14th Conference on Extreme Value Analysis, Probabilistic and Statistical Models and their Applications

EVA 2025 Chapel Hill

Conference Program
and
Book of Abstracts

Conference Program

Sunday, 22 June 2025

Short course

The short course consists of two tutorials and is organized by Miguel de Carvalho.

Statistical Modeling of Extremes

- **Anthony Davison** (EPFL)

Tutorial on Statistical Learning and Extreme Value Analysis

- **Anne Sabourin** (Université Paris Cité)

Monday, 23 June 2025

09:00 – 09:30 Conference opening

09:30 – 10:30 Plenary Lecture 1

TBA

- **Steve Sain** (Jupiter Intelligence)
  Challenges and applications associated with multivariate extremes
  (Chair: TBA)

11:15 – 12:45 Invited Session 1

IS1.1: Causality (Organizer: V. Chavez-Demoulin, Chair: P. Naveau)

TBA

- **S. Volgushev**: Inference for Hüsler-Reiss Graphical Models.
- **M. Krali**: Structural Equation Models for Multivariate Extremes.
  (Participant in "Best Student Paper Award")
- **N. Gnecco**: Extremes of Structural Causal Models.

IS1.2: Asymptotic Theory (Organizer: A. Ferreira, Chair: J. Einmahl)

TBA

- **H. Drees**: Towards a General Approach to Testing for Multivariate Regular Variation.
- **L. Henriques-Rodrigues**: Refined Extreme Risk Estimation: Generalized Means and Bias Reduction.
- **J.J. Cai**: Clustering Extreme Value Indices in Large Panels.

IS1.3: Bayesian Statistics (Organizer/Chair: L. Belzile)

TBA

- **R. Majumder**: Semiparametric Estimation of the Shape of the Limiting Bivariate Point Cloud.
- **S. Padoan**: A Bayesian Approach to Extreme Events in Time Series: Inference and Prediction via the Peaks over Threshold Method.
- **M. de Carvalho**: Tails of Random Probability Measures.

14:00 – 15:30 Invited Session 2

IS2.1: Extreme Quantile Regression (Organizer/Chair: S. Tokdar)

TBA

- **B. Youngman**: Gaussian Markov Random Field Models for High-Resolution Environmental Extremes
- **O. Pasche**: Extreme Conformal Prediction: Reliable Intervals for High-Impact Events (Participant in "Best Student Paper Award")
- **H. Lam**: Start Safe: Configuring Optimization Algorithms for Predictive Model Training and Decision-Making Under Extreme Events

IS2.2: Extremes and Risk Management in Finance and Insurance (Organizer/Chair: Z. Zhang)

TBA

- **J. Einmahl**: Tail Copula Estimation for Heteroscedastic Extremes
- **S. Aka**: Discrete Multivariate Generalized Pareto Distribution for Drought Risk Assessment
- **U. Mueller**: Time-Varying Extremes

IS2.3: Spatial Extremes (Organizers/Chairs: E. Hector and B. Reich)

TBA

- **L. Zhang**: Integrating Mechanistic Dynamics to Characterize the Causes of Extremes in Environmental Processes
- **B. Shaby**: Spatial Extremes on Domains With Physical Barriers
- **L. Kakampakou**: Geometric Modelling of Spatial Extremes

16:00 – 17:40 Contributed Session 1

CS1.1: Best Student Paper Award I, "Statistical Inference" (Chair: T. Mikosch)

TBA

- **M. Speers**: Automated Threshold Selection for Conditional Extreme Value Models
- **J. de Groot**: Extremes With Random Covariates
- **A. Mourahib**: Extremal Graphical Models With Non-Standard Extreme Directions and Conditional Independence
- **T. Wixson**: A Proxy-Likelihood Estimator for Multivariate Extremes Models With Intractable Likelihoods

CS1.2: Weather Extremes (Chair: TBA)

TBA

- **P. Busababodhin**: Ensemble Machine Learning Approaches for Parameter Estimation of the Generalized Pareto Distribution in the Chi River Basin Thailand
- **W. Choi**: A Dipole Pattern Bias in Marine Heatwave Intensity in the Kuroshio Extension Simulated by the CMIP6 Models: Poleward Shift of the Kuroshio Current
- **A. Doizé**: Stochastic Rainfall Generator With Heavy Rainfalls and Sustained Dry Spells
- **K. Saunders**: Modelling the Temporal Clustering of Extreme Rainfall Events

CS1.3: Extremes and Machine Learning (Chair: TBA)

TBA

- **B. Glowacki**: Heavy-Tail Structure in Stochastic Gradient Descent
- **S. Lhaut**: Wasserstein-Aitchison Generative Adversarial Networks for Extremes
- **P. Yang**: Stochastic Gradient Descents on Manifolds With an Application on Weighted Low-Rank Approximation
- **X. Ye**: Extreme Value Analysis for the Learning of Control-Theoretic Properties

CS1.4: Regularly Varying Stochastic Processes (Chair: TBA)

TBA

- **C. Hirsch**: Random Connection Hypergraphs
- **T. Owada**: Limit Theorems Under Heavy-Tailed Scenario in the Age Dependent Random Connection Models
- **J. Wang**: Clustering of Large Deviations Events in Heavy-Tailed Moving Average Processes
- **M. Ziegenbalg**: An Alternative Approach to Power Law Dynamics in Preferential Attachment Models

18:00 Welcome Reception

Tuesday, 24 June 2025

09:00 – 10:00 Plenary Lecture 2

TBA

- **Sidney Resnick** (Cornell University)
  Classifying Forms of Asymptotic Dependence in Bivariate Multivariate Heavy-Tailed Data
  (Chair: T. Wang)

10:45 – 12:15 Invited Session 3

IS3.1: Machine Learning (Organizer/Chair: A. Sabourin)

TBA

- **R. Huser**: Neural Methods for Amortized Inference with Multivariate and Spatial Extremes
- **M. Taheri**: Adaptive Tail Index Estimation: Minimal Assumptions and Non-Asymptotic Guarantees
  (Participant in "Best Student Paper Award")
- **G. Buritica**: Progression: A Regression Extrapolation Principle

IS3.2: Extrapolation in Regression (Organizer/Chair: M. de Carvalho)

TBA

- **V. Carcaiso**: Extrapolation of Extreme Covariates with an Application to Wildfire Prediction
- **J. Richards**: Heavy-Tailed Density Regression Using the Blended Generalised Pareto Distribution and SPxQR
- **V. Palacios Ramirez**: Heavy-Tailed Bayesian Nonparametric Regression Models

IS3.3: Geometric Methods (Organizer/Chair: J. Wadsworth)

TBA

- **N. Nolde**: Asymptotics of Light-Tailed Risk Variables with Application to Estimation of Probabilities of Risk Regions
- **I. Papastathopoulos**: Geometric Extremes from Limit Poisson Point Processes: Radial Generalized Pareto Distributions
- **C. Murphy-Barltrop**: Deep Learning of Multivariate Extremes via a Geometric Representation

13:30 – 15:10 Contributed Session 2

CS2.1: Best Student Paper Award II, "Methodology, Algorithms, and Applications" (Chair: T. Mikosch)

TBA

- **L. De Monte**: Generative Modelling of Geometric Multivariate Extremes Using Normalising Flows
- **M. W. Lee**: BLAST: A Bayesian Lasso Tail Index Regression Model with an Application to Extreme Wildfires
- **M. Shi**: Spatial Scale-Aware Tail Dependence Modeling for High-Dimensional Spatial Extremes
- **J. Jouni**: Statistical Inference and Model Selection for Models Adapted to Record Series

CS2.2: Asymptotic Statistics (Chair: TBA)

TBA

- **B. Das**: Measuring and Testing Tail Asymmetry
- **J. Fukuchi**: Bootstrap for the Vector Tail Empirical Process: Extension of Ivanoff, Kulik and Loukrati (2023)
- **J. Kim**: A Goodness of Fit Test for Distributions of Extreme Events
- **C. Rohrbeck**: Testing for Time-Varying Extremal Dependence

CS2.3: Geometric Extremes (Chair: TBA)

TBA

- **M. Corradini**: Dimension Reduction for Multivariate Geometric Extremes
- **J. Lee**: Identifying Extremal Dependence Classes Using Additive Mixtures in the Geometric Framework
- **X. Song**: Statistical Inference for Extremes of Stochastic Processes with Radial Generalized Pareto Process

CS2.4: Risks of Temporal and Spatial Dependence (Chair: TBA)

TBA

- **J. Lee**: Seasonal Trend Assessment of U.S. Extreme Precipitation via Changepoint Segmentation
- **S. Moen**: Modeling Autoregressive Conditional Regional Extremes with Applications to Solar Flare Prediction
- **W. Yue**: Statistical Modelling of Earthquake Occurrence and Extreme Magnitudes for Seismic Risk Assessment
- **Z. Zhang**: Sparse Multivariate Autoregressive Conditional Fréchet Models for High-Frequency Extreme Risk

15:40 – 17:20 Contributed Session 3

CS3.1: Quantiles and Regression (Chair: TBA)

TBA

- **H. Raubenheimer**: Extremes in Risk Management – A Nonparametric Approach to the Estimation of the Quantiles of Compound Distributions
- **G. Samorodnitsky**: Influence of Extremes on the Quantile Treatment Effect Estimation
- **L. M. Vidagbandji**: Extreme Quantile Regression Using Generalized Random Forests and Block Maxima Approach

CS3.2: Challenges in Extremal Inference (Chair: TBA)

TBA

- **M. G. Mayala**: Downsampling for Imbalanced Classification Using Infinite Centered Random Forests
- **J. Nolan**: Taming Sample Moments
- **E. Simpson**: Block Maxima Modelling in the Presence of Missing Data
- **D. Thakur**: Local LASSO: Variable Selection in High-Dimensional Spatial Regression

CS3.3: Contributed Session "Recent Advances on Stationary and Non-Stationary Time Series" (Organizer/Chair: S. Padoan)

TBA

- **D. Carl**: Likelihood-Based Inference in Stationary Time Series: The Block Maxima Method
  (Participant in "Best Student Paper Award")
- **E. Haufs**: Extreme Value Analysis Based on Blockwise Top-Two Order Statistics
  (Participant in "Best Student Paper Award")
- **R. Kulik**: Extremes for Non-Stationary Time Series

CS3.4: Spatial Extremes (Chair: TBA)

TBA

- **B. Béranger**: Fast and Efficient Inference for Flexible Spatial Extremes Models
- **K. Bratkova**: Conditional Extreme Value Models for Multivariate and Spatial Applications
- **C. Forster**: Spatio-Temporal Statistical Modeling of the Occurrence of Extreme Events
- **M. Thannheimer**: Bayesian Inference for Functional Extreme Events Defined via Partially Unobserved Processes

17:20 – 18:30 Poster Session

TBA

The list of posters is provided below in alphabetical order of presenters. Names with an asterisk indicate participants in the Best Poster Award.

- **P. Besana***: Assessing Extreme Droughts: An Extreme Value Theory Analysis of Common Drought Indices
- **L. Bocquet-Nouaille***: Control Variates for Variance-Reduced Extreme Value Index Estimators
- **P. Chakraborty***: Multivariate Regular Variation and the Calibration of p-value Combination Tests
- **K. Gasser**: Heatwave Attribution over Europe
- **A. Gimenez Zapiola***: Space Mining with Extremes
- **Y. Gong***: Scalable Causal Discovery for Extremes via Partial Tail Correlation and the PC Algorithm
- **K. Grolmusova***: Structured Multivariate Extreme Value Models for Flood Risk Estimation
- **P. Hübner***: On Predicting the Likelihood of High-Frequency Extreme Price Movements
- **Y. Kim***: Exact Coordinate Descent for High-Dimensional Regularized Robust M-Estimators
- **M. Lescart***: A Flexible Multivariate Generalized Pareto Model for Extreme Dependence Structures
- **C. Pacifici***: Statistical Prediction of Extreme Economic Losses and Fatalities in Europe due to Climate-Related Hazardous Events
- **R. Paulus***: Refining European Extreme Precipitation Return Levels using Regionalized GEV Models
- **P. Scanzi***: Bootstrap in Extreme Value Theory
- **A. Tytgat***: Multi-Regional Analysis of Antarctic Sea Ice Record Lows
- **C.-C. Wang***: Spatial Extremes on Domains with Physical Barriers
- **B. White***: Combining Observational and Model Data for Spatial Extremes through a Multivariate Gaussian Latent Process
- **F. Yu**: In-Degree Distribution of some Directed Preferential Attachment graphs

Wednesday, 25 June 2025

09:00 – 10:00 Plenary Lecture 3

TBA

- **Jennifer Wadsworth** (Lancaster University)
  Geometric extremes: from theory to high-dimensional methodology
  (Chair: TBA)

10:30 – 12:00 Invited Session 4

IS4.1: Computational and Learning Methods for Extremes (Organizer/Chair: J. Blanchet)

TBA

- **C.-H. Rhee**: From Extreme Value Analysis to Global Dynamics of SGD
- **W. Yang**: Limit Theorems for Stochastic Gradient Descent with Infinite Variance
- **J. Tan**: Estimation of Treatment Effects in Extreme and Unobserved Data
  (Participant in "Best Student Paper Award")

IS4.2: Forecasting Extremes and Rare Events (Organizer/Chair: C. Dombry)

TBA

- **J. Segers**: Tail Calibration of Probabilistic Forecasts
- **P. Naveau**: Assessment of Binary Classifiers for Asymptotically Dependent and Independent Extremes
- **R. Wickramarachchi**: From Assembly Lines to the Open Road: Predicting Rare Events in Autonomous Systems

IS4.3: Dimension Reduction (Organizer/Chair: A. Janssen)

TBA

- **A. Sabourin**: XLASSO: High-Dimensional Regression With Heavy-Tailed Predictors
- **D. Cooley**: Characterizing Tail Dependence via Pairwise Measures
- **M. Avella-Medina**: Spectral Learning of Multivariate Extremes

Excursion

Thursday, 26 June 2025

09:00 – 10:30 Invited Session 5

IS5.1: Extreme Value Analysis in Sports (Organizer/Chair: J. Einmahl)

TBA

- **X. Leng**: Has Pan Zhanle Surpassed Human Limits?
- **Y. He**: Accurate Estimates of Ultimate 100-Meter Records
- **M. Kebe**: Fastest Marathon Times Achievable Based on Extreme Value Statistics

IS5.2: Extremes in Time Series and Machine Learning Algorithms (Organizer/Chair: R. Kulik)

TBA

- **M. Ghannam**: Block Maxima Methods for Heteroskedastic, Heavy-Tailed Time Series
- **S. Rizzelli**: Inference on Marginal Expected Shortfall Under Multivariate Regular Variation
- **J. Blanchet**: Inference in Stochastic Optimization With Heavy Tails

IS5.3: Networks and Heavy Tails (Organizer/Chair: T. Wang)

TBA

- **M. Olvera-Cravioto**: Local Limits and Degree Asymptotics for a Family of Dynamic Random Digraphs
- **S. Bai**: Structural Causal Models for Extremes: A Perspective via Exponent Measure
- **D. Cirkovic**: On Tail Inference in Scale-Free Inhomogeneous Random Graphs

11:00 – 12:30 Invited Session 6

IS6.1: Extremes in Environmental Sciences (Organizer/Chair: P. Naveau)

TBA

- **W. Huang**: A Physically Assisted, Data-Driven Approach to Uncertainty Quantification of Rare Geophysical Extremes
- **J. Koh**: Extreme-Value Modelling of Migratory Bird Arrival Dates: Insights from Citizen Science Data
- **F. Ragone**: Simulation of Extreme Events and Rare Transitions in Climate Models with Rare Event Algorithms

IS6.2: Sparsity and High Dimensions (Organizer: P. Wan, Chair: TBA)

TBA

- **J. Heiny**: Testing Significant Dependencies in High Dimensions via Bootstrapping Maxima of U-Statistics
- **C. Zhou**: Clustering Tails in High Dimension
- **M. Oesting**: Regularized Weighted Score Matching Estimators for Hüsler–Reiss Graphical Models

IS6.3: Graphical Models (Organizer/Chair: S. Volgushev)

TBA

- **M. Lalancette**: Extremal Latent Tree Models
- **A. Kiriliouk**: X-Vine Models for Multivariate Extremes
- **A. Farrell**: Conditional Extremes with Graphical Models

13:45 – 15:25 Contributed Session 4

CS4.1: Best Student Paper Award III, "Multivariate Analysis" (Chair: T. Mikosch)

TBA

- **H. Tang**: On Estimation and Order Selection for Multivariate Extremes via Clustering
- **R. Campbell**: Piecewise-Linear Modeling of Multivariate Geometric Extremes
- **C. Yin**: Causal Tail Coefficient for Compound Extremes in Multivariate Time Series
- **J. Chen**: Multiple Extremal Integrals

CS4.2: Contributed Session "Optimal Prediction of Extreme Events: Some Theory and Applications" (Organizer/Chair: S. Stoev)

TBA

- **B. Bobbia**: Optimal Homogeneous Predictors of Rare Events
- **S. Stoev**: Optimal Prediction of Extreme Events: Characterizations and Examples
- **V. Verma**: Optimal Prediction of Extreme Events in Heavy-Tailed Time Series with Applications to Solar Flare Forecasting

CS4.3: Bayesian Models and Their Applications (Chair: TBA)

TBA

- **L. André**: Neural Bayes Inference for Complex Bivariate Extremal Dependence Models
- **O. Barbaux**: Non-Stationarity and Uncertainty in Design Life Level for Extreme Temperatures

- **S. Kovach**: Projecting Climate-Driven Mortality Extremes with Nonstationary Generalized Extreme Value Models
- **A. Mabe**: A Statistical Analysis of Elite Marathon Performances

CS4.4: Graphical Models and Extremal Processes (Chair: TBA)

TBA

- **F. Brück**: Graphical Models for Stable Lévy Processes Based on the Hüsler-Reiss Exponent Measure
- **H. Flury**: Asymptotic Theory for the Estimation of Brown–Resnick Processes via Composite Likelihood
- **P. Zhong**: Structured Inference for Hüsler-Reiss Graphical Models Based on Intrinsic Gaussian Markov Fields

15:55 – 17:30 Data Challenge

TBA Conference Dinner

Friday, 27 June 2025

09:00 – 10:00 Plenary Lecture 4

TBA

- **Peter Glynn** (Stanford University)
  Sequential Stopping Rules for Estimating Means of Infinite Variance Random Variables
  (chair: M. Olvera-Cravioto)

10:45 – 12:25 Contributed Session 5

CS5.1: Contributed Session "Extremes, Dynamical Systems and Multi-Fidelity" (Organizer/Chair: V. Pipiras)

TBA

- **V. Belenky**: Extremes of Ship Motions and Loads in Irregular Waves: Specific Features
- **S. Edwards**: Extreme Ship Response Estimates with Neural Networks
- **M. Kim**: Multi-Fidelity Monte Carlo Estimation with Applications to Extremes

CS5.2: Peaks Over Threshold (Chair: TBA)

TBA

- **L. Belzile**: Choosing the Threshold in Extreme Value Analysis
- **Z. Haskell-Craig**: Fitting a Peaks-Over-Threshold Model with Survey Weights, with Application to Blood Pressure Control in the US
- **T. Opitz**: Assessing the Size of Spatial Extreme Events Using Local Statistics Based on Exceedance Regions
- **I. Scheffel**: Asymptotics of Peaks-Over-Threshold Estimators in Long Memory Linear Time Series

CS5.3: Dimension Reduction (Chair: TBA)

TBA

- **K. Broadhead**: A Normed Vector Space of Equivalence Classes of Tail-Equivalent Regularly Varying Random Variables
- **P. Kuiper**: Learning Extrapolating Representations
- **F. Reinbott**: Principal Component Analysis for Max-Stable Distributions
- **M. Schiavone**: Nonparametric Estimation of the Spectral Density of a High-Dimensional Extreme Value Distribution

Book of Abstracts

Short Course Abstracts

Tutorial on Statistical Learning and Extreme Value Analysis

Anne Sabourin

Université Paris Cité
anne.sabourin@math.cnrs.fr

Abstract: In recent years, there has been a surge of theoretical and methodological advancements aimed at bridging the gap between Extreme Value Analysis (EVA) and fields such as machine learning, statistical learning, and artificial intelligence. These developments offer new perspectives both theoretically and methodologically. The goal of this tutorial is twofold:

(a) to explain the thought process and main ideas underlying statistical learning frameworks in EVA in a rather non-technical way, and

(b) to provide a deeper understanding of the proof techniques used to derive non-asymptotic guarantees.

The presentation will be structured around three main themes:

(i) Supervised Learning with Extreme Covariates: Classification and regression tasks, Empirical risk minimization, Risk decomposition

(ii) Unsupervised Learning: Dimensionality reduction, Detecting sparsity patterns, Anomaly detection

(iii) Proof Techniques: Analysis of the k-largest order statistics and exceedances above the $(1 - k/n)$-quantile, Conditioning trick, Concentration inequalities for rare events.

This tutorial is designed for researchers and practitioners interested in integrating EVA into modern statistical learning and machine learning frameworks, as well as incorporating modern statistical learning tools into EVA.

Session, Time and Place: Tutorial on Statistical Learning and Extreme Value Analysis, Sunday June 22 2025, TBA, TBA.

Statistical Modeling of Extremes

Anthony Davison

EPFL
anthony.davison@epfl.ch

Abstract: Over the past two decades Stuart Coles' book has provided an excellent and widely-used introduction to the statistics of extremes for researchers, students and practitioners from a wide variety of disciplines. The area has developed rapidly since its publication in 2001, however, and, joint with Stuart, Anthony Davison and Miguel de Carvalho have tried to update the book to reflect some modern developments, without reducing its accessibility and clarity. This workshop will give an overview of the main changes, including something of an historical overview of the development of basic statistical modelling of extremes over the past quarter-century, including in multivariate and spatial extremes, with the updated contents illustrated through hands-on computing experience.

Session, Time and Place: Tutorial on Statistical Modeling of Extremes, Sunday June 22 2025, TBA, TBA.

Plenary Session Abstracts

Challenges and applications associated with multivariate extremes

Steve Sain

Jupiter Intelligence
steve.sain@jupiterintel.com

Abstract: The traditional peril-specific characterization of extremes is instrumental in the assessment of environmental risk. However, events such as tropical cyclones, can lead to damage and loss from wind and flood, simultaneously. In this talk, an ongoing effort at Jupiter will be described that involves simulations of tropical cyclones to characterize the joint risk from wind and flood. Thinking about joint extremes at specific locations and across a large portfolio of locations raises a number of technical issues and open questions, and several of these will be highlighted. Some future directions will also be discussed.

Session, Time and Place: Plenary Lecture 1, Monday 23 June 2025, 9:30–10:30, TBA.

Classifying Forms of Asymptotic Dependence in Bivariate Multivariate Heavy-Tailed Data

Sidney Resnick

Cornell University
sir1@cornell.edu

Abstract: The ability to unambiguously classify the asymptotic dependence structure of multivariate data is often beyond the capability of graphical, exploratory tools. We present a rigorous, justifiable and practical testing framework that allows dependence structures to be categorized into four cases: (i) asymptotic independence, (ii) strong dependence, (iii) full dependence, and (iv) weak dependence. For bivariate heavy tailed data, switch to polar coordinates with the L_1 norm and these four cases are characterized respectively by the concentration of the limit angular measure on $\{0, 1\}$, a proper subset of $[0, 1]$, a single point, and the whole interval $[0, 1]$. Based on bootstrap methods we arrive at a comprehensive and theoretically justified classification tool. We have applied this tool to understanding extremal dependence between sectors of the US and Chinese economies separately and also for for analyzing extremal dependence *between* the US and Chinese economies.

Session, Time and Place: Plenary Lecture 2, Tuesday 24 June 2025, 9:00–10:00, TBA.

Geometric extremes: from theory to high-dimensional methodology

Jennifer Wadsworth

Lancaster University
j.wadsworth@lancaster.ac.uk

Abstract: The framework of multivariate regular variation has enjoyed a special status within the field of extreme value analysis, being the dependence assumption that leads to convergence of componentwise maxima to a multivariate max-stable distribution. While it is increasingly uncommon to model componentwise maxima, many statistical methods are still underpinned by multivariate regular variation. Although the assumption holds very broadly, it fails to provide techniques for extrapolation in many important cases.

A much smaller body of literature has considered the extremes of light-tailed distributions and their convergence, under appropriate scaling, on to a *limit set*. This framework of *geometric extremes* is undergoing a revival as the connections with existing dependence concepts become clear, and is now forming a basis for statistical methodology in its own right.

This talk will give an introduction to the geometric extremes framework, and provide a brief overview of new statistical methods based on this assumption. For high-dimensional modelling, we require principled simplifications of the model structure. We will introduce the concept of geometric extremal graphical models, and outline some theoretical results based on block graphs. On the practical side, we will demonstrate some initial results employing these ideas to model river flows in the northwest of England.

Based on joint work with: Ryan Campbell, Kristina Grolmusova, Natalia Nolde, Ioannis Papastathopoulos, Thordis Thorarinsdottir

Session, Time and Place: Plenary Lecture 3, Wednesday 25 June 2025, 9:00–10:00, TBA.

Sequential Stopping Rules for Estimating Means of Infinite Variance Random Variables

Peter Glynn

Stanford University
glynn@stanford.edu

Abstract: In many estimation settings, both in statistics and Monte Carlo, it is natural to collect observations until a confidence interval shrinks to a prescribed length. This problem has a long history, going back to work of Chow and Robbins in the 1960s. When the observations are iid with finite variance, the problem has a straightforward solution that is asymptotically valid as the prescribed length of the interval shrinks. However, the problem becomes challenging in settings in which the variance is difficult to estimate (e.g. autocorrelated finite variance data) or infinite variance. In this talk, we discuss a new class of sequential stopping procedures that address this issue by "canceling" out the variance or self-normalizing the data. We start by discussing sequential stopping in regimes in which the standard fixed sample size procedure results in a normal limit, and then discuss extensions to settings in which the fixed sample size limit distribution is a finite mean stable law. The sequential procedures, in the normal limit regime, have interesting connections to Bessel processes. In the infinite variance setting, we will draw upon ideas from sub-sampling. This represents joint work with Jose Blanchet, Jing Dong, and Wenhao Yang.

Session, Time and Place: Plenary Lecture 4, Friday 27 June 2025, 9:00–10:00, TBA.

Invited Session Abstracts

Structural Equation Models for Multivariate Extremes

Mario Krali

EPFL
mario.krali@epfl.ch

Co-authors: Anthony Davison

Abstract: This work attempts to fully describe joint extremal dependence in high dimensions for max-linear structural equation models supported on directed, but not necessarily acyclic, graphs. We work with the max-linear representation of such models and provide an identifiability result that establishes a connection between the parameters of the coefficient matrix and the node variables and enables applications to causal discovery and clustering. Under the framework of regularly varying and pairwise asymptotically independent innovations, we develop novel statistical machinery to estimate the extremal dependence structure based on the empirical angular measure, and show consistency for the estimated dependence parameters. We apply the proposed method to datasets of dimensions up to seventy, and see that it performs satisfactorily.

Session, Time and Place: Invited Session 1.1, Monday 23 June 2025, 11:15–12:45, TBA.

Extremes of Structural Causal Models

NICOLA GNECCO

Imperial College London
nicola.gnecco@gmail.com

Co-authors: Sebastian Engelke, Frank Röttger

Abstract: The behavior of extreme observations is well-understood for time series or spatial data, but little is known if the data generating process is a structural causal model (SCM). We study the behavior of extremes in this model class, both for the observational distribution and under extremal interventions. We show that under suitable regularity conditions on the structure functions, the extremal behavior is described by a multivariate Pareto distribution, which can be represented as a new SCM on an extremal graph. Importantly, the latter is a sub-graph of the graph in the original SCM, which means that causal links can disappear in the tails. We further introduce a directed version of extremal graphical models and show that an extremal SCM satisfies the corresponding Markov properties. Based on a new test of extremal conditional independence, we propose two algorithms for learning the extremal causal structure from data. The first is an extremal version of the PC-algorithm, and the second is a pruning algorithm that removes edges from the original graph to consistently recover the extremal graph. The methods are illustrated on river data with known causal ground truth.

Session, Time and Place: Invited Session 1.1, Monday 23 June 2025, 11:15–12:45, TBA.

Towards a General Approach to Testing for Multivariate Regular Variation

Holger Drees

University of Hamburg, Department of Mathematics
holger.drees@uni-hamburg.de

Abstract: Many statistical methods for analyzing the extreme value behavior of a sample of d-dimensional random vectors rely on the assumption that the observed vectors are multivariate regularly varying (perhaps after a marginal transformation). Despite its importance, surprisingly few statistical tests for this hypothesis have been proposed. In this talk, we discuss a general approach to tackle this problem, taking up an idea from Drees and Müller (2008). A particular challenge arises if the angular measure does not have full support.

References:

1. Drees, H., & Müller, P. (2008). Fitting and validation of a bivariate model for large claims. *Insurance: Mathematics and Economics*, 42, 638–650.

Session, Time and Place: Invited Session 1.2, Monday 23 June 2025, 11:15–12:45, TBA.

Refined Extreme Risk Estimation: Generalized Means and Bias Reduction

Lígia Henriques-Rodrigues

Universidade de Évora (UE) and Centro de Investigação em Matemática e Aplicações (CIMA), Évora, Portugal
ligiahr@uevora.pt

Co-authors: Frederico Caeiro, M. Ivette Gomes

Abstract: Extreme value theory is essential for assessing tail risk in distributions and portfolios. We focus on the conditional tail expectation (CTE), a more informative measure than Value at Risk (VaR). Assuming a Pareto tail and heavy-tailed models with a positive extreme value index (EVI), we leverage the link between EVI and CTE estimation. Using generalized means (GM) and reduced bias (RB) EVI-estimators, we derive new CTE-estimators, extending the work of Necir *et al.* (2010) and Henriques-Rodrigues *et al.* (2025). We provide the asymptotic behavior and comparison of the new CTE estimators and conduct a Monte Carlo simulation experiment to assess their finite-sample performance.

References:

1. Necir, A., Rassoul, A., & Zitikis, R. (2010). Estimating the Conditional Tail Expectation in the Case of Heavy-Tailed Losses. *Journal of Probability and Statistics*, 596839, 17 pages.
2. Henriques-Rodrigues, L., Gomes, M.I., Figueiredo, F., & Caeiro, F. (2025). A New Class of Conditional Tail Expectation Estimators. In Henriques-Rodrigues, L. *et al.* (eds)*New Frontiers in Statistics and Data Science*, Springer Proceedings in Mathematics & Statistics 469, 295–308.

Session, Time and Place: Invited Session 1.2, Monday 23 June 2025, 11:15–12:45, TBA.

Clustering Extreme Value Indices in Large Panels

Juan Juan Cai

Vrije University Amsterdam
j.cai@vu.nl

Co-authors: Chenhui Wang, Yicong Lin, Julia Schaumburg

Abstract: We analyze a large panel of units grouped by shared extreme value indices (EVIs) and aim to identify these unknown groups. To achieve this, we order the Hill estimates of individual EVIs and segment them by minimizing the total squared distance between each estimate and its corresponding group average. We show that our method consistently recovers group memberships, and we establish the asymptotic normality of the proposed group estimator. The group estimator attains a faster convergence rate than the individual Hill estimator, leading to improved estimation accuracy. Simulation results reveal that our method achieves high empirical segmentation accuracy, and the resulting group EVI estimates substantially reduce mean absolute errors compared to individual estimates. We apply the proposed method to analyze a rainfall dataset collected from 4735 stations across Europe, covering the winter seasons from January 1, 1950, to December 31, 2020, and find statistically significant evidence of an increase in the highest and a decrease in the lowest group EVI estimates, suggesting growing variability and intensification of extreme rainfall events across Europe.

Session, Time and Place: Invited Session 1.2, Monday 23 June 2025, 11:15–12:45, TBA.

Semiparametric Estimation of the Shape of the Limiting Bivariate Point Cloud

Reetam Majumder

University of Arkansas
reetamm@uark.edu

Co-authors: Benjamin A. Shaby, Brian J. Reich, Daniel S. Cooley

Abstract: We propose a model to flexibly estimate joint tail properties by exploiting the convergence of an appropriately scaled point cloud onto a compact limit set. Characteristics of the shape of the limit set correspond to key tail dependence properties. We directly model the shape of the limit set using Bézier splines, which allow flexible and parsimonious specification of shapes in two dimensions. We fit the Bézier splines to data in pseudo-polar coordinates using Markov chain Monte Carlo sampling, utilizing a limiting approximation to the conditional likelihood of the radii given angles. We propose a novel prior on the shape of the limit set via constraints on the parameters of the Bézier splines. A direct advantage of our Bayesian approach is that the support of this prior guarantees that each posterior sample is a valid limit set boundary, allowing direct posterior analysis of any quantity derived from the shape of the curve. Furthermore, we obtain interpretable inference on the asymptotic dependence class by using mixture priors with point masses on the corner of the unit box. Finally, we apply our model to bivariate datasets of extremes of variables related to fire risk and air pollution.

Session, Time and Place: Invited Session 1.3, Monday 23 June 2025, 11:15–12:45, TBA.

A Bayesian Approach to Extreme Events in Time Series: Inference and Prediction via the Peaks over Threshold Method

Simone Padoan

Bocconi University
simone.padoan@unibocconi.it

Co-authors: David Carl, Stefano Rizzelli

Abstract: In this work, we consider a (strictly) stationary time series $(X_t)_{i \geq t}$ whose one-dimensional distribution F belongs to the domain of attraction of a Generalized Extreme Value (GEV) distribution. We assume that the serial dependence between observations decays sufficiently fast as their time separation increases. Our focus is on the Peaks over Threshold (PoT) method, where exceedances above a threshold are asymptotically Generalized Pareto (GP) distributed. As a preliminary step, we examine the likelihood function of the GP model under the assumption of independent data and demonstrate that, even when applied to dependent data, it retains desirable theoretical properties for inference. In particular, the Maximum Likelihood Estimator (MLE) remains valid without efficiency loss. We then develop a Bayesian procedure for inferring the GP model parameters and estimating extreme quantiles of the underlying stationary distribution through posterior distributions. We propose an appropriate adjustment to the posterior distributions that ensures desirable theoretical properties. We investigate the its finite sample properties. Finally, we explore the Bayesian approach for predicting future extreme events through the posterior predictive distribution.

Session, Time and Place: Invited Session 1.3, Monday 23 June 2025, 11:15–12:45, TBA.

Tails of Random Probability Measures

Miguel de Carvalho

University of Edinburgh; University of Aveiro
Miguel.deCarvalho@ed.ac.uk

Abstract: While random probability measures have a long-standing tradition in probability and statistics, their tail behavior remains poorly understood. This gap is concerning and has been overlooked from the perspective of extreme value theory, as well as from that of Bayesian nonparametrics, where random probability measures are fundamental components of a wide range of models. By drawing on regular variation and subordinator theory, I will present novel results on the tails of two well-known classes of random probability measures (generalized Gamma and Pitman–Yor). Incidentally, the analysis of the problem of interest leads to a class of functions bounded by regularly varying functions with a common index. I refer to such functions as possessing *M-variation*, as they exhibit properties analogous to regular variation, including a representation theorem.

Keeping in mind the scope of the session, I will highlight the implications of the main results from both extreme value and Bayesian perspectives.

References:

1. Gil-Leyva, M. F., Palacios Ramírez, V. & de Carvalho, M. (2025). On the tails of Pitman–Yor processes. Submitted.
2. Cadena, M., Kratz, M. and Omey, E. (2017). On the order of functions at infinity, *Journal of Mathematical Analysis and Applications*, 452(1), 109–125.
3. Palacios Ramírez, V., de Carvalho, M. & Gutierrez, L. (2025, to appear). Heavy-tailed NGG-mixture models *Bayesian Analysis*.

Session, Time and Place: Invited Session 1.3, Monday 23 June 2025, 11:15–12:45, TBA.

Gaussian Markov Random Field Models for High-Resolution Environmental Extremes

Ben Youngman

University of Exeter
b.youngman@exeter.ac.uk

Co-authors: Ayu Shabrina

Abstract: The return level still remains one of the most heavily reported results of an extreme value analysis. This is especially true when we model environmental data, as we can convey risk to non-specialists in a relatively straightforward way. This study considers risk estimates for extreme rainfall in the UK based on 2.2km gridded rainfall data, which are the highest resolution UK-wide rainfall data currently available in the UK. We model these data with extreme value distributions with parameters distributed as Gaussian Markov random fields (GMRF). Our motivation is to consider the GMRF as a relatively assumption-free smoother, whereas other smoothing approaches, such as generalized additive models, typically require that we assume relationships with covariates (albeit semi-parametrically) to capture parameter variation. To fit GMRF models for extremes, we develop `evgmrf`, an end-user-friendly `R` package, which utlizes linear algebra and sparse matrices to allow fitting to large grids of data at moderate computational cost. We present risk estimates for extreme rainfall for the UK, in particular quantifying potential changes between 2000–2020, 2040–2060 and 2060–2080.

Session, Time and Place: Invited Session 2.1, Monday 23 June 2025, 14:00–15:30, TBA.

Extreme Conformal Prediction: Reliable Intervals for High-Impact Events

Olivier C. Pasche

University of Geneva
olivier.pasche@unige.ch

Co-authors: Henry Lam, Sebastian Engelke

Abstract: Conformal prediction is a popular method to construct prediction intervals for black-box machine learning models with marginal coverage guarantees. In applications with potentially high-impact events, such as flooding or financial crises, regulators often require very high confidence for such intervals. However, if the desired level of confidence is too large relative to the amount of data used for calibration, then classical conformal methods provide infinitely wide, thus, uninformative prediction intervals. In this paper, we propose a new method to overcome this limitation. We bridge extreme value statistics and conformal prediction to provide reliable and informative prediction intervals with high-confidence coverage, which can be constructed using any black-box extreme quantile regression method. The advantages of this extreme conformal prediction method are illustrated in a simulation study and in an application to flood risk forecasting.

Session, Time and Place: Invited Session 2.1, Monday 23 June 2025, 14:00–15:30, TBA.

Start Safe: Configuring Optimization Algorithms for Predictive Model Training and Decision-Making under Extreme Events

Henry Lam

Columbia University
khl2114@columbia.edu

Co-authors: Wasin Meesena

Abstract: We consider stochastic optimization where the goal is not only to optimize an average-case objective, but also mitigate the occurrence and impact of extreme catastrophic events. This problem is motivated from safety-aware decision-making and predictive model training. In particular, in the presence of a simulation model, variance reduction techniques are naturally employed to control estimation errors due to event rarity. We argue, however, that natural attempts to integrate variance reduction into optimization, even executed in a reasonable adaptive fashion, encounters fundamental challenges in terms of guaranteeing realistic runtime when using common stochastic gradient descent algorithms. On a high level, the challenge arises from the extreme sensitivity of tail-based objectives with respect to the decision variables, which renders the failure of traditional Lipschitz-based analyses. We offer remedies based on a notion of "safe initialization" that dissects algorithmic configurations that would allow for finite-time error control. We discuss implications of our findings on safe decision search and extremal predictive modeling.

Session, Time and Place: Invited Session 2.1, Monday 23 June 2025, 14:00–15:30, TBA.

Tail Copula Estimation for Heteroscedastic Extremes

John Einmahl

Tilburg University
j.h.j.einmahl@uvt.nl

Co-authors: Chen Zhou

Abstract: Consider independent multivariate random vectors that follow the same copula, but where each marginal distribution is allowed to be non-stationary. This non-stationarity is for each marginal governed by a scedasis function that is the same for all marginals. The usual rank-based estimator of the stable tail dependence function, or, when specialized to bivariate random vectors, the corresponding estimator of the tail copula, is shown to be asymptotic normal. Notably, the heteroscedastic marginals do not affect the limiting process. Next, in the bivariate setup, nonparametric tests for testing whether the scedasis functions for both marginals are the same are developed. Detailed simulations show the good performance of the estimator for the tail dependence coefficient as well as that of the new tests. In particular, novel asymptotic confidence intervals for the tail dependence coefficient are presented and their good finite-sample behavior is shown. Finally an application to the S&P500 and Dow Jones indices reveals that their scedasis functions are about equal and that they exhibit strong tail dependence.

Session, Time and Place: Invited Session 2.2, Monday 23 June 2025, 14:00–15:30, TBA.

Discrete Multivariate Generalized Pareto Distribution for Drought Risk Assessment

Samira Aka

ESSEC/Université Paris Saclay
samira.aka@square-management.com

Co-authors: Marie Kratz, Philippe Naveau

Abstract: Droughts are among the most severe and growing risks associated with climate change, with major consequences for agriculture, water resources, and financial systems. Managing these risks requires contributions from multiple disciplines, including climate sciences, probability and statistics. International organizations such as the Food and Agriculture Organization (FAO) and the United Nations Office for Disaster Risk Reduction (UNDRR) highlight the importance of proactive, risk-based approaches. As stated by the FAO, "risk-based drought management is multifaceted and requires the involvement of a variety of stakeholders" [2].

Our work (see [1]) contributes to this agenda by extending extreme value theory (EVT) to discrete multivariate settings, through the introduction of Multivariate Discrete Generalized Pareto Distributions (MDGPDs). These models bridge the gap between continuous EVT and discrete count data, offering a flexible approach to threshold exceedances for events such as dry spells. Rooted in Generalized Pareto theory [3,4], MDGPDs provide a principled framework for representing rare and compound events in a variety of applied contexts.

We detail the theoretical construction of MDGPDs, present simulation methods, and develop likelihood-free Bayesian inference techniques tailored to this discrete multivariate framework. A case study on European drought events illustrates the practical relevance of the model for climate-related risk assessment. The tools developed support decision-makers—such as insurers, policymakers, and climate risk analysts—in better understanding, anticipating, and pricing the impacts of extreme dry periods.

References:

1. Aka, S., Kratz, M., Naveau, P. (2025). Discrete multivariate generalized Pareto distribution with application to dry spells. *Preprint arXiv.*
2. Food and Agriculture Organization of the United Nations (2013). Drought risk management: A strategic approach. *FAO Land and Water Division.*
3. Hitz, A., Davis, R., Samorodnitsky, G. (2024). Discrete extremes. *Journal of Data Science*, 1–13.
4. Rootzén, H., Segers, J., Wadsworth, J. L. (2018). Multivariate peaks over thresholds models. *Extremes*, 21(1), 115–145.

Session, Time and Place: Invited Session 2.2, Monday 23 June 2025, 14:00–15:30, TBA.

Time Varying Extremes

Ulrich Mueller

Princeton
umueller@princeton.edu

Co-authors: Ulrich Mueller, Mark Watson

Abstract: Standard extreme value theory implies that the distribution of the largest observations of a large cross section is well approximated by a parametric model, governed by a location, scale and shape parameter. The extremes of a panel of independent cross sections are all governed by the same parameters as long as the underlying distribution as well as the size of the cross sections are time invariant. We derive inference about these parameters, and tests of the null hypothesis of time invariance, under asymptotics that do not require the number of extremes or the number of time periods to increase. We further apply Hamiltonian Monte Carlo techniques to estimate the path of time-varying parameters. We illustrate the approach in four examples of U.S. data: damages from weather-related disasters, financial returns, city sizes and firm sizes.

Session, Time and Place: Invited Session 2.2, Monday 23 June 2025, 14:00–15:30, TBA.

Integrating Mechanistic Dynamics to Characterize the Causes of Extremes in Environmental Processes

Likun Zhang

University of Missouri-Columbia
likun.zhang@missouri.edu

Co-authors: Megha Pailan, Christopher K. Wikle

Abstract: Clusters of extreme environmental events, such as floods and wildfires, often exhibit complex spatio-temporal patterns driven by underlying mechanistic processes. These dynamics involve interactions among multiple variables, which may act linearly or non-linearly, generating extreme behavior through both internal mechanisms (e.g., transient growth) and external forcings. To effectively characterize these complex dependencies, we introduce a novel statistical framework that integrates mechanistic dynamics with flexible spatio-temporal modeling. Our approach accommodates diverse types of forcing and captures behavior across the full range of the distribution, including typical conditions and extremes exhibiting varying asymptotic dependence structures. Using real-world data, we illustrate the capability of this framework to model event-level extremes without reliance on predefined maxima or threshold exceedances, enhancing our understanding of the causes and characteristics of extreme environmental processes.

Session, Time and Place: Invited Session 2.3, Monday 23 June 2025, 14:00–15:30, TBA.

Spatial Extremes on Domains with Physical Barriers

Benjamin Shaby

Colorado State University
bshaby@colostate.edu

Co-authors: Chien-Chung Wang

Abstract: We propose a model for spatial extremes what realistically accommodates nonconvex domains. Most existing models assume that tail dependence between locations depends only on their separation. Even if extended to be anisotropic or nonstationary, they still cannot accommodate features between points that may mediate the dependence. This assumption limits their applicability in settings where physical barriers like mountains, coastlines, or islands may attenuate or even block dependence in extremes events.

In this work, we use the idea of a visibility graph to modify an existing spatial random scale model. This results in a tail dependence model that naturally accommodates nonconvex spatial domains, where the nonconvex features are interpreted as physical barriers. Furthermore, the model inherits key properties from the scale mixture like the ability to represent asymptotic independence at long ranges and either asymptotic dependence or asymptotic independence at short ranges.

To assess the effectiveness of our model, we conduct a comprehensive study of tail dependence coefficients under various scenarios. We also evaluate parameter estimation performance, including coverage properties. Finally, we apply our framework to real-world data, demonstrating its utility in capturing spatial extremes in complicated nonconvex domains.

Session, Time and Place: Invited Session 2.3, Monday 23 June 2025, 14:00–15:30, TBA.

Geometric Modelling of Spatial Extremes

Lydia Kakampakou

Lancaster University
l.kakampakou1@lancaster.ac.uk

Co-authors: Jennifer L. Wadsworth

Abstract: Recent developments in extreme value statistics have established the so-called geometric approach as a powerful modelling tool for multivariate extremes. We adapt these tools to the case of spatial modelling and examine their efficacy at inferring extremal dependence and performing extrapolation. The geometric approach is based around a limit set described by a gauge function, which is a key target for inference. We consider a variety of spatially-parameterised gauge functions and perform inference on them by building on the framework of Wadsworth and Campbell (2024), where extreme radii are modelled via a truncated gamma distribution. We also consider spatial modelling of the angular distribution, for which we propose two candidate models. Estimation of extreme event probabilities is possible by combining draws from the radial and angular models respectively. We apply the methodology to a space weather dataset of daily geomagnetic field fluctuations.

References:

Wadsworth, J. L., & Campbell, R. (2024). Statistical inference for multivariate extremes via a geometric approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(5), 1243–1265.

Session, Time and Place: Invited Session 2.3, Monday 23 June 2025, 14:00–15:30, TBA.

Neural Methods for Amortized Inference with Multivariate and Spatial Extremes

Raphael Huser

King Abdullah University of Science and Technology (KAUST)
raphael.huser@kaust.edu.sa

Co-authors: Lidia Andre, Noel Cressie, Jordan Richards, Matthew Sainsbury-Dale, Jennifer Wadsworth, Andrew Zammit-Mangion

Abstract: Neural Bayes estimators are neural networks that approximate Bayes estimators. Once trained, these estimators are therefore extremely fast to evaluate and amenable to rapid uncertainty quantification, while also (approximately) inheriting the appealing large-sample properties of Bayes estimators. Moreover, they can be easily constructed from model simulations without computing the likelihood function. Neural Bayes estimators thus have compelling advantages when used with multivariate or spatial extremes models that have a computationally intractable likelihood function. In this talk, I will summarize our research progress in neural Bayes estimation and explain how the estimators can facilitate inference with multivariate or spatial-extremes data that may involve varying sample sizes, varying censoring levels used in peaks-over-threshold modeling, varying spatial configurations of observed locations, and/or missing data with varying missingness patterns. To conclude the talk, our user-friendly software will be presented, and real environmental applications will illustrate our proposed methodology.

Session, Time and Place: Invited Session 3.1, Tuesday 24 June 2025, 10:45–12:15, TBA.

Adaptive Tail Index Estimation: Minimal Assumptions and Non-asymptotic Guarantees

Mahsa Taheri

Research Associate, University of Hamburg
mahsa.taheri@uni-hamburg.de

Co-authors: Johannes Lederer, Anne Sabourin

Abstract: A notoriously difficult challenge in extreme value theory is the choice of the number $k \ll n$, where n is the total sample size, of extreme data points to consider for inference of tail quantities. Existing theoretical guarantees for adaptive methods typically require second-order assumptions or von Mises assumptions that are difficult to verify and often come with tuning parameters that are challenging to calibrate. This paper revisits the problem of adaptive selection of k for the Hill estimator. Our goal is not an 'optimal' k but one that is 'good enough', in the sense that we strive for non-asymptotic guarantees that might be sub-optimal but are explicit and require minimal conditions. We propose a transparent adaptive rule that does not require preliminary calibration of constants, inspired by 'adaptive validation' developed in high-dimensional statistics. A key feature of our approach is the consideration of a grid for k of size $\ll n$, which aligns with common practice among practitioners but has remained unexplored in theoretical analysis. Our rule only involves an explicit expression of a variance-type term; in particular, it does not require controlling or estimating a bias term. Our theoretical analysis is valid for all heavy-tailed distributions, specifically for all regularly varying survival functions. Furthermore, when von Mises conditions hold, our method achieves 'almost' minimax optimality with a rate of $\sqrt{\log \log n}\, n^{-|\rho|/(1+2|\rho|)}$ when the grid size is of order $\log n$, in contrast to the $(\log \log(n)/n)^{|\rho|/(1+2|\rho|)}$ rate in existing work. Our simulations show that our approach performs particularly well for ill-behaved distributions.

Session, Time and Place: Invited Session 3.1, Tuesday 24 June 2025, 10:45–12:15, TBA.

Progression: A Regression Extrapolation Principle

Gloria Buritica

AgroParisTech
gloria.buritica@agroparistech.fr

Co-authors: Sebastian Engelke

Abstract: Non-parametric and machine learning regression methods are popular because they can fit complex data during training; however, they are only reliable if new test points are well represented in training data. The problem of regression extrapolation, or out-of-distribution generalization, arises when predictions are required at test points beyond the training data range. In such cases, the non-parametric guarantees for regression methods from both statistics and machine learning typically fail. Based on the theory of tail dependence, we propose a novel statistical extrapolation principle. After a suitable, data-adaptive marginal transformation, our principle assumes the relationship between predictors and the response simplifies at the boundary of the training predictor samples. This assumption holds for a wide range of models, including the additive noise models for a broad family of non-parametric regression functions. Our semi-parametric method: progression, leverages this extrapolation principle and offers guarantees on the approximation error beyond the training data range. We demonstrate how this principle can be effectively integrated with existing approaches, such as random forests and additive models, to improve extrapolation performance on out-of-distribution samples.

1. Buriticá, G. & Engelke, S. (2024). Progression: an extrapolation principle for regression. *arXiv:2410.23246*.

Session, Time and Place: Invited Session 3.1, Tuesday 24 June 2025, 10:45–12:15, TBA.

Extrapolation of Extreme Covariates with an Application to Wildfire Prediction

Viviana Carcaiso

INRAE, Avignon, France
viviana.carcaiso@inrae.fr

Co-authors: Thomas Opitz, Sebastian Engelke, Juliette Legrand

Abstract: We propose methods to enhance the predictive performance of generalized additive models (GAMs) in the context of covariate extrapolation, where predictions rely on covariates beyond their observed range. When using predictive models such GAMs, shifts in the covariate distribution between training and prediction datasets can occur. Ignoring this issue may lead to inaccurate predictions near the extremes of the covariate distributions. For example, this problem is particularly critical in climate-change scenarios, where covariates simulated from climate models are likely to contain more extreme conditions. Our approach integrates GAMs for the bulk of covariate distributions with methods for modeling conditional extremes at high covariate values. Here, we consider binary responses and assume that an event occurs if an unobserved latent continuous variable exceeds a certain threshold. For large values of the covariates, this latent variable is assumed to depend linearly on the covariates with an additive error term, for which the distribution is characterized by the link function of the regression model. In an application to occurrence data of extreme wildfires in Europe, we explore how the new method can improve predictions, using weather-based fire danger indices and other environmental and meteorological variables.

Session, Time and Place: Invited Session 3.2, Tuesday 24 June 2025, 10:45–12:15, TBA.

Heavy-tailed Density Regression Using The Blended Generalised Pareto Distribution and Spxqr

Jordan Richards

University of Edinburgh
jordan.richards@ed.ac.uk

Co-authors: Reetam Majumder

Abstract: Semi-parametric quantile regression (SPQR) is a flexible and computationally-efficient approach to regression that exploits neural networks to learn a spline-based representation of conditional density functions. As it makes no parametric assumptions about the underlying density function, SPQR performs incredibly well for in-sample testing and interpolation. However, it can perform poorly when faced with heavy-tailed data and when asked to extrapolate beyond the range of observations, as it fails to satisfy any of the asymptotic guarantees of extreme value theory (EVT).

To build semi-parametric density regression models that can be used for reliable tail extrapolation, we create the blended generalised Pareto (GP) distribution, which i) provides a model for the entire range of data and, via a smooth and continuous transition, ii) benefits from exact GP upper-tails without the need for threshold selection. We combine SPQR with the blended GP to create extremal semi-parametric quantile regression (xSPQR), which provides a flexible semi-parametric approach to heavy-tailed density regression that is compliant with traditional EVT. We handle the interpretability of our approach through the use of model-agnostic variable importance scores, which provide the relative importance of covariates for determining the bulk and tails of the conditional density function. The efficacy of xSPQR is illustrated on simulated and real data examples.

Session, Time and Place: Invited Session 3.2, Tuesday 24 June 2025, 10:45–12:15, TBA.

Heavy-tailed Bayesian Nonparametric Regression Models

Karla Vianey Palacios Ramirez

Newcastle University
nkvp1@ncl.ac.uk

Co-authors: Miguel de Cavalho, Luis Gutierrez

Abstract: In this talk, I will present new methodological advances in modeling heavy-tailed data within a regression framework using Bayesian nonparametric (BNP) mixture models. This work pioneers the development of priors over the space of distributions with heavy-tailed margins. In order to set a prior in this space, it is crucial to define a random probability measure that can preserve heavy tails. I will first characterize such a measure, focusing on the normalized generalized gamma (NGG) process, and show that it maintains heavy tails when the centering distribution is heavy-tailed. Building on this foundation, I will introduce different BNP mixture models with predictor-dependent structures, allowing us to learn about the multivariate behavior of a distribution with heavy-tailed marginals via covariates and enabling flexible tail index regression modeling. I will compare two classes of heavy-tailed BNP mixture models, highlighting the distinction between scale and shape mixtures in capturing heavy-tailed behavior. I will exemplify by modeling how stimuli influence heavy-tailed alpha and beta brainwaves through covariates, using data from a neuroscience application.

References:

1. Palacios Ramírez, V., de Carvalho, M. & Gutiérrez, L. "Heavy-Tailed NGG-Mixture Model" *Bayesian Anal. Advance Publication 1 - 29, 2024.*
2. Miguel de Carvalho, Vianey Palacios R., Lígia Rodrigues & Myung Won Lee (2025) "Regression Models for Extreme Events" In: Handbook of Statistics of Extremes. *To appear*

Session, Time and Place: Invited Session 3.2, Tuesday 24 June 2025, 10:45–12:15, TBA.

Asymptotics of Light-tailed Risk Variables with Application to Estimation of Probabilities of Risk Regions

Natalia Nolde

University of British Columbia
natalia@stat.ubc.ca

Co-authors: Guus Balkema

Abstract: For random vectors with light-tailed densities whose level sets asymptotically have the same shape, we derive explicit asymptotic expressions for tail probabilities and moments of conditional excesses over a limiting threshold for homogeneous functionals of the coordinates, which we refer to as risk variables. These expressions depend on the shape of the level sets at an extreme point as well as that of the risk region induced by a given risk variable. On the basis of these asymptotic results, we construct estimators of probabilities of risk regions and conditional excesses and assess their performance in finite sample situations through a simulation study.

Session, Time and Place: Invited Session 3.3, Tuesday 24 June 2025, 10:45–12:15, TBA.

Geometric Extremes From Limit Poisson Point Processes: Radial Generalized Pareto Distributions

Ioannis Papastathopoulos

University of Edinburgh
i.papastathopoulos@ed.ac.uk

Co-authors: Lambert De Monte, Ryan Campbel, Haavard Rue

Abstract: We introduce a novel class of limit multivariate distributions, termed radial generalized Pareto distributions, which emerge as non-degenerate limits of radially recentered and rescaled exceedances above direction-dependent thresholds. This framework leverages the recently introduced framework of geometric extremes and a novel convergence to Poisson point processes, providing an overarching stochastic foundation for constructing multivariate distributions that allow extrapolation of risk in any direction within multivariate spaces. Our approach naturally leads to the notion of quantile and return sets which closely parallel related notions of optimal-transport based quantile regions, while also allowing for the constrution of isotropic return sets that are exceeded with equal probability along any direction, offering an improved understanding of extreme event risk. We develop a fully Bayesian inference framework for these multivariate distributions, using latent Gaussian processes, and construct novel diagnostics, based on properties of rotationally invariant random point measures, for assessing the convergence to the limit distribution. Our methods are applied to real-world data from hydrology and oceanography, demonstrating their broad applicability in risk analysis and their potential to inform decision-making about yet unobserved extreme events.

Session, Time and Place: Invited Session 3.3, Tuesday 24 June 2025, 10:45–12:15, TBA.

Deep Learning of Multivariate Extremes Via A Geometric Representation

Callum Murphy-Barltrop

TU Dresden
callum.murphy-barltrop@tu-dresden.de

Co-authors: Reetam Majumder, Jordan Richards

Abstract: Geometric extremes is becoming an increasingly popular modelling tool for extremal dependence. Recent work from Nolde and Wadsworth (2022) demonstrated that the limiting shapes of scaled sample clouds and their so-called *limit sets* are connected to a wide range of representations for multivariate extremes. Consequently, many recent works have introduced limit set estimators, with the resulting estimates being used for inference on the joint tail. However, these existing approaches are limited to low-dimensional settings and tend to require strong modelling assumptions.

In this talk, we introduce DeepGauge - the first deep learning approach for limit set estimation. By leveraging the predictive power and computational scalability of neural networks, we construct asymptotically justified yet highly flexible semi-parametric models for extremal dependence. Unlike existing techniques, DeepGauge can be applied in high-dimensional settings and requires few assumptions. We showcase the efficacy of our deep approach by modelling the complex extremal dependence between metocean variables sampled from the North Sea. These results demonstrate that deep learning can assist with the study of extremal dependence, opening new avenues for multivariate risk assessments.

References:

1. Nolde, N., & Wadsworth, J. L. (2022). Linking representations for multivariate extremes via a limit set. *Advances in Applied Probability*, 54(3), 688-717.

Session, Time and Place: Invited Session 3.3, Tuesday 24 June 2025, 10:45–12:15, TBA.

From Extreme Value Analysis to Global Dynamics of Sgd

Chang-Han Rhee

Northwestern University
chang-han.rhee@northwestern.edu

Co-authors: Xingyu Wang, Sewoong Oh

Abstract: In this talk, I will present a mathematical framework that enables precise characterization and control of the global dynamics of Stochastic Gradient Descent (SGD) within the complex non-convex loss landscapes typical in deep learning. These developments build on heavy-tailed large deviations formulation and a local stability analysis framework we recently introduced. Together, they give rise to heavy-tailed counterparts of the classical Freidlin-Wentzell and Eyring-Kramers theories. Our machinery elucidates how to manipulate heavy-tailed noises during training so that SGD avoids sharp minima almost completely, thereby achieving better generalization performance on test data.

Session, Time and Place: Invited Session 4.1, Wednesday 25 June 2025, 10:30–12:00, TBA.

Limit Theorems for Stochastic Gradient Descent with Infinite Variance

Wenhao Yang

Stanford University
yangwh@stanford.edu

Co-authors: Jose Blanchet, Aleksandar Mijatović

Abstract: Stochastic gradient descent algorithm is a classic algorithm that gains significant popularity from both empirical and theoretical perspectives. While its probability propositions are well-studied when the randomness is assumed to have a finite variance, there is a scarcity of research addressing its theoretical propositions in the case of infinite variance. In this paper, we establish the asymptotic behavior of stochastic gradient descent algorithm when the stochastic gradient has an infinite variance, specifically assuming the stochastic gradient is regular varying with index $\alpha \in (1, 2)$. The most recent limit theorems in this context were established in 1969 [1], but they only consider one-dimension case and assume the stochastic noises belongs to a restrictive class. We extend it into multi-dimension and a more general case, covering a broader class of infinite variance distributions. This extension requires entirely different techniques, as the original method does not apply to the multi-dimensional case. Our results indicate that the asymptotic distribution of the stochastic gradient descent algorithm aligns with the Ornstein-Unlenbeck process driven by an additive process instead of a Brownian motion process. Additionally, we explore the applications of these results in linear regression and logistic regression models.

References:

1. Tatiana Pavlovna Krasulina (1969). On stochastic approximation processes with infinite variance. *Theory of Probability & Its Applications, 14(3):522–526, 1969.Smith, J. (2021).*

Session, Time and Place: Invited Session 4.1, Wednesday 25 June 2025, 10:30–12:00, TBA.

Estimation of Treatment Effects in Extreme and Unobserved Data

Jiyuan Tan

Stanford University
jiyuantan@stanford.edu

Co-authors: Jose Blanchet, Vasilis Syrgkanis

Abstract: Causal effect estimation seeks to determine the impact of an intervention from observational data. However, the existing causal inference literature primarily addresses treatment effects on frequently occurring events. But what if we are interested in estimating the effects of a policy intervention whose benefits, while potentially important, can only be observed and measured in rare yet impactful events, such as severe wildfires, hurricanes, and tsunamis? The standard causal inference methodology is not designed for this type of inference since the events of interest may be scarce in the observational data. Extreme Value Theory (EVT) provides methodologies for analyzing these rare occurrences. We introduce a novel framework for assessing treatment effects in extreme data to capture the causal effect at the occurrence of rare events of interest. We employ the theory of multivariate regular variation to model extremities. We develop a consistent estimator for extreme treatment effects and present a rigorous non-asymptotic analysis of its performance, exploring the minimax optimality of our estimator under a suitable semi-parametric class. We illustrate the performance of our estimator using both synthetic and semi-synthetic data.

Session, Time and Place: Invited Session 4.1, Wednesday 25 June 2025, 10:30–12:00, TBA.

Tail Calibration of Probabilistic Forecasts

Johan Segers

KU Leuven
jjjsegers@kuleuven.be

Co-authors: Sam Allen, Jonathan Koh, Johanna Ziegel

Abstract: Probabilistic forecasts comprehensively describe the uncertainty in the unknown future outcome, making them essential for decision making and risk management. While several methods have been introduced to evaluate probabilistic forecasts, existing evaluation techniques are ill-suited to the evaluation of tail properties of such forecasts. However, these tail properties are often of particular interest to forecast users due to the severe impacts caused by extreme outcomes. In this work, we introduce a general notion of tail calibration for probabilistic forecasts, which allows forecasters to assess the reliability of their predictions for extreme outcomes. We study the relationships between tail calibration and standard notions of forecast calibration, and discuss connections to peaks-over-threshold models in extreme value theory. Diagnostic tools are introduced and applied in a case study on European precipitation forecasts.

Session, Time and Place: Invited Session 4.2, Wednesday 25 June 2025, 10:30–12:00, TBA.

Assessment of Binary Classifiers for Asymptotically Dependent and Independent Extremes

Philippe Naveau

CNRS (LSCE)
philippe.naveau@lsce.ipsl.fr

Co-authors: Juliette Legrand and Marco Oesting

Abstract: Machine learning classification methods usually assume that all possible classes are sufficiently present within the training set. Due to their inherent rarities, extreme events are always under-represented and classifiers tailored for predicting extremes need to be carefully designed to handle this under-representation. In this paper, we address the question of how to assess and compare classifiers with respect to their capacity to capture extreme occurrences. This is also related to the topic of scoring rules used in forecasting literature. In this context, we propose and study a risk function adapted to extremal classifiers. The inferential properties of our empirical risk estimator are derived under the framework of multivariate regular variation and hidden regular variation. A simulation study compares different classifiers and indicates their performance with respect to our risk function. To conclude, we apply our framework to the analysis of extreme river discharges in the Danube river basin. The application compares different predictive algorithms and test their capacity at forecasting river discharges from other river stations.

Session, Time and Place: Invited Session 4.2, Wednesday 25 June 2025, 10:30–12:00, TBA.

From Assembly Lines to The Open Road: Predicting Rare Events in Autonomous Systems.

Ruwan Wickramarachchi

AI Institute, University of South Carolina
ruwan@email.sc.edu

Abstract: In the age of embodied AI and smart automation, autonomous agents are increasingly deployed in high-stakes, real-world environments. Ensuring the robustness and resilience of these systems in the face of rare but critical failure events is essential for their safe and reliable operation. Accurate forecasting of such rare events is particularly crucial, as a single overlooked anomaly can lead to catastrophic consequences. In manufacturing, for instance, unplanned downtime due to rare failures costs industries over $50 billion annually, with sectors like automotive losing more than $2 million per hour—even with preventive maintenance systems in place.

However, the extreme rarity and complexity of these events pose significant challenges for AI methods. The scarcity of high-quality data, methodological gaps in the literature, and limited practical experience with multimodal signals further complicate rare event prediction and detection.

In this talk, I explore two real-world domains critical for autonomous systems. First, I focus on smart manufacturing/ Industry 4.0, presenting the complete lifecycle of rare event prediction – ranging from analog and multimodal dataset development to addressing data scarcity, improving data quality, and building and evaluating robust predictive models.

In the second part, I introduce the problem of unobserved entity prediction in autonomous driving, which involves identifying potentially missing entities due to perceptual failures in edge-case scenarios. I discuss how contextual and structured knowledge can be leveraged to develop Neurosymbolic AI solutions, validated across multiple real-world driving datasets.

I will share the datasets and resources we have developed and made publicly available for both cases to support and encourage further research in rare event prediction for autonomous systems.

Session, Time and Place: Invited Session 4.2, Wednesday 25 June 2025, 10:30–12:00, TBA.

XLASSO: High-dimensional Regression with Heavy-tailed Predictors

Anne Sabourin

Université Paris Cité, MAP5
anne.sabourin@u-paris.fr

Co-authors: Stephan Clémençon

Abstract: The prediction of extreme covariates has only recently been explored from a theoretical perspective within multivariate Extreme Value Theory (EVT). The challenge lies in constructing a prediction function that performs well in the tail regions of the covariates. This problem has been formalized relatively recently within a statistical learning framework, initially in an idealized context where the class of predictors is of low complexity, requiring no penalization.

In this work, I present an extension of this framework to high-dimensional settings, utilizing an ℓ_1-penalty and taking as input an appropriately rescaled version of the covariates. The proposed algorithm, XLASSO, offers non-asymptotic guarantees for prediction, ensuring consistency. Additionally, I will discuss how this theoretical framework can be applied to predict unbounded, heavy-tailed targets.

References:

1. Clémençon, S., & Sabourin, A. (2025). Weak Signals and Heavy Tails: Machine-learning meets Extreme Value Theory. arXiv preprint arXiv:2504.06984.

Session, Time and Place: Invited Session 4.3, Wednesday 25 June 2025, 10:30–12:00, TBA.

Characterizing Tail Dependence Via Pairwise Measures

Dan Cooley

Colorado State University
cooleyd@rams.colostate.edu

Co-authors: Kenneth Broadhead

Abstract: In high dimensions, fully characterizing the extremal dependence structure (via the angular measure or similar structure) is difficult. However, summarizing extremal dependence via pairwise measures, whose estimation is straightforward, can provide actionable dependence information for modeling. Assuming the framework of multivariate regular variation, this talk will begin by reviewing the tail pairwise dependence matrix, and show that when $\alpha = 2$ the TPDM has the properties of an inner product. In earlier work, a vector space was constructed via linear combinations of regularly varying random vectors, and combined with the inner product provided by the TPDM, we performed linear prediction at extreme levels.

In ongoing work, we are working to generalize the vector space beyond the aforementioned linear construction. Given a normalizing function $b(s)$, we have developed a normed vector space of equivalence classes of regularly varying random variables . This vector space will be the subject of Kenneth Broadhead's talk at EVA2025. Collections from this vector space cannot be assumed to be jointly regularly varying, and thus, dependence between vectors cannot be described by the TPDM as previously defined. We will present what we have learned about extending the vector space to be an inner product space.

Session, Time and Place: Invited Session 4.3, Wednesday 25 June 2025, 10:30–12:00, TBA.

Spectral Learning of Multivariate Extremes

Marco Avella Medina

Columbia University
marco.avella@columbia.edu

Co-authors: Richard Davis, Gennady Samorodnitsky

Abstract: We propose a spectral clustering algorithm for analyzing the dependence structure of multivariate extremes. More specifically, we focus on the asymptotic dependence of multivariate extremes characterized by the angular or spectral measure in extreme value theory. Our work studies the theoretical performance of spectral clustering based on a random k-nearest neighbor graph constructed from an extremal sample, i.e., the angular part of random vectors for which the radius exceeds a large threshold. In particular, we derive the asymptotic distribution of extremes arising from a linear factor model and prove that, under certain conditions, spectral clustering can consistently identify the clusters of extremes arising in this model. Leveraging this result we propose a simple consistent estimation strategy for learning the angular measure. Our theoretical findings are complemented with numerical experiments illustrating the finite sample performance of our methods.

Session, Time and Place: Invited Session 4.3, Wednesday 25 June 2025, 10:30–12:00, TBA.

Has Pan Zhanle Surpassed Human Limits?

Xuan Leng

Xiamen University
xleng@xmu.edu.cn

Co-authors: Yi He, HuifenXie

Abstract: We employ extreme value inference methods to assess whether Pan Zhanle's world-record time of 46.40 seconds in the men's 100-meter freestyle at the 2024 Paris Olympics exceeds the human limit. Our dataset comprises swimming records from 5,973 top male athletes spanning the years 1924 to 2025, with approximately three observations per athlete. We define the ultimate world record as the minimum of the left endpoints of the individual time distributions across all top swimmers. Based on Hill estimator for heterogeneous data, the ultimate record is estimated at 45.9seconds. We construct an asymptotic 95% lower confidence bound of 45.3 seconds, which is faster than the current world record of 46.40 seconds, indicating that Pan Zhanle's achievement remains within the statistically inferred human capability.

Session, Time and Place: Invited Session 5.1, Thursday 26 June 2025, 09:00–10:30, TBA.

Accurate Estimates of Ultimate 100-meter Records

Yi He

University of Amsterdam
y.he2@uva.nl

Co-authors: John H.J. Einmahl

Abstract: We employ the novel theory of heterogeneous extreme value statistics to accurately estimate the ultimate world records for the 100-m running race, for men and for women. For this aim we collected data from 1991 through 2023 from thousands of top athletes, using multiple fast times per athlete. We consider the left endpoint of the probability distribution of the running times of a top athlete and define the ultimate world record as the minimum, over all top athletes, of all these endpoints. For men we estimate the ultimate world record to be 9.56 seconds. More prudently, employing this heterogeneous extreme value theory we construct an accurate asymptotic 95% lower confidence bound on the ultimate world record of 9.49 seconds, still quite close to the present world record of 9.58. For the women's 100-meter dash our point estimate of the ultimate world record is 10.34 seconds, somewhat lower than the world record of 10.49. The more prudent 95% lower confidence bound on the women's ultimate world record is 10.20.

Session, Time and Place: Invited Session 5.1, Thursday 26 June 2025, 09:00–10:30, TBA.

Fastest Marathon Times Achievable Based on Extreme Value Statistics

Malick Kebe

Howard University
malick.kebe@bison.howard.edu

Co-authors: Saraleesan Nadarajah

Abstract: This study applies extreme value statistics to predict the fastest achievable marathon times for ten major marathons worldwide (Olympic, New York, Boston, Amsterdam, Berlin, Paris, Tokyo, Dubai, Rome, and Toronto) for both men and women. Using historical winning times and the generalized extreme value (GEV) distribution, we model the theoretical minimum marathon times achievable at each venue, focusing on the potential of each marathon course rather than individual runner performance. Our analysis reveals distinct patterns in fastest achievable times across different marathons, with Tokyo showing the greatest potential for men's record times and Paris for women after 2020. The study identifies significant differences in time trends between venues, with some marathons showing linear decreases in fastest achievable times while others exhibit non-linear patterns. We find no evidence of common gender or course effects across marathons, suggesting that each venue has unique characteristics affecting performance potential. The models demonstrate good fit with observed data and successfully predict record times within estimated bounds.

Session, Time and Place: Invited Session 5.1, Thursday 26 June 2025, 09:00–10:30, TBA.

Block Maxima Methods for Heteroskedastic, Heavy-tailed Time Series

Mai Ghannam

University of Toronto
mai.ghannam@utoronto.ca

Co-authors: Rafał Kulik, Stanislav Volgushev

Abstract: Extreme value theory (EVT) provides theoretical and methodological tools to study the probability and intensity of rare events. While the impact of heteroskedasticity in the PoT framework has recently been studied, there is no literature addressing BM methods under non-stationarity. In this work, we present the first theoretical analysis of BM methods under heteroskedasticity. We show that the classical block maxima maximum likelihood estimator (MLE) for the tail index is asymptotically inconsistent in the presence of heteroskedasticity, even when the underlying tail index is time-invariant. To address this, we propose two novel estimators and prove their consistency and asymptotic normality. We also introduce inference procedures to test for heteroskedasticity in extreme value data in the context of BM methods. Our findings broaden the applicability of BM methods to non-stationary settings and offer new tools for reliable tail inference in practice.

Session, Time and Place: Invited Session 5.2, Thursday 26 June 2025, 09:00–10:30, TBA.

Inference on Marginal Expected Shortfall Under Multivariate Regular Variation

Stefano Rizzelli

University of Padova, Department of Statistical Science
stefano.rizzelli@unipd.it

Co-authors: Simone Padoan, Matteo Schiavone

Abstract: Expected shortfall is arguably one of the most popular measures of risk and an extensive literature is devoted to its study, encompassing several univariate time series models and long- or short-range dependence setups. In multivariate contexts, marginal expected shortfall is a key measure of systemic risk, and the study of its extreme behaviour is particularly relevant for mitigating the impact of severe downturns in global financial markets. In this context, statistical inference is typically based on bivariate extreme-value models (tail copulas) for a given financial variable of interest and another that incorporates information on financial system's risk. In this work, we show that when multivariate regular variation can be plausibly assumed, it allows to explicitly account for complex extremal dependence structures among a large number of financial institutions, of which the market is composed, enabling more refined statistical modelling and inference. Specifically, we derive an approximation formula for the extreme marginal expected shortfall and derive an estimator, of which we also propose a bias-corrected version. In a general beta-mixing context, that accommodates popular time series models with heavy-tailed innovations, we establish proposed estimators' contraction rates and prove their asymptotic normality, which in turn allows the derivation of confidence intervals. A simulation study shows that the new estimators significantly improve upon the performance of the existing ones. This is a joint work with Simone Padoan (Bocconi University) and Matteo Schiavone (University of Padova) and will be presented in the session "Extremes in Time Series and Machine Learning Algorithms".

Session, Time and Place: Invited Session 5.2, Thursday 26 June 2025, 09:00–10:30, TBA.

Inference in Stochastic Optimization with Heavy Tails

Jose Blanchet

Stanford University
jose.blanchet@stanford.edu

Co-authors: Peter Glynn, Joost Jorritsma, Aleks Mijatovic, Wenhao Yang, Bert Zwart,

Abstract: In this talk, we study various inference problems arising in stochastic optimization with heavy-tailed input data. Our focus is on how the use of extreme value theory for heavy tailed random variables can be leveraged in various settings. Including, for example, performance guarantees for stochastic gradient descent with infinite variance noise or sharp approximations for chance constrained optimization problems. If time permits we will also discuss the development of optimal regularization in machine learning guided by distributionally robust optimization for infinite variance heavy tailed models.

References:

1. Blanchet, J., Mijatovic, A., Yang, W. (2024). *Limit Theorems for Stochastic Gradient Descent with Infinite Variance.* https://arxiv.org/abs/2410.16340
2. Blanchet, J., Jorritsma, J., Zwart, B. (2024) *Optimization Under Rare Events: Scaling Laws for Linear Chance-Constrained Programs.* https://arxiv.org/abs/2407.11825

Session, Time and Place: Invited Session 5.2, Thursday 26 June 2025, 09:00–10:30, TBA.

Local Limits and Degree Asymptotics for A Family of Dynamic Random Digraphs

Mariana Olvera-Cravioto

University of North Carolina at Chapel Hill
molvera@email.unc.edu

Co-authors: Yu, Fuwei

Abstract: In this talk we consider a dynamic directed random graph model for networks where individuals arrive in families/groups and attach to the network according to a general attachment function. Specifically, the the graph is obtained by collapsing a continuous-time branching process into random-sized families of nodes. The attachment function in the branching process is assumed to be regularly varying, giving rise to a wide range of possible in-degree distributions. The out-degree distribution is determined by the family sizes. We describe the local weak limit for this family of models, as well as provide exact asymptotics for the resulting degree distributions, which remarkably, can be determined by either the family sizes only, the attachment mechanism only, or a combination of both, depending on the relative heavy-tailedness of the attachment function and the family size distribution.

Session, Time and Place: Invited Session 5.3, Thursday 26 June 2025, 09:00–10:30, TBA.

Structural Causal Models for Extremes: A Perspective Via Exponent Measure

Shuyang Bai

University of Georgia, US
bsy9142@uga.edu

Co-authors: Fei Fang, Tiandong Wang

Abstract: Structural causal models, a recursive system of stochastic equations organized by a directed acyclic graph, play important roles in the analysis of causality. The appropriation of them to extreme value analysis has been an active topic of research.

In this talk, we introduce a new formulation of structural causal models tailored for extremes, which we refer to as extremal structural causal models (eSCMs). In contrast to the conventional structural causal model, whose randomness is governed by a probabilistic law, we propose a formulation of structural equations whose "randomness" is governed by an exponent measure, an infinite-mass law that naturally emerges in the analysis of multivariate extremes.

Key ingredients of this approach include the introduction of activation variables, which can be viewed as an abstraction of the single-big-jump principle, as well as a further randomization that enriches the laws of eSCMs. Such a formulation turns out to cover all possible laws of directed graphical models based on the recently introduced notion of extremal conditional independence.

We further identify a type of asymmetry inherent to eSCMs under natural assumptions, which facilitates identifiability of causal directions — a key challenge in causal inference. Finally, we propose a method to exploit this causal asymmetry and validate its effectiveness through both simulations and real benchmark datasets.

Session, Time and Place: Invited Session 5.3, Thursday 26 June 2025, 09:00–10:30, TBA.

On Tail Inference in Scale-free Inhomogeneous Random Graphs

Daniel Cirkovic

Texas A& M University
cirkovd@stat.tamu.edu

Co-authors: Tiandong Wang, Daren Cline

Abstract: Both empirical and theoretical investigations of scale-free network models have found that large degrees in a network exert an outsized impact on its structure. However, the tools used to infer the tail behavior of degree distributions in scale-free networks often lack a strong theoretical foundation. In this paper, we introduce a new framework for analyzing the asymptotic distribution of estimators for degree tail indices in scale-free inhomogeneous random graphs. Our framework leverages the relationship between the large weights and large degrees of Norros-Reittu and Chung-Lu random graphs. In particular, we determine a rate for the number of nodes $k(n) \to \infty$ such that for all $i = 1, \ldots, k(n)$, the node with the i-th largest weight will have the i-th largest degree with high probability. Such alignment of upper-order statistics is then employed to establish the asymptotic normality of three different tail index estimators based on the upper degrees. These results suggest potential applications of the framework to threshold selection and goodness-of-fit testing in scale-free networks, issues that have long challenged the network science community.

Session, Time and Place: Invited Session 5.3, Thursday 26 June 2025, 09:00–10:30, TBA.

A Physically Assisted, Data-driven Approach to Uncertainty Quantification of Rare Geophysical Extremes

Whitney Huang

Clemson University
wkhuang@clemson.edu

Co-authors: Katherine Kreuser

Abstract: This work aims to develop a rigorous and coherent statistical framework to improve risk assessment methodologies for rare geophysical events, particularly in cases where direct observations of extreme events are scarce and must be supplemented by computer simulations. The framework addresses three key tasks: (1) estimating the joint distribution of input variables that characterize rare events; (2) developing efficient statistical emulators to generate large synthetic input-output datasets, enabling forward propagation to approximate output distributions; and (3) comprehensively quantifying uncertainties by identifying multiple sources and assessing their impacts on downstream extreme value analysis. The proposed framework will be applied to the study of storm surges and volcanic eruptions, demonstrating its practical value in evaluating the risks associated with geophysical extremes.

Session, Time and Place: Invited Session 6.1, Thursday 26 June 2025, 11:00–12:30, TBA.

Extreme-value Modelling of Migratory Bird Arrival Dates: Insights From Citizen Science Data

Jonathan Koh

ETH Zurich
jonathan.koh@stat.math.ethz.ch

Co-authors: Thomas Opitz

Abstract: Citizen science mobilises many observers and gathers huge datasets but often without strict sampling protocols, resulting in observation biases due to heterogeneous sampling effort, which can lead to biased predictions. We develop a spatiotemporal Bayesian hierarchical model for bias-corrected estimation of arrival dates of the first migratory bird individuals at their breeding sites. Higher sampling effort could be correlated with earlier observed dates. We implement data fusion of two citizen-science datasets with fundamentally different protocols (BBS, eBird) and obtain posterior distributions of the latent process, which contains four spatial components endowed with Gaussian process priors: species niche; sampling effort; position and scale parameters of annual first arrival date. The data layer consists of four response variables: counts of observed eBird locations (Poisson); presence-absence at observed eBird locations (Binomial); BBS occurrence counts (Poisson); first arrival dates (Generalised Extreme-Value). We devise a Markov Chain Monte Carlo scheme and check by simulation that the latent process components are identifiable. We apply our model to several migratory bird species in the northeastern United States for 2001–2021 and find that the sampling effort significantly modulates the observed first arrival dates. We exploit this relationship to effectively bias-correct predictions of the true first arrivals.

Session, Time and Place: Invited Session 6.1, Thursday 26 June 2025, 11:00–12:30, TBA.

Simulation of Extreme Events in Numerical Models with Rare Event Algorithms

Francesco Ragone

University of Leicester
fr120@leicester.ac.uk

Abstract: The analysis of extreme events is an important area of application of numerical models in many different scientific fields. Studying these events on a robust statistical basis with complex numerical models is however computationally challenging, as very long simulations and/or very large ensembles are necessary to sample a sufficient number of events to have acceptable levels of statistical accuracy. This problem can be tackled using rare event algorithms, numerical tools designed to reduce the computational effort required to sample rare events in numerical models. These methods typically take the form of genetic algorithms, where a set of suppression and cloning rules are applied to the members of an ensemble simulation, in order to oversample trajectories leading to the events of interest. In this talk I will show recent applications of these methods to different classes of events, focusing in particular on extremes of surface temperature and sea ice cover. Finally I will discuss the relevance of these techniques for different applications, in particular for the analysis of tipping points and the validation of early warning indicators.

Session, Time and Place: Invited Session 6.1, Thursday 26 June 2025, 11:00–12:30, TBA.

Testing Significant Dependencies in High Dimensions Via Bootstrapping Maxima of U-statistics

Johannes Heiny

Stockholm University
johannes.heiny@math.su.se

Co-authors: Patrick Bastian, Holger Dette

Abstract: In this talk, we take a different look on the problem of testing the mutual independence of the components of a high-dimensional vector. Instead of testing if all pairwise associations (e.g. all pairwise Kendall's tau) between the components vanish, we are interested in the (null)-hypothesis that all pairwise associations do not exceed a certain threshold in absolute value. The consideration of these hypotheses is motivated by the observation that in the high-dimensional regime, it is rare, and perhaps impossible, to have a null hypothesis that can be exactly modelled by assuming that all pairwise associations are precisely equal to zero.

The formulation of the null hypothesis as a composite hypothesis makes the problem of constructing tests non-standard and in this talk we provide a solution for a broad class of dependence measures, which can be estimated by U-statistics. In particular we develop an asymptotic and a bootstrap level alpha-test for the new hypotheses in the high-dimensional regime. We also prove that the new tests are minimax-optimal and investigate their finite sample properties by means of a small simulation study and a data example.

The talk is based on joint work with Patrick Bastian and Holger Dette.

Session, Time and Place: Invited Session 6.2, Thursday 26 June 2025, 11:00–12:30, TBA.

Clustering Tails in High Dimension

Chen Zhou

Erasmus University Rotterdam
zhou@ese.eur.nl

Co-authors: Liujun Chen, Marco Oesting

Abstract: One potential solution to combat the scarcity of tail observations in extreme value analysis is to integrate information from multiple datasets sharing similar tail properties, for instance, a common extreme value index. In other words, for a multivariate dataset, we intend to group dimensions into clusters first, before applying any pooling techniques. This paper addresses the clustering problem for a high dimensional dataset, according to their extreme value indices.

We propose an iterative clustering procedure that sequentially partitions the variables into groups, ordered from the heaviest-tailed to the lightest-tailed distributions. At each step, our method identifies and extracts a group of variables that share the highest extreme value index among the remaining ones. This approach differs fundamentally from conventional clustering methods such as using pre-estimated extreme value indices in a two-step clustering algorithm.

We show the consistency property of the proposed algorithm and demonstrate its finite-sample performance using a simulation study and a real data application.

Session, Time and Place: Invited Session 6.2, Thursday 26 June 2025, 11:00–12:30, TBA.

Regularized Weighted Score Matching Estimators for Hüsler-reiss Graphical Models

Marco Oesting

University of Stuttgart
marco.oesting@mathematik.uni-stuttgart.de

Co-authors: Johannes Lederer

Abstract: Hüsler–Reiss distributions, often regarded as the natural analogue to Gaussian distributions in the context of extremes, have gained considerable popularity in recent years. Much like in Gaussian graphical models, their associated precision matrix encodes conditional independencies, making sparsity in this matrix both meaningful and desirable. In this talk, we explore weighted score matching estimators for the entries of the precision matrix, incorporating an ℓ_1 penalty to promote sparsity. We present both computational and theoretical insights into the estimator, including finite-sample guarantees and asymptotic properties such as sparsistency.

Session, Time and Place: Invited Session 6.2, Thursday 26 June 2025, 11:00–12:30, TBA.

Extremal Latent Tree Models

Michaël Lalancette

Université du Québec à Montréal
lalancette.michael@uqam.ca

Co-authors: Galiane Charbonneau

Abstract: Latent tree models are latent variable statistical models in which the joint distribution of the observed and unobserved variables is a tree graphical model. They can therefore be seen as certain multivariate marginals of tree graphical models, but this class nevertheless includes complex dependence structures such as one-factor, hidden Markov, and several phylogenetic models.

This work introduces a fully nonparametric class of multivariate Pareto distributions termed *extremal latent tree models*. While these distributions are not, strictly speaking, latent tree models, it is argued that they share a similar interpretation and are appropriate models for the extremal dependence of thresholded multivariate data in the presence of latent variables. A characterization of the class of extremal latent tree models is obtained, positioning it within the wider realm of extremal graphical models. Under standard assumptions on the underlying graph topology, leveraging the theory of phylogenetic trees and the work of Engelke & Volgushev (2022), it is shown that the number of latent variables and the latent tree structure can be learned through a generalized neighbor-joining algorithm. If time allows, exact and approximate nonparametric inference for extremal latent tree models will be discussed.

Reference:

1. Engelke, S. & Volgushev, S. (2022). Structure learning for extremal tree models. *J. R. Stat. Soc. Ser. B.*, 84(5), 2055-2087.

Session, Time and Place: Invited Session 6.3, Thursday 26 June 2025, 11:00–12:30, TBA.

X-vine Models for Multivariate Extremes

Anna Kiriliouk

University of Namur
annakiriliouk@gmail.com

Co-authors: Johan Segers, Jeongjin Lee, Lídia André

Abstract: Regular vine sequences permit the organization of variables in a random vector along a sequence of trees. Vine-based dependence models have become greatly popular as a way to combine arbitrary bivariate copulas into higher-dimensional ones, offering flexibility, parsimony, and tractability. We use regular vine sequences to decompose and construct exponent measure densities associated with multivariate extreme value distributions. Our approach sheds new light on existing parametric extreme-value models and facilitates the construction of new ones, called X-vines. Computations proceed via recursive formulas in terms of bivariate model components. We develop simulation algorithms for X-vine multivariate Pareto distributions as well as methods for parameter estimation and model selection on the basis of threshold excesses. The methods are illustrated by Monte Carlo experiments and a case study on US flight delay data. Finally, we discuss how X-vines can be used for quantile regression in extreme regions of the covariate space, allowing for asymptotic independence between the response variable and the covariates.

References:

1. Kiriliouk, A., Lee, J., & Segers, J. (2024). X-vine models for multivariate extremes. *Journal of the Royal Statistical Society Series B: Statistical Methodology, qkae105.*

Session, Time and Place: Invited Session 6.3, Thursday 26 June 2025, 11:00–12:30, TBA.

Conditional Extremes with Graphical Models

Aiden Farrell

Lancaster University
a.farrell1@lancaster.ac.uk

Co-authors: Emma F. Eastoe, Clement Lee

Abstract: Utilising graphical models for multivariate extreme value analysis is currently restricted to asymptotically dependent random variables only. The theory has been extended to asymptotically independent random variables, but no statistical methodology has been developed to accompany it. To fill this gap, we extend the conditional multivariate extreme value model with a new multivariate residual distribution that allows the dependence structure to be represented by *any* undirected graph. In addition, the model can be used to learn the graphical structure when it is unknown *a priori*. To support inference in high dimensions, we propose a stepwise inference procedure that is computationally efficient and loses no information or predictive power. We show that our method is flexible and accurately captures the extremal dependence for river discharges in the upper Danube River basin.

Session, Time and Place: Invited Session 6.3, Thursday 26 June 2025, 11:00–12:30, TBA.

Contributed Session Abstracts

Automated Threshold Selection for Conditional Extreme Value Models

Matthew Speers

Lancaster University
m.speers@lancaster.ac.uk

Co-authors: Jonathan Tawn

Abstract: The conditional extreme value model of (Heffernan & Tawn, 2004) provides a framework capable of modelling all multivariate extremal dependence types. For a chosen conditioning variable, it describes the joint behaviour of all variables given this conditioning variable exceeds a threshold. Whilst this approach has been widely applied and iterated on, there exist no methods for the automatic selection of the conditioning threshold value. We develop a metric which captures the bias in a conditional extremes model fitted above a given threshold. Studying this metric for a range of candidate thresholds allows for informed, semi-automatic threshold selection. We also extend the metric by applying a weighted variance penalty. After tuning, this formulation allows for fully-automatic selection of the conditioning threshold through minimisation of the variance-weighted metric. We demonstrate the performance of our method on synthetic datasets which represent a range of extremal dependence structures, comparing to random threshold threshold choices and the method of (Wan & Davis, 2019), who develop a threshold selection under multivariate regularisation.

Session, Time and Place: Contributed Session 1.1, Monday 23 June 2025, 16:00–17:40, TBA.

Extremes with Random Covariates

Just de Groot

Erasmus University Rotterdam
justdegroot@outlook.com

Co-authors: Chen Zhou

Abstract: Classical extreme value statistics assumes that observations are independent and identically distributed. In this case, the tail distribution of the focal variable is approximated by a Generalized Pareto Distribution (GPD). In applications where the tail distributions may vary according to covariates, the parameters of the GPD are often modeled by parametric functions of the covariates. In this study, we introduce the *proportional tail model*, which extends the heteroskedastic extremes model in Einmahl et al. (2016) by incorporating multiple random covariates into the tail distribution based on generic parametric models. We provide theoretical justifications for a likelihood based estimation of such parametric models.

Einmahl et al. (2016) introduces the scedasis function to model the scale variation in heavy-tailed distributions with respect to a single deterministic covariate, time. Mefleh et al (2020) models the scedasis as a parametric function of time. By contrast, the *proportional tail model* allows for negative extreme value index with endpoints depending on multiple random covariates. We show asymptotic properties for the maximum likelihood estimator of the extreme value index and the parameters in a generic family of parametric scedasis functions.

References:

1. Einmahl, J., Haan L., & Zhou, C. (2016). Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society Series B*, Volume 78(1), 31-51.
2. Mefleh, A., Biard, R., Dombry, C., & Khraibani, Z. (2020). Trend detection for heteroscedastic extremes. *Extremes*, Volume 23, 85-115.

Session, Time and Place: Contributed Session 1.1, Monday 23 June 2025, 16:00–17:40, TBA.

Extremal Graphical Models with Non-standard Extreme Directions and Conditional Independence

Anas Mourahib

UCLouvain
anas.mourahib@uclouvain.be

Co-authors: Sebastian Engelke, Johan Segers

Abstract: Consider a random vector representing risk factors. Extremal graphical models provide a framework for encoding conditional independence among these risks in extreme scenarios. Existing models, such as the Hüsler–Reiss graphical model, primarily address situations where all risks become large simultaneously. However, this assumption is unrealistic when some risks exhibit near independence. To overcome this limitation, we introduce a novel approach to constructing extremal graphical models that accommodates scenarios where some risks are large while others are not. As an example, we propose the mixture Hüsler–Reiss graphical model, which generalizes the existing Hüsler–Reiss graphical model.

1. Mourahib, Anas, Anna Kiriliouk, and Johan Segers. "Multivariate generalized Pareto distributions along extreme directions." Extremes (2024): 1-34.
2. Engelke, Sebastian, and Adrien S. Hitz. "Graphical models for extremes." Journal of the Royal Statistical Society Series B: Statistical Methodology 82.4 (2020): 871-932.

Session, Time and Place: Contributed Session 1.1, Monday 23 June 2025, 16:00–17:40, TBA.

A Proxy-likelihood Estimator for Multivariate Extremes Models with Intractable Likelihoods

Troy Wixson

Colorado State University
twixson@rams.colostate.edu

Co-authors: Daniel Cooley

Abstract: Many multivariate extremes models have intractable likelihoods and thus practitioners must use alternative fitting methods. Furthermore, likelihood-based methods for uncertainty quantification and model selection are unavailable. We develop a proxy-likelihood estimator for multivariate extremes models. Our method is based on the tail pairwise dependence which is a summary measure of the dependence in the tail of any multivariate extremes model. We employ the Hüsler-Reiss distribution as a proxy for the likelihood in a composite likelihood approach. Our desired model is linked to the proxy model (the Hüsler-Reiss) through the one-to-one relationship between the tail pairwise dependence parameter and the dependence parameter of the proxy model. The method, including model selection, is demonstrated using the transformed linear extremes time series models of Mhatre and Cooley (2024).

References:

1. Mhatre N. and Cooley D. (2024). Transformed-linear models for time series extremes. *Journal of Time Series Analysis.* 45(5):671–690.

Session, Time and Place: Contributed Session 1.1, Monday 23 June 2025, 16:00–17:40, TBA.

Ensemble Machine Learning Approaches for Parameter Estimation of the Generalized Pareto Distribution in the Chi River Basin Thailand

Piyapatr Busababodhin

Mahasarakham University
piyapatr.b@msu.ac.th

Co-authors: Tossapol Phoophiwfa, YingYing Song

Abstract: Accurate parameter estimation is central to extreme value analysis, particularly for hydrological modeling and flood risk assessment. This study presents an adaptive strategy for estimating parameters of the Generalized Pareto Distribution (GPD) using ensemble machine learning techniques, including Artificial Neural Networks (ANN), AdaBoost, and XGBoost. The objective is to enhance estimation accuracy under both stationary and non-stationary conditions.

The study focuses on the Chi River Basin in Northeast Thailand, utilizing meteorological and environmental data—such as the Normalized Difference Vegetation Index (NDVI), rainfall, runoff, and climate variables—collected from satellite sources and 92 meteorological stations between 2010 and 2024. Correlation-based feature selection was applied to identify significant predictors for the GPD's location and scale parameters.

Stationary models were estimated using maximum likelihood, while non-stationary conditions were addressed through ensemble learning applied to monthly maximum rainfall. Model performance was evaluated via the Nash–Sutcliffe Efficiency (NSE) coefficient. Results demonstrate that non-stationary models outperformed stationary ones in 82 of 92 stations, with NSE values ranging from 0.6 to 0.9, indicating high predictive accuracy. ANN-based models revealed that NDVI, geographic coordinates, and meteorological variables significantly influence the GPD's shape parameter.

Spatial maps of return levels for 2-, 5-, 10-, 20-, 50-, and 100-year periods were produced, offering actionable insights for flood risk mitigation and climate adaptation. These findings highlight the utility of ensemble machine learning in advancing extreme value modeling in flood-prone regions.

Session, Time and Place: Contributed Session 1.2, Monday 23 June 2025, 16:00–17:40, TBA.

A Dipole Pattern Bias in Marine Heatwave Intensity in The Kuroshio Extension Simulated by the CMIP6 Models: Poleward Shift of The Kuroshio Current

Wonkeun Choi

Korea Institute of Ocean Science and Technology
dnjsrms@kiost.ac.kr

Co-authors: Heeseok Jung, Minkyoung Bang, Zhenya Song, Namyoung Kang, Chan Joo Jang

Abstract: Marine heatwaves (MHWs) are extreme climate events that substantially increase ocean temperatures, leading to significant ecological and socio-economic consequences. Climate models are essential tools for investigating and projecting MHWs, but their inherent biases pose challenges for accurate projection. Understanding the characteristics of these biases and identifying their underlying causes are crucial for improving MHW simulations and enhancing model reliability. In this study, we apply a statistical approach to evaluate and classify bias patterns in the mean MHW intensity across 30 CMIP6 climate models over the North Pacific Ocean. To analyze the underlying causes of the biases, hierarchical clustering is performed based on the spatial correlation of mean MHW intensity biases in the Kuroshio Extension region. Additionally, ocean current simulations were assessed by identifying high-velocity grid points and quantifying their spatial displacement using a cubic polynomial regression. Our results reveal that over 80% of the models show a dipole-pattern bias in the mean MHW intensity, characterized by overestimation to the north and underestimation to the south of the Kuroshio Extension. This pattern is closely associated with the northward overshooting of the Kuroshio Current, which may be partly attributed to the coarse resolution of climate models. These findings suggest that improving the representation of ocean circulation, particularly in the western boundary current region including the Kuroshio region, is key to reducing biases in MHW projections.

Session, Time and Place: Contributed Session 1.2, Monday 23 June 2025, 16:00–17:40, TBA.

Stochastic Rainfall Generator with Heavy Rainfalls and Sustained Dry Spells

Antoine Doizé

Sorbonne-Université
antoine.doize@sorbonne-universite.fr

Co-authors: Denis Allard, Philippe Naveau, Olivier Wintenberger

Abstract: In light of increasing climate changes and its profound impacts, it is critical to accurately assess phenomena such as heavy rainfall and sustained dry or rainy spells. These events have severe consequences, including infrastructure damage or agricultural failures. Climate models have become an essential tool for projecting these variables in future scenarios. However, their high computational cost and resolution limitations raise significant challenges.

Stochastic weather generators, are a cost-effective means of simulating climate trajectories over long time scales. They find pattern in historical data and reproduce them. However, if emulating regular patterns is not necessarily a challenge, these generators must be designed carefully to also extrapolate and produce realistic extreme rainfall intensity, sustained dry spells and rainfall episodes. This aspect will be core in our study.

This work presents a daily rainfall generator producing realistic scenarios of rainfall and dry spells. For this we develop and study a Markov based model that both integrates mathematical waiting time representation of discrete distributions, and extreme value analysis. The use of the extended Generalized Pareto distribution let us model rainfall intensity and occurrence both in the bulk and the tail of the distribution. We study the sensitivity of our inference scheme throughout various numerical experiments and apply our model to different French weather stations with different climate characteristics.

Session, Time and Place: Contributed Session 1.2, Monday 23 June 2025, 16:00–17:40, TBA.

Modelling the Temporal Clustering of Extreme Rainfall Events

Kate Saunders

Monash University
kate.saunders@monash.edu

Co-authors: Jarryd Chapman, Michael Lydeamore, Kim Reid

Abstract: Extremes rainfall events that occur sequentially risk compounding in their impacts, magnifying social and economic costs compared to a single event. To effectively mitigate this compound risk, we need to understand the probability of multiple, independent extremes occurring in a window of time, such as week or a season. Yet much of the existing research does not devote sufficient attention to the statistics of the arrivals process. Instead, a constant rate or arrivals is often assumed as default. Here we develop a model for the arrivals process of extreme events. We use a non-homogeneous Poisson process. The risk of compound events can then be estimated over different windows of time easily and with a single statistical model. We demonstrate our approach using the east coast of Australia as a case study, motivated by the temporal clustering that occurred during 2022 floods. We fit models of increasing complexity to compare the influence short-term weather to long-term atmospheric drivers on the arrival process of extreme events. We find that short-term drivers are the leading factor influencing the arrival of extreme events, but that long-term atmospheric drivers like ENSO, can increase baseline risk during the La Nina phase. This methodology contributes a flexible modelling approach that can be adapted to other types of weather and climate extremes, and in different regions. Importantly this work shows to understand climate risk from extreme rainfall we need models of the arrival process, and this work contributes the first of such models for Australia.

References:

1. Chapman, J., Lydeamore, M., Reid, K. and Saunders, K.R (2025) Temporal Clustering of Extreme Rainfall Events *In preparation*.

Session, Time and Place: Contributed Session 1.2, Monday 23 June 2025, 16:00–17:40, TBA.

Heavy-Tail Structure in Stochastic Gradient Descent

Bartosz Glowacki

University of Ottawa
bglowack@uottawa.ca

Abstract: Certain machine learning (ML) algorithms are often regarded as black-box procedures: they tend to perform well empirically, yet their theoretical foundations remain largely unexplored. Their probabilistic and statistical properties are a terra incognita—with no comprehensive theoretical framework or established methodology currently available. Notably, extreme events, frequently driven by heavy-tailed behavior, arise naturally in various ML algorithms and may even play a beneficial role. For instance, in the standard Stochastic Gradient Descent (SGD) optimization method, heavy tails have been observed empirically and are believed to help the algorithm "escape" undesirable solutions. There have been efforts to support these empirical findings through mathematical theory. Given that many ML algorithms exhibit a stochastic recurrence structure, there exists a natural connection to the established theory of extreme value theory (EVT) for time series and Markov chains, which applies to online SGD algorithms. However, offline algorithms do not retain the Markov chain structure, rendering the existing theory inapplicable. Even the fundamental question regarding the tail behavior of outputs from offline SGD remains unresolved. Temporal dependence, clustering of extremes, and the statistical and limiting properties of heavy-tailed SGD iterates are still open questions, despite being well-studied in the EVT literature for time series and stochastic processes. In this presentation, used tools from multivariate regular variation for time series, we present theoretical results on heavy tailed behaviour of offline Stochastic Gradient Descent. The tail behaviour is influenced by a batch size and learning rates.

Session, Time and Place: Contributed Session 1.3, Monday 23 June 2025, 16:00–17:40, TBA.

Wasserstein-Aitchison Generative Adversarial Networks for Extremes

Stéphane Lhaut

UCLouvain
stephane.lhaut@uclouvain.be

Co-authors: Holger Rootzén, Johan Segers

Abstract: Modeling of the tail dependence of a random vector $\mathbf{X}$ in $\mathbf{R}^d$ is known to be of crucial importance in many applications such as climatology, finance and risk assessment. Traditional approaches rely on parametric statistical models, while recent methods leverage generative models—such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs)—which are well-suited for high-dimensional distributions. However, few, if any, of these methods efficiently incorporate the multivariate max-stable structure inherent to the problem.

In this work, we propose a Wasserstein-GAN (WGAN) architecture designed to generate samples of the angular measure Φ with respect to the L_1-norm of a marginally standardized version of $\mathbf{X}$. To facilitate learning, we draw on compositional data analysis to transform the angular components of large observations in the dataset into coordinates in an orthonormal basis of $\Delta_{d-1} = \{\mathbf{x} \in [0,1]^d : \|\mathbf{x}\|_1 = 1\}$, the support of Φ, before their entrance in the WGAN. This transformation simplifies the distribution, improving the learning process and mitigating challenges posed by the typically small sample size of real-world datasets. A simple post-processing step then converts these angles into samples from the Multivariate Generalized Pareto (MGP) distribution of $\mathbf{X}$, which turns out to be very useful in estimating tail probabilities related to $\mathbf{X}$. We demonstrate the method's performance on both simulated and real datasets.

Session, Time and Place: Contributed Session 1.3, Monday 23 June 2025, 16:00–17:40, TBA.

Stochastic Gradient Descents on Manifolds with an Application on Weighted Low-rank Approximation

PEIQI YANG

George Washington University
pqyang@gwu.edu

Co-authors: Conglong Xu, Hao Wu

Abstract: We prove a convergence theorem for stochastic gradient descents on manifolds with both adaptive and deterministic learning rate, and apply it to the weighted low-rank approximation problem.

Session, Time and Place: Contributed Session 1.3, Monday 23 June 2025, 16:00–17:40, TBA.

Extreme Value Analysis for The Learning of Control-theoretic Properties

Xiuzhen Ye

North Carolina State University
xye7@ncsu.edu

Co-authors: Xiuzhen Ye, Wentao Tang

Abstract: Statistical learning underpins recent advances in data-driven methods for nonlinear systems. To provide stability and performance guarantees, main interest for learning lies on properties related to system-wide behaviors (in the form of Lyapunov-like inequalities) [1]. For systems under thermodynamic laws, *dissipative properties* are sought for, which refers to the behavior that a storage function whose increment *never exceeds* a supply rate function dependent on system inputs and outputs [2]. While dissipativity learning can be performed using typical learning algorithms [3, 4], the corresponding learning theory (based on concentration inequalities) only certifies dissipativity for "the majority of observations" instead of as a strict system-wide behavior. Therefore, we adopt extreme value theory (EVT) to investigate the distribution of the dissipation rate of the system. The minimum dissipation of the system serves as a bound that establishes the achievable stability properties for control. To this end, a generalized extreme value distribution with an estimated extreme value index is used for describing the tail behavior. This work is an exemplification of using EVT as a generic tool for the machine learning of system behaviors and physical laws in science and engineering.

1. Sontag (1998). *Mathematical control theory.* Springer.
2. Brogliato et al. (2007). *Dissipative systems analysis and control.* Springer.
3. Tang & Daoutidis (2019). *Comput. Chem. Eng.*, 130, 106576.
4. Tang & Daoutidis (2021). *Syst. Control Lett.*, 147, 104831.
5. LoCicero et al. (2024). *arXiv:2411.13404.*

Session, Time and Place: Contributed Session 1.3, Monday 23 June 2025, 16:00–17:40, TBA.

Random Connection Hypergraphs

Christian Hirsch

Aarhus University
hirsch@math.au.dk

Co-authors: Morten Brun, Peter Juhász, Moritz Otto

Abstract: In this talk, I will describe a model for random hypergraphs based on weighted random connection models. In accordance with the standard theory for hypergraphs, this model is constructed from a bipartite graph. In this stochastic model, both vertex sets of this bipartite graph form marked Poisson point processes and the connection radius is inversely proportional to a product of suitable powers of the marks. Hence, the model is a common generalization of weighted random connection models and AB random geometric graphs.

For this hypergraph model, I will discuss the large-window asymptotics of graph-theoretic and topological characteristics such as higher-order degree distribution, Betti numbers of the associated Dowker complex as well as simplex counts. In particular, for the latter quantity there are regimes of convergence to normal and to stable distribution depending on the heavy-tailedness of the weight distribution. I will conclude by a simulation study and an application to the collaboration network extracted from the arXiv dataset.

Session, Time and Place: Contributed Session 1.4, Monday 23 June 2025, 16:00–17:40, TBA.

Limit Theorems Under Heavy-tailed Scenario in the Age Dependent Random Connection Models

Takashi Owada

Purdue University
owada@purdue.edu

Co-authors: Christian Hirsch

Abstract: We consider limit theorems associated with subgraph counts in the age-dependent random connection model. First, we identify regimes where the count of sub-trees converges weakly to a stable random variable under suitable assumptions on the shape of trees. The proof relies on an intermediate result on weak convergence of associated point processes towards a Poisson point process. Additionally, we prove the same type of results for the clique counts. Here, a crucial ingredient includes the expectation asymptotics for clique counts, which itself is a result of independent interest.

Session, Time and Place: Contributed Session 1.4, Monday 23 June 2025, 16:00–17:40, TBA.

Clustering of Large Deviations Events in Heavy-tailed Moving Average Processes

Jiaqi Wang

Cornell University
jw2382@cornell.edu

Co-authors: Gennady Samorodnisky

Abstract: How do large deviation events in a stationary process cluster? The answer depends not only on the type of large deviations, but also on the length of memory in the process. Somewhat unexpectedly, it may also depend on the tails of the process. In this paper we work in the context of large deviations for partial sums in moving average processes with short memory and regularly varying tails. We show that the structure of the large deviation cluster in this case markedly differs from the corresponding structure in the case of exponentially light tails, considered in [1]. This is due to the difference between the "conspiracy" vs. the "catastrophy" principles underlying the large deviation events in the light tailed case and the heavy tailed case, correspondingly.

Extending the "catastrophe" principle from the i.i.d. case to dependent data, we establish that the number of consecutive rare events grows on the order of $O(n)$ with a uniform limit distribution. These results provide new insights into the occurrence and structure of extreme events in dependent heavy-tailed systems.

References:

1. Chakrabarty, A., Samorodnitsky, G. (2023). Clustering of large deviations in moving average processes: The short memory regime. *Stochastic Processes and their Applications*, Volume(163), p. 387-423.

Session, Time and Place: Contributed Session 1.4, Monday 23 June 2025, 16:00–17:40, TBA.

An Alternative Approach to Power Law Dynamics in Preferential Attachment Models

Max Ziegenbalg

Otto von Guericke University Magdeburg
max.ziegenbalg@ovgu.de

Co-authors: Anja Janssen

Abstract: A common feature observed in large real-world networks are degree distributions that resemble power laws. Since this has serious practical implications, many models have been proposed over time that aim to reflect this property. One of these is the class of preferential attachment models, which gained popularity soon after their introduction by Barabási and Albert in 1999. They describe a discrete-time growing graph process in which, at each time step, a newly added vertex randomly establishes a certain number of edges to existing vertices with a probability that is an affine function of their degrees. This 'rich-get-richer' dynamic provides an intuitive explanation for power-law distributions and has furthermore been proven to lead to these distributions asymptotically.
In this talk, we provide a complementary approach aimed at analysing individual vertices and their interactions in large networks. To this end, we select a fixed number of the oldest vertices and let them evolve for a heavy-tailed random number of time steps. Utilising tools from extreme value theory, such as the tail coefficient and spectral measure, we can then make predictions for the chosen degree vector in large networks, which we interpret as just the extremal realisations of our model. We discuss several model specifications, such as finite versus infinite dimensions, and fixed versus random numbers of outgoing edges per vertex.

Session, Time and Place: Contributed Session 1.4, Monday 23 June 2025, 16:00–17:40, TBA.

Generative Modelling of Geometric Multivariate Extremes Using Normalising Flows

Lambert De Monte

University of Edinburgh
l.demonte@ed.ac.uk

Co-authors: Raphaël Huser, Ioannis Papastathopoulos, Jordan Richards

Abstract: We develop a novel deep-learning statistical method for multivariate extremes by combining geometric extremes with normalising flows on hyperspheres. Relying on a radial-angular decomposition of a random vector of interest, we model the distribution of radial exceedances of a high quantile of the distribution of radii given their angles as well as that of the angles themselves. We develop models bridging parsimony and flexibility by exploiting links between parameters that determine the limit distribution upon which our framework is built. Leveraging the generative properties of normalising flows on hyperspheres, we efficiently generate samples from our fitted models and perform probability estimation for arbitrary extremal regions of the multivariate space via Monte Carlo integration. The merits of our method are demonstrated via a simulation study in up to 10 dimensions and a case study of low and high wind extremes in relation to electricity production across the Pacific Northwest region of the United States.

Session, Time and Place: Contributed Session 2.1, Tuesday 24 June 2025, 13:30–15:10, TBA.

BLAST: A Bayesian Lasso Tail Index Regression Model with an Application to Extreme Wildfires

Myung Won Lee

University of Edinburgh
johnny.myungwon.lee@ed.ac.uk

Co-authors: Miguel de Carvalho, Daniel Paulin, Soraia Pereira, Ricardo Trigo, Carlos Da Camara

Abstract: In this talk, we propose a novel regression-based approach to model extreme events and their underlying risk factors. We leverage Bayesian regularisation within a generalised additive framework for tail index regression, enabling a more flexible approach to analysing extreme values. Our framework revolves around a conditional Pareto-type specification, enriched by the inclusion of Bayesian Lasso-type shrinkage priors and further refined through low-rank thin plate splines basis expansion. The proposed approach admits a neural model representation and it balances parsimony and flexibility through a prior that favours linear effects by penalising complexity, while allowing nonlinear effects when justified by the data. The performance is validated through a simulation study and applied to investigate Portugal's devastating wildfires of October 2017 and analyse how magnitude of such occurrences relates with other risk factors.

References:

1. de Carvalho, M., Pereira, S., Pereira, P. and de Zea Bermudez, P. (2022). An extreme value Bayesian Lasso for the conditional left and right tails. *Journal of Agricultural, Biological and Environmental Statistics* 27 222–239.
2. Kyung, M., Casella, G., Ghosh, M. and Gill, J. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis* 5(2): 369–411.
3. Wang, H. & Tsai, C. L. (2009). Tail index regression. *Journal of the American Statistical Association* 104(487) 1233–1240.

Session, Time and Place: Contributed Session 2.1, Tuesday 24 June 2025, 13:30–15:10, TBA.

Spatial Scale-Aware Tail Dependence Modeling for High-Dimensional Spatial Extremes

Muyang Shi

Colorado State University
muyang.shi@colostate.edu

Co-authors: Likun Zhang, Mark Risser, Ben Shaby

Abstract: Extreme events over large spatial domains may exhibit highly heterogeneous tail dependence characteristics, yet most existing spatial extremes models yield only one dependence class over the entire spatial domain. To accurately characterize "data-level dependence" in analysis of extreme events, we propose a mixture model that achieves flexible dependence properties and allows high-dimensional inference for extremes of spatial processes. We modify the popular random scale construction that multiplies a Gaussian random field by a single radial variable; we allow the radial variable to vary smoothly across space and add non-stationarity to the Gaussian process. As the level of extremeness increases, this single model exhibits both asymptotic independence at long ranges and either asymptotic dependence or independence at short ranges. We make joint inference on the dependence model and a marginal model using a copula approach within a Bayesian hierarchical model. Three different simulation scenarios show close to nominal frequentist coverage rates. Lastly, we apply the model to a dataset of extreme summertime precipitation over the central United States. We find that the joint tail of precipitation exhibits non-stationary dependence structure that cannot be captured by limiting extreme value models or current state-of-the-art sub-asymptotic models.

Session, Time and Place: Contributed Session 2.1, Tuesday 24 June 2025, 13:30–15:10, TBA.

Statistical Inference and Model Selection for Models Adapted to Record Series

Jinane Jouni

Mines Saint-Etienne, Univ Clermont Auvergne, CNRS, UMR 6158 LIMOS, Institut Henri Fayol
jinane.jouni@emse.fr

Co-authors: Anis S. Hoayek, Gilles R. Ducharme, Aline Mefleh, Mireille Batton-Hubert

Abstract: We investigate the statistical properties of upper records observed in a time series $\{X_t\}_{t\geq 1}$ under four widely studied record models: the Classical Record Model, the Yang-Nevzorov Model, the Linear Drift Model, and the Discrete-Time Random Walk Model. Our study focuses on the likelihood function constructed from the pairs of record values and their occurrence indicators, denoted as $\{R_n, L_n\}$. Using this likelihood function, we derive estimators for the key parameters governing these models. Furthermore, we explore statistical inference techniques for model selection, enabling the identification of the most suitable record model for a given series. To validate our theoretical findings, we conduct a comprehensive simulation study, illustrating the behavior of the estimators and their performance under different model settings.

References:

1. Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1998). *Records.* John Wiley & Sons.
2. Ballerini R and Resnick S 1985 *J. Appl. Probab.* 22 487–502
3. Hoayek, A. (2016). *Estimation des paramètres pour des modèles adaptés aux séries de records.* PhD thesis, Université Montpellier
4. Hoayek, A. S., Ducharme, G. R., and Khraibani, Z. (2017). Inférence fondée sur la vraisemblance pour des modèles de records. *Comptes Rendus. Mathématiques*, 355(10):1099–1103.
5. Majumdar, S. N. and Ziff, R. M. (2008). *Universal record statistics of random walks and Lévy flights.* Physical review letters, 101(5):050601.
6. Nevzorov, V. (1990). *Records for nonidentically distributed random variables.* In Proceedings of the Fifth Vilnius Conference, volume 2, pages 227–233

Session, Time and Place: Contributed Session 2.1, Tuesday 24 June 2025, 13:30–15:10, TBA.

Measuring and Testing Tail Asymmetry

Bikramjit Das

Singapore University of Technology and Design
bikram@sutd.edu.sg

Co-authors: Xiangyu Liu

Abstract: Measures of asymptotic tail dependence between two variables often tends to be symmetric in nature. We define a measure of *extreme tail association* between two variables, which is asymmetric, easily computable in sample data, and consistent for the population measure. We show that the sample tail association measure is weakly consistent under mild conditions on the underlying distribution. We also propose a test for bivariate tail asymmetry, and compute asymptotic distributions of the test statistic under both null and alternative hypotheses. Data from movement in cryptocurrency prices is used to exhibit the efficacy of the tools developed.

Session, Time and Place: Contributed Session 2.2, Tuesday 24 June 2025, 13:30–15:10, TBA.

Bootstrap for The Vector Tail Empirical Process : Extension of Ivanoff, Kulik and Loukrati (2023)

Jun-ichiro Fukuchi

Gakushuin University
junichiro.fukuchi@gakushuin.ac.jp

Abstract: Ivanoff, Kulik and Loukrati (*Extremes* (2023), 26), hereafter abbreviated to IKL, proved consistency of Efron's bootstrap for the tail empirical process and its integral functional. We extend their results to the vector tail empirical process derived from the sample of multivariate population. Let $\{X_k, k = 1, 2, \ldots,\}$ be a sequence of i.i.d. d-dimensional random vectors with a distribution function F and one-dimensional marginal distribution function F_i. Write $X_k = (X_{1,k}, \ldots, X_{d,k})^\mathsf{T}$. Let $\bar{F}_i(x) = 1-F_i(x)$ be the survivor function and assume that $\bar{F}_i$ is regularly varying with index $-\alpha_i > 0$. Write $T_i(t) = t^{-\alpha_i}$ and $\tilde{T}_{i,n}(t) = (n\bar{F}_i(u_n))^{-1} \sum_{k=1}^{n} \mathrm{I}(X_{i,k} > u_n t)$, $T_{i,n}(t) = \frac{\bar{F}_i(u_n t)}{\bar{F}_i(u_n)}$ where u_n is a sequence of real numbers such that $u_n \to \infty$ and $\min_{1\le i\le d} nF_i(u_n) \to \infty$. Define the tail empirical process (TEP) of i-th element by $G_{i,n}(t) = \sqrt{k_{i,n}} \left\{\tilde{T}_{i,n}(t) - T_{i,n}(t)\right\}, t > 0$. We call $\boldsymbol{G}_n(\cdot) = (G_{1,n}(\cdot), \ldots, G_{d,n}(\cdot))^\mathsf{T}$ the vector tail empirical process (VTEP). Let $\widetilde{G}_{i,n}(t)$ denote the TEP in which $T_{i,n}(t)$ is replaced with $T_i(t)$. Write $\widetilde{\boldsymbol{G}}_n(t) = \left(\widetilde{G}_{1,n}(t), \ldots, \widetilde{G}_{d,n}(t)\right)^\mathsf{T}$. First, We proved that VTEP $\boldsymbol{G}_n(\cdot)$ converges weakly to a Gaussian process if $1-F$ has multivariate regularly varying condition. We also proved that $\widetilde{\boldsymbol{G}}_n(t)$ converges weakly to a Gaussian process if $1 - F_i$ satisfies the second order regularly varying condition for each $i = 1, \ldots, d$. Second, as in IKL, weak convergence of vector tail empirical process generated by triangular arrays of random vectors is proved. The consistency of bootstrap for VTEP and its integral functionals follows from this result. As an application, we discuss a subset selection method of the component whose high quantile is the largest among d components of a multivariate distribution.

Session, Time and Place: Contributed Session 2.2, Tuesday 24 June 2025, 13:30–15:10, TBA.

A Goodness of Fit Test for Distributions of Extreme Events

Jiwoong Kim

University of South Florida
jwboys26@gmail.com

Abstract: This paper investigates computational aspects of a goodness-of-fit test for non-Gaussian distributions of extreme events based on Khmaladze martingale transformation, proposed by Khmaladze (1981), when the location and scale parameters of the distribution are unknown. In probability theory, Cauchy and Gumbel distributions gained attention since they were proposed as alternatives to a normal distribution to explain certain extreme events in economics and hydrology such as the sudden collapse of stock prices and a radical change in daily rainfall. Other applications of the Cauchy and Gumbel distributions can be found in various fields, including nuclear science. Crucially, however, research on the goodness-of-fit test for the Cauchy and Gumbel distributions has not garnered as much attention from statisticians as research on its real-world applications. We started this study against the backdrop of a rarity of tests for distributions of extreme events. In this paper, we rigorously investigate whether Khmaladze martingale transformation can be a viable option to test for distributions of extreme events. As a result of the investigation, we obtain useful by-products, including Mill's ratios of certain non-Gaussian distributions. Simulation studies demonstrate that the proposed goodness-of-fit test compares favorably with other well-celebrated tests.

References: 1. Khmaladze, E.V. (1981). A martingale approach in the theory of goodness- of-fit tests. Theory of Probability and Its Application, 26, 240-257.

Session, Time and Place: Contributed Session 2.2, Tuesday 24 June 2025, 13:30–15:10, TBA.

Testing for Time-Varying Extremal Dependence

Christian Rohrbeck

University of Bath
cr777@bath.ac.uk

Co-authors: Matthew Pawley, Euan McGonigle, Jordan Richards

Abstract: When fitting multivariate extreme value models, we often assume that the tail dependence structure is stationary, that is, it does not change over time. However, such an assumption may be questionable in practice. For instance, the spatial structure of climate extremes may change over time, and new regulatory frameworks can cause structural changes in the joint tail behaviour of financial asset prices. Therefore, we need tests to determine whether stationarity in the tail dependence structure is a reasonable assumption or not.

In this talk we introduce a statistical test for changes in the tail dependence of a multivariate random vector. We start by defining a time-dependent extension of the tail pairwise dependence matrix (TPDM) and establish its basic properties. Next, we derive the asymptotic null distribution of functionals of this object, which we use to test for deviations over time. Compared to other methods, such as Drees (2023), our test scales better to high dimensions, which we illustrate across a range of simulated data sets. We further apply our approach to Red Sea surface temperature data. Finally, we will outline how our method can be embedded into a change point detection algorithm to identify the time points at which the tail dependence of a random vector has undergone a change.

References:

1. Drees, H. (2023). Statistical inference on a changing extreme value dependence structure. *Annals of Statistics*, 51(4), 1824-1849.

Session, Time and Place: Contributed Session 2.2, Tuesday 24 June 2025, 13:30–15:10, TBA.

Dimension Reduction for Multivariate Geometric Extremes

Michela Corradini

Lancaster University
m.corradini@lancaster.ac.uk

Co-authors: Jeongjin Lee, Jennifer Wadsworth

Abstract: A recent geometric approach based on the convergence of the scaled sample cloud onto the limit set allows for capturing complex extremal dependence structures. The shape of the limit set describes the dependence structure, characterized by the gauge function. In order to perform statistical inference using the gauge function, a transformation to radial-angular components of the random vector $\mathbf{X}$ is convenient. However, effectively modelling the angular components in higher dimensions remains a challenge. We explore dimension reduction possibilities, facilitated by translating the angular components from the $d-1$-dimensional simplex to $\mathbb{R}^{d-1}$, as in compositional data analysis. Our aim in doing this is to preserve key features of the gauge function and angular distribution.

Session, Time and Place: Contributed Session 2.3, Tuesday 24 June 2025, 13:30–15:10, TBA.

Identifying Extremal Dependence Classes Using Additive Mixtures in the Geometric Framework

Jeongjin Lee

Lancaster University
j.lee58@lancaster.ac.uk

Co-authors: Jennifer Wadsworth

Abstract: The framework of geometric extremes is based on the convergence of scaled sample clouds onto a limit set, characterised by a gauge function, with the shape of the limit set determining its extremal dependence structure. While it is known that a blunt limit set implies asymptotic independence, the absence of bluntness can be linked to both asymptotic dependence and independence. Focusing on the bivariate case, under a truncated gamma modelling assumption, we show that a "pointy" limit set implies asymptotic dependence, thus offering practical geometric criteria for identifying extremal dependence class. Suitable models for the gauge function offer the ability to capture asymptotically independent or dependent data structures, without requiring prior knowledge of the true extremal dependence structure. The geometric approach thus offers a simple alternative to various parametric copula models that have been developed for this purpose in recent years. We consider two types of additively mixed gauge functions that provide a smooth interpolation between asymptotic dependence and asymptotic independence. We derive their explicit forms, explore their properties, and establish connections to the developed geometric criteria. Through a simulation study, we evaluate the effectiveness of the geometric approach with additively mixed gauge functions, comparing its performance to existing methodologies that account for both asymptotic dependence and asymptotic independence. This geometric approach is computationally efficient and yields reliable performance across various extremal dependence scenarios.

Session, Time and Place: Contributed Session 2.3, Tuesday 24 June 2025, 13:30–15:10, TBA.

Statistical Inference for Extremes of Stochastic Processes with Radial Generalized Pareto Process

Xindi Song

University of Edinburgh
s2150450@ed.ac.uk

Co-authors: Ioannis Papastathopoulos

Abstract: Accurately estimating the probabilities of rare events that surpass historical records is vital for risk assessment and prevention strategies. Motivated by a novel geometrical framework, in this work we develop functional extreme value models, suitable for asymptotically dependent and asymptotically independent stochastic processes, with a special emphasis on modelling the extremes of temporal stochastic processes. The method is based on breaking each realisation down to its magnitude and its shape, and on capturing how these variables depend in a suitably defined joint tail region. Specifically, we develop scalar-on-function regression models that capture non-linear effects of the process's shape on both high directional quantiles of the magnitude and the mean excess above these quantiles. Our methodology leads to extreme value models that permit extrapolation along any direction in function space, allowing for the analysis and occurrence of unprecedented event shapes and magnitudes. We showcase our framework through a simulation study on extremes of asymptotically independent processes. We also demonstrate the practical application of our framework by estimating return periods of extreme heat events in central England. This involves analyzing ($n = 147$) daily temperature data ($d = 365$), where each annual temperature profile is treated as a continuous curve, enabling functional regression to capture the shape and magnitude of extreme events.

Session, Time and Place: Contributed Session 2.3, Tuesday 24 June 2025, 13:30–15:10, TBA.

Seasonal Trend Assessment of U.S. Extreme Precipitation via Changepoint Segmentation

Jaechoul Lee

Northern Arizona University
jaechoul.lee@nau.edu

Co-authors: Mintaek Lee, Thea Sukianto

Abstract: Most climate trend studies analyze long-term trends as a proxy for climate dynamics. However, when examining seasonal data, it is unrealistic to assume that long-term trends remain consistent across all seasons. Instead, each season likely experiences distinct trends. Additionally, seasonal climate time series, such as seasonal maximum precipitation, often exhibit nonstationarities, including periodicities and location shifts. Failure to rigorously account for these features in modeling may lead to inaccurate trend estimates. This study quantifies seasonal trends in the contiguous United States' seasonal maximum precipitation series while addressing these nonstationarities. To ensure accurate trend estimation, we identify changepoints where the seasonal maximum precipitation shifts due to factors like measurement device changes, observer differences, or location moves. We employ a penalized likelihood method to estimate multiple changepoints, incorporating a generalized extreme value distribution with periodic features. A genetic algorithm based search algorithm efficiently explores the vast space of potential changepoints in both number and timing. Additionally, we compute seasonal return levels for extreme precipitation. Our methods are illustrated using two selected stations, and the results for the U.S. are summarized through maps. We find that seasonal trends vary more when changepoints are considered than in studies that ignore them. Our findings also reveal distinct regional and seasonal patterns, with increasing trends more prevalent during fall in the South and along the East Coast when changepoints are accounted for.

Session, Time and Place: Contributed Session 2.4, Tuesday 24 June 2025, 13:30–15:10, TBA.

Modeling Autoregressive Conditional Regional Extremes with Applications to Solar Flare Prediction

Steven Moen

The University of Wisconsin–Madison
stmoen@wisc.edu

Co-authors: Jili Wang, Zhengjun Zhang

Abstract: This paper studies big data streams with regional-temporal extreme event (REE) structures and solar flare prediction. An autoregressive conditional Fréchet model with time-varying parameters for regional and its adjacent regional extremes (ACRAE) is proposed. The ACRAE model can quickly and accurately predict rare REEs (i.e., solar flares) in big data streams. The ACRAE model, with some mild regularity conditions, is proved to be stationary and ergodic. The parameter estimators are derived through the conditional maximum likelihood method. The consistency and asymptotic normality of the estimators are established. Simulations are used to demonstrate the efficiency of the proposed parameter estimators. In real solar flare prediction, with the new dynamic extreme value modeling, the occurrence and climax of solar activity can be predicted earlier than with existing algorithms. The empirical study shows that the ACRAE model outperforms the existing prediction algorithms with sampling strategies.

Session, Time and Place: Contributed Session 2.4, Tuesday 24 June 2025, 13:30–15:10, TBA.

Statistical Modelling of Earthquake Occurrence and Extreme Magnitudes for Seismic Risk Assessment

WANCHEN YUE

Lancaster University
w.yue1@lancaster.ac.uk

Co-authors: Jonathan Tawn, Israel Martinez Hernandez

Abstract: The rise of human-induced seismicity, driven by industrial activities, such as gas extraction and carbon dioxide storage, poses significant risks to public safety and infrastructure. These problems affect communities globally, with severe consequences for urban development and environmental safety. At the Groningen gas field in the Netherlands, earthquakes triggered by gas extraction have damaged infrastructure and raised concerns about seismic events affecting local residents' livelihoods. Accurate statistical modelling is crucial for assessing risks and guiding public policy and in particular it is important to incorporate spatial-temporal variations from stresses due to changing gas extraction rates, and from the evolution of the quality of the measurement network. Our research leverages statistical methodologies to address key challenges in seismic risk assessment. We develop novel techniques to estimate the upper tail of earthquake magnitude distributions, incorporating expert knowledge of the maximum possible magnitude and addressing measurement error in the magnitude distribution. This approach enhances the reliability of extreme magnitude predictions, which are vital for effective risk preparedness and mitigation. We will also present work on the spatio-temporal occurrence of extreme earthquakes. Unlike other environmental extreme value applications, there are well-established models describing this clustering of earthquakes, namely Epidemic-Type Aftershock Sequence (ETAS) models, but we develop extensions of these models for human-induced seismicity.

Session, Time and Place: Contributed Session 2.4, Tuesday 24 June 2025, 13:30–15:10, TBA.

Sparse Multivariate Autoregressive Conditional Fréchet Models for High-frequency Extreme Risk

Zhengjun Zhang

University of Chinese Academy of Sciences
zhangzhengjun@ucas.ac.cn

Co-authors: Yongjoon Kim

Abstract: We introduce a novel class of multivariate Fréchet models that integrate dynamic cross-sectional and temporal dependencies to effectively capture time-varying marginal characteristics. Building on the classical multivariate maxima of moving maxima (M4) framework, we propose a sparse variant of the M4 random coefficient model (vSM4R) to address estimation challenges in high-dimensional settings. For marginal processes, we extend the autoregressive conditional Fréchet (AcF) model to an mmm-dependent structure (mAcF), enabling richer dependency patterns while maintaining a direct link to the vSM4R through standardization. We establish key probabilistic properties and develop composite likelihood estimation methods with proven statistical guarantees. Through simulations and high-frequency cryptocurrency return data, we demonstrate the robustness of our approach in capturing tail risks and modeling extreme co-movements in financial markets. Our results underscore the necessity of flexible multivariate extreme value models for risk management in volatile environments.

Session, Time and Place: Contributed Session 2.4, Tuesday 24 June 2025, 13:30–15:10, TBA.

Extremes in Risk Management – A Nonparametric Approach to the Estimation of the Quantiles of Compound Distributions

Helgard Raubenheimer

Centre for BMI, North-West University, South Africa
Helgard.Raubenheimer@nwu.ac.za

Co-authors: Tertius de Wet, Charl Pretorius, Riaan de Jongh

Abstract: Estimation of operational risk reserves is still widely done using the loss distribution approach. The accuracy of the estimation depends heavily on the accuracy with which the extreme quantiles of the aggregate loss distributions are estimated. Various approaches have been proposed to estimate the extreme quantiles of this compound distribution, including estimators based on the single-loss and perturbative approximations that rely on estimating an even more extreme quantile of the underlying severity distribution. However, estimation of these extreme quantiles may be inaccurate due to fitting a parametric severity distribution that fits the bulk of the data well but struggles to capture the tail behavioural characteristics of the distribution that generated the loss data. To circumvent this problem, we propose estimating nonparametrically a less extreme or lower quantile of the severity distribution, hopefully with better accuracy, and then multiplying this lower quantile by a certain factor to obtain an estimate of the required extreme quantile of the compound distribution. The factor or multiplier is derived by using extreme value theory and the single loss and perturbative approximation and estimated nonparametrically. This approach is evaluated by means of a simulation study which suggest that the second-order multiplier estimator based on the second-order perturbative approximation, is a good choice for practical applications.

Session, Time and Place: Contributed Session 3.1, Tuesday 24 June 2025, 15:40–17:20, TBA.

Influence of Extremes on the Quantile Treatment Effect Estimation

Gennady Samorodnitsky

Cornell University
gs18@cornell.edu

Co-authors: Marco Avella Medina, Richard Davis

Abstract: We consider the problem of estimating quantile treatment effects without assuming strict overlap , i.e., we do not assume that the propensity score is bounded away from zero. More specifically, we consider an inverse probability weighting (IPW) approach for estimating quantiles in the presence of missing data and pay special attention to scenarios where the propensity scores can tend to zero as a regularly varying function. In this case small propensity scores becomes extremes affecting estimation. Our approach effectively considers a heavy-tailed objective function for estimating the quantile process. We introduce a truncated and debiased IPW estimator that is shown to outperform the standard quantile IPW estimator when strict overlap does not hold.

Session, Time and Place: Contributed Session 3.1, Tuesday 24 June 2025, 15:40–17:20, TBA.

Extreme Quantile Regression Using Generalized Random Forests and Block Maxima Approach.

Lucien Mahutin Vidagbandji

Université Le Havre Normandie, LMAH UR 3821, France.
mahutin-lucien.vidagbandji@univ-lehavre.fr

Co-authors: Alexandre Berred, Cyrille Bertelle, Laurent Amanton.

Abstract: Quantile regression estimates conditional quantiles, enabling the exploration of relationships between independent and dependent variables. However, it faces extrapolation difficulties at high probability levels due to the scarcity of data in the tails of the distribution, as well as challenges related to the complexity of quantile functions and the potentially large number of predictors. To address these issues, we propose a method that models block maxima using the conditional generalized extreme value (GEV) distribution, whose parameters depend on covariates, thus enhancing extrapolation in the tails. We then use an approach based on generalized random forests (GRF) to estimate these parameters. Specifically, we maximize a penalized likelihood weighted by GRF-derived weights. This penalization overcomes the limitations of the maximum likelihood estimator (MLE) for small samples while maintaining its optimality for large samples. The performance of our method is demonstrated through simulation studies and an application to U.S. wage data from 1980.

1. Gnecco, N., Terefe, E.M., Engelke, S. (2024). Extremal Random Forests. *Journal of the American Statistical Association*, 1–24.
2. Coles, S.G., Dixon, M.J. (1999). Likelihood-Based Inference for Extreme Value Models. *Extremes*, 2(1), 5–23.

Session, Time and Place: Contributed Session 3.1, Tuesday 24 June 2025, 15:40–17:20, TBA.

Downsampling for Imbalanced Classification Using Infinite Centered Random Forests

Moria Grace Mayala

Sorbonne université (Laboratoire de Probabilités, Statistique et Modelisation)
moria.mayala@sorbonne-universite.fr

Co-authors: Olivier Wintenberger, Charles Tillier and Erwan Scornet

Abstract: Imbalanced data in classification tasks is widespread across many data mining domains and has been identified as a top problem in machine learning. Numerous solutions have been proposed to address the challenges posed by imbalanced data, ranging from data-level approaches to algorithm-level approaches. We study classification with downsampling, which is one of the popular data-level approaches using Centered Random Forests (CRFs). We show that while this method effectively addresses the imbalance, it introduces a substantial bias compared to the classical Random Forests (RFs) approach. To address this issue, we propose a debiasing method based on the two-stage(multistage) sampling. Furthermore, we establish the asymptotic normality of the estimators in both cases, providing a rigorous theoretical framework for evaluating their performance. Our theoretical results are supported by simulations demonstrating the effectiveness of our approach.

Session, Time and Place: Contributed Session 3.2, Tuesday 24 June 2025, 15:40–17:20, TBA.

Taming Sample Moments

John Nolan

American University
jpnolan@american.edu

Abstract: Computing the sample mean and other moments can give unreliable estimates in the presence of heavy tailed data, even with very large data sets. We describe a new technique that guards against extremes by splitting a data set into two subsets: small/moderate values and extreme values. Statistics are computed for each subset and then the results are combined to give a more reliable estimate of the desired moment.

Session, Time and Place: Contributed Session 3.2, Tuesday 24 June 2025, 15:40–17:20, TBA.

Block Maxima Modelling in the Presence of Missing Data

Emma Simpson

University College London
emma.simpson@ucl.ac.uk

Co-authors: Paul Northrop

Abstract: Modelling block maxima using the generalised extreme value (GEV) distribution is a common method for studying univariate extremes. One practical challenge in applying this methodology, which is often overlooked, is handling datasets containing missing values. In this case, one cannot be sure whether the true maximum has been observed in each block. If the issue is ignored, one may obtain biased GEV parameter estimates and encounter subsequent underestimation of return levels, which is clearly undesirable when these are to be used for practical decision making. An approach that is often used to handle missing data in this setting is to discard blocks where the proportion of missingness exceeds some specified threshold. The idea behind this is that the chance of having recorded the maximum in each remaining block is reasonably high, and the risk of obtaining biased return level estimates is therefore reduced. While this is a sensible approach, extreme value modelling already comes with an intrinsic lack-of-data issue, so it would generally be preferable not to lose any of the information contained in these discarded blocks. We propose an alternative approach where the standard block maxima methodology is extended to overcome missing data issues. Within the estimation of the GEV model parameters, we explicitly account for the proportion of missing values in each block. Inference is carried out using likelihood-based techniques, with return level estimates and associated confidence intervals obtained through profiling. The proposed methodology is demonstrated on air pollution data from Bloomsbury, U.K.

Session, Time and Place: Contributed Session 3.2, Tuesday 24 June 2025, 15:40–17:20, TBA.

Local LASSO: Variable Selection in High Dimensional Spatial Regression

Debjoy Thakur

Washington University in St. Louis
debjoy@wustl.edu

Co-authors: Nwakanma Sidney, Soumendra N. Lahiri

Abstract: In spatial regression, a challenging problem is to select a correct set of non-zero covariates in many real-life studies. For example in the election result, two closer counties may have a similar set of relevant covariates whereas for two distant counties, this similarity is very rare to behold. In the literature, the researchers wanted to solve this problem by grouped LASSO but here the problem is that if a covariate is zero in any location that does not imply the covariate will be zero in the entire spatial surface. In this context, we first formulate local LASSO such that two closer locations will have a similar set of non-zero covariates. We capture the non-zero spatial signal in the coefficient function by the tensor product of the wavelet transformation in the prototype set of a spatial surface. As a result, if a spatial covariate surface is non-zero then there will exist at least one non-zero signal in its prototype set in any resolution. We give the idea of how to generalize local LASSO under the SCAD penalty. We provide the exact post-selection inference of the selected covariates for every location. We theoretically validate the variable selection consistency under some regularity conditions. We have justified our results with two real-life election results.

Session, Time and Place: Contributed Session 3.2, Tuesday 24 June 2025, 15:40–17:20, TBA.

Likelihood-based Inference in Stationary Time Series: The Block Maxima Method

David Carl

Bocconi University
david.carl@phd.unibocconi.it

Co-authors: Simone A. Padoan, Stefano Rizzelli

Abstract: In this work, we investigate the block maxima method in the context of stationary time series. We begin by extending aspects of likelihood asymptotic theory for the estimation of the marginal parameters of the Generalized Extreme Value (GEV) distribution from the case of independence to scenarios involving serial dependence. Once the likelihood framework is established at a suitable level of generality, we shift our focus to its Bayesian counterpart, studying the corresponding asymptotic properties. Frequentist and Bayesian inference is then employed to estimate marginal parameters of the GEV, the extremal index, return levels and extreme quantiles of the underlying stationary distribution.

Session, Time and Place: Contributed Session 3.3, Tuesday 24 June 2025, 15:40–17:20, TBA.

Extreme Value Analysis Based on Blockwise Top-two Order Statistics

Erik Haufs

Ruhr-University Bochum
erik.haufs@rub.de

Co-authors: Axel Bücher

Abstract: Extreme value analysis for time series is often based on the block maxima method, in particular for environmental applications. In the classical univariate case, the latter is based on fitting an extreme-value distribution to the sample of (annual) block maxima. Mathematically, the target parameters of the extreme-value distribution also show up in limit results for other high order statistics, which suggests estimation based on blockwise large order statistics. It is shown that a naive approach based on maximizing an independence log-likelihood yields an estimator that is inconsistent in general. A consistent, bias-corrected estimator is proposed, and is analyzed theoretically and in finite-sample simulation studies. The new estimator is shown to be more efficient than traditional counterparts, for instance for estimating large return levels or return periods.

References:

1. Bücher, A., & H. *Extreme Value Analysis based on Blockwise Top-Two Order Statistics*. 2025. arXiv: 2502.15036 [math.ST].

Session, Time and Place: Contributed Session 3.3, Tuesday 24 June 2025, 15:40–17:20, TBA.

Extremes for Non-stationary Time Series

RAFAL KULIK

University of Ottawa
rafal.kulik@uottawa.ca

Co-authors: Mai Ghannam and Stanislav Volgushev

Abstract: In this talk we will review current results for extremes of non-stationary time series. We will start with presenting different non-stationary models used in the literature. We will give limiting results for sample maxima. These results will reveal different behaviour of the maxima for the entire sample and for blocks. It has an influence on statistical inference. I will continue with statistical methodology. Some results for Peak-over-Threshold method have been developed recently, however, almost nothing is known for the Block Maxima method. Based on the recent work with Mai Ghannam and Stanislav Volgushev, we will show that the Block Maxima methodology may lead to inconsistency. We will show how to fix this problem.

Session, Time and Place: Contributed Session 3.3, Tuesday 24 June 2025, 15:40–17:20, TBA.

Fast and Efficient Inference for Flexible Spatial Extremes Models

Boris Béranger

UNSW Sydney
b.beranger@unsw.edu.au

Co-authors: Peng Zhong, Scott A. Sisson

Abstract: Statistical modelling of spatial extreme events has gained increasing attention over the last few decades with max-stable processes, and more recently r-Pareto processes, becoming the reference tools for the statistical analysis of asymptotically dependent data. Although inference for r-Pareto processes is easier than for max-stable processes, there remains major hurdles for their application to very high dimensional datasets within a reasonable timeframe. In addition, both approaches have almost exclusively considered the Brown-Resnick model for its Gaussian-based foundations and the continuity of its exponent measure. In this talk, we derive a class of models for which this continuity property holds and present the *skewed Brown-Resnick* model, an extension of the Brown-Resnick that allows for non-stationarity in the dependence structure, and the *truncated extremal-t*, a refinement of the well-known extremal-t model. We use an inference methodology based on the intensity function of the process which is derived from the exponent measure, and demonstrate the statistical and computational efficiency of this approach. Applications to two real-world problems illustrate valuable gains in flexibility from the proposed models as well as appealing computational gains over reference methodologies.

Session, Time and Place: Contributed Session 3.4, Tuesday 24 June 2025, 15:40–17:20, TBA.

Conditional Extreme Value Models for Multivariate and Spatial Applications

Kristina Bratkova

Lancaster University, United Kingdom
k.bratkova@lancaster.ac.uk

Co-authors: Jonathan A. Tawn, Simon Brown, Ingrid Hobæk Haff

Abstract: Extreme environmental events such as the occurrence of a combination of high values of air pollutants and heatwaves over sites exhibit different forms of extremal dependence. Conditional extreme value models provide a versatile framework that allows for either asymptotic dependence or asymptotic independence pairwise. We extend the conditional extreme value model framework in two settings: multivariate (Heffernan and Tawn, 2004) and spatial (Wadsworth and Tawn, 2022). In a multivariate case, we look at extensions of existing models for modelling the residuals to address concerns about the curse of dimensionality and the limitations of the modelling of the marginal distribution of the residuals. Specifically, we look at an asymmetric generalised Gaussian distribution for the margins and vine copulas for the dependence. In a spatial case, we address spatial non-stationarity and spatial-temporal extreme events. We illustrate the methods with applications to air pollutants and gridded UK temperature data from a climate model.

Session, Time and Place: Contributed Session 3.4, Tuesday 24 June 2025, 15:40–17:20, TBA.

Spatio-Temporal Statistical Modeling of the Occurrence of Extreme Events

Carolin Forster

University of Stuttgart
Carolin.Forster@mathematik.uni-stuttgart.de

Co-authors: Marco Oesting

Abstract: In this work, we aim to model spatio-temporal heavy precipitation events, which enhances the understanding of their causes and their prediction. Therefore, we derived a theoretical result confirming that it is reasonable to model the occurrence of extremes in space-time with a point process as the point process of extremes converges in distribution and its limit distribution can be determined explicitly. Similarly to Koh et al. (2023), we propose to use a spatio-temporal point process with covariates. Here we model the intensity measure of the spatio-temporal point process differently adapting an approach from Baddeley et al. (2012) who estimate the intensity measure nonparametrically. This point process model can be extended by adding marks such as the spatial extent (see Oesting & Huser, 2022).

References:

1. A. Baddeley, Y.-M. Chang, Y. Song, & R. Turner (2012). Nonparametric estimation of the dependence of a spatial point process on spatial covariates. *Statistics and its Interface*, 5(2):221-236.
2. J. Koh, F.Pimont, J.-L. Dupuy, & T. Opitz (2023). Spatiotemporal wildfire modeling through point processes with moderate and extreme mark. *The Annals of Applied Statistics*, 17(1):560-582.
3. M. Oesting & R. Huser (2022). Patterns in spatio-temporal extremes. arXiv preprint arXiv:2212.11001

Session, Time and Place: Contributed Session 3.4, Tuesday 24 June 2025, 15:40–17:20, TBA.

Bayesian Inference for Functional Extreme Events Defined via Partially Unobserved Processes

Max Thannheimer

University of Stuttgart
max.thannheimer@mathematik.uni-stuttgart.de

Co-authors: Marco Oesting

Abstract: In order to describe the extremal behaviour of some stochastic process X approaches from univariate extreme value theory are typically generalized to the spacial domain. Besides max-stable processes, that can be used in analogy to the block maxima approach, a generalized peaks-over-threshold approach can be used, allowing us to consider single extreme events. These can be flexibly defined as exceedances of a *risk functional* r, such as a spatial average, applied to X.Inference for the resulting limit process, the so-called r-*Pareto process*, requires the evaluation of $r(X)$ and thus the knowledge of the whole process X.

In practical application we face the challenge that observations of X are only available at single sites. To overcome this issue, we propose a two-step MCMC-algorithm in a Bayesian framework. In a first step, we sample from X conditionally on the observations in order to evaluate which observations lead to r-exceedances. In a second step, we use these exceedances to sample from the posterior distribution of the parameters of the limiting r-*Pareto process*. Alternating these steps results in a full Bayesian model for the extremes of X.

We show that, under appropriate assumptions, the probability of classifying an observation as r-exceedance in the first step converges to the desired probability. Furthermore, given the first step, the distribution of the Markov chain constructed in the second step converges to the posterior distribution of interest. Our procedure is compared to the Bayesian version of the standard procedure in a simulation study.

Session, Time and Place: Contributed Session 3.4, Tuesday 24 June 2025, 15:40–17:20, TBA.

On Estimation and Order Selection for Multivariate Extremes via Clustering

He Tang

University of Georgia
ht11145@uga.edu

Co-authors: Shiyuan Deng, Shuyang Bai

Abstract: We investigate the estimation of multivariate extreme models with a discrete spectral measure using spherical clustering techniques. The primary contribution involves devising a method for selecting the order, that is, the number of clusters. The method consistently identifies the true order, i.e., the number of spectral atoms, and enjoys intuitive implementation in practice. Specifically, we introduce an extra penalty term to the well-known simplified average silhouette width, which penalizes small cluster sizes and small dissimilarities between cluster centers. Consequently, we provide a consistent method for determining the order of a max-linear factor model, where a typical information-based approach is not viable. Our second contribution is a large-deviation-type analysis for estimating the discrete spectral measure through clustering methods, which serves as an assessment of the convergence quality of clustering-based estimation for multivariate extremes. Additionally, as a third contribution, we discuss how estimating the discrete measure can lead to parameter estimations of heavy-tailed factor models. We also present simulations and real-data studies that demonstrate order selection and factor model estimation.

Session, Time and Place: Contributed Session 4.1, Thursday 26 June 2025, 13:45–15:25, TBA.

Piecewise-linear Modeling of Multivariate Geometric Extremes

Ryan Campbell

Lancaster University
r.campbell3@lancaster.ac.uk

Co-authors: Jennifer Wadsworth

Abstract: A recent development in extreme value modeling uses the geometry of the dataset to perform inference on the multivariate tail. A key quantity in this inference is the gauge function, whose values define this geometry. Methodology proposed to date for capturing the gauge function either lacks flexibility due to parametric specifications, or relies on complex neural network specifications in dimensions greater than three. We propose a semiparametric gauge function that is piecewise-linear, making it simple to interpret and provides a good approximation for the true underlying gauge function. This linearity also makes optimization tasks computationally inexpensive. The piecewise-linear gauge function can be used to define both a radial and an angular model, allowing for the joint fitting of extremal pseudo-polar coordinates, a key aspect of this geometric framework. We further expand the toolkit for geometric extremal modeling through the estimation of high radial quantiles at given angular values via kernel density estimation. We apply the new methodology to air pollution data, which exhibits a complex extremal dependence structure.

Session, Time and Place: Contributed Session 4.1, Thursday 26 June 2025, 13:45–15:25, TBA.

Causal Tail Coefficient for Compound Extremes in Multivariate Time Series

Cathy Yin

Imperial College London
l.yin24@imperial.ac.uk

Co-authors: Almut Veraart, Adam Sykulski

Abstract: Extreme events are multivariate in nature. A compound extreme occurs when a combination of variables together produces a significant impact, even if individual components are not marginally extreme. Most methods for detecting extremal causality focus primarily on individual extreme values and lack the flexibility to capture causal relationships between compound extremes. This work introduces a novel framework for detecting causal dependencies between extreme events, including compound extremes. We define the compound causal tail coefficient that captures the extremal dependance of compound events between pairs of time series. Based on a consistent estimator of this coefficient, we develop a bootstrap hypothesis test to evaluate the presence and direction of causal relationships. Our method can accommodate nonlinearity and latent confounding variables in autoregressive time series models. We demonstrate the effectiveness of our method through simulations and an application to real-world space-weather data.

Session, Time and Place: Contributed Session 4.1, Thursday 26 June 2025, 13:45–15:25, TBA.

Multiple Extremal Integrals

Jiemiao Chen

University of Georgia
jc17876@uga.edu

Co-authors: Shuyang Bai

Abstract: We introduce the new concept of multiple extremal integrals as an extension of single extremal integrals, which have played important roles in extreme value theory. The multiple extremal integrals are formulated in terms of a product random sup measure derived from the α-Fréchet random sup measure. We establish a LePage-type representation similar to the one used for multiple sum-stable integrals, which have been extensively studied in literature. This approach allows us to investigate the integrability, tail behavior, and independence properties of multiple extremal integrals. Additionally, we discuss an extension of a recent stationary model that exhibits an unusual extremal clustering phenomenon, now constructed based on the multiple extremal integral.

Session, Time and Place: Contributed Session 4.1, Thursday 26 June 2025, 13:45–15:25, TBA.

Optimal Homogeneous Predictors of Rare Events

Benjamin Bobbia

ISAE-SUPAERO
benjamin.bobbia@isae.fr

Co-authors: Stilian Stoev, Victor Verma

Abstract: In this work, we are interested in the estimation of events of the form $\{Y > q\}$ for Y a real random variable and q – a quantile of Y which becomes extreme. We focus on the case where Y is not observed, while we observe a vector of covariates $X = (X_i)_{i=1}^d$. Assuming that the vector $(X, Y) \in R^d \times R$ is regularly varying, we seek optimal predictors of the form $\{h(X) > \tau\}$, where h is a non-negative 1-homogeneous function.

Following Verma et *al*[1] (2025), the function h is chosen in order to maximize the asymptotic precision, or equivalently, the tail-dependence coefficient

$$\lambda(Y, h(X)) := \lim_{\tau\to\infty} P(Y > \tau | h(X) > \tau),$$

under the calibration constraints $P(Y > \tau) \sim P(h(X) > \tau)$, $\tau \to \infty$. We obtain h as solution of a calculus of variations optimization problem involving an integral functional of the spectral (aka angular) measure of (X, Y). The general results are illustrated with the spectral Dirichlet and multivariate logistic models, as well as, related spatial random measure and sup-measure models, where the optimal predictors can be derived explicitly.

Session, Time and Place: Contributed Session 4.2, Thursday 26 June 2025, 13:45–15:25, TBA.

Optimal Prediction of Extreme Events: Characterizations and Examples

Stilian Stoev

University of Michigan, Ann Arbor
sstoev@umich.edu

Co-authors: Victor Verma, Yang Chen

Abstract: Let (X, Y) be a pair of a random vector $X = (X_i)_{i=1}^d$ and a scalar random variable Y, interpreted as the covariates and response, respectively. The *optimal* prediction of the event $\{Y > F_Y^{\leftarrow}(p)\}$ in terms of X, for a fixed or an extreme quantile level $p \in (0, 1)$ is a fundamental problem. For a fixed $p \in (0, 1)$, we establish a general Neyman-Pearson type characterization of the predictors $\{h(X) > F_{h(X)}^{\leftarrow}(p)\}$, which maximize the *precision*

$$\lambda_p(Y, X) := P[Y > F_Y^{\leftarrow}(p) | h(X) > F_{h(X)}^{\leftarrow}(p)].$$

These *optimal predictors*, as $p \uparrow 1$, yield a natural notion of asymptotic optimality in extreme event prediction. In fact, the limit, referred to as the *optimal extremal precision*:

$$\lambda(Y, h(X)) = \lim_{p \uparrow 1} \lambda_p(Y, X)$$

maximizes the tail-dependence coefficient between Y and all measurable functions $g(X)$ of the covariates. We illustrate this framework by deriving closed-form optimal (extremal) predictors in linear time series, bi-variate max-stable, and Archimedean copula models with completely monotone generators.

References:

1. Verma, V., Stoev, S., & Chen, Y. (2025). On the optimal prediction of extreme events in heavy-tailed time series with applications to solar flare forecasting, *Journal of Time Series Analysis*, to appear. https://arxiv.org/abs/2407.11887.

Session, Time and Place: Contributed Session 4.2, Thursday 26 June 2025, 13:45–15:25, TBA.

Optimal Prediction of Extreme Events in Heavy-tailed Time Series with Applications to Solar Flare Forecasting

Victor Verma

University of Michigan
vkverma@umich.edu

Co-authors: Stilian Stoev, Yang Chen

Abstract: The prediction of extreme events in time series is a fundamental problem arising in many financial, scientific, engineering, and other applications. In this talk, I will present a general Neyman-Pearson-type characterization of optimal extreme event predictors in terms of density ratios. This characterization yields new insights and several closed-form optimal extreme event predictors for additive models. These results naturally extend to time series; I will discuss optimal extreme event prediction for heavy-tailed autoregressive and moving average models. Using a uniform law of large numbers for ergodic time series, we have established the asymptotic optimality of an empirical version of the optimal predictor for autoregressive models. Using multivariate regular variation, we have obtained expressions for the optimal extremal precision in heavy-tailed infinite moving averages, which provide theoretical bounds on the ability to predict extremes in this general class of models. I will close by describing the application of our theory and methodology to the important problem of solar flare prediction. Our results demonstrate the success and limitations of long-memory autoregressive as well as long-range dependent heavy-tailed FARIMA models for the prediction of extreme solar flares.

Session, Time and Place: Contributed Session 4.2, Thursday 26 June 2025, 13:45–15:25, TBA.

Neural Bayes Inference for Complex Bivariate Extremal Dependence Models

Lídia André

University of Namur
lidia.andre@unamur.be

Co-authors: Jennifer Wadsworth, Raphaël Huser

Abstract: Generating data from a model is often straightforward and computationally efficient, even when the likelihood function cannot be formulated or when its evaluation is computationally costly. This is the case for several models in the multivariate extremes literature, particularly for the most flexible tail models, including those that interpolate between the two key dependence classes of 'asymptotic dependence' and 'asymptotic independence'. To aid their implementation, we investigate likelihood-free approaches which leverage neural networks to performing inference. We explore the properties of neural Bayes estimators for parameter inference for several flexible yet computationally expensive models. Owing to the absence of likelihood evaluation in the inference procedure, classical information criteria cannot be used to select the most appropriate model. To overcome this, we propose using neural networks as neural Bayes classifiers for model selection. We provide an amortised likelihood-free model selection and inference toolkit, whereby the best model is selected, and its parameters are subsequently estimated using neural point estimation. We show that our classifiers can obtain fast and reliable estimates of the pairwise extremal behaviour of changes in horizontal geomagnetic field fluctuations at three distinct locations.

Session, Time and Place: Contributed Session 4.3, Thursday 26 June 2025, 13:45–15:25, TBA.

Non-stationarity and Uncertainty in Design Life Level for Extreme Temperatures

Occitane Barbaux

Autorité de sûreté nucléaire et de radioprotection (ASNR),
PSE-ENV/SCAN/BEHRIG, France
occitane.barbaux@umr-cnrm.fr

Co-authors: Philippe Naveau, Aurélien Ribes, Nathalie Bertrand

Abstract: This study investigates the inference of design life levels for extreme temperatures, with two primary objectives.

First, it introduces a single index that captures large values in time series, even in non-stationary contexts where traditional risk indicators, like return levels, are hard to interpret. The proposed equivalent reliability (ER) level corresponds to classical return levels in stationary settings but differs in non-stationary ones. It ensures the probability of all observations staying below the ER during a given design period, offering interpretability from a safety perspective. Second, it quantifies uncertainties related to this index, including inferential uncertainties and those arising from future climate change, as these uncertainties can increase under extreme conditions. These are integrated into a Bayesian predictive distribution, which remains underused in climatological analyses.

The study then analyzes annual temperature maxima over France from 1850 to 2100 under various emission scenarios using a non-stationary Bayesian hierarchical extreme value model combining 26 climate models with an observational record. Our results suggest that, in a high emission scenario, there is a noticeable risk that temperature over France will exceed 55°C, ie a level that had not been observed so far in any location around the world, suggesting uncontrolled climate change could completely redraw the risk associated with extreme temperature. Future work will extend this analysis to global annual temperature maxima.

Session, Time and Place: Contributed Session 4.3, Thursday 26 June 2025, 13:45–15:25, TBA.

Projecting Climate-driven Mortality Extremes with Nonstationary Generalized Extreme Value Models

Shaleni Kovach

UNC Graduate Student
shaleni@ad.und.edu

Co-authors: Richard Smith

Abstract: The intensification of extreme heat events under climate change poses growing risks to human health, particularly in urban areas. This work leverages Extreme Value Theory (EVT) to characterize and project the evolution of heat-related mortality across the 21st century in the United States. We employ nonstationary Generalized Extreme Value (GEV) models to estimate the annual maxima of daily mortality attributed to heat stress, using historical mortality and meteorological data from 30 U.S. cities and regions.

We use a Bayesian factor model to condition CMIP6 climate projections on regional observations, accounting for inter-model correlations and uncertainty to produce smoothed, region-specific temperature projections under various SSPs. These feed into a nonstationary GEV framework, linking temperature anomalies to changes in extreme event distributions. Our analysis shows substantial increases in extreme mortality by 2100, especially under SSP5-8.5, with likelihood ratio tests and sensitivity analyses confirming the robustness of the nonstationary GEV approach across cities and mortality metrics. This underscores the value of EVT for assessing climate-related public health risks.

Session, Time and Place: Contributed Session 4.3, Thursday 26 June 2025, 13:45–15:25, TBA.

A Statistical Analysis of Elite Marathon Performances

Abigail Mabe

University of North Carolina at Chapel Hill
aemabe@ad.unc.edu

Co-authors: Richard Smith

Abstract: The current running climate is characterized by speculations about elite marathon performances. Recent topics of debate include the legitimacy of Ruth Chepngetich's world record at the 2024 Chicago Marathon, as well as the impact of "super shoes" and other emerging technologies. Although previous journalistic articles and biological studies have covered these topics, they have not yet been analyzed through a statistical lens. In this study, we gathered approximately 30 years of data from the top 20 finishers of the Chicago, Boston, London, and Berlin marathons. Utilizing the r largest order statistics method, derived from the Generalized Extreme Value distribution, we modeled various linear and changepoint trends on these elite marathon performances. The statistical likelihood of each trend was evaluated through the method of maximum likelihood, likelihood ratio tests, and Bayesian analysis. After analyzing the most likely trends, we calculated endpoints that support the legitimacy of Ruth Chepngetich's world record, while providing a lack of evidence that super shoes are significantly changing the trajectory of elite marathon performances.

Session, Time and Place: Contributed Session 4.3, Thursday 26 June 2025, 13:45–15:25, TBA.

Graphical Models for Stable Lévy Processes Based on The Hüsler-reiss Exponent Measure

Florian Brück

University of Geneva
florian.brueck@unige.ch

Co-authors: Sebastian Engelke, Stanislav Volgushev

Abstract: We propose a flexible dependence model for stable distributions and their corresponding Lévy processes, whose construction is based on the Hüsler-Reiss exponent measure. Our model is a graphical model for the associated Lévy process, where the graph encodes conditional independence of the sample paths. We show that the restrictions of the Lévy process to an arbitrary orthant are in the domain of attraction of the Hüsler-Reiss distribution and propose a sparsistent estimator for its parameters. A similar model can also be used to describe extremes with the possibility of very small or large observations, as often encountered in financial applications.

Session, Time and Place: Contributed Session 4.4, Thursday 26 June 2025, 13:45–15:25, TBA.

Asymptotic Theory for The Estimation of Brown-resnick Processes Via Composite Likelihood

Hank Flury

University of North Carolina at Chapel Hill
fluryh@unc.edu

Co-authors: Richard Smith, Jan Hannig

Abstract: The Brown-Resnick process describes the limit of the maxima of scaled Gaussian processes. Our work assumes the use of the block maxima approach, in which a sample of size n is broken in to k blocks of size m, and the maximum is taken of each block. We assume a Gaussian process is observed, and that we apply the block maxima method in an attempt to estimate the limiting Brown-Resnick process. We extend previous results on the asymptotic distribution of the MLE of the Hüsler-Reiss distribution, which is the distribution of any two points of a Brown-Resnick process. We utilize a composite likelihood approach, in which the full likelihood of the data is replaced by a product of pairwise likelihoods. This allows us to estimate parametric models for the variogram of the Brown-Resnick process by the composite maximum likelihood estimator. We extend previous results by allowing for $N > 2$ sites, as well as estimation under an infill asymptotics regime. We characterize the asymptotic bias and variance in each of these settings and provide a Central Limit Theorem result for the asymptotic distribution of the estimator. We also provide simulations which support our theoretical derivations, and provide guidance for the practical application of our results. Possible data applications for our work are to spatial modeling of extreme weather events and natural phenomena.

Session, Time and Place: Contributed Session 4.4, Thursday 26 June 2025, 13:45–15:25, TBA.

Structured Inference for Hüsler-Reiss Graphical Models based on Intrinsic Gaussian Markov Fields

Peng Zhong

University of New South Wales, Sydney
peng.zhong@unsw.edu.au

Co-authors: Boris Beranger, Scott Sisson

Abstract: Since the introduction of extremal graphical models, various methods have been proposed to learn the underlying dependence structure using graphical lasso-based algorithms. Unlike conditional independence in Gaussian graphical models, extremal conditional independence is defined through the L_∞ risk functional by conditioning on single sites being large. Moreover, the extremal graph must be connected, as full independence is not permitted. However, current inference methods for graphical extremes often fail to guarantee connectivity, frequently yielding multiple disconnected components.

Additionally, these graphical lasso-based approaches rely on empirical estimates of either the variogram matrix or a aggregated covariance matrix, which can be unintuitive. In this talk, we introduce a novel construction for the Hüsler-Reiss graphical model based on intrinsic Gaussian Markov fields of order 1. We demonstrate that our proposed model effectively serves as a proxy for learning the graphical dependence structure, while our inference method ensures connectivity through penalization. Furthermore, when full independence is assumed between disconnected components, our approach can accommodate it yielding the exactly disconnected components. We validate the performance of our method through simulation studies and a real data application.

Session, Time and Place: Contributed Session 4.4, Thursday 26 June 2025, 13:45–15:25, TBA.

Extremes of Ship Motions and Loads in Irregular Waves: Specific Features

Vadim Belenky

David Taylor Model Basin (NSWCCD)
vadim.belenky.civ@us.navy.mil

Co-authors: Kenneth Weems, Vladas Pipiras

Abstract: The talk is intended to give a general overview of application of extreme value analysis for data associated with motion of a ship in irregular waves and associated processes of acceleration and loads. Characterization of extreme values of these processes is critically important for engineering decision for both design and operation of a ship.

Ship motion is a six-dimensional stochastic process characterizing solid body movement in space. Acceleration at a given point is a three-dimensional stochastic process, derived from the motions. Loads are internal forces at a given section of a ship (characterized by a value of force that needs to be applied to keep dynamic equilibrium if a part of a ship on one side of the given section is mentally removed.) Practical engineering, however, usually considers one or two (rarely three) components of these processes simultaneously. More details on how these processes are defined will be given in the talk.

The most prominent feature of both motions and loads processes is substantial changing of underlying physics with departure from the mean: i.e. substantially different mathematical models are required for a single standard deviation vs. three standard deviations. The problem is exacerbated by unpractical computational costs of brute force application of time-domain numerical simulation of required engineering fidelity. The talk will give a brief overview of these models and other sources of data on motions and loads.

As a result of this change of underlying physics, direct characterization of the tail of distribution encounters significant difficulties, manifesting themselves in a confidence interval of unpractical width. To overcome this difficulty, a physics-inform approach is used, i.e. reduced-order mathematical model, describing basic physics of the process, is included in statistical procedure. It could be as simple as just prescribing a heavy tail for large amplitude roll motion, based on consideration of well-known to mariners a "hanging roll phenomenon", described with a piecewise-linear ordinary differential equation. The talk will include several examples of application of physics-informed approach.

A separate problem is a validation of extreme value methods, appropriate for confident engineering application. The idea is to use a reduce-order numerical simulation tool that incorporate only basic physics (i.e. only qualitatively valid) to produce a sample of large volume where extreme values can be observed. A small subset of this data is then used as an input for physics-informed technique; the result of both computations are compared. The talk will give several examples of such validation campaigns.

Finally the talk will be completed with an overview of regulatory and procedural application of some of the applications of extremes to ship motions in waves.

Session, Time and Place: Contributed Session 5.1, Friday 27 June 2025, 10:45–12:25, TBA.

Extreme Ship Response Estimates with Neural Networks

Samuel Edwards

Naval Surface Warfare Center - Carderock Division
samuel.j.edwards33.civ@us.navy.mil

Abstract: In both ship design and operation, one of the most important considerations is extreme events. Although there is generally no impact on day-to-day routine operations in calm seas, knowledge of ocean wave conditions that may lead to extreme responses is imperative for safe ship operation. Addressing extreme event behavior involves understanding of the stochastic wave conditions as well as non-linear ship hydrodynamics. Consequently, prediction of extreme responses is challenging due to both the stochastic and nonlinear nature of rare events.

The most straightforward approach to estimating extremes of stochastic non-linear systems is through Monte Carlo simulations. However, most simulation tools with adequate fidelity have significant computational cost particularly when considering potential rare events over long time durations. Extrapolation methods based on Weibull distributions can be explored with a limited dataset. However, this approach requires knowledge of the response with particular focus on the tail of the distribution.

Other methods to identify extreme behavior efficiently without overextending assumptions have been developed. One such method is the Design Loads Generator (DLG) [1, 2]. DLG was initially developed for linear systems with stochastic Gaussian input, along with modified phase distributions based on Extreme Value Theory, to generate ensembles of extreme realizations for a given return period.

A recent method employs a fast lower-fidelity simulation tool that retains major nonlinearities to identify extreme conditions, and then simulates these conditions with a higher-fidelity tool [3]. In this framework, a surrogate model does not need to be identified but requires a high level of correlation at the peaks between the two simulation tools.

In a recent study [4], fast lower-fidelity simulations were augmented with a neural network. Large events were identified in each lower-fidelity realization with a short time window applied around each large value. A neural network was then trained to transform these lower-fidelity "snippets" to a corresponding higher-fidelity realization. The time series maxima probability distributions generated with the higher-fidelity simulation tool were captured by the neural network approach with only using a fraction of the data.

The study proposed here seeks to further develop and investigate the "snippet" based approach in a multi-fidelity framework with neural network corrections. The importance of location and availability of the rare events within the snippets will be examined. This study will more formally investigate the limits of the snippet based approach.

Session, Time and Place: Contributed Session 5.1, Friday 27 June 2025, 10:45–12:25, TBA.

Multi-Fidelity Monte Carlo Estimation with Applications to Extremes

Minji Kim

University of North Carolina at Chapel Hill
mkim5@unc.edu

Co-authors: Brendan Brown, Vladas Pipiras

Abstract: In a multi-fidelity setting, data are often available from multiple models of varying fidelity, those with higher fidelity carrying higher costs. In such contexts, a larger amount of low- fidelity data can be obtained and leveraged to make more efficient inference about quantities of interest, e.g. the mean, for high-fidelity variables. In this talk, we focus on methods aimed at more efficiently fitting parametric models to high-fidelity data. Specifically, we consider three multi-fidelity parameter estimation methods that utilize moment-based and likelihood-based approaches. The proposed methods are illustrated on several parametric models, with the focus on parametric families used in extreme value analysis. An application is also provided concerning quantification of occurrences of extreme ship motions generated by two computer codes of varying fidelity. Time permitting, we shall also discuss the setting of non-parametric distribution estimation, leveraging the larger amount of available low-fidelity data.

Session, Time and Place: Contributed Session 5.1, Friday 27 June 2025, 10:45–12:25, TBA.

Choosing the Threshold in Extreme Value Analysis

Léo Belzile

HEC Montréal
leo.belzile@hec.ca

Co-authors: Anthony C. Davison, Sonia Alouini

Abstract: We provide an extensive review of threshold selection mechanisms for peak over threshold analysis, including semiparametric methods based on Hill's estimator, visual diagnostics, goodness of fit tests, and extended generalized Pareto models. Starting with the statistical properties underlying the various proposals, we provide critical assessment of methods strengths and weaknesses. We consider the problem of multiple testing of nested hypothesis and automation of various threshold selection procedures. An extensive simulation study under various tail regimes, with serial dependence and varying sample sizes, allows us to identify the most promising methodologies.

Session, Time and Place: Contributed Session 5.2, Friday 27 June 2025, 10:45–12:25, TBA.

Fitting a Peaks-over-Threshold Model with Survey Weights, with Application to Blood Pressure Control in the US

Zoe Haskell-Craig

New York University School of Global Public Health
zjh235@nyu.edu

Co-authors: Yajun Mei, Rebecca Betensky

Abstract: After decades of improvement in population mean blood pressure (BP) in the U.S., researchers have observed an uptick in mean BP values in recent years. Understanding these trends requires accurate modeling of both the center and the tail of the BP distribution; a shift in the entire distribution may indicate socioeconomic trends affecting the whole population, while changes in the high-BP tail may represent changing access to clinical care. The Peaks-over-threshold (POT) approach is well-suited for modeling distribution tails parametrically, yet, BP data is often collected using complex survey designs. Failing to account for survey weights can lead to biased parameter estimates.

We address this challenge by employing pseudo maximum likelihood estimation (PMLE) to incorporate the survey weights in estimating the parameters of the Generalized Pareto Distribution (GPD). We develop this approach both for procedures that fit the GPD to the distribution tail (above a fixed threshold) and for mixture models that treat the threshold as an unknown parameter separating the tail distribution (GPD) from a bulk distribution below. Analytically, we determine conditions under which neglecting survey weights may or may not lead to bias in GPD parameter estimates. In particular, estimates of the shape parameter may still be unbiased if tail observations share similar weights. We demonstrate the PMLE approach through simulations and application to BP data in NHANES.

Session, Time and Place: Contributed Session 5.2, Friday 27 June 2025, 10:45–12:25, TBA.

Assessing The Size of Spatial Extreme Events Using Local Statistics Based on Exceedance Regions

Thomas Opitz

INRAE
thomas.opitz@inrae.fr

Co-authors: Ryan Cotsakis, Elena di Bernardino

Abstract: Extreme events arising in georeferenced processes can take various forms, such as occurring in isolated patches or stretching contiguously over large areas, and can further vary with the spatial location and the extremeness of the events. We use excursion sets above threshold exceedances in data observed over a two-dimensional grid of rectangular pixels to propose a general family of coefficients that assess spatial-extent properties relevant for risk assessment, and study candidate coefficients from this family. These coefficients are defined locally and interpreted as a spatial distance from a reference site where the threshold is exceeded. We develop statistical inference and discuss robustness to boundary effects and resolution of the pixel grid. To statistically extrapolate coefficients towards very high threshold levels, we formulate a semiparametric model and estimate a parameter characterizing how coefficients scale with the quantile level of the threshold. Such coefficients can be useful for exploratory analysis but also as covariates in generative models for spatial random fields with non-stationary dependence. The new coefficients are illustrated through simulated data, as well as in an application to gridded daily temperature in continental France. We find notable differences in estimated coefficient maps between climate model simulations and observation-based reanalysis.

Session, Time and Place: Contributed Session 5.2, Friday 27 June 2025, 10:45–12:25, TBA.

Asymptotics of Peaks-over-threshold Estimators in Long Memory Linear Time Series

Ioan Scheffel

University of Stuttgart
ioan.scheffel@mathematik.uni-stuttgart.de

Co-authors: Marco Oesting, Gilles Stupfler

Abstract: In this talk, we consider peaks-over-threshold (POT) estimators for extremes of long memory linear time series. As these time series are not beta-mixing, classical asymptotic results on POT estimators are not applicable. We adapt a reduction principle for subordinated long memory linear time series to our setting. Thus we prove a central limit theorem for POT estimators, including the Hill estimator. We obtain convergence to stable limit distributions with different rates for light and heavy tails.

Session, Time and Place: Contributed Session 5.2, Friday 27 June 2025, 10:45–12:25, TBA.

A Normed Vector Space of Equivalence Classes of Tail Equivalent Regularly Varying Random Variables

Kenneth Broadhead

Colorado State University
kenneth.broadhead@colostate.edu

Co-authors: Daniel Cooley

Abstract: We construct a normed vector space consisting of equivalence classes of tail equivalent regularly varying random variables on a probability space. We begin with a treatment of univariate regular variation whose starting point is a suitable normalizing function, $b(s)$, that can normalize the tail of a (classically) regularly varying random variable. This gives rise to the notion of an asymptotic scale with respect to a given normalizing function. For a fixed normalizing function $b(s)$, we then consider the space V_b, of all random variables on a probability space whose tails can be normalized (even trivially) by $b(s)$. The notion of asymptotic scale can then be used to form a quotient vector space, whose non-zero elements are equivalence classes of tail equivalent regularly varying random variables. Additionally, it is shown that this same structure can be achieved by further developing the notion of asymptotic scale, giving rise to a semi-norm on V_b, and hence a proper norm on the quotient space.

Session, Time and Place: Contributed Session 5.3, Friday 27 June 2025, 10:45–12:25, TBA.

Learning Extrapolating Representations

Patrick Kuiper

Duke University
patrick.kuiper@duke.edu

Co-authors: Ali Hasan, Jose Blanchet, Vahid Tarokh

Abstract: High-dimensional images and measurements may indicate extreme behavior, such as advanced stages of cancer or material failure, without existing values of the data exhibiting large magnitudes. The brain of a practitioner can extract latent features from these images that indicate the underlying extreme behavior. Classically, extreme value theory provides a theoretical framework for modeling extremeness of magnitude of measurements and is used to model the tail events of a distribution using limited observations. Motivated by the above, we propose a framework for learning representations from images that are amenable to analysis using classical extreme value theory. Specifically, we use the max-stability property of extreme value distributions to design representation learning techniques such that the resulting representations in the feature domain lend themselves to analysis by extreme value theory and can model images of extreme characteristics in the image domain. This enables our method to extrapolate to observations of extreme behavior in the image domain. We then extend our method to infinite dimensional observations such as time series. Numerical results and theoretical analysis are provided demonstrating the performance of our modeling approach.

Session, Time and Place: Contributed Session 5.3, Friday 27 June 2025, 10:45–12:25, TBA.

Principal Component Analysis for Max-stable Distributions

Felix Reinbott

Otto-von-Guericke University Magdeburg
felix.reinbott@ovgu.de

Co-authors: Anja Janßen

Abstract: Principal component analysis (PCA) is one of the most popular dimension reduction techniques in statistics and is especially powerful when a multivariate distribution is concentrated near a lower-dimensional subspace. Multivariate extreme value distributions have turned out to provide challenges for the application of PCA since their constraint support impedes the detection of lower-dimensional structures and heavy-tails can imply that second moments do not exist, thereby preventing the application of classical variance-based techniques for PCA. We adapt PCA to max-stable distributions using a regression setting and employ max-linear maps to project the random vector to a lower-dimensional space while preserving max-stability. We also provide a characterization of those distributions which allow for a perfect reconstruction from the lower-dimensional representation. Finally, we demonstrate how an optimal projection matrix can be consistently estimated and show viability in practice with a simulation study and application to a benchmark dataset.

Session, Time and Place: Contributed Session 5.3, Friday 27 June 2025, 10:45–12:25, TBA.

Nonparametric Estimation of The Spectral Density of A High-dimensional Extreme Value Distribution

Matteo Schiavone

University of Padova
matteo.schiavone@phd.unipd.it

Co-authors: Simone Padoan, Stefano Rizzelli

Abstract: The spectral measure provides an infinite-dimensional characterization of dependence structures in multivariate extreme value theory, yet its non-parametric density estimation remains notoriously challenging. Existing methods are effective in low dimensions but struggle to scale efficiently as dimensionality increases.

We propose a novel inferential procedure for estimating spectral density in arbitrary dimensions. Our approach unfolds in two key stages: first, we establish an analytical relationship between the unit-Pareto spectral density and its counterpart under multivariate regular variation, thereby circumventing the moment constraint. Second, we leverage the isometric log-ratio transformation from compositional data analysis to map the problem from the geometrically constrained simplex to an unconstrained Euclidean space R^{d-1}.

This methodology preserves the intrinsic geometry of the problem while enhancing computational feasibility. It retains the flexibility of non-parametric approaches while overcoming the dimensional barriers that have historically constrained the analysis of multivariate extremal dependence.

Session, Time and Place: Contributed Session 5.3, Friday 27 June 2025, 10:45–12:25, TBA.

Poster Abstracts

Assessing Extreme Droughts: An Extreme Value Theory Analysis of Common Drought Indices

Paolo Besana

UCLouvain
paolo.besana@uclouvain.be

Co-authors: Hugues Goosse, Johan Segers, Francesco Ragone

Abstract: Extreme droughts represent a significant threat to both ecosystems and socio-economic systems. Although numerous indices have been developed for drought monitoring and prediction, robust risk quantification remains challenging.
Given their widespread use as drought indices, we focus our analysis on the Standardized Precipitation Index (SPI) and the Standardized Precipitation- Evapotranspiration Index (SPEI) to examine their tail behaviour within the framework of Extreme Value Theory. Specifically, we evaluate the performance of these indices across multiple datasets using a suite of goodness-of-fit tests—including Anderson–Darling and Cramér–von Mises—as well as memory and correlation analyses. A particular emphasis is placed on the implications of the underlying assumptions used to construct these indices on the estimation of key quantities, such as the intensity and return times of extreme drought events.
We quantify the associated uncertainties and discuss how these influence the interpretation of results based on different indices. Finally, we propose targeted methodological refinements to mitigate potential biases and enhance the accuracy of risk quantification.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Control Variates for Variance-Reduced Extreme Value Index Estimators

Louison Bocquet-Nouaille

ONERA/DTIS & Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, Toulouse, France
louison.bocquet_nouaille@onera.fr

Co-authors: Benjamin Bobbia, Jérôme Morio

Abstract: The estimation of the Extreme Value Index (EVI) is fundamental in extreme value analysis but suffers from high variance due to reliance on a few extreme observations.

We propose a control variates based transfer learning approach in a semi-supervised framework: a small set of coupled target and source observations is available, combined with a large amount of unpaired source data. By expressing the Hill estimator on the target variable as a ratio of expectations, we apply approximate control variates to both the numerator and denominator to reduce variance without introducing bias. The scaling parameters of both control variates are jointly optimized to minimize the overall variance of the ratio.

The proposed method can be extended to other EVI estimators with smaller bias, such as the moment estimators, by applying the control variates technique to both the first and second-order moments. Strong variance reduction can be achieved when target and source data are tail dependent.

References:

1. Kim, M., Brown, B., & Pipiras, V. (2024). Parametric multi-fidelity Monte Carlo estimation with applications to extremes. *arXiv preprint* arXiv:2410.08523.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Multivariate Regular Variation and the Calibration of p-Value Combination Tests

Parijat Chakraborty

The University of Michigan
cparijat@umich.edu

Co-authors: Stilian Stoev, Richard Guo, Kerby Shedden

Abstract: Combining many dependent p-values into a single one is a popular methodology in multiple testing. It turns out that a very general family of heavy-tailed combination tests can be thoroughly understood in the framework of multivariate regular variation. The precise (asymptotic) calibration properties of combination tests can be characterized in terms of the exponent measure of regular variation. We first recover all existing results on the Cauchy combination test, a test which is widely used to control the family wise error rate while ensuring robustness to dependence. We prove that these results correspond to the regime of asymptotic independence and that asymptotic dependence can break the calibration of the Cauchy combination test, though it always remains honest. To amend for that we introduce the Pareto combination test and show that it is universally calibrated under both tail-independence and tail-dependence. Simulation studies illustrate the calibration results and the power properties of the new Pareto combination test.

References:

1. Liu, Y., & Xie, J. (2020). Cauchy combination test: a powerful test with analytic p-value calculation under arbitrary dependency structures. *Journal of the American Statistical Association*, 115(529), 393–402.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Heatwave Attribution over Europe

Kamal Gasser

LIDAM, ISBA, UCLouvain
kamal.gasser@uclouvain.be

Co-authors: Francesco Ragone, Johan Segers

Abstract: We are investigating if the rise in global mean surface temperature caused by human forcing has resulted in a change in temperature distribution, hence increasing the frequency of heatwaves. Heatwaves manifest themselves in both spatial and temporal dimensions ; nevertheless, the majority of attribution studies concentrate on individual locations and analyze heatwaves by calculating specific indices to capture their spatiotemporal dependency. However, this technique seems inadequate since indices do not accurately represent the real events. We aim to address this gap by using advanced statistical models and concentrating our attribution analysis on the whole of Europe. We use a clustering technique using a distance measure suitable for extremes to pinpoint spatial regions where concurrent severe temperatures occur simultaneously. Then, we formulate a flexible non-stationary space-time extreme value model that accommodates diverse forms of asymptotic dependency. Lastly, we use our approach to diverse climate model outputs to calculate the proportion of attributable risk.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Space Mining with Extremes

Alfredo Gimenez Zapiola

Politecnico di Milano & EPFL (visiting)
alfredo.gimenez@polimi.it

Co-authors: Anthony C. Davison, Simone Vantini

Abstract: Before launching a mission to mine some celestial body, it is imperative to possess a reliable mineralogical map of it. We propose to solve this by applying (Spatial) Extremes to an available hyper-spectral image of the body of interest. We work with the Cuprite image (NASA). At every pixel, a discretely sampled reflectance function is observed, so correlations are obtained with respect to the known spectra of candidate minerals. Exceedance over a high threshold is indicative of their presence. Therefore, we see the data as a single realisation of a multivariate extremal process, where the exceedances of different minerals may be observed at different pixels. This poses an interesting challenge as the field of Spatial Extremes typically works with n i.i.d realisations of an univariate extremal process. We address this in a two-fold manner: (1) GPD-GAMs are used to model (spatially) the exceedances of each candidate mineral; (2) with such margins, D-vine copulae constructions are employed with bivariate Wadsworth et. al, Tawn, and other copulas to model asymptotic dependence, asymptotic independence and independence across the $d = 12$ candidate minerals that present exceedances. Finally, mineralogical maps are outputted using the estimated joint distribution. Model diagnostics as well as a qualitative comparison with an available mapping of the same image verify model adequacy. Main references:

1. Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). *Pair-copula constructions of multiple dependence.* Insurance: Mathematics and economics, 44(2):182–198.
2. Wadsworth, J. L., Tawn, J. A., Davison, A. C., and Elton, D. M. (2016). *Modelling across extremal dependence classes.* Journal of the Royal Statistical Society Series B: Statistical Methodology, 79(1):149–175.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Scalable Causal Discovery for Extremes via Partial Tail Correlation and the PC Algorithm

Yan Gong

Harvard T.H. Chan School of Public Health, Boston, United States
yangong@hsph.harvard.edu

Abstract: The PC algorithm (Spirtes et al., 2000) is a cornerstone for causal structure learning in high-dimensional directed acyclic graphs (DAGs), typically using Gaussian conditional independence tests. However, its performance is compromised under extreme-value conditions where dependencies are driven by heavy-tailed phenomena. In response, we present a scalable extension of the PC algorithm by incorporating the partial tail correlation coefficient (PTCC) (Lee and Cooley, 2022; Gong et al., 2024), an innovative metric designed to assess conditional dependence in tail regions. This enhancement maintains the computational efficiency of traditional PC-type algorithms while markedly improving accuracy in identifying tail-driven dependencies. The PTCC-augmented PC algorithm significantly outperforms standard Gaussian methods in identifying sparse causal structures under heavy-tailed noise, as confirmed by extensive simulations and illustrated by real-world applications in cryptocurrency markets, cancer gene regulatory networks, teleconnections and global climate extremes. This work leverages extreme-value statistics with graphical causal inference, providing a robust tool for high-dimensional tasks where extremes carry crucial signals.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Structured Multivariate Extreme Value Models for Flood Risk Estimation

Kristina Grolmusova

STOR-i CDT, Lancaster University
k.grolmusova1@lancaster.ac.uk

Co-authors: Jennifer Wadsworth and Thordis Thorarinsdottir

Abstract: We exploit the new framework of multivariate geometric extremes to develop statistical methodologies that will allow analysis of extremes of river flows at multiple locations on a river network simultaneously. Current methodologies within this framework are limited to a relatively low number of dimensions. This is insufficient for the purposes of flood risk estimation, since the number of gauging stations on a river network is often of the order 10-20. In order to create a parsimonious model in higher dimensions, we use recent work to define the gauge function which describes the limit shape of the scaled sample cloud in a structured way, using the river network. We compare the performance of the structured gauge functions to alternatives for capturing the extreme flows of a river network near Preston, UK.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

On Predicting the Likelihood of High-Frequency Extreme Price Movements

Philippe Hübner

HEC-Liège, University of Liège, Belgium
phubner@uliege.be

Co-authors: Julien Hambuckers, Philippe Hübner

Abstract: The ability to better understand and predict high-frequency extreme price movements (EPM) is crucial for market stability and investor confidence. In this paper, rather than focusing on the EPM themselves, we study the time-varying characteristics and determinants of their distribution. To do so, we assume that it can be well approximated by the distribution for the maxima of high-frequency negative returns over small time-intervals. Thus, relying on extreme value theory, we specify for these maxima a Generalized Autoregressive Score (GAS) model with a generalized extreme value (GEV) response distribution and lagged predictors driving the parameters. Moreover, to perform covariate selection, we introduce a L_0-penalized likelihood procedure. We investigate the properties of our estimation procedure in a Monte Carlo simulation. In our empirical application, we use limit order book and transaction data from 2009 to 2019 for 50 NASDAQ-traded stocks and consider a large set of predictors, including liquidity measures, to explain and forecast the distribution of block-maxima of 1-minute negative returns. We show that our model performs better than other benchmark specifications, mainly due to the incorporation of important conditioning variables. From the variable selection procedure and corresponding Shapley values, we show that liquidity measures play a significant role for determining the distribution of future EPM, along with daily uncertainty indicators. We also investigate the dependence of the EPM distribution between the 50 stocks.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Exact Coordinate Descent for High-Dimensional Regularized Robust M-Estimators

Younghoon Kim

Cornell University
yk748@cornell.edu

Co-authors: Po-Ling Loh, Sumanta Basu

Abstract: We develop an exact coordinate descent algorithm for high-dimensional adaptive robust M-estimators. This algorithm ensures fast and guaranteed convergence for high-dimensional penalized robust regression, particularly when the model is assumed to be sparse. Unlike existing composite gradient methods or second-order approximation methods, the proposed algorithm does not require initial values close to the optimal solution. In line with conventional coordinate descent algorithms, we also propose a screening rule to accelerate convergence. We formally establish the convergence of the algorithm to the global optimum in the case of Huber loss and to local stationary points for other robust loss functions. Simulation studies on various cases of heavy-tailed predictors and noise and real data application are conducted.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

A Flexible Multivariate Generalized Pareto Model for Extreme Dependence Structures

Mirco Lescart

UNamur
mirco.lescart@unamur.be

Co-authors: Anna Kiriliouk, Philippe Naveau

Abstract: Extreme value analysis in multivariate settings requires models capable of handling both asymptotic dependence and asymptotic independence. We introduce a flexible framework that extends the multivariate generalized Pareto distribution, capable of representing both asymptotic dependence and asymptotic independence. The model is based on a hierarchical construction with independent Gamma-distributed latent variables, allowing for a smooth transition between different dependence regimes. Parameter estimation is performed using the Neural Bayes Estimator, a likelihood-free deep learning-based method that efficiently handles the model's complexity. We illustrate the versatility of the model through an application to extreme rainfall data in France, highlighting its adaptability to heterogeneous dependence structures.

References:

1. Kiriliouk, A., Segers, J., & Rootzén, H. (2019). Peaks over thresholds modelling with multivariate generalized Pareto distributions. *Technometrics*, 61(1), 123-135.
2. Sainsbury-Dale, J., Gandy, A., & Hartigan, J.A. (2024). Likelihood-free inference via neural Bayes estimation. *Journal of the Royal Statistical Society: Series B*, to appear.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Statistical Prediction of Extreme Economic Losses and Fatalities in Europe due to Climate-Related Hazardous Events

Carlotta Pacifici

Bocconi University
carlotta.pacifici@unibocconi.it

Co-authors: Jaroslav Mysiak, Simone A. Padoan

Abstract: In accordance with the European Community's objective of hedging against climate-related risks, we analyze extreme daily economic losses and fatalities occurred in Europe following climate hazardous events between 1980 and 2023. The aim is to disclose the uncertainty behind future extreme events, by providing short-term precautionary indications of tail risks. Extreme losses and fatalities are defined as values exceeding a sufficiently high threshold and are modeled through the continuous and discrete Generalized Pareto distribution (GPD), respectively. GPD parameters are estimated through a Bayesian inferential procedure, whose benefit is to naturally allow to derive a predictive distribution, forming the basis for predicting future extreme events. We estimate the predictive distribution of future extreme losses and fatalities using the posterior predictive distribution and evaluate both the plausibility of the maximum observed sample value and the likelihood of similar or even more extreme events occurring again. We account for nonstationarity by allowing the tail of the conditional distribution of losses and fatalities to vary with the frequency of extremes over time. Both stationary and nonstationary analyses indicate that extreme events comparable to the maximum observed sample value are likely to recur as early as 2026, with even more severe events becoming increasingly plausible over time.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Refining European Extreme Precipitation Return Levels using Regionalized GEV Models

Robert Paulus

UNamur
robert.paulus@unamur.be

Abstract: Extreme precipitation events have increased in frequency and intensity across Europe, notably with days exceeding 150mm of rainfall in countries such as Belgium, Spain, and Slovakia, in 2021 and 2024. This trend has led to substantial human and economic losses, particularly in unprepared urban areas. This study conducts a spatio-temporal analysis of precipitation maxima to refine return level estimates, crucial for risk assessment. We compare bayesian and frequentist statistical models, including techniques like Bayesian hierarchical models and profile likelihoods, to assess their effectiveness in extreme precipitation predictability. Rainfall exhibits a localized nature with low spatial dependency; we take advantage of this by applying regionalized, univariate GEV models. By applying this framework to European precipitation data, which captures localized rainfall patterns, we aim to enhance the precision of return level estimates in different sub-regions. This, in turn, will improve predictive reliability for extreme precipitation events, aiding in better risk assessment and preparedness strategies across Europe.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Bootstrap in Extreme Value Theory

Pietro Scanzi

Bocconi University
pietro.scanzi@phd.unibocconi.it

Co-authors: Simone Padoan, Stefano Rizzelli

Abstract: An important goal in the univariate extreme value theory (EVT) is to estimate the tail index, extreme quantiles and the upper-endpoint of the original distribution. In the literature, several estimators have been proposed and their asymptotic theory investigated (De Haan & Ferreira, 2006). Unfortunately, in some cases the asymptotic variance of these estimators is intricate or unavailable in closed form, making uncertainty quantification hard to achieve. Recently, De Haan & Zhou (2024) have explored the non-parametric bootstrap in both block maxima (BM) and peaks over threshold (POT) for assessing the variance of the probability weighted moment (PWM) estimator of the tail index. They developed the bootstrap analog of the asymptotic expansion of the block maxima quantile process and of the tail quantile process to show the consistency of the bootstrapped estimators. In this work, we want to derive parallel results for other tail index estimators, such as the maximum likelihood estimators (MLE), and transfer these findings to the extreme quantile and distribution endpoint estimators. Then, having examined more broadly the strength and weaknesses of non-parametric bootstrap inference, we investigate the theory and simulations also for the parametric bootstrap case.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Multi-Regional Analysis of Antarctic Sea Ice Record Lows

Alexandre Tytgat

Université Catholique de Louvain
alexandre.tytgat@uclouvain.be

Co-authors: François Massonnet, Anna Kiriliouk

Abstract: Antarctic sea ice has reached historically low extents in recent years, with 7 of the 10 lowest measurements occurring in the past decade since satellite monitoring began in 1979. This important decline threatens both regional ecosystems and global climate stability. To shed light on this situation, we propose a multi-regional statistical extreme value analysis of Antarctic sea ice extent minima. Specifically, we take a block-minima approach in which the location and scale are parametrized using climate covariates linked to hypothesized sea ice dynamics. By applying this methodology to key Antarctic sectors and synthesizing the regional results, we obtain estimations for return levels of regional record lows and identify key drivers of sea ice decline within the wider Antarctic region.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Spatial Extremes on Domains with Physical Barriers

Chien-Chung Wang

Colorado State University
ccwang@colostate.edu

Co-authors: Ben Shaby

Abstract: We propose a model for spatial extremes what realistically accommodates nonconvex domains. Most existing models assume that tail dependence between locations depends only on their separation. Even if extended to be anisotropic or nonstationary, they still cannot accommodate features between points that may mediate the dependence. This assumption limits their applicability in settings where physical barriers like mountains, coastlines, or islands may attenuate or even block dependence in extremes events.

In this work, we use the idea of a visibility graph to modify an existing spatial random scale model. This results in a tail dependence model that naturally accommodates nonconvex spatial domains, where the nonconvex features are interpreted as physical barriers. Furthermore, the model inherits key properties from the scale mixture like the ability to represent asymptotic independence at long ranges and either asymptotic dependence or asymptotic independence at short ranges.

To assess the effectiveness of our model, we conduct a comprehensive study of tail dependence coefficients under various scenarios. We also evaluate parameter estimation performance, including coverage properties. Finally, we apply our framework to real-world data, demonstrating its utility in capturing spatial extremes in complicated nonconvex domains.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

Combining Observational and Model Data for Spatial Extremes through a Multivariate Gaussian Latent Process

Brian White

UNC Chapel Hill, Department of Statistics and Operations Research
bnw@unc.edu

Co-authors: Brian Blanton, Rick Luettich, Richard L. Smith

Abstract: Extreme coastal events pose significant risks to communities worldwide. Given the sparse nature of in situ sea-level observations, numerical models such as ADCIRC are often employed to simulate these events and derive quantities—like return levels—that summarize the magnitude of rare occurrences. However, while model-derived estimates are valuable on their own, integrating them with observational data is critical for generating realistic spatial fields of extreme value estimates. In this paper, we extend the work of Russell et al. (2020) by introducing a novel data assimilation framework that first computes generalized extreme value parameter estimates at both observational and model sites, and then jointly models these estimates as a six-dimensional Gaussian process. Our approach yields return level estimates—such as the 100-year event—with quantified uncertainty that is uniformly superior to those obtained from either data set independently. This framework advances extreme value analysis by effectively fusing heterogeneous data sources for improved coastal hazard assessment and risk management.

References:

1. Russell, B., Risser, M., , Smith, R.L. and Kunkel, K.E. (2020). Investigating the association between late spring Gulf of Mexico sea surface temperatures and US Gulf Coast precipitation extremes with focus on Hurricane Harvey. *Environmetrics*, 31(2).

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.

In-Degree Distribution of some Directed Preferential Attachment Graphs

Fuwei Yu

University of North Carolina at Chapel Hill
fwy@unc.edu

Co-authors: Mariana Olvera-Cravioto

Abstract: We study the distribution of the in-degree in a broad family of directed pref- erential attachment graphs with random out-degrees, which depending on the shape of the attachment function and the distribution of the out-degrees can span the entire range of heavy-tailed distributions, from semi-exponential (e.g., heavy-tailed Weibull) to regularly varying (e.g. Pareto). The analysis is based on a characterization of the in-degree in terms of a random sum of i.i.d. Markovian birth processes stopped simultaneously at a random time. Our results provide exact asymptotics for the tail distribution which describe the extreme properties of vertices in the graph.

Session, Time and Place: Poster Session, Tuesday 24 June 2025, 17:20–18:30, TBA.