M2 Internship: Bayesian Inverse Problem for Spatially Varying Anisotropy in Stochastic PDEs

Contacts:
Charlie Sire – charlie.sire@agroparistech.fr
Lucia Clarotto – lucia.clarotto@agroparistech.fr

Internship Laboratory :
UMR MIA Paris-Saclay, Palaiseau Campus, 22 place de l'agronomie, 91120 Palaiseau

Keywords: SPDEs, anisotropy, Bayesian inverse problem, spatio-temporal statistics.

To apply send a CV and a motivation letter with subject "Application for SPDE internship".

1 Context

When modeling and predicting spatial or spatio-temporal phenomena, a classical approach in geostatistics consists in representing the observed data as realizations of a Gaussian random field (GRF), which is fully characterized by its mean function and covariance kernel. Although physical constraints can be incorporated into the specification of the prior covariance kernel (Padilla-Segarra et al., 2025), an alternative approach consists in considering GRFs defined as solutions to stochastic partial differential equations (SPDEs).

For instance, Clarotto et al. (2024) propose modeling solar radiation data using a spatio-temporal SPDE that incorporates both advection and diffusion:

$$\left[\frac{\partial}{\partial t} + \frac{1}{c}\left(\kappa^2 - \nabla \cdot \mathbf{H}(\mathbf{s})\nabla\right)^{\alpha} + \frac{1}{c}\boldsymbol{\gamma} \cdot \nabla\right] X(t, \mathbf{s}) = \frac{\tau}{\sqrt{c}} Z(t, \mathbf{s}), \tag{1}$$

where Z denotes a stochastic forcing term. The advection term $\boldsymbol{\gamma} \cdot \nabla$ accounts for the transport of the process across the spatial domain, while the spatial diffusion is modeled by the term $\nabla \cdot \mathbf{H}(\mathbf{s})\nabla$. The matrix-valued function $\mathbf{H}(\mathbf{s})$ encodes spatially varying anisotropy, specifying the preferred directions of the diffusion at each location $\mathbf{s}$. The additional parameters α, c, κ, $\boldsymbol{\gamma}$, and τ further determine the covariance structure of the solution and are not detailed here.

To perform statistical inference for this model from observed data, the work of Lindgren et al. (2011) provides a foundational framework. Their approach relies on a finite element approximation of the SPDE solution over the spatial domain, leading to a Gaussian Markov random field (GMRF, see Rue and Held, 2005) with a sparse precision matrix. This sparsity makes it possible to handle large datasets, overcoming a classical computational limitation of Gaussian process regression.

A central question when developing this type of approach is the estimation of the unknown parameters. This is classically achieved through maximum likelihood estimation, but evaluating the likelihood for a given parameter configuration can be computationally demanding and may require approximations (see Pereira et al., 2022). Inference becomes even more challenging when local anisotropies are considered, meaning that the matrix-valued function $\mathbf{H}(\mathbf{s})$ in Equation 1 is not constant, modeling a situation when the preferred direction of diffusion varies across space.

This internship therefore aims to model and estimate these local anisotropies.

2 Scientific goals of the internship

The main objective of this internship is to develop a framework to estimate spatially varying anisotropy parameters. A naive approach would consist in estimating the anisotropy parameters at every node of the mesh used for the finite element approximation. However, this would lead to a number of parameters proportional to the mesh size, making maximum likelihood optimization computationally infeasible.

The idea is therefore to model the spatially varying anisotropy parameters (e.g., the function $\mathbf{H}(\mathbf{s})$ in Equation 1) using spatial random fields, thereby accounting for uncertainty in their estimation within a hierarchical Bayesian framework. Although related hierarchical approaches have been proposed in the literature (see Ingebrigtsen et al. 2014; Fuglstad et al. 2015), their idea is to express the anisotropy parameters as linear combinations of predefined basis functions, whose choice can be challenging. An alternative parametrization to be investigated is to specify the anisotropy parameters at a set of selected pilot points and treat their values as uncertain quantities. The spatial covariance structure governing the resulting random fields of anisotropy parameters and the location of the pilot points can then be chosen based on physical considerations.

3 Provisional work plan

During the internship, the student is expected to carry out the following tasks:

- Review recent literature on SPDEs and Bayesian hierarchical models;
- Develop a fully Bayesian framework for modeling and simulating spatially varying anisotropy parameters in a simple spatial SPDE, using simulated data;
- Apply the proposed approach to a more complex spatio-temporal model for real environmental data and compare it with existing methods;
- Explore improvements in likelihood computation, for instance through adaptive mesh refinement.

4 Profile and practical information

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 : Charlie Sire and Lucia Clarotto are researchers in spatio-temporal statistics, with expertise in Bayesian modeling and SPDE approach.
- Duration and starting date: 6 months, starting in February 2027 or after.
- Salary: The intern will receive a gratification which is capped around 700 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 and environmental sciences.

References

Lucia Clarotto, Denis Allard, Thomas Romary, and Nicolas Desassis. The SPDE approach for spatio-temporal datasets with advection and diffusion. *Spatial Statistics*, 62:100847, 2024.

Geir-Arne Fuglstad, Finn Lindgren, Daniel Simpson, and Håvard Rue. Exploring a new class of non-stationary spatial gaussian random fields with varying local anisotropy. *Statistica Sinica*, pages 115–133, 2015.

Rikke Ingebrigtsen, Finn Lindgren, and Ingelin Steinsland. Spatial models with explanatory variables in the dependence structure. *Spatial Statistics*, 8:20–38, 2014.

Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(4):423–498, 2011.

Adrian Padilla-Segarra, Pascal Noble, Olivier Roustant, and Éric Savin. Physics-informed, boundary-constrained gaussian process regression for the reconstruction of fluid flow fields. *International Journal for Uncertainty Quantification*, 2025.

Mike Pereira, Nicolas Desassis, and Denis Allard. Geostatistics for large datasets on riemannian manifolds: a matrix-free approach. *Journal of Data Science*, 20(4):512–532, 2022. doi: 10.6339/22-JDS1075.

Havard Rue and Leonhard Held. *Gaussian Markov random fields: theory and applications*. Chapman and Hall/CRC, Boca Raton, FL, USA, 2005.