

Sujet de stage M2
Mathématiques appliquées / Statistique
Année universitaire 2018-2019

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 de 2011 à 2016 pour 2300 parcelles du réseau DEPHY, réparties sur le Nord et l'Ouest de la France. Les conditions climatiques sont connues soit au travers des relevés journaliers de température et de précipitation réalisés par des stations météorologiques équitablement réparties sur le territoire considéré, soit en accédant à des données de ré-analyse issue 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.

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, boosting aléatoire, réseaux de neurones...) pour rechercher dans les données les profils spatio-temporels déterminant les variations de rendement. Une première étape sera de détecter les combinaisons des indicateurs éco-climatiques pour expliquer le plafonnement ou la baisse de rendement de ces dernières années. L'analyse sera menée sous le logiciel R.

Ce projet sera encadré par Denis Allard (INRA BioSp) et Florent Bonneau (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 (Nicolas Urruty) du groupe InVivo.

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

Durée : 4 à 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 : Si vous êtes intéressé, envoyer un CV et une lettre de motivation à Denis Allard (denis.allard@inra.fr) et Florent Bonneau (florent.bonneu@univ-avignon.fr).