Postdoc position @ Mines Paris – PSL University and ANDRA

Statistical modeling of spatio-temporal data distributed over surfaces : Application to the monitoring of nuclear waste storage cells

This research project is motivated by a collaboration between the Geostatistics team at the Geosciences center of Mines Paris - PSL and the French National Agency for Radioactive Waste Management (ANDRA), as part of the Geolearning chair (https://chaire-geolearning.org/en/). The aim of the chair is to develop methods in geostatistics, extreme event theory and machine learning for data analysis in support of the climate transition.

Background and objectives

For over 25 years, ANDRA has been carrying out research for the Cigeo project, a deep geological disposal facility for high-level (HL) and long-lived intermediate-level (ILW) radioactive waste to be built in France. Located at a depth of 500 meters in the Callovo-Oxfordian (COX) clay, Cigeo is designed to accommodate waste produced by France's current fleet of nuclear facilities, until they are dismantled, as well as waste from the reprocessing of spent fuel from nuclear power plants. The facility's monitoring strategy includes the comparison of in situ data (sensors) with predictive numerical simulations, to ensure that the repository remains within its expected operating range. The design of the Cigeo underground facility is based on a number of guiding principles, in particular the monitoring of reference structures that can be used to trace the overall behavior of structures of the same type. These control structures are equipped with numerous monitoring devices, limiting the need to place large numbers of sensors in all of the structures of the same type.

In this post-doctoral project, we are focusing on cells designed for the storage of high-activity (HA) waste. The aim is to propose methods for estimating the temperature and deformation fields of a group of HA cells, including cells with an abundance of sensors and others with a limited number of sensors. Using synthetic sensor data (i.e. generated from demonstrators implemented at the Meuse/Haute-Marne underground laboratory), the scientific challenge is to develop geostatistical algorithms capable of reliably reconstructing the thermal and mechanical conditions at all points of the cell lining and over time (the HA cell lining being cylindrical in shape).

Research goals

The aim of the project is to develop statistical methods for interpolating and predicting spatio-temporal data distributed over surfaces. These methods will be applied to data from temperature and deformation sensors placed at some points on the surface of a cell in order to monitor its evolution. The methods developed will meet four main objectives :

- Take into account spatial and temporal correlations observed in the data.
- Reconstruct temperature and deformation fields defined over the entire surface of a cell, predict their future evolution, and provide a means of quantifying the uncertainty of these.
- Identify and model the links between temperature and deformation fields for a given cell and between several cells, in order to propose multivariate methods for reconstructing (and inferring) these fields.
- Be robust to missing data (e.g. from sensor failures) and identify potential sensor drifts.

Considering a geostatistical approach to the problem, a starting point for the project is the so-called "SPDE" approach, whereby observed spatio-temporal data are modeled as samples of a Gaussian field defined as the solution of a stochastic partial differential equation (cf. [2]). This approach has led to the development of efficient methods for inferring non-stationary Gaussian fields, defined on surfaces (see e.g. [3]) and even spatio-temporal fields (see e.g. [1]). One of the aims of the project will therefore be to adapt these methods to the case of modeling spatio-temporal fields defined on open surfaces (such as cylindrical shape of a cell).

Applicant profile

- PhD in statistics, statistical learning or applied mathematics.
- Aptitude for topics combining theory (statistical, mathematical and/or physical modeling) and practice (development of numerical algorithms to answer the problem).
- Proficiency in Python and/or R.

Informations complémentaires

The person recruited will be part of the Geostatistics team at the Geosciences center of Mines Paris - PSL (Fontainebleau, 77) and may spend time at the ANDRA site (Châtenay-Malabry, 92). She will be supervised by Mike Pereira, in the Geostatistics team, and by Julien Cotton at ANDRA.

- Planned start date : From February 2024
- Duration: 24 months
- Salary: Between €2640 and €3340 gross (depending on experience)
- Opportunity to take part in courses given to engineering students at Mines Paris - PSL

How to apply ?

Send a CV, a cover letter and contact details for two references to the addresses below:

mike.pereira@minesparis.psl.eu and julien.cotton@andra.fr

For more information

Do not hesitate to write to us:

- Mike Pereira, Mines Paris – PSL : mike.pereira@minesparis.psl.eu
- Julien Cotton, ANDRA : julien.cotton@andra.fr

You can also consult the following document for further information on the proposed topic (in French): https://mike-pereira.github.io/files/postdoc-andra.pdf

References

(1) Clarotto, L., Allard, D., Romary, T., & Desassis, N. (2022). The SPDE approach for spatio-temporal datasets with advection and diffusion. *arXiv preprint arXiv:2208.14015*.

(2) Lindgren, F., Bolin, D., & Rue, H. (2022). The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running. *Spatial Statistics*, *50*, 100599.

(3) Pereira, M., Desassis, N., & Allard, D. (2022). Geostatistics for Large Datasets on Riemannian Manifolds: A Matrix-Free Approach. *Journal of Data Science*, *20*(4), 512-532.