GAME 2026 – Generative AI Modelling for Extreme Events

11–12 June 2026
Room "Sala 30" – Botte B4 – DAMA Tecnopolo, Bologna

Detailed Programme

Day 1 – Thursday, 11 June 2026

Time	Speaker / Details	Title
13:00–13:30	**Registration**	
13:30–14:00	**Opening**	Introduction to the Workshop
14:00–14:30	Miguel de Carvalho	Generative Cascades of Multivariate Extremes
14:30–15:00	Giorgia Ramponi	Multi-agent Learning without Rewards
15:00–15:30	Luis Gutiérrez Inostroza	Bayesian Model Selection for Heavy-Tailed Data via Transformations: Exploring Possibilities for Amortized Inference
15:30–16:00	**Coffee break**	**Group Picture!**
16:00–16:30	Sebastian Engelke	Statistical Extrapolation and Applications to AI Weather Forecasting
16:30–17:00	Maud Thomas	Non-Parametric Simulation of Multivariate Extreme Events via Spectral Bootstrap
17:00–17:30	Lambert De Monte	Multivariate Extremal Index and Temporal Clustering in Multivariate Extremes: A Case Study of Dutch Hydrological Extremes

Day 2 – Friday, 12 June 2026

Time	Speaker / Details	Title
09:30–10:00	Antonio Tirri & Melchiorre Danilo Abrignani	Title TBC
10:00–10:30	Antonello Squintu	Exploring AI-enhanced Methods for the Improvement of Sea Level Warnings in Venice
10:30–11:00	Matteo Angelinelli	HPC for Extreme Events: Use Cases and Opportunities at CINECA
11:00–11:30	**Coffee break**	
11:30–12:00	Olivier Lopez	Wasserstein GAN with Data Augmentation for the Generation of Multivariate Extremes
12:00–12:30	Alison Peard	Simulating Spatially Resolved Multi-Hazards with Generative Deep Learning

12:30–13:00	Ashley Turner	Scale-Preserving Autoencoders for High-Dimensional Heavy-Tailed Data
13:00–14:00	**Light lunch**	
14:00–14:30	Amir Khorrami Chokami	Understanding High-Risk Adolescent Obesity Through Variable Importance and Data Augmentation
14:30–15:00	Nathan Huet	Regression for Extremes: An Application to the Prediction of Sea Levels
15:00–15:30	Jean Pachebat	Heavy Tails in Generative Modeling: HT-GAN and Stochastic Interpolants
15:30–16:00	**Coffee break**	
16:00–16:30	Jordan Richards	Generating Multivariate Extremes via a Radial-Angular Decomposition
16:30–17:00	Daniela Castro-Camilo	Learning Extremal Dependence with Normalising Flows
17:00–17:30	Johnny Lee	Generative Modelling for Cascading Extremes via Autoregressive Transform

Day 1 – Thursday, 11 June 2026

14:00–14:30 Talk 1

Speaker: Miguel de Carvalho
Affiliation: University of Edinburgh
Title: *Generative Cascades of Multivariate Extremes*

In this talk, I first motivate the notion of cascades of extremes—as recently introduced by de Carvalho et al. (2026, to appear)—such as a tsunami followed by an earthquake or a cascade of stock market losses. I then argue that the sequential nature of cascading extremes makes them particularly appealing for the setting of generative extremes, and introduce an EVT-grounded generative framework that synthesizes cascades through causal attention. Within this framework, a causal Transformer maps the history of extremes into a latent state that governs direction, magnitude, and inter-event timing. Simulated and real data examples will be used to illustrate the scope of the proposed approaches.

14:30–15:00 Talk 2

Speaker: Giorgia Ramponi
Affiliation: University of Zurich
Title: *Multi-agent Learning without Rewards*

This talk studies how to learn equilibrium behavior in multi-agent systems without relying on hand-designed rewards. Instead, it focuses on learning from demonstrations and human knowledge through multi-agent imitation learning. I will present recent results in mean-field games and Markov games showing that standard single-agent imitation methods are insufficient for recovering equilibria, and that equilibrium-aware objectives and interaction can be necessary. In particular, I will discuss improved guarantees for mean-field settings, hardness results for non-interactive learning in Markov games, and a reward-free interactive method, MAIL-WARM, that achieves rate-optimal learning of equilibria from data.

15:00–15:30 Talk 3

Speaker: Luis Alberto Gutiérrez Inostroza
Affiliation: Pontificia Universidad Católica de Chile
Title: *Bayesian Model Selection for Heavy-Tailed Data via Transformations: Exploring Possibilities for Amortized Inference*

In this talk, we will discuss model selection for predictor-dependent regularly varying distributions. The goal is to identify which predictors influence the distribution of a heavy-tailed response while accurately estimating both its central behavior and tail extremes. Standard spike-and-slab priors are not directly applicable in this setting, as they rely on analytical integration of model parameters. To address this, we propose a transformation-based framework in logarithmic space, where certain heavy-tailed distributions exhibit light-tailed behavior. Because the transformation is injective, posterior inference can be mapped back to the original space coherently. We establish posterior consistency for the resulting nonparametric model, extending Bayesian model selection methods to regularly varying response domains. Possible settings for amortized inference in this model will be discussed.

16:00–16:30 Talk 4

Speaker: Sebastian Engelke
Affiliation: University of Geneva
Title: *Statistical Extrapolation and Applications to AI Weather Forecasting*

Machine learning methods perform well on prediction tasks within the range of their training data, but often struggle with extrapolation, that is, prediction in regions of the predictor space with few or no training observations. In this talk, we argue that extreme value theory provides the mathematical foundation for developing statistical methodology for extrapolation across a wide range of applications. We present our progression method, which leverages extreme value principles to enable statistically grounded extrapolation in regression tasks. We illustrate the societal relevance of such methods through AI weather forecasting. While AI weather models have shown superior performance over traditional physics-based systems on standard benchmarks, they remain challenged when extrapolating to extreme events. Using a benchmark dataset composed exclusively of record-breaking weather events, we show that, in this regime, physics-based models continue to outperform leading AI systems. We further introduce a novel Turing test for physical consistency in state-of-the-art AI weather forecasts, which leverages predictive differences on extrapolation points across an ensemble of pre-trained AI models and is calibrated using principles from conformal prediction.

16:30–17:00 Talk 5

Speaker: Maud Thomas
Affiliation: Université Lyon 1
Title: *Non-Parametric Simulation of Multivariate Extreme Events via Spectral Bootstrap*

Inference in extreme value theory relies on a limited number of extreme observations, making estimation challenging. To address this limitation, we propose a non-parametric simulation scheme, the multivariate extreme events spectral bootstrap simulation procedure, relying on the spectral representation of multivariate generalized Pareto-distributed random vectors. Unlike standard bootstrap methods, our approach preserves the joint tail behaviour of the data and generates additional synthetic extreme data, thereby improving the reliability of inference. We demonstrate the effectiveness of our procedure on the estimation of tail risk metrics, under both simulated and real data. The results highlight the potential of this method for enhancing risk assessment in high-dimensional extreme scenarios.

17:00–17:30 Talk 6

Speaker: Lambert De Monte
Affiliation: University of Edinburgh
Title: *Multivariate Extremal Index and Temporal Clustering in Multivariate Extremes: A Case Study of Dutch Hydrological Extremes*

The relatively recent re-emergence of a geometric approach to multivariate extreme value theory has led to new theoretical and methodological considerations for the study of extreme multivariate events. However, most of these recent methods have neglected the important issue of temporal dependence which may induce clustering of extreme events. In this talk, I will introduce the directional multivariate extremal index (MEI), a measure of the severity of clustering of extreme events which is applicable to all of these recent methods, and propose a Bayesian inference method based on the SPDE/GMRF framework. The usefulness of the MEI is demonstrated on an in-depth case study of compound hydrological extremes of the Rhine and the Meuse rivers for which both low and high discharges have important downstream impacts on Rotterdam in the Netherlands. This is joint work with Bas Matsuura and Ioannis Papastathopoulos.

Day 2 – Friday, 12 June 2026

09:30–10:00   Talk 7

Speaker: Antonio Tirri & Melchiorre Danilo Abrignani
Affiliation: Leithà – Unipol Group
Title: *From Pixels to Policies: AI for Climate Risk and E2E Reproducibility*

Extreme climate events strike Italy ever more frequently, with billion-euro impacts — the Emilia-Romagna flood, the northern hailstorms: for insurance, understanding them means estimating risk as Hazard × Exposure × Vulnerability. We show how Earth Observation and AI support prevention, assessment and post-event analysis: hazard mapping from reanalysis data, nowcasting with generative networks, customer alerts, flood detection from optical/SAR data, and loss sizing. Finally, we reflect on how we made these products reliable and verifiable over time: end-to-end reproducibility as a balance between engineering and fast experimentation, MLOps/IaC as epistemic tools, traceability and lineage, and a minimal, pragmatic set of practices — rigour per unit of cost.

10:00–10:30   Talk 8

Speaker: Antonello Squintu
Affiliation: CMCC
Title: *Exploring AI-enhanced Methods for the Improvement of Sea Level Warnings in Venice*

The city of Venice (Italy) is highly vulnerable to weather-driven sea level height (SL) surge, which causes serious disruptions to city services, including water-ambulances and water-buses, and damage to commercial and cultural assets. Early detection of these events is crucial for increasing the preparedness of citizens and stakeholders and for optimizing the organization of major events. The increased frequency and intensity of high-water events is linked to the rise in average global sea level and to the combination of astronomical tide and weather-driven sea level surge. While these components can be accurately determined via observations and astronomical calculations, the meteorological contribution requires weather forecasts as inputs. The MedEWSa project aims to improve early warning systems by enhancing forecasts of weather-driven sea level anomalies using AI algorithms. The work began with the application of the evolutionary algorithm PCRO-SL on ERA5 reanalysis data to detect relevant lagged drivers of sea level anomaly in the Euro-Mediterranean domain. These drivers were used to train multiple neural networks and tree-based models, with in-situ observations as target series. The algorithms were fine-tuned and evaluated, investigating predictor roles through explainability tools and identifying the best performing model, which has been implemented for daily application. As a result, the Venice Municipality Control Room can benefit from sub-seasonal predictions of sea level, enriching the standing early warning system and increasing the resilience of the city to extreme high-water events.

10:30–11:00   Talk 9

Speaker: Matteo Angelinelli
Affiliation: CINECA HPC
Title: *HPC for Extreme Events: Use Cases and Opportunities at CINECA*

The increasing frequency and intensity of extreme weather events in recent years has emphasised the importance of advanced tools for analysing, simulating and predicting these events. In this context, high-performance computing (HPC) is pivotal in enabling researchers to handle the complexity of atmospheric models and vast amounts of data. This talk will present several use cases in which CINECA has contributed to the study of extreme weather phenomena. It will describe the methodologies adopted, the computational platforms used and the results achieved. Particular emphasis will be placed on how HPC enhances model resolution, accelerates simulation workflows, and facilitates multi-scale, data-intensive analysis. Finally, we will introduce a project through which researchers and institutions can access CINECA's computational resources and specialised technical support. This project aims to promote the use of HPC technologies in various fields, including environmental and climate-related research.

11:30–12:00 Talk 10

Speaker: Olivier Lopez
Affiliation: CREST – ENSAE IP Paris
Title: *Wasserstein GAN with Data Augmentation for the Generation of Multivariate Extremes*

In this work, we introduce a version of Wasserstein Generative Adversarial Networks whose generative structure is adapted to the analysis of multivariate extremes. To compensate for the lack of data when dealing with the tail of the distribution, we train the network based on a mixture of real observations, and synthetic ones, based on a simple preliminary generative model. We derive theoretical results that legitimate this data augmentation strategy, quantifying the bias introduced by the approach, and the benefits in terms of variance.

12:00–12:30 Talk 11

Speaker: Alison Peard
Affiliation: University of Oxford
Title: *Simulating Spatially Resolved Multi-Hazards with Generative Deep Learning*

When natural hazards coincide or spread across large areas they can create major disasters. For accurate risk analysis, it is necessary to simulate many spatially resolved hazard events that capture the relationships between extreme variables, but this has proved challenging for conventional statistical methods, particularly in high-dimensional settings. In this article we show that generative deep learning models—when combined with specific transformations to the training data—offer a useful alternative method for stochastically sampling realistic multi-hazard events. Our framework combines generative adversarial networks with extreme value theory in a hybrid approach that can capture complex dependence structures in gridded multivariate weather data, while providing a theoretical basis for extrapolation to new extremes. We apply our method to jointly model fields of strong winds, heavy precipitation, and low atmospheric pressure (approximately 12,000 variables) during storms in the Bay of Bengal, demonstrating that our model learns the spatial and multivariate extremal dependence structures of the underlying data and captures the distribution of storm severities. For the Bay of Bengal case study, we validate our approach against a popular model for multivariate climate extremes, and demonstrate improved performance in capturing the extremal correlation structure.

12:30–13:00 Talk 12

Speaker: Ashley Turner
Affiliation: Imperial College London
Title: *Scale-Preserving Autoencoders for High-Dimensional Heavy-Tailed Data*

High-dimensional and heavy-tailed data arise in fields such as climate science, finance, insurance and astronomy. Such data often exhibit low-dimensional structure, motivating dimension reduction and latent-space modelling. However, when the downstream objective is magnitude-focused, such as extreme-value analysis, generic nonlinear dimension reduction can destroy the radial scaling and tail behaviour of the original data distribution. This talk will begin with motivating dimension reduction in high-dimensional statistical modelling, recalling some relevant extreme-value theory, including regular variation, tail indices and tail measures, and a review of standard autoencoders as nonlinear dimension-reduction architectures. Homogeneous Autoencoders are then presented as a scale-preserving autoencoder architecture for heavy-tailed data. We show why unconstrained autoencoders may reconstruct data accurately while substantially distorting latent radial properties and tail behaviour and propose an architecture combining a homogeneous encoder with an asymptotically homogeneous decoder. Together, these maps transport radial magnitudes analytically between observation and latent space, yield explicit tail-index and tail-measure transformations under regular variation and preserve the original radial tail index under the encoder-decoder round trip. The methodology is demonstrated on synthetic heavy-tailed manifold-supported data and ERA5 atmospheric fields, where Homogeneous Autoencoders match the reconstructive performance of parameter-matched standard autoencoders while preserving latent and round-trip radial tail diagnostics by construction.

14:00–14:30 Talk 13

Speaker: Amir Khorrami Chokami
Affiliation: University of Cagliari
Title: *Understanding High-Risk Adolescent Obesity Through Variable Importance and Data Augmentation*

Child and adolescent obesity represents an increasing public health concern, and identifying the variables that play a major role in its development is of primary interest. Shapley effects provide a measure of variable importance. However, since obesity is defined through extreme values of the body mass index (BMI), the available data are often limited, particularly when focusing on specific age groups, leading to potentially unstable estimates. To mitigate this issue, we propose a data augmentation strategy based on vine copulas, to generate synthetic observations concentrated in the obese region of the BMI distribution. This augmented dataset allows for more stable and reliable estimation of Shapley effects.

14:30–15:00 Talk 14

Speaker: Nathan Huet
Affiliation: Ca' Foscari University of Venice
Title: *Regression for Extremes: An Application to the Prediction of Sea Levels*

In this presentation, we first introduce a probabilistic framework for regression in regions where the input variable is extreme, in contrast to classical approaches that typically focus on regions where the output variable is extreme. We establish theoretical results on risks and regression functions in these extreme regions, and propose an adapted learning algorithm. In a second part, we apply this algorithm to the prediction of extreme sea levels along the French Atlantic coast. The goal is to predict sea levels at a tide gauge station with a short observation record, by leveraging extreme data from nearby stations with longer historical records. This approach allows us to increase the available extreme data and then to reduce the uncertainty of extreme estimates at the target station. Alongside our regression approach, we also implement another method based on multivariate extreme value theory, which can, for example, be used to generate synthetic samples of extreme sea levels.

15:00–15:30 Talk 15

Speaker: Jean Pachebat
Affiliation: École Polytechnique
Title: *Heavy Tails in Generative Modeling: HTGAN and Stochastic Interpolants*

This talk covers two contributions on structure-preserving generative models for heavy-tailed distributions. The first addresses the reproduction of tail dependence in generative models with a quantification made on the error on the stable tail dependence function (stdf). The second studies heavy-tailed generation within the framework of stochastic interpolants, showing how the interpolant construction can be adapted to converge to prescribed heavy-tailed regimes while preserving training stability. Applications focus on financial and risk data.

16:00–16:30 Talk 16

Speaker: Jordan Richards
Affiliation: University of Edinburgh
Title: *Generating Multivariate Extremes via a Radial-Angular Decomposition*

Leveraging the recently emerging geometric approach to statistical modelling of multivariate extremes, we create a framework for fast and cheap stochastic generation of extreme event sets that comply with the standard regularity conditions of extreme value analysis. In this work, we explore how deep extremal regression models—neural networks designed for extreme value regression and extrapolation—can be couched within the semi-parametric angular-radial (SPAR) model for multivariate extremes. Our resulting deepSPAR model can flexibly capture marginal and joint extremes of multivariate data with minimal modelling assumptions and without the need for a two-stage inference procedure. Model estimation is fast and simulation from fitted models is immediate, meaning that millions of events can be produced in seconds for use in downstream impact models. New visual diagnostics for extreme value regression models are presented alongside a case study of five metocean variables at a site in the Celtic Sea: wind speed, wind direction, wave height, wave period, and wave direction. We show that deepSPAR can capture well the complexity in the joint and marginal extremes of the data and allows for reliable extrapolation far into the joint tails.

16:30–17:00 Talk 17

Speaker: Daniela Castro-Camilo
Affiliation: University of Glasgow
Title: *Learning Extremal Dependence with Normalising Flows*

The multivariate generalised Pareto distribution (mGPD) is a standard framework for modelling threshold exceedances in environmental and financial risk analysis. However, the practical use of mGPD models is often limited due to the infinite possible parametrisations of their dependence function, with only a few parametric models available in practice. We introduce GPDFlow, an innovative mGPD model that leverages normalising flows to flexibly represent the dependence structure. Unlike traditional parametric mGPD approaches, GPDFlow does not impose explicit parametric assumptions on dependence, resulting in greater flexibility and enhanced performance. Additionally, GPDFlow allows direct inference of marginal parameters, providing insights into marginal tail behaviour. We derive tail dependence measures for GPDFlow, including extensions to higher dimensions and partial exceedances. Through simulations and an application to systemic financial risk across five major US banks, we show that GPDFlow offers substantially greater flexibility and improved performance than classical parametric mGPD models.

17:00–17:30 Talk 18

Speaker: Johnny Lee
Affiliation: University of Edinburgh
Title: *Generative Modelling for Cascading Extremes via Autoregressive Transform*

The propagation of extreme shocks—often manifesting as systemic crises or domino effects—is a persistent challenge for dynamic modelling. In this talk, we propose a novel generative framework for modelling cascading extremes using an autoregressive, decoder-only transformer. We view a p-th order Markov chain as the starting point for a d-dimensional stochastic process. Then, we threshold multivariate exceedances to identify extreme events and encode the resulting patterns as a finite vocabulary of extremal tokens. This recasts the modelling problem into a discrete sequence-to-probability task, where the transformer learns the conditional transition dynamics of the next extremal token given a preceding cascade history. The model uses masked multi-head self-attention to estimate transition probabilities, thereby capturing both dependence and directional propagation patterns when extreme events are triggered. In our numerical studies, we present the amortised transformer model to evaluate token generation performance by comparing its posterior predictive distribution and the conditional probabilities against the truth.