M2 Internship: Inference of spatio-temporal stochastic recurrence models for extreme oceanographic data

Contacts:
Gloria Buriticá - gloria.buritica@agroparistech.fr
Lucia Clarotto - lucia.clarotto@agroparistech.fr

2026

Keywords: Extreme value theory, Spatio-temporal models, Stochastic Recurrence Equation, Inference

To apply, send a CV, a motivation letter and academic transcript with subject "Application for extremes internship" to both supervisors.

1 Context

Climate services provide weather forecasts and climate projections that are essential for anticipating the risks associated with climate hazards and that also play a major role as predictors in models used in the energy, agricultural and industrial sectors.

Despite recent advances in modeling climate variables, forecasts for oceanographic variables such as precipitation, wind speeds, energy flows, and wave heights still tend to underestimate the risks associated with extreme events (Gorse et al., 2025). Indeed, these variables regularly reach extreme levels and therefore their stochastic modeling requires the use of tools from heavy tail theory (Resnick, 2007).

The topic of this internship focuses on improving stochastic modeling of extreme oceanographic conditions, which is key to ensuring the effectiveness of risk prediction plans associated with climate hazards. An important aspect is ensuring spatial adequacy, as reliable predictions require models that accurately capture the spatial variability of extreme events.

The model is given by the stationary solution of the spatial stochastic recurrence equation (SRE):

$$X_t = A_t X_{t-1} + B_t, \qquad t \in \mathbb{Z}, \tag{1.1}$$

for independent and identically distributed innovations $((A_t, B_t))$ with values in $\mathbb{R}^{N\times(N+1)}$ where N denotes the number of spatial locations of the field (X_t).

In the univariate case, the model (1.1) has been extensively studied in the field of time series (Buraczewski et al., 2013). The advantage of this model is that it allows the description of heavy-tailed phenomena of the Pareto type through the introduction of multiplicative noise, even when the innovations $((A_t, B_t))$ are light-tailed.

More precisely, the Kesten-Goldie model (Kesten, 1974; Goldie, 1991) establishes sufficient conditions on $((A_t, B_t))$ such that

$$\mathbb{P}(X_t > x) \sim cx^{-\alpha}, \qquad x \to \infty, \quad x > x_0,$$

with $c, \alpha > 0$. Here, the tail index $\alpha > 0$ describes the properties of the heavy tails of the data distribution.

The internship aims to extend this model to the spatial framework based on Gaussian innovations, and to implement the numerical tools necessary for its inference. This work also aims to overcome the computational barriers associated with the transition from a univariate model to a spatial model, capable of ensuring both fitting and forecasting on a spatial grid.

Regarding inference and forecasting, we propose relying on maximum likelihood methods (such as Kalman Filter) for state space models (Brockwell and Davis, 2002). The challenge will be to extend these methods to take into account random matrices $((A_t, B_t))$, relying on Expectation-Maximization (EM) algorithm, and to design numerical solutions that allow inference and forecasting on datasets with a high number of spatial locations. Solving this problem would enable a comprehensive comparison between traditional spatio-temporal models, which are often based on light tail assumptions, and multiplicative noise models, more adapted to heavy tails.

The intern will adjust and compare these models on data from coastal meteorological stations on oceanographic conditions such as wind speeds and energy flows.

2 Scientific goals of the internship

In order to highlight the advantages of the stochastic model (1.1) for modeling oceanographic phenomena with heavy tails, it is essential to develop effective numerical tools for estimating and adjusting this model to the data.

During the internship the student is expected to perform the following tasks:

- Conduct a literature review on recent work using the SRE model (1.1) for extremes;
- Develop approaches for estimating the parameters of model (1.1) for spatial data;
- Adjust the model (1.1) to oceanographic data (Raillard et al., 2023);
- Compare the model (1.1) to spatio-temporal models in terms of consistency in spatial and temporal structure, and modeling of extremes.

The outcome of this internship is an operational code that would enable adjustment to different climate science datasets. This work would highlight the advantages of the multiplicative noise model in terms of both heavy tail modeling and spatial and temporal consistency.

3 Profile & environment

The candidate should be a 2nd year master or last year engineer student, in Statistics, Mathematics or Machine Learning. Scientific programming skills in Python and/or R are required.

- Location : UMR MIA Paris-Saclay, Palaiseau Campus, 22 place de l'agronomie, 91120 Palaiseau.
- Supervision : Gloria Buritica is a researcher in statistics for extremes and Lucia Clarotto is a researcher in spatio-temporal statistics.
- Starting date: flexible, starting in February 2026 or after.
- Duration: 6 months.
- Salary: as an intern, you'll receive a gratification which is capped around 650 euros/month net.

The candidate will have an office, and benefit from the work environment of the MIA Paris-Saclay laboratory, with many PhD students & postdocs working on statistical modeling and machine learning for the life sciences.

References

Peter J Brockwell and Richard A Davis. *Introduction to time series and forecasting*. Springer, 2002.

D. Buraczewski, E. Damek, T. Mikosch, and J. Zienkiewicz. Large deviations for solutions to stochastic recurrence equations under kesten's conditions. *The annals of Probability*, 41:2755–2790, 2013.

Charles M Goldie. Implicit renewal theory and tails of solutions of random equations. *The Annals of Applied Probability*, pages 126–166, 1991.

Nathan Gorse, Olivier Roustant, Jérémy Rohmer, and Déborah Idier. Simulation of extreme functionals in meteoceanic data: Application to surge evolution over tidal cycles. *arXiv preprint arXiv:2508.13687*, 2025.

Harry Kesten. Renewal theory for functionals of a markov chain with general state space. *The Annals of Probability*, pages 355–386, 1974.

Nicolas Raillard, Simon Chabot, Christophe Maisondieu, David Darbynian, Gregory Payne, and Louis Papillon. Resourcecode: A python package for statistical analysis of sea-state hindcast data. *Journal of Open Source Software*, 8(83):4366, 2023.

S. I. Resnick. *Heavy-tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York, 2007.