
INFÉRER ET PRÉDIRE LES DYNAMIQUES D'ESPÈCES INVASIVES, FOCUS SUR *XYLELLA FASTIDIOSA*

Thèse de Candy **ABBOUD**¹

Analysée par Philippe **STOOP**²

Directeurs de thèse : Samuel SOUBEYRAND (Directeur de recherche INRA - Avignon) et Éric PARENT (Professeur AgroParisTech/INRA - Paris)

Cette thèse de mathématiques appliquées porte sur un sujet particulièrement préoccupant pour l'agriculture actuelle : l'apparition récurrente d'espèces ravageuses invasives, favorisée par le développement des échanges internationaux. L'importance historique de ce phénomène est apparue dès le 19^{ème} siècle, avec l'apparition dans les vignes européennes du mildiou et du phylloxera, importés d'Amérique. Depuis cette époque, de nombreux autres cas ont été observés et ce phénomène n'impacte plus seulement l'agriculture, comme en témoignent les cas récents, et bien connus du grand public, de la pyrale du buis, du frelon asiatique ou du moustique-tigre.

Face à ces invasions récurrentes, et souvent dévastatrices pour les cultures concernées, il devient urgent de développer des modèles épidémiologiques permettant de prévoir l'extension géographique de ces espèces à partir de leurs foyers d'apparition. Cela permettrait de mieux cibler les opérations de suivi et de prophylaxie (laquelle se réduit le plus souvent à la destruction des plantes atteintes, d'où la nécessité d'un bon ciblage), et, si possible, de prévoir la direction et la vitesse de l'extension géographique de ces foyers. C'est l'objectif ambitieux de cette thèse, qui vise à développer dans ce but des modèles génériques, mais en partant du cas d'étude de *Xylella fastidiosa*, une bactérie particulièrement redoutable pour l'olivier, apparue récemment en Italie, puis en Corse, alors qu'elle épargnait l'Europe jusqu'à présent.

L'introduction montre bien toute la difficulté du sujet. En effet, cette bactérie est particulièrement polyphage : plus de 300 espèces végétales peuvent en être l'hôte. De plus, pas moins d'une douzaine d'espèces d'insectes sont soupçonnées d'en être les vecteurs. Le peu de données biologiques connues ne permet donc pas d'identifier *a priori* les variables environnementales qui vont influencer sur la diffusion de la bactérie. De plus, la variété des hôtes et vecteurs potentiels ne permet pas d'exclure l'hypothèse de contaminations multiples. Enfin, l'historique des données disponibles est réduit, puisque les suivis sur le terrain dont disposait la chercheuse n'ont commencé qu'en 2015. Avoir si peu de recul dans le temps complique singulièrement la tâche, d'autant plus que, comme on le découvre dans le courant de la thèse,

¹ Thèse de doctorat de l'Université d'Aix Marseille Université, Ecole doctorale 184 – Mathématiques et Informatique de Marseille, UR INRA Biostatistique et Processus spatiaux (UR 546, BioSP), soutenue le 11/12/2019.

² Membre correspondant de l'Académie d'agriculture de France, section 9, « Agrofournitures ».

il est probable que la bactérie a été introduite en Corse plusieurs décennies avant d'y être détectée. La gestion des incertitudes est donc un point crucial de ce travail de modélisation.

Face à la difficulté de son sujet, Candy Abboud mobilise méthodiquement des outils mathématiques très variés, afin d'élargir progressivement les problématiques traitées et d'explorer de nouvelles hypothèses, dans un aller-retour constant entre les considérations théoriques et les résultats appliqués. Précisons tout de suite, que compte tenu du haut niveau de ce travail, juger de la pertinence et de l'originalité de la démarche mathématique suivie demande des compétences dont ne dispose pas l'auteur de cette note. Sur ce point, nous nous reposons sur les appréciations élogieuses des encadrants de la thèse. Ce qui a retenu notre attention, en tant qu'agronome, est la façon dont la doctorante a su concilier la valeur théorique de son travail et son caractère appliqué, clairement exposés :

- D'abord par ses questions de recherche :
 - Comment inférer les conditions initiales (lieu et date de la première introduction de l'espèce invasive) à partir d'un modèle à composantes mécanistes et stochastiques ?
 - Comment étendre ce cadre de modélisation et d'inférence, en combinant différents modèles dans une approche de Bayesian Model averaging ?
 - Comment introduire dans ce cadre le traitement d'introductions multiples de l'espèce invasive ?
- Et par ses questions épidémiologiques :
 - Où et quand ont été introduites les différentes souches de *X. fastidiosa* observées en Corse ?
 - Quelle est la vitesse de propagation de la bactérie ?
 - Quel est l'impact des températures hivernales sur sa dynamique ?
 - Comment va évoluer dans le futur sa distribution spatiale ?

Dans un premier temps (chapitre 3), C. Abboud se concentre sur les données du sud-ouest de la Corse, lieu de la première détection de *X. fastidiosa*. Cela lui permet de faire l'hypothèse d'une contamination initiale unique, dont elle modélise l'expansion avec une approche mixte (mécaniste et statistique) de réaction - diffusion - absorption, conditionnée par un seuil de températures minimales hivernales, qu'elle valide avec les données de suivi de 2015 à 2017. Cette première étape lui permet de confirmer l'importance des températures minimales hivernales comme critère de la possibilité d'installation de *X. fastidiosa*. Son étude lui permet d'établir que seules des températures minimales de janvier et février supérieures à 5°C permettent l'implantation de la bactérie, 5,4°C étant la valeur qui permet le meilleur accord entre le modèle et les observations de terrain. A l'intérieur de la zone favorable définie par ce seuil, les zones atteintes ont progressé en moyenne de 155 m par mois. Cela l'amène à considérer que la contamination initiale aurait eu lieu près d'Ajaccio vers 1959 : une estimation beaucoup plus ancienne que les dates précédemment avancées d'après l'analyse génétique des populations observées (1965 et 1980), mais qui reste compatible avec elles, toutes ces estimations ayant des intervalles de confiance très larges. Il en résulte que le calcul sur la vitesse d'expansion du front épidémiologique est lui-même soumis à de fortes incertitudes, ce qui a de fortes incidences sur les mesures à prendre en matière de suivi de l'épidémie. Cela amène C. Abboud à tenter d'améliorer la précision des paramètres épidémiologiques, question qu'elle aborde au chapitre suivant.

Dans le chapitre 4, la doctorante utilise la technique du Bayesian Model Averaging (BMA) pour rechercher les meilleures valeurs des différents paramètres de modélisation de l'expansion de *X. fastidiosa*, issues de la comparaison de différentes familles de modèles mécanistes et statistiques. Cette nouvelle approche confirme globalement les résultats précédents : la meilleure température seuil trouvée est de 5,5°C, et la date de contamination initiale obtenue est encore plus ancienne (1952) que dans le chapitre précédent, mais reste compatible avec les estimations précédentes, l'intervalle de confiance étant une fois encore très large. D'un point de vue théorique, le principal enseignement de ce chapitre est en effet que, contrairement aux attentes, le BMA n'a pas permis d'obtenir de meilleurs résultats que l'optimisation d'un modèle unique réalisée au chapitre 3.

Dans le 5^{ème} chapitre, C. Abboud passe à une nouvelle étape, avec un modèle permettant de prendre en compte des introductions multiples du pathogène. Elle s'inspire pour cela des modèles de colonies coalescentes, couramment utilisés pour décrire ce type d'invasions biologiques, et démontre qu'ils peuvent être traités comme des Processus de Markov Déterministes par Morceaux (PDMP) spatio-temporels. Elle applique cette fois ce cadre à l'ensemble des données corses (y compris la Corse du Nord), où deux souches différentes de *X. fastidiosa* ont été identifiées. Ce chapitre reste un peu inachevé, dans la mesure où les résultats ne sont pas parfaitement concluants (il reste des divergences notables entre le modèle et la réalité, en particulier sur la côte orientale de la Corse), mais c'est bien compréhensible vu la complexité du cas d'étude qui lui a été fixé. Ce chapitre apporte néanmoins des enseignements intéressants : ce nouveau modèle trouve des paramètres de réaction-diffusion-absorption proches des valeurs obtenues au chapitre 3, ce qui laisse penser que ces valeurs sont assez robustes. L'incertitude sur ces paramètres est inférieure à celle obtenue au chapitre 3, mais il n'a pas été possible de vérifier si cela tient au modèle lui-même, ou simplement au fait que plus de données ont pu être utilisées sur cette nouvelle étude plus tardive, et sur une aire géographique plus large.

Dans son chapitre 6, C. Abboud détaille les nombreuses perspectives ouvertes par son travail, tant sur le plan méthodologique que sur les applications pratiques pour la surveillance et la gestion de cette espèce invasive. Cela va de l'amélioration des cartes de risque à l'élaboration de nouvelles stratégies d'échantillonnage pour un suivi mieux ciblé sur le terrain. Cette question est d'autant plus d'actualité que, depuis le début de ce travail, *X. fastidiosa* a atteint la Provence, et va donc nécessiter une biosurveillance à bien plus grande échelle.

Ce compte-rendu rapide peut donner l'impression que cette thèse ouvre plus de questions qu'elle n'en résout, mais cela tient à la richesse du travail théorique réalisé, par rapport au caractère encore assez lacunaire et à la faible profondeur temporelle des données disponibles sur ce ravageur d'arrivée récente. Son apport majeur est de confirmer, par plusieurs méthodes différentes, une hypothèse assez contre-intuitive : *X. fastidiosa* est très probablement présente en Corse depuis plusieurs décennies, suite à un très petit nombre d'introductions restées longtemps inaperçues. La coïncidence temporelle de son apparition en Corse avec sa première détection en Italie n'en est que plus fascinante, et il sera passionnant de voir ce que donnent les modèles développés ici sur les données épidémiologiques italiennes. On peut aussi penser que leur application à des espèces moins polyphages, et d'apparition plus ancienne, permettront d'affiner ces modèles sur des cas d'étude moins difficiles. Beaucoup de questions restent donc à éclaircir, mais la balle est maintenant dans le camp des spécialistes de pathologie végétale et des entomologistes agricoles, qui pourront, par exemple, utiliser les seuils de température hivernale définis ici, afin de resserrer leur filet autour des espèces d'insectes vecteurs potentiels.

Il est important de souligner ici que, malgré la grande complexité des concepts mathématiques mobilisés dans ce travail, la lecture de cette thèse reste agréable, grâce à sa rédaction très pédagogique. On apprécie en particulier l'introduction de chaque chapitre par un résumé graphique, qui expose de façon très claire le raisonnement suivi, et, en fin de chapitre, les perspectives ouvertes qui montrent bien comment les résultats de chaque étape conditionnent le chapitre suivant. On voit très clairement ici une démarche scientifique se construire progressivement, par la confrontation permanente entre les concepts théoriques et les données à interpréter. Cette approche respectueuse des données, et donc des faits, tranche très agréablement avec de trop nombreuses publications, où des experts en mathématiques appliquées plaquent mécaniquement à l'agronomie des méthodes statistiques mises au point pour d'autres disciplines, sans jamais s'interroger sur la plausibilité biologique ou agronomique de leurs résultats. Cette thèse constitue donc une lecture intellectuellement très stimulante pour les agronomes spécialistes de la protection des cultures, et ouvre de nombreuses pistes pour des collaborations plus poussées entre biologistes et mathématiciens. De plus, l'actualité récente lui donne une résonance particulière : à l'heure où tous les citoyens sont sensibilisés par l'épidémie de Covid-19 à l'importance d'une modélisation fiable de la dynamique des épidémies, il est agréable de voir que l'agriculture dispose dans ce domaine de compétences qui n'ont rien à envier à celles de l'épidémiologie humaine.

A ce titre, l'approche novatrice, originale de cet excellent travail de thèse mérite qu'il soit valorisé par la mise sur le site de l'Académie d'agriculture de France de cette analyse.