



Avis de Soutenance

Monsieur Bifeng HU

Sciences de la Terre

Soutiendra publiquement ses travaux de thèse intitulés

Modélisation spatio-temporelle conjointe des propriétés des sols à partir des données d'inventaire et de surveillance en France : approche bayésienne appliquée au phosphore du sol à travers la France métropolitaine

dirigés par Monsieur Hocine BOURENNANE

Ecole doctorale : Energie, Matériaux, Sciences de la Terre et de l'Univers - EMSTU

Unité de recherche : INRA Sc du Sol - Unité Recherche Science du Sol

Soutenance prévue le **mardi 17 novembre 2020** à 14h00

Lieu : Faculté Des Sciences, Rue de Chartres, 45100 Orléans

Salle : des theses - En visioconférence

Composition du jury proposé

M. Hocine BOURENNANE	INRAE Orléans	Directeur de thèse
M. Philippe LAGACHERIE	INRAE Montpellier	Rapporteur
M. Christian WALTER	Agrocampus Ouest	Examineur
M. Didier CHAUVEAU	Université d'Orléans	Examineur
M. Thomas OPITZ	INRAE Avignon	Examineur
M. Nicolas SABY	INRAE Orléans	Co-encadrant de thèse
Mme Liliane BEL	Université Paris-Saclay	Rapporteuse
M. Vincent CHAPLOT	IRD Marseille	Examineur

Mots-clés : harmonisation, propriétés du sol, INLA-SPDE, données anciennes, spatio-temporelle,

Résumé :

Cette thèse vise à développer un cadre statistique combinant une approche bayésienne et les approches de cartographie numérique des sols et se basant sur des données existantes issues du système d'information national français afin de cartographier et détecter les variations spatio-temporelles des propriétés des sols. Ce cadre a été mis en œuvre pour explorer la variabilité spatio-temporelle de la teneur en phosphore extractible (P2O5) dans les sols agricoles de la France métropolitaine de 1991 à 2014 à l'aide de la collecte de données dans le cadre du réseau de mesure de la qualité des sols et du programme d'analyses des terres (BDAT). La première partie du mémoire définit le cadre scientifique,

l'approche développée et les bases de données utilisées. La deuxième partie concerne l'harmonisation des teneurs en P2O5 du sol, mesurées par différentes méthodes analytiques à travers la France en utilisant des méthodes de régression par les moindres carrés partiels et d'apprentissages automatiques. La dernière partie traite de la modélisation spatio-temporelle des teneurs en P2O5 assimilable des sols agricoles en France par une méthode d'inférence bayésienne qui combine une approximation analytique des lois appelée INLA et l'approche utilisant la géostatistique et des fonctions aléatoires vues comme solutions d'équations aux dérivées partielles stochastiques (SPDE). Ce travail s'est focalisé sur la modélisation de tendances temporelles non linéaires variant dans l'espace avec une évaluation des incertitudes. Les observations sont ainsi modélisées à travers un processus lisse latent dont les composantes sont additives et dotées de connaissances a priori de type gaussien. J'ai enfin proposé une approche pour explorer graphiquement les tendances des interpolations spatio-temporelles du modèle en utilisant une analyse multivariée de type clustering. Concernant l'harmonisation des mesures de P2O5, les résultats montrent une bonne capacité des fonctions de pédotransfert à estimer séparément P2O5 Olsen en utilisant différentes combinaisons de proxys de P2O5 et des propriétés du sol, et ce dans les sols calcaires et non calcaires. Ces fonctions sont ensuite appliquées sur la BDAT pour obtenir un ensemble conséquent d'observations harmonisées afin de conduire l'analyse spatio-temporelle. L'approche INLA-SPDE a d'abord été construite et testée à l'échelle régionale en Bretagne, puis à l'échelle nationale. La qualité globale du modèle, évaluée par une validation croisée à 5 itérations, reste assez faible à l'échelle régionale ($R^2 = 0,23$) et légèrement meilleure pour l'échelle nationale ($R^2 = 0,29$). Cela dit, quelle que soit l'échelle, les modèles ajustés ont fourni des prédictions spatio-temporelles continues de P2O5 Olsen dans l'espace et dans le temps ainsi que les incertitudes associées. Par ailleurs, la structure du modèle nous a permis de mener des inférences statistiques sur ses différentes composantes. Ces inférences ont confirmé la significativité des évolutions temporelles des structures spatiales de P2O5 Olsen qui pourraient être liées à des changements connus dans les pratiques agricoles. Dans un deuxième temps, j'ai conduit une étude de sensibilité à la résolution de la discrétisation nécessaire à la méthode SPDE sur l'étape d'ajustement du modèle. J'ai observé des résultats très variables se traduisant par des estimations très différentes des paramètres des variables latentes. Cela pourrait s'expliquer par la forte variabilité des valeurs mesurées dans la base de données. En effet, pour un site de mesure, la teneur mesurée d'Olsen P2O5 peut fortement varier entre deux dates successives. Des recherches supplémentaires seront naturellement nécessaires pour améliorer le cadre de cette modélisation permettant de stabiliser davantage les résultats actuels. Enfin, l'approche par clustering semble prometteuse car elle permet de communiquer facilement la variabilité des trajectoires non linéaire dans le temps des prédictions du modèle.