# Statistics and Mathematical Modelling in Combination

**La Trobe City Campus**

**Level 2, 360 Collins Street Melbourne, 3000**

**16-18 November 2022**

https://www.mathematics.org.au/sys/public/home.php?conf_id=57

Institute of Mathematics for Industry
Kyushu University

AMSI
AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

LA TROBE UNIVERSITY AUSTRALIA

# Workshop Information

Deterministic modellers and statisticians have much to gain by working in teams applying approaches from both fields. The combination of statistics and classical dynamics has long been a fertile area, tracing back to statistical mechanics from the end of the 19th century and stochastic differential equations from the 1920s. The need to combine the two modelling approaches has never been greater and neither has the opportunity for affordable high-performance computation.

As one application, during the COVID-19 pandemic, agreement has been found between agent-based models and differential compartment equations in modelling infection numbers. Each approach gives confidence to the other, and this suggests scope for new hybrid models.

Other relevant areas include:

- Model identification from a set of output data
- Inverse problems to determine transport coefficients from noisy data
- Evolution equations with uncertain model parameters
- Transport in random media
- PDE for financial derivatives based on fluctuating stocks
- Stationary distributions of ergodic dynamical systems
- Boltzmann equation based on statistical dynamics
- Transport in fractal media
- Complex systems in ecology and economics
- Probabilistic interpretation of deterministic Schroedinger equation

## Organising committee

- Professor Marcel Jackson – La Trobe University (Chair)
- Emeritus Professor Philip Broadbridge – La Trobe University (Deputy Chair)
- Dr Christopher Lenard – La Trobe University (Conference Co-Director)
- Dr Rebecca Chisholm – La Trobe University (Conference Co-Director)
- Dr Luke Bennetts – University of Adelaide
- Associate Professor Joel C. Miller – La Trobe University
- Associate Professor Winston Sweatman – Massey University, New Zealand
- Professor Luke Prendergast – La Trobe University
- Professor Osamu Saeki – Kyushu University, Japan
- Professor Kenji Kajiwara, Kyushu University, Japan
- Associate Professor Pierluigi Cesana, Kyushu University, Japan
- Dr Anja Slim – Monash University
- Dr Melanie Roberts – Griffith University
- Dr Peter van den Kamp – La Trobe University
- Diana Heatherich – La Trobe University (secretariat)

## Plenary speakers

- Andrea Bertozzi
- iadine Chadès
- Jukka Corander
- Kei Hirose
- Shizuo Kaji
- Emma McBryde
- Oliver Maclaren
- David Price
- Melanie Roberts
- Michael Stumpf
- Natalie Thawwattana

## Acknowledgements

The Organising Committee gratefully acknowledges the assistance of John Banks. Generous funding is provided by La Trobe University, the Australian Mathematical Sciences Institute, and the Institute of Mathematics for Industry–Kyushu.

# Workshop Timetable

## Wednesday 16 Nov

| Time | Name | Title |
|---|---|---|
| **9:15–9:30** | | **Arrival and brief introduction** |
| 9:35–10:25 | Jukka Corander | Robust and scalable inference for simulator-based models |
| **10:30–10:45** | | **Morning Tea** |
| 10:45–11:35 | Shizuo Kaji | Modelling preference with hyperplane arrangement |
| 11:40–12:00 | Yvonne Stokes | Chemical signalling and tissue response: a moving boundary problem in biology |
| 12:00–12:20 | Rahil Valani | Dynamics of inertial particle focusing in curved ducts |
| 12:20–12:40 | Soukaina Hadiri | Some results on the mixed bifractional Brownian motion |
| **12:40–2:00** | | **Lunch** |
| 2:00–2:50 | David Price | Supporting government response to COVID-19 through model-based situational assessment |
| **2:55–3:10** | | **Afternoon Tea** |
| 3:10–4:00 | Melanie Roberts | MERGE and the role of gully erosion modelling to protect water quality on the Great Barrier Reef |

# Thursday, 17 Nov

| Time | Name | Title |
| --- | --- | --- |
| 9:15–10:05 | Andrea Bertozzi | |
| 10:10–10:30 | Malay Banerjee | Effect of slow-fast time scale on spatio-temporal pattern formation |
| **10:30–10:45** | | **Morning Tea** |
| 10:45–11:35 | Oliver Maclaren | Data, Models and Uncertainty: Perspectives and Tools |
| 11:40–12:00 | Sarah Vollert | A sequential method for efficiently parameterising ensemble ecosystem models |
| 12:00–12:20 | Jordan Pitt | Model predictions of wave overwash extent into the marginal ice zone |
| 12:20–12:40 | Adeshina Adekunle | A new mathematical modelling framework for capturing and forecasting Australia COVID-19 waves: transitioning from Delta wave into Omicron wave |
| **12:40–2:00** | | **Lunch** |
| 2:00–2:50 | Natalie Thamwattana | Interaction between nanostructures: relation between their atomic distributions and modelling approaches |
| **2:55–3:10** | | **Afternoon Tea** |
| 3:10–4:00 | Kei Hirose | Penalized likelihood approach in multivariate regression with missing values and its application to materials properties prediction |
| | | **Conference Dinner** |

# Friday, 18 November

| Time | Name | Title |
|---|---|---|
| 9:15–10:05 | Michael Stumpf | The Mathematics and Statistics of CellMaps and Whole Cell Modelling |
| 9:10–9:30 | Chris Drovandi | Likelihood-Free Methods and Model Misspecification |
| **10:30–10:45** | | **Morning Tea** |
| 10:45–11:35 | iadine Chadès | Developing AI Decision Tools for Conservation |
| 11:40–12:00 | Ton Viet Ta | Fish Schooling |
| 12:00–12:20 | Matthew Adams | Analysis of model sloppiness: what can it do, and what's next? |
| 12:20–12:40 | Saddam Abbasi | Identifying state of the process using ML algorithms |
| **12:40–2:00** | | **Lunch** |
| 2:00–2:50 | Emma McBryde | The application of mathematics to pandemics: some examples of modelling used during COVID-19 |
| **2:55–3:10** | | **Afternoon Tea** |
| 3:10–3:30 | Komal Singla | Symmetry Analysis and Exact solutions of fractional order (2+1)-dimensional Burgers system |
| 3:30-3:50 | Manoj Kumar | Analysis of Diffusive Size-Structured Population Model and Optimal Birth Control |

## Selected talks from the Forum Math for Industry

Courtesy of funding from AMSI, some talks at the Forum Math for Industry (the week following the conference) are available to conference participants through zoom.

| Time | Name | Title |
|---|---|---|
| **21 November** 10:00–10:50 | Wil Schilders | Mathematics: key enabling technology for scientific machine learning |
| 2:00–2:50 | Emma McBryde | The application of mathematics to pandemics: some examples of modelling used during COVID-19 |
| 3:10–4:00 | Oliver Maclaren | Identifiability analysis and predictive uncertainty for complex mathematical and simulation models |
| 6:00–7:00 | Malay Banerjee | **Christie Eliezer Memorial Lecture** Epidemic to immuno-epidemic models of COVID-19 |
| **22 November** 9:15–10:05 | Jukka Corander | Advances in likelihood-free inference with applications to evolutionary epidemiology |
| 10:30–11:20 | Natalie Thamwattana | Modelling clogging in granular assembly when treating acidic groundwater |
| 11:20–12:10 | Luke Bennetts | Modelling flexural strains at the outer margins of Antarctic ice shelves caused by ocean waves |
| 2:00–2:50 | Masayo Hirose | An Assessment of Prediction Error under Area Level Model with Arc-Sin Transformation |
| **23 November** 9:15-10:05 | Andrea Bertozzi | TBC |
| 11:20–12:10 | Stephen Taylor | Mathematical Modelling of nitrogen management on dairy farms |
| 2:00–2:50 | Shizuo Kaji | Homological features of 3D medical images |
| **24 November** 9:15-10:05 | Hugh Possingham | Decision science thinking applied to nature conservation |
| 10:30–11:20 | iadine Chadès | Challenges of developing decision tools to guide conservation decisions |
| 3:10–4:00 | Melanie Roberts | The effect of sediment heterogeneity on sediment yield during gully erosion |

# Plenary Abstracts

## Developing AI Decision Tools for Conservation

iadine Chadès

I will give an overview of the AI research we have been conducting to help make better decisions in the field of conservation (e.g. adaptive management) over the last 10 years. In particular, we have developed algorithms to solve and increase interpretability of Markov decision models and stochastic dynamic programming. I will be highlighting that the current bottleneck is not necessary our ability to solve complex decision problems, rather it is our ability to make solutions easy to interpret and more likely to be trusted.

## Robust and scalable inference for simulator-based models

Jukka Corander

Simulator-based models are becoming increasingly popular in many research domains across academia and industry. Calibration of such models in the light of data and quantification of uncertainty about model parameters are key challenges for practical applications and the topic has received accelerating attention during the past decade. We will discuss various inference techniques for simulator-based models that improve computational feasibility by adopting techniques from machine learning to build surrogate models for the approximate likelihood or posterior. Several of these approaches are available in the open-source software platform Engine for Likelihood-Free Inference (ELFI): elfi.ai.

## Likelihood-Free Methods and Model Misspecification

Chris Drovandi

Likelihood-free (LFI) methods such as approximate Bayesian computation are now widely adopted in many scientific disciplines for calibrating complex statistical models with intractable likelihood functions. In this talk I will describe some LFI methods developed by the computational statistics and machine learning communities, and discuss their performance under the commonly-encountered situation of model misspecification.

## Penalized likelihood approach in multivariate regression with missing values and its application to materials properties prediction

Kei Hirose

In the field of materials science and engineering, statistical analysis has recently been used to predict multiple material properties from an experimental design. These material properties correspond to response variables in the multivariate regression model. We conduct a penalized maximum likelihood procedure to estimate model parameters. In some cases, there may be a relatively large number of missing values in the response variables, owing to the difficulty of collecting data on material properties. We, therefore, propose a method based on the expectation-maximization (EM) algorithm to incorporate a correlation structure among the response variables into a statistical model.

## Modelling preference with hyperplane arrangement

Shizuo Kaji

A person's preference on a set of options, such as political parties and film genres, can be modelled by a (partial) order on the set. Modelling preference data collected from many individuals with various tastes is a subject of preference learning. There are two major approaches to modelling preference data; based on the distance on orders and based on a utility function defined over the set of options. These approaches lack flexibility (or are biased) since too much structure is forced on the preference data to be modelled by the mathematical structure that the models utilise. Instead, we rely on a geometric entity, hyperplane arrangement, to model preference data. The geometric and combinatorial structure of hyperplane arrangement provides a good balance of flexibility and regularisation.

## Data, Models and Uncertainty: Perspectives and Tools

Oliver Maclaren

Complex mathematical and simulation models are central to science, engineering, and policymaking. However, model sophistication frequently outpaces available data while interpretational issues can make drawing causal conclusions difficult. This talk will cover my research interests in methods for bridging the gap between data and complex mathematical models representing scientific and engineering understanding. I will illustrate with examples such as my work on COVID-19 modelling for policymakers, parameterisation of cell, tissue, and population dynamics models in biology, and large-scale uncertainty quantification for geothermal reservoir simulation models. I will discuss the challenges of model development, model interpretation, parameter identifiability, prediction vs. parameter estimation and mechanistic understanding, and how new and old statistical uncertainty quantification methods can help us use complex mechanistic models more effectively. I will highlight high- level conceptual issues and promising, practical, computationally-efficient methods. A key theme throughout will be the interplay between traditional applied mathematical modelling, statistics, and newer areas such as causal inference.

A key element of epidemic decision-making is situational awareness — that is, knowing the current and potential future status of the epidemic. Outputs from mathematical and statistical models have provided enhanced situational awareness to governments throughout the course of the COVID-19 pandemic. Key analyses include estimation of the effective reproduction number (Reff), forecasting of epidemic activity, and forecasts of ward- and ICU-bed demand. Accurate and timely estimation of Reff enables the tracking and planning of progress towards the control of outbreaks. Short-term forecasts of daily case incidence and hospital bed occupancy provide information on future health system requirements, which supports both clinical and public health planning.

In this talk, I will describe Australia's situational awareness modelling program for COVID-19 through 2020–21. I will provide an overview of the modelling outputs reported to key government decision-making committees on (at least) a weekly basis since April 2020, and highlight some challenges with providing near-real-time analytic support.

## MERGE and the role of gully erosion modelling to protect water quality on the Great Barrier Reef

Melanie Roberts

Gully erosion is the majority source of fine sediment that reaches the Great Barrier Reef (GBR), degrading water quality and contributing to poor outcomes including coral death. Consequently, gully remediation is a significant focus of activities to improve GBR water quality.

The MERGE gully erosion model was developed in partnership with Queensland Government and the Queensland Water Modelling Network to provide a process-based model to inform gully rehabilitation actions at specific sites. In this talk, I introduce the MERGE model and share the outcomes of a pilot study to explore the on-ground application of the model.

## The Mathematics and Statistics of CellMaps and Whole Cell Modelling

Michael Stumpf

The recently announced ARC Centre of Excellence for the Mathematical Analysis of Whole Cell Models aims to deliver the mathematics required to compute life. The Centre will establish in silico biology alongside in vivo and in vitro biology. These models will allow us to understand the complexity of life at the cellular level and enable new ways of combining diverse and heterogenous data. This will allow us to understand the mechanisms underlying cellular behaviour, and to apply rational design engineering methods in order to control the dynamics of biological systems. Here I will outline some of the exciting opportunities for mathematics and statistics that come out of this research program.

## Interaction between nanostructures: relation between their atomic distributions and modelling approaches

Natalie Thamwattana

Applications of nanomaterials are found in many areas, including renewable energy, electronics, textiles, food technology, environment, health care and medicine. Understanding mechanics of these materials is important as it can help to optimise their performances. In this talk, we consider structures of nanomaterials, the role that molecular structures play in determining an approach for modelling intermolecular interactions and their applications.

# Contributed Abstracts

## Identifying state of the process using ML algorithms

Saddam Akber Abbasi (Qatar University), Mohammed Ahmed (Qatar University) and Adegoke Nurudeen (The University of Sydney)

A process working under the random cause of variation is considered in-control whereas if special cause variations are in effect, the process is considered to be out-of-control. To identify the state of the process, control charts are widely used as a tool of the Statistical Process Control tool-kit. Control chart functionality mostly depends on a set of assumptions that may not be valid for many real-life processes. In this study, we will be applying a set of ML algorithms to identify the state of the process. The process begins by providing a number of examples (in-control and out-of-control) for training the ML models. Once the models are trained, their performance will be assessed and compared using the probability of correct detection. This study will provide an alternative to the control charts for efficient and robust process monitoring.

## Analysis of model sloppiness: what can it do, and what's next?

Matthew Adams, Gloria Monsalvo-Bravo, Sarah Vollert, Imke Botha, Christopher Drovandi

When performing statistical fitting of deterministic models to data, using either frequentist or Bayesian approaches, uncertainty in model parameters is often estimated. However, this uncertainty can have a complex, non-linear structure which has implications for how best to interpret and/or improve the model. This talk discusses "analysis of model sloppiness", a tool for unveiling this parameter uncertainty structure. Applications of this analysis include: (1) uncovering controlling mechanisms underlying the system being modelled, (2) informing which parameter measurements need to be prioritised in future experiments, (3) guiding strategic model reduction, and (4) diagnostically comparing the accuracy of different model-data fitting algorithms.

## A new mathematical modelling framework for capturing and forecasting Australia COVID-19 waves: transitioning from Delta wave into Omicron wave.

Adeshina I. Adekunle, Mingmei Teo, August Hao, Gerry Ryan, Nick Golding, Rob Moss and Peter Dawson

Covid-19 pandemic may be subsiding, but the damage caused by the SARS-CoV2 virus will take long time to amend. Many Governments adopted series of public health control measures to reduce the burden of this disease during different SARS-CoV2 strain specific waves. These measures include lockdown, isolations, quarantine, and facemask wearing. To proffer these measures in Australia, the Australian Government relied on forecasting outputs from the National Situational Assessment Team. In the work, we provide particle filter forecasts for the Delta and Omicron strain waves in Australia using an auto stochastic compartmental model. The reliability of the forecasting approach is demonstrated.

## Effect of slow-fast time scale on spatio-temporal pattern formation

Malay Banerjee

Spatiotemporal pattern formation in interacting population models is an interesting field of study as it can capture the stationary as well as dynamic patchy distribution of population within their habitat. Introduction of nonlocal interaction in the spatiotemporal model can produce stationary pattern by a spatiotemporal model with Rosenzweig-MacArthur reaction kinetics [MBVV]. On the other hand, it is evident that growth of various prey and their predators take place at different rates when measured with respect to a fixed time scale. This fact is incorporated in to the mathematical model by introducing different time-scales into the growth equations. The resulting models (with temporal reaction kinetics only), in general, exhibit two different types of oscillatory behavior, namely, canard oscillation and relaxation oscillation [PRCSPMB]. The main objective of this talk is to describe a spatiotemporal model for interacting population with nonlocal interaction term and slow-fast time scale, and discuss various scenarios of stationary and non-stationary pattern formation.

**References**

[MBVV] Banerjee, M., Volpert, V. (2017). Spatio-temporal pattern formation in Rosenzweig-MacArthur model: Effect of nonlocal interactions. *Ecological Complexity* **30**: 2–10.

[PRCSPMB] Ray Chowdhury, P., Petrovskii, S., Banerjee, M. (2021). Oscillations and pattern formation in a slow-fast prey-predator system. *Bulletin of Mathematical Biology* **83**: 110.

## Some results on the mixed bifractional Brownian motion

Soukaina Hadiri

In this work, firstly, we introduce a new gaussian process as an extension of the well known bifractional Brownian motion as a linear combination of a finite number of independent bifractional Brownian motions. We have chosen to call this process the mixed bifractional Brownian motion. Secondly, we study some stochastic properties and characteristics of this process: The Holder continuity, the self similarity, the quadratic variation, the Markov property and the differentiability of the trajectories, the long-range dependence, the stationarity of the increments and the behavior of the noise generated by the increments of this process. We believe that our process can be a possible candidate for models which involve self similarity, long range dependence and non-stationarity of increments.

## Analysis of Diffusive Size-Structured Population Model and Optimal Birth Control

Manoj Kumar

This work addresses the optimal birth control problem for invasive species in a spatial environment. We apply the method of semigroups to qualitatively analyze a size-structured population model in which individuals occupy a position in a spatial environment. With insect population in mind, we study the optimal control problem which takes fertility rate as a control variable. With the help of adjoint system, we derive optimality conditions. We obtain the optimality conditions by fixing the birth rate on three different sets. Using Ekeland's variational principle, the existence, and uniqueness of optimal birth controller to the given population model which minimizes a given cost functional is shown. A concrete example is also given to see the behaviour of population density.

## The application of mathematics to pandemics: some examples of modelling used during COVID-19

Emma McBryde, James Cook University

Mathematical models have been applied to explain infectious diseases outbreaks for over a century, but have never been taken so seriously as during the recent COVID-19 pandemic, during which they were used to synthesize evidence and inform public health action. This talk will discuss some of the models used at state level, national level and globally. It will discuss limitations and future directions for modelling.

When Australia closed its borders it did so on advice based on layered transmission and mobility models. This work suggested that by February 2020, several countries had already had cases of SARS CoV-2, without knowing it. Flight mobility suggested where the likely epidemic would spread and correctly identified changes in epicentre to Europe and later South America. I will discuss some of the modelling results that suggested this, and the importance of the early international travel restrictions.

One of the first pieces of available evidence about COVID-19 was its very specifically age-based effects, with children both less likely to acquire COVID and less likely to spread it. We used this information -along with age-specific contact matrices - to assess the potential risks and benefits of school closure. We also used age-based matrices to investigate optimal vaccine distribution at a time when vaccines were scarce. Results show that prioritising the most vulnerable (older age) was almost always a better strategy than prioritising the highest transmitters (20-30 year olds).

I will finish by discussing model refinements that are being made currently and a vision for open-science in the modelling emerging infectious diseases space.

## Model predictions of wave overwash extent into the marginal ice zone

Jordan Pitt and Luke Bennetts

Overwash is an important aspect of the dynamics in the marginal ice zone where sea ice and ocean waves interact. Overwash dissipates wave energy, and the presence of water on top of sea ice can drive growth or melting, depending on the local thermodynamic conditions. The presence of water on floes is also important for biologic and chemical processes. While overwash has been observed and investigated under experimental conditions, it has not yet been studied in the marginal ice zone. One reason for this lack of in-situ measurements and observations is due to the marginal ice zone being highly dynamic, and the onset of overwash only occurring under specific and sensitive conditions. To facilitate future observations we have produced a stochastic model of the extent of overwash into fields of sea ice by combining a new model of the onset of overwash and a standard transmission model. This combined transmission and overwash model is validated against experimental observations and is used to provide the extent of overwash for various realistic ice field and wave field conditions.

## Symmetry Analysis and Exact solutions of fractional order (2+1)-dimensional Burgers system

Komal Singla

The exact solutions of fractional order (2+1)-dimensional Burgers system are determined by using symmetry approach and power series technique. Also, the graphical behavior of obtained solutions is provided for better interpretation.

## Chemical signalling and tissue response: a moving boundary problem in biology

Yvonne Stokes

## Fish schooling

Ton Viet Ta

In this talk, we introduce our mathematical models of stochastic differential equations for fish schooling. Structural stability of models against noise is then studied numerically. Patterns obtained from the models which are consistent with real observations are presented.

## Dynamics of inertial particle focusing in curved ducts

Rahil Valani

Particles suspended in fluid flow through a curved duct can focus to stable equilibrium positions in the duct cross-section due to the balance of two dominant forces - inertial lift force and secondary drag force. Such particle focusing is exploited in various medical and industrial technologies aimed at separating particles by size. In this talk, I will present results of our numerical investigation of the dynamics of neutrally buoyant particles in fluid flow through curved ducts with rectangular cross-sections. I will show that rich bifurcations take place in the particle equilibria as a function of the duct bend radius. I will also offer insights on how these bifurcations in combination with particle dynamics can be exploited to separate particles of different sizes in circular and spiral ducts.

## A sequential method for efficiently parameterising ensemble ecosystem models

Sarah Vollert, Chris Drovandi, and Matthew Adams

Ensemble ecosystem models are valuable decision-making tools for understanding the effects of conservation actions and human impacts on threatened species. Models parameterised with dynamic systems constraints help us understand ecosystems with limited data availability. However existing methods are computationally inefficient, preventing larger networks from being studied. Using Bayesian approaches, we build on current methods to overcome this technical obstacle. Compared with the existing method, we find that using a sequential Monte Carlo approach yields similar parameter inferences and model predictions while being significantly faster. Consequently, we can study larger and more realistic networks, improving ecosystem modelling capabilities.

# Speaker Index