



Vous êtes cordialement invité(e) à la soutenance de thèse de

Mike PEREIRA

**Champs aléatoires généralisés définis sur des variétés riemanniennes :
théorie et pratique**

Soutenance prévue **jeudi 28 novembre 2019 à 14h00**
Amphithéâtre Coléou - AgroParisTech 16 rue Claude Bernard 75005 Paris

Composition du jury proposé :

Denis ALLARD	Directeur de recherche	INRA PACA	Rapporteur
Finn LINDGREN	Professeur	University of Edimburgh Old College	Rapporteur
Annika LANG	Associate Professor	Chalmers University of Technology	Examineur
Emilio PORCU	Professeur	Trinity College	Examineur
Liliane BEL	Professeur	AgroParisTech	Examineur
Nicolas DESASSIS	Chargé de recherche	MINES ParisTech	Examineur
Hans WACKERNAGEL	Directeur de recherche	MINES ParisTech	Examineur
Cédric MAGNERON	Ingénieur	ESTIMAGES	Invité

Résumé :

La géostatistique est la branche des statistiques s'intéressant à la modélisation des phénomènes ancrés dans l'espace au travers de modèles probabilistes. En particulier, le phénomène en question est décrit par un champ aléatoire (généralement gaussien) et les données observées sont considérées comme résultant d'une réalisation particulière de ce champ aléatoire. Afin de faciliter la modélisation et les traitements géostatistiques qui en découlent, il est d'usage de supposer ce champ comme stationnaire et donc de supposer que la structuration spatiale des données se répète dans le domaine d'étude. Cependant, lorsqu'on travaille avec des jeux de données spatialisées complexes, cette hypothèse devient inadaptée. En effet, comment définir cette notion de stationnarité lorsque les données sont indexées sur des domaines non euclidiens (comme des sphères ou autres surfaces lisses)? Quid également du cas où les données présentent structuration spatiale qui change manifestement d'un endroit à l'autre du domaine d'étude? En outre, opter pour des modèles plus complexes, lorsque cela est possible, s'accompagne en général d'une augmentation drastique des coûts opérationnels (calcul et mémoire), fermant alors la porte à leur application à de grands jeux de données. Dans ce travail, nous proposons une solution à ces problèmes s'appuyant sur la définition de champs aléatoires généralisés sur des variétés riemanniennes. D'une part, travailler avec des champs aléatoires généralisés permet d'étendre naturellement des travaux récents s'attachant à tirer parti d'une caractérisation des champs aléatoires utilisés en géostatistique comme des solutions d'équations aux dérivées partielles stochastiques. D'autre part, travailler sur des variétés riemanniennes permet à la fois de définir des champs sur des domaines qui ne sont que localement euclidiens, et sur des domaines vus comme déformés localement (ouvrant donc la porte à la prise en compte du cas non stationnaire). Ces champs généralisés sont ensuite discrétisés en utilisant une approche par éléments finis, et nous en donnons une formule analytique pour une large classe de champs généralisés englobant les champs généralement utilisés dans les applications. Enfin, afin de résoudre le problème du passage à l'échelle pour les grands jeux de données, nous proposons des algorithmes inspirés du traitement du signal sur graphe permettant la simulation, la prédiction et l'inférence de ces champs par des approches "matrix-free".

Mots-clés : Champ aléatoire généralisé, Variété riemannienne, Traitement du signal sur graphe, Equation aux dérivées partielles stochastique, Méthode des éléments finis

Vous êtes cordialement invité(e) au pot amical qui suivra la soutenance
espace Dubos (4ème étage, aile Arbalète)

Generalized random fields on Riemannian manifolds: theory and practice

Abstract:

Geostatistics is the branch of statistics attached to model spatial phenomena through probabilistic models. In particular, the spatial phenomenon is described by a (generally Gaussian) random field, and the observed data are considered as resulting from a particular realization of this random field. To facilitate the modeling and the subsequent geostatistical operations applied to the data, the random field is usually assumed to be stationary, thus meaning that the spatial structure of the data replicates across the domain of study.

However, when dealing with complex spatial datasets, this assumption becomes ill-adapted. Indeed, how can the notion of stationarity be defined (and applied) when the data lie on non-Euclidean domains (such as spheres or other smooth surfaces)? Also, what about the case where the data clearly display a spatial structure that varies across the domain? Besides, using more complex models (when it is possible) generally comes at the price of a drastic increase in operational costs (computational and storage-wise), rendering them impossible to apply to large datasets.

In this work, we propose a solution to both problems, which relies on the definition of generalized random fields on Riemannian manifolds. On one hand, working with generalized random fields allows to naturally extend ongoing work that is done to leverage a characterization of random fields used in Geostatistics as solutions of stochastic partial differential equations. On the other hand, working on Riemannian manifolds allows to define such fields on both (only) locally Euclidean domains and on locally deformed spaces (thus yielding a framework to account for non-stationary cases). The discretization of these generalized random fields is undertaken using a finite element approach, and we provide an explicit formula for a large class of fields comprising those generally used in applications. Finally, to solve the scalability problem, we propose algorithms inspired from graph signal processing to tackle the simulation, the estimation and the inference of these fields using matrix-free approaches.

Keywords: Generalized random field, Riemannian manifold, Graph signal processing, Stochastic partial differential equation, Finite element method