

Sujet de stage M2
Mathématiques appliquées / Statistique
Année universitaire 2019-2020

Apprentissage statistique pour la détermination de profils climatiques spatio-temporels à l'origine des variations de rendements agricoles

Contexte scientifique :

Les rendements agricoles dépendent de plusieurs facteurs, dont (mais pas seulement) la variété cultivée, les pratiques agricoles, les choix techniques et les conditions climatiques qui fluctuent d'une année à l'autre. Sur l'un des rares jeux de données contenant plusieurs milliers de parcelles, l'objectif de ce stage est d'identifier les profils spatio-temporels des conditions climatiques à l'origine des variations de rendement. Cette meilleure compréhension des phénomènes climatiques doit permettre de mieux déterminer et prédire les risques pour les agriculteurs afin de leur proposer des solutions adaptées comme des améliorations variétales, des modifications des pratiques agricoles ou le recours à des assurances, etc.

Données à analyser et méthodologie :

Nous disposons des relevés de rendement de blé tendre de 2011 à 2016 pour 2300 parcelles du réseau DEPHY, réparties sur le territoire métropolitain français. Les conditions climatiques sont connues au travers de données de ré-analyse issues de la base SAFRAN de Météo-France. Les conditions environnementales et les pratiques agricoles seront raisonnablement supposées inchangées durant la période 2011-2016, afin de se concentrer sur l'influence des conditions climatiques. Nous disposons également d'indicateurs éco-climatiques résumant l'information météorologique lors de différentes phases du cycle de croissance du plant de blé. Toutes ces données sont contenues dans une seule base de données prête pour analyse.

Face aux limites des modèles spatiaux (champs, processus) et temporels pour détecter des phénomènes complexes inconnus, l'objectif majeur de ce stage sera de s'appuyer sur des techniques d'apprentissage statistique (forêts aléatoires, réseaux de neurones, boosting aléatoire...) pour rechercher dans les données les profils spatio-temporels déterminant les variations de rendement. Une première étape sera de détecter les profils spatio-temporels climatiques associées aux stades de croissance du plant de blé pour expliquer les fluctuations et notamment la baisse de rendement de ces dernières années. Selon l'avancement du projet, un développement possible sera de proposer des modèles couplés alliant les techniques de machine learning et des modèles de statistique spatiale. L'analyse sera principalement menée sous le logiciel R et éventuellement sous Python.

Ce projet sera encadré par Denis Allard (INRA BioSp) et Florent Bonneu (Avignon Université, LMA), en collaboration avec Davide Martinetti (INRA BioSp), des collègues de l'unité AgroClim (Patrick Bertuzzi et Inaki Garci-De-Cortazar) et de l'entreprise Agrosolutions (Geoffrey Monier).

Lieu : INRA Avignon, unité BioSP (Biostatistique et Processus Spatiaux).

Durée : 6 mois entre mars et octobre 2019.

Rémunération : la gratification de stage sera de 577,50€/mois net, conformément à la grille officielle en vigueur en 2019 en France (base de 22 jours de présence/mois et 35h/semaine).

Profil du candidat : Data science, statistique spatiale, séries temporelles, programmation R, intérêt pour le climat, l'agriculture et l'environnement.

Contacts : CV, lettre de motivations et relevés de notes de Master à :

Denis Allard (denis.allard@inra.fr) et Florent Bonneu (florent.bonneu@univ-avignon.fr).