

## OFFRE DE STAGE EN MATHÉMATIQUES

"Interpolation spatiale d'une variable régionalisée par méthode de krigeage"

- Durée : 3 mois entre mai et juillet 2022.
- Lieu : Laboratoire de Mathématiques et Applications, Université de Poitiers

Pour candidater, envoyez votre CV et vos relevés de notes de Licence et Master à :  
[yousri.slaoui@univ-poitiers.fr](mailto:yousri.slaoui@univ-poitiers.fr) et [arnaud.poinas@univ-poitiers.fr](mailto:arnaud.poinas@univ-poitiers.fr).

## CONTEXTE DU STAGE

Le Laboratoire de Mathématiques et Applications (LMA) de l'université de Poitiers et du CNRS (UMR 7348) développe un programme d'étude et de recherche autour de l'apprentissage Statistique dans un cadre des données spatiales. Un financement de stage de recherche de niveau master (M1 ou M2) a été obtenu pour ce sujet dans le cadre d'un projet CPER-FEDER.

Vous serez encadré(e) par Yousri Slaoui (maître de conférences HDR, spécialiste en Statistique et Probabilités) et Arnaud Poinas (maître de conférences, spécialiste en Statistique et Probabilités).

## PROFIL ET COMPÉTENCES RECHERCHÉES

- On recherche un(e) étudiant(e) issu(e) d'un master 1 ou master 2 en mathématiques, ayant suivi des cours avancés en probabilité et statistique et avec des connaissances de bases en programmation python et/ou R.
- Capacité à travailler en collaboration, prise d'initiative, autonomie.

## CONTEXTE SCIENTIFIQUE

On considère un ensemble de  $n$  points sur le plan  $\{x_1, \dots, x_n\}$ . En chacun de ces points est observé une réalisation d'une variable aléatoire  $Z(x_1), \dots, Z(x_n)$ . On considère que ces variables aléatoires sont les valeurs au points  $x_1, \dots, x_n$  d'une fonction aléatoire  $Z(\cdot)$  dont on cherche à estimer la valeur en tout point  $x$ .

Ce problème est très important en géostatistique où il est courant de prélever des données à divers sites d'observations. Par exemple, en glaciologie il est courant de prélever des carottes de glace à plusieurs emplacements du glacier pour connaître son épaisseur à ces emplacements et ensuite vouloir estimer l'épaisseur du glacier en tout point. En foresterie, il est aussi courant de prélever des carottes de terres à plusieurs emplacements d'une forêt pour connaître la composition du sol (pH, concentration en minéraux, ...) à ces emplacements et vouloir ensuite estimer la composition du sol en tout point.

La méthode d'estimation classique, appelée **krigeage**, consiste à estimer la valeur de  $Z(x)$  en tout point  $x$  par  $\hat{Z}(x) = \sum_{i=1}^n \lambda_i Z(x_i)$ , une combinaison linéaire des valeurs observées de  $Z(x_1), \dots, Z(x_n)$ . Les coefficients  $\lambda_i$  sont alors choisis de sorte que  $\hat{Z}(x)$  soit non-biaisé et de variance minimale. Sous certaines hypothèses, ces coefficients optimaux dépendent alors de la valeur de la fonction de corrélation  $C(h) = \mathbb{E}[Z(x+h)Z(x)] - \mathbb{E}[Z(x+h)]\mathbb{E}[Z(x)]$  (que l'on suppose ne pas dépendre de  $x$ ). En pratique, cette fonction n'est pas connue et elle doit être estimée de façon paramétrique ou non-paramétrique.

Des travaux de recherche plus récents [1] s'intéressent notamment au problème de choisir les sites d'observations  $x_1, \dots, x_n$  de façon à minimiser les erreurs d'estimation de la fonction de covariance. Une méthode populaire consiste à choisir les  $x_i$  aléatoirement mais de sorte qu'ils soient bien répartis spatialement et pas trop proche les uns des autres. Il s'agit alors de la réalisation de ce qu'on appelle un **processus ponctuel répulsif**.

## OBJECTIFS DU STAGE

- Étudier les méthodes de krigeage simple et ordinaire et les appliquer sur des exemples jouets en R ou en Python.
- Étudier les méthodes d'estimation de la fonction de corrélation afin de pouvoir faire du krigeage sur des vrais jeux de données.
- Étudier mathématiquement et numériquement l'utilisation de processus ponctuels répulsifs pour sélectionner les sites d'observations.

## Références

- [1] François Bachoc. Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of gaussian processes. *Journal of Multivariate Analysis*, 125 :1–35, 2014.
- [2] Ricardo A. Olea. *Geostatistics for Engineers and Earth Scientists*. Springer, 1999.