Learning Spatio-Temporal Climate Extremes

Date: Wednesday, May 27, 2026
Venue: Foyer Royal, Louvain-la-Neuve
Address: Place Raymond Lemaire 1, 1348 Louvain-la-Neuve
Webpage: https://exalt-project.github.io/workshop-2026/

Programme

Time	Speaker	Title
09:00	–	Welcome
09:10	Sebastian Engelke	AI weather forecasting: assessing extrapolation and physical consistency
09:50	Hans Van de Vyver	Beyond the observed: a stochastic simulation approach for impact-based storylines
10:30	–	Coffee break
11:00	Ben Youngman	Extreme value models for environmental data
11:40	Melinda Galfi	A large-deviation view on persistent extremes
12:20	–	Lunch
13:50	Alexandre Tytgat	Detecting changes in Antarctic sea ice extent annual minima
14:10	Paolo Besana	Bayesian POT analysis of drought extremes: quantifying the impact of drought index definition
14:30	Kamal Gasser	A spatio-temporal statistical framework for heatwave attribution under climate change
14:50	Robert Paulus	Adaptive regionalization for extreme precipitation: a neural network-weighted independence likelihood approach
15:10	–	Coffee break
15:40	Daniela Castro-Camilo	Attributing changes in extreme events: a causal approach to the tails
16:20	Alexis Boulin	Linear factor models for tail dependence in high dimensions with applications to wind turbine cut-in risk
17:00–18:30	–	Wrap-up & drinks

Sebastian Engelke (University of Geneva)

AI weather forecasting: assessing extrapolation and physical consistency

Abstract

Recent AI weather models have demonstrated superior performance over traditional physics-based numerical weather prediction systems on a range of standard benchmarks. However, important limitations remain, particularly regarding their ability to extrapolate to extreme events and to produce physically consistent forecasts. In this work, we evaluate model performance on a benchmark dataset composed exclusively of record-breaking events and show that, in this regime, physics-based models continue to outperform leading AI models such as GraphCast and Pangu-Weather. We further introduce a novel Turing test for physical consistency in AI forecasts. The proposed framework leverages the predictive spread across an ensemble of pre-trained AI models and is calibrated using principles from conformal prediction. Through experiments on the Lorenz attractor, we demonstrate the effectiveness of the approach in detecting physically implausible dynamics. We apply the method to assess the physical consistency of outputs from state-of-the-art AI weather models.

Hans Van de Vyver (Royal Meteorological Institute of Belgium)

Beyond the observed: a stochastic simulation approach for impact-based storylines

Abstract

Extreme weather events are increasing in frequency and intensity, necessitating robust frameworks for event attribution. However, traditional probability-based attribution often faces challenges such as high uncertainty, subjectivity, and out-of-distribution issues. An alternative strategy involves the use of "storylines"—narratives of future weather that inform stakeholders and the public on how extreme events arise and how severe they may become under different climate scenarios. This approach prioritizes plausibility over exact probability.

We present a stochastic rare-event algorithm based on Fourier resampling, incorporating key concepts from multivariate spatial extremes modelling. A central feature of this algorithm is that the generated datasets maintain the same spatial tail dependence as observed events, accurately reproducing the statistics of damaged regions and storm sizes. The algorithm is fast, easy to implement, and scalable to high-dimensional, dense spatio-temporal datasets.

Finally, we explore new avenues for event attribution by introducing changes in large-scale circulation dynamics, moving beyond the purely thermodynamical focus of the standard World Weather Attribution (WWA) protocol.

Ben Youngman (University of Exeter)

Extreme value models for environmental data

Abstract

Most environmental extremes coincide with rare events. As a result, they typically have limited data representing them. A particularly useful tool in the modelling of extremes is to incorporate physics, or to use statistical approaches that can capture physical behaviour. For example, on a given day in summer, we expect the probability of exceeding 40°C to be similar on consecutive days. From a statistical perspective, this means we expect the extreme value distributions that represent extreme temperatures to be similar on consecutive days. This talk will introduce generalised additive models for extremes, which provide an intuitive and flexible approach to building in such constraints when modelling data. This construction allows more data to be coherently used, partially overcoming data scarcity. Various weather-related examples will be used for illustration, including examples of modelling in R.

Melinda Galfi (VU Amsterdam)

A Large-Deviation View on Persistent Extremes

Abstract

Extreme value theory (EVT) is a well-established statistical framework for estimating the likelihood of rare events. However, because EVT focuses on instantaneous extremes, its applicability is limited when persistence is central - as in heatwaves, cold spells, and other prolonged extreme events.

In this talk, I introduce an alternative viewpoint based on Large Deviation Theory (LDT), applied to the statistics of persistent extremes. Rather than analysing instantaneous maxima or minima, LDT quantifies the probabilities of extremes of finite-time averages - periods during which the mean anomaly remains unusually high or low. This approach naturally captures a defining characteristic of many high-impact climate extremes: their duration.

I will compare EVT and LDT and discuss a recently identified property of extreme events related to large-scale atmospheric circulation anomalies: the typicality of extremes, meaning that persistent extremes exhibit a stronger similarity and coherence in their circulation patterns than moderate anomalies, indicating that they arise from recurrent dynamical states.

Alexandre Tytgat (UCLouvain)

Detecting Changes in Antarctic Sea Ice Extent Annual Minima

Abstract

The recent rapid decline in Antarctic summer sea ice, marked by successive record-breaking minima over the last decade, has fueled the debate over whether the Antarctic sea ice system has undergone a regime shift. Existing studies of Antarctic sea ice focus on monthly and seasonal averages, using methods tailored to the bulk of the distribution to assess changes in mean, variability, and the rarity of recent record lows. Extreme-value modeling of annual minima, and the associated probabilities of record events, remains largely unexplored. At the regional level, analyses further rely on standard sector boundaries that may misrepresent the true spatial structure of sea ice signals, and this probabilistic framework has not been extended to the regional scale.

To address this, we develop a clustering algorithm using the F-madogram distance to partition the Antarctic into regions whose sea ice extent minima exhibit strong dependence. Within each region and for the whole Southern Ocean, we fit both stationary and non-stationary GEV models to characterize distributional changes, test for the presence of breakpoints, and estimate the probability of record-low events. At the Southern Ocean scale, instead of a structural break, our results reveal a sustained increase in variability throughout the observational record, which explains the succession of record lows in the last decade. At the regional level, our clustering reveals robust regions with distinct tail behavior. While most experienced low sea ice extents in the last decade, these events are not individually anomalous within their regional historical context. What is exceptional is their simultaneous co-occurrence across regions.

Paolo Besana (UCLouvain)

Bayesian POT analysis of drought extremes: quantifying the impact of drought index definition

Abstract

Extreme drought events represent a major source of socio-economic and environmental risk, affecting water resources, agriculture, ecosystems, and energy systems across multiple climatic regimes. As such, Extreme Value Theory provides a powerful framework for drought risk assessment through return levels and related tail-based metrics. Numerous indices have been developed to summarise drought characteristics through scalar quantities intended to capture the magnitude and persistence of drought conditions, yet the role of these indices in shaping estimates of extreme behaviour remains poorly understood. In this study, we quantify how the choice of drought index affects inferred extreme behaviour and associated risk metrics over Europe.

Using E-OBS observations and a high-resolution CMIP6 model, we compare two indices based on the same physical variable, namely climatic water balance: a parametric index (SPEI3) and a non-parametric standardised anomaly (S3). A consistent Peaks-over-Threshold framework is applied, combining local Generalized Pareto Distribution fits with Bayesian spatial regularisation via Gaussian Markov Random Fields, allowing robust and spatially

coherent estimation of return levels across the domain.

The obtained results allow us to quantify how alternative definitions of drought measures affect Extreme Value Theory based inference of extremes, including their impact on return level estimation, tail characterisation, and the spatial structure of inferred risk within the Generalized Pareto framework. The analysis provides a statistically consistent basis for comparing these effects across the domain and across different climatic regimes.

These findings have implications for climate risk analysis and policy-relevant applications. We assess how different drought measures influence estimates of extreme risk and discuss the range of behaviours and discrepancies that can emerge under alternative formulations of drought indices. We argue that drought index selection should be treated as an explicit modelling choice rather than a neutral preprocessing step, and that the uncertainty associated with this choice should be explicitly accounted for in drought risk assessment, interpretation, and communication.

Kamal Gasser (UCLouvain)

A spatio-temporal statistical framework for heatwave attribution under climate change

Abstract

We develop a unified statistical framework for attributing heatwaves as spatio-temporal phenomena under climate change. We quantify the impact of anthropogenic forcing on the probability and persistence of heatwaves not captured by standard marginal extreme-value approaches. Our methodology constructs a generative model for daily temperature fields that separates marginal nonstationarity from spatio-temporal dependence. We combine three components: a Bayesian spatial quantile regression model for the bulk of the data; a nonstationary spatial generalized extreme value model for tail behavior; and a copula-based model capturing both asymptotic dependence and independence in the extremes. The framework is applied to the CMIP6 MRI-ESM2 climate model, contrasting factual and counterfactual scenarios for probabilistic attribution. Our results show that the approach captures key heatwave characteristics inaccessible to traditional methods, enabling direct estimation of event-level attribution metrics. Overall, it provides a flexible basis for analyzing and attributing complex climate extremes as space-time objects.

Robert Paulus (UCLouvain)

Adaptive regionalization for extreme precipitation: a neural network-weighted independence likelihood approach

Abstract

Recent European floods and the damages associated with them highlight the societal cost of underestimating extreme rainfall events. Infrastructure design therefore depends on reliable estimates of precipitation return levels, that is, the rainfall amounts associated with rare events. The difficulty is that observational records contain only a limited number of extremes, so fitting an extreme value model separately at each site often produces unstable estimates and wide confidence intervals. To reduce uncertainty it is common to borrow strength across sites through pooling approaches. However, such methods can be biased when extremal behaviour is spatially heterogeneous and spatio-temporally non-stationary, as is common in hydroclimatically diverse regions and under climate change.

We propose a flexible adaptive pooling strategy for space-time non-stationary precipitation extremes. For each target site, we fit a generalized extreme value model using a neural network-weighted independence likelihood, in which the network assigns weights to neighbouring observations based on distributional similarity and spatial distance. This allows the effective region to be learned from the data and to be smooth rather than fixed in space. Through simulation studies, we show that the learned weights adapt well to the strength and scale of the underlying space-time generating process. During training, the method moves from broad pooling towards more localized estimation, exposing a clear bias-variance trade-off. We quantify this trade-off using the RMSE of the predicted spatial GEV parameters across epochs together with Kish's effective sample size, which measures the amount of information effectively borrowed through the learned weights. Because the GEV parameters and regionalization weights are estimated jointly, the resulting regions of influence are spatially smooth, stable across random seeds, and close to the generating fields in simulation.

Daniela Castro-Camilo (University of Glasgow)

Attributing changes in extreme events: a causal approach to the tails

Abstract

How are extreme climate events changing, and why? Answering this requires methods that focus on the tails of distributions rather than their averages. In this talk, I introduce a causal framework based on extreme quantile treatment effects (eQTEs) to quantify how anthropogenic forcing shifts the severity of rare events. Our proposed approach, called the Tail-Calibrated Inverse Estimating Equation (TIEE) framework, integrates causal inference and extreme value theory within a single, stable estimation procedure. By borrowing strength across quantile levels and directly modelling tail behaviour, TIEE enables reliable inference even for very rare events. An application to Alpine precipitation provides evidence of an overall intensification in extreme rainfall after adjusting for atmospheric circulation.

Alexis Boulin (Ruhr University Bochum)

Linear Factor Models for Tail Dependence in High Dimensions with Applications to Wind Turbine Cut-In Risk

Abstract

We develop an interpretable class of high-dimensional tail-dependence models tailored to applications such as assessing joint low-wind risks in wind-energy. A standard tool for describing extremal dependence in high dimensions is the stable tail dependence function. In our framework, it is induced from a latent K-factor model with $K \ll d$, yielding explicit dimension reduction. Allowing for possibly dependent heavy-tailed factors, the model is parametrized by a sparse loading matrix and the lower-dimensional spectral measure of the factors. Under a mild pure-variable structural assumption, we establish identifiability, propose algorithms to estimate both the number of factors and the loading structure, and develop practical rank-based estimation methods via a margin-free tail pairwise dependence matrix. A spatial wind-energy case study demonstrates how the latent factor structure enables scalable estimation of the probability that a large fraction of turbines simultaneously fall below their cut-in wind-speed thresholds, a key quantity for operational planning.