtian Liu, Linlin Ge, Xiaojing Li	Coseismic Deformation Inferred from DInSAR and Model Results for 2008 Sichuan and 2009 L'Aquila Earthquakes
Philippe Lorrain	Philippe Lorrain, Stefan Brieschenk, Russell R. Boyce	Effects of uncertainties in shock tunnel test conditions on the analysis of scramjet combustion experiments for access-to-space applications
Ken Lynn	Ken Lynn	The Relevance of Ionospheric Autoscaling Programs to Ionospheric Research
Helen Maynard-Casely	Helen Maynard-Casely	Probing the planets with crystallography

Name of Presenter	Name of all authors	Title
Shunsuke Miyazaki	Shunsuke Miyazaki, Eric Shear, Isaac DeSouza, Jagnnath Kshtriya	York-RMIT: Mars Inspiration Team Mission Design
Dean Mossveld	Dean Mossveld	Additive Manufacturing in Rocket Propulsion Applications
David Netherway	David Netherway, Robert Gardiner-Garden, Andrew Heitmann, Charlie Williams	Sporadic E based Antenna Model Gain Validation
Garry Newsam	Garry Newsam and Mark Graham	Key Performance Determinants for Sensors Conducting Surveillance of Space
Thanh Nguyen	Thanh Nguyen, Emma Bland, John Devlin, Darrell Elton, Guang Deng and Edhem Custovic	The Spectrum Difference Function technique for velocity cross-checking in TIGER application
Brett Northey	Brett Northey, Robert Gardiner-Garden, Andrew Heitmann & David Netherway	Producing a regional Real-time Ionospheric Model (RTIM)
Scott O'Brien	Scott O'Brien, Kevin Parkinson, Eamonn Glennon, Joon Wayn Cheong, and Andrew Dempster	Multi-Antenna Switching for Spaceborne GPS Receivers
Elias Aboutanios	Barnaby Osborne, Jendi Kepple, Naomi Tsafnat, Elias Aboutanios, Andrew Dempster	Design of a space compatible 3d printed satellite structure
Elias Aboutanios	Barnaby Osborne, Elias Aboutanios, Andrew Dempster	Impact of the Australian space hardware regulatory requirements on small satellite missions
Jonathan Paxman	J. P. Paxman, P. A. Bland, M. C. Towner, E. K. Sansom, M. Cupak	Fireballs in the Sky: Using Smartphone Technology to Enhance Citizen Reports of Fireball Observations
Tim Payne	Tim Payne	A Review of Space Programs and Associated R&D Developments in Defence
Lenard Pederick	Lenard Pederick	Empirical Climatology of the Sporadic E Layer (ECSEL) from Radio Occultation Measurements
Xiuying Peng	Xiuying Peng, Shijie Fan	Rapid Determination of Tidal Heights Based on GPS Precise Point Positioning
Li Qiao	Li Qiao, Barnaby Osborne, Andrew G Dempster	Design and simulation of attitude determination of UNSW ECO CubeSat using EKF
Li Qiao	Li Qiao, Ediz Cetin, Zhaoxu Hu, Andrew G Dempster	Prompt response for space collision assessment: Detection of Space Debris Using Space-borne Platforms
Shasidran Raj	Shasidran Raj, R Fleddermann, L E Roberts, R L Ward, A J Sutton, D M R Wuchenich, D E McClelland, D A Shaddock	Space Debris Tracking and Manoeuvring using Continuous Wave Lasers
William Reid	William Reid, Ali Haydar Göktoğan	A Model-Based Software Development Architecture for Control of Experimental Planetary Rovers
Mark Rice	Mark Rice, Gottfried Lechner, John Ophel	Australian Trials of Second Generation Distress Beacons over MEO Satellites
Josh Richards	Josh Richards	Laughing All The Way To Mars: Comedians & Poets in Space Science Outreach
Gemma Roberts	Gemma Roberts, Patrice Fey, Helen Brand	Alteration minerals in the Great Artesian Basin: Implications for weathering processes on Mars
Eleanor Sansom	E. K. Sansom, P. A. Bland, J. Paxman, M. C. Towner	Predicting a Meteoroid's Path to the Ground

Name of Presenter	Name of all authors	Title
Murray Sciffer	M. D. Sciffer, C. L.Waters, G. D. Sciffer and R. Lysak	Coupling of global magnetospheric wavemodes to field line resonances
Igor Shardakov	Irina O.Glot, Lyudmila A. Golotina, Victor N. Terpugov and Igor N. Shardakov	Thermo-Mechanics of Semi-Crystalline Polymers for space radiation protection: Theory and Experiments
Danielle Shean	Danielle A. Shean	Using the Advanced Science Institute to increase retention of high performing students
Jack Soutter	Jack Soutter, Jonti Horner, John Kielkopf, Leigh Brookshaw, Carolyn Brown, Brad Carter, Stephen Marsden, Mathew Mengel, Belinda Nicholson, Ian Waite	Research and Education at the University of Southern Queensland's Mt. Kent Observatory
Paul Stewart	P. N Stewart, P. D. Nicholson, P. G. Tuthill	Cassini for Space Based Stellar Spectra
Samira Tasnim	S. Tasnim and Iver H. Cairns	Generalized Theory for the Solar Wind
Paul G. ten Boom	Paul G. ten Boom	Doubting Dark Energy by way of a non-mundane Pioneer anomaly explanation
Viktor Terpugov	Viktor Terpugov, Denis Efremov, Viacheslav Chudinov, Irina Osorgina, Andrey Merzlyakov, Gennadii Bashin, Sergei Rusakov, Alexander Svistkov, Alexey Kondyurin	Stratospheric experiments on curing of composite materials
Martin Towner	M. C. Towner, P. A. Bland, M. Cupak, R. M. Howie, J. P. Paxman, E. K. Sansom, G. K. Benedix, S. J. Tingay, J. A. Harrison, M. J. Galloway and T. Jansen-Sturgeon	Data pipeline for a highly automated Desert Fireball Network
Anne Unewisse	Anne Unewisse, Manuel Cervera and Andrew Cool	Observations of a mesospheric bore over Edinburgh, Adelaide
Kehe Wang	Kehe Wang, David Neudegg, Colin Yuile, Campbell Thomson, Garth Patterson, Richard Marshall, Michael Terkildsen and Mike Hyde	The Space Weather Data Managed by IPS Radio and Space Services of Australia
Peter Ward	Peter Ward	The Rare Earth Hypothesis in 2014
Graeme Wren	Graeme Wren	Space Surveillance Telescope (SST)
Joel Younger	Joel Younger	Meteor Radar Including First Observations of the Camelopardalids Shower from Comet 209P/LINEAR
Joel Younger	Joel Younger	Atmospheric Density Measurements in the Upper Atmosphere Using Meteor Radar
Yunchao Zhang	Yunchao Zhang, Christine Charles and Rod Boswell	Ion beam in an annular helicon thruster
Kefei Zhang	Kefei Zhang	CRC-SEM - a new horizon of Australian space tracking research
Kefei Zhang	Zhang K., Norman R., Wu S., Le Marshall, Carter B., Yan S., Zhang S., Rohm W., Yan Y., Manning T.,Wang X., Choy S	Recent RMIT Quest for GNSS Atmospheric Sounding

Posters

Posters are listed in alphabetical order of first author

Name of Presenter	Name of All Authors	Title
Ruken Alac Barut	Ruken Alac Barut, Linlin Ge, Jean Xiaojing Li	Integration of InSAR with GPS and Geophysical Seismic Modelling
Prateek Badiger	Prateek Badiger, Lachlan Thompson, Thomas Baum	APAGPR: A space based ground penetrating RADAR system using CubeSat platform.
Graziella Caprarelli	Emily Bathgate, Helen Maynard-Casely, Graziella Caprarelli, Linda Xiao, Barbara Stuart, Kate Smith	Raman and XRD Study of Icelandic Tephra minerals: implications for Mars
Annalea Beattie	Annalea Beattie	Constitutional Draft for Lunar Settlement
Sanat Biswas	Sanat K. Biswas, Li Qiao, Andrew Dempster	Space-borne GNSS based orbit determination using a SPIRENT GNSS simulator
Ethan Chang	Ethan Chang, Linlin Ge, Xiaojing Li	Building Footprint Extraction from Airborne LIDAR Data
Bruna Contro de Godoy	Bruna Contro, Rob A. Wittenmyer, Jonti Horner, Jonty Marshall	The dynamical structure of the inner HR 8799 debris disk
Daniel Cotton	Daniel V. Cotton, Lucyna L. Kedziora-Chudczer, Kimberly Bott and Jeremy Bailey	Forward Modelling for Neptune's D:H Ratio – Preliminary Results
Lucy Forman	L. V. Forman, P. A. Bland, N. E. Timms, & G. K. Benedix	Impact-Induced Compaction in CV Chondrites: Exploring a Hidden Record with EBSD
Eamonn Glennon	Eamonn Glennon, Joon Wayn Cheong, Kevin Parkinson, and Andrew Dempster	Fast Cold-Start Acquisition of GPS Signals Using the Delay Doppler Map Accelerator
Ali Haydar Göktoğan	Ali Haydar Göktoğan	Mobility Performance Analysis of the Experimental Mars Rover Mawson
Zhaoxu Hu	Zhaoxu Hu, Li Qiao, Ediz Cetin, Andrew Dempster	A Simplified Model for Estimating Orbits of Small Space Debris
Alessandro Ippolito	A.Ippolito, C. Scotto, M. Francis, M. Layoun, M. Parkinson, D. Neudegg	IPS oblique ionograms automatic scaling
Champlain Kenyi	Champlain Kenyi, Jeremy Bailey and Daniel V. Cotton	Evaluation of a Ground-based Fabry-Perot Instrument for Total-Column CO2 and O2 Measurements in the Near-infrared
Seungho Lee	Seungho Lee, Linlin GE, Xiaojing Li	Hybrid Image Fusion using standard method
Lavender Liu	Qingxiang Liu, Linlin Ge, Xiaojing Li	Coastal Erosion Detection Using Multispectral Images with Accurate Tidal Model
Jonathan Horner	Belinda Nicholson, Leigh Brookshaw, Carolyn Brown, Brad Carter, Stephen Marsden, Matthew Mengel, Jack Soutter, Ian Waite, Jonti Horner	The "STARWINDS" Project: Space Weather Impacts On Planetary Systems
Kefei Zhang	R. Norman, K. Zhang, J. Le Marshall, W. Rohm, B. Carter and S. Alexander	The simulated impact of a severe troposphere weather event on GNSS signal propagation paths
Dale Potts	Dale Potts, Jan-Peter Muller, Ady James, Dave Walton, Edward Melina	Assessment of existing satellite capacity for methane gas seep detection from orbit

Name of Presenter	Name of All Authors	Title
Ali Remezan Nejad	Ali Remezan Nejad, Thomas Baum	Innovative Power Generation and Management Methods for a Space-Based Earth Observation System Using RADAR
William Reid	William Reid, Ali Haydar Göktoğan	Digging a Trench: Uncovering Martian Analogue Sub-Surface Regolith Using a Reconfigurable Robotic Rover
Alexey Kondyurin	Sergei Rusakov, Liudmila Komar, Alexander Svistkov, Viktor Terpugov, Alexey Kondyurin	A model: influence of mass loss on the curing process of epoxy composition in free space environment
Fubara Warmate	F.G. Warmate, Linlin Ge , Jean Xiaojing Li	Evaluation of neural network algorithms for implementing optimised application in image analysis and processing
Jonathan Horner	Jeremy R Wood, Jonathan Horner	The Dynamics of Centaurs in the 2:1 Mean Motion Resonance of Neptune