#BIBC2025

International Biometrics Society Australasian Region

Biometrics in the Bush Capital

Canberra, Australia
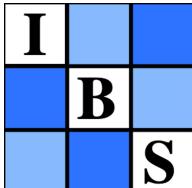
24 – 28th November 2025

Table of contents

About

The International Biometric Society (IBS) is devoted to the development and application of statistical and mathematical theory and methods in the Biosciences, including agriculture, biomedical science and public health, ecology, environmental sciences, forestry, and allied disciplines. It welcomes as members statisticians, mathematicians, biological scientists, and others devoted to interdisciplinary efforts in advancing the collection and interpretation of information in the biosciences.

The IBS is organised into geographically-defined regions, each with at least 50 members. The Australasian Region (IBS-AR) is one of over 30 regions under the IBS umbrella. It provides to members of the IBS who reside in Australia, New Zealand and nearby countries a local focus and support for biometrics undertaken in this region.

The Australasian Region Biometrics conference is a biannual conference sponsored by the IBS-AR. The conference is a forum for biometricians, statisticians, and those interested in the development and application of statistical and mathematical theory and methods to problems in the biosciences. The conference is open to members and non-members of the Society.

Society website: biometricsociety.org.au

Conference website: biometricsociety.org.au/conference2025

IBS-AR Council

- **Alison Kelly** (University of Queensland) as President
- **James Curran** (University of Auckland) as Vice-President
- **Louise McMillan** (Victoria University of Wellington) as Secretary
- **Ian Renner** (University of Newcastle) as Treasurer
- **Sam Rogers** (University of Adelaide) as Membership Officer
- **Emi Tanaka** (Australian National University) as Social Media Officer
- **Warren Muller** (CSIRO) as Public Officer
- **Charley Budgeon** (University of Western Australia) as Biometric Bulletin Correspondent
- **Peter Green** (AgResearch Ltd) as Website manager
- **Samuel Muller** (Macquarie University) as IBS Council Member
- **Kevin Murray** (University of Western Australia) as IBS Council Member
- **Chris Triggs** (University of Auckland) as IBS Executive Board

Last updated: 2025-11-05 17:59:10

Organisers

Local Organising Committee

- **Emi Tanaka** (Australian National University) as Chair
- **Alan Welsh** (Australian National University)
- **Fonti Kar** (Australian National University)
- **Joanne Thandrayen**
- **Warren Muller** (CSIRO)
- **Weihao (Patrick) Li** (Australian National University)
- **Yidi Deng** (Australian National University)

Social Program Committee

- **Quan Vu** (Australian National University) as Chair

Scientific and Program Committee

- **David Warton** (UNSW Sydney) as Chair
- **Carolyn Huston** (CSIRO)
- **Emi Tanaka** (Australian National University)
- **Jean Yang** (University of Sydney)
- **Julian Taylor** (Adelaide University)
- **Nicola Armstrong** (Curtin University)
- **Thomas Lumley** (University of Auckland)

Conference Allies

- **Fonti Kar** (Australian National University)
- **Louise McMillan** (Victoria University of Wellington)
- **Scott Foster** (CSIRO)

Sponsors

We gratefully acknowledge our sponsors for their generous support, which has made this conference possible.

Platinum

Gold

Supporter

ACT Branch

Schedule

Monday, 24 November

08:30-09:00	Workshop	Haydon Allen	Emailed to participants
09:00-17:00	Workshop		
17:30-18:00	General	Ovolo Nishi	Ovolo Nishi
18:00-20:30	Social		Ovolo Nishi

Tuesday, 25 November

08:15-08:45	General	Foyer	Foyer		
08:45-09:05	General	Arc Cinema	Arc Cinema		
09:05-09:55	Keynote	Arc Cinema Chair: Chris Brien	Arc Cinema	Chris Brien	
10:00-11:00	Talk	Chair: Garth Tarr	Arc Cinema	Garth Tarr	
			Arc Cinema	Garth Tarr	
			Arc Cinema	Garth Tarr	

		Theatrette		
10:00-11:00	Talk	Chair: Scott Foster	Theatrette	Scott Foster
			Theatrette	Scott Foster
			Theatrette	Scott Foster
11:00-11:30	Social	Gallery	Gallery	
11:30-12:50	Talk	Arc Cinema Chair: Alan Welsh	Arc Cinema	Alan Welsh
			Arc Cinema	Alan Welsh
			Arc Cinema	Alan Welsh
			Arc Cinema	Alan Welsh
11:30-12:50	Talk	Theatrette Chair: James Curran	Theatrette	James Curran
			Theatrette	James Curran
			Theatrette	James Curran
			Theatrette	James Curran
12:50-13:50	Social	Arc Cinema Gallery	Gallery	
13:50-15:20	Invited Session	Chair: Francis Hui	Arc Cinema	Francis Hui
			Arc Cinema	Francis Hui
			Arc Cinema	Francis Hui

13:50-15:10	Talk	Theatrette Chair: Ruth Butler	Theatrette	Ruth Butler
			Theatrette	Ruth Butler
			Theatrette	Ruth Butler
			Theatrette	Ruth Butler
15:20-15:50	Social	Arc Cinema Gallery	Gallery	
15:50-16:50	Talk	Chair: Zhanglong Cao Theatrette	Arc Cinema	Zhanglong Cao
			Arc Cinema	Zhanglong Cao
			Arc Cinema	Zhanglong Cao
15:50-16:50	Talk	Chair: Graham Hepworth	Theatrette	Graham Hepworth
			Theatrette	Graham Hepworth
			Theatrette	Graham Hepworth

17:00-19:00	Poster	Gallery	Gallery	
			Gallery	
			Gallery	
			Gallery	
			Gallery	
			Gallery	
			Gallery	
			Gallery	
			Gallery	

Wednesday, 26 November

08:30-08:50	General	Foyer	Foyer	
08:50-08:55	General	Arc Cinema	Arc Cinema	
08:55-09:45	Keynote	Arc Cinema Chair: Thomas Lumley	Arc Cinema	Thomas Lumley
09:50-10:50	Talk	Chair: David Warton	Arc Cinema	David Warton
			Arc Cinema	David Warton
		Theatrette	Arc Cinema	David Warton
09:50-10:50	Talk	Chair: Sam Rogers	Theatrette	Sam Rogers
			Theatrette	Sam Rogers
			Theatrette	Sam Rogers
10:50-11:20	Social	Arc Cinema	Gallery	
11:20-12:50	Invited Session	Chair: Robert Clark	Arc Cinema	Robert Clark
			Arc Cinema	Robert Clark
			Arc Cinema	Robert Clark

11:20-12:40	Talk	Theatrette Chair: Linh Nghiem	Theatrette	Linh Nghiem
			Theatrette	Linh Nghiem
			Theatrette	Linh Nghiem
			Theatrette	Linh Nghiem
12:50-13:50	Social	Gallery	Gallery	
13:00-17:00	Social			

Thursday, 27 November

08:30-08:50	General	Foyer	Foyer	
08:50-08:55	General	Arc Cinema	Arc Cinema	
08:55-09:45	Keynote	Arc Cinema Chair: Warren Muller	Arc Cinema	Warren Muller
09:50-10:50	Talk	Chair: Vanessa Cave Theatrette	Arc Cinema	Vanessa Cave
			Arc Cinema	Vanessa Cave
			Arc Cinema	Vanessa Cave
09:50-10:50	Talk	Chair: Alice Richardson	Theatrette	Alice Richardson
			Theatrette	Alice Richardson
			Theatrette	Alice Richardson
10:50-11:20	Social	Gallery	Gallery	

			Arc Cinema	
11:20-12:40	Talk	Chair: Matthew Schofield	Arc Cinema	Matthew Schofield
			Arc Cinema	Matthew Schofield
			Arc Cinema	Matthew Schofield
			Arc Cinema	Matthew Schofield
11:20-12:40	Talk	Theatrette Chair: Chris Triggs	Theatrette	Chris Triggs
			Theatrette	Chris Triggs
			Theatrette	Chris Triggs
			Theatrette	Chris Triggs
12:40-13:40	Social	Gallery	Gallery	
13:00-13:30	General	Arc Cinema / Theattre	Theattre	
13:40-15:10	Invited Session	Chair: Emi Tanaka	Arc Cinema	Emi Tanaka
			Arc Cinema	Emi Tanaka
			Arc Cinema	Emi Tanaka
13:40-15:00	Talk	Theatrette Chair: Sam Mason	Theatrette	Sam Mason
			Theatrette	Sam Mason
			Theatrette	Sam Mason
			Theatrette	Sam Mason

15:10-15:40	Social	Gallery / Arc City	Gallery	
15:40-16:40	Talk	Chair: Julian Taylor Theatrette	Arc Cinema	Julian Taylor
			Arc Cinema	Julian Taylor
			Arc Cinema	Julian Taylor
15:40-17:00	Talk	Chair: Oyelola Adegboye	Theatrette	Oyelola Adegboye
			Theatrette	Oyelola Adegboye
			Theatrette	Oyelola Adegboye
			Theatrette	Oyelola Adegboye
18:30-22:00	Social	Banquet Hall	Banquet Hall	

Friday, 28 November

Time	Type	Location / Chair	Location	Person
08:40-09:00	General	Foyer	Foyer	
09:00-09:05	General	Arc Cinema	Arc Cinema	
09:05-09:55	Keynote	Arc Cinema Chair: Louise McMillan	Arc Cinema	Louise McMillan
10:00-11:00	Talk	Chair: Yidi Deng Theatrette	Arc Cinema	Yidi Deng
			Arc Cinema	Yidi Deng
			Arc Cinema	Yidi Deng
10:00-11:00	Talk	Chair: David Baird	Theatrette	David Baird
			Theatrette	David Baird
			Theatrette	David Baird
11:00-11:30	Social	Gallery	Gallery	
11:30-12:20	Keynote	Arc Cinema Chair: James Curran	Arc Cinema	James Curran
12:20-12:30	General	Arc Cinema	Arc Cinema	
12:30-13:30	Social	Outdoor	Outdoor	

DRAFT

Keynote talks

Optimizing research impact through interdisciplinary and collaborative research

Charmaine B. Dean (University of Waterloo)

Interdisciplinary collaborative research is a key component of data science and for some of us, plays an important part of our roles as statisticians. It is not unusual that we become accustomed to vertical thinking whereby we use existing tools and methods in our own specialty to problem solve, losing sight of the larger interdisciplinary context of data science, and the context of the scientific challenge. The Government of Canada - Science and Technology branch has identified several priority research challenge topics that involve cross-disciplinary work. Although statistical tools and analytics are identified in these research challenge priority areas, additionally, the development of fundamental transformative and enabling technological tools specifically for statistical methods and analytics to support research and societal advancement is also seen as a priority. This talk shares insights about the challenges and opportunities for statistics in interdisciplinary research. Specifically, monitoring viral signals in wastewater and assessing forest fire risk are given as complex, case studies that use a collaborative and interdisciplinary approach to solve difficult problems. This approach will demonstrate the significant benefits for not only optimizing research impact but for training students to become horizontal problem solvers across a wide range of research methods which will benefit them in navigating complex problems and in the development of appropriate tools for their analysis

Saddlepoint approximations for likelihoods

Jesse Goodman (University of Auckland)

Co-authors: Godrick Oketch and Rachel Fewster; Alice Hankin, Sixiang Shan and George Fan

Classically, the saddlepoint approximation has been used as a systematic method for converting a known generating function into an approximation for an unknown density function. More recently, it has been used instead as an approximation to the *likelihood function*. In this viewpoint, it is the underlying data-generating process whose generating function is used, and the saddlepoint approximation can be maximized to compute an approximate saddlepoint MLE for given observed data. This talk will explain how the saddlepoint approximation can be interpreted with a statistical lens, including common features for those otherwise intractable models where we can compute a generating function but not a likelihood. Many of these models come from statistical ecology, including sitations where we gather population-wide observations only. In addition, the talk will describe a class of models having simple theoretical guarantees for the effect of using the saddlepoint MLE as a substitute for the unknown true MLE, and will demonstrate new tools to visualize the saddlepoint approximation intuitively, to simplify and automate the computation of saddlepoint MLEs, and to quantitatively assess the amount of approximation error introduced by using an approximate likelihood as a substitute for an intractable true likelihood.

Modularizing Biometric Models Facilitates Multistage Computing

Mevin B. Hooten (The University of Texas at Austin)

Bayesian modeling has become invaluable in biometrics. It allows us to formally consider unobserved processes while accommodating uncertainty about data collection and our understanding of biological and ecological mechanisms. Several excellent software packages are available for fitting Bayesian models to data and are being applied every day to analyze biometric data. These methods allow us to answer questions using data in ways that has never before been possible. The adoption of Bayesian methods has led to bigger models necessary to answer tough questions using large and varied data sets. Bigger models and data sets lead to computing bottlenecks. Fortunately, a solution to Bayesian computing roadblocks sits in plain sight. The structure of Bayesian models allows us to rearrange them so that we can perform computing in stages. We can break big models into pieces, fit them separately, and then recombine them in later computing stages. Recursive Bayesian approaches can save us time by leveraging the parallel architecture of modern computers. A modular perspective allows us to see Bayesian models in a way that facilitates multistage computing. I will demonstrate the procedure with a set of biometric examples. These include geostatistical models in marine science, capture-recapture models for abundance estimation, and spatial point process models for species distributions.

On Finding Good Experiments

Cheng Soon Ong (CSIRO and ANU)

One of the key choices we have as scientists is to design informative experiments. With computational methods like AI promising accurate predictions, we revisit the question of adaptively designing new measurements that take previous data into account. Using examples from genomics, we illustrate some recent ideas on using machine learning to recommend experiments. Then we discuss potential impacts on choosing measurements in spatiotemporal problems. We conclude by outlining some opportunities and challenges of including machine learning in the scientific discovery process.

Uses of gnm for Generalized (Non)linear Modelling

Heather L. Turner (University of Warwick)

The R package {gnm} was designed as a unified interface to fit Generalized Nonlinear Models: *generalized* to handle responses with restricted range and/or a variance that depends on the mean, and *nonlinear* to allow the predictor for the mean to be nonlinear in its parameters. This framework covers several models that were proposed in the literature and adopted in practice before {gnm} was released, but used to require a mixed bag of specialised software to fit.

With {gnm} celebrating its 20th birthday this year, it's a good time to review how the package is being used. I'll highlight some of the applications we were aware of when {gnm} was first developed, that remain in common use, and explore more recent applications, particularly in the field of biometrics.

We'll discover one motivation for using {gnm}, is for the "eliminate" feature that efficiently estimates stratification parameters. This can be useful even when the predictor is linear, as in the case of using conditional Poisson models to analyse case-crossover studies in epidemiology. We'll also look at two of the packages that have built on {gnm}. The first, {multgee}, uses {gnm} to fit multiplicative interactions for certain correlation structures when modelling categorical data, with applications in public health, agriculture, and psychology. The second, {VFP}, is a more specialised package that uses {gnm} to model the mean-variance relationship for in-vitro diagnostic assays.

Through these use cases we'll see how different features of {gnm} can be applied, demonstrating the versatility of this software.

Invited talks

Inferring the rate of undetected contamination using random effects modelling of biosecurity screening histories

Sumonkanti Das (Australian National University)

Group testing plays a vital role in biosecurity operations worldwide, particularly in minimising the risk of introducing exotic pests, contaminants, and pathogens through imported agricultural products. A common screening strategy involves pooling items from consignments and testing each group for contamination presence, with consignments typically rejected if any group tests positive. Although screening designs often target a high probability of detection assuming a fixed minimum prevalence, analysing the historical results of these tests to infer the extent of contamination in non-rejected consignments (referred to as leakage) is less common.

This study advances censored beta-binomial (BB) models to address contamination risk in frozen seafood imports into Australia, incorporating imperfect tests. Motivated by the characteristics of our case study, we develop a new class of BB models that impose a minimum positive consignment propensity threshold, capturing scenarios where contamination is either absent or exceeds a known minimum level. To fit these models, we propose a Metropolis-Hastings (MH) algorithm conditioned on prior distributions for sensitivity and specificity, allowing efficient estimation of quantities related to contamination levels. We analyse historical testing data under multiple scenarios using the proposed MH algorithm, yielding novel insights into both contamination risk and leakage.

Finally, we use model-based simulations to communicate risk levels, providing key insights into potential undetected contamination.

Predication of Daily Weight Gain with Cattle Behaviour and Daily Activity Using Triaxial Accelerometer Data

Shuwen Hu (RMIT)

Co-authors: Reza Arablouei, Greg Bishop-Hurley, Neil Bagnall and Aaron Ingham

Live body weight gain is an important measurement of animal performance. In this study, we predict the cattle's daily weight gain from five core cattle behaviours and a measure of total daily activity based on accelerometer data. To collect data, we conducted an experiment equipping a herd of 60 Brahman steers with research collars containing triaxial accelerometers over nearly one month in Australia. We used the accelerometer data, which represents the intensity of animal movement, to compute an activity metric within a five-minute window. In addition, we use pre-trained accelerometer-based machine learning models to classify cattle behaviour into grazing, ruminating, resting, walking, drinking or other classes over five-second time windows. Daily behaviour profiles were constructed for each animal and experiment day by aggregating the behaviour predictions over every calendar day. Our objective was to explore how to use behaviours and activity metrics to predict the cattle's daily weight gain. The daily activity values ranging from 5.44g to 23.69g. The average daily time spent grazing, ruminating, resting, walking and drinking was 8.97±1.12, 7.78 ±1.03, 5.83±1.05, 1.00±0.4, and 0.18±0.12 hours, respectively. Some weather information data are combined in the model to predict the cattle live-weight gain. The best R-squared value is 0.467, with a minimum root mean

square error (RMSE) of 0.867 from the linear regression model. The live-weight gain could not be fully explained by measurements taken in this study, but we showed how these factors can influence the variability in cattle performance.

The difficulties of clustering categorical or mixed data

Louise McMillan (Victoria University of Wellington)

Co-authors: Daniel Fernández, Shirley Pledger, Richard Arnold, Ivy Liu, Eleni Matechou, Adam Glucksman

I will discuss techniques for clustering categorical data that go beyond treating the data as numerical or converting the categories to dummy variables. I will talk about a range of approaches to clustering categorical data, including the `clustord` R package for clustering ordinal data, which uses model-based clustering via likelihoods and the EM algorithm. Then I will discuss a Bayesian approach from population genetics that may be extendable to general mixed datasets and is the subject of my latest research.

Evaluating the impact of trait measurement error on genetic analysis of computer vision-based phenotypes

Gota Morota (The University of Tokyo)

Quantitative genetic analysis of image- or video-derived phenotypes is increasingly being performed for a wide range of traits. Pig body weight values estimated by a conventional approach or a computer vision system can be considered as two different measurements of the same trait, but with different sources of phenotyping error. Previous studies have shown that trait measurement error, defined as the difference between manually collected phenotypes and image-derived phenotypes, can be influenced by genetics, suggesting that the error is systematic rather than random and is more likely to lead to misleading quantitative genetic analysis results. Therefore, we investigated the effect of trait measurement error on genetic analysis of pig body weight (BW). Calibrated scale-based and image-based BW showed high coefficients of determination and goodness of fit. Genomic heritability estimates for scale-based and image-based BW were mostly identical across growth periods. Genomic heritability estimates for trait measurement error were consistently negligible, regardless of the choice of computer vision algorithm. In addition, genome-wide association analysis revealed no overlap between the top markers identified for scale-based BW and those associated with trait measurement error. Overall, the deep learning-based regressions outperformed the adaptive thresholding segmentation methods. This study showed that manually measured scale-based and image-based BW phenotypes yielded the same quantitative genetic results. We found no evidence that BW trait measurement error could be influenced, at least in part, by genetic factors. This suggests that trait measurement error in pig BW does not contain systematic errors that could bias downstream genetic analysis.

Optimal allocation of resources between control and surveillance for complex eradication scenarios

Mahdi Parsa (Department of Agriculture, Fisheries and Forestry)

Effective eradication of invasive species over large areas requires strategic allocation of resources between control measures and surveillance activities. This study presents an analytical Bayesian framework that integrates stochastic modelling and explicit measures of uncertainty to guide decisions in complex eradication scenarios. By applying Shannon entropy to quantify uncertainty and

incorporating the expected value of perfect information (EVPI), the framework identifies conditions under which investment into control or surveillance becomes worthwhile. Findings show that strategies which hedge against uncertainty can markedly improve the robustness of eradication outcomes with only marginal increases in expected costs. This approach offers practical tools for designing more cost-effective and reliable eradication programs and for prioritising data collection to reduce uncertainty where it has the greatest impact.

Surveillance for counterfactual scenarios of invasive species - why it's useful and a convenient way to do it.

Andrew P. Robinson (The University of Melbourne)

Co-authors: John Kean, Melissa Welsh, and Kuo-Szu Chiang

The impact of invasive species is affected by a range of factors, many of which can be anticipated in advance – for example, the prevalence of host material, climate suitability, the size of affected agricultural resources, and so on. One factor that cannot be anticipated in advance is the size of the incursion at the time of its detection. Unfortunately, the impact of an incursion is tightly tied to its maturity at detection, ranging from a single seed, for example, to a 50,000 hectare infestation of plants.

We propose a simple probability model for the detection of an invasive species that can either capture or integrate out the consequent uncertainty of the maturity of the incursion. We represent the relationship between surveillance and the detection of the organism using survival analysis: the detection of the incursion is analogous to the survival event; it is a binary occurrence that happens at some point in time, and once it has happened it does not happen again.

Under such a model, we can connect the distribution of the size of the infestation at the time to detection to the probability of detecting the incursion given that it has not already been detected, namely, the hazard function. For example, a popular model for the detection of an incursion of size x with number of traps t and probability of detecting a single pest p is:

$$h(x,t,p) = 1 - (1-p)^{tx}$$

Algebra leads us to a size-at-detection pdf. Other corrections are also applied as needed. The outcome is a pdf that is a function of process parameters, enabling straightforward assessment of different surveillance choices. Parameter estimates for the distribution can be derived from first principles, field experiments, or expert elicitation.

In this presentation we will derive and demonstrate the use of the survival-based incursion size at detection pdf and discuss its implications and challenges.

Fishing for Heritability in the Gill Microbiome: Why Statisticians Should get out into the field

Elle Saber (Australian National University)

Host‐associated microbiomes are increasingly recognised as integral to health, yet the extent to which host genetics shapes these communities remains unclear. While heritable components of gut and skin microbiomes have been documented in several vertebrates, evidence for the gill microbiome of fish is scarce. This exploratory study sampled the gill microbiome of Atlantic salmon within a Tasmanian breeding program to investigate potential genetic influences. Despite careful planning, study participants did not always behave, and the practical constraints of a commercial operation prevented complete data capture. Having participated in the data collection I was better prepared to understand the limitations of the dataset when doing the downstream analysis. The take home message is that, even if we feel like a fish out of water, time spent among the data can be just as

valuable as time spent our desks.

Bayesian clustered ensemble prediction for multivariate time series

Shonosuke Sugasawa (Keio University)

We propose a novel methodology called the mixture of Bayesian predictive syntheses (MBPS) for multiple time series count data and apply the methodology to predict the numbers of COVID-19 inpatients and isolated cases in Japan and Korea at the subnational level. MBPS combines a set of predictive models and partitions the multiple time series into clusters based on their contribution to predicting the outcome. In this way MBPS leverages the shared information within each cluster and avoids using a multivariate count model, which is generally cumbersome to develop and implement. Our data analyses demonstrate that the proposed MBPS methodology has improved predictive accuracy and uncertainty quantification.

Optimal sampling in border biosecurity: Application to skip-lot sampling

Raphael Trouve (The University of Melbourne)

Border biosecurity faces mounting pressure from increasing global trade, requiring cost-effective inspection strategies to reduce the risk of importing pest and diseases. Current international standards recommend inspecting all incoming consignments (full census) with fixed sample sizes (e.g., 600 units) for high-risk pathways, but this may be overkill for lower-risk pathways with established compliance records. When should agencies use skip-lot sampling (SLS, sometimes called continuous sampling plan), which adaptively reduces inspections based on recent compliance history, over full census inspection?

We developed a propagule pressure equation for SLS in overdispersed pathways and used Lagrange multipliers to derive a solution. Results show the choice depends on pathway overdispersion, sampling costs, and budget constraints. Optimal sample sizes are typically smaller than current recommendations, with better returns from inspecting a larger proportion of consignments rather than larger samples per consignment. This framework provides biosecurity agencies with data-driven guidance for implementing adaptive sampling strategies.

Species archetype models for presence-only data

Skipton N.C. Woolley (CSIRO)

Co-authors: David Warton, Piers Dunstan

Joint species distribution modelling is a recently emerged and potentially powerful statistical method for analysing biodiversity data. Despite the plethora of presence-only occurrence data available for biodiversity analysis, there remain few examples of presence-only multiple-species modelling approaches. We present a mixture-of-regressions model for understanding how groups of species are distributed based on presence-only data. Our approach extends Species Archetype Models using a point process framework and incorporates joint estimation of sighting bias based on all species occurrences included in the model. We demonstrate that our method can accurately recapture mean and variance of parameters from simulated data sets and provide better estimates than those generated from multiple single species presence-only species distribution models. We apply our approach to a Myrtaceae presence-only occurrence dataset from New South Wales, Australia. We show that presence-only Species Archetype Models allow for the propagation of variance and uncertainty from the data through predictions, improving inference made on presence-only biodiversity data for multiple species.

Contributed talks

Handling Missingness in Prevalence Estimates from National Surveys

Oyelola Adegboye (Menzies School of Health Research)

Co-authors: Li Siyu, Tomoki Fujii and Denis HY Leung

Population-based surveys, such as the Demographic and Health Surveys (DHS), are pivotal for estimating the prevalence of important diseases, particularly in low-resource settings. However, missing data, such as person non-response, particularly refusals, poses a serious challenge, potentially introducing substantial bias in prevalence estimates. Using HIV data sets from Malawi's 2004 DHS, Antenatal Clinics surveillance, and the Malawi Diffusion and Ideational Change Project (MDICP), this paper evaluates existing estimators and proposes novel approaches to adjust for refusal bias. These include complete case analysis, mean score imputation, inverse propensity score weighting, bounding techniques using longitudinal or sentinel data and Manski's partial identification approach. The study distinguishes between refusals and non-contacts and examines how prior knowledge of HIV status influences participation. Estimates of HIV prevalence varied notably across methods, with refusal-adjusted approaches generally yielding higher prevalence rates than complete case estimates. The divergence is more pronounced among women, largely due to higher refusal rates among them. Bounding methods provide credible intervals for prevalence under weak assumptions. In the case of HIV estimates, refusal bias, particularly when linked to prior testing knowledge, can significantly distort HIV estimates. Integrating multiple data sources and using methodologically transparent adjustment techniques are critical for robust HIV surveillance in low-resource settings.

Reporting Odds Ratios under Fluctuating Reporting Rates in Spontaneous Reporting Systems

Tatsuhiko Anzai (Institute of Science Tokyo)

Co-authors: Kunihiko Takahashi

Spontaneous adverse event reporting systems, including the Japanese Adverse Drug Event Report (JADER) database, the FDA Adverse Event Reporting System (FAERS), and VigiBase, play a critical role in post-marketing drug safety surveillance. The reporting odds ratio (ROR) is a commonly used measure for detecting suspected adverse drug reactions as "signals", based on disproportionality in reporting between a specific drug and all others, as summarized in a two-by-two contingency table. However, during events such as the COVID-19 pandemic, reporting rates of adverse drug reactions can fluctuate, potentially introducing bias into ROR. This study proposes a method for deriving the ROR that incorporates four parameters, each corresponding to fluctuations in reporting ratios for the cells in the two-by-two contingency table. These parameters can be estimated using the divergence between observed and predicted values, where the predicted values are derived from a regression model under the assumption that the reporting rates for the drug and the adverse reaction remain stable. We evaluate the properties and performance of the proposed method through both real-world data analysis and simulation studies, demonstrating its effectiveness for signal detection.

Pooled testing with penalized regression models

Christopher Bilder (University of Nebraska-Lincoln, Department of Statistics)

Co-authors: Pranta Das, Joshua Tebbs, and Christopher McMahan

Pooled testing (also known as group testing) is a widely used procedure to test individuals for infectious diseases. Rather than testing each specimen separately, multiple specimens are pooled together and tested as one. The pool test outcome, along with further tests as needed, are used to determine which individuals are positive or negative for an infectious pathogen. The COVID-19 pandemic especially highlighted the importance of pooled testing, with laboratories adopting it worldwide to increase their testing capacity. Pooled testing traditionally relies on observing the positive/negative test outcomes alone. However, during the pandemic, new pooled testing algorithms were developed that utilize the viral load information from a pool test. These new algorithms use penalized regression models to predict the viral load of each individual, which lead to individual positive/negative predictions. The purpose of our presentation is to provide a comparison of these new algorithms relative to standard ones. Both gains and losses are quantified by applying algorithms under fair comparison settings.

Nested-factorial treatment models: types, their uses and examples

Chris Brien (The Plant Accelerator, University of Adelaide)

Most commonly factorial experiments involve several crossed treatment factors and the treatment model includes terms for all possible combinations of the treatment factors. However, treatment models that employ nested treatment factors are beneficial in a number of situations and could be used more often. Applicable situations include those in which (i) treatment factors are intrinsically nested, (ii) treatments involve one or more factors plus a control and (iii) treatments have been systematically allocated and/or are pseudoreplicated. Four basic types of nested factors that start with a single nesting factor are (i) one nested factor, (ii) multiple, crossed, nested factors, (iii) multiple, independent, nested factors, and (iv) a hierarchy of multiple, nested factors. Examples involving the different situations and types of nesting will be presented. Constructing the factors using the R package dae (Brien, 2025, https://cran.r-project.org/package=dae) will be described and the properties of the models discussed for the examples via their anatomies produced using dae.

Bayesian Ordinal Regression for Crop Development and Disease Assessment

Zhanglong Cao (Centre for Crop and Disease Management, Curtin University)

Co-authors: Rose Megirian, Matthew Nguyen and Adam Sparks

Accurate assessment of crop development and disease severity is essential for informed agronomic decision-making. This study presents a Bayesian framework for analysing ordinal data from field trials, including growth scale progression and disease ranking scores. Using the brms package in R, we applied cumulative logit models to evaluate the effects of sowing depth and treatment combinations on cereal growth stages, measured via the Zadok's scale, across two Western Australian sites (Merredin and Wickepin). The same framework is being extended to model disease severity scores, demonstrating its versatility across categorical biological measurements. Our workflow incorporates rigorous model testing and evaluation, including posterior predictive checks and leave-one-out cross-validation (LOO-CV), to ensure robust inference and model fit. Rather than relying on p-values from linear mixed models, the Bayesian approach provides interpretable probabilities of achieving specific growth stages or disease severity levels. This shift enables more nuanced understanding of treatment effects and supports decision-making under uncertainty.

Do Mice Matter? The Impact of Mice on a New Zealand Ecosanctuary

Vanessa Cave (VSNi)

Mammal-resistant fences have enabled the successful eradication of exotic mammals from ecosanctuaries in New Zealand. However, preventing the re-invasion of mice remains a challenge. Indeed, mice are still present in many fenced ecosanctuaries and can reach high population densities.

Scientists at Manaaki Whenua – Landcare Research have been studying the impact of mice on biodiversity at Sanctuary Mountain Maungatautari. Two independently fenced sites within the sanctuary were managed to create contrasting mouse populations: one with high mouse numbers, and the other with undetectable levels. After two years, management protocols were reversed, with mice eradicated from the first site and allowed to increase at the second. Data on the abundance of invertebrates, seedlings, and fungi were collected throughout the duration of the study.

Temporal trends in abundance were analysed and compared using linear mixed models with smoothing splines. The findings suggest that mice can have severe impacts on native biodiversity, particularly invertebrates, posing a significant threat to ecological recovery efforts in fenced ecosanctuaries.

Automatic debiased machine learning (autoDML) for causal inference: implementation and evaluation in real-world observational studies

Tong Chen (Murdoch Children's Research Institute)

Co-authors: Stijn Vansteelandt and Margarita Moreno-Betancur

The estimation of average treatment effect (ATE) is a central goal of observational health research. Causal machine learning (CML) methods are increasingly popular for ATE estimation, as they leverage data-adaptive algorithms to account for complex relationships amongst confounders, exposure and outcome while ensuring valid inference. Major CML methods, such as double machine learning and targeted maximum likelihood estimation, offer robustness and efficiency advantages but they require first estimating the propensity score and then calculating treatment divided by the propensity score, known as the Riesz representer. Standard machine learning algorithms for the propensity score estimation are suboptimal as they are designed to minimise prediction error, ignoring that equally large estimation errors in small and large propensity scores propagate very differently into the ATE estimates upon taking the reciprocal. Automatic debiased machine learning (autoDML), introduced by Chernozhukov et al. (2021), addresses this issue by directly estimating the Riesz representer using machine learning methods, yielding a more stable alternative to standard CML methods. Despite its theoretical advances, autoDML remains largely unknown in biostatistical research and has not been compared to standard CML methods in realistic simulation studies. In this talk, we bridge these gaps by introducing autoDML and highlighting its advantages. We describe a practical, step-by-step guide for implementing autoDML, and compare autoDML with standard CML methods using comprehensive, realistic simulation studies. We further illustrate the utility of autoDML using data from the Longitudinal Study of Australian Children to evaluate the impact of overweight or obesity on cardiovascular outcomes in adolescence.

A Set of Precise Asymptotics for Gaussian Variational Approximation in Generalised Linear Mixed Models

Nelson J. Y. Chua (Australian National University)

Co-authors: Francis K. C. Hui, Luca Maestrini, A. H. Welsh

Generalised linear mixed models (GLMMs) are used to model data with a clustered or hierarchical

structure, capturing intra-cluster dependence through the use of (latent) random effects. However, this dependence structure results in a likelihood function that is typically computationally intensive to evaluate, and so approximate likelihood approaches are often used instead for model fitting and inference. One such approach which has grown in popularity over the past decade is Gaussian variational approximation (GVA). In addition to estimates of the model parameters, GVA concurrently produces random effects predictions and associated uncertainty measures.

In this talk, we formulate a set of precise asymptotic properties for both the parameter estimates and random effects predictions obtained from GVA, focusing on independent cluster GLMMs. We find that these properties change substantially depending on whether or not the random effects are conditioned on. By comparing GVA's asymptotic properties and empirical finite-sample performance with that of other commonly used approximate likelihood approaches, we highlight situations in which the use of GVA would be advantageous in practice.

Estimating abundance in small populations using pedigree reconstruction

Sarah Croft (University of Otago)

Co-authors: Jamie Sanderlin, Michael Black, Richard Barker and Matthew Schofield

Accurate measures of abundance are essential for successful monitoring of animal populations, and for assessing the efficacy of conservation interventions. In previous work, genotypic information has been incorporated into Mark-Recapture models to enable the identification of individuals, as well as determination of kinship between observed individuals in Close-Kin Mark-Recapture (CKMR) models. Generally, CKMR models make large sample assumptions limiting their application to many endangered and at-risk species. We have developed Bayesian methodology to estimate population abundance and dynamics for small, isolated populations using pedigree reconstruction. The true underlying pedigree completely describes the abundance and population structure over time, however the true relationships between individuals in wild populations are rarely known. Given a set of observed genotypes, along with supplementary data, our methodology is able to successfully reconstruct the pedigree without the need for large sample assumptions. Prior knowledge of the mating structure and reproductive dynamics of the population can also be incorporated in the model. In this talk I will present our pedigree reconstruction approach for population estimation using dead recovery data, and will discuss the challenges associated with the full pedigree approach.

Genstat Markdown: Reproducible Research with Genstat

James M. Curran (University of Auckland)

Co-authors: Yuxiao Wang, Simon Urbanek, and David Baird

** This talk should not be accepted unless you are desperate :-) **

Reproducible research ensures that scientific findings can be independently verified by others using the same data and methods. Markdown plays a key role in supporting reproducibility by allowing researchers to combine code, results, and narrative text in a clear, lightweight format. Tools like R Markdown or Jupyter Notebooks use Markdown to embed executable code alongside explanations and outputs, making it easy to document workflows, share analyses, and regenerate results consistently across different environments.

Genstat is a statistical software package designed for data analysis, particularly in biometrical research and applications. Many users know it through its user-friendly interface. However, it also It offers a powerful scripting language.

In this talk I will describe a collaborative project between students and staff at the University of

Auckland and VSNi Ltd, that allows Genstat to be used in conjunction with R Markdown to provide flexible reproducible research capability to the Genstat user community.

Multivariate meta-analysis methods for high-dimensional data

Alysha M. De Livera (La Trobe University)

Co-authors: Luke Prendergast, Robert Staudte, and Jayamini Liyanage

Multivariate meta-analysis methods for high-dimensional data

Meta-analysis is a statistical method that combines quantitative results from multiple independent studies on a particular research question or hypothesis, with the goal of making inference about the population effect size of interest. Traditional meta-analysis methods have focused on combining results from multiple independent studies, each of which has measured an effect size associated with a single outcome of interest. Modern studies in evidence synthesis, such as those in biological studies have focused on combining results from studies which have measured multiple effect sizes associated with multiple correlated outcomes. We will present a novel, multivariate meta-analysis method for obtaining summary estimates of the effect sizes of interest for high-dimensional data, with applications to real and simulated high-dimensional data. We will discuss advantages, disadvantages and the statistical challenges, and present a CRAN-based R package for implementation of the methods.

StableMate: a regression framework for selecting stable predictors across heterogeneous data environments

Yidi Deng (Melbourne Integrative Genomics; Department of Anatomy and Physiology, The University of Melbourne; Research School of Finance, Actuarial Studies and Statistics, The Australian National University)

Co-authors: Jiadong Mao, Jarny Choi, Kim-Anh L√™ Cao

Inferring reproducible relationships between biological variables remains a challenge in the statistical analysis of omics data where $p > 10{,}000$ and $n < 500$. Methods that identify statistical associations lack interpretability or reproducibility. We can address these limitations by inferring stable associations that are robust to external perturbation on the data. Stable associations can be an implication in causality since causal relationships are necessarily stable in some sense (Pearl et al. 2009). Unstable associations can also be of interests in certain biological applications to study functional heterogeneity in a biological system.

We developed a new regression framework, StableMate based on the concept of stabilised regression (SR), which utilise heterogenous data to enforce stability (Pfister et al. 2021). Given datasets generated from different environments, such as experiments or disease states, StableMate 1. identifies stable predictors with consistent functional dependency with the response across environments. 2. builds a robust regression model with stable predictors to enforce generalisable prediction in unseen environments. The ultimate aim is to build selection ensembles. However, unlike SR that selects stable predictors by performing stability tests on every possible predictor subset, StableMate optimizes efficiency with a greedy search based on our improved stochastic stepwise selection algorithm. In a simulation study, we show that StableMate outperformed SR for both variable selection and prediction and significantly reduces running time. In three case studies of cancer with different omics data types, we show that StableMate can also address a wide range of biological questions.

Fitting integrated species distribution models using mgcv

Elliot Dovers (UNSW Sydney)

Co-authors: Jakub Stoklosa, David L. Miller, David I. Warton

Integrated species distribution models (ISDMs) are a useful tool for ecologists, allowing them to use multiple sources of data to infer the distribution of species across geographic regions. Recent studies have shown that including latent spatial terms to account for dependence structures within and between datasets is often crucial to the performance these models. However, the inclusion of these latent terms can make the ISDMs technically challenging to fit, and often require users to learn/adopt bespoke software.

We describe how ISDMs can be fitted using `mgcv` on `R` by using the grouped family feature (`gfam`). This permits multiple likelihoods - and hence multiple data types - to be fitted within the same model by using each dataset to inform the estimation of common parameters. We additionally show how smoothers over the geographic coordinates that approximate Gaussian random fields can be used to incorporate the necessary latent spatial effects. We use presence/absence and presence-only data to demonstrate that ISDMs fitted via `mgcv` have comparable performance to other available software - through both simulations and in application, where we predict the occurrence of a tree species in NSW, Australia.

Continental-Scale Bayesian Analysis of Acacia Flowering Phenology: A Novel Framework Integrating Phylogenetic Signal and Circular Statistics

Owen Forbes (CSIRO)

Co-authors: Peter H. Thrall, Andrew G. Young

Climate change is driving widespread shifts in plant phenology globally, with critical implications for ecosystem functioning and species interactions. While most phenological studies focus on temperate deciduous species at local scales, Australia's diverse Acacia genus offers unique opportunities to understand phenological responses to variations in rainfall and temperature across continental gradients. Leveraging digitised herbarium collections, we present the largest phylogenetically-informed analysis of acacia flowering phenology undertaken to date, spanning 548 species across Australia's full climatic spectrum.

We developed a novel phenological modelling approach using a Bayesian hierarchical framework that makes several key innovations. First, we extend traditional phylogenetic least squares (PGLS) approaches by incorporating phylogenetic signal as a structured random effect within a fully Bayesian context, enabling simultaneous estimation of phylogenetic covariance alongside spatial and temporal dependencies. Second, we employ von Mises circular distributions to model flowering time, naturally accommodating seasonal cyclicity while avoiding linear boundary artifacts. This modular approach also integrates spatial autocorrelation structure, time-lagged climate covariates, and phylogenetic relationships within a unified probabilistic framework. Our analysis of 20,998 herbarium specimens spans 110 years (1910 – 2020) across Australia, using climate observations from the Australian Gridded Climate Data, with Köppen climate zones included in the model to stratify climate-phenology relationships.

This framework successfully quantifies species-specific climate sensitivities while accounting for phylogenetic constraints. Expected effect of time-lagged temperature and rainfall effects vary significantly across Köppen climate zones and species. Species vulnerability patterns could help identify taxa which could be potentially at higher risk under projected climate scenarios. This modular Bayesian approach demonstrates the research potential of natural history collections for continental-scale ecological inference, providing a statistical template for global phenological studies addressing critical climate

change impacts.

Using a linear mixed model based wavelet transform to model non-smooth trends arising from designed experiments

Clayton Forknall (Queensland Department of Primary Industries)

Co-authors: Alison Kelly, Yoni Nazarathy and Ari Verbyla

The linear mixed model (LMM) representation of the cubic smoothing spline is a powerful tool for modelling smooth trends arising from designed experiments. However, when trends arising from such experiments are non-smooth, meaning that they are characterised by jagged features, spikes and/or regions of rapid change approximating discontinuities, smoothing spline techniques prove ineffective. A technique that has proven useful for the modelling of non-smooth trends is the wavelet transform. Existing methods to incorporate the wavelet transform into the LMM framework are varied, but often share a common limitation; that is, a reliance on classical wavelet approaches that require observations to be equidistant and dyadic ($\log_2(n)$ is an integer) in number. More recently, second generation wavelet methods have been developed, which overcome the limiting constraints imposed by classical wavelet approaches, enabling the wavelet transform to be applied to sets of non-equidistant observations, of any number. We present a method for the incorporation of these second generation wavelets, namely second generation B-spline wavelets, into an LMM framework to facilitate the wavelet transform. Furthermore, using the structure implicit in the resulting B-spline wavelet basis, we propose extensions to the LMM framework to enable heterogeneity of the associated wavelet variance across wavelet scales. This provides a new LMM based method which enables the flexible modelling of non-smooth trends arising in the conduct of designed experiments. The proposed method is demonstrated through application to a data set exhibiting non-smooth characteristics, that arises from a designed experiment exploring the proteome of barley malt.

Koala Distirbution and Abundance

Scott D. Foster (Data61, CSIRO)

Co-authors: Scott D. Foster, Wen-Hsi Yang, Peter Caley, John McEvoy, David Ramsey, Drew Terasaki Hart, David E. Uribe-Rivera, Romane H. Cristescu, Eric Vanderdys, Adam McKeown, Corey J. A. Bradshaw, Emma Lawrence, Andrew Hoskins

The koala (*Phascolarctos cinereus*) is a well-known and studied Australian marsupial, but the species presents a complex case for conservation. Currently, most koala conservation efforts focus on local-scale population estimates, which are often based on expert opinion or anecdotal evidence, and either precede or ignore most (or all) available data. In contrast, conservation listing advice and associated recovery plans require population estimates at the species-range scale. A data-driven, national-scale population estimate including all available information will therefore guide effective management of koala populations by providing high-quality and objective information. To this end, we have designed nationally consistent survey, implemented it (with the aid of partners) and analysed the resulting data (and others). The design uses recently developed techniques (clustered spatially-balanced designs) whilst the analysis uses emerging models that incorporate multiple data types (e.g. point process, binary and count) that are often-enough collected using different equipment, protocols and staff. Integrated species distribution models (ISDMs) have, at their heart, a simple point process but this simplicity still allows for some complexity in terms of how different types of data can inform the point process. The model, when fitted to koala data, indicates that the distribution of koalas is patchy throughout much of eastern Australia. It also infers that there are more koalas than previously guessed. Our estimates provide unprecedented evidence to support nationally consistent and spatially explicit decision-making for koala conservation, and do so with relevant measures of uncertainty.

A semi-supervised framework for diverse multiple hypothesis testing scenarios

Jack Freestone (Macquarie University)

Co-authors: William S. Noble and Uri Keich

Standard multiple testing procedures are designed to report a list of discoveries, or suspected false null hypotheses, given the hypotheses' p-values or test scores. Recently there has been a growing interest in enhancing such procedures by combining additional information with the primary p-value or score. In line with this idea, we develop RESET (REScoring via Estimating and Training), which uses a unique data-splitting protocol that subsequently allows any semi-supervised learning approach to factor in the available side information while maintaining finite sample error rate control. Our practical implementation, RESET Ensemble, selects from an ensemble of classification algorithms so that it is compatible with a range of multiple testing scenarios without the need for the user to select the appropriate one. We apply RESET to both p-value and competition based multiple testing problems and show that RESET is (1) power-wise competitive, (2) fast compared to most tools and (3) able to uniquely achieve finite sample false discovery rate or false discovery exceedance control, depending on the user's preference.

Fitting Generalised Linear Mixed Models using Sequential Quadratic Programming

Peter Green (Tadorna Data Science)

Finding maximum likelihood estimates in generalised linear mixed models (GLMMs) can be difficult due to the often-intractable integral over the random effects.

Ensuring convergence can be tricky, especially in binomial GLMMs, and often multiple optimisers and settings need to be tried to get satisfactory results. This is exacerbated if you want to use parametric bootstrap for your inference.

Sequential quadratic programming (SQP) is a method for solving optimisation problems with non-linear constraints. SQP offers an alternative option for maximising the Laplace approximation to the GLMM likelihood, bypassing the need for an inner penalised iterative reweighted least squares (PIRLS) loop.

This talk discusses the implementation of SQP for GLMMs and compares its performance to other common approaches.

Modelling Species Diversity from Citizen-Science Bird Counts

Graham Hepworth (Statistical Consulting Centre, The University of Melbourne)

Co-authors: Jeremy Silver, William Steele, Al Danger

Bird count data from the Greater Melbourne region over a 25-year period were recorded by volunteer observers, based on 20-minute surveys. Counts for a large range of bird species were reported either as counts or simply as "present". Our investigation focused on whether there has been an increase in the presence and abundance of noisy miners, an endemic species known for its aggressive and territorial behaviour, and whether this could partly explain apparent declines in other species.

The dataset was augmented with weather, geospatial and ecological information, and a mixed-effects modelling framework was adopted. Seasonal effects and long-term trends were accounted for via regression splines. The influence of different observers and locations were represented as random effects. GAMMs were fitted for (a) the presence/absence of noisy miners, (b) the abundance of noisy miners, and (c) species richness for particular groupings of birds. In modelling species richness, noisy

miner abundance was included as a predictor; we used multiple imputation to account for the missing data when this species was reported as "present" but no count was recorded.

There was a small, positive association found between noisy miner abundance and all-species diversity – a surprising result. A stronger negative association was found for some species groupings. For many species groups, diversity tended to decrease with increasing distance from the nearest native vegetation feature, and with increasing wind speed. The work provided several statistical and computational challenges.

Data-Adaptive Automatic Threshold Calibration for Stability Selection

Martin Huang (The University of Sydney)

Co-authors: Samuel Muller, Garth Tarr

Stability selection has gained popularity as a method for enhancing the performance of variable selection algorithms while controlling false discovery rates. However, achieving these desirable properties depends on correctly specifying the stable threshold parameter, which can be challenging. An arbitrary choice of this parameter can substantially alter the set of selected variables, as the variables' selection probabilities are inherently data-dependent. To address this issue, we propose Exclusion Automatic Threshold Selection (EATS), a data-adaptive algorithm that streamlines stability selection by automating the threshold specification process. EATS initially filters out potential noise variables using an exclusion probability threshold, derived from applying stability selection to a randomly shuffled version of the dataset. Following this, EATS selects the stable threshold parameter using the elbow method, balancing the marginal utility of including additional variables against the risk of selecting superfluous variables. We evaluate our approach through an extensive simulation study, benchmarking across commonly used variable selection algorithms and static stable threshold values.

Model-based assessment of functional and phylogenetic diversity

Shaoqian Huang (School of Mathematics and Statistics, UNSW)

Co-authors: David Warton

Topic Biodiversity is a quantity of fundamental interest, but measuring it remains challenging. Two key aspects of diversity are functional diversity (FD)—the extent to which species in a community differ in their ecological functions—and phylogenetic diversity (PD)—the extent to which species differ in their evolutionary histories.

Limitations Commonly used distance-based diversity measures, such as Rao's Q and Faith's PD, are typically affected by sampling intensity, which refers to the level of effort invested in collecting samples. Specifically, these measures confound true diversity changes with changes in species richness driven by different sampling intensities. In addition, there is no established statistical framework for analyzing how diversity changes along environmental gradients.

Our Model We propose a model-based assessment of PD and FD changes along gradients. It is robust to variations in sampling intensity, as the model explicitly captures the main trend in abundance change—namely, changes in species richness—and then isolates the residual variation, the β-diversity change. We then measure PD or FD change using linear contrasts of the β-diversity parameters, where the contrasts correspond to the main axes of the phylogenetic or functional similarity matrix.

Simulation Analysis Based on the presence/absence pattern, we conducted simulations in which we designed phylogenetic distance matrices with different numbers of species, modelled on real data. The results show that our approach tends to outperform methods currently used in ecology.

Scalable finite mixture of regression models for clustering species responses in ecology

Francis KC Hui (The Australian National University)

When modeling species assemblages in ecology, clustering species with similar responses to the environment can facilitate a more parsimonious understanding of the assemblage, and improve prediction by borrowing strength across species within the same cluster. One statistical method for achieving the above is species archetype models (SAMs), a type of finite mixture of regression model where species are clustered according to the *shape* of their environmental response.

In this talk, we introduce approximate and scalable SAMs or asSAMs, which overcomes some of the current computational drawbacks in fitting SAMs. We show how asSAMs promotes fast uncertainty quantification via bootstrapping, along with fast variable selection on archetypal regression coefficients courtesy of a sparsity-inducing penalty. Simulation studies and an application to presence-absence records of over 230 species from the Great Barrier Reef Seabed biodiversity project demonstrate asSAMs can achieve similar to or better estimation, selection, and predictive performance than several existing methods in the literature.

Using point cloud data to discover genomic regions associated with dynamic height

Colleen H Hunt (University of Queensland)

Co-authors: David R. Jordan and Emma S. Mace

Plant height in grain sorghum (*Sorghum bicolor L. Moench*), influences yield, lodging resistance, and harvestability, yet conventional methods usually measure plant height at maturity, but this single-time-point measurement misses the growth dynamics across the season. Tracking and analysing height over time can reveal more about the genetic factors controlling development rate and pattern. Data from 881 sorghum lines from a diverse population were planted in a partially replicated design with 1190 plots. UAV-based high-throughput phenotyping was used to generate weekly plant height measurements from emergence to flowering. These aerial images were collected and processed using photogrammetric software to create dense point clouds and canopy height models for each plot. A logistic regression model within a linear mixed model framework was applied to effectively depict the sigmoidal growth pattern typical of sorghum, with parameters indicating maximum height, growth rate, and inflection point. These biologically relevant parameters offered a straightforward, quantitative summary of each genotype's growth trajectory. The logistic growth parameters were then used in a genome-wide association study (GWAS) to identify loci linked to the dynamic aspects of height development, with several significant associations detected. This is among the first studies to combine logistic growth modelling of high-throughput UAV data with GWAS to dissect the genetic control of growth dynamics in sorghum. Our findings underline the usefulness of logistic regression for modelling crop growth and showcase the effectiveness of combining UAV phenotyping, statistical growth modelling, and GWAS to identify genomic regions that influence developmental trajectories in sorghum.

Group Sampling with Imperfect Testing for Biosecurity Applications

Adele Jackson (ABARES)

Co-authors: Belinda Barnes and Mahdi Parsa

Group sampling, also known as pooled or batch sampling, is a standard technique in biological sciences and the health sector to use limited resources efficiently. Common objectives include detecting pest species and inferring prevalence of disease in communities, livestock or wildlife. The purpose of this paper is to support the design of robust group sampling strategies when testing processes are

imperfect. We formulate analytical distributions and statistics for grouped hypergeometric sampling and its binomial approximation that incorporate a variety of types of imperfect test. These include tests that respond to the presence or absence of contaminated material in a group, as well as PCR and serological testing processes where sensitivity depends on the number of contaminated items in each group. We also formulate the Hellinger information of a sampling scheme, which allows us to develop group sampling design strategies that increase the accuracy of inferred prevalence. This is an essential component in decision-making during outbreaks of disease and in operational biosecurity applications. Based on this work, we estimate leakage through a grouped sampling scheme and show how accounting for leakage can alter sampling strategies and improve risk management decisions.

The geometry of diet: using projections to quantify the similarity between sets of dietary patterns

Beatrix Jones (University of Auckland)

Co-authors: Shiyu Fan, Karen Mumme, John Thompson

Food consumption is complex and high dimensional. Nutrition researchers measure (or attempt to measure) consumption using food frequency questionnaires, food recalls, or food records. This generates high dimensional data which is frequently summarized using principle components, principle components with rotation, or factor analysis. The resulting linear combinations are called "dietary patterns." In this talk we consider quantifying how similar two different sets of dietary patterns are, assuming the same set of underlying variables has been collected. We explore using a multivariate extension of Tucker's congruence coefficient for this purpose. Tucker's congruence coefficient can be thought of as the length of the projection of one factor direction onto another, or equivalently the (absolute) cosine of the angle between them; thus it ranges from 0 to 1. In (eg) two dimensions, we consider the square root of the area of a unit square projected from one space onto another; this can be generalized for any number of dimensions. The measure is symmetric and invariant to rotation. To contextualize scores on this measure, we compute this measure of agreement for several datasets from the dietary pattern validation literature, where the same food questionnaire is given to the same people a few weeks apart. We also consider the effect of truncating small loadings, as is common when describing dietary patterns.

The performance of Yu and Hoff's confidence intervals for treatment means in a one-way layout

Paul Kabaila (La Trobe University)

Consider a one-way layout and suppose that we have uncertain prior information that the treatment population means are equal or close to equal. Yu & Hoff (2018) extended the "tail method" for finding a confidence interval for a scalar parameter of interest that has (a) specified coverage probability and (b) relatively small expected length when this parameter takes values in some given set. They used this extension to find confidence intervals for these treatment means that have (a) specified coverage probability individually and (b) relatively small expected lengths when this uncertain prior information happens to be correct. They assessed the expected lengths of these confidence intervals, over the whole parameter space, using a semi-Bayesian analysis. I describe a revealing alternative assessment of these expected lengths using a fully frequentist analysis.

Yu, C. & Hoff, P. (2018) Adaptive multigroup confidence intervals with coverage. *Biometrika*, 105, 319-335.

`heritable` An R package for heritability calculations for plant breeding trials

Fonti Kar (Australian National University)

Co-authors: Emi Tanaka

Understanding how much biological variation is heritable - or passed down from one generation to next - is central to plant breeding. Heritability is a useful indictor to assess the genetic gain of desirable traits such as yield or disease resistance in crop trials. We created `heritable`, an R package to streamline the calculation of heritability from `asreml` and `lme4` mixed model outputs in a user-friendly manner. The package provides up to six different types of heritability measures and helper functions to compare between these. Our goal is to support decision makers by providing an intuitive, open and reproducible workflow in calculating heritability in R.

Visualize your fitted non-linear dimension reduction model in the high-dimensional data space

P. G. Jayani Lakshika (Monash University, Australia)

Co-authors: Dianne Cook, Paul Harrison, Michael Lydeamore, Thiyanga S. Talagala

Non-linear dimension reduction (NLDR) techniques such as t-SNE, UMAP, PHATE, PaCMAP and TriMAP provide a low-dimensional representation of high-dimensional data by applying a non-linear transformation. The methods and parameter choices can create wildly different representations, so much so that it is difficult to decide which is best, or whether any or all are accurate or misleading. NLDR often exaggerates random patterns, sometimes due to the samples observed, but NLDR views have an important role in data analysis because, if done well, they provide a concise visual (and conceptual) summary of high-dimensional distributions. To help evaluate the NLDR, we have developed a way to take the fitted model, as represented by the positions of points in 2D, and turn it into a high-dimensional wireframe to overlay on the data, viewing it with a tour. Viewing a model in the data space is an ideal way to examine the fit. One can see whether it fits the points everywhere or fits better in some places, or simply mismatches the pattern. It is used here to help with the difficult decision on which 2D layout is the best representation of the high-dimensional distribution, or whether the 2D layout is displaying mostly random structure. It can also help to see how different layouts made by different methods are effectively the same summary, or how the different methods have some particular quirks. This methodology is available in the R package `quollr`. We will demonstrate the technique using single-cell data, particularly to understand cluster structure.

A covariate-adaptive test for replicability across multiple studies with false discovery rate control

Dennis Leung (University of Melbourne)

Co-authors: Ninh Tran

Replicability is a gold standard for credible scientific discoveries. The partial conjunction (PC) p-value proposed by Benjamini and Heller (Biometrics, 2008), which summarizes individual base p-values obtained from multiple similar studies, can gauge whether a signal of interest is replicable. However, when a large set of features are examined by these studies (such as comparable genomewide association studies conducted by different labs), testing for their replicated signals simultaneously can pose a very underpowered problem, due to both the multiplicity correction required and the inherent limitations of PC p-values. This power deficiency is particularly severe when replication is demanded for all studies, the most natural benchmark a practitioner performing meta-analyses may request.

We propose ParFilter, a procedure that combines the ideas of filtering and covariate-adaptiveness to power up large-scale testing for replicated signals as described above. It validly reduces the multiplicity burden by partitioning the studies into smaller groups and borrowing between-group information to filter out unpromising features. Moreover, harnessing side information offered by available covariates, it trains hypothesis weights to encourage rejections of features more likely to

exhibit replicated signals. We prove its finite-sample false discovery rate control under standard assumptions on the dependence of the base p-values across features. In simulations as well as a real case study on autoimmunity based on RNA-Seq data obtained from thymic cells, the ParFilter has demonstrated competitive performance against other existing methods for such replicability analyses.

Running Human Subject Experiments via Online Crowdsourcing

Patrick Li (Australian National University)

Crowdsourcing platforms such as Amazon Mechanical Turk and Prolific provide scalable and accessible tools for running online human subject experiments. This talk offers an overview of how to design, set up, and manage such studies effectively. Drawing on experience from multiple projects, I will walk through the key steps for obtaining ethics approval, compare platform workflows, and discuss considerations for participant recruitment, screening, task design, quality control, cost estimation, and common pitfalls. This session is intended for researchers planning to conduct behavioural, perceptual, or decision-making experiments, as well as those developing data annotation pipelines for machine learning or applied research.

Risk of Guillain-Barré Syndrome after COVID-19 vaccination and SARS-CoV-2 infection: A multinational self-controlled case series study

Han Lu (University of Auckland)

Co-authors: Sharifa Nasreen, Yannan Jiang, Arier Lee, Helen Petousis-Harris, Jim Buttery, Steve Black, and Jeffrey Kwong

Introduction: The Global Vaccine Data Network (GVDN) is a multinational collaboration established to conduct globally coordinated epidemiological studies on vaccine safety using health data. This study aimed to assess the risk of Guillain-Barré syndrome (GBS) within 42 days after exposure to any COVID-19 vaccines and SARS-CoV-2 infection, combing multisite data using a common GVDN protocol and data model.

Methods: We used a self-controlled case series (SCCS) design and identified GBS cases via electronic data sources (EDS) from 20 GVDN sites globally. Fifteen sites performed medical chart review using Brighton Collaboration case definitions to determine the level of diagnostic certainty (LOC). The relative incidence (RI) between pre-defined risk and control windows was calculated using conditional Poisson regression, controlling for seasonality. De-identified case-level data were combined for individual patient data (IPD) analysis, and the estimates were aggregated with otherwise locally analysed site-level results to calculate the overall RI with 95% confidence interval (95%CI) using random effects meta-analysis.

Results: We identified 2086 GBS cases (4329 vaccine doses) from EDS and 410 cases were chart-reviewed (12% with LOC 1-2). We observed an association between Vaxzevria/Covishield vaccine and GBS among LOC 1-2 cases within 42 days post-exposure (RI 3.10, 95%CI 1.12-8.62). No increased risk of GBS was observed with mRNA and other vaccines. Among 489 EDS -identified cases post-infection, significant risk was found between SARS-CoV-2 and GBS (RI 3.35, 95%CI 1.83-6.11). The SCCS method is an efficient approach to collect minimal data for IPD meta-analysis. Full results using different SCCS models will be discussed.

Crossvalidation for predictive models in complex survey data

Thomas Lumley (University of Auckland)

Co-authors: Amaia Iparragirre

Cross-validation is a standard technique for choosing models with good out-of-sample prediction error in statistical learning. Chopping data up at random makes sense for independent observations, but not for observations from a multistage survey. For example, when a single cluster is split between test and training sets there is the potential for data leakage and underestimation of prediction error. I will describe an approach to cross-validation for complex survey data using replicate weights and describe its implementation in the R survey package.

Building Trust Without Peer Review: Establishing Reproducibility Standards in Industrial Statistical Consulting

Dean Marchiori (Wave Data Labs)

In academic research, peer review serves as a cornerstone for ensuring the credibility and reproducibility of findings. However, in industry, statistical consultants often operate without the formal structures of peer review, posing gaps in ensuring the reproducibility and trustworthiness of their analyses.

This talk explores how statistical consultants for industry can proactively establish and adhere to standards that promote reproducibility and foster trust, even in the absence of traditional peer review mechanisms.

Researchers, data scientists and statisticians often engage with and provide expertise to government, industry and other groups outside of academia. This expertise is often trusted implicitly and relied on to make important decisions. On the other hand, there are many well known criticisms of conventional peer review systems both in academic publishing and in commercial work.

Drawing on practices from various industries, we will propose a pragmatic framework for the development and implementation of standardized reporting formats, the use of open-source tools for reproducible analysis, and the adoption of best practices for documentation and code sharing. We will also examine the role of professional organizations and industry consortia in setting guidelines that encourage transparency and accountability.

Attendees will gain insights into practical strategies for implementing reproducibility standards in their own consulting and research practices and contribute to a broader movement towards transparency and accountability in industrial statistics.

Accounting for heterogeneous detection rates when inferring eradication of an invasive species

Sean A. Martin (QUT School of Mathematics/SAEF)

Co-authors: Leah South, Zachary T. Carter, Michael Bode

Eradication of invasive species from islands is an expensive but often crucial activity for maintaining the ecosystems of these islands. A key challenge in limiting the costs of an eradication campaign is knowing the size of the remaining population of invasive animals; this information informs both ongoing resource planning and when to end a campaign. Campaign progress is measured by the decline in detections of the target species. However, detection of a population is imperfect and highly variable, obscuring the true rate of decline. As a result, uncertainty arises when detections cease - have all individuals truly been eradicated, or do some remain undetected? Statistical models have been developed to infer eradication success based on a record of detections (sighting records), however, most models ignore natural heterogeneity in detectability. Our work examines the influence of variance in detection rates on the false positive and false negative error rates of hypothetical eradication campaigns and relates this to overall costs of island conservation. We use a combination of Bayesian likelihood and likelihood-free models to overcome simultaneous inference limits on

abundance and detectability within this highly stochastic study system.

Spatio-Temporal Species Distribution Modelling

Sam Mason (UNSW)

Detecting species response to climate change is a critical concern in ecology requiring relevant climatic predictors over long temporal windows and large spatial extents. To date this has been technologically challenging resulting in the widespread use of 30 year averaged datasets which, while readily accessible, actually mask species responses by effectively treating the climate as temporally static.

Recent advances in data acquisition technology have made it easier to obtain a wide range of environmental predictors at appropriate spatio-temporal scales and at consistent resolutions needed for effective research.

By using dynamic predictors and anomalies from their mean values in spatio-temporal species distribution models we see if we can detect species response to climate change and if this yields better predictive performance than current practice. Interestingly, we have not found appreciable spatio-temporal signal, in contrast to expectations and other work which used naïve approaches to analysis.

Multi-environment trial analysis of count data with complex variance structures using generalised linear mixed models

Michael H. Mumford (Queensland Department of Primary Industries)

Co-authors: Alan Huang, Alison M. Kelly

The analysis of response data from agricultural field experiments conducted in multi-environment trials (METs) is typically performed using linear mixed models (LMMs). The strength of the LMM framework is in the ability to seamlessly model experimental design terms and to include complex variance structures, especially for unbalanced data. When the response variable is count data, the assumptions underpinning the LMM are violated and it is necessary to extend to a generalised linear mixed model (GLMM) approach. Statistical modelling using GLMMs introduces additional complexities arising from a combination of (i) susceptibility to large estimation biases due to approximations of the marginal likelihood, (ii) accounting for heterogeneity of variance/dispersion, and (iii) the increase in computational resources required. In this talk, a statistical methodology is proposed for the MET analysis of count data. The analysis approach uses a GLMM framework, assuming an underlying mean-parameterised Conway-Maxwell Poisson distribution, that can account for arbitrarily under and over-dispersed count data. This framework enables partitioning of residual variation from genetic and other extraneous sources of variation, and adopts a factor-analytic model for the genotype by environment interaction effects. The proposed methodology is applied to a series of common bean trials, where the aim is genotype selection for the response variable pod count per plant. The analysis is implemented using the `glmmTMB` R-package, which uses automatic differentiation to enhance computational speed, the Laplace approximation estimation method to reduce estimation biases, and a residual maximum likelihood (REML)-like correction to further reduce estimation biases for variance components.

Better Conversations, Better Support: Strengthening Consulting through Practical Education and Community

Sharon G. Nielsen (Adelaide University)

Co-authors: Sam Rogers and Annie Conway

Effective statistical consulting in agriculture depends not just on technical expertise, but on shared understanding, clear communication, and the confidence of those seeking advice. Through its series of practical agronomic workshops and the linked Community of Practice, the Biometry Hub has shown how targeted education can transform the way consulting support is delivered and received.

The workshops, typically four core sessions delivered over five days, cover the design of robust experiments, sound data collection, practical analysis, and clear interpretation of results. By equipping researchers and industry staff with these practical skills, the workshops lay the groundwork for more productive conversations between statisticians and collaborators.

The Community of Practice then extends this support beyond the workshop room. Monthly meetings held over Zoom create an informal space for participants to revisit concepts, tackle real-world questions, and build confidence alongside peers and mentors. This talk will share how combining hands-on training, tailored tools like the biometryassist R package, and an ongoing network of support strengthens the consulting relationship, helping ensure that good statistical practice is not only taught, but actively put into practice across the grains industry.

Teaching Meta-Analysis for Systematic Reviewers with Mixed Statistical Training

Xu Ning (University of New South Wales)

Co-authors: David K E Chan, Coralie Williams, and Peter Humburg

Higher-degree research (HDR) students often consider conducting a systematic review of their research domain, with the goal of synthesising their findings with a meta-analysis. However, the statistical training of HDR students and their supervisors can range from little to comprehensive. Consequently, we found in our consultations with these students that they may struggle in specifying their meta-analytical model, misinterpret their model outputs and diagnostics, or both, especially with complex meta-analytical models. Hence, we have developed a short course on meta-analysis to address these knowledge gaps. The course aims to develop students' mastery in meta-analytical methods. We will present empirical findings on whether the course's aim was met and discuss the pedagogical considerations made to accommodate the range of statistical training in our target cohort.

A circular hidden Markov model for directional time series data

A.A.P.N.M. Perera (The Australian National University)

Co-authors: Francis K.C. Hui, Carolyn Huston, A. H. Welsh

Modeling directional time series data such as wind or ocean current direction presents several interesting challenges. Standard linear time series techniques do not account for the circularity of the observations, while existing circular modeling approaches typically work best when the data span a small arc.

Motivated by a series of fire burn experiments collecting wind direction in southeastern Australia, we propose a new method for directional time series data that combines the flexibility of hidden Markov models for capturing different latent states at different periods of time, with a conditional von Mises distribution given the latent state to explicitly account for the circular nature of the responses. The resulting circular hidden Markov model (cHMM) can allow for multimodality and/or varying amounts of circular dispersion over time.

Furthermore, by utilizing a von Mises distribution whose mean direction depends on previous observations, we can accommodate serial correlations within a specific hidden state. We employ direct maximum likelihood estimation to fit the cHMM, and examine three approaches to perform forecasting

based on extrapolating the latent state sequence and then direction observations conditional on this sequence. An application to the motivating wind direction datasets reveals that cHMMs produce similar or better point/probabilistic forecasting performance compared with several established time series methods.

Tales from the jungle: a personal perspective of statistical consulting since COVID

Alice Richardson (Australian National University)

The period since COVID-19 has brought unique challenges to statistical consulting, with researchers venturing into unfamiliar statistical territory. The activities of an academic statistical consulting practice have therefore very much taken on the feel of a journey through a statistical jungle.

In this talk I will describe some of the more alarming encounters I have experienced in this jungle over the last five years. Examples will be selected from my client interactions within the broad biometric research field. Names and topics will be altered to protect the participants, but not so much as to alter the key learnings from the encounters.

These examples will cover the entire range of statistical endeavour. From power and sample size calculation through data collection to modelling, presentation of results and visualisation, no area of statistical activity is untouched by surprising methodological suggestions from researchers. The simultaneous emergence of large language models has also brought unexpected consequences in consulting practice, as researchers increasingly turn to AI tools for statistical guidance. I'll draw out the patterns of common traps and highlight how the biometric community can assist and advise.

I will also discuss some of the practical strategies I have developed to carve out a safe path. These strategies can be used for navigating similar consulting challenges and for fostering better statistical practices in other collaborative research environments beyond academia.

Speed: An R package for Spatially Efficient Experimental Designs

Sam Rogers (University of Adelaide)

Co-authors: Wasin Pipattungsakul, Julian Taylor

Agricultural field trials are typically designed using robust statistical randomisation, and best practice agricultural field trials also consider the spatial layouts of the experiment, and how that interacts with the treatments of interest. However, software providing access to spatially optimised experimental designs for field trials is not readily available and can be difficult for new users to get started with due to sparse documentation.

In this talk, we discuss the newly developed *speed* R package which provides easy access to a fully open-source package with comprehensive documentation to enable the design of spatially optimal experiments. The package offers model-free spatially optimal designs via a simulated annealing optimisation algorithm. It can produce a spatially optimal version of many types of experimental designs commonly used in agricultural research as well as many more complex designs such as incomplete block designs and partially replicated designs. It provides multiple objective functions out of the box, with the additional flexibility to choose or enable custom optimisation metrics, depending on the objective of the researcher. It also provides some helper functions for plotting and evaluating experimental designs either produced via *speed* or alternative design packages.

To demonstrate the package's capabilities, we present spatially optimal designs for challenging scenarios including two-dimensional blocking and partially replicated designs. We also show that speed is significantly faster compared to alternative software.

Simulation-based Inference about Variance Components

Farwa Saleem (School of Mathematics and Statistics, UNSW Sydney)

Co-authors: David Warton

Variance components in generalized linear mixed models (GLMMs) are used to quantify random effects, for example, they can be used in ecological studies to quantify the size of the effect of an environmental variable on species composition. However, constructing accurate confidence intervals for these components remains challenging with small samples or near-boundary estimates, where asymptotic methods (e.g., profile likelihood, Wald intervals) do not perform well. To address this, we propose a novel simulation-based framework using bootstrap test inversion coupled with quantile regression. Our approach efficiently estimates confidence limits by leveraging the continuous distribution of the estimator, eliminating discretization steps inherent in conventional bootstrap procedures. This enhances robustness in low-information scenarios while adapting to boundary constraints. Simulation studies demonstrate that our method achieves superior coverage accuracy compared to existing asymptotic and bootstrap alternatives under small-sample conditions and near-boundary parameters. We show the application of the technique to an ecological case study analysing response of species composition to environmental gradients. This work delivers a flexible tool for more defensible inference in applications demanding precise uncertainty quantification for random effects.

Keywords: Variance components, GLMM, bootstrap, quantile regression, confidence intervals, ecological statistics.

The 4S method for the longitudinal analysis of multidimensional questionnaires: application to Parkinson's disease progression from patient perception

Tiphaine Saulnier (New Zealand Brain Research Institute, Christchurch, New Zealand and Institute of Neurodegenerative Diseases (IMN), University of Bordeaux, Bordeaux, France)

Co-authors: C√©cile Proust-Lima, Daniel J. Myall, Toni L. Pitcher, Leslie Livingston, Rachel Nolan, Sophie Grenfell, John C. Dalrymple-Alford, Alexandra Foubert-Samier, Wassilios G. Meissner, Tim J. Anderson and Campbell J. Le Heron

Introduction In health research, questionnaires are widely used to assess clinical states or quality of life (QoL). Longitudinal analysis of these tools offers valuable insights into disease progression. However, this raises multiple challenges, such as handling repeated ordinal responses, capturing the multidimensional traits underlying all the items, and accounting for informative dropout due to death during follow-up. In this work, we present the 4S method – a comprehensive strategy developed to address these challenges, and illustrate it describing health-related QoL (Hr-QoL) changes in Parkinson's disease (PD).

Methods The 4S method comprises four successive steps: (*1 – structuring*) identify questionnaire dimensions through factor analyses; (*2 – sequencing*) describe each dimension progression and associated predictors with a joint latent process model combining an item-response mixed model and a proportional hazard for death risk; (*3 – staging*) compare progression across dimension continuums by projecting clinical stages; (*4 – selecting*) highlight stage-specific most informative items based on Fisher information.

Application We analyzed longitudinal data from the New Zealand Parkinson Progression Programme (NZP[3]), following over 400 PD patients for up to 16 years. Hr-QoL was measured via the PDQ-39 questionnaire, covering motor and non-motor spheres. Five dimensions were identified: mobility, daily activities, psycho-social, stigma, and cognition/communication/bodily discomfort. All, except stigma, showed progressive decline, with patterns notably varying by sex and onset age. Items

related to walking, dexterity, anxiety, and communication appeared as particularly sensitive during PD stages, offering guidance to enhance patient-centered care.

Conclusion The 4S method is a comprehensive statistical strategy suited to analyze repeatedly-collected questionnaires in health studies.

Enhancing Fraud Detection in Banking through Random Survival Forests: Addressing Data Imbalance and Model Transparency

Arjun Sekhar (Macquarie University)

Co-authors: Samuel Muller and Maurizio Manuguerra

With the rise in financial fraud amid growing transaction volumes, in this presentation we share our insights from the use of Random Survival Forests (RSFs) for enhancing fraud detection in banking systems. While many machine learning techniques prioritise accuracy, they overlook two crucial issues: the imbalance of fraud versus legitimate transaction data, and the transparency of the model decision-making process.

We position RSFs as a potential solution, leveraging their time-to-event structure to enable dynamic fraud prediction and their ensemble nature to handle high-dimensional, skewed datasets. Using real-world transactional data, the presentation addresses model imbalance through resampling strategies and evaluates the model's ability to detect minority-class fraud events without inflating false positives.

Furthermore, the interpretability of RSFs is explored through explainable AI frameworks, offering transparency essential for regulatory compliance and institutional trust. By investigating RSFs to deliver fraud detection that is both accurate and accountable, this presentation bridges algorithmic robustness with transparency, advancing a practical and governance-aligned solution for responsible AI deployment in today's increasingly regulated financial services landscape.

Variable Selection in a Joint Model for Huntington's Disease Data

Rajan Shankar (The University of Sydney)

Co-authors: Tanya Garcia and Garth Tarr

Huntington's disease is a neurodegenerative disease caused by a defective Huntingtin gene, with symptoms that progressively worsen and eventually lead to a clinical diagnosis. Identifying the clinical and demographic factors that influence symptom severity and time-to-diagnosis is critical for understanding disease progression so that early-intervention strategies can be timely implemented. We propose a joint model to relate symptom severity y and time-to-diagnosis x, conditional on clinical and demographic predictor variables **z**. However, it may be that certain predictor variables are important for y but not for x and vice-versa, so we use regularisation techniques to select different sets of predictor variables for y and x. Since x is a time-to-event variable, there is the added challenge that many of its values are right-censored due to individuals who did not develop the disease during the study. Therefore, to fit the joint model, we apply the expectation-maximisation (EM) algorithm to alternate between parameter estimation and imputation of the right-censored values until convergence. We demonstrate our method on Huntington's disease patient data, showcasing how users can choose appropriate values for the regularisation tuning parameters.

An allometric differential equation model quantifies energy trade-offs between growth and reproduction under temperature variation

Hideyasu Shimadzu (Kitasato University)

Individual variation in growth and reproduction is common even within a single species. Such differences play a crucial role in shaping life-history strategies, particularly in response to environmental factors such as temperature. This talk introduces a novel system of allometric differential equations that models internal energy allocation between two fundamental biological processes, growth and reproduction, under different thermal regimes.

Climate change is not only driving increases in mean temperatures but also intensifying temperature variability, resulting in greater environmental unpredictability. While temperature is known to regulate resource allocation in ectotherms, the consequences of stochastic thermal fluctuations for life-history traits remain poorly understood. Using *Daphnia magna* as a model organism, we present a novel allometric growth model that characterises energy allocation dynamics in fluctuating thermal environments. Our framework quantifies the effects of randomly varying temperatures on lifetime patterns of growth and reproduction. Our model reveals that exposure to unpredictable thermal regimes can elicit life-history responses similar to those observed under persistently elevated temperatures. Importantly, our energy-based approach identifies patterns of reproductive investment that are not discernible from growth data alone, highlighting the pivotal role of temperature variability in shaping life-history trajectories. As climate change increasingly entails unpredictable environmental conditions, the model and findings presented here offer valuable tools for anticipating and managing biological responses to future climate scenarios.

Zero-inflated Tweedie distribution for abundance of rare ecological species

Nokuthaba Sibanda (Victoria University of Wellington)

Abundance data for rare species can be extremely zero-inflated, where percentage of zeros can be over 90%. This poses a challenge even for the standard Tweedie model which naturally allows for a probability mass at zero with continuous non-negative values. We investigate use of a zero-inflated Tweedie distribution when modelling non-negative continuous abundance values with zeros for rare species. Despite their significance, research on zero-inflated models has predominantly focused on count models such as zero-inflated Poisson or negative binomial regressions, with only recent studies exploring zero-inflated Tweedie models in insurance claims (Zhou, Qian, and Yang 2022; Gu2024; So and Valdez 2024; So and Deng 2025).

The zero-inflated Tweedie model uses a mixture model approach that integrates a Tweedie model with a binary model to distinguish between excess zeros - those resulting from an independent process, and true zeros those resulting from the Tweedie model itself. We use a Bayesian approach to estimate the model parameters. We model means of the Tweedie model using a log-link function with covariates and unobserved random effects. The spatial association between observations is accounted for using a conditionally auto-regressive prior.

A missing data detective story – how I navigated through a perfect storm of drop-out's, COVID and informative missingness.

Eve Slavich (University of New South Wales)

Co-authors: Rebecca Young, Anita Pfeiffer, Emmanuel Karantanis

Real-world clinical studies often present missing data challenges that leave the textbook educated scratching their heads. Through consulting, I become involved in a study where multiple correlated

missing data mechanisms created analytical challenges. Following implementation of new birthing protocols aimed at reducing birth-related injuries, in a Sydney health district, a follow-up study was designed to evaluate incontinence outcomes. However, the missing data pattern proved to be informatively related to both treatment and outcome variables: women who experienced successful outcomes (no incontinence) were systematically less likely to attend follow-up appointments. The situation was further complicated when the COVID-19 pandemic struck mid-study, dramatically reducing attendance rates and introducing an additional layer of correlated missingness that varied by treatment timing and baseline characteristics. Working within the constraints of the study design and the budget for a statistician, I developed a model for missingness which I will outline in this talk.

While missing data complications can severely challenge study validity, thoughtful statistical approaches can still yield meaningful insights when the caveats (caveats galore!) are properly acknowledged and addressed.

References: Young, R et al, 2025, Outcomes for obstetric anal sphincter injuries and anal incontinence following introduction of a perineal bundle, Continence, Volume 14

Keywords: Missing data analysis, informative missingness, multiple imputation, statistical consulting, statistical practice

Rate-optimal sparse gamma scale mixture detection

Michael Stewart (The University of Sydney)

Co-authors: Qikun Chen

We consider a model where observations from a known gamma distribution are possibly contaminated by observations from another gamma distribution with the same shape but a different mean. Such a model has been considered for times between neurotransmitter releases based on a Markov chain with amalgamated indistinguishable states. We focus on the case where the contaminating component occurs rarely, the so-called sparse gamma scale mixture detection problem. Due to the irregularity of such models theoretical results concerning detectability bounds are non-standard. Nonetheless in recent years a body of theory has been developed which covers the case when the mean of the unknown contaminating component is smaller than the null mean, but not when it is larger. We present some recent results filling this gap in the literature. In particular we describe a test which attains the optimal rate of convergence in various local alternative scenarios which is a Bonferroni-type test combining three different tests.

False Discovery Rate Controlled Robust Variable Selection under Cellwise Contamination

Xiaoya Sun (School of Mathematical and Physical Sciences, Macquarie University)

Co-authors: Houying Zhu, Garth Tarr and Samuel Muller

The increasing complexity and dimensionality of modern datasets make cellwise outliers a frequent and expected phenomenon, complicating tasks such as variable selection. An mRNA-Seq expression profiling data set of human brain tissue for Huntington's Disease (HD) illustrates this practically, where we work to identify key genes involved in HD by comparing affected individuals with neurologically normal controls. Our objective is to obtain reliable results while minimizing the impact of cellwise outliers. While many robust variable selection methods have been developed, few effectively handle high-rate cellwise contamination, and it remains an open problem to what extent such methods can control error rates. We introduce GRALF, the Gaussian-Ranked Adaptive Lasso with False discovery rate (FDR) control, a robust method designed for high-dimensional variable selection under cellwise contamination, taking advantage of FDR control to increase statistical reliability. GRALF builds on

the framework of GR-ALasso, an effective robust variable selection method using the Gaussian-Rank estimator in the Adaptive Lasso, and integrates FDR control by generating fake counterparts of the Gaussian-rank transformed variables and estimating the number of false discoveries in the optimization process. Simulation studies demonstrate GRALF's desirable performance in terms of empirical power and FDR control under various conditions. The analysis of the HD data set identified potential gene markers associated with the development of HD, which overlap with findings from previous research, further validating the effectiveness of GRALF.

A Test for Detecting Multiple Clusters with Hotspot Spatial Properties

Kunihiko Takahashi (Institute of Science Tokyo)

Co-authors: Hideyasu Shimadzu

Various statistical tests have been widely used in spatial epidemiology to investigate regional patterns in disease occurrence, particularly to assess whether disease risk is significantly elevated in specific areas compared to neighboring regions or adjacent time periods. One such method is the cluster detection test (CDT), which identifies non-random spatial distributions of diseases and highlights high-prevalence regions without prior assumptions. Among CDT methods, scan statistics are compelling and use a maximum likelihood framework to search across spatial and/or temporal windows for potential clusters. Examples include Kulldorff's circular scan statistic and the flexibly shaped scan statistic by Tango and Takahashi. More recently, Takahashi and Shimadzu proposed a scan-based method that simultaneously detects multiple clusters by integrating generalized linear models and an information criterion to determine the optimal number of clusters. Traditional scan-based tests often assume that disease risk is uniformly elevated within a single cluster. However, they may mistakenly combine multiple adjacent hotspots—each with potentially different risk levels—into one, thereby masking meaningful spatial heterogeneity. In this study, we propose a new test procedure that more accurately identifies adjacent hotspot clusters as distinct entities. Our approach enhances the scan statistic framework by incorporating Cochran's Q-statistic to assess heterogeneity within clusters. We demonstrate the effectiveness of the proposed method through real-world applications and compare its performance with conventional scan-based tests.

Integrating Spatial Data and On-Farm Experimentation to Understand Wheat Variety Performance Across Western Australia

Sandra K. Tanz (Centre for Crop and Disease Management, Curtin University)

Understanding spatial variability in crop performance is critical to improving the statistical design and interpretation of variety trials at the farm scale. This pilot study investigates the integration of large-scale on-farm experimentation (OFE) with spatial data analysis to better understand genotype-by-environment interactions in wheat. The trial spans five grower-managed sites across Western Australia, incorporating two core wheat varieties—Scepter (a widely grown commercial line) and IGW6993 (a near-release InterGrain line)—with some growers also including Rockstar and Tomahawk CL Plus based on local relevance.

Each trial is embedded within commercial paddocks and implemented using strip trial designs aligned with paddock production zones identified from historical yield maps. This design allows statistical comparisons of varietal performance across contrasting spatial zones defined by environmental variation in soil, topography, and management history. The study aims to (1) develop robust methods for spatially aware experimental design in commercial cropping systems, and (2) assess the potential for OFE to complement small-plot trials in supporting variety selection and agronomic decision-making.

While the trial was only recently seeded (May 2025), preliminary outputs include trial design validation

and early NDVI imagery from emergence. Future analyses will incorporate additional covariates such as soil electrical conductivity (EM mapping), NDVI time series, and topographic indices to model spatial responses. This work highlights the value of integrating spatial data into trial workflows and explores the role of predictive breeding tools in commercial agriculture.

Outlier-robust estimation of state-space models using a penalised approach

Garth Tarr (The University of Sydney)

Co-authors: Rajan Shankar, Ines Wilms and Jakob Raymaekers

State-space models are a broad class of statistical models for time-varying data. The Gaussian distributional assumption on the disturbances in the model leads to poor parameter estimates in the presence of additive outliers. Whilst there are ways to mitigate the influence of outliers via traditional robust estimation methods such as M-estimation, this issue is approached from a more modern perspective that utilises penalisation. A shift parameter is introduced at each timepoint, with the goal being that outliers receive a non-zero shift parameter while clean timepoints receive a zero shift parameter after estimation. The vector of shift parameters is penalised to ensure that not all shift parameters are trivially non-zero. Apart from making it feasible to fit accurate and reliable time series models in the presence of additive outliers, other benefits of this approach include automatic outlier flagging and visual diagnostics to provide researchers and practitioners with better insights into the outlier structure of their data. We will demonstrate the utility of this method on animal tracking data.

A Proportional Random Effect Block Bootstrap for Highly Unbalanced Clustered Data

Zhi Yang Tho (Australian National University)

Co-authors: Alan H. Welsh and Raymond Chambers

Clustered data arise naturally in a wide range of scientific and applied research settings where units are grouped within cluster, and are commonly analyzed using linear mixed model to account for within-cluster correlations. This article focuses on the scenario in which cluster sizes might be highly unbalanced, and proposes a proportional random effect block bootstrap that is applicable in such case and is robust to misspecification of the stochastic assumptions of the linear mixed model. The proposed method is a generalization of the random effect block bootstrap, originally designed for the balanced case, and can be used to perform inferences on parameters of the linear mixed model or functions thereof. We establish asymptotic consistency of the proposed bootstrap under general cluster sizes scenario, showing that the original random effect block bootstrap is only consistent when cluster sizes are balanced. A modified random effect block bootstrap is also proposed which enjoys similar asymptotic consistency properties as the proportional random effect block bootstrap. Simulation study demonstrates the strong finite sample inferential performance of the proposed bootstraps, particularly compared with the random effect block bootstrap and several existing bootstrap methods for clustered data. We apply the proposed bootstraps to the Oman rainfall enhancement trial dataset with cluster sizes ranging from 1 to 58. Results show that the bootstrap confidence intervals based on our proposed bootstraps are more adequate than those of random effect block bootstrap and that the employed ionization technology has a statistically significant effect in increasing the amount of rainfall.

Functional Data Analysis for the Australian Grains Industry

Braden J. Thorne (Curtin University)

Co-authors: Adam H. Sparks

When performing statistical analysis on time series data with regular sampling, it is common to apply filtering or windowing methods to reduce the complexity of the task. While this process enables many classical approaches, it inevitably leads to compression of the full information available. An alternative approach is to treat observations as samples of a continuous mathematical function and focus analysis on the curves these functions produce rather than the samples. This is the underlying idea of functional data analysis, a statistical analysis approach that has seen growing attention in recent years. In this talk I will offer an introduction to functional data analysis and detail our exploration of these methods for application to the grains industry across Australia. Specifically, I will present two case studies; estimating charcoal rot prevalence in sixteen years of data from in-paddock soybean experimental trials using weather data, and analysing frost risk in broadacre crops with varying stubble management practices using sensor data.

Extending diagnostic validity meta-analysis to several diagnostic guidelines

Alain C. Vandal (The University of Auckland/Waipapa Taumata Rau)

Co-authors: Henry C. K. Kwok, Fransiska R. M. Falconer, Andrew G. Hill, Andrew D. MacCormick

Optimal management of acute biliary disease should include an assessment for possible choledocholithiasis (stones in the bile duct). Various diagnostic guidelines have been developed by expert bodies for this purpose, but uncertainties remain about their performance in wider practice. We carried out a meta-analysis to compare three of these guidelines. Each guideline yields a score, for which different thresholds are considered in terms of sensitivity and specificity. Of 1,892 records, thirty-one studies were identified for this meta-analysis. All studies focused on one or more of three international guidelines, namely the ASGE guidelines from 2010, the revised ASGE guidelines from 2019 and the ESGE guidelines from 2019. Each study reported on the diagnostic validity of one or more of these guidelines at one or more score cutpoint. The analysis was carried out using the *diagmeta* R package [Rücker G, Steinhauser S, Kolampally S, Schwarzer G (2022)]. Their approach allows the meta-analytical estimation of the area under the receiver operating curve, a quantity that summarises the performance of a diagnostic tool. The approach consists in fitting transformed values of sensitivity and specificity at various cutpoints, which can differ between studies, accounting for study heterogeneity using linear mixed effects modelling. While the *diagmeta* approach applies to the meta-analysis of a single diagnostic procedure, we show how its framework can be extended to jointly and severally test for performance similarity between guidelines. We also show how it can be extended to identify outlier studies, for the purpose of carrying out sensitivity analyses.

Integrated Species Distribution Models: A Single-Index Approach

Quan Vu (Australian National University)

Co-authors: Francis Hui, Alan Welsh, Samuel Muller, Eva Cantoni and Chris Haak

In ecology and fisheries, species abundance data are often collected using a number of surveys that have different characteristics. Spatial statistical models can be used to integrate information from these datasets to improve both interpretability and prediction of the abundance. In this talk, we introduce the single-index integrated species distribution model, based on a single index representing the latent spatial distribution of the species abundance, and survey-specific link functions representing the different catchability properties of each survey. We demonstrate the use of the model through an analysis of scallop data collected from two surveys: a bottom trawl survey which covers a wide spatial domain but is less efficient, and a hydraulic dredge survey which is more efficient and spatially targeted. Results show that our model offers meaningful interpretations of covariate effects, spatial distribution, and survey catchability differences, and achieves superior predictive performance compared to contemporary species distribution models.

Extension of the corrected score estimator in a Poisson regression model with a measurement error

Kentarou Wada (Tokyo University of Science)

Co-authors: T. Kurosawa

Kukush et al. (2004) discussed the bias of the naive estimator for the regression parameters in a Poisson regression model with a measurement error for the case where the explanatory variable and measurement error follow normal distributions. Wada and Kurosawa (2023) proposed the corrected naive (CN) estimator as a consistent estimator for a Poisson regression model with a measurement error for the case where the explanatory variable and measurement error are general distributions. The CN estimator directly calibrates the bias of the naive estimator. The CN estimator is given by the solution of the estimation equation of the Poisson regression model under the error-in-variables framework. However, the CN estimator does not always have an explicit expression under the condition that the explanatory variable and measurement error follow general distributions. On the other hand, Kukush et al. (2004) considered the corrected score (CS) estimator as a consistent estimator for the true parameter of the Poisson regression model with a measurement error. In this research, we extend the CS estimator to the case where the explanatory variable and measurement error are general distributions. The new estimator can be applied for the condition that the CN estimator does not have an explicit expression. As illustrative examples, we give simulation studies to verify the effectiveness of the new estimator.

Simultaneous Inference for Latent Variable Predictions in Factor Analytic Models

Zhining Wang (Research School of Finance, Actuarial Studies and Statistics, Australian National University)

Co-authors: Emi Tanaka, Francis K.C. Hui

Factor analytic models (also known as latent variable models) are fundamental tools in multivariate statistics, widely applied in fields such as psychology, economics, and social sciences. While considerable attention has been given to estimation and inference on parameters such as the loading matrix and the error variances, relatively less research has been done on how to perform inference on the predicted factors, e.g., how to construct prediction intervals for the latent variables jointly across all the clusters in a given dataset. We explore a framework for the simultaneous inference of the predicted latent variables in factor analytic models. We demonstrate the construction of simultaneous prediction intervals for the predicted factors, examining strategies such as the bootstrap and Monte Carlo simulation, and show how this also facilitates joint/multiple testing across different cluster-level predictions. We examine the practical feasibility and robustness of the proposed simultaneous inference methods with simulation studies and an application in the biosciences.

Estimating extinction time from the fossil record using regression inversion

David I. Warton (UNSW Sydney)

Co-authors: Victor Tsang

An important problem in palaeoecology is estimating the extinction or invasion time of a species from the fossil record - whether because this is of interest in and of itself, or in order to understand the causes of extinctions and invasions, for which we need to know when they actually happened. There are two main sources of error to contend with - sampling error (because the last time you see a species need not be the last time it was there) and measurement error (dating specimens, usually well known). The paleobiology literature typically ignores one or other of these sources of error, leading

to bias and underestimation of uncertainty to an extent that is often qualitatively important.

The problem is surprisingly difficult to address statistically, because while standard regularity conditions are technically satisfied, we are typically close to a boundary where they break down, and hence standard asymptotic approaches to inference typically perform poorly in practice. We propose using a novel method, which we call regression inversion, for exact inference, and we apply this technique to a compound uniform-truncated t (CUTT) model for fossil data. We show via simulation that this approach leads to unbiased estimators, and accurate interval inference, in contrast to its competitors. We show how to check the CUTT assumption visually, and provide software to apply all of the above in the reginv package.

The equalto covariance structure for meta-analysis using the `glmmTMB` **R package**

Coralie C. Williams (University of New South Wales)

Co-authors: Maeve McGillycuddy, David I. Warton, Shinichi Nakagawa

Meta-analysis is a widely used statistical method for synthesising quantitative results from related studies, enabling researchers to address broad questions and explore sources of heterogeneity. In R statistical software, several packages support meta-analytic modelling, including the actively maintained `metafor` package which is commonly used across all disciplines. But meta-analysis typically uses a mixed model, with the variance structure of study-level random effects known, so it should be possible to fit such models using standard mixed modelling software. We propose using the `glmmTMB` package for this purpose. In the last decade, `glmmTMB` has developed into a flexible package to fit generalised linear mixed models (GLMMs) via the Template Model Builder (TMB) framework. The package holds a similar interface as the `lme4` package but supports a broader range of distributions and covariance structures using Laplace approximation. Here, we introduce a new covariance structure, called "*equalto*", to the `glmmTMB` package, so that it can be used to fit meta-analytic models by allowing users to specify a known variance-covariance matrix for the sampling error. This enables explicit modelling of heteroscedasticity or dependence among effect sizes. The new implementation offers an alternative way to do a meta-analysis, convenient for users already familiar with fitting mixed models in `lme4` or `glmmTMB` packages, and capable of handling complex model structures. This novel implementation supports more flexible modelling of meta-analytical data, expanding the R toolkit available for evidence synthesis. We showcase its applicability with illustrative examples in ecology and evolution.

Disease cluster detection via functional additive models incorporating spatial correlation

Michio Yamamoto (The University of Osaka / RIKEN AIP / Shiga University)

Co-authors: Tatsuhiko Anzai, Kunihiko Takahashi

Detecting spatial clusters of diseases is crucial for understanding disease patterns and developing effective prevention and treatment strategies. Spatial scan statistics are powerful tools for detecting spatial clusters with a variable scanning window size. If covariates are related to an outcome and not geographically randomly distributed, searching for spatial clusters may require adjusting for the covariates. In addition, spatial correlation in the outcome, which is often overlooked during cluster detection, can affect the results. In this study, we propose a new spatial scan statistic that handles multiple functional covariates indicating past information over time and the spatial correlation of the outcome. Our method flexibly models these factors in the framework of functional additive models. We develop an optimization algorithm to estimate the model parameters for the normal outcome case. A simulation study and real data analysis indicate that the proposed method can detect disease clusters despite longitudinal covariates and spatial correlations compared to existing methods.

Childhood Risk and Resilience Factors for Pasifika Youth Respiratory Health: Accounting for Attrition and Missingness

Siwei Zhai (University of Auckland)

In New Zealand, 7% of deaths are related to respiratory diseases, with Pacific people at higher risk. Our work investigates the causal effects of early-life risks and resilience factors on early-adulthood lung function amongst Pacific Islands Families Study (PIFS) cohort members (n=1,398). 466 from the cohort participated in the respiratory study. Primary outcome was forced expiratory volume in 1 second (FEV1) z-score at age 18 years. FEV1 and healthy lung function (HLF), defined as the z-score being larger than -1.64, were secondary outcomes. A previous study had evaluated the effects of early-life nutrition factors on the respiratory health of Pacific youth. The results suggested a positive impact of consuming more fruit and vegetables during childhood on respiratory health later in life. The follow-up study will continue to explore the effects of factors from relevant domains based on the PIFS cohort, where a new integrated model will be applied. A simulation will be conducted to determine this model.

Elastic Net Regularization for Vector Generalized Linear Models: A Flexible Framework for High-Dimensional Biomedical Data

Wenqi Zhao (University of Auckland, Auckland, New Zealand)

Co-authors: Dr.Thomas Yee

We introduce a novel implementation of elastic net regularization for vector generalized linear models (VGLMs), capable of fitting over 100 family functions and designed to support complex, high-dimensional modeling tasks commonly encountered in the biosciences. VGLMs extend classical GLMs by accommodating multivariate and multi-parameter responses, making them particularly well-suited for heterogeneous biomedical data.

Our method integrates sparse estimation techniques—such as lasso, ridge, and their convex combinations—into this broader modeling framework, enhancing model interpretability and stability in high-dimensional settings. The algorithm is implemented in the vglmnet function within the new VGAMplus package for R. It leverages a modified iteratively reweighted least squares (IRLS) procedure, combined with pathwise coordinate descent, Karush-Kuhn-Tucker (KKT) condition checks, and strong rules for variable screening to ensure computational efficiency.

This framework supports a wide range of models beyond the exponential family, including ordinal, categorical, and zero-inflated distributions commonly encountered in fields such as epidemiology, genomics, and clinical research. We illustrate the utility of our approach through comparisons with existing tools (glmnet, ordinalNet, mpath) and apply it to real-world datasets involving survival outcomes, count data, and bivariate binary responses. By uniting the structural flexibility of VGLMs with the benefits of regularization, our method provides a powerful and scalable solution for modern statistical modeling in the biosciences.

Extending Spatial Capture-Recapture with the Hawkes Process

Alec B. M. van Helsdingen (University of Auckland)

Co-authors: Charlotte M. Jones-Todd

Spatial capture-recapture (SCR) is a well-established method used to estimate animal population size from animal sighting or trapping data. Standard SCR methods assume animal movements are independent and consequently cannot incorporate site fidelity (attachment to a particular region)

nor the temporal correlation of an animal's location. Recent work has sought to solve these issues by explicitly modelling animal movement.

In this talk we propose an alternative solution for camera trapping surveys based on a multivariate self-exciting Hawkes process. Here the rates of detection of a given animal at a given camera are a function of not only the location and its proximity to the animal's activity center, but also where and when the animal was most recently detected.

Through a mixture of Gaussian distributions, our model expects more detections closer in space to the last detection, and reduces to SCR when an animal is yet to be detected. This formulation, we believe, better reflects animal behaviour because shortly after detection, we expect to see an individual close to where it was last seen. Thus, our model allows us to account for both site fidelity and the inherent temporal correlation in detections that have not previously been accounted for in SCR-type models.

In this talk, I will 1) give an overview of Self-Exciting Spatial Capture-Recapture (SESCR) models, and 2) demonstrate the additional inference that can be drawn from such models and 3) apply the framework using a few case studies to compare traditional SCR and SESCR.

Tuesday

Wednesday

Thursday

Friday

Posters

The "Theory of Sampling" (ToS) - have statisticians missed the boat?

Damian Collins (NSW DPIRD)

Co-authors: Anne Harris

The "Theory of Sampling" (ToS) framework originated with chemical engineer Pierre Gy over 50 years ago. It essentially deals with the sampling of materials, particularly soils, composts,and minerals. It enjoys a strong advocacy amongst chemical engineers, geologists and similar professionals. The ToS instructs practitioners to stop and think about the sampling procedure. It encourages them to consider all the sources of variation in the material they wish to sample rather than just choosing to take the most convenient sample such as a "grab sample" from the surface. However, one of its strongest advocates argues that, in order to save on analysis costs, only one "correct" composite sample is necessary. This single sample, deemed to be "representative" of the entire material, is all that is required to characterise a material (Esbenson (2017)). Composting sampling guidelines around the world use the ToS to justify this practice of single sampling. For example, it has currently been recommended in the AS4454 review in Australia (https://www.as4454review.com.au/). However, with only one sample analysed, contributions to measurement uncertainty from either the sampling or analysis processes, cannot be assessed. It is concerning that despite the apparent sophistication of ToS methodology and the breadth of its advocacy network, it perpetuates the scientifically unsound idea of a "representative sample." .Even the name, "Theory of Sampling", is a misnomer, suggesting an even broader application than just the sampling of materials. We review the ToS and consider implications to statistical practise and adoption.

Esbenson, K.H. (2017). Sampling: Theory and Practice. Alchemist 85 pages 3–6. https://www.lbma.org.uk/al-chemist/issue-85/sampling-theory-and-practice

Evaluating Remote Monitoring in Automated Peritoneal Dialysis: A Difference-in-Differences Analysis

Annie Conway (ANZORRG , ANZDATA , Australia and New Zealand Dialysis and Transplant Registry)

Co-authors: Stephen McDonald

During a cholera outbreak in 1850s London, physician John Snow famously traced the source of the outbreak to contaminated water supply. In what could be deemed a "natural experiment", he compared cholera mortality rates between areas served by different water companies before and after one company relocated its water intake, using an early form of the method *difference-in-differences* (DiD).

Recently, DiD has gained renewed popularity in health and epidemiological sciences for testing the effectiveness of technological or policy interventions. However, real-world applications can have complexities such as staggered roll-out of the intervention and heterogeneity between treated units. Additionally, these often involve count outcomes that violate the linear model assumptions underlying existing DiD methods.

In this project I consider the impact of a new technology for automated peritoneal dialysis (APD)

that allows patients to be remotely monitored by clinicians as they complete dialysis at home. This new APD cycler was rolled out gradually in PD centres across Australia and New Zealand between 2017 and 2023. Using data from the ANZDATA registry, I assess remote monitoring's impact across PD centres, counting the number of deaths, peritonitis infections and technique failures, comparing before and after the roll-out and between treated and untreated centres.

This analysis employs various DiD methods to overcome the challenges of staggered adoption and heterogeneous effects, to provide an evaluation of remote monitoring's impact on critical patient outcomes.

Using atmospheric transport models to predict species incursions in northern Australia

Zhenhua (Iris) Hao Dr (Department of Agriculture, Fisheries and Forestry)

Predicting the arrival of invasive species is critical for strengthening biosecurity, particularly in regions like northern Australia that are vulnerable to aerial incursions. This ongoing study compares atmospheric modelling with detection patterns from long-term field surveillance to evaluate how different models can be integrated to increase confidence in identifying likely source regions, arrival pathways and periods of elevated risk. We apply three complementary models: HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory), FLEXPART (FLEXible PARTicle dispersion model), and Lagrangian Coherent Structures (LCS). These models range in complexity - from simulating basic wind trajectories to identifying hidden transport patterns within the atmosphere. A rich empirical foundation is provided by two decades of fruit fly trapping data collected through the Northern Australia Quarantine Strategy (NAQS) in the Torres Strait. By identifying likely incursion windows for Oriental Fruit Fly (Bactrocera dorsalis) arrivals from Papua New Guinea, we evaluate and compare model performance and gain insights into potential dispersal mechanisms. While model implementation and case analysis are still underway, preliminary results suggest that combining outputs from different models improves the detection and interpretation of likely incursion events. Future work will focus on integrating these model outputs with additional environmental and biological datasets to produce dynamic risk maps. These maps aim to support NAQS surveillance planning and resource prioritisation. This framework can also be extended to other high-risk species to enhance predictive capacity and improve biosecurity strategies for Australia's north.

Cleaning Text Data with Large Language Models

Jiajia Li (Australian National University)

Co-authors: Emi Tanaka

Data cleaning – particularly the standardisation and feature extraction from unstructured text – is a time-consuming and often manual process in data science workflows. Traditional methods, such as approximate string matching, lack semantic understanding and still require inefficient, manual corrections. We leverage the transformative potential of large language models (LLMs) to automate these tedious tasks and significantly accelerate data preparation.

We introduce the `emend` R package, which leverages LLMs to perform various data cleaning tasks, including matching categorical variables, reordering ordinal variables, translating languages, and standardising dates and addresses. The package integrates with local LLMs via Ollama as well as cloud-based APIs via the R package ellmer, offering multiple options for using LLMs from different providers. A comprehensive simulation study evaluated emend's core functionality across 13 LLMs and a baseline string-matching method on real-world datasets of country names, institutional affiliations, and species names. Results demonstrate high accuracy for LLM-based methods in general text standardisation, with GPT-4o-mini and GPT-4o excelling on country names and affiliations, respectively.

The `emend` package streamlines data cleaning in R, reducing the need for extensive manual corrections for data analysts. This work highlights LLMs as powerful tools for increasing efficiency of cleaning text data.

Performance of Factor Analytic Mixed Models and Plackett–Luce Models for Genotype Ranking in Multi-Environment Trials

Jiazhe Lin (ANU)

Co-authors: Emi Tanaka, Fonti Kar

Multi-environment trials are central to plant breeding, providing evaluations of genotype performance across locations and time to capture environmental variability. The resulting data are often heterogeneous and unbalanced, motivating the use of advanced statistical models. The Factor Analytic Linear Mixed Model (FALMM) has become a benchmark approach, offering a parsimonious decomposition of genotype and genotype-by-environment effects and supporting inference on average genotype performance. Alternatively, the Plackett-Luce Model (PLM), originally developed in psychology, allows aggregation of environment-specific rankings into a global ranking that reflects overall genotype performance. We propose a two-stage framework that integrates FALMM and PLM, and through extensive simulation, evaluate conditions under which each method provides optimal inference.

ggincerta: An R Package for Uncertainty Visualisation with a Layered Grammar of Graphics

Xueqi Ma (Australian National University)

Co-authors: Emi Tanaka, Francis KC Hui, Weihao Li, Quan Vu

Uncertainty is inherent in all estimates derived from data. However, when visualising these estimates on spatial maps, uncertainty is often overlooked, leading to decision-making based solely on point estimates. The `Vizumap` R package addresses this gap by integrating both estimates and their uncertainties in areal data, offering four types of visualisations: bivariate choropleth maps, pixel maps, glyph maps, and exceedance probability maps.

Building upon `Vizumap`, we introduce the `ggincerta` R package, a ground-up reimplementation that recreates all of `Vizumap`'s visualisations, but also provides enhanced functionality, within the layered grammar of graphics framework provided by `ggplot2`. By leveraging `ggplot2`, `ggincerta` seamlessly integrates uncertainty visualisation into users' existing workflows, employing familiar syntax and conventions. This approach not only enhances accessibility for those already comfortable with `ggplot2`, but also exemplifies how existing systems can be extended to new purposes, rather than developing separate, ad-hoc solutions that require learning additional languages or interfaces.

Small Area Estimation of Fertility Rate in the Philippines using DHS and Census Data

George Manapat (University of the Philippines Diliman)

Accurate and timely estimation of fertility indicators at subnational levels is essential for effective health planning, especially in countries with geographically diverse populations and uneven data availability. In the Philippines, while national and regional fertility rates are routinely measured through national surveys such as the Demographic and Health Survey (DHS), local-level estimates, e.g., at the provincial level, remain sparse and unreliable due to limited sample sizes. Using the data from the 2022 Philippine DHS and auxiliary covariates from the 2020 Census of Population and Housing, small area estimates of fertility rate at the provincial level are obtained using a Bayesian hierarchical model accounting for spatial correlation. Preliminary findings show substantial differences in fertility

rates between provinces. Some areas, especially in more remote or underserved regions, have much higher fertility than the national average. The results underscore the necessity for area-specific health programs and policies to address the unique demographic challenges of each locality.

Bayesian inference for sparse Gaussian copula graphical model

Tomotaka Momozaki (Tokyo University of Science)

We consider Bayesian inference for sparse dependence structures among variables in multivariate mixed data, which consists of both continuous and discrete variables. Traditional graphical modeling approaches often struggle with mixed-type data, as they typically focus on either purely continuous or purely discrete settings. This limitation becomes particularly problematic in high-dimensional scenarios where identifying relevant variable relationships is crucial for interpretability and prediction accuracy. This study proposes a novel Bayesian approach that combines Gaussian copula graphical models with advanced shrinkage priors, including horseshoe priors and double exponential priors. The Gaussian copula framework enables unified modeling of mixed-type variables by transforming them to a common latent Gaussian scale, while shrinkage priors facilitate automatic variable selection by adaptively shrinking irrelevant dependencies toward zero. This combination addresses both the mixed-data challenge and the high-dimensionality issue simultaneously. We develop an efficient posterior sampling algorithm that maintains the positive definiteness of correlation matrices by employing a block Gibbs sampler with hit-and-run algorithm. Our approach utilizes rank information to handle mixed-type data within the copula framework, enabling unified treatment of continuous and discrete variables. The proposed method is particularly valuable for analyzing complex datasets in genomics, economics, and social sciences, where variables of different types naturally co-occur. We present comprehensive evaluation of our method's effectiveness through extensive simulation studies comparing with existing approaches and real data analysis demonstrating practical applicability.

Paired Comparison with Cyclic Dominance: An Extension of the Bradley–Terry Model

Yuki Ohno (Tokyo University of Science)

Co-authors: Kouji Tahata

The Bradley–Terry model has become a cornerstone for analyzing paired-comparison data, providing a principled approach to estimating "strength" parameters without imposing a latent, continuous ordering on teams, treatments, or evaluators. Its transitivity constraint—each pairwise preference must be consistent with a global ranking—yields excellent performance in many applications, including sports analytics, psychometric scaling, and clinical preference studies. However, empirical datasets sometimes exhibit cyclic dominance, rendering the classical model inadequate and leading to systematic misfit.

Real-world data often display cyclic patterns, resembling a "rock-paper-scissors" dynamic, which can violate the transitivity constraint typically expected in such analyses. To address this issue, we first review important applications of the Bradley-Terry model and highlight instances where non-transitive outcomes occur. Next, we introduce a minimal extension of the model that preserves the original likelihood but adds a single "cycle-strength" parameter. This parameter is designed to capture uniform three-way dominance among any trio of outcomes. When the cycle-strength parameter is set to zero, the model reverts to the classical Bradley-Terry framework. Importantly, we demonstrate that our cycle strength corresponds precisely to the quasi-asymmetry measure proposed by Tahata et al. (2004), establishing a clear quantitative relationship between model fit and deviations from transitivity.

Visualisation of multinomial multilevel time-series modelling with application to current smoking status

Alice Richardson (Australian National University)

Co-authors: Sumonkanti Das, Ashis Talukder, Mu Li, and Bernard Baffour

In Australia, tobacco smoking has declined overall over the last 25 years, yet substantial disparities exist across age, sex and state. In this talk we will present the results of multilevel multinomial time series modelling to estimate trends in smoking status (current, former and never) across small domains defined by seven age groups, two genders, and eight states and territories from 2001–2022.

The model is developed using ex-smokers as the reference category, as their proportion remains relatively stable over time. Statistically significant random intercepts and random slopes for linear time trends are identified at the state-age-sex level. The spatio-temporal extension of the multinomial multilevel logistic model yields detailed estimates, helping to uncover disparities in the trends of smoking decline. Temporal random effects at the state-sex and age-sex levels also substantially contribute to achieving numerically consistent trend estimates.

Visualisations of these models are especially important for conveying results in a compelling manner to health researchers and policy makers and program implementation teams. We will discuss the process of arriving at our visualisations using R for this particular application.

Judgement Post-Stratification for Covariate Adjustment in Pairwise Comparisons in Block Designs

Sam Rogers (University of Adelaide)

Co-authors: Omer Ozturk and Olena Y Kravchuk

The randomised complete block design (RCBD) remains one of the most widely used experimental designs in agronomy, agriculture, and controlled-environment research. In this study, we introduce a novel post-experimental method, judgement post-stratification, for improving the efficiency of pairwise treatment comparisons using plot-level covariate information. The method operates within the RCBD framework to focus on generalised designs with multiple replications of treatments within blocks. Using simulations based on uniformity trials, we show that when plots can be consistently ranked according to an expected outcome, independent of treatment assignment, this ranking can be used post-hoc to enhance the precision of estimated treatment differences. This approach offers a practical and flexible tool for incorporating auxiliary information into the analysis of block designs without compromising the integrity of the original randomisation.

Evaluating Diagnostic Performance via Bayesian F_1 Score Estimation without a Gold Standard

Jun Tamura (Yokohama City University)

Co-authors: Yusuke Saigusa, Kouji Yamamoto

In clinical medicine, diseases are often diagnosed with high accuracy based on characteristic signs and symptoms. However, the gold standard tests used to confirm diagnoses are frequently invasive, expensive, or impractical for population-level screening. This highlights the importance of developing and validating simple, non-invasive screening tests. Diagnostic performance is typically evaluated using metrics such as accuracy, sensitivity (Se), specificity (Sp), positive predictive value (PPV), and negative predictive value (NPV). However, when disease prevalence is low or class imbalance exists, accuracy and Sp may not adequately reflect test performance. F_1 score, which considers both Se and PPV, has been widely adopted in such settings as a balanced measure of diagnostic ability. A

limitation of the F_1 score is its reliance on the true disease status defined by a gold standard. In some fields, however, the disease mechanism may be unclear or the definition of clinical signs ambiguous, making it difficult to define a reliable gold standard. This presents challenges for evaluating diagnostic accuracy using conventional methods. To address this issue, we propose a Bayesian framework that enables estimation of the F_1 score without directly observing gold standard outcomes. The approach uses latent class analysis to estimate unobserved disease status and offers a promising method for assessing test performance in settings where the gold standard is unavailable or uncertain

Visualization for departures from symmetry with the power-divergence-type measure in square contingency tables

Wataru Urasaki (Department of Information Sciences, Tokyo University of Science)

Co-authors: Tomoyuki Nakagawa, Jun Tsuchida and Kouji Tahata

When the row and column variables consist of the same category in a two-way contingency table, it is called a square contingency table. Such tables often have an association structure concentrated along the main diagonal, making the analysis of symmetric relationships and transitions important. Various models and measures have been proposed to analyze these structures to understand the changes between two variables' behavior at two-time points or cohorts. This is necessary for a detailed investigation of individual categories and their interrelationships, such as shifts in brand preferences. We propose a novel correspondence analysis (CA) framework to evaluate departures from symmetry in square contingency tables with nominal scales, using a power-divergence-type measure. This approach ensures that well-known divergences can also be visualized and, regardless of the divergence used, the CA plot consists of two principal axes with equal contribution rates. Our visualization method enables the magnitude of each category's departure from symmetry to be assessed by its deviation from the origin, while asymmetric relationships between category pairs can be interpreted through corresponding triangle areas. Importantly, the scaling of the departures from symmetry provided by the measure is independent of sample size, allowing for meaningful comparisons and unification of results across different tables. This standardization supports broader applicability in empirical studies. We also present some considerations from the results of the analysis with actual data. Our framework thus offers an effective tool for studying structural shifts in categorical data.

Code of Conduct

The International Biometrics Society Australasian Region (IBS-AR) is dedicated to providing a welcoming and harassment-free experience for everyone. We do not tolerate unprofessional or inappropriate conduct, which includes discrimination and harassment, sexual or otherwise. Conference participants should always work towards maintaining an environment free of harassment and discrimination. At the discretion of the IBS-AR conference organisers and/or IBS-AR council, any participants violating these rules may be asked to leave the conference, without any fee refund.

Participants agree to:
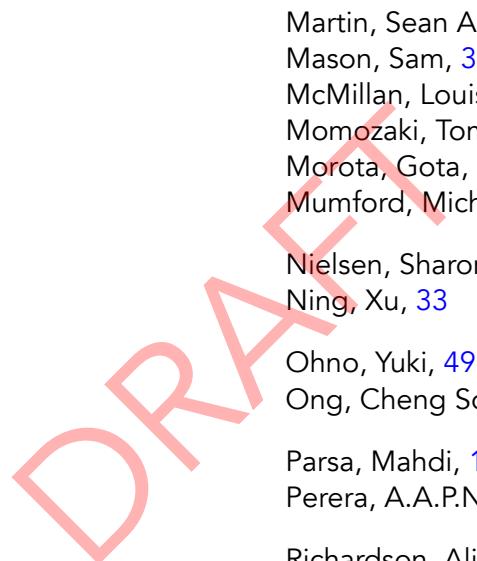
- Be considerate in speech and actions, and actively seek to acknowledge and respect the boundaries of fellow attendees at all times.
- Refrain from demeaning, discriminatory, or harassing behaviour and speech.
- Alert a member of the IBS-AR council or conference organisers if they notice a dangerous situation, someone in distress, or violation of this code of conduct.

Harassment is demeaning or threatening behaviour directed at an individual or a group of people. Harassment includes, but is not limited to,

- Offensive comments related to gender, gender identity and expression, sexual orientation, disability, mental illness, person's lifestyle choices and practices, physical appearance, age, race, religion (or lack thereof) or beliefs;
- Unwelcome sexual attention and behaviour including inappropriate sexual propositions; unnecessary touching; verbal or written comments or physical actions of a sexual nature; sexually degrading words used to describe an individual; a display of sexually suggestive objects or pictures; or sexually explicit jokes; and
- Threatening behaviour including bullying, threat of violence, deliberate intimidation, stalking or incitement of violence. Reporting

If you are being harassed by a member/guest/participant of/at the IBS-AR Conference, or notice that someone else is being harassed, or have any other concerns, please contact the designated conference allies in the conference team. The current conference allies can be found on page 4. If the person who is harassing you is one of the conference allies, they will recuse themselves from handling your incident and will not be given any information without your explicit consent. We will respond as promptly as we can.

Index

DRAFT

Sponsored by

Platinum

Australian National University

Gold

GRDC — GRAINS RESEARCH & DEVELOPMENT CORPORATION

Waipapa Taumata Rau
University of Auckland

Supporter

Statistical Society of Australia

DIAGRAM of THE CAUSES of MORTALITY
IN THE ARMY IN THE EAST.

2.
APRIL 1855 to MARCH 1856.

1.
APRIL 1854 to MARCH 1855.