



The Statistical Society of Australia News

SSAI 

SSAI GOLDEN JUBILEE TRAVEL GRANT REPORT

In July I was fortunate enough to attend the International Conference on Robust Statistics (ICORS) hosted in St Petersburg, Russia. There were around 100 delegates – the small size of the conference resulted in a real sense of community, which made it a very enjoyable and stimulating week.

The academic program was lead with strong keynotes and impressive talks from invited speakers. I particularly enjoyed the presentations of Roger Koenker, Hira Koul and Yanyuan Ma.

Mention needs to be made of a number of young researchers in robust statistics (robusticians? robustniks?) demonstrating the depth of the field and continued interest in this important area of statistics. In particular Viktoria Oelleler (student of Christophe Croux) presented some interesting and promising work on regression S-estimates in the case when you have cellwise contamination (as opposed restricting attention to contaminated observations). Also, William Aeberhard (student of Stephane Heritier) presented a very nice talk on robust approaches to negative binomial modeling motivated by a study into Parkinson's disease where a few subjects were outlying in the sense that they experienced many more falls than usual.

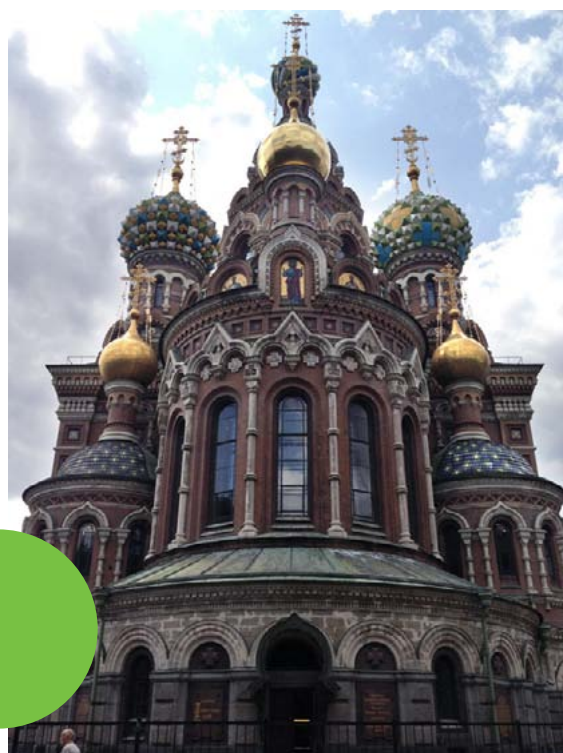
[> Continued on page 5](#)

In this issue

Editorial	2
Events	4
President's Column	6
Member news	7
MPE 2013 Conference	12
CSIRO takes charge of Australia's information overload	17
From the SSAI Office	21
NSW Branch	24
QLD Branch	26
SA Branch	27
VIC Branch	29
WA Branch	32



Garth Tarr



September 2013
Issue 144

SSAI

PO Box 213, Belconnen ACT 2616

We are located on the ground floor of
ABS House, room GN 311.

Phone 02 6251 3647

Fax 02 6251 0204

Email eo@statsoc.org.au

Website www.statsoc.org.au

Editors

Alice Richardson

School of ISE, University of Canberra

Michael Adena

Datalytics

Correspondence

Please direct all editorial

correspondence to Alice Richardson

Email eo@statsoc.org.au

Disclaimer

The views of contributors to this
Newsletter should not be attributed to
the Statistical Society of Australia, Inc.

Subscriptions

The Newsletter of the Statistical
Society of Australia is supplied free to
all members of the society. Any others
wishing to subscribe to the newsletter
may do so at an annual cost of
A\$30.00 for an issue of four numbers.

Advertising

Advertising will be carried in the
Newsletter on any matters which
the Editors feel are of interest to the
members of the Society.

For details of advertising rates, etc.
contact the SSAI Executive Officer at
eo@statsoc.org.au

DEADLINE FOR NEXT NEWSLETTER
10 November 2013

SSAI 

EDITORIAL

The International Year of Statistics is now more than half way through, and if the articles contributed to this newsletter are any indication, it's a busy year in the Branches and organisations that employ statisticians. National Mathematics Day, organized by the Australian Association of Mathematics Teachers, was held on 16 August and had a statistics focus in honour of IYS. If you know of or participated in any activities for [National Mathematics Day](#), the Editors would love to hear from you.

Just before the deadline of our previous issue (31 May - 2 June 2013), the extremely popular [GovHack](#) 2013 competition took place, with over 1000 participants Australia-wide and 130 competition entries. There were several statistics-related prizes, including Best Analysis awarded by CSIRO and Best Data Visualisation highlighting South Australia awarded by CIO South Australia. The prize-winning solution was a web app that aims to build awareness of the risk of early childhood development vulnerability by postcode and suggest risk-minimising interventions and strategies. The app is called "IMPROVING EVERY CHILD'S CHANCE IN LIFE". The statistical solution was implemented in R.

And as the year continues to unfold, the National Committee for Mathematical Sciences has been working hard on the [Decadal Plan for the Mathematical Sciences](#). Their website indicates that the Steering Committee will be identifying major strategic actions and recommendations in August, ready for a workshop in December 2013.

This issue we also welcome the first report of incoming National President, John Henstridge. In other Society news, a new look for the Society's website will be launched soon, so keep an eye on the [SSAI](#) website for a fresh, user-friendly site.

Finally, I'd like to encourage you to click on the links, not only here in the editorial but also in the Branch reports and other newsletter articles. Our correspondents supply the links in consultation with speakers. The links are an integral part of the newsletter, they enrich the content, and there are many delightful surprises to be found. Society members are also contributing to other forums and we rely on you to bring these to everyone's attention. [Ron Sandland](#) takes issue with a writer's advice "Calculus. Trust me. You'll never use it." [Brownwyn Harch](#) encourages her staff to introduce themselves not as mathematicians (statisticians) but as "working on global food security using maths" (stats). Click and enjoy!

Alice Richardson



and **Michael Adena**



SSAI CENTRAL COUNCIL

Executive Committee

President: John Henstridge
Secretary: Doug Shaw
secretary@statsoc.org.au

Branch Presidents and Branch Secretaries

Canberra

President: Ray Lindsay
Secretary: Warren Müller
secretary.actbranch@statsoc.org.au

New South Wales

President: Scott Sisson
Secretary: Ryan Defina
Ryan.Defina@abs.gov.au

Queensland

President: Helen Johnson
Secretary: Tania Patrao
Tania.Patrao@menzies.edu.au

South Australia

President: Richard Woodman
Secretary: Paul Sutcliffe
suffers@bigpond.net.au

Victoria

President: Lyle Gurrin
Secretary: Sandy Clarke
sjclarke@unimelb.edu.au

Western Australia

President: Anna Munday
Secretary: Ryan Admiraal
R.Admiraal@murdoch.edu.au

SECTION CHAIRS

Bayesian Statistics

Chair: Scott Sisson
Scott.Sisson@unsw.edu.au
Assistant Chair: Jannah Baker
Jannah.Baker@qut.edu.au
<http://www.statsoc.org.au/bayesian-statistics.htm>

Environmental Statistics

Chair: David Clifford
David.Clifford@csiro.au
<http://www.statsoc.org.au/environmental-statistics.htm>
Assistant Chair: Mayukh Samanta
Mayukh.Samanta@qut.edu.au
<http://www.statsoc.org.au/environmental-statistics.htm>

Social Sciences

Chair: Michele Haynes
M.Haynes@uq.edu.au
Assistant Chair: Jegar Pitchforth
Jegar.Pitchforth@qut.edu.au
<http://www.statsoc.org.au/social-statistics.htm>

Statistical Education

Co-Chair: Michael Martin
Michael.Martin@anu.edu.au
Co-Chair: Peter Howley
Peter.Howley@newcastle.edu.au
Assistant Chair: Su Yun Kang
s7.kang@qut.edu.au
<http://www.statsoc.org.au/statistical-education.htm>

Surveys and Management

Chair: Stephen Horn
Stephen.Horn@fahcsia.gov.au
Assistant Chair: Charisse Farr
A.Farr@qut.edu.au
<http://www.statsoc.org.au/surveys-and-management.htm>

Biostatistics

Lyle Gurrin
lgurrin@unimelb.edu.au
Assistant Chair: Nicholas Tierney
Nicholas.Tierney@gmail.com
<http://www.statsoc.org.au/medical-statistics>

Section for International Engagement

Mark Griffin
m.griffin@adasis-oz.com
<http://www.statsoc.org.au/IntEngagementSection>

Young Statisticians' Network

Chaturi Bhaskaran
u4681540@gmail.com
<http://www.statsoc.org.au/about-young-stats.htm>

Further contact details for Society Secretaries and Section Chairs can be obtained by contacting the Society on (02) 6251 3647

EVENTS

CPD43: ADVANCED ANALYSIS OF LINKED DATA

25-29 September 2013, Adelaide

INSTRUMENT DESIGN AND TESTING

26-27 September, Sydney

COGNITIVE INTERVIEWING

8 October, Sydney

INSTRUMENT DESIGN AND TESTING

10-11 October, Melbourne

INTERNATIONAL CONFERENCE ON STATISTICAL DISTRIBUTIONS AND APPLICATIONS (ICOSDA13)

10-12 October 2013, Mt Pleasant, MI, USA

SYMPOSIUM ON BAYES FOR OFFICIAL STATISTICS

16 October, Canberra

LONGITUDINAL STUDY OF AUSTRALIAN CHILDREN (LSAC) AND LONGITUDINAL STUDY OF INDIGENOUS CHILDREN (LSIC) RESEARCH CONFERENCE

13-14 November 2013, Melbourne

2013 NZSA CONFERENCE

24 – 27 Nov 2013, University of Waikato, New Zealand

BIOMETRICS BY THE CANALS

1-5 December 2013, Mandurah, WA

20TH INTERNATIONAL CONGRESS ON MODELLING AND SIMULATION (MODSIM2013)

1-6 December 2013, Adelaide

33RD WORKSHOP ON BAYESIAN INFERENCE AND MAXIMUM ENTROPY METHODS IN SCIENCE AND ENGINEERING («MAXENT 2013»)

15 - 20 December 2013, Canberra

FOURTH BIENNIAL INTERNATIONAL STATISTICAL ECOLOGY CONFERENCE (ISEC2014)

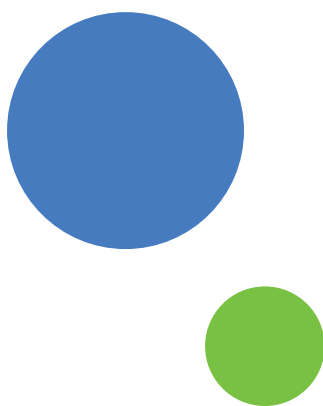
1-4 July 2014, Montpellier France

AUSTRALIAN STATISTICAL CONFERENCE (ASC2014)/IMS ANNUAL MEETING

7-10 July 2014, Sydney

ISBA 2014 - TWELTH WORLD MEETING OF ISBA

14-18 July 2014, Cancun, Mexico



My presentation was the final talk on the program - I was worried that people would not hang around until the end but that wasn't the case. Fortunately there was only one session running and I had a large captive audience. In my talk I described an approach to estimating inverse covariance matrices by combining a pairwise approach with a regularisation technique – a side effect of this particular regularisation technique is that the resulting inverse covariance matrix is guaranteed to be positive definite, even if the input matrix, a symmetric matrix whose elements are pairwise robust covariances, is not. An advantage of this approach is that it can cope with a high level of contamination, in particular if the contamination is scattered in a cellwise manner. After the talk I received positive feedback from a number of people including Mia Hubert and Peter Rousseeuw who commented that it appeared to be a useful contribution to an important problem.

In addition to the busy academic program, the social program was also packed. One evening we had a boat tour on the canals after which we watched them raise the bridges on the Neva. An afternoon was spent touring the gardens and palace at Peterhof just outside of St Petersburg. This was a great opportunity to casually chat with a number of eminent researchers including Elvezio Ronchetti and Maria-Pia Victoria-Feser.



As I was already in Europe, I also took the opportunity to present at the European Meeting of Statisticians at the end of July. The slides of both talks can be found here: <http://www.maths.usyd.edu.au/u/garht/>

I'd like to thank SSAI for helping get me to Europe to share my research and strengthen ties with international researchers. I'm hopeful that the connections I made will lead to fruitful collaboration and gainful employment in the near future!

Garth Tarr

If you are interested in applying for the SSAI Golden Jubilee Travel Grant, please turn to page 23 to find out how!



PRESIDENT'S COLUMN

The Future of Statistics in Australia

This is my first President's Column since starting as President of the Statistical Society of Australia. Taking over from Kerrie Mengersen and her predecessors as President, I have been set a high standard to work to. On behalf of all members of the Society, I thank Kerrie for all her work over the past two years and I look forward to continuing to work with her in her role as Vice president.

A particular achievement of the past two years has been the revision of the Society's strategic and operational plans, a process led by Kerrie. These plans clearly present a number of challenges for the Society, and I have no doubt that members will work to implementing them over the next two years. I find that the membership of the Society maintains a strong belief in the value and importance of statistics and the role of the Society. I was amazed at the support given to the Society several years ago when we had our financial problems. Similarly, the passion expressed in forums such as ANZSTAT indicates an active community. This gives me hope that the Society will be able to more than just meet the challenges we face.

While the statistical community in Australia has much to be proud of, sometimes statistics appears under threat. The membership of our Society is less than it was when I joined forty years ago, while Australia's population has almost doubled. Statistics groups in our universities, along with other mathematical sciences, have shrunk in size. Without wishing to be pessimistic, I want to highlight one specific challenge.

Both in Australia and internationally statistical work is often done by non-statisticians. This is not necessarily a problem in itself – our Society prides itself in being open to anyone with a genuine interest in statistics – except that it is often done poorly. A recent column by Marie Davidian, President of the American Statistical Association, was titled "Aren't We Data Science", highlighting that some developments in the world of big data are bypassing statisticians, a loss to both sides. In Australia, we have a National Committee for Data in Science, under the auspices of the Academy of Science, that contains no input from or representation of statisticians.

I don't believe there is any intention to exclude statisticians; rather there is a lack of consciousness of what statisticians can offer. This invisibility of statisticians affects other areas as well, particularly recruitment to our profession. Statisticians must take some of the blame for this – it is not enough to claim that we have a right to be included simply because we are who we are. It is essential for statisticians to be highly visible in promoting our discipline and profession. It is also necessary for statisticians to appreciate the issues of large datasets and how best to harness the computational power now available to us.

This challenge is one shared by the other mathematical sciences in Australia. It is timely that a Decadal Plan for the Mathematical Sciences is currently being prepared through the National Committee for the Mathematical Sciences. To me, statistics is the most applicable part of modern mathematical science and the Society has the opportunity for providing leadership in areas such as professional recognition and encouraging more students to study the mathematical sciences.

Finally, if members wish to contact me about these or other matters, please do. I cannot guarantee a quick response, but I will try to reply when I can.

John Henstridge

john@daa.com.au



John Henstridge

MEMBER NEWS

The SSAI would like to congratulate VIC Branch member Dr Duncan Ironmonger for being recognised in the 2013 Queen's Birthday Honours List for significant service to economics as a researcher, author and academic.

We would also like to congratulate NSW Branch member Professor Matt Wand on receiving the Hannan Medal for his long term work as well as VIC Branch member Dr Aurore Delaigle on having been awarded the Moran Medal as an early career researcher. Their citations (<http://science.org.au/awards/awardees/2013awards.html>) read:



Matt Wand

'Matt Wand's main research focus is non-linear statistical models and methodology for high-dimensional and complex data, in the face of rapid technological change. Much of this research incorporates ongoing developments in Machine Learning. His contributions are multifaceted and involve applications, theory, methodology and publicly available software. Whilst most of Wand's research is generic, areas of application that have driven some his research include public health, computational biology and the natural environment.'

'Dr Delaigle's has made influential contributions to contemporary statistical problems, including deconvolution, regression with measurement errors, functional data analysis, analysis of high dimensional data, group testing, and a wide variety of contributions to function estimation. She is remarkably adept at transforming complex and highly abstract methods into easy to understand concepts, and at developing fully applicable techniques that work in a wide variety of settings. An advantage of her approach to developing methodology is that her techniques apply at once to many practical problems, in both the biological and physical sciences.'



Aurore
Delaigle

Australian Statistical Conference in conjunction with the Institute of Mathematical Statistics Annual Meeting ASC-IMS 2014 CONFERENCE

7 – 10 July 2014
Australian Technology Park, Sydney

PROGRAM UPDATE
www.asc-ims2014.com



On behalf of the Statistical Society of Australia and the Institute of Mathematical Statistics, the organising committee invites you to register in the joint Australian Statistical Conference/IMS Annual Meeting, to be held 7–10 July 2014 in Sydney, Australia.

Delegates from all areas of statistics will join with world class Australian and International statisticians and mathematicians to develop, network and share their knowledge and expertise. In 2014 the Statistical Society of Australia will hold its biennial ASC in conjunction with the IMS Annual meeting. The Conference will provide opportunities for presentations on a wide range of topics and recognises the role that statistics plays in all aspects of modern life.

KEYNOTE SPEAKERS

ASC Keynote Speakers

James Brown, University of Southampton
Adrian Baddeley, University of Western Australia
Sheila Bird, Cambridge University
Rob Tibshirani, Stanford University

IMS Keynote Speakers:

Thomas G. Kurtz, University of Wisconsin-Madison
Peter Donnelly, University of Oxford
Terry Lyons, University of Oxford
Nina Gantert, Technische Universität München
Martin Hairer, University of Warwick
Timo Seppalainen, University of Wisconsin-Madison
Matthew Stephens, University of Chicago
Harrison Zhou, Yale University

MARK THE KEY DATES IN YOUR DIARY:

**Abstract Submission
Deadline:**
30 OCTOBER 2013

Author Notification:
30 NOVEMBER 2013

Early Bird Deadline:
28 FEBRUARY 2014

The conference objectives are to:

- Attract world class statisticians to share their knowledge and expertise
- Inform delegates about new work and developments in statistics, probability and mathematical statistics
- Provide an opportunity for professionals from all of these aforementioned areas to network, present and discuss ideas

Topics of interest include but are not limited to: spatial statistics, Bayesian statistics, computational and asymptotic statistics, sample surveys, methodology for official statistics, stochastic/statistical modelling, biostatistics, multivariate statistics, probability, mathematical statistics, econometrics and financial statistics.

The venue for this meeting is the Australian Technology Park in Sydney.

On behalf of the Program Committee and the Local Organizing Committee, we invite you to join us in Sydney for this exciting scientific event. Your participation will ensure that the 2014 ASC-IMS Conference will be a memorable meeting.

CALL FOR PROPOSALS

You are invited to submit an abstract for consideration for a contributed oral or poster presentation.

As this conference is a joint meeting between the Statistical Society of Australia and the Institute of Mathematical Statistics, an extensive and wide-ranging program will be available. As benefiting an event of this size, with approximately 12 Keynote presentations and 6 parallel streams, a large portion of the program will be by invitation. However, a substantial part of the program will be set aside for contributed presentations, both oral and poster. While there is no restriction on the topic or number of contributed presentations, the number of oral presentations is by nature limited. We encourage participants to submit their abstracts from May 2013.



ADDRESS FOR COMMUNICATIONS

Conference Managers



arinex pty limited
ABN 28 000 386 676

ASC-IMS 2014 Conference Managers

Address:
Level 10, 51 Druiett St,
Sydney NSW 2000, Australia
Ph: +61 2 9265 0700
Fax: +61 2 9267 5443
Email: asc-ims2014@arinex.com.au
Website: www.asc-ims2014.com

Visit the website for further updates: www.asc-ims2014.com



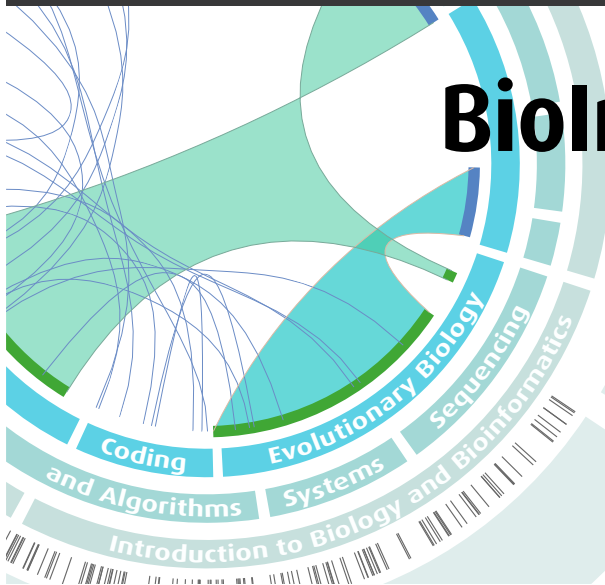
BioInfoSummer 2013

University of Adelaide 2–6 December 2013

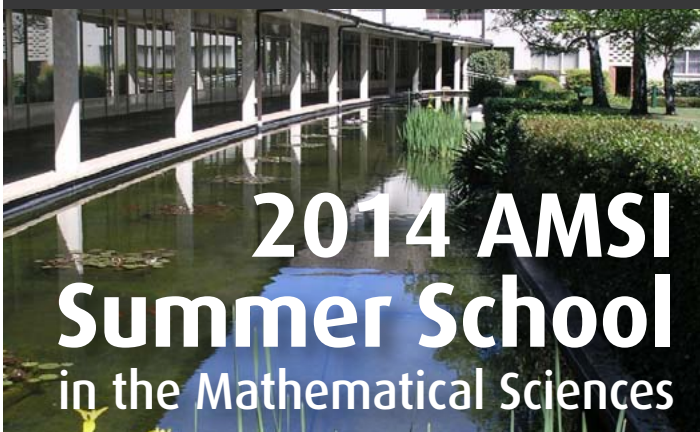
BioInfoSummer introduces students, researchers and others working in related areas to the bioinformatics.

The program features:

- Introduction to Biology and Bioinformatics
- Evolutionary Biology
- Systems Biology
- Next Generation Sequencing
- Coding and Algorithms for Bioinformatics



Register now: www.amsi.org.au/BIS



6–31 January 2014

Australian National University

7 Reasons to Attend:

- **Learn** from Australia's leading mathematicians and statisticians
- **Choose** from a wide range of courses to suit your specialty
- **Gain** credit towards your degree
- **Meet** future employers at the Careers Afternoon
- **Build** your networks at dinners, BBQs and special events
- **Broaden** your knowledge base with advanced coursework
- **Discover** the latest subjects in your discipline

Full travel and accommodation scholarships available!



Apply today: www.amsi.org.au/SS



In your 3rd year of maths & stats?
Find out what it's like to be a researcher

2013/14 AMSI Vacation Research Scholarship

Get paid to complete a six-week summer research project under the guidance of a supervisor!

Scholarships:

- Australia wide
- \$450/week for six-weeks
- Present at CSIRO's Big Day In
- Fully funded travel and accommodation
- Open to intending honours and masters students



Apply today: www.amsi.org.au/VRS

BIOMETRICS by the Canals



The International Biometric Society, Australasian Region Conference

The Sebel Mandurah Hotel, Western Australia, December 1-5, 2013

Photographed by Rexness

INVITATION

On behalf of the Australasian Region of the International Biometric Society we invite you to join us for our biennial conference which will be held in Mandurah, Western Australia, from 1-5 December, 2013. Share your successes and failures with other professional statisticians working in the biosciences, including agriculture, biomedical science and public health, bioinformatics, ecology, environmental sciences and forestry. Themes for the conference will be:

- Spatial and temporal statistics
- Linear mixed models
- Complex genetic mixed models
- Design of experiments
- Generalized linear and additive models
- Bayesian methods

The four-day conference, 2-5 December, will be preceded by two one-day short courses on Sunday, 1 December:

- Statistics for Spatio-Temporal Data* given by Prof. Noel Cressie
- Vector Generalized Linear and Additive Models* given by Dr. Thomas Yee

Further information will become available on the conference website:

<http://www.BiometricSociety.org.au/conferences/Mandurah2013>

Kind regards

The Local Organising Committee & the Scientific Programme Committee

INVITED SPEAKERS

Prof. Noel Cressie
(co-sponsored by NIASRA)
NIASRA, University of Wollongong
Spatio-temporal smoothing

Thomas Yee
Auckland University
Generalized linear models

Prof. Ian James
Murdoch University
Host-viral interactions

Prof. Christine Müller
Technische Universität Dortmund
Prediction of growth processes

Prof. Alan Welsh
Australian National University
Selecting linear mixed models

Ric Coe
World Agroforestry Centre
Option x Context interaction and design of multi-environment trials

Registrations and abstracts open 3 June 2013

CONFERENCE LOCATION

Mandurah is located 72 kilometres south of Perth on the edge of the Peel-Harvey estuary, an idyllic inland waterway of over 130 square kilometres, twice the size of Sydney Harbour, and blessed with abundant marine life and natural beauty. Recently declared Australia's fastest growing regional destination, the coastal town of Mandurah has become a vibrant tourist attraction with a multitude of activities and entertainment throughout the year. It has delightful restaurants and cafes, crystal clear blue water, and an abundance of activities for the whole family, Mandurah makes for a great getaway!

CONFERENCE VENUE

The conference will be held at **The Sebel Mandurah Hotel** which is located on the stunning Mandurah Estuary with spectacular water views and close to the attractions of Mandurah.



TRAVEL & ACCOMMODATION

Travel options from Perth to Mandurah will be publicised on the conference website at the time of registration

SOCIAL PROGRAMME

A conference is not complete without the opportunity to socialise and network with colleagues, providing a break from the more formal parts of the programme.

A welcome party and conference dinner will be the main social functions. In addition one afternoon will be set aside for local sightseeing and other activities.

MPE 2013 CONFERENCE

This year is not only the international year of Statistics, but it has also been dedicated to Maths of Planet Earth (MPE). In Australia the MPE 2013 conference took place at Rydges Hotel in Melbourne from 8 - 12 July. It was an action packed five days with an impressive array of plenary and contributed presentations, public lectures, round table discussions and a poster session Tuesday evening. The conference was very well organised by AMSI and was supported by several organisations. SSAI participated in this important event and sponsored the "Data-Based View of Our World" theme. Session organisers were Kerrie Mengersen and Bronwyn Harch. Wearing my hat as SSAI MPE 2013 representative I helped chair the parallel sessions.

SSAI invited two excellent plenary speakers: John Cook from the Global Change Institute at the University of Queensland, who gave a captivating talk about *The challenges of communicating the reality of climate change* with advice and insight into climate change deniers; and Bronwyn Harch, Chief of CSIRO's Mathematics, Informatics and Statistics division, who engaged the audience with her interesting talk about *Data Science: informing insights, understanding and decision making relevant to Planet Earth*. Both SSAI plenaries were very well received by conference attendees and had us all spellbound.

The three *Data-Based View of Our World* parallel sessions ran on Monday afternoon and all day Wednesday. They were well attended and there were some excellent presentations. In the first parallel session, our SSAI president, Kerrie Mengersen, did not disappoint with a fascinating whirlwind world tour of Bayesian Statistics in her talk *Bayesian statistics for planet earth*. The talks of the other two presenters in the session, who were from Monash University, complemented each other: Yanfei Kang spoke about *A shape extraction and classification method for atmospheric time series*; and Danijel Belusic posed the question 'Is geometry above physics in atmospheric boundary-layer flow?' in his talk entitled *A data-driven bottom-up approach for understanding the dynamics of the atmospheric boundary layer*.

> Continued on next page



John Cook,
plenary speaker
at MPE2013



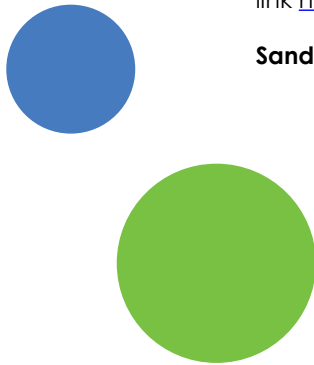
Bronwyn Harch,
plenary speaker
at MPE2013

In the second parallel session on Wednesday morning the three presenters were: Stuart Mead from Macquarie University who gave us an insight into his interesting research into volcanoes, which is sponsored by CSIRO and Risk Frontiers, in his presentation entitled *Determining change points in completeness of historical event databases*; Sandra Johnson from Queensland University of Technology who talked about *Developing a sustainability scorecard for Dairy Australia*; and Murray Aitkin from the University of Melbourne who gave an intriguing talk resulting from joint research with the University of Lancaster into modelling social networks, entitled *Statistical modelling of a terrorist network*.

The final SSAI parallel session also had three presenters. The first two were Stephen Horn (SSAI treasurer) from the Department of Families, Housing, Community Services and Indigenous Affairs who presented joint work with Raymond Czaplewski of US Forest Service State discussing *Space methods – a new paradigm for official statistics*; Lesley Wyborn from Geoscience Australia who told us about the Virtual Geophysics Laboratory, a joint project with CSIRO, in her talk called *Virtual laboratories: enabling reproducibility of collaborative scientific research*. The final presentation of the SSAI parallel sessions was by Jason Whyte from the University of Melbourne, who gave an entertaining and thought provoking talk on *Choosing between models for a time series? Data points us in the right direction!* A very apt note on which to end the SSAI sessions.

Our thanks go to all who participated in the SSAI sessions, to the MPE 2013 conference organisers, and the many other sponsors and contributors who made this conference a huge success. If you missed the chance to attend this important event, details of the MPE 2013 conference with links to the conference program, photos and presentations can be found at the following link <http://mathsofplanetearth.org.au/events/2013/>

Sandra Johnson



“The Mathematical Sciences in 2025” a report of the US National Research Council

Nalini Joshi

July 5, 2013

This article is published with the kind permission of the Australian Mathematical Society and has appeared in the AMS's "Gazette".

The US National Academies Press published the report “The Mathematical Sciences in 2025” earlier this year¹. This was a serendipitous discovery, as it happened around the same time as the call for submissions to the Decadal Plan for Mathematical Sciences in Australia. The report makes eight recommendations, which are listed below. Whilst the funding for and support for mathematical sciences is very different between USA and Australia, we share similar concerns, hopes and ambitions. In this column, I would like to ask you to consider the report’s recommendations and whether they are equally important for the Australian context.

The concerns of this report overlap with many of the themes of the Decadal Plan. I was pleased to read statements such as these:

In order for the whole mathematical sciences enterprise to flourish long term, the core must flourish. This requires investment by universities and by the government in the core of the subject.

(from the Summary, p. 2)

The overview to the report describes the vitality of the US mathematical sciences as “excellent”. Chapter 4 reports that there are eight institutes of mathematical sciences funded by the National Science Foundation in the US, not including various others such as the Clay Mathematics Institute, the Simons Center for Geometry and Physics and the Kavli Institute for Theoretical Physics. Such sentiment and level of funding is very different to the current state of concern expressed by many of us about the mathematical sciences in Australia. A major issue of concern, for example, is the decrease in the number of departments of statistics in Australia. Two of the recommendations below refer to “mathematics and statistics departments” in the US. It is sobering to reflect that in Australia there are only two departments of statistics left from about twenty or so that operated in the mid 1990s.

The report contains six chapters in addition to the summary and six appendices. The actions are encapsulated in eight recommendations, which are quoted below along with additional information (in italics) needed to make them self-contained for this column². As you read each recommendation, I would like to encourage you to try imagining the corresponding actions in Australia, with commensurate bodies such as the Australian Research Council replacing the National Science Foundation. What would your reactions to such recommendations be if they were to appear in the Decadal Plan? I would love to know your thoughts. (Please feel free to email me.)

- Recommendation 3-1 (p.68): The National Science Foundation should systematically gather data on such interactions (*i.e., data on graduate student training that crosses discipline boundaries with mathematics*) – for example, by surveying departments in the mathematical sciences for

¹This report can be downloaded freely, after registration, from the National Academies Press at http://www.nap.edu/catalog.php?record_id=15269

²Each recommendation has a number referring to the chapter in which it is made. I have also supplied the page number on which each recommendation can be found.

the number of enrollments in graduate courses by students from other disciplines, as well as the number of enrollments of graduate students in the mathematical sciences in courses outside the mathematical sciences. The most effective way to gather these data might be to ask the American Mathematical Society to extend its annual questionnaires to include such queries.

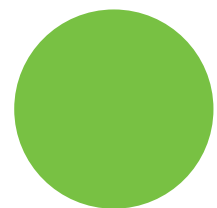
- Recommendation 3-2 (p.69): The National Science Foundation should assemble data about the degree to which research with a mathematical science character is supported elsewhere in the Foundation. (Such an analysis would be of greatest value if it were performed at a level above DMS (*i.e.*, *Division of Mathematical Sciences*).) A study aimed at developing this insight with respect to statistical sciences within NSF is under way as this is written, at the request of the NSF assistant director for mathematics and physical sciences. A broader such study would help the mathematical sciences community better understand its current reach, and it could help DMS position its own portfolio to best complement other sources of support for the broader mathematical sciences enterprise. It would provide a baseline for identifying changes in that enterprise over time. Other agencies and foundations that support the mathematical sciences would benefit from a similar self-evaluation.
- Recommendation 5-1 (p.127): Mathematics and statistics departments, in concert with their university administrations, should engage in a deep rethinking of the different types of students they are attracting and wish to attract and must identify the top priorities for educating these students. This should be done for bachelors, masters, and Ph.D.-level curricula. In some cases, this rethinking should be carried out in consultation with faculty from other relevant disciplines.
- Recommendation 5-2 (p.127): In order to motivate students and show the full value of the material, it is essential that educators explain to their K-12 and undergraduate students how the mathematical science topics they are teaching are used and the careers that make use of them. Modest steps in this direction could lead to greater success in attracting and retaining students in mathematical sciences courses. Graduate students should be taught about the uses of the mathematical sciences so that they can pass this information along to students when they become faculty members. Mathematical science professional societies and funding agencies should play a role in developing programs to give faculty members the tools to teach in this way.
- Recommendation 5-3 (p.127-8): More professional mathematical scientists should become involved in explaining the nature of the mathematical sciences enterprise and its extraordinary impact on society. Academic departments should find ways to reward such work. Professional societies should expand existing efforts and work with funding entities to create an organizational structure whose goal is to publicize advances in the mathematical sciences.
- Recommendation 5-4 (p.137): Every academic department in the mathematical sciences should explicitly incorporate recruitment and retention of women and underrepresented groups into the responsibilities of the faculty members in charge of the undergraduate program, graduate program, and faculty hiring and promotion. Resources need to be provided to enable departments to monitor and adapt successful recruiting and mentoring programs that have been pioneered at many schools and to find and correct any disincentives that may exist in the department.

- Recommendation 5-5 (p.144): The federal government should establish a national program to provide extended enrichment opportunities for students with unusual talent in the mathematical sciences. The program would fund activities to help those students develop their talents and enhance the likelihood of their pursuing careers in the mathematical sciences.
- Recommendation 6-1 (p.152): Academic departments in mathematics and statistics should begin the process of rethinking and adapting their programs to keep pace with the evolving academic environment, and be sure they have a seat at the table as online content and other innovations in the delivery of mathematical science coursework are created. The professional societies have important roles to play in mobilizing the community in these matters, through mechanisms such as opinion articles, online discussion groups, policy monitoring, and conferences.

The report goes into much more detail in its two hundred or so pages. There are also focused topics encapsulated in “Boxes” . For example, Box 5-1 “Mathematical Circles: teaching students to explore” describes a longstanding, engaging initiative in mathematical outreach. The appendices describe past strategic studies, such as the David report of 1990 which “led to striking increases in federal funding for mathematical sciences for a few years” (p. 155-6). Chapter 2 “Vitality of the Mathematical Sciences” provides a very thoughtful description and exploration of current and recent advances and strengths in the mathematical sciences, ranging from the topology of three-dimensional spaces to compressed sensing.

In the Decadal Plan, we need actionable goals as well as aspirational statements. For example, if Recommendations 5-1, 5-3 and 5-4 were adopted in the Decadal Plan, I would have urged the addition of specific pathways that would fund and enable academic departments to implement “deep rethinking”, rewards for outreach and recruitment initiatives.

In the Australian context, we also need actions that will bring mathematical sciences to the level of vitality that the USA and other western countries enjoy. My hope is that after the first Decadal Plan is put into action, the next one will be able to start from a baseline of revitalised mathematical sciences in Australia.



CSIRO TAKES CHARGE OF AUSTRALIA'S INFORMATION OVERLOAD

On 6 August 2013, CSIRO announced the appointment of Dr Bronwyn Harch as chief of its newest research division with a focus on data-driven computational and information sciences.

The new research division, CSIRO Computational Informatics (CCI), will work in partnership with the organisation's National Research Flagship program to address key national challenges across the information and decision making value chain. It will also work closely with a range of external partners and collaborators to create a capability hub in key research areas including next generation data analytics, autonomous robotics, complex systems modelling, and decision making under uncertainty.

"From pioneering the development of digital image analysis for agriculture in the 1980s to securing the landmark patent to Wireless LAN technology in the 1990s, CSIRO has always been at the forefront of technology-related research which tackles the nation's toughest challenges," said Dr Harch.

"The proliferation of smart devices and increasing access to next generation broadband has caused an explosion in the volume, velocity and variety of data and information. With predictions, that by 2020 the average person will own six different smart devices connecting us to over 37 billion 'things', from cows in the field to our car to our fridge door through the Internet, it is clear the amount of data we produce will continue to grow at an exponential rate," said Dr Harch.

"We have responded to the information and data challenges facing Australia with the formation of this new CSIRO Computational Informatics (CCI) Division. Our integrated and strengthened capabilities in this division will enable us to remain at the forefront of global developments in key research areas that transform the information and decision making workflows of industry, government and the innovation sectors."

A recent McKinsey Institute report on the future of 'disruptive technologies' shows a reduced focus on new gadgets and gizmos and an increased emphasis on the impact of technologies which require advanced data analytics. The same report predicts that by 2025 the potential economic contribution of new disruptive technologies such as mobile Internet, advanced robotics and 3D printing are expected to return between \$14 trillion and \$33 trillion globally each year.

Dr Harch believes the new Division's focus on data-driven science will help CSIRO to increase the economic and social well-being of the nation by tackling some of the biggest challenges facing Australia including our declining productivity and ageing population.

Dr Harch believes the new Division's focus on data-driven science will help CSIRO to increase the economic and social well-being of the nation by tackling some of the biggest challenges facing Australia including our declining productivity and ageing population.

"Through our Preventative Health Flagship, we are already leading the way in transforming advanced data analytics to pinpoint the genes that could lead to a simple blood screening test for Alzheimer's disease before it takes hold. We're also working on another project, with our Digital Productivity and Services Flagship and Australian Centre for Broadband Innovation (ACBI), which leverages next generation broadband networks to monitor health data



Bronwyn Harch

and help older Australians to live in their own homes longer, independently and safely."

However, it's not just the healthcare industry reaping the rewards from this explosion of data. In an impressive debut, research from the CSIRO Computational Informatics Division has earned a staggering 13 nominations across multiple categories from environmental science to the services industries at this year's iAwards, Australia's premier technology accolades.

These projects include:

- **Zebedee**: a handheld 3D mapping device which doesn't rely on GPS and can be used underground or inside buildings.
- **ReMoTe**: allows a remote expert to guide in real-time an onsite technician using gestures and voice to increase productivity and assist in up-skilling the workforce.
- **Smarter, Safer Homes**: a platform which uses home sensors and video conferencing to support older Australians to live in their own homes safely and independently.
- **Environmental Linked Data**: which produced Australia's first set of open-linked long-term climate data.
- **Sense-T Aquaculture Decision Support System**: a platform which helps to ensure public food safety by supplying the Tasmanian Shellfish Quality Assurance Program with real-time environmental awareness to simplify decisions regarding when to close shellfish farms.
- **Patient Flow Research**: analytics to predict requirements and optimise configurations for hospital beds.
- **Automated Disease Grading and Clinical Decision Support**: an automatic disease grading and clinical decision support system for diabetic retinopathy level grading.
- **Computer-aided Ocular Biomarker Suite**: an eye test for early detection of Alzheimer's disease 15-20 years before diagnosis is usually possible.
- **SensorDB Virtual Laboratory**: used for storing, analysing and looking for patterns in distributed sensor data for plant phenomics in the field. This work received two Australian Capital Territory categories iAwards.
- **Medtex**: a software platform that unifies the language across narrative medical reports and extracts clinically relevant information to aid decision support.
- **Lung Anatomy Trainer (eLANt)**: a web-based program that uses an online lung anatomy and nomenclature training tool for teaching the complex and variable human lung anatomy seen from a fibre-optic inspection tool.
- **Starbug**: Jeremy Breen, a University of Tasmania and CSIRO student, won a postgraduate tertiary award for autonomous analysis of marine sediment chemistry using an underwater robot.

The winners of this year's iAwards were announced during a ceremony in Melbourne on 8th August 2013 (www.iawards.com.au).

For more information about the CSIRO Computational Informatics Division please visit www.csiro.au/ci.

Read more media releases in CSIRO's [Media](#) section.



OUR INSURANCE GIVES YOUR BUSINESS A CRITICAL BIAS

As the appointed Insurance Brokers to the SSAI we are proud to offer their members a truly unique Professional Indemnity Insurance Policy: one that will cover your particular requirements perfectly and save you money.

As the preferred Professional Indemnity Insurance Broker for SSAI and it's Members we have brokered an exclusive arrangement with Chubb Insurance Australia Ltd, who are a market leader in specialty insurance coverages, which are uniquely designed for specific industries.

Benefits to SSAI Members:

1. Competitive premiums -- in a majority of cases, we guarantee to offer a cheaper premium than you paid last year, subject to minimum premium levels.
2. One of the broadest policy wordings available, including:
 - a. Full cover for Contractors, not just vicarious
 - b. Loss includes punitive and exemplary damages
 - c. Built in "contingent" bodily injury/property damage coverage
 - d. Worldwide coverage includes "duty to defend" anywhere in the World
 - e. "Formal investigations" cover available to the full Policy Limit
 - f. 100% of Defence Costs paid within the Limit even if only part of actual Loss if covered
 - g. Auto Reinstatement of Limit.

When it comes to finding an insurance broker, do you look for the widest choice of cover, the best value for money or an adviser that really understands your specific needs? Usually finding more than one of these qualities is rare. But IAA is in the unique position of offering you all three.

Our Commitment to You

As your appointed General Insurance Broker, we will:

- Conduct an analysis of your risk/needs and recommend the most appropriate covers to protect your assets and liabilities
- Provide alternatives on your risk retention through various excess options
- Allow you to pay premiums in monthly instalments
- Provide prompt and timely servicing of your accounts to ensure you are adequately informed of issues that may affect your business
- Provide competitive quotations for any other classes of insurance you may have.

In the event of a claim, we will:

- Provide you with the insurers claim form if needed
- Arrange for the loss assessor to be appointed if applicable
- Advise on the completion of the claim form
- Provide advice to ensure you have received your full entitlements.

As one of Australia's biggest General Insurance Brokers, our knowledge of the insurance industry is second to none. Our experience in handling insurance for many thousands of customers enables us to provide the best and most current advice. In other words, at IAA, we're committed to making insurance work for you.

CALL: (02) 9964 7303

Mike Aurisch

ANZIIF (Snr. Assoc) Dip Fin Serv (Brok) QPIB
Authorised Representative of Insurance Advisernet Australia Pty Limited
ABN 15 003 886 687 | AFSL No 240549 | AR No 403930
Level 31, 100 Miller Street, NORTH SYDNEY NSW 2059
T (02) 9964 7303 | M 0439 562 009 | F (02) 9954 1809 | E maurisch@iaa.net.au



OVERSEAS STATISTICIANS VISITING AUSTRALIA

We have an "Overseas Visitors" page on the SSAI website (<http://www.statsoc.org.au/OverseasVisitors>). The aim of this page is to provide a public database with the names of overseas visitors, giving other organisations the opportunity to benefit from the visit as well. If you or your organisation think that they would like to work with one of the visitors listed on the website, simply send an email to the SSAI office, explaining the details of your proposal, and the office will forward your email to the visitor in question.

If you know of statistical experts from overseas planning a visit to your organisation, please advise the SSAI by sending an email containing the name, details and travel dates of the visitor to eo@statsoc.org.au. We would also need the visitor's email address (not to be published on the website), so that we can ask for his or her permission to put their name up, as well as the name and details of a contact person in Australia.

YOUNG STATISTICIANS

Exciting news! On the 1st July 2013, Chaturi Bhaskaran became Chair of the SSAI Young Statisticians. Already, she is doing a wonderful job.

I wanted to express my appreciation to all the wonderful people I was privileged to work alongside for my two-year term, particularly the Branch YS representatives and YSC2013 Organising and Assisting committees. You are all incredible people!

I would also like to acknowledge the support for Young Statisticians provided by the SSAI Branches and Executive. Their commitment to invest in Australia's future statistical leaders is just one of the reasons SSAI has a bright future.

I am eagerly anticipating the future achievements of Young Statisticians under Chaturi's guidance!

Susanna Cramb

LOOKING FOR A JOB?

For a listing of current statistical vacancies in Australia and New Zealand visit:

<http://www.statsci.org/jobs>

Do you have a job to advertise on the website?

Email a position description to eo@statsoc.org.au. Listing is free!

35% SSAI Member Discount Promotion with Wiley

The SSAI members receive a special discount of 35% on online purchases with Wiley or Wiley-Blackwell (<http://www.wiley.com/statistics>). This offer excludes school and textbooks.



To activate this benefit, please go to <http://au.wiley.com/WileyCDA/Section/id-410891.html>.

Register on the SSAI/Wiley Landing Page and from then on you can purchase Wiley books at the 35% discount without use of a Promotional Code. For those members outside of Australia and New Zealand, please use the Promo Code SDP92 at checkout to activate the 35% discount. Members who registered on the Wiley site before the SSAI link was available may encounter problems when trying to get the discount. To get the 35% discount these members unfortunately need to register using another email address.

Toll free phone (from within Australia only) 1800 777 474

Toll free phone (from New Zealand only) 0800 448 200

Other overseas phone + 61 7 33548455

E-mail: custservice@johnwiley.com.au

FROM THE SSAI OFFICE

There is usually a lapse of about 3 weeks between the newsletter editors' meeting and the SSAI newsletter going out to the members. As I write my bit for the newsletter today, it gives me a great uplift knowing that, while at the moment I still have to warm my hands by holding a cup of tea, by the time the SSAI members read this, we will be well into Spring.

Winter is always a busy time at the SSAI Office with preparations for the SSAI and ASPAI Central Council AGMs as well as the SSAI and ASPAI AGMs in full swing. This year the latter will be held in Melbourne at the end of August. I always enjoy the opportunity to come face to face with our members. Working in an office on your own can be quite isolating and my visits to the AGMs and SSAI conferences are always a highlight for me. This year we will meet up with members of the VIC Branch before the AGMs to give them the opportunity to speak with representatives of the SSAI Executive Committee. This is a great chance for the members to give feedback to the Executive Committee and for the Executive Committee to hear what members want from their professional association. A study on professional associations recently conducted and prepared by Survey Matters, Australia found that

A stable or increasing demand for professional associations is predicted by an overwhelming majority of members.

The increase in demand is more likely to be forecast by young members – despite industry concerns around difficulties recruiting and retaining members in the younger age brackets.

The main reason for many members joining is for access to relevant information - they are looking to their association to be the leading, authoritative source of news about their profession.

Members want quality professional development and accreditation programs that enhance their career prospects and protect the reputation of their profession.

Advocacy is critical to members, yet only 63% of members are satisfied with their associations efforts in this area

Age, but more importantly career stage, plays a big part in how members want to receive information and interact with other members.

This benchmarking report is a helpful tool in kick-starting a discussion with our own members about their concerns. A member survey is at the top of the Executive Committee's to do list to give you a chance to voice your thoughts on your membership with SSAI.

The SSAI Office has also been following the "Scrap the Cap" initiative initiated as a result of the former Federal Government's proposal to introduce a \$2000 cap on tax deductibility of self education expenses. SSAI shared the view of many other organisations that such a tax would significantly impact the training and professional development of all Australians. The tax-exempt amount was too low to allow taxpayers to undertake any useful education, especially if overseas travel was required. It decided to get involved with the "Scrap the Tax" initiative, because it wants to support and encourage members to maintain their professional standards. The new measure was due to take effect on 1 July 2014. However, the former Government bowed to widespread opposition to

> Continued on next page



Marie-Louise Rankin

its planned imposition and in its pre-election Economic Statement, released on 2 August, the then treasurer, Chris Bowen, announced that introduction would be deferred until 1 July 2015 to allow for consultation on the impact of the change. If you or your organisation wish to support the "Scrap the Cap" initiative, you can do so by signing a petition or telling your story at <http://www.scrapthecap.com.au/>.

Before I close I want to let you know that the Accreditation Committee is currently looking for volunteers to join the committee. Accreditation Committee members normally hold the qualification of Accredited Statistician and will serve for three years. If you are interested in becoming a member of the Accreditation Committee, please email me at eo@statsoc.org.au and I will forward your expression of interest to the Committee Chair. Would you please note that the Accreditation Committee aims to be as representative as possible of the Branches and interests of the Society, the statistical profession and of areas of statistical expertise and not all applications to join it will be successful.

And while I am on the subject of volunteers: At the last Membership Officers' meeting it was decided that SSAI is in need of an in depth analysis of its membership database over the last years. This suggestion was taken to the Executive Committee and the committee members proposed finding a student member or member of the Young Statistician's Network willing to take this on. I will supply the data and hopefully we'll find a volunteer who can make sense of the figures. If you think that volunteer could be you, please email me at eo@statsoc.org.au.

Marie-Louise Rankin

SSAI GOLDEN JUBILEE TRAVEL GRANT

to provide overseas travel funds to SSAI student members, who can prove consecutive SSAI membership for a minimum of two years.

Last year the SSAI introduced a travel grant that offers limited travel funds to assist *student members* of the SSAI to attend overseas conferences at which they present a paper or poster.

A maximum of \$1000 is available per application, limited to a single trip during the course of the student's studies. Students will not be supported in their first year of study and will have had to be members of the Society for at least 2 years prior to the application deadline. Applications are required to be lodged in advance of travelling. In exceptional circumstances an application can be for post-conference support, but the application will then have to be made within 1 month of returning and the 2 year mandatory membership period prior to departure must still be met. Exceptional circumstances are limited to unforeseeable student out of pocket expenses arising from other funding sources not fulfilling their obligation or changes to the trip that could not have been avoided.

A complete application will consist of

- Information on the conference and its importance to student's work (2-3 lines)
- Details of the paper/s/poster student wants to present at the conference
- A list of other funds sought or promised, including student's home institution
- Student's out of pocket expenses expected
- Any other supporting material student feels is necessary
- A letter of support SIGNED by one of student's supervisors AND student's Departmental Head
- Student's CV



The application deadline is 31 March 2014.

If successful the student member is required to produce original receipts for amounts of equal or greater value than the grant. These receipts will be returned to the student marked with how much has been reimbursed. The student will therefore still be able to use the receipts for proof of attendance or to claim any funding shortfall from other organisations. The student member will also need to supply a report of his or her involvement in the conference to be published in the SSAI newsletter. This report should confirm the actual travel details and papers presented.

Recipients of the grant are asked to acknowledge the SSAI's support in the presentations and in any published version of the paper.

One travel grant is available per year. Assuming that more than one application will be received per year, either the Executive Committee or a special committee would help with the selection process.

For more information or to apply, please contact the SSAI Office (eo@statsoc.org).

With this travel grant program the SSAI seeks to underline its objective to further the study, application and good practice of statistical theory and methods in all branches of learning and enterprise. It has been implemented to confirm to members that the SSAI is willing to support student statisticians and their budding careers.

NSW BRANCH

Confidentiality and the use of confidential data in research

On May 22, 2013, Dr Christine O'Keefe from the Division of Mathematics, Informatics and Statistics at the CSIRO presented a seminar entitled, "Confidentiality and the use of confidential data in research". Dr O'Keefe Leads the Program in Business and Services Analytics at CSIRO. She began her presentation by talking about the increasing demand by for data that can be used for social and economic research and analysis. The challenge for data custodians such as the Australian Bureau of Statistics is to provide these data in such a way that does not violate legal and moral obligations to protect privacy and confidentiality. For example, the 1905 Census and Statistics Act which states that "... results or abstracts ... shall **not** be published or disseminated in a manner that is likely to enable the identification of a particular person or organization." Similarly, the federally funded Population Health Research Network has a mission to provide linkable, de-identified data for health research

It is easy to understand the importance of protecting privacy & confidentiality. Individuals don't want others to have access to sensitive data related to their health, employment or financial situation. Businesses want to be protected against releasing commercially sensitive material. However, ensuring confidentiality is easier said than done! Christine gave a couple of interesting examples illustrating how it is not enough to simply delete names and addresses from a database. It can be surprisingly easy to identify people based on just a few attributes. Christine then went on to compare and contrast some of the approaches that various agencies are using to assure the privacy and confidentiality of individuals in the datasets that they release for research purposes. Such approaches include de-identification strategies, using secure and virtual data centres, applying statistical disclosure control strategies, using remote analysis servers and finally, generating synthetic data. There are advantages and disadvantages to each of these approaches and in general there is a tradeoff between level of assured confidentiality and the amount of information that can be made available to researchers and policy makers. Christine spent some time describing how remote servers work. The concept here is that researchers or analysts never have access to the raw data, but only the results of analyses run on those data via a special server. An important aspect of such systems is that protections can be put in place to ensure that users don't try to "trick" the system into providing confidential data. CSIRO has developed a system called Privacy Preserving Analytics that can be used to create such an analysis framework. The advantage of such approaches is that they allow users access to data that would otherwise be too sensitive for release. There are interesting technical challenges in putting the necessary protections in place to ensure that analysis results don't inadvertently disclose private or confidential information. Interested readers should contact Christine at Christine.O'Keefe@csiro.au



Christine
O'Keefe

Louise Ryan

NSW BRANCH

Methodology Architecture – A Roadmap for New Methodological Directions in the ABS

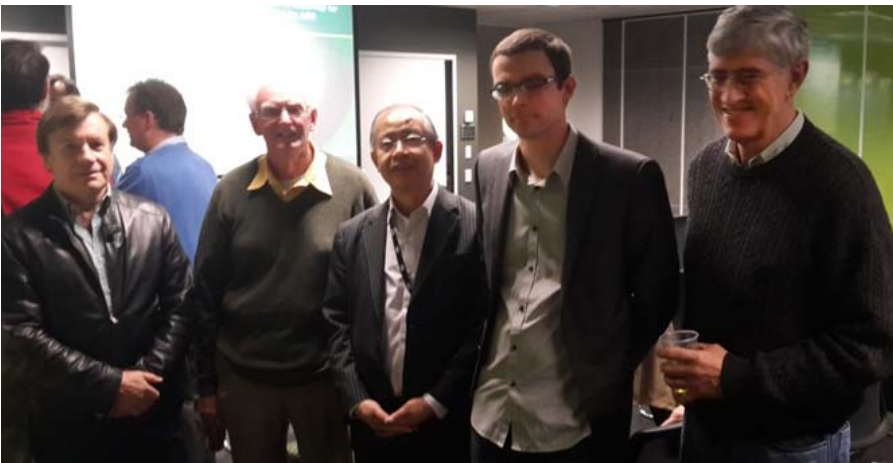
For July the NSW Branch were off to the ABS for a seminar by First Assistant Statistician and Chief Methodologist, Dr. Siu-Ming Tam. The topic of Methodology Architecture highlighted some of the key challenges being faced by the ABS in remaining relevant within a changing world. There was discussion about the use of 'plug and play' modules, industrialisation of methods and innovation within official statistics. Approaches once foreign to the field such as 'nowcasting', pattern recognition and the utilisation of big data were also raised as interesting future directions for the agency.

It wasn't all serious though, Siu-Ming provided some insightful (and often humorous) examples of occasions where approaches have and haven't worked. There were also some fun stories of his time in Hong Kong as a young statistician trying to make his way in academia, experiences submitting to journals, and how to do a PhD in one weekend!

Overall the evening was a lot for fun, and we thank both Siu-Ming and all attendees for making it a success.

Ryan Defina

Dr. Siu-Ming Tam addressing those in attendance



Dr Siu-Ming Tam with A/Prof. Scott Sisson, Dr Murray Cameron, Dr Nicholas Fisher and guest



QLD BRANCH

Visualising Big Data – Antony Unwin

The Queensland Branch had a double feature for the July meeting. The evening's talks began by joining the SSAI National Access Grid seminar series to hear Trevor Hastie's presentation on Sparse Linear Models. There were a few technical glitches; however, branch members and guests were able to see and hear most of Trevor's talk. This was followed by a talk on Visualising Big Data by Antony Unwin from the University of Augsburg in Southern Germany. Antony's research interests include large datasets, interactive statistical graphics and the development of statistical data visualisation software. He is the first professor of computer-oriented statistics and data analysis at the university and is co-author of two texts on graphical methods: "Graphics of large datasets" and "Handbook of data visualization". His presentation focused on the challenges of analysing large datasets and the contribution that can be made by graphical procedures to all stages of data analysis from initial exploration to understanding results from statistical models. Antony discussed some of the challenges to standard graphical procedures posed by large datasets and how these procedures could be amended to accommodate large datasets. He emphasised the importance of interactivity for data visualisation and gave a series of interesting demonstrations using the free software called Mondrian. Links to an array of graphics programs, including the Mondrian freeware, can be found on Antony's website at <http://stats.math.uni-augsburg.de/~unwin/>.



Antony Unwin

Elaine Pascoe

SA BRANCH

Imputation of skewed continuous data: advances from the use of Normal imputation

The well attended May meeting of the SA Branch welcomed Dr Murthy Mittinty from The University of Adelaide presenting on 'Imputation of skewed continuous data: advances from the use of Normal imputation'. Murthy's current interest in this area is motivated by work with the Australian Early Development Index (AEDI) – a collection of scales that measure children's development as they enter school. The scales are heavily positively skewed, but low scores in the long tails are particularly of interest as these scores represent the more vulnerable children. The total sample size in this data set is over 17,500, with up to 700 cases missing on particular scales. Murthy outlined the work he is currently doing to improve imputation of these skewed distributions.

After reviewing current practices in univariate and multivariate imputation based on the joint normal distribution, Murthy introduced methods specifically for univariate skewed distributions. This included Tukey's gh distribution (based on transformations of estimated standard normal deviates), beta and Weibull distributions, and Fleishman polynomials. In presenting the results of his simulation studies, Murthy illustrated that normal-based imputation often resulted in shifted and increased peaks in the imputed distributions and increased variability. This suggests caution is needed when interpreting results of normal-based imputation on potentially skewed variables. Imputation based on the beta distribution did a better job at preserving the peaks and tails of the original skewed distribution.

Extensions to the multivariate missing case are less well described, particularly the beta distribution based method. Murthy presented his current work in this area, specifically considering the case of missingness in multiple outcomes. Murthy detailed his suggested multiple imputation Bayesian approach to this problem, fitting marginal beta distributions to estimate the shape and scale parameters, and then using a Clayton copula to estimate the correlation between the two outcomes and hence derive their joint distribution. The resulting bivariate beta copula is then used to generate the required missing values. Murthy presented the results of testing this method on skewed HDL and LDL cholesterol data from the Aboriginal Birth Cohort. Results from his simulation studies showed that the method appears to preserve the tails of the distributions, even at 30% missingness.

In conclusion Murthy stressed that his message was not to discourage the use of multivariate normal methods, but rather to complement them with other methods. His future work will look at evaluating the multivariate beta method on the AEDI dataset and extensions for multivariate beta regression. The presentation was followed by much lively discussion on the technical challenges of multiple imputation as well as 'selling' imputation to our sometimes sceptical clients.

Kylie Lange

Clustering of High-Dimensional Mass Spectrometry Data and Applications to Ovarian Cancer

The speaker for the July meeting of the South Australian Branch was Lyron Winderbaum, a PhD student studying at Adelaide University. The topic of Lyron's talk was 'Clustering of High-Dimensional Mass Spectrometry Data and Applications to Ovarian Cancer'.

In the first part of his talk Lyron provided some background to proteomics, the large-scale study of proteins. Proteins are sequences of amino acids produced from the 'instructions' encoded in the genes of the living organism. The technology of mass spectrometry (MS) enables the characterisation of proteins in a sample by ionising the proteins, and ions then being separated according to their mass-to-charge ratio (m/z ratio). Due to the specialised ionisation process used (MALDI), ionised proteins can be fragmented through well-understood pathways. Inference can be drawn on the proteins present from detection of the mass-to-charge ratios of their fragment ions.

The data discussed in this talk was from a technology called Imaging Mass Spectrometry (IMS). Lyron explained how the ovarian cancer data of interest needs to be analysed by IMS instead of MS because the cancerous tissue is contained within a tissue sample consisting of multiple tissue types. The raw IMS data is processed into a set of observations where each observation corresponds to a single pixel in a high-resolution image. Each observation is a high-dimensional vector of binary variables representing peaks within fine intervals of mass-to-charge ratios in the sample spectrum.

Lyron demonstrated how the k-means clustering method is able to effectively distinguish the different tissue types in his ovarian cancer sample data. The pixels were classified into four clusters according to their mass spectra, with one cluster representing tissue affected by cancer. In the example the clusters were shown to be spatially localised, with the cancer-affected tissue concentrated in four particular areas of the sample image. Of particular interest to the biochemist and clinicians are molecular discriminators of the clusters. The analysis identified discriminators through a dimension-reduction process which places focus on the variables showing greatest difference between the clusters. The resulting sparse data produced by the dimension reduction were presented as a heatmap, which is a representation of the data which enhances interpretation.

In the final part of the talk Lyron discussed the problem of being able to distinguish the differences between patients from within-patient differences. An example was presented in which 11 samples were taken from samples from three patients. A type of similarity score was calculated for each pair of samples using the common variables resulting from the dimension-reduction process. There was clear differentiation between the samples from the different patients, a result which provides promise that such a differentiation technique could be used for the individualisation of cancer treatments.

Julian Whiting

VIC BRANCH

Review of 'Information Integration in Industry: A Case Study'



Benjamin Rubinstein

Benjamin Rubinstein presented the May seminar for the Victorian Branch. His talk covered work from his previous job at Microsoft Research. He has also worked at other large technology companies, including Intel and Google, and is currently based at IBM Research Australia.

The motivation for the research was to improve the Bing search engine. Microsoft wanted to provide more comprehensive search results for specific types of objects, by incorporating data from existing databases. An example used throughout the talk is searches for films. The well-known Internet Movie Database (IMDb) provides extensive information for a very large number of films. The goal was to collect data from many such sources, merge them together into a single database, and then use this to present detailed results directly in Bing whenever a user searches for a film (including also the ability to interact directly with the film, for example by streaming it online.)

This is a challenging task for many reasons. The source databases will usually vary substantially in size and quality, will have records with different attributes and no reliable common keys, and will have large numbers of missing values.

Ben explained that this type of 'data integration' problem is a very general one and has been tackled by different communities, each developing their own terminology. For example, *entity resolution* or *deduplication* in the databases literature, *coreference resolution* in natural language processing, and *record linkage* within statistics. Ben's experience and work has been mainly within the machine learning community, but he did an excellent job in his talk to translate it for a statistics audience, especially in mapping methodological concepts between these two 'dialects'.

Ben gave context to his work by first describing the 5 steps in a typical 'pipeline' for a data integration project: data preparation, blocking, scoring, matching and merging. Ben noted that statistical techniques can be usefully applied in all of them. However, his talk focused on only two of these, presenting the methods his team developed to solve them.

Typical pipeline

	Title	Year	Directors	Cast	Runtime	...
	Come Fly With Me	2004	Peter Kagan	Michael Bublé	63	...
	Michael Jordan: Come Fly With Me	1989		Michael Jordan, Jay Thomas	42	...



The first method, described as 'multi-task learning for matching multiple sources', tackles the scoring step. The starting point is:

1. a set of databases with records that need to be matched,
2. a set of *feature score functions* which give a similarity score for any pair of records (these are developed in step 3 of the pipeline),

> Continued on next page

3. a small set of records which have been matched manually (called the *labelled dataset*) to provide a gold standard to use for training and evaluation.

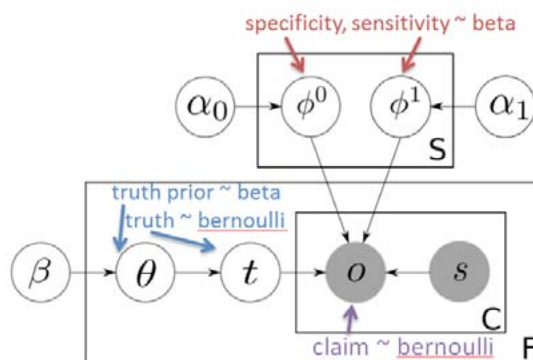
For example, IMDb and Metacritic are two separate databases with information on films. If we take a single record from each one, we will have information such as a film title, a director, length of the film, names of cast members, etc. We assume that we have a function for each such point of comparison, which takes these as input and calculates a similarity score ('feature score'). If the two records represent the same film, then we expect similar information in the two databases, meaning that most of these scores should be high. How do we combine all of these scores together to determine the overall similarity of two records? The labelled dataset gives us some information to help answer this, and the question is how to best use it to determine a good overall score (which is then later used as a basis for the matching step).

The canonical version of this problem only considers two databases. The Microsoft team was attempting to do it with many databases at once. A limiting factor is the size of the labelled dataset. Ideally you would obtain labelled data for each *pair* of databases, but that would need to grow quickly as the number of databases increases. A way was needed to cope with the inevitable shortage.

The approach Ben's team took was described as using *transfer learning* via *inter-task regularisation*. In more familiar terms, it was similar to a hierarchical model, and was fit by optimising a quadratic objective function with linear and quadratic penalty terms to induce sparsity. Ben showed that this performed better than standard approaches that don't borrow ('transfer') strength across the available labelled data, and that it can cope with simultaneously matching many databases (results were shown up to 30).

The second method that Ben presented tackles the final step: merging together matched records. In the example above, this means inferring the true film title, director, cast members, etc., based on a set of matched film records that come from step 4. The problem is to resolve any conflicting information from the matched records and come up with a consensus value for each attribute. Some methods already existed in the literature, but had various disadvantages. An important aspect of the problem was to allow for differing quality of information from the various sources.

The approach Ben's team took used a graphical model where the true values for each attribute were latent variables, and the underlying sensitivity and specificity of each source database were also encoded as latent variables. A Gibbs sampling scheme allowed efficient estimation of all parameters, simultaneously providing inference on the attribute values as well as the source quality.



> Continued on next page

Ben finished by briefly describing how the results of his team's research were communicated within Microsoft and actually implemented in products delivered to customers. It involved employees in many business divisions and multiple countries. Ben highlighted that this translational part of statistical work can often be very challenging. He identified the following as being key ingredients to their success:

- developing the models in small, focused stages,
- implementing them in software components that were easy to customise, configure and re-use,
- creating a code repository that was open to the whole company,
- using and/or creating support tools that automated repetitive tasks, such as administration of datasets and workflow management.

Their methods ultimately featured in over a dozen products. A standout example was the online media services platform included with Microsoft's Xbox 360 video game console. It became one of the most popular ways for people to access videos online in the US, and led to hundreds of million dollars in extra revenue.

Damjan Vukcevic

WA BRANCH

The 2013 calendar is shaping up to be an exciting one for the Western Australia Branch. In addition to the typical monthly seminars, a partnership with the University of Western Australia's Centre for Applied Statistics is opening a greater range of workshops to members. Also, the Young Statisticians will be holding their biennial Young Statisticians Workshop in October, a workshop that is well-attended by students and Young Statisticians in the area, and 2013 will see the first Frank Hansford-Miller Award recipient, Janet Godolphin of the University of Surrey, visit Perth to present a series of seminars for the society and local universities. This award, made possible by the bequest of the late Frank Hansford-Miller (see <http://www.statsoc.org.au/objectlibrary/109?filename=2002%20Nov%20-%20newsletter101.pdf> for an interesting read on Frank), makes possible the visit of a prominent UK statistician to Perth every two years.

The March seminar, the first of 2013, began with the Western Australia Branch's AGM and was followed by a seminar given by Ryan Admiraal of Murdoch University. Ryan's talk, entitled "Generating heterosexual partnership networks with degree distributions and mixing totals estimated from egocentric data," presented a problem in the modeling of contact-based infectious diseases, such as sexually transmitted diseases. In order to project the possible impacts of interventions, it is important that the models capture key network structures observed in the population of interest. In particular, they should match the momentary degree distribution (i.e. distribution of concurrent partnerships) and mixing (i.e. level of contact between people of various characteristics, such as ages, races, etc.) of the population of interest.

Ryan explained that most models can match desired momentary degree distributions or mixing totals for a population but not both because of the complex relationship between momentary degree distributions and mixing totals. He then described a three-part process that allowed one to generate networks matching both the degree distributions and mixing totals for a population. The first step of this process involves simultaneously estimating degree distributions and mixing totals from egocentric samples. This step requires a constrained maximum likelihood estimation approach, as there are a series of consistency constraints that must be specified to ensure agreement between the estimated degree distributions and mixing totals. (For example, the total number of partnerships as given by the momentary degree distributions must match that as given by the mixing totals, and such consistencies should be maintained when considering partnerships broken down by age, race, etc.)

After a consistent set of degree distributions and mixing totals have been estimated for a population, these then serve as the mean value parameters for an exponential-family random graph model (ERGM) which, in turn, can be used to find the corresponding natural parameters for the ERGM using Newton-Raphson (or some other appropriate root-finding algorithm). Finally, once the set of natural parameters for the ERGM is known, MCMC can be used to generate networks that match the desired momentary degree distributions and mixing totals.

Ryan concluded the talk by presenting an example using reported heterosexual partnership data from the National Longitudinal Study of Adolescent Health, demonstrating how consistent degree distributions and mixing totals based on race could be estimated and used to generate heterosexual partnership networks consistent with these degree distributions and mixing totals.

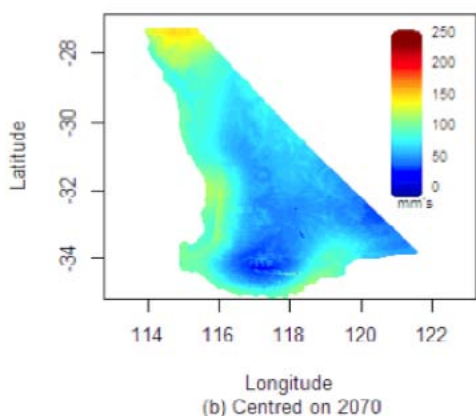
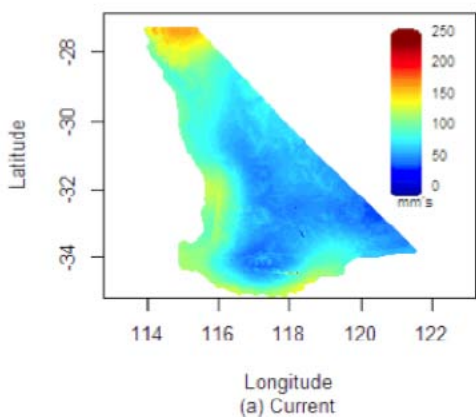
> Continued on next page

At the April meeting of the Western Australia Branch, Mark Palmer, senior statistician at CSIRO Mathematics, Informatics and Statistics, Western Australia, presented an interesting talk on the "Spatial-temporal modelling of extreme rainfall in west and eastern Australia."

In discussing climate and weather, changes in "averages" are often reported, such as mean or median rainfall. But, when assessing and adapting to the potential impact of climate change, it is often more important to consider climate or weather extremes. Therefore, questions such as, "will the frequency and amount of heavy rainfall increase in a particular region or location?" are often considered. Mark discusses that the relationship between frequency and amount of rainfall can be depicted graphically via intensity-duration-frequency (IDF), depth area (DA), and depth-duration-frequency (DDF) curves, and are useful in engineering specifications (such as the design of dams, or bridges), for farming practices, flooding, and the spread of water borne diseases.

The nature of extreme events is that they are rare, thus, only a small amount of data is available for each rainfall station (gauged location). Therefore, Mark proposes a spatial model to borrow strength from neighbouring rainfall stations, in order to increase accuracy for future forecasts of extreme rainfall events. The spatial model also gives forecasts of extreme rainfall at ungauged locations. Mark presented his framework for modelling rainfall extremes; he uses a Bayesian Hierarchical Model (BHM), with a Generalized Extreme Value (GEV) distribution. Mark's framework allows for the incorporation of climate change (in the form of climate drivers from climate models), and temporal aspects are included in his model by looking for trends within covariates, such as the Southern Oscillation Index (SOI), or the ocean heat content. Mark discusses that in contrast to methods currently used, his model allows for measures of variability for forecasts of future extreme rainfall at both gauged and ungauged locations.

Mark concluded his presentation by demonstrating the use of his framework in southwest and northwest of Western Australia, and the Sydney region of New South Wales. The graph below taken from Mark's presentation depicts the return level for the Wheatbelt region of Western Australia. In comparison to current rainfall levels, the graphs show that in 2070, rainfall is expected to decrease in this region.



Anna Munday presenting the 2013 WA Branch Honours Scholarship to Thomas Davidson.

The SSAI Western Australian Branch May meeting had a Young Statisticians theme with two Young Statisticians, Karl Beidatsch and Carol Wang, giving presentations based on their Honours research projects. Also during this meeting, the 2013 WA Honours scholarship was presented to the very worthy applicant, Thomas Davidson, who is undertaking Honours at Murdoch University this year.

Karl Beidatsch, a graduate consultant statistician with Data Analysis Australia, undertook his Honours year at Curtin University during 2012, and it was this research that he shared at the May seminar through his presentation titled "Support Vector Machines in Species Distribution Modelling". In this presentation, Karl described how he was able to assess the performance of a newly-developed presence-only modelling method, the

> Continued on next page

Least-Squares Support Vector Machine, using presence-only sightings of bottlenose dolphins in the Swan-Canning river park sourced from data that had been collected by volunteers for the DolphinWatch project since 2009. On top of the very interesting and informative presentation Karl gave, he was also able to conclude the question time with some tips on where the most common places for dolphin sightings in the Swan River were.

Carol Wang was the second speaker of the night, with her presentation titled "The Association Between Palliative Care and Hospital Usage in the Final Year of Life," which was based on the Honours research project that she undertook at the University of Western Australia in 2012. Through her research, Carol used generalised estimating equations to examine the impact of palliative care on the probability that a terminally ill patient will use a hospital emergency department in the final three months of life. This analysis used more than a thousand records of terminally ill Western Australians, where the matching of treatment and control was based on propensity scores. Carol's results agree with many studies that show there is a lower probability of visiting emergency departments for early specialist palliative care receivers.

At the June meeting, Matthew Knuiman presented a talk titled, "Estimating the Effect of a Time-Varying Predictor on a Time-Varying Outcome from Longitudinal Data: A Case Study Using RESIDE Study Data". The models and study presented in Matthew's talk are related to his work for the RESIDENTIAL Environmental Study (RESIDE) at the University of Western Australia Centre for Built Environment and Health. RESIDE aims to evaluate the impact of urban design on health, in particular the impact of urban design on use of public transport, cycling and walking. This study uses data collected from the Perth area based on a questionnaire that asks participants a variety of questions about things such as their height, weight, body mass index (BMI), and walking habits at specific time points (1, 3, and 6 years).

Matthew's talk outlined the issues and regression models involved in estimating the effect of a time-varying predictor on a time-varying outcome with longitudinal data and provided examples from RESIDE with the use of SAS. The issues discussed in Matthew's talk include that of choosing a model that can account for correlation between repeated measures on the same person. Moreover, the 'between-person' estimate is subject to confounding, while the 'within-person' effect is not subject to confounding.

Linear mixed models for quantitative outcomes and logistic mixed models for binary outcomes can be used to estimate the effects and compare the 'between-person' and 'within-person' estimates.

The study found that the more "walkable" a neighbourhood is, the more likely the subject (person) will walk for public transport. The study also found that for an increase in walking for recreation, the subject's BMI decreases.

Ryan Admiraal, Sylvia Soltyk, and Elyse Corless