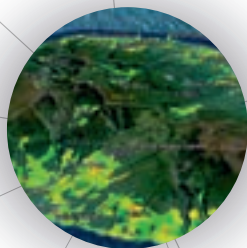


les dossiers
d'AGROPOLIS
INTERNATIONAL

*Compétences de la communauté scientifique
en région Occitanie*

SYSTÈMES COMPLEXES
de la biologie aux territoires



Numéro 23
Juin 2018

Modèles probabilistes et méthodes d'inférence statistique : identification de patterns de développement dans les données de phénotypage de plantes

Les progrès récents en agronomie reposent sur la révolution génomique couplée avec des avancées technologiques permettant d'acquérir de nouvelles données de phénotypage de plantes à des échelles variées. De nouvelles approches sont nécessaires pour caractériser les processus de développement de la plante, depuis les échelles tissulaires jusqu'à celle de la plante entière. Les plateformes robotisées de phénotypage produisent des données spatio-temporelles de développement de plantes à haut débit et permettent d'étudier l'effet de facteurs environnementaux et génétiques sur le développement des plantes dans un contexte renouvelé. Les méthodes d'analyse des données de phénotypage de plantes sont en pleine mutation du fait de la disponibilité de données massives à faible coût, de l'augmentation des puissances de calcul et de l'émergence de nouveaux paradigmes mathématiques à l'interface entre modélisation probabiliste, inférence statistique et reconnaissance des formes. Les données de phénotypage sont intrinsèquement structurées (spatio-temporelles, multi-échelles), ce qui nécessite de développer des modèles probabilistes et des méthodes d'inférence appropriés pour les analyser.

Ce nouveau paradigme pour analyser de telles données est caractérisé par :

- Des données structurées en séquences, séries temporelles, arborescences, graphes ou images 2D et 3D.
- Des modèles intégratifs qui combinent à la fois des variables réponses, des variables explicatives et des variables latentes pour modéliser des structures à différentes échelles (par ex. des stades de développement ou des phases de croissance). Des dépendances complexes entre ces variables souvent indexées par une structure comme une série temporelle ou une arborescence peuvent être représentées par des modèles graphiques probabilistes.
- Un cadre d'inférence étendu qui concerne l'estimation de paramètres mais aussi la structure du modèle (en utilisant des techniques de sélection de modèles) ainsi que des structures latentes (par ex. la segmentation en stades de développement ou en phases de croissance).

Contact : Y. Guédon (UMR AGAP), yann.guedon@cirad.fr

Modèles « structures-fonctions » des plantes

Les demandes sociétales pour une agriculture plus écologique et durable dans un contexte de changement climatique exigent une meilleure compréhension du fonctionnement et du développement des plantes. Pour cela, une approche prenant en compte explicitement la structure en croissance des plantes pour simuler l'interaction complexe entre les processus qui gouvernent cette croissance et la structure s'est développée depuis quelques années, appelée *Functional Structural Plant Model* (FSPM). Cette approche transdisciplinaire à l'interface de la biologie, l'informatique et la physique, permet d'intégrer, sous la forme d'un modèle 3D dynamique, des connaissances sur la morphologie et le fonctionnement des plantes. Elle permet notamment d'étudier le rôle de la structure des plantes sur différents processus écophysologiques tels que l'interception de la lumière ou la propagation de maladies, pour la partie aérienne, ainsi que les relations racine/sol. Elle nécessite donc d'assembler les différents outils et modèles hétérogènes développés par la communauté scientifique, en utilisant par exemple les méthodes

proposées par la plateforme OpenAlea. L'équipe Virtual Plants (Inria/Cirad) participe à l'élaboration de nombreux outils pour la simulation de la croissance et de différents processus écophysologiques tels que ceux mécaniques, de captation de la lumière, d'allocation de carbone, etc. Elle développe ou collabore également à la construction de différents modèles FSPM, représentant les parties aériennes ou souterraines et allant des herbacées (blé, maïs, sorgho) aux arbres fruitiers (manguier, pommier). Les défis en cours autour de ces problématiques sont de permettre une utilisation en ligne de ces outils complexes pour des simulations distribuées sur le web ainsi que la standardisation d'un certain nombre d'outils de simulation de processus écophysologiques. D'un point de vue applicatif, les défis à venir sont l'intégration de ces modèles à différentes échelles dans des modèles de culture, pour prendre en compte à la fois la variabilité génétique et l'hétérogénéité spatiale des populations simulées.

Contact : F. Boudon (UMR AGAP/Équipe-projet Virtual Plants), frederic.boudon@cirad.fr

◀ *Simulation stochastique d'un manguier. Différents stades phénologiques de la croissance sont représentés avec l'apparition de nouvelles pousses feuillées (en haut à droite), d'inflorescences (en bas à gauche) et la croissance des fruits (en bas à droite).*
© F. Boudon/Cirad/Inria

