

FORMULAIRE A RENSEIGNER
ET A RETOURNER AUX DIRECTIONS DES DEPARTEMENTS
POUR LE 9 JANVIER 2017

INTITULE DU SUJET DE THESE

Méthodes statistiques pour optimiser les stratégies d'épidémiologie végétale

CONTEXTE ET PROBLEMATIQUE

L'efficacité des réseaux d'épidémiologie repose sur l'utilisation de méthodes statistiques appropriées. Ces réseaux ont deux grands types d'objectifs : détecter précocement d'éventuelles maladies émergentes, et surveiller le développement de maladies pour aider à les contrôler. Lorsque l'objectif est la détection précoce, l'échantillonnage adaptatif en grappes peut être utilisé même si la maladie n'a pas encore été détectée dans la zone de surveillance. Il présente une meilleure efficacité que les plans classiques dans le cas d'événements rares et spatialement agrégés [1,2,3]. Cependant le choix des nouveaux sites à échantillonner après détection de la maladie sur un site ne prend en compte que les sites voisins et ne repose sur aucun modèle. Lorsque l'objectif est la surveillance du développement d'une maladie, il est possible d'échantillonner en priorité les sites à haut risque [4]. Les méthodes d'échantillonnage basé sur le risque reposent en général sur des modèles de dynamique épidémique qui incluent le cycle de vie du pathogène et sa fonction de dispersion. Le risque en chaque site peut être calculé au moyen d'un modèle dont les paramètres sont estimés à partir des données, ou bien en utilisant des cartes de risque [5,6]. Ces méthodes présentent cependant des lacunes. D'une part, elles considèrent le risque d'arrivée d'une maladie sur un site, mais pas le risque qu'un site devienne source de contamination ; or l'effort d'échantillonnage pourrait aussi porter sur les sites dont la contamination est susceptible d'avoir un impact important sur le développement ultérieur de la maladie. D'autre part, les cartes de risque ne sont pas issues d'un processus d'assimilation des données permettant la mise à jour du modèle en temps réel. Enfin, les plans d'échantillonnage pré- et post-émergence sont étudiés séparément et aucune étude ne porte sur l'allocation optimale des ressources entre ces deux phases.

RESUME DU TRAVAIL PROPOSE

Questions de recherche

Trois questions principales seront abordées :

- 1) A quelles conditions (en termes de ressources, types de paysage et paramètres épidémiologiques) les programmes d'épidémiologie génériques (sans *a priori*) ou ciblés sur un pathogène peuvent-ils apporter des informations opportunes sur un pathogène émergent ?
- 2) Dans quelle mesure les méthodes d'assimilation de données en temps réel permettent-elles d'améliorer l'efficacité de la surveillance basée sur le risque associé à un pathogène connu ?
- 3) Comment optimiser l'allocation spatio-temporelle des ressources entre surveillance pré- et post-émergence sous divers scénarios épidémiologiques ?

Objectifs

L'objectif de la thèse est de développer, adapter et tester des méthodes statistiques pour optimiser l'épidémiologie des maladies végétales. Les recherches s'appuieront sur des méthodes récemment développées non seulement dans le domaine végétal, mais aussi dans le domaine animal en adaptant ces dernières à des pathosystèmes végétaux.

Des méthodes d'échantillonnage seront développées et optimisées dans l'objectif de détecter l'arrivée d'une épidémie ou d'évaluer la prévalence d'une maladie. Des méthodes d'estimation de paramètres épidémiologiques (distances de dispersion, taux de reproduction de base R_0) après détection d'une maladie seront mises au point.

La prise en compte de l'information spatialisée permettra d'établir des cartes de risque, d'évaluer l'impact de la dynamique épidémique sur l'effort de surveillance (fréquence et taux d'échantillonnage) et d'optimiser le rapport entre l'effort de gestion et l'effort d'échantillonnage.

Les questions spécifiques aux épidémies multi-pathogènes seront abordées sous l'angle de la recherche de facteurs de risque commun et de l'évaluation de la robustesse de la méthode d'échantillonnage en cas d'arrivée d'un nouveau pathogène.

La stratégie de surveillance visera à optimiser les parts relatives consacrées à la gestion de la maladie et à l'acquisition de nouvelles données permettant de mieux connaître la maladie, incluant les covariables à observer.

Les méthodes développées reposeront en partie sur des résultats théoriques et utiliseront des techniques de modélisation et de simulation numérique pour évaluer l'efficacité des différentes stratégies d'échantillonnage en fonction des caractéristiques du pathosystème. Les méthodes seront génériques avec des études de cas sur certains pathosystèmes, en particulier le modèle Prunus / *Plum pox virus* et le modèle Bananier / *Mycosphaerella fijiensis*.

Méthodologie

Une première étape consistera à développer des modèles spatio-temporels permettant de simuler des épidémies selon divers scénarios avec des valeurs fixées des paramètres du pathogène (distance de dispersion, durée de la période de latence, R_0) et du paysage (caractéristiques spatiales des parcelles et des plants dans les parcelles, caractéristiques de résistance des variétés...). Le réalisme des données simulées sera évalué en les comparant à divers jeux de données disponibles (cercosporioses du bananier, sharka des *Prunus*). Les données simulées seront ensuite utilisées dans un double objectif. D'une part, elles serviront à évaluer et améliorer les procédures d'estimation des paramètres en se basant sur leur capacité à ré-estimer les paramètres épidémiologiques utilisés dans les simulations. D'autre part, comme les données auront été simulées avec des paramètres épidémiologiques et des types de paysage connus, elles serviront de base à l'évaluation de diverses stratégies d'épidémiosurveillance et donc au choix de la stratégie optimale selon les paramètres épidémiologiques et le type de paysage. Plus spécifiquement, les stratégies testées incluront différentes méthodes d'échantillonnage (simple, stratifié, adaptatif, basé sur le risque), différents modes d'allocation des ressources dans l'espace (localisation et densité des sites de prospection) et dans le temps (effort d'échantillonnage pré- et post-émergence, fréquence des prospections) et différentes probabilités de détection du pathogène. Des stratégies de surveillance basées sur des techniques d'assimilation de données permettant de réévaluer en temps réel les paramètres des modèles et le risque spatialisé seront développées et testées sur les données simulées.

Données disponibles

Le sujet est essentiellement méthodologique mais s'appuiera sur plusieurs jeux de données. Ces données permettront d'estimer les paramètres épidémiologiques relatifs aux pathosystèmes considérés.

Cercosporiose noire du bananier. Deux jeux de données sur cette maladie ont été recueillis dans les Caraïbes (Martinique et Guadeloupe). Ils sont constitués de relevés d'état sanitaire datés et géo-référencés et de données environnementales spatialisées (parcellaire et données météorologiques).

Sharka des Prunus. Cette maladie soumise à une lutte obligatoire fait l'objet de prospections systématiques. A l'échelle du paysage, les données disponibles comprennent les coordonnées géographiques de parcelles de pêcheurs ou d'abricotiers prospectées et les effectifs des arbres de chaque parcelle selon leur état sanitaire (symptomatique ou asymptomatique). A l'échelle de la parcelle, des données ont été collectées dans un domaine de plus de 300 ha durant 6 ans.

Positionnement dans la stratégie du Cirad et de l'unité

Ces recherches relèvent de l'axe stratégique 4 du Cirad "Comprendre, anticiper et gérer les risques liés aux bio-agresseurs pour renforcer la santé des animaux et des plantes" et plus spécifiquement de la thématique "Maîtriser durablement les risques liés à la santé des animaux et des plantes". Elles s'intègrent aux actions de modélisation en épidémiologie conduites au sein de l'axe « Processus évolutifs et épidémiologiques » du projet de l'UMR BGPI.

Références citées

[1] P. S. Ojiambo and H. Scherm, "Efficiency of adaptive cluster sampling for estimating plant disease

- incidence.," *Phytopathology*, vol. 100, no. 7, pp. 663–70, Jul. 2010.
- [2] S. A. Gattone and T. Di Battista, "Adaptive cluster sampling with a data driven stopping rule," *Stat. Methods Appl.*, vol. 20, no. 1, pp. 1–21, Oct. 2010.
- [3] S. A. Gattone, M. Esha, and J. W. Mwangi, "Application of Adaptive Cluster Sampling with a Data-Driven Stopping Rule to Plant Disease Incidence," *J. Phytopathol.*, vol. 161, no. 9, pp. 632–641, Sep. 2013.
- [4] S. Parnell, T. R. Gottwald, T. Riley, and F. Van Den Bosch, "A generic risk-based surveying method for invading plant pathogens," *Ecol. Appl.*, vol. 24, no. 4, pp. 779–790, 2014.
- [5] W. Luo, S. Pietravalle, S. Parnell, F. Van den Bosch, T. R. Gottwald, M. S. Irej, and S. R. Parker, "An improved regulatory sampling method for mapping and representing plant disease from a limited number of samples," *Epidemics*, vol. 4, no. 2, pp. 68–77, Jun. 2012.
- [6] S. Parnell, T. R. Gottwald, M. S. Irej, W. Luo, and F. van den Bosch, "A Stochastic Optimization Method to Estimate the Spatial Distribution of a Pathogen from a Sample," *Phytopathology*, vol. 101, no. 10, pp. 1184–1190, Oct. 2011.

LISTE DE QUELQUES PUBLICATIONS DE L'EQUIPE PROPOSANT LE SUJET

Rimbaud L. (2015) *Conception et évaluation assistée par la modélisation de stratégies de gestion d'une épidémie dans un paysage hétérogène*. Thèse de Doctorat, Montpellier SupAgro.

Rimbaud L., Dallot S., Gottwald T.R., Decroocq V., Soubeyrand S., Jacquot E., Thébaud G. (2015) Sharka epidemiology and worldwide management strategies: learning lessons to optimize the control of plant virus diseases. *Annual Review of Phytopathology* 53: 357-378.

Landry C. (2015) *Modélisation des dynamiques de maladies foliaires de cultures pérennes tropicales à différentes échelles spatiales : cas de la cercosporiose noire du bananier*. Thèse de Doctorat, Université des Antilles.

Rieux A., Soubeyrand S., Bonnot F., Klein E. K., Ngando J. E., Mehl A., Ravigné V., Carlier J., de Lapeyre De Bellaire L. (2014) Long-distance wind-dispersal of spores in a fungal plant pathogen: Estimation of anisotropic dispersal kernels from an extensive field experiment. *PLoS ONE* 9(8):e103225.

Morelli M.J., **Thébaud G.**, Chadœuf J., King D.P., Haydon D.T., Soubeyrand S. (2012) A Bayesian inference framework to reconstruct transmission trees using epidemiological and genetic data. *PLoS Computational Biology* 8(11): e1002768.

Oro F. Z., Bonnot F., Ngo-Bieng M.A., Delaitre E., Dufour B. P., Ametefe K. E., Mississo E., Wegbe K., Muller E., Cilas C. (2012) Spatiotemporal pattern analysis of Cacao swollen shoot virus in experimental plots in Togo. *Plant Pathology* 61(6):1043–1051.

Bonnot F., de Franqueville H., Lourenço E. (2010) Spatial and spatiotemporal pattern analysis of coconut lethal yellowing in Mozambique. *Phytopathology* 100(1):300–312.

ECOLE DOCTORALE D'INSCRIPTION PROPOSEE (nom et numéro)

GAIA (Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau) – ED 584

DIRECTEUR OU CO-DIRECTEUR DE THESE CIRAD (en précisant s'il a donné son accord)

La thèse sera encadrée par François Bonnot (CIRAD, UMR BGPI) (a donné son accord).

DIRECTEUR DE THESE NON CIRAD

Emmanuel Jacquot, INRA, UMR BGPI (a donné son accord)

MODALITES D'ACCUEIL AU CIRAD

Le (la) doctorant(e) sera accueilli(e) au sein de l'UMR BGPI.

COMPOSITION PREVISIONNELLE DU COMITE DE THESE

Emmanuel Jacquot (INRA), François Bonnot (CIRAD), Gaël Thébaud (INRA), Claire Neema (Montpellier SupAgro), Samuel Soubeyrand (INRA), Vladimir Grosbois (CIRAD), Pascal Hendrikx (ANSES), Jean-François Guégan (IRD).

CO-FINANCEMENTS ENVISAGES ET/OU OBTENUS (cofinancement de bourse)

Financement obtenu d'une demi-bourse de thèse par la Fondation Agropolis dans le cadre du projet étandard « E-SPACE » (2015-2019).

FINANCEMENT DU FONCTIONNEMENT DE LA THESE (nom du projet auquel est rattaché le sujet de la thèse, ...)

Financement par le projet étandard « E-SPACE » (Fondation Agropolis).

NOUVEAUX EQUIPEMENTS REQUIS

Un ordinateur.

PARTENARIAT NORD ET SUD ENVISAGE

Le doctorant sera en contact avec les structures de Guadeloupe et de Martinique (FREDON) pour l'accès aux données acquises. Le travail de thèse sera réalisé en France métropolitaine.

POSITIONNEMENT DU DOCTORANT SUR UN DISPOSITIF EN PARTENARIAT

NON

LIEN AVEC LA LPO DE L'UNITE DE RECHERCHE à préciser :

La thèse se positionne dans le projet étandard épidémiosurveillance « E-SPACE » (Fondation Agropolis) qui fait partie des projets de l'UMR dans le cadre de l'objectif opérationnel "Production de connaissances et expertises" de la LPO. Elle s'intègre dans l'axe stratégique 4 du Cirad "Comprendre, anticiper et gérer les risques liés aux bio-agresseurs pour renforcer la santé des animaux et des plantes", axe majeur auquel contribue l'UMR.