



Année universitaire 2018-2019

Proposition de stage M2

Exploration de méthodes statistiques multi-tableaux pour l'agro-écologie

Magali San Cristobal, Michel Goulard, Aude Vialatte, Nathalie Peyrard

INRA Toulouse, UMR DYNAFOR

magali.san-cristobal@inra.fr, michel.goulard@inra.fr, aude.vialatte@ensat.fr,
nathalie.peyrard@inra.fr

Le projet « Muesli » piloté par des scientifiques de l'UMR Dynafor de l'INRA d'Auzeville, vise à explorer les services écosystémiques en lien avec la production agricole. Comment le paysage dans sa composition, sa complexité, est lié ou non à la production agricole et forestière ? Divers éléments ont été mesurés, de manière plus ou moins précise : le rendement en céréale à paille de 30 parcelles agricoles, le rendement estimé de placettes forestières, des services écosystémiques de pollinisation, de prédation de ravageurs, de biodiversité, les traitements en produits phytosanitaires. Le paysage peut être appréhendé de diverses manières, nous en étudierons trois: (i) le paysage « anthropologique », c'est-à-dire perçu par les hommes en termes de composition et diversité d'éléments paysagers (types de cultures, bois, haies, routes, etc ...), (ii) le paysage perçu par les oiseaux (la présence des certaines espèces aviaires dénote un certain type de paysage, favorable à leur implantation), et (iii) le paysage vu par des capteurs (quelques centaines de bandes hyperspectrales, ainsi que du Lidar qui mesure la hauteur des éléments au sol).



L'objectif du stage sera d'intégrer cet ensemble de données hétérogènes, bruitées, composé aussi de mesures indirectes, afin d'obtenir une vue d'ensemble qui soit parlante pour les agroécologues. La littérature scientifique propose des outils statistiques, tels que l'Analyse Factorielle Multiple (AFM), la Regularized Generalized Canonical Correlation Analysis (RGCCA), la PLS Path Modelling (PLS-PM), les Structural Equation Modelling (SEM) par exemple. La PLS-PM a donné des résultats satisfaisants sur la partie « agricole » du projet Muesli, nous proposons dans ce stage de faire de même sur les données « forêts » de Muesli.

Dans un deuxième temps, selon les affinités de l'étudiant(e), nous approfondirons un des points suivants. La première piste concerne l'imputation des valeurs manquantes et leur influence dans l'analyse. La deuxième piste sera centrée sur les données hyperspectrales : comment proposer des statistiques résumant le paysage autour de chaque parcelle agricole à partir de centaines de milliers de pixels décrits par 300 bandes spectrales. Une troisième possibilité concerne la comparaison de plusieurs approches statistiques, notamment la PLS-PM et le SEM.

Un sujet de thèse sera déposé cette année sur ce sujet, en prolongement du stage.

CARACTERISTIQUES DU STAGE

Durée et conditions du stage : 6 mois (entre janvier - août 2018, dates négociables) / 580 euros par mois environ

Structure d'accueil : UMR 1201 Dynafor, INRA Toulouse.

PROFIL RECHERCHE

Nous recherchons un(e) étudiant(e) en M2 ayant de solides bases en statistiques, de l'expérience du logiciel R, ayant le goût pour l'écologie et l'agroécologie.

Merci d'envoyer CV, lettre de motivation et personnes référentes à michel.goulard@inra.fr et magali.san-cristobal@inra.fr.

Les candidatures seront considérées dès leur réception et jusqu'au 30 novembre au plus tard.

