2018 AUT Mathematical Sciences Symposium

Auckland University of Technology
Auckland, New Zealand

22nd – 23rd November 2018

Published by:
Mathematical Sciences Research Group
School of Engineering, Computer and Mathematical Sciences
Auckland University of Technology
http://www.aut.ac.nz/study-at-aut/study-areas/computer-mathematical-sciences/research-groups/mathematical-sciences-research-group

Contents

Welcome to the 2018 AUT Mathematical Sciences Symposium

On behalf of the Mathematical Sciences Research Group within the School of Engineering, Computer and Mathematical Sciences at Auckland University of Technology, we have much pleasure in welcoming you to the 2018 AUT Mathematical Sciences Symposium.

This is the fifth such Symposium and it is a continuation of our efforts to develop and promote the research being undertaken within the Department of Mathematical Sciences as part of our recently enlarged School. We are delighted to welcome a number of invited speakers to the Symposium with the aim of exploring collaborative opportunities and potential new areas of research that can be established with our research active staff.

The concept of this Symposium was a joint effort of us. We both appreciate the assistance of staff of the Department, in particular Dr Kate Lee, Dr Sarah Marshall, Dr Nuttanan Wichitaksorn and Dr Wenjun Zhang, who have each been involved in a variety of activities to ensure the continued success of this series.

As New Zealand's newest university we have recently had the opportunity to employ a number of new academic staff, all of whom have been developing research profiles. We are currently continuing to appoint additional staff to assist with our expanding Analytics program and develpoments in Engineering. We are putting in place a number of opportunities that will support and assist our academic staff extend and enhance their activities, with this meeting being one such effort.

Our growing postgraduate programme in the Mathematical Sciences at Honours, Masters and Doctoral levels has been enhanced with our Master of Analytics (MAnalytics) degree, now in its fourth year. The success of this programme, with close on forty students at various stages of completion of the degree, is leading to increased project supervision demands on our staff as well as leading to growing links with business and industry. We have an established arrangement with the SAS Institute that sees students in our MAnalytics degree gaining SAS Certification on graduation.

The Mathematical Sciences Research Group focuses on two main areas – Analytics and Applied Mathematics. We are very much focused on "research lead teaching" and we have developed a small number of research clusters within these areas to strengthen and support those academic staff working in these areas. Ideally we would like to foster collaborative activities and we thank those of you who have joined us at this meeting and we hope that we can facilitate some future joint research efforts.

We have kept the focus narrow so as to make the meeting meaningful and rewarding for those who participate. We hope that you enjoy your time with us and that you find the exercise a useful adjunct to the mathematical and statistical scene within New Zealand.

On behalf of the Mathematical Sciences Research Group

Jeffrey Hunter
Professor of Mathematical Sciences

Jiling Cao
Professor of Mathematics

Co-chairs of the 2018 AUT Mathematical Sciences Symposium

AUT CITY CAMPUS

55 Wellesley Street East, Auckland 1010

SCHOOLS

Art & Design – Level 3, WE building
Business – Level 1, WF building
Colab – Level 11, WG building
Communication Studies – Level 12, WG building
Engineering, Computer & Mathematical Sciences – Level 1, WT building and Level 3, WZ building
Hospitality & Tourism – Level 3, WH building
Language & Culture – Level 8, WT building
Law – Level 6, WY building
Science – Level 5, WZ building
Social Sciences & Public Policy – Level 14, WT building
Te Ara Poutama – Level 3, WB building

STUDENT HUB

Level 2, WA building

Phone: 0800 AUT UNI (0800 288 864)
Web: www.aut.ac.nz/studenthub

SERVICES AND FACILITIES

AUT International Centre – Ground Floor, WY building
AUTSA (Auckland University of Technology Student Association) – Level 2, WC building
Early Childhood Centre – Level 2, WA building via Gate 2
Estates Service Centre, Security – Corner St Paul & Wakefield St, WO building
Learning Lab – Level 3, WA building
Library – Level 4, WA building
PinkLime (print services) – Level 3, WA building
Student Counselling & Mental Health – Room 204, WB building
Student Medical Centre – Room 219, WB building
ubiq (formerly University Bookshop) – Room 122, WC building

- Student Hub
- Student lounge
- Cafés
- Library
- Early Childhood Centre
- Gym
- Conference facility
- Intercampus shuttle bus stop
- Mobility parks

Symposium Schedule

	Thursday 22nd November	
9:00- 9:30	**Registration** WF Level 7	
9:30 - 9:45	**Welcome** WF710	
9:45-10:30	**Renate Meyer** Bayesian Nonparametric Time Series Analysis and Its Applications WF710	
10:30-11:00	Morning Tea	
11:00-11:45	**Michael O'Sullivan** Analytics/OR for Social Investment Planning WF710	
11:50-12:15	**Victor Miranda Soberanis** WF710	**Graeme Wake** WF711
12:15-13:30	Lunch	
13:30-14:15	**Matt Parry** Stochastic modelling, smoothing splines and scoring rules WF710	
14:20-14:45	**Jelena Cosic** WF710	**Catherine Hassell Sweatman** WF711
14:50-15:15	**Oliver Stevenson** WF710	**Robert Borotkanics** WF711
15:15-15:45	Afternoon Tea	
15:45-16:10	**Simon Harris** WF710	**Ruanui Nicholson** WF711
16:15-16:40	**Armando Amaris** WF710	**Graham Weir** WF711
16:45-17:30	**Winston Sweatman** Symmetrical four-body problems in one to three dimensions WF710	

18:30	**Dinner** Four Seasons Restaurant, 55 Wellesley St E

	Friday 23rd November	
9:00-9:45	**Bob Durrant** Random Projections for Dimensionality Reduction WF710	
9:50-10:15	**Budhi Surya** WF710	**Robin Hankin** WF711
10:15-10:45	Morning Tea	
10:45-11:10	**Roy Costilla** WF710	**Gaurav Kapoor** WF711
11:15-11:40	**Jeffrey Hunter** WF710	**Rewat Khanthaporn** WF711
11:45 - 12:10	**Azam Asanjarani** WF710	**Jiling Cao** WF711
12:10-13:30	Lunch	
13:30-13:55	**Stuart Weston** WF710	**Wenjun Zhang** WF711
14:00 - 14:25	**Murray Jorgensen** WF710	**Maryam Hasannasab** WF711
14:30 - 14:55	**Sarah Marshall** WF710	**Shu Su** WF711
14:55-15:25	Afternoon tea	
15:25-16:10	**Gregoire Loeper** Reconstruction of missing data by optimal transport: applications in cosmology and finance WF710	
16:10-16:15	**Farewell**	

Symposium Information

Location

The AUT Mathematical Sciences Symposium will be held in **WF Building**, corner of Wakefield Street and Mayoral Drive, Auckland Central, rooms WF710 and WF711.

Registration

Registration will take place on level 7 of the WF building.

Presentations

Invited talks will be 40 minutes with 5 minutes for questions and contributed talks will be 20 minutes with 5 minutes for questions. There is a 5 minute break after each invited talk to allow delegates to move between rooms.

Refreshments

Morning tea and afternoon tea will be served on level 7 of the WF Building. There are a large number of choices for lunch within a short walking distance of the campus.

Dinner

The symposium dinner will begin at 6:30pm on Thursday 22nd November 2018. The venue for the dinner is Four Seasons, WH Building, corner of Mayoral Drive and Wellesley Street East, Auckland.

The cost of dinner will be covered by Mathematical Sciences Research Group (MSRG) for participants who are presenting at the symposium. Partners are welcome to attend the dinner, however unfortunately the cost of their meals will not be covered by MSRG. If you need to pay for your or your partner's dinner, please discuss this with a member of the organising committee when you register.

Further Queries

If you have any queries please do not hesitate to contact a member of the organising committee: Jiling Cao, Jeffrey Hunter, Kate Lee, Sarah Marshall, Nuttanan Wichitaksorn and Wenjun Zhang.

Sponsors

Thank you to the sponsors of this event.

List of Authors and Titles

Riemann surfaces, Human Migrations and Polytope Complexes.

Armando Amaris
Finance Department, Auckland University of Technology
armando.amaris@aut.ac.nz

We show how polytope complexes associated to cubic graphs can help us to understand both the structure of the moduli space of closed Riemann surfaces and the structure of the space of migrational patterns that arise while considering the problem of reconstructing ancient human migrations. In the context of Riemann surfaces of genus two, we can build its associated polytope complex in a canonical way, starting by projecting each Riemann surface of genus two to the hyperbolic sphere of dimension two. In the context of human migrations, migrational patterns arise naturally because we can model our earth as a two dimensional sphere and based in the location of several populations we can get a Voronoi diagram that determines a decomposition of the sphere. We have different geometries that play an important role in Riemann surfaces and human migrations, basically identical combinatorial behavior, and definitely a unified way for viewing two classification problems that initially looks unrelated.

Stability of a reward-observing queueing system

Azam Asanjarani, Yoni Nazarathy
Department of Statistics, The University of Auckland
azam.asanjarani@auckland.ac.nz

We consider a simple discrete-time controlled queueing system, where the controller has a choice of which server to use at each time slot and server performance varies according to a Markov modulated random environment. We explore the role of information in the system stability region. At the extreme cases of information availability, that is when there is either full information or no information, stability regions and maximally stabilizing policies are trivial. But in the more realistic cases where only the environment state of the selected server is observed, only the service successes are observed or only queue length is observed, finding throughput maximizing control laws is a challenge. To handle these situations, we devise a Partially Observable Markov Decision Process (POMDP) formulation of the problem and illustrate properties of its solution. We further model the system under given decision rules, using Quasi-Birth-and-Death (QBD) structure to find a matrix analytic expression for the stability bound.

Multi-collinearity in statistical prediction of disease status and potential effects on clinical decision making

Robert Borotkanics
DoBE, AUT
robert.borotkanics@aut.ac.nz

The broad effects of multi-collinearity are well described in the peer-reviewed literature. Collinearity increases the standard error of beta coefficients, increases the risk of covariate significance errors. This phenomenon also reduces study power. This means that potentially important covariates in prediction models could be unintentionally excluded from final models, for instance. The facets of collinearity in the context of clinical decision aids and therefore clinical decision making is not well described. Recent technological advances have resulted in the emergence of the systematic collection of individual health status via electronic health records. These health records are being embedded with greater frequency with clinical guidelines and clinical decision aids. These are often referred to as clinical decision support. There are instances where potentially collinear covariates collected in an electronic health record

– e. g., prostate specific antigen and age – could be collected and used in a clinical decision. The potential implications and significance of collinear covariates used in the face of clinical decision making is ever important as the effect of such collinearity in a predictive model and its associated diagnostic accuracy are poorly understood. Receiver operating characteristic (ROC) analyses are commonly applied to appraise the accuracy of clinical decision aids, like diagnostic tests. This approach is beneficial in that continuous covariates can be described in the context of probability of being disease positive or negative. Regulators, like the US Food and Drug Administration often require ROC analyses of software-based clinical decision aids, as in many scenarios the software serves as a form of device-based decision aid. A recent example is FDA's recent approval of Viz. AI. Therefore, effect of collinear clinical covariates in a software may have regulatory implications as well. Therefore, the author has undertaken a two-part study. The first part of the study evaluates the effect of collinear predictors via a series of simulation experiments, reporting the expected effect under defined forms of collinearity. These simulation experiments are followed by two case examples. These case examples form the study's second part. The first case example, based on data from a recent case-control study, appraises the effect of collinear biomarkers on the accurate identification of prostate cancer cases. The second case study is based on an intervention-based cohort study that evaluated pain 12 months after reconstructive knee surgery. In this study, predictors of long term pain, post-operative pain, many collinear, were evaluated and again, reported. In both examples the effect of collinearity on diagnostic accuracy and misclassification are described. Potential implications for future clinical decision support applications are summarized.

Pricing Variance Swaps under Hybrid CEV and Stochastic Volatility

Jiling Cao, Jeong-Hoon Kim, Wenjun Zhang
Department of Mathematical Sciences , Auckland University of Technology
jiling.cao@aut.ac.nz

In this paper, we consider the problem of pricing a variance swap whose underlying asset price dynamics is modeled under a hybrid framework of constant elasticity of variance and stochastic volatility (CEVSV). Applying the multi-scale asymptotic analysis approach, we obtain a semi-closed form approximation of the fair continuous variance strike. We conduct numerical experiments by applying this approximation formula to calculate the square root of the fair continuous variance strike with different values of parameters. The market data of S&P 500 options are used to obtain calibrations of the CEVSV model, and then the estimated parameters are further used to compute the values of the square root of fair continuous variance strike. In addition, we also analyse and compare the performance of the CEV model, the CEVSV model and the Heston stochastic volatility model.

Bayesian Network as a Modelling Tool for Increasing Knowledge on the Factors Influencing Vineyard Longevity and Sustainability

Jelena Cosic, Steffen Klaere, Matthew Goddard, Bruno Fedrizzi
School of Biological Sciences and Department of Statistics, University of Auckland
j.cosic@auckland.ac.nz

The long-term project 'Resilient and Profitable NZ wine industry' has the objective to study the impact of different vineyard management techniques on the vineyard longevity and profitability, and to increase the knowledge of the factors influencing longevity and profitability. To find meaningful answers appropriate quantifiable outcomes need to be obtained. Profitability of a vineyard can be quantified by its yield and quality of the end product, while health will be studied in a more holistic way by developing a vineyard ecosystems model incorporating the data obtained from different areas of interest. The

empirical nature of data collection makes a computational ecosystem modelling approach the most suitable. Such approaches are quite common and popular in ecology, and are promising for this project. Of particular interest are Bayesian Networks (BNs) which have received increased attention throughout several research fields for their ability to incorporate prior knowledge and to handle incomplete data. BN have also been shown to efficiently avoid overfitting the data, and avoiding the observation of 'chimeric' effects. We will use BN to model vineyard ecosystems incorporating microbial, fungal and eukaryotic molecular data,chemical profiles, meteorological information, and other markers at different points in the life cycle of vineyards, and discover the differences vineyard managements make with respect to resilience and profit. Some of the challenges that we see are: variables that have been measured on different time scales, a large amount of microbial data and uncertainty of the interactions of components included in our ecosystem.

Genetic control of temperament traits across species: association of autism spectrum disorder genes with cattle temperament

Roy Costilla, Kathryn Kemper, Enda Byrne, Naomi Wray, Ben Hayes
Institute for Molecular Bioscience, University of Queensland
r.costilla@imb.uq.edu.au

Temperament traits are relevant to both human and cattle well-being. In humans, they are genetically correlated with several psychiatric disorders and in cattle, they are of commercial importance as more docile animals have better performance. We hypothesised that genetic factors contributing to variation between individuals will be shared across these species and investigate the association of Autism Spectrum Disorder (ASD) genes with cattle temperament. We conducted a series of genome-wide association studies on cattle flight time, a temperament phenotype defined as time taken for an animal to cover a short fixed distance after being released from an enclosure, in two breeds of cattle and meta-analysed the results for a combined sample of 9223 animals and 28. 3 million SNPs. We tested the effects of SNPs surrounding 63 bovine orthologous genes associated with ASD in this meta-analysis. The resulting 180,000 SNPs within 100Kb showed a skewed QQ plot, providing evidence of a higher association than expected by chance. These SNPs also explained 7% of the total additive genetic variance in the biggest cattle cohort (randomized permutation test P value <0. 02). ASD genes showed enrichment in brain (up-regulation), mammary (down-regulation) and white skin (down-regulation) bovine tissues. Genes with the most significant associations with the cattle flight time phenotype in the ASD set were GABRB3, CUL3, and INTS6. These genes contribute to inhibition of neurotransmitters in the vertebrate brain, protein binding, and signalling receptor activity. These findings provide computational evidence that genetic control of temperament traits might be shared across these species.

Random Projections for Dimensionality Reduction

Bob Durrant
Department of Mathematics and Statistics, University of Waikato
bobd@waikato.ac.nz

Linear dimensionality reduction is a key tool in the statistician's toolbox, used variously to make models simpler and more interpretable, to deal with cases when n<p (e. g. to enable model identifiability), or to reduce compute time or memory requirements for large-scale (high-dimensional, large p) problems. In recent years, random projection ('RP'), that is projecting a dataset on to a k-dimensional subspace ('k-flat') chosen uniformly at random from all such k-flats, has become a workhorse approach in the machine learning and data-mining fields, but it is still relatively unknown in other circles. In this talk I

will review an elementary proof of the Johnson-Lindenstrauss lemma (JLL) which, perhaps rather surprisingly, shows that (with high probability, independent of the data dimensionality) RP approximately preserves the Euclidean geometry of projected data. This result has provided some theoretical grounds for using RP in a range of applications but from the perspective of statistical inference such guarantees are rather unsatisfying. In particular JLL-based guarantees for statistical inference on RP'ed data generally become weaker as the sample size increases, contrary to expectation and experience. Focusing on the problem of classification (a.k.a. discriminant analysis) I show that a more careful approach – that was inspired by a recreational problem concerning 'random triangles' – can avoid this issue and furnish dimension-free guarantees for linear classification that don't grow with the sample size and are meaningful even when n<p. I discuss some insights and consequences of this theory: What makes a classification problem easy?

Formula 1, Australian Masterchef, and Counterstrike

Robin Hankin
Department of Mathematical Sciences, AUT
robin.hankin@aut.ac.nz

The Bradley-Terry model for datasets involving paired comparisons or competition has wide uptake in the R community. However, existing functionality is restricted to pairwise comparison or conflict, such as between two boxers, chefs, or squash players. Here I present software and analysis for the more general case where individuals compete against one another as part of a team. Here, "compete" is broadly defined in terms of competitive situations that include racing, subjectively defined comparisons such as artistic performances, direct martial combat and e-sports. The method allows me to quantify difficult aspects of competition such as team cohesion and non-transitive sport tactics. I will demonstrate the new "hyper2" R package for generalized Bradley-Terry models and use it to estimate individual players' strengths in a range of competitive situations including Formula 1motor racing, Australian MasterChef, and e-sports such as Counterstrike.

The genealogy of uniform sample from a Galton Watson process

Simon Harris, Matthew Roberts, Samuel Johnston
Department of Statistics, University of Auckland
simon.harris@auckland.ac.nz

Consider a continuous-time Galton-Watson branching process. If the population survives until some large time T, then choose k particles uniformly from those alive. What does the ancestral tree drawn out by these k particles look like? Some special cases were known, eg. Durrett (1978), O'Connell (1995), Athreya (2012), but we will discuss some recent more complete answers.

Hyperbolic Efficiency Measurement: A Conic Programming Approach

Maryam Hasannasab
The University of Auckland
maryam.hasannasab@auckland.ac.nz

The hyperbolic distance function (HDF) model was first introduced in 1985 but despite its attractive features in allowing for greater flexibility for firms to modify inputs and outputs to increase efficiency and to readily model undesirable outputs, it has not been as popular in empirical applications. We believe this is in part due to the difficulty of solving the variable returns to scale model. We develop a computational procedure for the hyperbolic model within the nonparametric framework of data envelopment analysis. We convert the nonlinear hyperbolic model under variable returns to scale into an equivalent conic optimization problem with linear constraints plus a 'toppled ice cream' cone constraint that can be efficiently solved by specialized interior point methods in about the same time as a linear program. Applying the dual of an ice cream cone, we formulate a multiplier based (dual) HDF model. We elaborate on the structural details of both primal and dual HDF models through geometrical figures. We apply our method to measure the performance of a sample of US banks producing both desirable (e. g. loans, securities) and undesirable outputs (non-performing loans).

Mathematical model of diabetes and lipid metabolism : investigating the timescales corresponding to both low carbohydrate diets and low fat diets for weight loss

Catherine Hassell Sweatman
Department of Mathematical Sciences , Auckland University of Technology
catherine.sweatman@aut.ac.nz

Currently there is great debate on the relative merits and dangers of fats and sugars in our diet. Both low carbohydrate and low fat diets are recommended for those who wish to lose weight and avoid Type II diabetes. In order to investigate the relative merits of different diets, a mathematical model of glucose, insulin, glucagon, β-cell, leptin and fat dynamics and hepatic, peripheral and adipose insulin sensitivity is presented. The model includes plasma non-esterified fatty acids, muscle lipids, hepatic lipids and very low density lipoprotein triglycerides, allowing comparisons with recent health research. Decay constants for fat mass are predicted to vary widely, depending on the diet. Implications for experimental work are discussed.

Kemeny's Constant for Markov Renewal Processes

Jeffrey Hunter
Department of Mathematical Sciences, Auckland University of Technology
jeffrey.hunter@aut.ac.nz

Kemeny's constant, as derived for irreducible finite Markov chains in discrete time, is extended to Markov renewal processes. We first survey Markov chain derivations leading to two possible constants, one with the interpretation as the expected time to a randomly chosen state, ("expected time to mixing") and one as the "expected time to hitting" (when the chosen state happens to be sampled at the first transition). Such functions, as expected times, do not depend on the initial starting state and are thus constants (Kemeny's constants). Three alternative Kemeny's functions and their variants, under "mixing" and "hitting" conditions, are considered for Markov renewal processes. Following the derivation of specific expressions, it is shown that typically they lead to a constant if and only if the mean holding times between the states in the Markov renewal process are constant. However, one particular variant

(under hitting conditions) leads to a constant, independent of the initial state, analogous to the discrete time Markov chain result. Specifically, if the state space S is finite, the sum of each mean first passage time $m_{i,j}$ (omitting the mean return time $m_{j,j}$ when $i = j$) weighted by the stationary probability ω_j associated with the continuous time embedded semi-Markov process, is a constant, independent of i, for any Markov renewal process.

Scaling iterative parameter estimation to big data

Murray Jorgensen
Department of Mathematical Sciences, Auckland University of Technology
majmurr@gmail.com

In this talk we will consider an approach to carrying out statistical analyses on large data sets known as 'Divide and Conquer' or 'Divide and Recombine'. We are concerned with situations where the data is abundant and is divided into subsets. We assume that it is the analysis of the full data set that we are concerned about and that the partitioning of the data is simply a nuisance. In the 'big data' context it is not always the case that the analyst is in control of the division of the data. For example the data may be stored on a number of machines at various locations corresponding to nodes on a network. We might envisage some operations being performed in parallel on a number of nodes and others taking place at a central machine. In a different situation the data may be stored on a single, very large file, too large to be read into RAM for analysis. In this case a subset could consist of a 'chunk' of the file of appropriate size for reading into RAM. We will introduce some variants of this type of algorithm and then consider how well these perform when the analysis has to be carried out by an iterative algorithm such as Iteratively Reweighted Least Squares or Expectation Maximization.

Pricing Leveraged ETF Options Under Heston Dynamics

Gaurav Kapoor
Department of Mathematical Sciences, Auckland University of Technology
gk7927@gmail.com

Exchange traded funds (ETFs) and their derivatives have seen an enormous growth in the financial industry over the last two decades, presently trading over 1. 5 trillion USD in assets under management. Leveraged ETFs (LETFs) are a byproduct of regular ETFs that have seen both criticism and praise in the markets. This thesis derives a pricing formula for call options of LETFs with the assumption that the underlying asset follows Heston model dynamics. Stochastic volatility is an important factor in modelling LETFs, and the Heston model appropriately captures this characteristic. The model includes an additional process to reliably capture the path-dependence of LETFs. A relationship between the price of an LETF and the value of its underlying asset is established. This relationship is dependent on the leverage ratio of the LETF and the path-dependent volatility of the underlying asset. Through empirical analysis, the accuracy of this link is justified. Furthermore, this link provides useful information on the behaviour of LETFs, which is analyzed in depth. The option pricing formula is derived by defining the joint moment-generating function of the underlying asset and its volatility and linking this function to the characteristic function of an LETF. The Carr-Madan Fourier transform method is utilized to obtain a closed-form solution for call prices. The prices are compared with those obtained using Monte-Carlo simulations and the results are consistent, with the advantage of greatly reduced computational times. Sensitivity analysis is performed to study the effect of various parametric changes on the pricing formula.

On the Cryptocurrency and Risk and Return Performance

Rewat Khanthaporn
Department of Mathematical Sciences, Auckland University of Technology
rkhantha@aut.ac.nz

A new asset class, cryptocurrency, has been proposed after the first digital currency was invented by Satoshi Nakamoto in 2009. Currently, there are approximately 2,000 virtual currencies available in the market where its capitalization has exceeded USD 2 trillion. In this talk, the cryptocurrency is considered as an asset class for dynamic portfolio management in which its risk and return performance is assessed. Based on the empirical evidence and the numerical results, the cryptocurrency tends to have higher risk and return than the other asset classes. Moreover, an introduction to digital currency market and its characteristics, with Bitcoin as an example, will be presented to better understand the cryptocurrency's ecosystem.

Reconstruction of missing data by optimal transport: applications in cosmology and finance

Gregoire Loeper
School of Mathematical Sciences, Monash University
Gregoire.Loeper@monash.edu

Optimal Transport is an old optimisation problem that goes back to Gaspard Monge in 1781. I will give some historical perspective of the problem and its solutions, and then present some recent results where techniques from Optimal Transport can be used, going from the problem of reconstruction of the early universe, to a problem of model calibration in finance.

Modelling a renewing free repair warranty using an alternating geometric process

Sarah Marshall, Richard Arnold, Stefanka Chukova, Yu Hayakawa
Department of Mathematical Sciences, Auckland University of Technology
sarah.marshall@aut.ac.nz

The cost of rectifying warranty claims can be significant and therefore accurate estimation of the warranty cost is important. We model the product life cycle using an alternating geometric process (AGP), in which there is an alternating sequence of operational and repair times. To accommodate the ageing of the product and repair equipment, we use a decreasing geometric process to model the consecutive operational times and an increasing geometric process to model the consecutive repair times. We model the warranty claims process and evaluate the warranty servicing costs under non-renewing, renewing, and restricted renewing free repair warranties. In this talk, theoretical results relating to the renewing and restricted renewing free repair warranties will be presented. Properties of the model will be demonstrated using a simulation study.

Bayesian Nonparametric Time Series Analysis and Its Applications

Renate Meyer, Claudia Kirch, Matthew Edwards, Alexander Meier
Department of Statistics, University of Auckland
meyer@stat.auckland.ac.nz

Nonparametric Bayesian inference has seen a rapid growth over the last two decades and various nonparametric Bayesian approaches to time series analysis have been developed. Most notably, Carter and Kohn (1997), Gangopadhyay (1998), Choudhuri et al. (2004), and Rosen et al (2012) used Whittle's likelihood for Bayesian modeling of the spectral density as the main nonparametric characteristic of stationary time series. As shown in Contreras-Cristan et al. (2006), the loss of efficiency of the nonparametric approach using Whittle's likelihood approximation can be substantial. On the other hand, parametric methods are more powerful than nonparametric methods if the observed time series is close to the considered model class but fail if the model is misspecified. Therefore, we suggest a nonparametric correction of a parametric likelihood that takes advantage of the efficiency of parametric models while mitigating sensitivities through a nonparametric amendment. We use a nonparametric Bernstein polynomial prior on the spectral density with weights induced by a Dirichlet process. Contiguity and posterior consistency for Gaussian stationary time series have been shown in by Kirch et al (2018). Bayesian posterior computations are implemented via a MH-within-Gibbs sampler and the performance of the nonparametrically corrected likelihood is illustrated in a simulation. We illustrate this approach through applications in physiology, ecology and astrophysics: analysing heart rate variability in ECG time series, the Southern Oscillation Index (one of the key atmospheric indices for gauging the strength of El Nino events and their potential impacts on the Australian region) and LIGO gravitational wave data.

Vector Generalized Linear Time Series Models

Victor Miranda Soberanis, Thomas Yee
Department of Statistics, University of Auckland
vmsoberanis@gmail.com

In this talk I will provide an overview of the recently proposed class of time series (TS) models called Vector Generalized Linear Time Series Models (VGLTSMs), which can be thought of as an extension of the GLM-type regression models for TS analysis using vector generalized linear models (VGLMs). VGLTSMs are greatly motivated by the plethora of models detached from the ARMA–GARCH class of TS, having pockets of substructure but little overriding framework. The crucial VGLM ideas are constraint matrices, vector responses and covariate-specific linear predictors, and estimation by iteratively reweighted least squares and Fisher scoring. The only addition to the VGLM framework is a log-likelihood that depends on past values. It will be shown how several classes of TS models and cointegrated TS are accommodated as special cases of VGLTSMs. Time permitting, algorithmic details of its implementation in R, that are to be compared to other software for TS analysis, and properties such as stationarity and the expected information matrices are surveyed.

Use of the Bayesian Approximation Error Approach to Account for Model Discrepancy: The Robin Problem Revisited

Ruanui Nicholson, Noemi Petra, Jari Kaipio
Department of Engineering Science, University of Auckland
rnic052@aucklanduni.ac.nz

We address the problem of accounting for model discrepancy by the use of the Bayesian approximation error (BAE)approach in the context of inverse problems. In many inverse problems when one wishes to infer some primary parameter of interest there are other secondary parameters which are also uncertain. In the standard Bayesian (or deterministic) approach such nuisance parameters are either inverted for or are ignored (perhaps by assignment of some nominal value). However, it is well understood that the ill-posedness of general inverse problems means they do not handle modelling errors well. The BAE approach has been developed as an efficient means to approximately pre-marginalize over nuisance parameters so that one can systematically incorporate the effects of neglecting these secondary parameters at the modelling stage. We motivate the method through an application to the Robin problem governed by the Poisson equation.

Analytics/OR for Social Investment Planning

Mike O'Sullivan
Department of Engineering Science, University of Auckland
michael.osullivan@auckland.ac.nz

"Social Investment is about improving the lives of New Zealanders by applying rigorous and evidence-based investment practices to social services." – treasury. govt. nz. "Social investment contrasts with traditional approaches to funding government activity, which focus more heavily on what was delivered and ensuring good value for money in the delivery of those services, rather than on the value of outcomes achieved." – State of the State New Zealand 2016, Deloitte. This presentation will: 1) discuss how social investment was utilised in Oranga Mahi, a cross-agency partnership between the Ministry of Social Development, the University of Auckland and multiple District Health Boards (DHBs) throughout NZ; 2) summarise two infrastructure investment planning projects from the UK and Auckland respectively; and 3) present a proposed generic framework for investment that generalises all of the previous models and incorporates BigData, Analytics, and Operations Research to enable robust social investment planning.

Stochastic modelling, smoothing splines and scoring rules

Matt Parry
Dept of Mathematics & Statistics, University of Otago
mparry@maths.otago.ac.nz

If the answer to uncertainty is modelling, then the answer to certainty is scoring rules. In this talk, I pull together several strands of my recent work on the spread of disease, trajectory reconstruction, and the evaluation of probabilistic forecasts. In all cases, probability is the key ingredient that allows us to quantify uncertainty in the inputs and the outputs of our models. I also discuss how scoring rules can play a role in both assessing the quality of a model's output and in inferring the model's inputs.

Gaussian process models for predicting batting ability in cricket

Oliver Stevenson
Department of Statistics, University of Auckland
o.stevenson@auckland.ac.nz

In the sport of cricket, variations in a player's batting ability can usually be measured on one of two scales. Short-term changes in ability that are observed during a single innings, and long-term changes that are witnessed between matches, over entire playing careers. To measure short-term, within-innings variation, a Bayesian survival analysis method is derived and used to fit a model which predicts how the batting abilities of professional cricketers evolve during an innings. The results from the within-innings model provide evidence to support the cricketing belief of 'getting your eye in', whereby batsmen are more vulnerable early in their innings, but improve as they adapt to the specific match conditions. A second model is then fitted to explain how player batting ability changes between-innings, from match to match. To account for both recent performances and the element of randomness associated with cricket, the model uses a Gaussian process to measure and predict current and future batting abilities. Generally speaking, the results from the between-innings model support an anecdotal description of a typical sporting career. Young players tend to begin their careers with some raw but undeveloped ability, which improves over time as they gain experience and participate in specialised training and coaching regimes. Eventually players reach the peak of their careers, after which ability tends to decline. The results provide more accurate quantifications of a player's batting ability at any given point of their career, compared with traditional cricketing metrics, and have practical implications in terms of player comparison, talent identification and team selection policy.

Pricing VIX derivatives with infinite-activity jumps

Shu Su, Jiling Cao, Xinfeng Ruan, Wenjun Zhang
ECMS, AUT
sus.bearwith@gmail.com

We propose a group of models with infinite-activity jumps for pricing VIX options and futures and conduct an extensive empirical analysis of the effect of the different types of jumps. We consider two types of pure Levy jump processes in our jump specification. The processes include variance gamma process (VG) and normal inverse Gaussian process, which are based on Brownian motion with a different subordinator, gamma process and inverse Gaussian process respectively. We apply a deterministic sampling technique, Unscented Kalman filter (UKF), and quasi-maximum log-likelihood estimate method (QMLE) to calibrate our models. To compare with the models with finite jumps and without jumps, we find that the models with infinite-activity jumps improve the accuracy of prediction, especially in the VIX options valuation.

Parisian excursion below a fixed level from the last record maximum of Levy insurance risk process

Budhi Surya
School of Mathematics and Statistics, Victoria University of Wellington
budhi.surya@vuw.ac.nz

In this talk I present some new results on Parisian ruin under Levy insurance risk process, where ruin occurs when the process has gone below a fixed level from the last record maximum, also known as the high-water mark or drawdown, for a fixed consecutive periods of time. The law of ruin-time and the position at ruin is given in terms of their joint Laplace transforms. Identities are presented semiexplicitly in terms of the scale function and the law of the Levy process. They are established using recent developments on fluctuation theory of drawdown of spectrally negative Levy process. In contrast to the Parisian ruin of Levy process below a fixed level, ruin under drawdown occurs in finite time with probability one.

Symmetrical four-body problems in one to three dimensions

Winston Sweatman
Institute of Natural and Mathematical Sciences, Massey University
w.sweatman@massey.ac.nz

The dynamics of several point masses moving under mutual gravitational attraction is a challenging mathematical problem. To provide insight into more general few-body cases, and to aid our understanding, it is useful to begin our study with some simpler examples. The symmetrical four-body problem is one such example with a symmetrical arrangement of masses. Providing we deal with collisions, another simplification is to start in one dimension before moving to higher dimensions.

Why are our pine trees going red? The problem of Red Needle Cast.

Graeme Wake, Nari Williams, Penelope Bilton, Rebecca Turner, Tony Pleasants
Institute of Natural and Mathematical Sciences, Massey University
g.c.wake@massey.ac.nz

Red needle cast was first detected in New Zealand in 2008 but it was probably present in forests for a few years before that. It is caused by a strain of phytohthora that results in pine needles turning red and being shed prematurely. A tree without needles does not grow very much. The disease can cause up to forty percent growth loss in a year. The origin of red needle cast in New Zealand has been traced to Oregon in the United States. It is believed to have been transported here in plant material and on forestry machinery. A simple systems model has been proposed to understand and quantify the onset and epidemiology of red needle cast in radiata pine. This disease is impacting much of the New Zealand forestry estate being driven through the production of self-replicating spores which are dispersed with water. The first model is at present deterministic, not spatially or age-structured, and initially not including seasonal or environmental effects. This model showed the clear existence of calculable thresholds for disease proliferation and elimination, thereby capturing the essential components of the biological mechanisms. It is to be used to identify thresholds for infection to spread or to disappear. Further it provides an easy means to predict outcomes for different scenarios. In this paper the established model which previously had neglected the effects of the environment, is generalised to include seasonal effects. The weather cycle drives the solution to produce in some cases quite different long-term outcomes, depending on the external parameters. The system is of course now non-autonomous, with weather imposed, almost yearly periodicity. Coexisting stable long-term solutions are also then driven to exhibit this periodicity.

Magnetic field from a 3D block magnet

Graham Weir, George Chisholm, Jerome Leveneur
Institute of Fundamental Sciences, Massey University
grahamweir@xtra.co.nz

Neodymium magnets were independently discovered in 1984 by General Motors and Sumitomo. Today they are the strongest type of permanent magnet commercially available. They are the most widely used industrial magnet, with many applications, including in hard disk drives, cordless tools and magnetic fasteners. We derive a mathematical model of the 3D magnetic field for a neodymium magnet, assuming an idealised block geometry and uniform magnetisation. For each field or observation point, the 3D solution involves twenty four non-dimensional quantities, arising from the eight vertex positions of the magnet, and the three components of the magnetic field. The only unknown in the model is the value of magnetisation, with all other model quantities defined in terms of field position and magnet location.

Applying the Likelihood Ratio for Astro Physics

Stuart Weston, Nicolas Seymour, Sergei Gulyaev
Institute for Radio Astronomy & Space Research (IRASR), Auckland University of Technology
sweston@aut.ac.nz

Large multi-wavelength astronomical surveys are an important method to gain statistically significant datasets for the understanding of galaxy formation and evolution. One key problem is determining which sources (galaxies) are truly associated with one another and which are unrelated between these different wavelength surveys. Where surveys have similar wavelengths, resolutions and sensitivity this matching of sources can be achieved with a simple nearest neighbour match; however where the surveys are dissimilar with very different resolutions and possibly sensitivity such as between radio and infrared, a nearest neighbour approach becomes unreliable. To date this has been a manual process, however with new surveys from facilities such as the Australian Square Kilometre Array Pathfinder (ASKAP) detecting many tens of millions of radio sources such an approach is no longer feasible. This work presents cross-identifying radio sources in the Australia Telescope Large Area Survey (ATLAS) with infrared (IR) counterparts from the FUSION Spitzer catalogue, using a refinement of the Likelihood Ratio (LR). The algorithm and pipeline use Python and a Relational Database to automate the process of cross-identifying between the two catalogues. The developed LR algorithm has been extended for more complex cases such as two or more IR source contributing to one radio source, and also to identify possible complex radio sources (where multiple radio components are matched to one IR source). We examine refinements to previous LR cross-matching techniques where we incorporate the LR and the derived reliability.

Inferring information from the S&P 500 and CBOE indices

Wenjun Zhang, Jiling Cao, Xinfeng Ruan
Department of Mathematical Sciences, Auckland University of Technology
wzhang@aut.ac.nz

The Chicago Board Options Exchange (CBOE) updated the CBOE Volatility Index (VIX) in 2003 and further launched the CBOE Skew Index (SKEW) in 2011, in order to measure the 30-day risk-neutral volatility and skewness of the S&P 500 Index (SPX). This paper mainly compares the information extracted from the SPX and CBOE indices in terms of the SPXoption pricing performance. Based on our empirical analysis, VIX is a very informative index for option prices. Whether adding the SKEW or the VIX term structure can improve the option pricing performance depends on the model we choose. Roughly speaking, the VIX term structure is informative for some models, while, the SKEW is very noisy and does not contain much important information for option prices.

Index of Authors