Deep learning generative models for sampling high-dimensional extreme precipitation maps

Host institution: BioSP Unit, INRAE, Avignon

Supervisors: Nicolas Lafon (BioSP/INRAE), Thomas Opitz (BioSP/INRAE), Gabriel Victorino Cardoso (Mines Paris - PSL)

Starting date: From June 2026
Duration: 18 months
Salary: from 3550 to 3720 € gross, monthly, based on past experience.

How to apply? Please send your CV and cover letter to the following address nicolas.lafon@inrae.fr, thomas.opitz@inrae.fr

---

Scientific context

Over the past decade, generative approaches originating from machine learning have rapidly expanded and are increasingly applied to the modeling of high-dimensional environmental variables. Extreme events associated with such variables often correspond to high-impact natural phenomena, requiring particularly accurate modeling.

Generating extreme precipitation maps using generative models therefore represents a major applied challenge, especially since physical models based on global circulation simulations still reproduce such extremes imperfectly and cannot serve as fully reliable emulators.

However, modeling extremes with generative methods raises specific challenges, including extrapolation in distribution tails and the coherent prediction of extremely rare — or even unobserved — quantiles. Incorporating results from extreme value theory is therefore essential. While recent approaches attempt to combine these frameworks, few allow heterogeneous tail behaviors across marginal distributions, as most impose a common tail index.

Separating marginal modeling from dependence modeling thus appears to be a promising strategy but remains largely unexplored.

Objectives

The overall goal of this postdoctoral project is to develop and evaluate state-of-the-art generative models (diffusion models, flow matching, stochastic interpolants) for the generation of high-dimensional precipitation maps.

Methodological objectives include:

- generating distributions with heterogeneous tails,
- enforcing asymptotic dependence structures between extreme events,
- the explicit separations of marginal distributions and dependence modelling
- designing evaluation criteria grounded in multivariate extreme value theory

Applied objectives will consist in faithfully reproducing the specific statistical properties of precipitation fields, such as intermittency, as well as, modelling precipitation accumulations across multiple spatial and temporal scales. Finally, achieving a suitable balance between statistical performance and computational efficiency will be necessary.

Geolearning Chair framework

The Geolearning Chair aims to develop innovative approaches combining geostatistics, extreme value theory, and machine learning for the modeling of environmental data. In this context, the simulation of spatial extreme events using generative models constitutes a key scientific focus.

A PhD project, initiated in September 2025, investigates the generation of heavy-tailed distributions using generative approaches. The present postdoctoral project builds upon this general objective while exploring complementary research directions.

Required skills

Applied statistics, climate sciences, deep learning, Python/R programming.