Sponsors and participating organisations

Conference organisers:
Kerrie Mengersen,
Cheryle Blair,
Nicole White, Paul Wu & Xing Lee
### Conference Program: Monday 7\(^{th}\) December 2015

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Chair</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 – 10:15</td>
<td><strong>Registration</strong></td>
<td></td>
</tr>
<tr>
<td>10:15 – 10:30</td>
<td><strong>Welcome and Conference Opening</strong></td>
<td>Kerrie Mengersen</td>
</tr>
<tr>
<td>10:30 – 11:30</td>
<td><strong>Keynote Presentation:</strong> Professor Fabrizio Ruggeri, CNR-IMATI, Milan</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Title: New classes of priors based on stochastic orders and distortion functions</td>
<td></td>
</tr>
<tr>
<td>11:30 – 13:00</td>
<td><strong>Tutorial 1:</strong> Professor Richard Boys, Newcastle University, Newcastle upon Tyne, UK</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Title: Bayesian inference for the normal linear model</td>
<td></td>
</tr>
<tr>
<td>13:00 – 15:15</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>15:15 – 15:45</td>
<td><strong>Invited Presentation 1:</strong> Neil Gordon, Defence Science and Technology Group, Department of Defence, Australia</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Title: Bayesian methods and the search for MH370</td>
<td></td>
</tr>
<tr>
<td>15:45 – 16:15</td>
<td><strong>Invited Presentation 2:</strong> Beatrix Jones, Massey University</td>
<td>Nicole White</td>
</tr>
<tr>
<td></td>
<td>Title: Decoupled shrinkage and selection for Gaussian graphical models</td>
<td></td>
</tr>
<tr>
<td>16:15 – 16:30</td>
<td>Afternoon Tea</td>
<td></td>
</tr>
<tr>
<td>16:30 – 17:00</td>
<td><strong>Invited Presentation 3:</strong> Murray Aitkin, University of Melbourne</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Title: How many classes? Statistical modelling of a social network and a terrorist network, with a latent class model and Bayesian model comparisons</td>
<td></td>
</tr>
<tr>
<td>17:00 – 18:00</td>
<td><strong>Workshop 1</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interactive problem solving – case studies in statistical modelling and computation</td>
<td></td>
</tr>
<tr>
<td>18:00 – 20:00</td>
<td>Dinner</td>
<td></td>
</tr>
<tr>
<td>20:00 – 22:30</td>
<td><strong>Poster session 1</strong></td>
<td></td>
</tr>
</tbody>
</table>
### Conference Program: Tuesday 8th December 2015

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Chair</th>
</tr>
</thead>
</table>
| 9:00 – 10:00 | **Keynote Presentation:** Professor Antonietta Mira, Università della Svizzera italiana (USI-University of Lugano) and Insubria University, Como, Italy  
  *Title: Inference in statistical and mechanistic network models* | Julie Vercelloni |
| 10:00 – 10:50| **Tutorial 2:** Professor Fabrizio Ruggeri, CNR-IMATI, Milan  
  *Title: Bayesian inference for Poisson processes and applications in reliability* |                        |
| 10:50 – 11:10| Morning Tea                                                                            |                      |
| 11:10 – 12:00| **Tutorial 2 (cont.):** Professor Fabrizio Ruggeri, CNR-IMATI, Milan  
  *Title: Bayesian inference for Poisson processes and applications in reliability* | Daniel Kennedy      |
| 12:00 – 12:30| **Invited Presentation 4:** Craig Anderson, University of Technology Sydney  
  *Title: Bayesian cluster detection via adjacency modelling* |                      |
| 12:30 – 13:00| **Invited Presentation 5:** Nial Friel, University College Dublin, Ireland  
  *Title: Calibrating mis-specified Bayesian exponential random graph models* |                      |
| 13:00 – 15:15| Lunch                                                                                   |                      |
| 15:15 – 16:15| **Keynote Presentation:** Professor Dianne Cook, Monash University  
  *Title: Is what you see really there? Combining statistical inference with exploratory data analysis* | Earl Duncan         |
| 16:15 – 16:30| Afternoon Tea                                                                          |                      |
| 16:30 – 17:00| **Invited Presentation 6:** Aldo Saavedra, University of Sydney  
  *Title: A Bayesian approach to the spectral analysis of non-stationary spatial-temporal data* |                      |
| 17:00 – 18:00| **Workshop 2**  
  *Interactive problem solving – case studies in statistical modelling and computation* |                      |
| 18:00 – 20:00| Dinner                                                                                   |                      |
| 20:00 – 22:30| **Poster session 2**                                                                   |                      |
### Conference Program: Wednesday 9th December 2015

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
</table>
| 9:00 – 10:00  | **Keynote Presentation**: Professor Richard Boys, Newcastle University, Newcastle upon Tyne  
Title: *Parameter inference for a stochastic kinetic model of expanded polyglutamine proteins* |
|               | Chair Sam Clifford                                                                  |
| 10:00 – 11:10 | **Tutorial 3**: Professor Antonietta Mira, Università della Svizzera italiana (USI-University of Lugano) and Insubria University, Como, Italy  
Title: *MCMC: short introduction and selected advanced topics* |
|               | Chair Sam Clifford                                                                  |
| 11:10 – 11:30 | Morning Tea                                                                         |
| 11:30 – 12:00 | **Invited Presentation 7**: Russell Millar, University of Auckland  
Title: *WIC and importance sampling approximations to cross-validation for models with observation-level latent variables* |
|               | Chair Miles McBain                                                                  |
| 12:00 – 12:30 | **Invited Presentation 8**: Paul Blackwell, University of Sheffield  
Title: *Exact Bayesian inference for continuous-time models of animal movement* |
|               | Chair Miles McBain                                                                  |
| 12:30 – 13:30 | Lunch                                                                              |
| 13:30 – 15:00 | **Tutorial 4**: Professor Dianne Cook, Monash University  
Title: *Making effective plots to explore your data, and adding interactivity* |
|               | Chair Nicholas Tierney                                                              |
| 15:00 – 15:30 | **Invited Presentation 9**: Antony Overstall, University of Glasgow, UK  
Title: *Bayesian optimal design for ordinary differential equation models* |
<p>|               | Chair Nicholas Tierney                                                              |
| 15:30 – 15:45 | Conference close                                                                    |
| 15:45 – 16:15 | Afternoon Tea                                                                       |</p>
<table>
<thead>
<tr>
<th>Presenter</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kate Devitt</strong>, Queensland</td>
<td>Can coherence solve prior probabilities for Bayesianism?</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Renate Meyer</strong>, University of</td>
<td>Bayesian semiparametric spectral analysis of locally-stationary time series with applications to gravitational wave data</td>
</tr>
<tr>
<td>Auckland</td>
<td></td>
</tr>
<tr>
<td><strong>Chatanya Joshi</strong>, Waikato</td>
<td>Fast Bayesian inference using low discrepancy sequences</td>
</tr>
<tr>
<td>University</td>
<td></td>
</tr>
<tr>
<td><strong>Aleysha Thomas</strong>, Queensland</td>
<td>Statistical analysis of persistent organic pollutants in Queensland</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Richard Woodman</strong>, Flinders</td>
<td>Confirmatory factor analysis (CFA) using Bayesian estimation achieved better model fit than with maximum likelihood and thereby</td>
</tr>
<tr>
<td>University</td>
<td>provided construct validity for a lymphedema knowledge survey.</td>
</tr>
<tr>
<td><strong>Chris Oates</strong>, University of</td>
<td>Probabilistic integration</td>
</tr>
<tr>
<td>Technology Sydney</td>
<td></td>
</tr>
<tr>
<td><strong>Jessie Roberts</strong>, Queensland</td>
<td>Communicating statistical uncertainty in disease mapping to non-expert audiences.</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Marcela Cespedes</strong>, Queensland</td>
<td>Statistical methods for spatio-temporal modelling of longitudinal neuroimaging data</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Xing Lee</strong>, Queensland University of Technology</td>
<td>ABC model selection for max-stable models</td>
</tr>
<tr>
<td><strong>Carla Chen</strong>, Australian Institute of Marine Sciences</td>
<td>Understanding the boom-and-bust cycle of crown-of-thorns Starfish using Bayesian semi-individual agent based model with ABC</td>
</tr>
<tr>
<td><strong>Nicholas Hughes</strong>, University of Queensland</td>
<td>Estimating cortical feature maps with dependent Gaussian processes</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Daniel Kennedy</strong>, Queensland</td>
<td>Proposed probabilistic deconvolution of DNA methylation in heterogeneous samples</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Martine Barons</strong>, University of</td>
<td>Coherent frameworks for integrated decision support systems</td>
</tr>
<tr>
<td>Warwick</td>
<td></td>
</tr>
<tr>
<td><strong>Julie Vercelloni</strong>, Queensland</td>
<td>Cumulative effects of disturbances on the coral cover of the Great Barrier Reef: Crown-of-Thorns starfish is trumps</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Fadhah Alanazi</strong>, Queensland</td>
<td>Modelling 3D spatial and spatio-temporal data using pair-copulas</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>Cathy Hargrave</strong>, Queensland</td>
<td>Development of a feature alignment score for cone beam computed tomography based image guided radiotherapy</td>
</tr>
<tr>
<td>University of Technology</td>
<td></td>
</tr>
<tr>
<td>Presenter</td>
<td>Title</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Amy Cook</strong>, Queensland University of Technology</td>
<td>Machine learning and visual analytics for consulting business decision support</td>
</tr>
<tr>
<td><strong>David Smith</strong>, Flinders University</td>
<td>Exploring the measurement structure of the Gambling Related Cognitions Scale (GRCS) in treatment-seekers: a Bayesian structural equation modelling approach</td>
</tr>
<tr>
<td><strong>Wen-Hsi Yang</strong>, CSIRO</td>
<td>Bayesian lattice filters for time-varying autoregression and time-frequency analysis</td>
</tr>
<tr>
<td><strong>Earl Duncan</strong>, Queensland University of Technology</td>
<td>Comparison of re-labelling algorithms applied to Bayesian mixture models: A simulation study</td>
</tr>
<tr>
<td><strong>Paul Wu</strong>, Queensland University of Technology</td>
<td>Absorbing Markov chains and persistence in complex ecosystem models</td>
</tr>
<tr>
<td><strong>Michael Spence</strong>, University of Sheffield</td>
<td>Coupling random inputs for improved ABC parameter estimation in complex models</td>
</tr>
<tr>
<td><strong>Yi-Fan Lin</strong>, University of Newcastle</td>
<td>Bayesian interval estimation and performance measurement</td>
</tr>
<tr>
<td><strong>Vasileios Giagos</strong>, University of New England</td>
<td>LNA-based ABC for stochastic kinetic models</td>
</tr>
<tr>
<td><strong>John Xie</strong>, Griffith University</td>
<td>Modelling QMRA dose-response processes using OpenBUGS: challenges and solutions</td>
</tr>
<tr>
<td><strong>Sarini Sarini</strong>, Queensland University of Technology</td>
<td>Modelling falls in early Parkinson’s disease</td>
</tr>
<tr>
<td><strong>Ester Meenken</strong>, The New Zealand Institute for Plant and Food Research</td>
<td>Improved credible intervals for a latent variable using MCMC and a deterministic computer simulation model for wheat development</td>
</tr>
<tr>
<td><strong>Tomasz Bednarz</strong>, Queensland University of Technology</td>
<td>Immersive Mixed Reality environments as modern interactive visualisation platforms</td>
</tr>
<tr>
<td><strong>June Kim</strong>, Queensland University of Technology</td>
<td>Interactive mirroring: Being and thing become space - Science and art interplays</td>
</tr>
<tr>
<td><strong>Ramethaa Jeyapalan</strong>, Queensland University of Technology</td>
<td>Eliciting and encoding expert knowledge on variable selection for species distribution models</td>
</tr>
<tr>
<td><strong>Nicholas Tierney</strong>, Queensland University of Technology</td>
<td>The sharpest tool in the shed: Applied Bayesian data analysis for industry</td>
</tr>
<tr>
<td><strong>Sama Low-Choy</strong>, Griffith University</td>
<td>Verifying Bayesian networks</td>
</tr>
<tr>
<td><strong>Cameron Williams</strong>, Queensland University of Technology</td>
<td>Evaluating framing effects in surveys of climate change beliefs: an iterative modelling strategy for integrating and refining psychological &amp; cultural theory</td>
</tr>
<tr>
<td><strong>Shovanur Haque</strong>, Queensland University of Technology</td>
<td>Assessing the accuracy of record linkages with Markov Chain based simulation approach</td>
</tr>
</tbody>
</table>
Fabrizio Ruggeri is Research Director at CNR-IMATI, Milan, Italy, and Adjunct Professor at Queensland University of Technology, Brisbane, Australia. His main research interests include reliability, general statistical modelling with applications in industry, and Bayesian statistics. He is a former President of the European Network for Business and Industrial Statistics (ENBIS) and of the International Society for Bayesian Analysis (ISBA). He is the author of over 100 research articles, co-author of a book on Bayesian Inference of Stochastic Processes and he edited two volumes on Bayesian Robustness and one on Statistics in Healthcare. He also served as Editor of the ISBA Bulletin and Bayesian Analysis and is currently Editor-in-Chief of Applied Stochastic Models in Business and Industry and Wiley StatsRef.

Fabrizio is Co-Director of the Applied Bayesian Statistics summer school (ABS) and Chair of the Scientific Committee of the Bayesian Inference in Stochastic Processes (BISP) series of workshops. He is a faculty in the Ph.D. programme in Mathematics at the Universities of Pavia and Milano-Bicocca, Italy, and the Ph.D. programme in Statistics at the University of Valparaiso, Chile. He has been member of many committees, including Savage Award, De Groot Award, Zellner Medal and Box Medal. He is Fellow of the American Statistical Association and ISBA, besides being one of the first two recipients of the Zellner Medal, the most prestigious award by ISBA.

**Keynote presentation**

*New classes of priors based on stochastic orders and distortion functions*
Authors: Pablo Arias-Nicolas, Fabrizio Ruggeri and Alfonso Suarez-Llorens

In the context of robust Bayesian analysis, we introduce a new class of prior distributions based on stochastic orders and distortion functions. We provide the new definition, its interpretation and the main properties and we also study the relationship with other classical classes of prior beliefs. We consider also Kolmogorov and Kantorovich metrics to measure the uncertainty induced by such class, as well as its effect on the set of corresponding Bayes actions. Finally, we conclude the talk with some numerical examples.

**Tutorial**

*Bayesian inference for Poisson processes and applications in reliability*

The tutorial, based on two chapters of the book "Bayesian Analysis of Stochastic Process Models" by Rios Insua, Ruggeri and Wiper, presents basic estimation of parameters in Poisson processes and their use in reliability problems.
Keynote Speaker: Professor Antonietta Mira, Università della Svizzera italiana (USI-University of Lugano) and Insubria University, Como, Italy

Professor Mira is co-founder and co-director of the InterDisciplinary Institute of Data Science – IDIDS at the Università della Svizzera italiana (USI-University of Lugano) where she also serves as the Vice-Dean in the Faculty of Economics and a Professor of Statistics. She is also Professor of Statistics at Insubria University, Como, Italy. She is a Visiting Fellow of the Isaac Newton Institute for Mathematical Sciences at Cambridge University and has been a visiting professor at Université Paris-Dauphine, the University of Western Australia, and the University of Bristol, UK. She has won awards for excellence in both research and teaching. She is the principal investigator on several projects at the Swiss National Science Foundation and a member of multiple scientific committees representing her areas of expertise: Bayesian Computation, and Markov Chain Monte Carlo Computation and Theory. Her current research focuses on methodological and computational statistics, both of which have a clear interdisciplinary scope across social science, finance, economics and industry.

**Keynote presentation**

**Inference in Statistical and Mechanistic Network Models**

**Tutorial**

**MCMC: short introduction and selected advanced topics**
Keynote Speaker: Professor Dianne Cook, Monash University

Professor Cook is a Fellow of the American Statistical Association. Her research is in data visualization, exploratory data analysis, multivariate methods, data mining and statistical computing. She has experimented with visualizing data in virtual environments, participated in producing software including xgobi, ggobi, cranvas and several R packages. Methods development include tours, projection pursuit, manual controls for tours, pipelines for interactive graphics, a grammar of graphics for biological data, and visualizing boundaries in high-d classifiers. Her current work is focusing on bridging the gap between statistical inference and exploratory graphics. She is doing experiments using Amazon's Mechanical Turk, and eye-tracking equipment. Some of the applications that she has worked on include backhoes, drug studies, mud crab growth, climate change, gene expression analysis, butterfly populations in Yellowstone, stimulus funds spending, NRC rankings of graduate programs, technology boom and bust, election polls, soybean breeding, common crop population structures, insect gall to plant host interactions, soccer and tennis statistics.

Keynote presentation

Is what you see really there? Combining statistical inference with exploratory data analysis

Plots of data often provoke the response "is what we see really there". In this talk we will discuss ways to give visual statistical methods an inferential framework. Statistical significance of "graphical discoveries" is measured by having the human viewer compare the plot of the real dataset with collections of plots of null datasets: plots take on the role of test statistics, and human cognition the role of statistical tests, in a process modeled after the "lineup", popular from criminal legal procedures. This is a simple but rigorous protocol that provides valid inference, yielding p-values and estimates of the test power, for graphical findings.

Amazon's Mechanical Turk is used to implement the lineup protocol and crowd-source the inference. Turk is a resource where people are employed to do tasks that are difficult for a computer, in this case, evaluating structure in plots of data. With a suite of experiments, the lineup protocol was run head-to-head against the equivalent conventional test, yielding results that mirror those produced by classical inference. This talk will describe these results, and show how the lineup protocol is used for assessing graphical findings and designing good data plots.

Joint work with Heike Hofmann, Mahbubul Majumder, Niladri Roy Chowdhury, Lendie Follett, Susan Vanderplas, Adam Loy, Yifan Zhao, Nathaniel Tomasetti

Tutorial

Making effective plots to explore your data, and adding interactivity

This tutorial will focus on plotting data using R, in a reproducible way, focusing on how to understand what is your data. We will cover the grammar of graphics, mapping elements of data to plots, and learn to make simple plots like scatterplots, bar charts, time series, profiles, boxplots, adding color, symbol, lines, models. We will also cover layering different data sets, drawing maps, and using cognitive principles to produce quality publication graphics. The R packages used include ggplot2, maps, ggmap, GGally, gridExtra, ggthemes, knitr, rmarkdown.
Keynote Speaker: Professor Richard Boys, Newcastle University, Newcastle upon Tyne

Professor Boys research interests include: Bayesian statistics, statistical bioinformatics and stochastic systems biology. He is a Member of Engineering and Physical Science Research Council (EPSRC) Peer Review College and Chairman of the Royal Statistical Society, Graduate Training Programme Committee. Professor Boys is from the School of Mathematics & Statistics, Newcastle University, Newcastle upon Tyne, UK.

Keynote presentation

Parameter inference for a stochastic kinetic model of expanded polyglutamine proteins

The presence of protein aggregates in cells is a known feature of many human age-related diseases, such as Huntington’s disease. Simulations using fixed parameter values in a model of the dynamic evolution of expanded polyglutaime (PolyQ) proteins in cells have been used to gain a better understanding of the biological system, how to focus drug development and how to construct more efficient designs of future laboratory-based in vitro experiments. However, there is considerable uncertainty about the values of some of the parameters governing the system. Currently, appropriate values are chosen by ad hoc attempts to tune the parameters so that the model output matches experimental data. The problem is further complicated by the fact that the data only offer a partial insight into the underlying biological process: the data consist only of the proportions of cell death and of cells with inclusion bodies at a few time points, corrupted by measurement error.

Developing inference procedures to estimate the model parameters in this scenario is a significant task. Ideally the chemical species in the model would be observed continuously in time, or at least some of them observed at a reasonable number of time points. The model probabilities corresponding to the observed proportions cannot be evaluated exactly and so they are estimated within the inference algorithm by repeatedly simulating realisations from the model. In general such an approach is computationally very expensive and therefore we construct Gaussian process emulators for the key quantities and reformulate our algorithm around these fast stochastic approximations. We conclude by examining the fit of our model and highlight new insights into appropriate values of the model parameters.

Tutorial

Bayesian inference for the normal linear model

Much of Bayesian inference nowadays analyses complex hierarchical models using computer-intensive methods. However, there was a time, not that long ago, when most analyses used a conjugate analysis of say the normal linear model or techniques such a Gaussian quadrature or Laplace’s method to evaluate the required integrals. This tutorial will give an overview of the basics underpinning an analysis of data assuming a normal linear model.
Bayesian methods and the search for MH370
Presenter: Neil Gordon
E-mail: Neil.Gordon@dsto.defence.gov.au
Authors: Neil Gordon, Sam Davey, Ian Holland, Mark Rutten and Jason Williams
Affiliation: Defence Science and Technology Group, Department of Defence, Australia

On 7th March 2014 Malaysian Airlines flight MH370 carrying 239 passengers and crew from Kuala Lumpur to Beijing lost contact with Air Traffic Control and was subsequently reported missing. Over the following days an extensive air and sea search was made around the last reported location of the aircraft in the Gulf of Thailand without success. Signals transmitted by the aircraft’s satellite communications terminal to Inmarsat’s 3F1 Indian Ocean Region satellite indicated that the aircraft continued to fly for several hours after loss of contact.

In this talk we show how Bayesian methods have been used to produce a probability distribution of MH370 flight paths. This probability distribution defines the search zone in the southern Indian Ocean. We describe in detail how the probabilistic models of aircraft flight dynamics, satellite communication system measurements, environmental effects and radar data were constructed and calibrated. A particle filter based numerical calculation of the aircraft flight path probability distribution is described and the method is demonstrated and validated using data from several previous flights of the accident aircraft. The authors are members of the Australian Transport Safety Bureau MH370 search team.
Decoupled shrinkage and selection for Gaussian graphical models

Presenter: Beatrix Jones
E-mail: m.b.jones@massey.ac.nz
Authors: Beatrix Jones\textsuperscript{1} Gideon Bistricer\textsuperscript{1}, Carlos Carvalho\textsuperscript{2} and P Richard Hahn\textsuperscript{3}
Affiliations: \textsuperscript{1}Institute of Natural and Mathematical Sciences, Massey University, Albany, New Zealand
\textsuperscript{2}McCombs School of Business, University of Texas, Austin TX, USA
\textsuperscript{3}Booth School of Business, University of Chicago, Chicago IL, USA

Even when a Bayesian analysis has been carefully constructed to encourage sparsity, conventional posterior summaries with good predictive properties (eg the posterior mean) are typically not sparse. An approach called Decoupled Shrinkage and Selection (DSS), which uses a loss function that penalizes both poor fit and model complexity, has been used to address this problem in regression. This talk extends that approach to Gaussian graphical models. In this context, DSS not only provides a sparse summary of a posterior sample of graphical models, but allows us to obtain a sparse graphical structure that summarises the posterior even when the (inverse) covariance model fit is not a graphical model at all. This potentially offers huge computational advantages. We will examine simulated cases where the true graph is non-decomposable, a posterior is computed over decomposable graphical models, and DSS is able to recover the true non-decomposable structure. We will also consider computing the initial posterior based on a Bayesian factor model, and then recovering the graph structure using DSS. Finally, we illustrate the approach by creating a graphical model of dependencies across metabolites in a metabolic profile—in this case, a data set from the literature containing simultaneous measurements of 151 metabolites (small molecules in the blood), for 1020 subjects.
How many classes? Statistical modelling of a social network and a terrorist network, with a latent class model and Bayesian model comparisons

Presenter: Murray Aitkin
E-mail: Murray.aitkin@unimelb.edu.au
Authors: Murray Aitkin\textsuperscript{1}, Duy Vu\textsuperscript{1} and Brian Francis\textsuperscript{2}
Affiliations: \textsuperscript{1}University of Melbourne
\textsuperscript{2}University of Lancaster, UK

This talk discusses the assessment of the number of classes in a social network, through the latent class model extension of the exponential random graph model. The assessment uses a new Bayesian method of model comparisons, based on the posterior distribution of the deviance for each of the competing models.

The approach is applied to a well-known social network in which the number of classes is known a priori, and to the Noordin Top terrorist network, analysed at length in the book by Everton (2012).

The performance of the posterior deviance model comparison method is illustrated with simulations from single population models and normal mixture models.

References


**Bayesian cluster detection via adjacency modelling**

Presenter: **Craig Anderson**  
E-mail: craig.anderson@uts.edu.au  
Authors: **Craig Anderson**¹, Duncan Lee² and Nema Dean²  
Affiliations: ¹University of Technology Sydney  
²University of Glasgow

Disease mapping is the field of spatial epidemiology interested in characterising how disease risk varies across different geographical regions. A key aim is to identify regions which exhibit significantly elevated disease risk levels, thus allowing public health interventions to be focused on these areas. Bayesian models utilising a Conditional Auto-Regressive (CAR) structure are typically used in these settings. These models usually assume a spatially smooth risk surface across the entire region, but this is not necessarily realistic in practice. Using a case study of respiratory hospital admissions in Glasgow, Scotland, a city with many localised inequalities, I will present an alternative approach which uses clustering techniques to allow for discontinuities in the spatial structure.
Calibrating mis-specified Bayesian exponential random graph models

Presenter: Nial Friel
E-mail: nial.friel@ucd.ie
Author: Nial Friel
Affiliation: University College Dublin, Ireland

Exponential random graph models are widely used for the analysis of social network data. However, parameter estimation is very challenging, since the resulting likelihood function involves an intractable normalising constant. There have been several tractable likelihood approximations presented in the literature, dating back to the pseudolikelihood approximation. It is tempting to consider an approximate posterior distribution where one replaces the likelihood function with the pseudolikelihood approximation. However, this leads to a mis-specified posterior which can result in very misleading inference. We address this issue by calibrating this mis-specified posterior and illustrate through several examples that this can lead accurate estimation. Equally importantly we show that inference can be carried out at a fraction of the computational cost compared to the state-of-the-art.
A Bayesian approach to the spectral analysis of non stationary spatial-temporal data
Presenter: Aldo F. Saavedra
E-mail: Aldo.Saavedra@sydney.edu.au
Authors: Aldo F. Saavedra¹, Sally Wood² and Ori Rosen³
Affiliations: ¹Data Science Hub, University of Sydney
²Business School, University of Sydney
³Department of Mathematical Sciences, University of Texas at El Paso, USA

A mixture-of-experts approach is proposed to modelling the spectral density of a non-stationary time series, where the non-stationarity is with respect to space, time and other covariates. This research is motivated by the need to identify and predict influenza ‘signatures’ across Western Australia. This is achieved by developing a novel nonparametric spatial-temporal model of the number of notified cases of influenza in Western Australia. The underlying spectra are modelled as functions of time, space, and covariates such as the type of influenza virus, weather conditions and the demographic data, using a mixture formulation. The number of components in the mixture is assumed to be unknown but finite. The Whittle likelihood is used as an approximation to the true likelihood and Gaussian process priors for the mixture components are used to model the spectra as functions of frequency. The frequentist properties of the technique are examined via simulation. Initial analysis of the real example reveals that there are distinct influenza signatures in WA. These signatures depend not only on space and time but also on covariates such as workforce movement patterns.
The Watanabe information criterion (WIC) has formal derivations from the theory of algebraic geometry, and as an approximation to leave-one-out cross-validation (LOO-CV). Importance sampling (IS) also provides an alternative approximation to LOO-CV.

In the context of models with observation-level latent variables, the existing literature provides examples of application of WIC and IS using observation-level likelihood.

It is shown here that this approach can be fraught with peril and great care is needed.

In particular, the theoretical justifications of WIC are invalid at the observation level, and its behaviour is application specific.

In particular, it is shown that observation-level WIC is erroneous when applied to over-dispersed count data.

Observation-level IS does continue to be theoretically valid, but in practice can be so numerically unstable as to give misleading results.

It is recommended that WIC and IS be applied using likelihood marginalized over the latent variables.
Exact Bayesian inference for continuous-time models of animal movement
Presenter: Paul Blackwell
E-mail: p.blackwell@sheffield.ac.uk
Authors: Paul Blackwell and Mu Niu
Affiliation: University of Sheffield, UK

Realistic models of wildlife movement present substantial challenges for statistical inference. The models need to take account of variation in behaviour over time, usually through an unobserved behavioural state, and need to be formulated in continuous time, to allow coherent interpretation, comparison and handling of irregular observations. A suitable class of models, which are ecologically and intuitively appealing, can be defined by representing an animal’s behaviour as a continuous-time Markov chain, with transition rates that depend on time and on habitat, and representing its movement as a two-dimensional diffusion process, with parameters determined by its behaviour. The complex relationship between location, behaviour and movement in these models makes inference difficult, and often a simplistic discrete-time approximation is used. In this talk I will describe recent progress in exact Bayesian inference for such models, using Markov chain Monte Carlo (Blackwell et al, 2015; Methods in Ecology and Evolution, doi 10.1111/2041-210X.12460). This method uses a uniformization approach to represent the unobserved changes in behaviour as a thinned Poisson process, avoiding any time-discretization error.

These ideas will be illustrated using data on individual fishers, Martes pennanti, (courtesy of Scott LaPoint, Max Planck Institute for Ornithology, Konstanz) and wild boar, Sus scrofa, (courtesy of Mark Lambert, Animal and Plant Health Agency, York). If time permits, I will also look at extensions to multiple animals, using data on a group of simultaneously tracked reindeer, Rangifer tarandus, (courtesy of Anna Skarin, Swedish University of Agricultural Sciences, Uppsala).
Bayesian optimal design for ordinary differential equation models

Presenter: Antony Overstall
E-mail: Antony.Overstall@glasgow.ac.uk
Author: Antony Overstall
Affiliation: University of Glasgow

Bayesian optimal design is considered for physical models derived from the (intractable) solution to a system of ordinary differential equations (ODEs).

Bayesian optimal design requires the minimisation of the expectation (over all unknown and unobserved quantities) of an appropriately chosen loss function. This can be non-trivial due to 1) the high dimensionality of the design space; and 2) the intractability of the expected loss. In this case, a further complication arises from the intractability of the solution to the system of ODEs.

We propose a strategy that employs a modification of the continuous coordinate exchange algorithm where a statistical emulator is employed to approximate the expected loss function, and a probabilistic solution to the system of ODEs. The strategy is demonstrated on several illustrative examples from the biological sciences.
Can coherence solve prior probabilities for Bayesianism?

Presenter: S. Kate Devitt
E-mail: kate.devitt@qut.edu.au

Coherence between propositions promises to fix the vexing circumstance of prior probabilities for subjective Bayesians. This paper examines the role of coherence as a source of justification for Bayesian agents, particularly the argument that all propositions must cohere within an agent’s ‘web of belief’, aka confirmational holism. Unfortunately, confirmational holism runs across a potentially devastating argument that a more coherent set of beliefs resulting from the addition of a belief to a less coherent set of beliefs is less likely to be true than the less coherent set of beliefs. In response, I propose confirmational chorism (CC) to avoid this troubling outcome. CC posits that coherence adds epistemic justification by limited, logically consistent sets of beliefs exhibiting a satisficing degree of strength, inferential and explanatory connection. Limited coherence may resolve the above argument, but raises the need for another kind of justification: coordination (integration across sets of beliefs). Belief coordination requires suppressing some beliefs and communicating other beliefs to ensure convergence on the right action for performance success. Thus, a Bayesian formed belief in any particular context is justified not just because it is reliably formed and coherent, but also because of how it is coordinated between local and holistic goals.

Bayesian semiparametric spectral analysis of locally-stationary time series with applications to gravitational wave data

Presenter: Renate Meyer
E-mail: meyer@stat.auckland.ac.nz

Astronomy is on the verge of an exciting new era, with the second generation of ground-based gravitational wave (GW) interferometers expected to be operational this year for Advanced LIGO [1], and soon thereafter for Advanced Virgo and KAGRA [2]. GWs, i.e. ripples in spacetime caused by accelerating non-symmetric massive objects, have been predicted by Einstein’s general relativity theory but until today have not yet been directly observed. The detectors should observe GWs that produce changes to their 4 km length arms that are about ten thousand times smaller than the size of a proton, making their measurement a formidable task. Of paramount concern for reliable signal detection and accurate parameter estimation is not only a concise waveform model that characterises the GW signal [3] but also a statistically rigorous model of the time series noise. Standard GW data analysis assumes detector noise is stationary and Gaussian distributed, with a known power spectral density (PSD) that is usually estimated using clean off-source data. Real GW data often depart from these assumptions, and misspecified parametric models of the PSD could result in misleading inferences. We propose a Bayesian nonparametric approach to improve this. We put a nonparametric Bernstein polynomial prior on the PSD [4] and update this using Whittle’s likelihood [5]. Posterior samples are obtained using a blocked Metropolis-within-Gibbs sampler. We simultaneously estimate the reconstruction parameters of a rotating core collapse supernova GW burst that has been embedded in simulated Advanced LIGO noise. We also discuss an approach to deal with non-stationary data [6].

References
[1] https://www.advancedligo.mit.edu
**Fast Bayesian inference using low discrepancy sequences**

Presenter: Chaitanya Joshi  
E-mail: cjoshi@waikato.ac.nz  
Authors: Chaitanya Joshi, Paul Brown and Stephen Joe  
Affiliation: Faculty of Computing and Mathematical sciences, University of Waikato, New Zealand

For models with latent Gaussian Markov random field (GMRF) structure, methods such as the Integrated Nested Laplace Approximation (INLA) can provide fast and accurate inference. These methods explore the parameter space using a predetermined grid. Therefore, the computational advantage may be lost if the number of hyper parameters is moderately large (6 or above). Instead, if the parameter space is explored using the low discrepancy sequences (LDS), it may be possible to extend the computational benefit to models with up to 10 or 12 hyper parameters without losing the accuracy. We explain how this can be done and illustrate using a few examples.

---

**Statistical analysis of persistent organic pollutants in Queensland**

Presenter: Aleysha Thomas  
E-mail: a12.thomas@qut.edu.au  
Authors: Aleysha Thomas, Kerrie Mengersen, Nicole White, Leisa Maree-Toms, Jochen Mueller and Fiona Harden  
Affiliations: 1Queensland University of Technology, 2University of Queensland

Persistent Organic Pollutants (POPs) are toxic, carbon-based chemicals used for industrial/pesticide purposes. Many POPs are banned all over the world. Still, these chemicals remain in the environment, are distributed via land/air/water and tend to accumulate in the body tissue. This poster focuses on organochlorine pesticides (OCPs) which were used to protect crops, livestock and buildings in Australia between the mid-1940s and 1960s. The levels of OCPs in the Australian population are regularly assessed and compared to international observations. Usually researchers utilise summaries, analysis of variance followed by multiple comparison tests to describe the levels of OCP in the population. However, more complex methods can be used to infer relationships and trends. Pesticide levels are averaged from pools of individuals' blood serum from urban areas in Queensland, Australia over ten years. As the same individuals are not sampled over time, the pooled observations are treated as replicates of sub-populations. Complex linear models such as classification and regression trees (CARTs) and linear mixed models are used to identify trends and associations in the data.
Confirmatory factor analysis (CFA) using Bayesian estimation achieved better model fit than with maximum likelihood and thereby provided construct validity for a lymphedema knowledge survey.

Presenter: Richard J Woodman
E-mail: Richard.Woodman@flinders.edu.au
Authors: Richard J Woodman 1, Agnik Sarkar 2, Changcheng Zhu 3, Zoe Bogner 3 and Neil Piller 2
Affiliations: 1Flinders Centre for Epidemiology and Biostatistics, School of Medicine, Flinders University, South Australia
2Lymphoedema Research Unit, Department of Surgery, School of Medicine, Flinders University, South Australia
3Faculty Learning Technologies Unit, Faculty of Medicine, Nursing and Health Sciences, Flinders University, South Australia

Background: Model fit for a CFA is used to determine construct validity including the hypothesised underlying dimensionality. Frequentist CFA approaches such as maximum likelihood (ML) estimation often provide poor model fit as a result of unrealistically constraining most item-item residual covariances to zero. Bayesian CFA allows greater flexibility and maintains substantive theory by allowing item-item covariances to vary from zero using uninformed priors with zero means.

Methods: We used data from an 11-item survey testing knowledge of lymphedema to estimate a hypothesised one-factor solution using both Bayesian and ML CFA with Mplus software.

Results: Model fit was poor for both Bayesian and ML CFA when all covariances were constrained to zero. The addition of 6 covariances based on identification of 4 closely related survey items improved model fit but was still sub-optimal for both approaches. A Bayesian CFA with all item-item covariances allowed to vary around zero provided excellent fit and the only 3 large (beta>0.3) and significant (p<0.05) covariances were a subset of the previously identified 6 covariances. In contrast, good model fit was only obtained using ML when 8 item-item covariances were included which were statistically identified using modification indices rather than substantive knowledge.

Conclusion: Bayesian estimation provided a one factor model with excellent fit and maintained substantive theory. ML estimation required inclusion of several item-item covariances which were not all identifiable using survey item knowledge alone. Bayesian CFA provides a more flexible and perhaps more realistic method for assessment of construct validity than ML methods.

Probabilistic integration
Presenter: Chris Oates
E-mail: christopher.oates@uts.edu.au
Authors: Chris Oates 1, François-Xavier Briol 2,3, Nicolas Chopin 4, Mark Girolami 3 and Michael Osborn 2
Affiliations: 1University of Technology Sydney,
2University of Oxford,
3University of Warwick,
4École nationale de la statistique et de l’administration économique (ENSAE)

This presentation is about an exciting new area of research called Probabilistic Numerics. I will provide an overview of recent developments and explain how statisticians have a central role to play in contemporary numerical analysis. Focussing on integration, due to its frequent usage in Bayesian statistics, I will present probabilistic integrators that build upon, and have some advantages over, well-studied Monte Carlo methods.

Web: http://oates.work
Communicating statistical uncertainty in disease mapping to non-expert audiences
Presenter: Jessie Roberts
E-mail: jessie.roberts@qut.edu.au
Author: Jessie Roberts
Affiliation: Queensland University of Technology

Communicating the uncertainty and reliability of scientific insights to non-expert decision-makers is essential to the application of scientific discoveries in industry, government and business. In a world where the future is unknown and we never have perfect information, uncertainty is critical to making the most of the information that we do have. It is an indicator of the reliability of information, a measure of how our knowledge is changing over time, is critical to comparing and evaluating the accuracy of different research outputs or methodologies, and guides future research by highlighting current gaps in our knowledge.

In tomorrow's data driven landscape, where non-experts and all industries are demanding insights from the data emerging around us, uncertainty information becomes increasingly important. The development of strategies and methodologies for effectively communicating the uncertainty and reliability of data derived and scientific insights is a key challenge facing science communication today. This research focuses on extending the field of data visualisation and disease mapping to include uncertainty information. While data visualisation has been proven to be a highly effective medium for communicating scientific and data derived insights to non-expert audiences, there is currently no accepted or standard methods for visualizing uncertainty.

This research focuses on the National Cancer Atlas as a case study and platform for developing methodologies and frameworks for designing uncertainty communications. The research extends the current approaches to disease mapping and geospatial health statistics as well as presents a structured design process for designing targeted uncertainty communication material that engages project stakeholders within the process.

Statistical methods for spatio-temporal modelling of longitudinal neuroimaging data
Presenter: Marcela Cerpedes
E-mail: marcela.cerpedes@hdr.qut.edu.au
Authors: Marcela Cespedes¹, James McGree¹, Christopher Drovandi¹, Kerrie Mengersen¹, James Doecke² and Jurgen Fripp²
Affiliations: ¹Mathematical Sciences School, Queensland University of Technology ²CSIRO Digital Productivity and Services/Australian E-Health Research Centre

The neuropathological process of Alzheimer’s disease (AD) can begin decades prior to the manifestation of symptoms. This complex process often overlaps with natural brain degeneration from healthy ageing, which makes it difficult to distinguish early disease onset. The degenerative process occurs in multiple regions over time resulting in a reduction of cortical thickness, which can be monitored via resting state Magnetic Resonance Imaging (MRI). Modelling the multi-region degeneration process over time requires the application of spatio-temporal models, however challenges remain in modelling spatial dependence, as associations among regions do not necessarily decrease as a function of growing separation between measured locations. In this research we demonstrate methodology to discover groups or patterns of region-wise degeneration over time in a mixed effects model setting. The initial methodology derived lays the foundation for novel spatio-temporal neurodegeneration modelling over time.
ABC model selection for max-stable models
Presenter: Xing Lee
E-mail: xj.lee@qut.edu.au
Authors: Xing Lee, Chris Drovandi and Tony Pettitt
Affiliation: Queensland University of Technology

Max-stable models are a common choice for spatial extreme data as they arise naturally as infinite dimensional generalisation of multivariate extreme value theory. Statistical inference for such models is complicated by the intractability of the multivariate density function. Nonparametric and composite likelihood based approaches have been proposed to address this difficulty. More recently, a simulation-based approach using approximate Bayesian computation (ABC) has been proposed for estimating parameters of max-stable models.

The work here expands on the ABC method to address the issue of model selection between spatial extreme models. ABC algorithms rely on evaluation of discrepancies between model simulations and the observed data rather than explicit evaluations of computationally expensive or intractable likelihood functions, as is the case for max-stable models. In particular, this work considers six max-stable models: the Smith model, the Schlather model with either Whittle-Matérn, Cauchy, powered exponential or Bessel covariance function, and the Brown-Resnick model.

The method was applied to annual maximal temperature data from 26 weather stations dispersed around South Australia. Agriculture and horticulture are strong industries in South Australia which depend heavily on the dry and hot weather of the state. However, extreme weather events can be detrimental to these industries. For example, bushfires, a consequence of extremely high temperatures coupled with low humidity, have been reported on numerous occasions in South Australia. It is thus important to be able to confidently estimate climate conditions particularly in areas further away from weather stations in order to better manage and safeguard such industries.

Understanding the boom-and-bust cycle of crown-of-thorns starfish using Bayesian semi-individual agent based model with ABC
Presenter: Carla Chen
E-mail: c.ewels@aims.gov.au
Author: Carla Chen
Affiliation: Australian Institute of Marine Sciences

Every seven to ten years, outbreaks of the crown-of-thorns starfish (COTS) on the Great Barrier Reef have a detrimental impact on the coral cover at affected locations. A recently published longitudinal study estimated the outbreak of COTS has contributed 42% of coral loss in GBR. The outbreak of COTS often lasts around five years, before a sudden decline in the population. Although the GBR has experienced four major outbreaks since the early 1960’s, and various hypotheses have been proposed for the cause of the outbreaks, the definitive causes of the boom-and-bust cycle still remains unknown. In contrast, the biology of COTS is relatively well understood. Therefore, the aim of this study is to examine the possibility of explaining the boom-and-bust cycle which occurred at four reefs using a stochastic Bayesian individual agent based model, building upon existing knowledge of the biology and ecological evolutionary processes of COTS. In contrast to most process models, the model developed in the study 1) used an agent based approach which allows the inclusion of finer resolution species-environment interaction, and 2) is an empirical model. Due to the complexity of the evolutionary process, the likelihood of the model is intractable; therefore we used an approximate Bayesian computation (ABC) for parameter estimation.
Estimating cortical feature maps with dependent Gaussian processes
Presenter: Nicholas J. Hughes
E-mail: nicholas.hughes1@uqconnect.edu.au
Authors: Nicholas J. Hughes and Geoffrey J. Goodhill
Affiliation: Queensland Brain Institute and School of Mathematics and Physics, The University of Queensland

Bayesian approaches have found widespread applicability in neuroscience. Here we generalise Bayesian methods for studying functional maps in the visual brain. Cells in mammalian visual cortex are arranged in stereotypical patterns based on their preferences for the orientation of edges and contours in the visual input. These orientation preference maps are striking examples of spatial organisation in the brain, and have highly consistent statistical properties across both individuals and species. However the fine-scale, quantitative analysis of these properties requires accurate reconstructions of map structure from noisy imaging data. A Bayesian approach to this reconstruction problem utilising Gaussian process regression was recently proposed, which produces substantially improved map estimates compared to classical techniques. However this work has so far ignored the several maps of preference for other visual features which coexist with the orientation preference map, and that these maps have mutually dependent structures. Here we extend the Gaussian process framework to consider multiple outputs, using filtering kernel convolutions to define the relationships between maps, allowing multiple maps to be modelled simultaneously. We demonstrate that this model encodes the well-known correlations between maps and also improves estimates of multiple maps compared to both classical techniques and the single output approach. The use of a Bayesian prior to encode map structure and dependencies is integral to this improved estimation. This provides the first principled approach for studying the spatial relationships between feature maps in visual cortex, and demonstrates a general method for the estimation of multiple, correlated structures from noisy data.

Cellular heterogeneity in DNA methylation: A new approach for the estimation of cellular proportions in whole blood
Presenter: Nicole White
E-mail: nm.white@qut.edu.au
Authors: Nicole White, Miles Benton, Rodney Lea, Lyn Griffiths and Kerrie Mengersen
Affiliation: Queensland University of Technology

DNA methylation is a commonly studied epigenetic mechanism characterised by the chemical addition of a methyl group (CH3) to the 5’ position of a cytosine, commonly occurring at CpG dinucleotides. With a recognised role in many fundamental cellular processes, including transcriptional regulation and cell differentiation, DNA methylation is an active area of research with increasing emphasis placed on its association with biological processes and disease etiology. This research has been made possible by advances in high-throughput technology, namely the Illumina Infinium 450K microarray, that enable the measurement of methylation at the genome scale.

Cellular heterogeneity has been acknowledged as a key complicating factor in the analysis of DNA methylation, in light of observed differences in methylation profiles between cell populations. Failure to account for the cellular composition of individual samples bears consequences for association studies based on DNA methylation, where one seeks to identify systematic variation attributable to a chosen trait. Methodology to account for cellular heterogeneity is now established in the literature; however the approaches proposed are limited in terms of the number of isolated cell populations considered and the representation of uncertainty in estimated proportions. In this work, an alternative approach for the estimation of cell proportions is proposed, and benefits over existing approaches are outlined. Our methodology is demonstrated on a publicly available dataset containing cell-sorted DNA methylation measurements on six healthy male volunteers. The performance of the proposed methodology is also compared with the current gold standard for cell proportion estimation in whole blood.
Proposed probabilistic deconvolution of DNA methylation in heterogeneous samples

Presenter: Daniel Kennedy
E-mail: d.w.kennedy1992@gmail.com
Authors: Daniel Kennedy, Rod Lea, Nicole White, Miles Benton, Lyn Griffiths and Kerrie Mengersen
Affiliation: Queensland University of Technology

Methylation is an epigenetic phenomenon in which a methyl group attaches to a cytosine base on the DNA molecule. It is known to have a role in gene silencing, and as a result there is great interest in determining methylation patterns associated with cell type differentiation and disease. Due to the difficulty in obtaining cell-sorted data, in silico techniques for deconvolving mixed cell samples are critical for making inferences at the level of the cell type.

This work first investigates current methods for deconvolving mixed cell samples, evaluating their performances against available cell-sorted data. It is found that these methods are unreliable and ineffective due to incorrect assumptions of Gaussian error and minimal variation on the cell-type specific level, especially given the low sample sizes generally observed for methylation data.

A probabilistic approach based on latent variable modelling is proposed to provide deconvolution at the cell-type specific level, built around methods for estimating cell-type proportion. The model will allow for non-Gaussian variation around mean parameters, and will be flexible to incorporate discrete and continuous predictors of interest.

Coherent frameworks for integrated decision support systems

Presenter: Martine J. Barons
E-mail: Martine.Barons@warwick.ac.uk
Authors: Martine J. Barons, Jim Q. Smith and Manuele Leonelli
Affiliation: Department of Statistics, University of Warwick

In today’s increasingly interconnected world, designing policies to effect change when systems are large and intimately connected is increasingly desirable, but often infeasible. In such a scenario, a subjective expected utility policy maker who manages complex, dynamic systems needs to draw on the expertise of a variety of disparate panels of experts and integrate this information coherently. To achieve this, diverse supporting probabilistic models have to be networked together; the output of one model providing the input to the next. Such a network can then be used as framework for an integrating decision support system that can inform policy choice or be used to manage a crisis. In Smith, Barons & Leonelli (2015), we develop a statistical methodology underpinning Bayesian inference for such a tool. In particular, we derive sufficient conditions that ensure inference remains coherent before and after relevant evidence is accommodated into the system. In this presentation, we review some recent work on Bayesian inference underlying integrated decision support for huge processes, illustrated using examples drawn from work with local government in the UK designed to support decisions impacting on household food security.

Cumulative effects of disturbances on the coral cover of the Great Barrier Reef: Crown-of-Thorns starfish is trumps

Presenter: Julie Vercelloni  
E-mail: j.vercelloni@qut.edu.au

Authors: **Julie Vercelloni**¹, Kerrie Mengersen¹ and M. Julian Caley²

Affiliations: ¹School of Mathematical Sciences, Queensland University of Technology  
²Australian Institute of Marine Science, Townsville, Queensland, Australia

Although intermediate rates and intensities of natural disturbances can support the existence of high diversity ecosystems, anthropogenic disturbances currently threaten the continued existence of some major habitat-forming species. Recently, such disturbances have resulted in worldwide declines of coral populations. The complexity of the spatio-temporal dynamics of coral populations, disturbances, and their interactions, increase the difficulty of understanding their effects. Here, using a Semi-Parametric Bayesian Hierarchical Model, we quantify the strength of cumulative effects of disturbances on Great Barrier Reef's coral cover trajectories by estimating coral abundance and the uncertainties associated with the model structure and its parameters. The effects of coral predation by Crown-of-Thorns starfish (CoTS), the total number of disturbances and time between two consecutive disturbances were the main factors contributing to increased long-term cumulative impacts on coral dynamics. CoTS out-ranked all other disturbance types in their contribution to coral cover decline and were associated with reduced resilience.

Modelling 3D spatial and spatio-temporal data using pair-copulas

Presenter: Fadhah Alanazi  
E-mail: fadhahamer.alanazi@hdr.qut.edu.au

Authors: Fadhah Alanazi, Helen Thompson, Chris Drovandi and Kerrie Mengersen

Affiliation: School of Mathematical Sciences, Queensland University of Technology

Spatial and spatio-temporal data are becoming of increasingly more interest, since analysis of such data informs policy in many areas such as the environment, agriculture and mining. Such data may exhibit complex spatial and temporal dependence, exhibiting both non-linear and non-Gaussian dependence structures. In conventional geostatistical models, these dependencies are often ignored, which may result in inaccurate estimation of, and prediction from, the random field. Pair-copulas are able to capture these complex dependence structures. The poster will show the results of modelling 3D spatial and spatiotemporal data using pair-copulas.
Development of a feature alignment score for cone beam computed tomography based image guided radiotherapy

Presenter: Kerrie Mengersen
Authors: Catriona Hargrave\textsuperscript{1,2,3}, Timothy Deegan\textsuperscript{1}, Tomasz Bednarz\textsuperscript{3}, Michael Pouslen\textsuperscript{1}, Fiona Harden\textsuperscript{2} and Kerrie Mengersen\textsuperscript{3}
E-mail: Cathy.Hargrave@health.qld.gov.au
Affiliations:
\textsuperscript{1}Radiation Oncology Mater Centre, QLD Health, South Brisbane, Australia
\textsuperscript{2}School of Clinical Sciences, Queensland University of Technology, Brisbane Australia
\textsuperscript{3}School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia

Image guided radiotherapy (IGRT) uses pre-treatment images of cancer patients on the treatment unit couch to determine if corrections to patient positioning are required prior to treatment delivery. Cone beam computed tomography (CBCT) images are compared to reference images, demonstrating the required position of the radiation fields with respect to the patient’s tumour and surrounding anatomy. CBCT-based IGRT requires clinicians to weigh up many clinical factors within tight time constraints. A feature alignment score correlated with the probability of correct radiation dose delivery to tumours and organ at risks (OARs) is being developed for inclusion in an IGRT decision-support Bayesian network. An actionable score in the clinical IGRT environment should incorporate expert clinician opinion and be appropriate for use alongside IGRT technology. Radiation therapists (RTs) and radiation oncologists (ROs) participated in a workshop to elicit current prostate IGRT practices. Randomly selected CBCTs from 36 prostate cancer patients were retrospectively contoured to identify tumour and OAR volumes, then registered with the radiotherapy treatment plan to simulate varying degrees of positioning related dose errors. Exported contours and dose volume histogram (DVH) data were compared to standard volume variation metrics. Reduction in measured dose errors were poorly correlated with the Dice similarity score; maximum Hausdorf distances performed better, but were not spatially specific. Using Molleweide projections to quantify volume differences relative to the patient coordinate system provides the required spatial information. These data could be incorporated into a tool for visually tracking residual errors that may exist after the application of positioning corrections.
Exploring the measurement structure of the Gambling Related Cognitions Scale (GRCS) in treatment-seekers: a Bayesian structural equation modelling approach
Presenter: David Smith
E-mail: david.smith@flinders.edu.au
Author: David Smith
Affiliation: Flinders Human Behaviour and Health Research Unit, Department of Psychiatry, Flinders University

Knowledge of an individual’s underlying gambling related cognitions plays an important role in treatment approaches for Gambling Disorder. The Gambling Related Cognitions Scale (GRCS) is therefore frequently used in clinical settings for screening and evaluation of treatment outcomes. However, GRCS validation studies have generated conflicting results regarding its latent structure using traditional confirmatory factor analyses (CFA). This may partly be due to the rigid constraints imposed on cross-factor loadings with traditional CFA. The aim of this investigation was to determine whether a Bayesian structural equation modelling (BSEM) approach to examination of the GRCS factor structure would better replicate substantive theory and also inform model re-specifications. Participants were 454 treatment-seekers at first presentation to a gambling treatment centre between January 2012 and December 2014. Model fit indices were well below acceptable standards for CFA. In contrast, the BSEM model which included small informative priors for the residual covariance matrix in addition to cross-loadings produced excellent model fit for the original hypothesized factor structure. The results also informed re-specification of the CFA model which provided more reasonable model fit. These conclusions have implications that should be useful to both clinicians and researchers evaluating measurement models relating to gambling related cognitions in treatment-seekers.

Bayesian lattice filters for time-varying autoregression and time-frequency analysis
Presenter: Wen-Hsi Yang
E-mail: Wen-Hsi.Yang@csiro.au
Authors: Wen-Hsi Yang, Scott H. Holan and Christopher K. Wikle
Affiliations: 1Commonwealth Scientific and Industrial Research Organisation (CSIRO)
2Department of Statistics, University of Missouri

Modeling nonstationary processes is of paramount importance to many scientific disciplines including environmental science, ecology, and finance, among others. Consequently, flexible methodology that provides accurate estimation across a wide range of processes is a subject of ongoing interest. We propose a novel approach to model-based time-frequency estimation using time-varying autoregressive models. In this context, we take a fully Bayesian approach and allow both the autoregressive coefficients and innovation variance to vary over time. Importantly, our estimation method uses the lattice filter and is cast within the partial autocorrelation domain. The marginal posterior distributions are of standard form and, as a convenient by-product of our estimation method, our approach avoids undesirable matrix inversions. As such, estimation is extremely computationally efficient and stable. To illustrate the effectiveness of our approach, we conduct a comprehensive simulation study that compares our method with other competing methods and find that, in most cases, our approach performs superior in terms of average squared error between the estimated and true time-varying spectral density. Lastly, we demonstrate our methodology through three modeling applications; namely, insect communication signals, environmental data (wind components), and macroeconomic data (US gross domestic product (GDP) and consumption.
Comparison of re-labelling algorithms applied to Bayesian mixture models: A simulation study
Presenter: Earl Duncan
E-mail: earl.duncan@qut.edu.au
Authors: Earl Duncan, Nicole White and Kerrie Mengersen
Affiliation: ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology (QUT), Australia

Mixture models accommodate unobservable heterogeneity in the data by specifying the likelihood function as a mixture of distributions, with each component distribution corresponding to a homogenous subpopulation. A prevailing problem inherent to all mixture models is the issue of label-switching, whereby the labels used to identify the mixture components have the potential to switch at each iteration of the algorithm used to update the posterior distribution. This has the effect of permuting the sample of parameter values comprising the posterior, making estimation of the parameters inaccurate and prohibiting sensible inferences from the model. The most common strategy to deal with this issue is to apply a re-labelling algorithm to the posterior sample to determine the permutations necessary to reverse the effect of label switching, thus making the component labels consistent. This study compares the effectiveness, computational efficiency, and accuracy of several re-labelling algorithms for a variety of models, datasets, and theoretical label switching situations. The results of this study debunk some misconceptions about relabelling algorithms, and demonstrate the need for caution when applying relabelling algorithms.

Absorbing Markov chains and persistence in complex ecosystem models
Presenter: Paul Wu
E-mail: p.wu@qut.edu.au
Authors: Paul Wu, Kerrie Mengersen and Julian Caley
Affiliations: Queensland University of Technology and Australian Institute of Marine Science

Predicting the persistence of species under anthropogenic and environmental disturbances is a key component in the management of complex ecosystems. One way to capture the various ecological, biological and environmental processes and their interactions under uncertainty is to use a probabilistic state space. Dynamic Bayesian Networks (DBNs), which are considered general forms of Hidden Markov Models and Kalman filters, provide a method for forwards and backwards inference on such state space models. Thus, it is possible to compute the probability of persistence under different disturbance regimes, or alternatively, predict the regimes under which persistence or a desired population level is achieved.

However, due to interdependencies between variables, absorbing states or even absorbing classes (sets of states) can arise. When the system enters an absorbing state, it can never leave that state, such as if an individual organism enters a deceased state. In an ecosystem, there can be more complex relationships between population, loss and recovery (such as with recruitment and growth) resulting in a zero absorbing class. As a result, following changes to the population arising from disturbances, the modelled system is unable to transition out of the zero states even with positive recovery. The poster illustrates the problem with an example seagrass ecosystem DBN model. Through the introduction of variable transition probabilities, effectively producing an inhomogeneous Markov Chain, it is possible to overcome the challenge of absorbing states. Consequently, it is possible to more accurately capture loss and recovery and thus the probability of persistence of an ecosystems for the purposes of decision support.
Coupling random inputs for improved ABC parameter estimation in complex models
Presenter: Michael A Spence
E-mail: m.a.spence@sheffield.ac.uk
Author: Michael A Spence
Affiliation: University of Sheffield, UK

Approximate Bayesian Computation (ABC) gives a means of Bayesian inference when the likelihood is intractable, as is often the case with complex stochastic models, such as individual- or agent-based models. It can be combined with the idea of Markov chain Monte Carlo, to give the ABC–MCMC algorithm. Unfortunately, for some models this method does not perform well, and some alternatives have been proposed including the f+sMCMC algorithm (Neal and Huang, 2015; Scandinavian Journal of Statistics 42:378–396) that explores the space of random inputs as well unknown model parameters. In this presentation I will present the Coupled Gibbs ABC algorithm (CG-ABC) introduced by Spence and Blackwell (2015; Statistics and Computing, doi 10.1007/s11222-015-9593-2), a further development of the ABC–MCMC algorithm which makes moves within the joint parameter and random input space in order to get better mixing of the Markov chain. CG-ABC includes Gibbs steps that condition on aspects of the currently accepted model and allow parameters and random inputs to move efficiently under this conditioning. I will show empirically that CG-ABC improves the efficiency of the ABC-MCMC algorithm on a range of models, including an individual-based model of a group-living bird, the woodhoopoe, and a multi-species size-spectrum model of marine population dynamics. Parts of this work are joint with Paul Blackwell and Joe Hunt.

Bayesian interval estimation and performance measurement
Presenter: Yi-Fan Lin
E-mail: C3192283@uon.edu.au
Author: Yi-Fan Lin
Affiliation: University of Newcastle

The use of Bayesian methods within statistical process control is an important pursuit. This presentation will describe a new Bayesian control chart for monitoring clinical indicator data in hospitals based on the beta binomial posterior predictive (BBPP) distribution. The BBPP chart has been tested against the traditional Bernoulli CUSUM chart under the conditions that the parameters are either known or need to be estimated, and by constructing the BBPP chart’s limits using two methods; one a more traditional “central” interval approach and the other using the “highest posterior density” (HPD) interval. Preliminary results indicate the BBPP chart with HPD-based control limits provides better run length performance than the traditional central interval, and compared with the Bernoulli CUSUM chart, the BBPP chart has a better run length performance when the true underlying proportion is unknown, with the BBPP chart detecting signals faster than the Bernoulli CUSUM chart. It was also found that in the event of underlying parameters requiring estimation, the Bernoulli CUSUM chart’s performance may be improved by using Bayesian estimation for the limits.
LNA-based ABC for stochastic kinetic models
Presenter: Vasileios Giagos
E-mail: vgiagos@une.edu.au
Author: Vasileios Giagos
Affiliation: University of New England

Kinetic models are used to describe the interactions of a population (e.g. ecological species, genetic components, epidemiological states) which, in turn, can be expressed as Markov Jump Process (MJP) to describe their time evolution. The parameters of interest are the kinetic parameters which govern the rate of these interactions (e.g. birth rate, transcription rate, infection rate).

The Linear Noise Approximation (LNA) is an asymptotic stochastic approximation to the MJP dynamics which offers a tractable likelihood depending on the numerical solution of a system of Ordinary Differential Equations. Drawing inspiration from Indirect Inference we use the LNA-derived density as a metric for a Sequential Approximate Bayesian Computation algorithm. We note that the proposed algorithm scales seamlessly in a parallel computation environment and can also be used as a diagnostic to assess the LNA approximation under specific experimental conditions. Finally, we demonstrate our methodology in examples from ecology and system biology.

Modelling QMRA dose-response processes using OpenBUGS: challenges and solutions
Presenter: Gang Xie
E-mail: g.xie@griffith.edu.au
Authors: Gang Xie1,2, Anne Roiko1,3, Peter Dunn2, Sonya Kozak3 and Kerrie Mengersen4
Affiliations: 1Smart Water Research Centre, Griffith University
2University of the Sunshine Coast
3School of Medicine, Griffith University
4 Queensland University of Technology

Quantitative microbial risk assessment (QMRA) is widely accepted for characterization of the microbial risks associated with food, water and wastewater. The exponential and beta-Poisson single-hit models are the two most commonly used dose-response models in QMRA. The implementation of these dose-response models by following a Markov Chain Monte Carlo (MCMC) sampling approach could provide greater flexibility for statistical inference and may open up a new path for development of the new dose-response models based on mechanistic assumptions. Performing MCMC sampling for the QMRA dose-response models, however, is not a trivial exercise. In this paper/poster, we share our experience of modelling the QMRA dose-response processes using OpenBUGS (version 3.2.2). The exponential and beta-Poisson single-hit dose-response models were coded in BUGS language in two different ways, respectively. The first method was to code the dose-response models based on the analytic model formulas; the second method was to code the models based on the underlying dose-response processes. Analysis results were compared between both methods to verify the MCMC sampling results for the second method which was more prone to errors in both coding process and the MCMC implementation. The method based on the underlying dose-response processes has the potential/advantage for development of the new dose-response models based on mechanistic assumptions. By focusing on the unique OpenBUGS function ‘density’ (which is not available in the more popular WinBUGS software), we explored the possibility for implementation of a generalized dose-response model using OpenBUGS. Discussions are given for a potential limitation of MCMC implementation related to the OpenBUGS distribution function ‘dbin’. Further research is required to solve the problem of implementation of the generalized dose-response model using OpenBUGS.
Modelling falls in early Parkinson’s disease
Presenter: Sarini Sarini
E-mail: sarini@hdr.qut.edu.au
Authors: Sarini Sarini¹, James McGree¹, Nicole White¹, Kerrie Mengersen¹, Graham Kerr²
Affiliations: ¹School of Mathematical Sciences, QUT
²Institute of Health and Biomedical Innovations, QUT

Background: Falling is common for people with Parkinson’s disease (PD), giving negative consequences to their quality of life. While many studies have been conducted for the identification of the risk factors associated with whether a patient falls or not, the identification of factors contributing to the frequency of falls has not been adequately addressed due to failing to account for the excess zero counts observed in early PD patients. To this end, we explored the use of the zero inflated Poisson (ZIP) and hurdle models to describe fall count data.

Methods: Measurements from instruments for clinical assessment of PD, postural balance, general well-being and socio-demographic taken on 50 patients with early stage PD were considered. Falls were assessed prospectively using monthly fall diaries for 6 months following baseline assessment. The ZIP and hurdle models were considered for the description of these data. Models were assessed based on the deviance information criterion (DIC), and important risk factors related to falls were selected using stochastic search variable selection (SSVS).

Results: Fifty per cent of the subjects experienced at least one fall throughout the 6 month period. Upon implementing SSVS, the Unified Parkinson’s Disease Rating Scale (UPDRS) and PD Questionnaire (PDQ39) appeared most useful in describing the collected data. Risk factors relating to whether a patient fell or not, as well as those contributing to the frequency of falls were identified. These included: previous falls, social support, depression, postural instability and gait disturbance, types of PD and rigidity. The ZIP and hurdle models yielded similar results in terms of the identified risk factors. However, the ZIP model was preferred under DIC. Further, goodness-of-fit as judged by posterior predictive checks confirmed that the ZIP model provided a reasonable fit to the data.

Conclusions: Amongst many instruments to assess PD, UPDRS seemed to be the most informative to predict falls, and PDQ39 was comparable to the UPDRS. In addition to the motoric assessment, mental health assessment seemed to be important as well in relation to the falls incidence.

Improved credible intervals for a latent variable using MCMC and a deterministic computer simulation model for wheat development
Presenter: Ester D. Meenken
E-mail: esther.meenken@plantandfood.co.nz
Authors: Ester D. Meenken¹,², Chris M. Triggs² and Hamish E. Brown¹
Affiliations: ¹The New Zealand Institute for Plant and Food Research, New Zealand
²Department of Statistics, The University of Auckland, New Zealand

Bio-physical agricultural computer simulation models are used to inform and support agronomic research and government policy for applications including food security and climate change adaptation (Rosenzweig, Jones et al. 2013). They are simplified mathematical representations of physiological and physical processes that occur in plants and soils in response to environmental and management drivers. Most are process based, and update at some discrete time-unit (e.g. hourly, daily, weekly). Assimilating real-time data to update the model predictions can provide valuable insights to their behaviour. The Bayesian hierarchical model framework provides the best paradigm to date to achieve this (Cressie and Wikle 2011):

The general form of the joint filtering model is

\[ P(Z_t, \theta \mid C_{12}) \propto P(C_t \mid Z_t, \theta)P(Z_t \mid C_{t-1}, \theta)P(\theta) \]

where \( C \mid Z, \theta \) is the data model, \( Z \mid \theta \) is the state model, and \( \theta \) is the input parameter model.

Early and accurate estimation of the final number of leaves (\( f_{ln} \)) a wheat plant will develop can provide valuable information about when the crop will be ready for harvest. However, real-time data cannot be observed for this latent variable. However, data relating to the current number of leaves (\( ln \)) is available. We describe an approach to reduce the credible intervals of \( f_{ln} \) as simulated by an agricultural computer simulation model using observed data for \( ln \) via the Bayesian hierarchical model framework.

References
Immersive Mixed Reality environments as modern interactive visualisation platforms
Presenter: Tomasz Bednarz
E-mail: tomasz.bednarz@qut.edu.au
Authors: Tomasz Bednarz, June Kim, Ross Brown, Allan James, Steven Psaltis, Erin Peterson, Gavin Winter, Ian Turner and Kerrie Mengersen
Affiliation: Queensland University of Technology

Nowadays, industry and academia are interested in novel visualisation systems to improve operational efficiency, communicate their message, enhance expert elicitation, etc. Technologies to enhance participants' experience are advancing but still there is a lack of evidence supporting the extent to which these emerging technologies positively affect user experience and performance. In this poster, we describe an initiative of immersive panoramic display systems that could be used in various scenarios in a more effective way compared to traditional systems. This idea represents a step towards new platforms utilising 360-degree video experiences with annotations of the information coming from (big data) analytics engines. At the poster session, we will show a few use cases and demos, where the Mixed Reality technologies are effectively used to bring new levels to human-data interactions.

Interactive mirroring: Being and thing become space – Science and art interplays
Presenter: June Kim
E-mail: june.kim@qut.edu.au
Authors: June Kim and Tomasz Bednarz
Affiliation: Queensland University of Technology

Our poster showcases interactive art installation that conceptualizes reflective behaviour on ourselves through a virtual mirror. Spectator(s) reflect and manipulate themselves on the projected physical model of a balloon, in this case a bunny or sack shaped 3D meshes acting as a mirror. During development of this project, we investigated two key modes of interaction, which are through human breath and through hand actions: grabbing, pulling, and releasing. Hand movements are also used to create a force fields that drives balloon movements. Our Virtual Reality system, connects physical simulations, interactions with abstract artistic expression. It equally inspires artists and scientists. During the BoB conference, portable version of this interactive art installation will be showcased. This work demonstrates effective use of complex scientific simulations to drive artistic expressions. Art of interactive visualisation is a great tool for communicating science.
Hierarchical clustering of mixture component allocations: exploring data-driven subgroups in a case-control study of Alzheimer’s disease
Presenter: Zoé van Havre
E-mail: zoevanhavre@gmail.com
Authors: Zoé van Havre, Victor Villemagne and James Doecke
Affiliation: Commonwealth Scientific and Industrial Research Organisation (CSIRO)

The neurological drivers of Alzheimer’s disease (AD) begin years prior to the development of noticeable symptoms, and as such AD can lie undetected for up to 20 years. Improving early detection is crucial, leading to earlier access to preventative treatment options. A well-known feature of AD progression is a build-up of β-amyloid in the brain, a starch-like protein which interferes with the connections between neurone. In this study, the levels of β-amyloid in 33 regions of the brain (n=507) is analysed to determine the number of normally distributed mixture components present within each region, using overfitted Bayesian mixture models and the Zmix method. The allocations of each individual patient and region to the resulting clusters is modelled using a secondary hierarchical clustering procedure, to identify subgroups with similar patterns of β-amyloid levels across all regions. These new subgroups are investigated given known disease status, age, and other available demographics related to disease progression.

Eliciting and encoding expert knowledge on variable selection for species distribution models
Presenter: Ramethaa Pirathiban
E-mail: r.jeyapalan@qut.edu.au
Authors: Ramethaa Pirathiban1, Kristen J Williams2, Tony Pettitt1 and Samantha J Low-Choy1,3
Affiliations: 1School of Mathematical Sciences, Queensland University of Technology
2CSIRO Land and Water Flagship
3Griffith Social and Behavioural Research College, Griffith University

In species distribution models (SDM), the number of potential variables available in spatial datasets, imagery and acoustic sensing are increasing rapidly. With the increase in the number of potential variables, a problem which remains is deciding which environmental variables to be included into the model. Appropriate selection of environmental variables may greatly affect the performance of the model and its potential application for prediction or explanation. It also affects the cost of compilation and storage of future data. In the last decade, the challenge of variable selection for model fitting has increasingly been addressed through post-hoc model comparisons, via indicators such as AIC, and has been strongly advocated from an information-theoretic perspective. Nevertheless, some researchers may purposefully limit their analyses to a few ecologically justifiable or easily measured predictors with the resultant potential for under-specified models. Others may draw upon a more comprehensive set of variables, but risk over-fitting. The risk of selecting too many variables is exacerbated by the current popularity of data-driven approaches such as Boosted Regression Trees and the so-called MAXENT.

Several elicitation methods have been constructed to capture expert knowledge within a Bayesian statistical framework of SDMs for input into explanatory models such as logistic regression, classification trees and hierarchical models such as conditional probability networks. Yet, to date, most elicitation methodology has focused on elicitation of model parameters, such as regression coefficients, rather than variable choice. Here we develop and refine implementation of an elicitation protocol on variable selection for SDM. We demonstrate how variable importance can be obtained then used to guide variable selection in SDMs, or to define priors within Bayesian SDMs. Furthermore this study will emphasise to consider explanatory power and ecological basis for choosing variables and select appropriate methods of assessing the quality of candidate variables.
Employee health is becoming more important as research continues to demonstrate that healthy workers are more efficient, and that a return on investment for improving health can be attained. In this study we explore case study questions from an industry partner. Example questions are:

- “Has there been a link between the banning of smoking tobacco on site and improvement/decreasing health of employees? Can we continue to track the health and wellbeing to see if there has been a long-term improvement?”
- “Can we track when the wellness programs were initiated on site and the trends of employees health – are they improving, getting worse? What elements of their health, which SEGs?”
- “Can we use the trends from previous medical data to reflect possible future risks [risk profile] for certain individuals/groups of workers on site?”
- “Over the course of the mine life/data collection have there been changes or trends in increasing hearing loss with correlation to key site changes eg plant or equipment changes, trends between maintenance or shut downs etc. Can this be looked at within SEGs and overall site exposure?”

Standard frequentist statistical approaches for each question are applied. However, these standard approaches do not fully account for the complexities of the dataset, such as missing data, hierarchical structure, and repeated measurements. A Bayesian approach to these case study questions allows for these challenges to be approached more thoroughly. Bayesian approaches that accommodate these features are then described. Applications and inferences for industry are then discussed.

Typically Bayesian networks (BNs) are verified through calibration, which in ecology involves ground-truthing to field data. However where BNs are developed for risk assessment, calibration is not possible. Often the reason for relying on expert knowledge to estimate components of risk is that empirical data is not yet available. This was the situation encountered in the case study that motivated this research. A participatory approach was used to compile a BN, which was a regulatory requirement for evaluating the chance that No More Harm would occur on release of a genetically modified organism (GMO). Mosquitos are the vector for Dengue virus, which can have devastating effect on human health. By modifying the males of the species it was possible to prevent reproduction of the mosquitos, and thereby control the population. However before release of any GMO, a risk assessment is required.

The usual process for compiling BNs involves a preparatory phase to formulate the problem, a qualitative phase to propose the model structure, a quantitative phase to estimate the marginal and conditional probabilities in the network, followed by a verification phase to check this complex model. In this work, we have explored two methods for verification. The first involved pooling expert estimates of the conditional probability tables, and using a double Rasch model to separate out the effects of experts and hazards. In addition, three of the authors (SLC, JM, PdB) were involved in design and delivery of a specific phase of verification. Since all nodes in the BN evaluated the chance that a hazard would occur (either marginally or depending on parent hazards), we were able to help experts to evaluate the order of magnitude of hazards by eliciting the marginal probability of each node (edge nodes as well as internal nodes within the graph). This second analysis showed how a double Rasch model could be adapted to separate out the effects of experts and hazards on these marginal probabilities.

This work adds two tools to the toolbox for verifying Bayesian networks, by providing feedback to experts and modellers on the sources of consensus and dissensus (disagreement or vagueness) about BN nodes. This is achieved whilst accommodating multiple sources of variation, due to elicitation error (as presented in earlier work by Albert et al, 2012) as well as pooling across experts and items (based on earlier work by Low-Choy, Mittinty and Mengersen, 2011). It exploits the Rasch model (Rasch 1960) to provide a more informative alternative to the standard approach of linear pooling for combining expert opinion (O’Hagan et al 2006; Genest and Zidek 1986).
Evaluating framing effects in surveys of climate change beliefs: an iterative modelling strategy for integrating and refining psychological & cultural theory

Presenter: Cameron Williams
E-mail: cameron.williams@connect.qut.edu.au
Authors: Cameron Williams¹, Kristy Kitto¹ and Sama Low-Choy¹,²
Affiliations: ¹Queensland University of Technology
²Griffith Social & Behavioural Research College, Griffith University

In 2013, CSIRO conducted a survey to investigate the relationships between Climate Change beliefs (CCB) and respondents' attitudes and psychological profiles. The questions in the survey were presented to respondents in one of four orders, that re-arranged questions on Climate Change, Attitudes to Science and Political Beliefs. The aim of this research project was to determine whether order affected the reported CCBs, as suggested by general results from cognitive science, expert elicitation and psychometric testing. This has important implications for surveying populations about CCB as well as beliefs in general.

With a complex dataset of 102 questions aggregated into 22 constructs and a sample of size 533, an iterative strategy was used to build an appropriate model. Initial models that harnessed no previous information on model structure performed quite poorly. A conceptual model of the relationships between variables was elicited from the survey creators, and used as the basis for hierarchical models. These models were later expanded to include elicited priors. Order effects did not appear significant as main effects and required hierarchical models to assess interactions.

During the modelling process, the role of previously validated constructs for Right Wing Authoritarianism and Social Dominance Orientation were re-examined. As a result the questions within these constructs were remapped to an alternative scale aligned with Grid-Group cultural theory. Bayesian models addressed two opposite logical perspectives. A multivariate normal model was used to characterize grid-groups for CC believers and nonbelievers, and a Bernoulli regression used to predict CCB, knowing grid-group.

Assessing the accuracy of record linkages with Markov Chain based simulation approach

Presenter: Shovanur Haque
E-mail: shovanur.haque@hdr.qut.edu.au
Authors: Shovanur Haque and Steven Stern
Affiliation: School of Mathematical Sciences, Queensland University of Technology

Data linkage or record linkage is the process of finding matches and linking records from one or more data sources such that the linked records represent the same entity. An entity might be a business, a person, or some other type of listed unit. Record linkage is increasingly being used by statistical organisations to link administrative, survey and census files to create a robust file. Despite an increasing number of applications of record linkage, there has been little work done assuring the quality of linked files. To ensure that the matched records in the combined file actually correspond to the same individual is crucial for the validity of any analyses based on the combined data. The main focus of the research is to assess the likelihood that records matched from the two files actually belong to the same individual. This paper proposes a Markov Chain based simulation approach for assessing standard errors of estimates deriving from the analysis of the linked data. An agreement array will be created from all linking fields across all records in the two linked files and then simulate following a defined algorithm developed maintaining internal consistency patterns of agreement while preserving underlying probabilistic linking structure.
List of Participants

Aldo Saavedra  Aldo.Saavedra@sydney.edu.au
Aleysha Thomas  a12.thomas@qut.edu.au
Amy Cook  a20.cook@student.qut.edu.au
Anthony Ebert  21732327@student.uwa.edu.au
Anthony Pettitt  a.pettitt@qut.edu.au
Antonietta Mira  antonietta.mira@usi.ch
Antony Overstall  Antony.Overstall@glasgow.ac.uk
Beatrix Jones  m.b.jones@massey.ac.nz
Borek Puza  borek.puza@anu.edu.au
Brigitte Colin  b.colin@qut.edu.au
Cameron Williams  cameron.williams@connect.qut.edu.au
Carla Chen  c.ewels@aims.gov.au
Cathy Hargrave  c.hargrave@qut.edu.au
Chatanya Joshi  cjoshi@waikato.ac.nz
Chris Oates  christopher.oates@uts.edu.au
Christopher Drovandi  c.drovandi@qut.edu.au
Craig Anderson  craig.anderson@uts.edu.au
Daniel Kennedy  daniel.kennedy@hdr.qut.edu.au
David Smith  david.smith@flinders.edu.au
Dianne Cook  dicook@monash.edu
Earl Duncan  earl.duncan@qut.edu.au
Ester Meenken  esther.meenken@plantandfood.co.nz
Fabrizio Ruggeri  fabrizio@mi.imati.cnr.it
Fadhah Alanazi  fadhahamer.alanazi@hdr.qut.edu.au
Gang Xie  g.xie@griffith.edu.au
Irit Aitkin  irit.aitkin@unimelb.edu.au
James Lewis  j21.lewis@qut.edu.au
Jenna Thompson  jenna.thompson@qut.edu.au
Jessie Roberts  jessie.roberts@qut.edu.au
Julie Vercelloni  j.vercelloni@qut.edu.au
June Kim  june.kim@qut.edu.au
Kate Devitt  kate.devitt@qut.edu.au
Kathleen Byrnes  byrneskem@bigpond.com
Kerrie Mengersen  kmengersen@qut.edu.au
Leah Price  leah.south@qut.edu.au
Marcela Cespedes  marcela.cespedes@hdr.qut.edu.au
Maria Adamou  M.Adamou@soton.ac.uk
Martine Barons  Martine.Barons@warwick.ac.uk
Matt Shepit  matt@panalysis.com
Maxwell King  Max.King@monash.edu
Michael Spence  m.a.spence@sheffield.ac.uk
Miles McBain  m.mcba@qut.edu.au
Murray Aitkin  Murray.aitkin@unimelb.edu.au
Neil Gordon  neil.gordon@dsto.defence.gov.au
Nial Friel  nial.friel@ucd.ie
Nicholas Hughes  nickjhughes@gmail.com
<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicholas Tierney</td>
<td><a href="mailto:nicholas.tierney@qut.edu.au">nicholas.tierney@qut.edu.au</a></td>
</tr>
<tr>
<td>Nicole White</td>
<td><a href="mailto:nm.white@qut.edu.au">nm.white@qut.edu.au</a></td>
</tr>
<tr>
<td>Paul Blackwell</td>
<td><a href="mailto:p.blackwell@sheffield.ac.uk">p.blackwell@sheffield.ac.uk</a></td>
</tr>
<tr>
<td>Paul Wu</td>
<td><a href="mailto:p.wu@qut.edu.au">p.wu@qut.edu.au</a></td>
</tr>
<tr>
<td>Peter Baker</td>
<td><a href="mailto:p.baker1@uq.edu.au">p.baker1@uq.edu.au</a></td>
</tr>
<tr>
<td>Ramethaa Jeyapalan</td>
<td><a href="mailto:r.jeyapalan@qut.edu.au">r.jeyapalan@qut.edu.au</a></td>
</tr>
<tr>
<td>Renate Meyer</td>
<td><a href="mailto:renate.meyer@auckland.ac.nz">renate.meyer@auckland.ac.nz</a></td>
</tr>
<tr>
<td>Richard Woodman</td>
<td><a href="mailto:Richard.Woodman@flinders.edu.au">Richard.Woodman@flinders.edu.au</a></td>
</tr>
<tr>
<td>Richard Boys</td>
<td><a href="mailto:richard.boys@ncl.ac.uk">richard.boys@ncl.ac.uk</a></td>
</tr>
<tr>
<td>Russell Millar</td>
<td><a href="mailto:r.millar@auckland.ac.nz">r.millar@auckland.ac.nz</a></td>
</tr>
<tr>
<td>Sama Low-Choy</td>
<td><a href="mailto:s.low-choy@griffith.edu.au">s.low-choy@griffith.edu.au</a></td>
</tr>
<tr>
<td>Samuel Clifford</td>
<td><a href="mailto:samuel.clifford@qut.edu.au">samuel.clifford@qut.edu.au</a></td>
</tr>
<tr>
<td>Samuel Rathmanner</td>
<td><a href="mailto:samuel.rathmanner@qut.edu.au">samuel.rathmanner@qut.edu.au</a></td>
</tr>
<tr>
<td>Sarini Sarini</td>
<td><a href="mailto:sarini@hdr.qut.edu.au">sarini@hdr.qut.edu.au</a></td>
</tr>
<tr>
<td>Shovanur Haque</td>
<td><a href="mailto:shovanur.haque@hdr.qut.edu.au">shovanur.haque@hdr.qut.edu.au</a></td>
</tr>
<tr>
<td>Susanna Cramb</td>
<td><a href="mailto:susannacramb@cancerqld.org.au">susannacramb@cancerqld.org.au</a></td>
</tr>
<tr>
<td>Tomasz Bednarz</td>
<td><a href="mailto:tomasz.bednarz@qut.edu.au">tomasz.bednarz@qut.edu.au</a></td>
</tr>
<tr>
<td>Vasileios Giagos</td>
<td><a href="mailto:vgiagos@une.edu.au">vgiagos@une.edu.au</a></td>
</tr>
<tr>
<td>Wen-Hsi Yang</td>
<td><a href="mailto:Wen-Hsi.Yang@csiro.au">Wen-Hsi.Yang@csiro.au</a></td>
</tr>
<tr>
<td>Xing Lee</td>
<td><a href="mailto:xj.lee@qut.edu.au">xj.lee@qut.edu.au</a></td>
</tr>
<tr>
<td>Yi-Fan Lin</td>
<td><a href="mailto:C3192283@uon.edu.au">C3192283@uon.edu.au</a></td>
</tr>
<tr>
<td>Yiolanda Englezou</td>
<td><a href="mailto:ye1g14@soton.ac.uk">ye1g14@soton.ac.uk</a></td>
</tr>
<tr>
<td>Zoé van Havre</td>
<td><a href="mailto:zoevanhavre@gmail.com">zoevanhavre@gmail.com</a></td>
</tr>
</tbody>
</table>