MATRIX Research Program:
Multivariate Dependence Modelling:
Theory and Applications

Titles and Abstracts

Elif Acar (University of Manitoba)

Title: Conditional copula survival models for incomplete data
Abstract: Incomplete data is a common feature in large biomedical databases, including among others those of phenotyping experiments and genetic studies. A fundamental challenge in these settings is to develop flexible modelling strategies for clustered censored survival data. To this end, the effectiveness of copula survival models has been studied extensively for right censored data. However, interval-censored and semi-competing risk settings are considerably less explored. In this talk, I will discuss and demonstrate the use of conditional copula models for clustered survival data under different censoring mechanisms, focusing on the cases with interval-censoring and semi-competing risks.

Mohamed Chaouch (Qatar University)

Title: Semiparametric copula-based multivariate M-estimators
Abstract: This paper introduces copula-based estimation of a class of conditional multivariate M-estimators for stationary and ergodic processes. This class encompasses, as special cases, multivariate regression, spatial median, geometric quantiles and multivariate expectiles. We investigate several asymptotic properties of the estimator. That includes uniform almost sure convergence as well as asymptotic distribution. Application to multivariate data outliers detection as well as multistep forecasting are also discussed.

Jarryd Chapman (Monash University)

Title: Modelling the Temporal Clustering of Eastern Australian Rainfall Extremes
Abstract: Extreme rainfall events can cause considerable social and economic costs to the impacted communities. These costs are heightened when extreme events cluster together in time, creating a compounding hazard that may result in a natural disaster such as flooding.

The East Coast of Australia was subject to repeated rainfall extremes and flooding events during the 2020-2022 "triple dip" La Nina, with the February 2022 event resulting in $4.8b in insurance claims and 27 deaths. This research uses a Non-Homogeneous Poisson Point Process to model the temporal clustering in the extreme arrival process and to analyse the spatial distribution of extremes clustering on Australia's East Coast, finding evidence that extreme rainfall events across the East Coast exhibit temporally clustered behaviour and that this clustering is spatially dependent. The research has applications for policy purposes, risk/disaster management and the insurance industry.

Tingjin Chu (University of Melbourne)

Title: A geostatistical analysis of metallicity variations in galaxies

Abstract: The metallicity of diffuse ionised gas (DIG) cannot be determined using strong emission line diagnostics, which are calibrated to calculate the metallicity of Hii regions. Because of this, resolved metallicity maps from integral field spectroscopy (IFS) data remain largely incomplete. In this work, we introduce the geostatistical technique of universal kriging, which allows the complete 2D metallicity distribution of a galaxy to be reconstructed from metallicities measured at Hii regions, accounting for spatial correlations between nearby data points. We apply this method to construct metallicity maps of the local spiral galaxy NGC 5236 using data from the TYPHOON/PrISM survey.

Luciana Dalla Valle (University of Plymouth)

Title: Bayesian Multivariate Nonlinear State Space Copula Models for Air Pollution

Abstract: A novel flexible class of multivariate nonlinear non-Gaussian state space models, based on copulas, is proposed. Specifically, it is assumed that the observation equation and the state equation are defined by copula families that are not necessarily equal. Inference is performed within the Bayesian framework, using the Hamiltonian Monte Carlo method. Simulation studies show that the proposed copula-based approach is extremely flexible, since it is able to describe a wide range of dependence structures and, at the same time, allows us to deal with missing data. The application to atmospheric pollutant measurement data in specific geographical regions shows that the approach is suitable for accurate modelling and prediction of data dynamics in the presence of missing values. Comparison to traditional approaches shows the superior performance of the proposed model with respect to predictive accuracy.

Gery Geenens (University of New South Wales)

Title: Universal Copulas

Abstract: The early 21st century has seen the advent of copulas as primary statistical tools when it comes to model dependence between numerical random variables. A copula is classically understood as a cumulative distribution function on the unit hypercube with standard uniform margins – we refer to such distributions as "Sklar's copulas", owing to their central role in the decomposition of multivariate distributions established by the celebrated Sklar's theorem. The argument habitually put forward for outlining the appeal of copula models is that they allow pulling apart the dependence structure of a bivariate vector, characterised by its copula, from the individual behaviour of its marginal components. Though, this interpretation can only be justified in the continuous framework, as copulas lose their "margin-free" nature outside of it, making Sklar's copula models unfit for modelling dependence between non-continuous variables. In this work, we argue that the very notion of copula should not be imprisoned into Sklar's theorem, and we propose an alternative definition of copulas which follows from approaching their role and meaning more broadly. This definition coincides with Sklar's copulas in the continuous framework, but leads to different concepts in other settings. We call this construction "universal copulas" and show that these maintain all the pleasant properties (in particular: 'margin-freeness') which make Sklar's copulas sound and effective in the continuous case. We illustrate our findings with some examples of "universal copula modelling" between two discrete variables, and between one continuous variable and one binary (Bernoulli) variable.

Bahareh Ghanbari (RMIT University)

Title: A Non-parametric estimator for high-dimensional factor copula models with the use of proxy variables

Abstract: Factor copula models have emerged as parsimonious tools for capturing dependence in observed variables based on one or multiple unobserved common factors. These models represent canonical vine copulas, integrating both observed and latent variables, thereby facilitating the representation of tail, asymmetric, and non-linear dependence structures. Most of the literature on factor copulas focuses on parametric modeling, but most of the time parametric assumptions are made solely for practical convenience. However, relying on parametric assumptions in many cases results in a misspecified model, leading to a loss of consistency in estimators. This issue is particularly critical when one or more linking copulas deviate from popular parametric copula families. Although nonparametric approaches offer a potential solution, they typically involve more intricate and computationally intensive methods. In light of this, we propose a novel

nonparametric kernel estimator for such models and compute its rate of convergence. The efficacy of our estimation method is demonstrated through an extensive simulation study.

Clara Grazian (University of Sydney)

Title: Efficient estimation of Bayesian Factor Copula Models
Abstract: We introduce innovative models for capturing the interdependence among economic variables through a Bayesian approach. These models, rooted in a latent factor framework, offer a compelling solution for high-dimensional scenarios. By integrating semiparametric marginal distributions, they yield versatile multivariate distributions. Lacking closed-form densities, we demonstrate the reliability of simulation-based estimation; we present the development of an efficient approximate Bayesian computation algorithm, enhancing computational scalability and facilitating model fitting. Our findings reveal pronounced tail dependence, diverse dependence patterns, and asymmetric behaviours, notably stronger during market downturns. Furthermore, our results showcase the superiority of Bayesian factor copula models in estimating systemic risk metrics.

Jialing Han (University of York)

Title: ARCH and GARCH-type Processes Replicated by Vt-s-vine models
Abstract: Parametric and semi-parametric method for modelling and forecasting the time series based on the vt-s-vine copula approach for financial time series developed in Bladt and McNeil (2021) is applied in this paper. By combining the v-transform to describe the stochastic volatility and s-vine copulas to capture the serial dependence with non-Gaussian margins and copulas, the models offer an improvement of the replication and prediction of ARCH and GARCH-type processes.

Arnab Hazra (Indian Institute of Technology Kanpur)

Title: Spatial wildfire risk modeling using a tree-based multivariate generalized Pareto mixture model
Abstract: Wildfires pose a severe threat to the ecosystem and economy, and risk assessment is typically based on fire danger indices such as the McArthur Forest Fire Danger Index (FFDI) used in Australia. Studying the joint tail dependence structure of high-resolution spatial FFDI data is thus crucial for estimating current and future extreme wildfire risk. However, existing likelihood-based inference approaches are computationally prohibitive in high dimensions due to the need to censor observations in the bulk of the distribution. To address this, we

construct models for spatial FFDI extremes by leveraging the sparse conditional independence structure of Hüsler–Reiss-type generalized Pareto processes defined on trees. These models allow for a simplified likelihood function that is computationally efficient. Our framework involves a mixture of tree-based multivariate generalized Pareto distributions with randomly generated tree structures, resulting in a flexible model that can capture nonstationary spatial dependence structures. We fit the model to summer FFDI data from different spatial clusters in Mainland Australia and 14 decadal windows between 1999–2022 to study local spatiotemporal variability with respect to the magnitude and extent of extreme wildfires. Our proposed method fits the margins and spatial tail dependence structure adequately and is helpful in providing extreme wildfire risk measures. Our results identify a significant increment in spatially aggregated fire risk across a substantially large portion of Mainland Australia, which indicates serious climatic concerns.

Wenqing He (Western Ontario University)

Title: Feature Screening with Large Scale and High Dimensional Censored Data
Abstract: Data with a huge size present great challenges in modelling, inferences, and computation. We will present a screening method for large-sized survival data, where the sample size is large, and the dimension of covariates is of non-polynomial order of the sample size. We rigorously establish theoretical results and conduct numerical studies to assess the performance of the proposed method. The method capitalizes on the connections among useful regression settings and offers a computationally efficient screening procedure.

Jaehong Jeong (Hanyang University)

Title: Wind vector modeling using space-time cross-covariance models
Abstract: Wind energy has significant potential for future energy portfolios without negative environmental impacts. We analyze wind vectors in East Asia from the fifth-generation ECMWF atmospheric reanalysis. To model the wind vectors, we consider non-Gaussian processes based on the Tukey g-and-h transformation, along with multivariate covariance functions. The proposed model can account for non-Gaussian features and some nonstationary dependence structures of the wind vectors. We show that the proposed method with a covariance structure reflecting the nonstationarity due to the latitude and the land-ocean differences, leads to better predictions of wind speed and wind potential, which is crucial for wind power generation planning.

Harry Joe (University of British Columbia)

Title: Vine copulas with latent variables
Abstract: Consider the case of a large number of variables where latent variables can explain the dependence in different groups of variables. Theory and algorithms are given for estimating the latent variables via proxies, and using the proxy variables in vine copula estimation.

Pavel Krupskiy (University of Melbourne)

Title: Non-central squared copulas with flexible tail dependence
Abstract: In this talk, we introduce a new class of non-central squared copulas used to model non-monotonic dependence. We study dependence properties and discuss some special cases that have more flexibility in the tails comparing to the existing squared copula models. We also discuss the inference methods for this class of models and conduct an extensive simulation study to assess the performance of the inference methods in large dimensions. Finally, we discuss some applications of this class of copulas to spatial data modelling.

Haibo Li (University of Melbourne)

Title: Subspace projection regularization for high-dimensional Bayesian linear inverse problems
Abstract: The Bayesian statistical framework provides a systematic approach to enhance the regularization model by incorporating prior information about the desired solution. For the linear Bayesian inference inverse problems with Gaussian prior, we propose a new iterative regularization algorithm based on subspace projection regularization (SPR). The SPR projects the high-dimensional problem to a series of lower-dimensional problems by iteratively constructing solution subspaces that incorporate information about the prior distribution of the underlying solution. With the new designed early stopping rules, this iterative algorithm can obtain a regularized solution with a satisfied accuracy. Several theoretical results about the algorithm are established to reveal the regularization properties of it. Both low- and high-dimensional Bayesian inference inverse problems are used to test the proposed algorithm and demonstrate its robustness and efficiency. The most computationally intensive operations in the proposed algorithm only involve matrix-vector products, making it highly efficient for high-dimensional problems.

Xiaoting Li (University of British Columbia)

Title: A Bayesian factor-vine copula model for extreme flood insurance losses

Abstract: Statistical inference on the dependence of multivariate extremes poses notable challenges, particularly in contexts characterized by large dimensions and sparse extreme observations. While multivariate copula models, notably the vine copula and the factor copula, provide flexible parametric methods for dependence modelling, caution is warranted when using them for extremal dependence inference or tail extrapolation. In this study, we introduce a novel class of factor-vine copula models. It is designed for modelling the dependence of extreme insurance losses within the context of the National Flood Insurance Program (NFIP), with broader applicability to spatio-temporal dependence modelling in multivariate time series data featuring a clustering structure. The factor-vine copula model is a variant of vine copula with a latent factor structure. It integrates the advantages of both vine and factor copulas by allowing for great flexibility in tail dependence modelling while maintaining interpretability through a parsimonious latent structure. Within the factor-vine model, we demonstrate how the incorporation of univariate extreme value margins and tail-weighted dependence measures can address the existing challenges associated with using parametric copulas for extreme inference. Furthermore, we discuss the practical applications of the proposed model in the context of the NFIP, focusing on its efficacy in evaluating the risks associated with extreme weather events.

Xueming Li (University of Melbourne)

Title: Bayesian blend of satellite precipitation estimates with rain gauge observations across Australia

Abstract: Precipitation analysis is an important component of meteorological and hydrological applications. High-quality precipitation data is essential for understanding rainfall variations over time and space, as well as developing reliable predictions across Australia. There are two primary sources of precipitation data: rain gauge station observations, known for accuracy but sparse coverage, and remote sensing data, offering broader spatio-temporal coverage but are less accurate. In this study, we propose a Bayesian method that integrates rain gauge observations with data from multiple satellite estimates, resulting a continent-wide fused precipitation dataset covering Australia. In reality, even with the introduction of an uninformative prior, the true posterior distribution is hard to obtain. To address this challenge, we employ variational Bayesian inference to find an approximate distribution that minimizes the distance from the true posterior distribution. Another significant advantage of variational Bayesian inference is its capability to handle a substantial amount of missing values by generating predicted values. This approach allows

us to effectively blend the strengths of both rain gauge and satellite data, leveraging the accuracy of the former and the broader coverage of the latter. The resulting fused precipitation dataset aims to provide a comprehensive and reliable representation of precipitation patterns across the Australian region, enhancing our ability to analyze and predict rainfall variations.

Guilherme Ludwig (University of Campinas)

Title: A sparse non-stationary spatio-temporal covariance model with advection effects for rainfall data

Abstract: We propose a sparse, non-stationary model constructed using a mixture of spatiotemporal covariance models with advection (Gupta and Waymire, 1987; Cox and Isham, 1988); namely, models that have larger covariance values along an orientation vector in the spatio-temporal index set, that simulate wind direction and cloud movement. We show that a mixture of such models can allow for wind direction change in data, and produce seasonal variation. We construct a MCMC procedure for Bayesian estimation and illustrate the problem with rainfall data from the southeastern coast of Brazil. This is a joint work with Pedro Nasevicius Ramos (UNICAMP).

Hans Manner (University of Graz)

Title: Testing the equality of changepoints

Abstract: Testing for the presence of changepoints and determining their location is a common problem in time series analysis. Applying changepoint procedures to multivariate data results in higher power and more precise location estimates, both in online and offline detection. However, this requires that all changepoints occur at the same time. We study the problem of testing the equality of changepoint locations. One approach is to treat common breaks as a common feature and test, whether an appropriate linear combination of the data can cancel the breaks. We propose how to determine such a linear combination and derive the asymptotic distribution resulting CUSUM and MOSUM statistics. We also study the power of the test under local alternatives and provide simulation results of its finite sample performance. Finally, we suggest a clustering algorithm to group variables into clusters that are co-breaking.

Bouchra Nasri (University of Montreal)

Title: Test of serial dependence for multivariate time series with arbitrary distributions

Abstract: In this talk, one studies tests of serial independence using a fixed number of consecutive observations from a stationary time series, first in the univariate case, and then in the multivariate case, where even high-dimensional vectors can be used. The common distribution function is not assumed to be continuous, and the test statistics are based on the multilinear copula process. A bootstrap procedure based on multipliers is also proposed and shown to be valid. Tests based on Spearman's rho and Kendall's tau statistics are also considered, extending the results known for the case of continuous distributions. Contiguity results are obtained for some specific models and sufficient conditions for consistency of test statistics are stated. Also, numerical experiments are performed to assess the finite sample level and power of the proposed tests. A case study using a time series of Arctic sea ice extent images is used to illustrate the usefulness of the proposed methodologies. The R package MixedIndTests includes all the methodologies proposed in this article.

Sévérien Nkurunziza (University of Windsor)

Title: On a class of tensor shrinkage estimators and recent identities

Abstract: In this talk, we present a class of tensor shrinkage estimators as well as some challenges related to the risk analysis of such estimators. We also present some recent identities which are useful in establishing the risk dominance of tensor shrinkage estimators.

Reza Nosratpour (RMIT University)

Title: Machine Learning approaches to classify and estimate precipitation

Abstract: Precipitation as a key hydrological component has a significant impact on climate, agriculture, and public safety. Hence, precise estimation of precipitation is important for hydrological and climate studies. In recent decades many projects such as the Global Precipitation Measurement Mission (GPM), Tropical Rainfall Measurement Mission (TRMM), and Climate Prediction Center MORPHing technique (CMORPH) have been developed for estimation and forecasting of precipitation with a global and quasi-global coverage using Remotely sensed data. However, these satellite datasets have systematic and random errors, they provide useful information for improving our understanding of precipitation regimes. Moreover, Machine Learning approaches have shown significant performance in modelling complex interactions of precipitation and other atmospheric variables such as

surface temperature and Precipitable Water Vapor (PWV). Multi-Layer Perceptron, Convolution Neural Networks, and Rand Forest are some of the most popular ML techniques used in different studies. By harnessing the power of machine learning (ML), precipitation models were developed that significantly improved the accuracy of satellite-based precipitation estimates. In addition, Due to the extensive spatial coverage, satellite precipitation estimates have provided researchers with an excellent opportunity to investigate various precipitation regimes. To this end, Unsupervised ML algorithms such as K-means++ have shown capabilities to classify and analysis of spatial regimes of precipitation. The results of these studies can be used as inputs for flood management models and Extreme Precipitation Early Warning systems. Ultimately, these research have significant value for water resource management, climate studies, and improvement of satellite rainfall estimates.

Karim Oualkacha (University of Quebec in Montreal)

Title: Multiple-Output Excpectile Regression via Copulas
Abstract: Expectiles summarize a distribution in a manner similar to quantiles. Expectile regression has recently gained great popularity, in part due to its attractive statistical and computational properties. Despite the renewed interest, it unfortunately remains limited to single-output problems. To enhance on this and to gain insight into multivariate data, we build on a class of multivariate expectile loss functions to develop a unified and flexible copula-based multivariate expectile regression framework. Our approach provides a new class of multiple-output expectile regression estimators, which are unique solutions to convex risk minimization problems. We model the joint distribution of the multiple-output and the regressors using a copula model, which separates modelling the dependence and the marginal distributions. Then, we rewrite the multivariate expectile regression loss function in terms of the copula and the marginal distributions. We prove the asymptotic properties of our estimators (weak convergence and i.i.d. representation). We demonstrate the effectiveness of our approach through simulation studies and by analyzing the Fourth Dutch Growth data.

Guoqi Qian (University of Melbourne)

Title: Mapping Australia's precipitation – harnessing the synergies of multi-satellite remote sensing and gauge network data

Abstract: We develop Precipitation Profiler-Observation Fusion and Estimation (PPrOFusE), a tool to deliver high-quality gauge and multi-satellite fused precipitation data. PPrOFusE consists of two steps. Step 1, the relationship among the multiple sources of data is modelled by multiple linear regression at each rain gauge location, returning the least squares estimates for the associated regression coefficient vector. Step 2, such regression coefficient vector estimates for all rain gauge locations are fitted by a spatial autoregression model, whereafter the multiple linear regression coefficient vectors for those locations void of rain gauges are predicted by spatial interpolation. PPrOFusE is applied to the precipitation system in Australia over years 2000 to 2022 and has generated interesting results.

Rachael Quill (State Electricity Commission, Victoria/Department of Energy, Environment and Climate Action)

Title: Modelling Dependence in the Renewable Energy Sector
Abstract: To meet our renewable energy targets in Australia, multi-modelling approaches are required to understand the interdependence between large-scale systems, such as optimising the production of energy from wind and solar generation along-side grid-scale battery storage. From modelling the spatial dynamics of wind fields to predicting the temporal dynamics of the energy market, a multitude of techniques are required to address problems across the sector. This talk will provide an overview of some of these, many of which require hybrid techniques, blending physical approaches with statistical modelling and machine learning.

Bruno Remillard (HEC Montreal)

Title: Copula models for time series
Abstract: In this presentation, I will present fundamental results for using copula-based models in time series. In particular, I will cover modelling, estimation, goodness-of-fit and tests of independence/randomness.

Huiyan Sang (Texas A&M University)

Title: Graph-Split-Based Decision Trees and Their Ensembles for Spatial and Network Clustering, Regression and Classification

Abstract: Decision trees and their ensembles such as CART, BART, XGBoost, and RF have become extremely popular in data science due to their exceptional performance in machine-Learning tasks such as clustering, regression, and classification. We present a novel ensemble decision tree method on graphs to handle spatial and network data with complex geometries. The new method utilizes a highly flexible graph split-based decision rule to relax the axis-parallel split rule assumption found in most existing ensemble decision tree methods. We describe a scalable recursive algorithm on graphs for model estimations. The superior performance of the proposed method is demonstrated through various clustering, regression, and classification tasks on spatial and network data.

Kate Saunders (Monash University)

Title: An extreme value post-mortem of the 2022 Australian extreme rainfall
Abstract: Extreme rainfall caused flooding in multiple Australia states in 2022; Queensland (Feb –Mar 2022), New South Wales (Mar, Jul, Oct 2022), Victoria (Oct 2022), Tasmania (Oct 2022) and the year rounded out with an ex-tropical cyclone in Western Australia (Dec –Jan 2023). In addition to the extreme and widespread nature of the flooding, some communities experienced back-to-back events. One community was flooded five times in 18 months.

In this talk, I'll undertake a post-mortem of the extreme rainfall events that impacted Australia in 2022. The goal will be to determine where we can reliably assign a probability to the recent events, and where we need concessions to make the research questions scientifically tractable. This will involve going beyond just the methods and considering the practical limitations imposed by observational records. Of particular concern in this post-mortem, is understanding whether the third consecutive La Nina contributed to the likelihood of spatially compounding and temporally compounding events in Australia, and what recommendations we could reasonably make to policy makers concerned about future extreme rainfall events.

Michael Smith (University of Melbourne)

Title: Cutting Feedback in Misspecified Copula Models
Abstract: In copula models the marginal distributions and copula function are specified separately. We treat these as two modules in a modular Bayesian inference framework, and propose conducting modified Bayesian inference by "cutting feedback". Cutting feedback limits the influence of potentially misspecified modules in posterior inference. We consider

two types of cuts. The first limits the influence of a misspecified copula on inference for the marginals, which is a Bayesian analogue of the popular Inference for Margins (IFM) estimator. The second limits the influence of misspecified marginals on inference for the copula parameters by using a rank likelihood to define the cut model. We establish that if only one of the modules is misspecified, then the appropriate cut posterior gives accurate uncertainty quantification asymptotically for the parameters in the other module. Computation of the cut posteriors is difficult, and new variational inference methods to do so are proposed. The efficacy of the new methodology is demonstrated using both simulated data and a substantive multivariate time series application from macroeconomic forecasting. In the latter, cutting feedback from misspecified marginals to a 1096 dimension copula improves posterior inference and predictive accuracy greatly, compared to conventional Bayesian inference.

Dietrich von Rosen (Swedish University of Agricultural Sciences)

Title: Safety belt regression, a mixture of classical regression and ridge regression
Abstract: Penalized estimation methods will be considered (Ridge estimation). If there is a model including parameters and an estimation function, e.g., the least squares function or likelihood function, so called penalized estimators can be obtained. This means that the estimation function is modified by adding some "fitting" term. However, in this presentation of the subject, restrictions are put on the parameters instead of the estimation function which in turn also will lead to a penalized estimation function. Indeed the two different approaches are similar but the estimators differ. In some way, from a likelihood point of view, it is more natural to put restrictions on the parameters in a model than on the estimation function. The approach is based on convex optimization theory.

Jenny Wadsworth (University of Lancaster)

Title: Statistical inference for multivariate extremes via a geometric approach
Abstract: A geometric representation for multivariate extremes, based on the shapes of scaled sample clouds in light-tailed margins and their so-called limit sets, has recently been shown to connect several existing extremal dependence concepts. However, these results are purely probabilistic, and the geometric approach itself has not been fully exploited for statistical inference. We outline a method for parametric estimation of the limit set shape, which includes a useful non/semi-parametric estimate as a pre-processing step. More fundamentally, our approach provides a new class of asymptotically-motivated

statistical models for the tails of multivariate distributions, and such models can ccommodate any combination of simultaneous or non-simultaneous extremes through appropriate parametric forms for the limit set shape. Extrapolation further into the tail of the distribution is possible via simulation from the fitted model. A simulation study confirms that our methodology is very competitive with existing approaches and can successfully allow estimation of small probabilities in regions where other methods struggle. This talk will review the basic ideas of this new framework for multivariate extremes and illustrate recent updates to the methodology.

Yan Wang (RMIT University)

Title: Joint Species Distribution Modelling

Abstract: Understanding species interactions and their responses to environmental variables is fundamental to ecology. Joint Species Distribution Modelling (JSDM) has emerged as a powerful tool for investigating the complex relationships between species in ecological communities. By considering multiple species simultaneously, JSDM provides valuable insights into community assembly, biodiversity conservation, and ecosystem functioning. These modelling approaches offer a comprehensive framework for ecologists to disentangle the intricacies of species interactions and to make more accurate predictions about community responses to changing environments. In this presentation, I will introduce the main approaches of the JSDM and their applications, highlighting their utilities in unravelling the co-occurrence patterns, community assembly, and the impacts of environmental factors on biodiversity.

Grace Y. Yi (Western Ontario University)

Title: Estimation and variable selection under the function-on-scalar linear model with covariate measurement error

Abstract: Function-on-scalar linear regression has been widely used to model the relationship between a functional response and multiple scalar covariates. Its utility is, however, challenged by the presence of measurement error, a ubiquitous feature in applications. Naively applying the usual function-on-scalar linear regression to error-contaminated data often yields biased inference results. Further, the estimation of model parameters is complicated by the presence of inactive variables, especially when handling data with a large dimension. Building parsimonious and interpretable function-on-scalar linear regression models is in urgent demand to handle error-contaminated functional data.

In this talk, I will discuss this problem and investigate measurement error effects. I will talk about procedures to simultaneously estimate functional coefficients and select salient predictors. Numerical studies will be presented to illustrate the performance of the proposed method.

Archer Yang (McGill University)

Title: Regularized Estimating Equation: Some New Perspectives
Abstract: We make some observations about the equivalences between regularized estimating equations, fixed-point problems and variational inequalities: (a) A regularized estimating equation is equivalent to a fixed-point problem, specified via the proximal operator of the corresponding penalty. (b) A regularized estimating equation is equivalent to a (generalized) variational inequality. Both equivalences extend to any estimating equations with convex penalty functions. To solve large-scale regularized estimating equations, it is worth pursuing computation by exploiting these connections. While fast computational algorithms are less developed for regularized estimating equation, there are many efficient solvers for fixed-point problems and variational inequalities. In this regard, we apply some efficient and scalable solvers which can deliver hundred-fold speed improvement. These connections can lead to further research in both computational and theoretical aspects of the regularized estimating equations. The applications in Multivariate Regularized Estimating Equations will be discussed.