

Prise en compte de la structure du paysage pour la surveillance épidémiologique du HLB sur l'île de La Réunion et exploration d'une nouvelle méthode de diagnostic par Spectroscopie Proche Infra-Rouge



Présenté par **Nathan Créquy**

Dans le cadre d'un **stage d'année de césure** effectué du **06/01/2020 au 14/08/2020**

Au sein du **CIRAD, UMR PVBMT** (Peuplements Végétaux et Bioagresseurs en Milieu Tropical)

Encadrement :

Virginie Ravigné, maître de stage

Olivier Pruvost, maître de stage

Karine Boyer

Christophe Le May, professeur tuteur de césure à Agrocampus Ouest

Remerciements :

Je tiens à remercier les personnes qui ont contribué au bon déroulement de ce stage :

- Virginie RAVIGNE, ma responsable de stage, pour sa gentillesse, ses précieux conseils et pour m'avoir donné la liberté de développer une thématique au cours de mon stage
- Olivier PRUVOST, mon responsable de stage, pour son aide en microbiologie et son encadrement au laboratoire
- Karine BOYER, pour m'avoir formé aux techniques de laboratoire
- Christophe LE MAY, mon tuteur de césure, pour son soutien en dépit de la rareté des nouvelles que je lui ai donné
- Ismaël Houillon, doctorant, pour ses conseils avisés et son amitié partagée
- Simon Hoarau, agriculteur sur la commune de Petite-Ile, pour m'avoir autorisé à mener des expérimentations sur ses parcelles
- Marine WEISHAAR, qui m'a précédée dans ce stage, pour son amitié et pour m'avoir présenté les méthodes de travail sur le terrain et au laboratoire
- Frédéric CHIROLEU et Anna DOIZY, pour leur aide en statistiques
- Emmanuel TILLARD, directeur du pôle élevage, pour la confiance qu'il m'a témoignée en me permettant d'utiliser les instruments de mesures de spectrométrie
- Serge NABANEZA, pour ses conseils et sa formation à l'utilisation des instruments de mesures de spectrométrie
- Tous les stagiaires au Cirad, pour leur accueil et leur amitié qui ont rendu ce stage et ce confinement inoubliables !

Table des matières :

1. Introduction	4
2. Matériels et méthodes	9
2.1. Prédiction du parcellaire agrumicole réunionnais	9
2.1.1. Connaissance du parcellaire agrumicole réunionnais	9
2.1.2. Constitution de la base de données	9
2.1.3. Prédiction de la nature des parcelles	9
2.2. Cartographier le risque de présence du HLB	10
2.2.1. Plan de surveillance et échantillonnage	10
2.2.2. Analyse statistique : probabilité de présence de la maladie	11
2.3. Prédiction du statut infecté des arbres par Spectroscopie Proche Infra-Rouge	11
2.3.1. Principe de la Spectroscopie Proche Infra-Rouge	11
2.3.2. Echantillonnage	11
2.3.3. Prise de données	12
2.3.4. Analyses moléculaires et sérologiques	15
2.3.4.1. Détermination du statut HLB des lots par PCR quantitative	15
2.3.4.2. Diagnostic du virus de la tristeza (CTV)	16
2.3.5. Classification par <i>machine learning</i> sous R	16
3. Résultats	17
3.1. Prédiction du parcellaire réunionnais	17
3.2. Cartographier le risque de présence du HLB	19
3.3. Signature spectrale de la maladie du HLB	21
3.3.1. Spectres proche-infrarouge	21
3.3.2. Prédiction du statut infecté des arbres par Spectroscopie Proche Infra-Rouge (SPIR)	22
3.3.3. Identification des longueurs d'onde caractéristiques de la maladie	23
4. Discussion	26
4.1. Prédiction du parcellaire agrumicole	26
4.2. Prédiction de l'état sanitaire des parcelles et cartographie du risque HLB	26
4.3. Prédiction de l'état sanitaire des arbres par SPIR	27
5. Conclusion	29
6. Bibliographie	30

1. Introduction

L'émergence de nouvelles maladies chez les plantes cultivées est au cœur des préoccupations actuelles. En effet, l'intensification des échanges commerciaux et des flux de voyageurs facilite le transport de matériel végétal infecté vers des bassins de productions jusque-là considérés comme indemnes (Jones & Baker, 2007; Brasier, 2008). Certains pathogènes voient également leur niche écologique modifiée en réponse au changement climatique (Anderson et al., 2004; Pautasso et al., 2012) et leurs impacts sur la production et la sécurité alimentaire de certains pays peuvent être alarmants. Par exemple, la culture du manioc, 4^{ème} source de calories à l'échelle mondiale, est aujourd'hui grandement menacée en Afrique par la propagation de deux virus émergents, le virus de la striure brune du manioc et celui de la mosaïque du manioc (Graziosi et al., 2016; Legg et al., 2014). L'augmentation des rendements agricoles est nécessaire pour répondre aux besoins d'une population mondiale grandissante (Gilland, 2002). C'est pourquoi il est nécessaire d'améliorer ou de stabiliser la protection des cultures à son niveau de performance actuel. Le recours massif aux produits phytosanitaires a déjà par le passé montré ses limites. Leurs impacts négatifs sur la santé humaine, la faune et la flore et même sur la durabilité de la production agricole, notamment par la destruction des auxiliaires des cultures ont été démontrés (Wilson & Tisdell, 2001). De plus, la protection qu'ils offrent n'est pas pérenne en raison de la rapide sélection de résistances chez les organismes nuisibles cibles (Gullino et al., 2000; Ma & Michailides, 2005; Wyenandt et al., 2018). La protection des cultures doit donc se pencher sur de nouveaux leviers au premier rang desquels la prévention, qui consiste à anticiper et surveiller les risques épidémiologiques afin de les traiter précocement.

L'épidémiologie étudie l'effet sur la dynamique spatio-temporelle d'une maladie de la structure du paysage, c'est-à-dire à la fois la composition, la disposition et les liens entre les diverses entités paysagères. La transmission d'un pathogène d'un hôte infecté vers un hôte sain implique des facteurs biotiques (via un vecteur arthropode par exemple) ou abiotiques (pluie, vent par exemple). Dans la plupart des cas, la probabilité de transmission diminue avec la distance entre les hôtes, c'est pourquoi la localisation spatiale des hôtes, du pathogène et des vecteurs ont une influence sur la dynamique spatiale de la maladie (Ostfeld et al., 2005). En identifiant les impacts du paysage sur la dynamique épidémique, l'épidémiologie spatiale offre la possibilité de calibrer la surveillance et orienter la gestion des maladies de façon optimale dans le paysage agricole.

L'épidémiologie du paysage puise ses concepts dans l'écologie du paysage. Elle s'inspire de la théorie des métapopulations (Hanski, 2001) selon laquelle la distribution spatiale d'une espèce peut être décrite comme un ensemble de sous-populations distinctes connectées entre elles par la migration d'individus, et qui peuvent éventuellement s'éteindre et être recolonisées. Cette théorie est appliquée en épidémiologie végétale en considérant le pathogène comme la population étudiée et les plantes hôtes comme les différents îlots d'habitats au sein du paysage (Grenfell & Harwood, 1997). Les modèles épidémiologiques de métapopulation s'appuient sur la théorie des graphes en considérant les sous-populations (les hôtes ou les parcelles hôtes) comme un réseau de nœuds connectés entre eux par des liens représentant la probabilité de transmission. En épidémiologie du paysage, chaque lien peut

par exemple être affecté par un indice de connectivité, qui tient compte de la taille des sous-populations hôtes et des distances entre elles. Par exemple, il a été montré chez des silènes *Lychnis alpina* atteintes par le champignon castrateur *Microbotryum violaceum*, responsable du charbon des anthères, que la connectivité entre les sous-populations favorisait la propagation de la maladie dans un plus grand nombre de sites, mais diminuait l'impact de la maladie au sein de ces mêmes sites, en favorisant également la diffusion de génotypes hôtes résistants dans la métapopulation de silènes (Carlsson-Granér & Thrall, 2002). Ainsi ces modèles de métapopulation n'utilisent pas directement la localisation des sous-populations dans l'espace, c'est pourquoi ces méthodes sont qualifiées de « spatialement implicites ».

Au-delà de la prise en compte de la connectivité entre sous-populations, l'épidémiologie du paysage peut aussi tenir compte de la localisation et de la forme exactes des sous-populations d'hôtes. Ces approches sont alors qualifiées de « spatialement explicites ». En effet, l'idée est d'identifier les caractéristiques du paysage, à la fois les facteurs biotiques et abiotiques, associés à la présence et à la propagation d'un agent pathogène dans un environnement (Plantegenest et al., 2007; Nnadi Nnaemeka et al., 2011; Meentemeyer et al., 2012). Ces approches nécessitent de modéliser la niche écologique de l'agent pathogène et à produire des cartes de risque de sa présence en établissant des corrélations entre données de présence/absence d'une espèce avec des données caractérisant l'environnement (Martinetti & Soubeyrand, 2018). Elles peuvent également être complétées par un couplage de ces cartes de risque avec des modèles de propagation épidémiologique (Stephen Parnell et al., 2017; Martinetti & Soubeyrand, 2018).

La surveillance épidémiologique, c'est-à-dire le suivi de la propagation d'agents infectieux sur un territoire, est essentielle pour prévenir les risques sanitaires et prendre des mesures de gestion et de contrôle efficaces. Pour optimiser la surveillance à moindre coût, il est nécessaire de cartographier le risque de présence de la maladie. On peut alors se focaliser sur la détection de nouveaux foyers d'infection dans les zones à plus haut risque (Stephen Parnell et al., 2017). Les chances de contrôler voire d'éradiquer l'épidémie sont augmentées avec la précocité, la fréquence et l'efficacité des détections (Cook et al., 2008). Cependant, cette surveillance est limitée par des facteurs socio-économiques que sont les ressources, la compétence des acteurs et l'état de préparation (Rimbaud et al., 2015). Par exemple, l'état de préparation peut être très faible dans le cadre d'une maladie émergente. De même, les territoires avec peu de ressources auront de fait moins de moyens à allouer à la surveillance. Différentes méthodes pour établir des cartes de risque sont employées en fonction des caractéristiques du paysage (homogène ou fragmenté) et de la spécificité d'hôte du pathogène (généraliste ou spécialiste).

Le cadre statistique permettant de réaliser ces cartes de risque a été développé par Parnell et coll. (2014). Cette méthode requiert de pouvoir estimer le nombre de reproduction de base R_0 de la maladie (le nombre moyen de contaminations causées par chaque hôte infecté) en chaque point du paysage et la probabilité de contamination entre deux points. Les auteurs ont appliqué cette méthode sur la dynamique de la maladie du Huanglongbing affectant l'agrumiculture agroindustrielle en Floride, c'est-à-dire sur une bactérie pathogène spécialisée sur les agrumes, dans un paysage très homogène à grande échelle, et surtout très surveillé. Plus récemment, Martinetti et Soubeyrand (2018) ont réussi l'application de cette méthode à

la propagation de *Xylella fastidiosa*, bactérie pathogène très généraliste et moins connue, sur le territoire Corse, et en région PACA, des paysages très hétérogènes. Ils ont découpé le paysage en cellules de 1km² au sein desquelles ils ont compilé des variables pédoclimatiques. Ils ont ensuite mis en corrélation ces données avec la présence/absence de la bactérie afin de cartographier la probabilité de présence du pathogène. Cette cartographie du risque a ensuite été couplée à un modèle épidémiologique afin d'étudier la dynamique de propagation de la bactérie.

Dans ces études, la connaissance du parcellaire agricole, c'est-à-dire, la position exacte et la forme des parcelles agricoles, n'était pas forcément nécessaire pour étudier la dynamique épidémiologique, soit parce que le parcellaire était continu et homogène (comme en Floride), soit parce que l'agent pathogène étudié circulait aussi largement dans le compartiment sauvage (pathogène généraliste). A l'inverse, dans le cas de paysages agricoles plus hétérogènes et de pathogènes spécialistes, la connaissance du parcellaire agricole peut se révéler cruciale pour prédire la propagation de l'agent pathogène dans le paysage.

Certaines études se basent sur un parcellaire issu de bases de données officielles (Haight et al., 2011; Papaix et al., 2011). Par exemple, un modèle épidémiologique de la maladie du flétrissement américain du chêne causé par le champignon *Ceratocystis fagacearum* a été établi en s'appuyant sur des données officielles de densité en chêne et de présence de la maladie dans des forêts du Minnesota (Haight et al., 2011). L'établissement d'un modèle épidémiologique a permis aux auteurs d'étudier la propagation de la maladie selon différents scénarios de gestion et d'évaluer les pertes économiques occasionnées. D'autres études établissent des modèles épidémiologiques en se basant sur un parcellaire prédit grâce à des méthodes statistiques (Rimbaud et al., 2018). Des simulations de la propagation du virus de la *sharka* sur un parcellaire de vergers dynamique dans le temps, établi grâce à des prédictions des dates d'arrachage et de replantation des vergers connus, ont été réalisées (Rimbaud et al., 2018). Dans tous les cas, le choix du parcellaire sur lequel travailler n'est pas trivial et peut, dans certaines zones d'études moins administrées, se révéler compliqué. Des étapes de reconstitution du parcellaire peuvent alors être nécessaires avant même de pouvoir procéder à une étude de risque.

La production agrumicole mondiale est aujourd'hui menacée par l'émergence de la maladie du Huanglongbing (HLB) aussi appelée « greening » (Wang & Trivedi, 2013). Le HLB a été signalé pour la première fois en Chine dans les années 1930 et a rapidement envahi l'ensemble des bassins de production d'agrumes du globe, à l'exception du bassin méditerranéen (J. M. Bové, 2014). Des bactéries Gram – non cultivables appartenant au genre *Candidatus Liberibacter* sont responsables de la maladie. Il en existe plusieurs 'espèces' : *Ca. L. asiaticus*, *Ca. L. africanus* et *Ca. L. americanus*. La bactérie circule dans le phloème des agrumes et obstrue les vaisseaux conducteurs de sève (Sechler et al., 2009), entraînant la mort de l'arbre entre 2 et 6 ans après son infection (J. Bové, 2006). Dans l'Union Européenne, les espèces de *Ca. Liberibacter* ont été incluses dans la liste des 20 organismes de quarantaine prioritaires établie en 2019 (commission delegated regulation EU 2019/1702).

La détection visuelle de la maladie est rendue difficile par la longue phase cryptique de 6 mois à 2 ans qui suit l'infection et par la confusion entre des carences en nutriments et les

symptômes observés : jaunissement asymétrique des feuilles et des flushs (J. Bové, 2006), fruits asymétriques de petite taille et à coloration inversée (Gottwald, 1989) avec des graines avortées (J. Bové, 2006). La symptomatologie n'est donc pas suffisante pour établir un diagnostic, des analyses en laboratoire par extraction d'ADN et PCR quantitative (Li et al., 2006) sont nécessaires, la bactérie n'étant pas cultivable *in vitro*. La bactérie est vectée par des psylles, dont les espèces *Diaphorina citri* et *Trioza erytreae*. La bactérie présente une spécificité d'hôtes pour les plantes de la famille des Rutaceae, principalement *Citrus* avec une présence sur plusieurs espèces d'autres genres de rutacées, dont le genre *Murraya* fortement attractif pour *D. citri* (Halbert & Manjunath, 2004). Cependant, les taux de transmission de la bactérie par les psylles de *M. paniculata* vers des agrumes sont très faibles (Cifuentes-Arenas et al., 2019). Ces autres Rutacées peuvent servir de réservoir aux vecteurs mais d'un point de vue épidémiologique, le réservoir de la bactérie se limite donc aux agrumes.

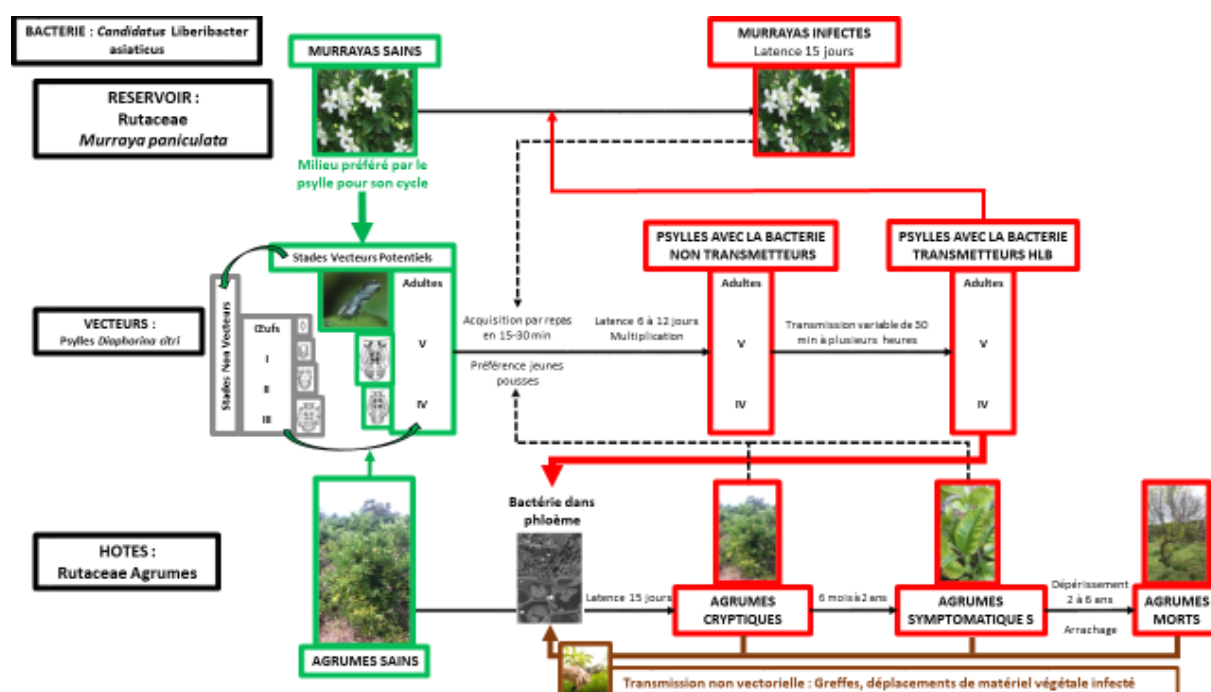


Figure 1 Modalités de dispersion de la bactérie *Ca. Liberibacter asiaticus*. En vert, les compartiments sains du pathosystème et en rouge ceux infectés. D'après Houillon & Pruvost (Comm. Pers.).

A ce jour, aucun traitement utilisable à grande échelle n'est disponible. La gestion de la maladie se limite à l'arrachage systématique des arbres infectés, au contrôle des vecteurs par des méthodes de lutte chimique ou biologique, à la surveillance des pépinières et à l'utilisation de combinaisons variétés et porte-greffes tolérants (Aubert et al., 1996) bien que de nouvelles pistes soient prometteuses (Munir et al., 2018). L'utilisation irraisonnable d'antibiotiques par le passé a vite montré son inefficacité à contrôler plusieurs maladies des plantes et les inquiétudes quant à son impact sur la santé humaine sont aujourd'hui grandissantes (« Spraying Diseased Citrus Orchards with Antibiotics Could Backfire », 2019). Néanmoins, un

programme massif d'application annuelle de plusieurs centaines de tonnes de streptomycine pour lutter contre le HLB aux USA a été approuvé par les autorités américaines (Taylor & Reeder, 2020).

A la Réunion, département ultramarin français de l'Océan Indien, le HLB a causé une épidémie sévère dans les années 1970-1980 avant d'être contrôlée grâce à un programme d'arrachage massif des arbres malades, d'injection d'antibiotique et à une lutte biologique contre les vecteurs grâce aux insectes parasitoïdes *Tamarixia dryi* et *Tamarixia radiata* (Aubert et al., 1996; J. Bové, 2006). Cependant, on assiste depuis 2015 à une réémergence du HLB, menaçant la filière agrumicole de l'île. Le HLB étant classé comme un organisme nuisible réglementé par arrêté préfectoral, les travaux de recherche effectués suscitent un intérêt direct pour la filière agrumicole.

Avec près de 200 hectares de vergers, la filière agrumicole réunionnaise est de taille restreinte et la culture des agrumes reste assez traditionnelle avec souvent des parcelles de petite surface, plusieurs cultures associées sur le même terrain et une grande diversité de cultivars (25 variétés d'agrumes sont cultivées sur l'île). Ce paysage agrumicole très morcelé s'explique en partie par la forte variabilité des conditions pédoclimatiques sur l'île. En effet, la Réunion possède un fort gradient altitudinal (les parcelles s'échelonnent de 0 à 1200 mètres) et de fortes différences d'hygrométrie entre la côte au vent à l'Est et la côte sous le vent à l'Ouest, ce qui contribue à créer des centaines de micro-climats.

Les défis posés par ce système sont ceux que l'on retrouve dans les agricultures non industrielles, fréquentes dans de nombreux pays du sud : (i) le paysage est fragmenté et hétérogène donc potentiellement très structurant pour la propagation épidémique, (ii) pour diverses raisons, la connaissance du paysage agricole est très imparfaite, (iii) à la Réunion comme dans de nombreux territoires, la surveillance épidémiologique se heurte à la question des moyens qui contraignent fortement la quantité de données disponibles quant à la dynamique épidémique. Comment améliorer, dans ce contexte contraint la surveillance épidémiologique ? Au cours de ce stage, j'ai exploré trois pistes complémentaires. Tout d'abord j'ai contribué à améliorer la connaissance du parcellaire agrumicole de la Réunion en complétant les bases de données officielles par une combinaison de photo-interprétation d'images aériennes et de missions de terrain et en prédisant la position probable de parcelles d'agrumes grâce à des algorithmes de « Machine Learning ». Deuxièmement, j'ai caractérisé l'influence de la structure du paysage (parcellaire et variables pédoclimatiques) sur la présence de la maladie afin d'établir une cartographie du risque permettant de focaliser la surveillance sur les zones à plus haut risque. Enfin, j'ai souhaité travailler sur l'augmentation de la production de données de suivi de la propagation de maladie en explorant le potentiel d'une nouvelle méthode de détection par Spectroscopie Proche Infra-Rouge (SPIR), présentant les avantages d'être applicable rapidement, à moindre coût et à grande échelle.

2. Matériels et méthodes

2.1. Prédiction du parcellaire agrumicole réunionnais

2.1.1. Connaissance du parcellaire agrumicole réunionnais

Au début de ce stage, je disposais d'une base de données issue de la fusion de bases de données officielles et non officielles réalisée par mon équipe d'accueil, comportant la position de 2920 vergers sur l'île parmi lesquels 640 étaient classés comme « agrumes » et 827 comme « non agrumes ». Le statut des vergers restants était indéterminé. Tout au long de mon stage, dans la même démarche que les travaux précédents, j'ai régulièrement identifié des parcelles candidates par photo-interprétation et organisé des campagnes de terrain afin de vérifier leur statut, collecter des informations complémentaires telles que les variétés cultivées, l'accessibilité du verger, ainsi que d'autres caractéristiques utilisées avec les SIGs et effectuer des prélèvements dans le but d'isoler de l'ADN bactérien de *Candidatus Liberibacter asiaticus* afin d'étudier le lien entre la diversité génétique de la bactérie et sa circulation dans le paysage. Par ailleurs j'ai analysé des photos satellites sur Google Earth pour déterminer les dates de plantation et dans certains cas d'arrachage des vergers d'agrumes.

2.1.2. Constitution de la base de données

Les variables pédoclimatiques (réserve utile, exposition, type de sol et couleur, pluviométrie, altitude, pente, ...) ont été calculées sur un logiciel de SIG (QGIS.org, 2020) à partir de couches d'information géographiques disponibles sur la plateforme AWARE (Cirad AWARE, 2020). Afin de prendre en compte la composante historique et sociale dans la décision d'installer une parcelle d'agrumes, les densités environnantes en agrumes ont été également calculées pour différents rayons sous R (R Development Core Team, 2005). Celles-ci correspondent au rapport entre la somme des surfaces cultivées en agrumes dans un certain rayon r autour d'une parcelle et la surface du disque de rayon r . Une classification de l'occupation du sol réalisée par télédétection à partir d'images satellites (S. Dupuy, 2019) a également été utilisée pour calculer des densités environnantes pour chaque catégorie d'entité du paysage (forêts, vergers de litchi ou longani, prairies fauchées par exemple) autour de chaque parcelle et pour différents rayons.

2.1.3. Prédiction de la nature des parcelles

Afin de prédire la nature des parcelles dont le statut reste à ce jour indéterminé et de cartographier le parcellaire « prédit » sur QGIS, j'ai appliqué des méthodes de classification supervisées au jeu de données dont je disposais.

Les méthodes utilisées sont des Support Vector Machine (SVM) et Random Forest.

Les SVM sont des algorithmes de classification binaire par apprentissage supervisé. Ils présentent l'avantage de limiter les risques de sur-apprentissage en estimant leur complexité au cours de la phase d'apprentissage. Un SVM recherche un classificateur ou hyperplan qui sépare les données en maximisant la distance entre les deux classes de points. On cherche ici à prédire la nature des parcelles en fonction d'un ensemble de variables pédoclimatiques. Une transformation non linéaire des données par différents *kernels* peut être utilisée. Les modèles

de Random Forest effectuent un apprentissage sur de multiples arbres décisionnels construits aléatoirement et entraînés sur des sous-ensembles du jeu de données différents.

Le jeu de données utilisé comporte les parcelles dont le statut est déterminé, soit 827 parcelles classées en « non-agrume » et 640 classées en « agrume », caractérisées par des variables descriptives (variables pédoclimatiques et occupation du sol dans des rayons autour de chaque parcelle).

Dans un premier temps, 75% des données servent à l'apprentissage du modèle, c'est-à-dire à l'estimation des paramètres pour chaque variable prédictive, et les 25% restants servent à la validation du modèle. Par la suite, on réalise une validation croisée à 10 répétitions. Le jeu de données est alors séparé selon un ratio 90/10 à dix reprises, l'ensemble des données sert donc à l'apprentissage et à la validation du modèle.

2.2. Cartographier le risque de présence du HLB

2.2.1. Plan de surveillance et échantillonnage

Depuis 2015, la DAAF (Direction de l'Alimentation, de l'Agriculture et de la Forêt) organise une surveillance du parcellaire agrumicole. La FDGDON (Fédération Départementale des Groupements de Défense contre les Organismes Nuisibles) réalise les prélèvements sur le terrain. L'ANSES (Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du travail) au travers de son unité LSV-RAPT et le CIRAD (Centre de Coopération internationale en Recherche Agronomique pour le Développement) apporte leur expertise au sein du comité de pilotage. Le plan d'échantillonnage a considérablement évolué au cours du temps, aujourd'hui sous l'impulsion des travaux du CIRAD et de l'ANSES, 60 arbres par hectare sont prélevés et les analyses moléculaires sont réalisées en poolant les arbres par 5 ou 10 arbres. Ce protocole a permis de suivre environ 100 parcelles en 2018 (cf figure 2) et a été poursuivi les années suivantes.

Année	2015	2016	2017	2018-2020
Nombre de parcelles suivies	12	30	30	100
Nombre de prélèvements par parcelle	0 à 4	2 à 5	5	60 arbres/ ha en poolant par lot de 5 à 10 arbres

Figure 2 Modalités d'échantillonnage des parcelles suivies dans le cadre de l'épidémiosurveillance du HLB à la Réunion depuis 2015.

Aux prélèvements officiels de la FDGDON s'ajoutent 684 prélèvements réalisés par le CIRAD lors de sorties terrain. Dans ce cas, seuls les arbres très symptomatiques sont prélevés, l'objectif étant de maximiser le nombre d'arbres détectés positifs afin d'étudier la diversité

généétique de la bactérie par typage de 12 marqueurs microsatellites. Cela permet d'étudier les liens entre différentes souches bactériennes et paramètres environnementaux et de détecter d'éventuelles nouvelles introductions.

2.2.2. Analyse statistique : probabilité de présence de la maladie

Les deux jeux de données de prélèvements ont été combinés et une fusion par localisation avec la couche des parcelles d'agrumes a été réalisée sous QGIS. On dispose alors de 226 parcelles d'agrumes pour lesquelles on a le résultat d'un test HLB (171 parcelles positives et 55 négatives). Des variables pédoclimatiques et de densités environnantes ainsi qu'un indice de connectivité (IC) ont été calculées sur QGIS et R de la même manière que pour la reconstruction du parcellaire. On calibre le modèle Random Forest sur ces données et on en évalue la prédiction.

Ensuite, on utilise ce modèle pour prédire le statut HLB des parcelles qui ont été prédites en « agrume » précédemment.

2.3. Prédiction du statut infecté des arbres par Spectroscopie Proche Infra-Rouge

2.3.1. Principe de la Spectroscopie Proche Infra-Rouge

La SPIR est une technique d'analyse basée sur le principe d'absorption des rayonnements infra-rouge par la matière organique (Bertaux, 2014). Lorsqu'un rayonnement rencontre la matière, un transfert d'énergie peut avoir lieu à condition que l'énergie qu'il transporte (dépendant de sa longueur d'onde) soit égale à la différence d'énergie potentielle entre deux niveaux d'énergie d'excitation d'un atome ou d'une molécule. La quantité d'énergie absorbée par l'échantillon dépend donc de sa composition chimique. La mesure de l'intensité énergétique transmise I après traversée de l'échantillon permet de calculer une valeur d'absorbance à chaque longueur d'onde (l'absorbance correspond au logarithme décimal du rapport entre l'intensité énergétique I_0 à une longueur d'onde donnée, avant traversée du milieu, et l'intensité énergétique transmise I).

2.3.2. Echantillonnage

L'échantillonnage a eu lieu les 3 et 4 juin 2020 entre 8h et 10h sur la parcelle de Monsieur Simon Hoarau située à Petite-Ile (cf figure 3). Notre choix s'est orienté vers cette parcelle car sa superficie de 2,35 ha est suffisamment grande pour garantir de trouver des arbres de la même variété à la fois sains et atteints du HLB. En effet, une prospection réalisée par la FDGDON avait révélé la présence de 8 lots positifs sur les 15 lots testés en mars 2019. De plus, la conduite de culture est de bonne qualité ce qui permet d'éviter au mieux des facteurs confondants.

Quatorze arbres d'une même variété d'oranger, *Citrus sinensis* var. Washington Navel, ont été échantillonnés. Les arbres ont été choisis selon leurs caractères phénotypiques, sept étaient supposés sains et sept autres étaient supposés atteints du HLB. Sur chaque arbre, trois lots de 10 feuilles ont été prélevés de manière aléatoire et ont été conservés à une température comprise entre 8 et 14 °C jusqu'à ce que les mesures de SPIR aient été effectuées (cf figure 4).

Localisation des arbres échantillonnés



Figure 3 Localisation des arbres échantillonnés sur la parcelle de Simon Hoarau à Petite-Ile

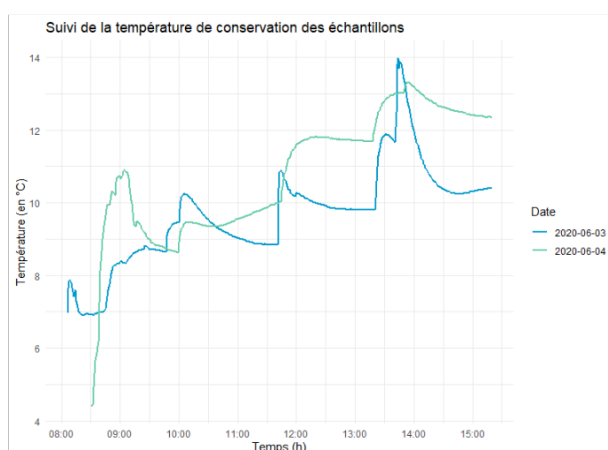


Figure 4 Suivi de la température de conservation des échantillons lors de l'expérimentation

2.3.3. Prise de données

Les feuilles ont été nettoyées à l'eau avec un papier absorbant afin d'enlever les dépôts de sulfate de cuivre et de fumagine présents sur le limbe. Un spectromètre ASD Labspec 5000® et sa sonde ASD Contact Probe® ont été utilisés pour recueillir sur chaque feuille 6 spectres proche-infrarouges compris entre 350 et 2500 nm avec un pas de 1 nm.

Lors d'une expérience préliminaire portant sur quatre arbres de l'espèce, *Citrus sinensis* var. Washington Navel, on a déterminé le nombre minimal de spectres qu'il fallait collecter pour maximiser la justesse de la prédiction du statut des arbres. Trois lots de 10 feuilles par arbre ont été échantillonnés, 10 mesures de SPIR ont été réalisées sur chacune des feuilles. Une prédiction du statut des arbres a été effectuée en utilisant d'une à dix mesures par feuille. Dans chacun des cas, différents paramètres ont été calculés à partir de la matrice de confusion du modèle (cf figure 5). La prédiction a été évaluée grâce au calcul de la justesse de la

prédiction (accuracy), de la sensibilité (sensi) et de la précision (pospred). A partir de six mesures de SPIR par feuille, on observe un plateau dans la justesse de la prédiction (cf accuracy figure 6) et dans la détection des positifs (cf sensibility figure 6).

	Actual Positive	Actual Negative
Predicted Positive	True Positive(TP)	False Positive(FP) (Type 1 Error)
Predicted Negative	False Negative(FN) (Type 2 Error)	True Negative(TN)

Accuracy = $\frac{\text{True Positive} + \text{True Negative}}{\text{Total Population}}$	Error Rate/Misclassification rate = $\frac{\text{False Positive} + \text{False Negative}}{\text{Total Population}}$
Precision = $\frac{\text{True Positive}}{\text{Predicted Positive(TP+FP)}}$	Sensitivity/Recall = $\frac{\text{True Positive}}{\text{Actual Positive(TP+FN)}}$
Specificity = $\frac{\text{True Negative}}{\text{Actual Negative(FP+TN)}}$	F1 Score = $\frac{2 * (\text{Recall} * \text{Precision})}{\text{Recall} + \text{Precision}}$

Confusion Matrix or Error Matrix along with different metrics

Figure 5 Calcul de différents paramètres à partir de la matrice de confusion du modèle pour en évaluer la qualité de la prédiction

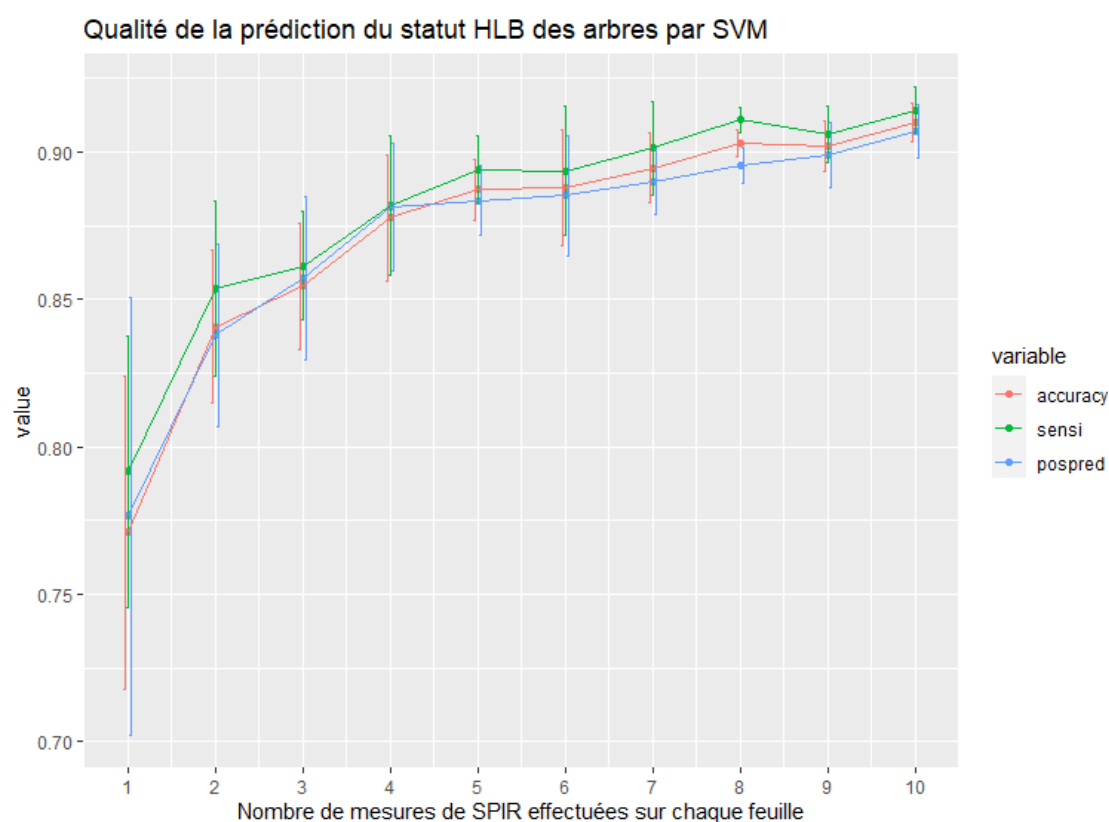


Figure 6 Evaluation de la qualité de prédiction du statut HLB des arbres par SVM en fonction du nombre de mesures de SPIR effectuées sur chaque feuille

Afin de réaliser des mesures de SPIR en réflectance, les feuilles ont été disposées sur un disque en Spectralon® d'absorbance quasi-nulle, la quasi-totalité de la lumière transmise à travers la feuille est donc réfléchi et retransverse la feuille avant d'atteindre le capteur. Une grande attention a été portée à la propreté du support pour conserver sa valeur de réflectance constante au cours du temps. Chaque spectre correspond à la réflectance moyenne de 50 mesures de réflectance entre 350 et 2500 nm. Les mesures de réflectance ont été réalisées à température ambiante (22°C). Une calibration de l'appareil a été effectuée toutes les 15 minutes en réalisant une mesure de SPIR sur le support en Spectralon® seul.

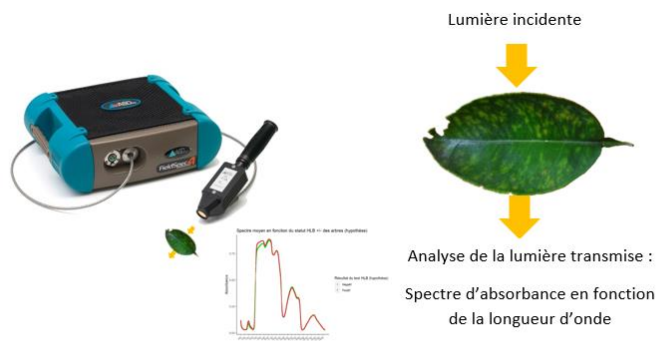


Figure 7 Principe de fonctionnement simplifié d'un spectromètre proche infra-rouge

Des variables phénotypiques ont également été relevées sur chaque feuille, telles que le statut en flush (cf figure 8), la sévérité liée au chancre citrique (cf figure 9) et la présence de chloroses (cf figure 10). Des échelles de notation ont été construites pour ces deux dernières variables.



Figure 8 Illustration des modalités de la variable feuille en flush

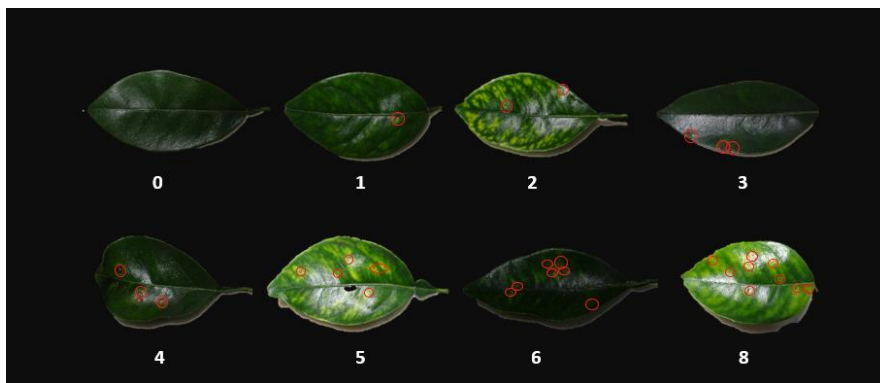


Figure 9 Echelle de notation des symptômes associés au chancre citrique



Figure 10 Echelle de notation associée aux chloroses

2.3.4. Analyses moléculaires et sérologiques

2.3.4.1. Détermination du statut HLB des lots par PCR quantitative

Les nervures et pétioles ont été découpés puis broyés par lot de 10 au broyeur à billes Bioreba® Homex (Lenze) dans des sachets Bioreba® pour chacun des lots. 3,78 mL de tampon de broyage Tris/EDTA/SDS (méthode officielle ANSES MOA0033) ont été utilisés par sachet, ce volume a été déterminé en pesant la biomasse par sachet et en respectant le rapport 1 mL de tampon de broyage pour 0,2 g de matière végétale. L'ADN a ensuite été extrait des broyats grâce au kit DNeasy Plant Kit (Qiagen, Valencia, CA).

La présence de la bactérie *Ca. Liberibacter asiaticus* a été déterminée par qPCR selon la méthode utilisée par Li et al. (2006). Pour cela, deux amorces, HLBr (spécifique du genre *Liberibacter*, s'hybridant sur le brin *reverse*) et HLBas (spécifique de *Ca. Liberibacter asiaticus*, s'hybridant sur le brin *forward*) ont été utilisées. Une sonde HLBp est spécifique du genre *Liberibacter* permet de visualiser le signal dans le cadre de cette qPCR selon la méthode *Taqman*.

Les paramètres suivants ont été utilisés pour la PCR en temps réel : préchauffage à 95°C pendant 10 min, puis 45 cycles 95°C pendant 15 secondes suivi de 1 min à 58°C, excitation par une longueur d'onde à 495 nm et mesure de l'émission d'une longueur d'onde à 515 nm.

Les échantillons sains et malades ont été distingués selon leur valeur de C_T (Cycle Treshold). Le C_T correspond au nombre de cycles PCR à partir duquel le produit de la PCR est détectable,

plus le C_T est faible plus il y a d'ADN présent initialement et inversement. Deux valeurs seuils de C_T , 32 et 36, ont été utilisées pour distinguer les échantillons sains et malades. Les échantillons présentant un C_T inférieur ou supérieur à cette valeur seuil ont été respectivement considérés comme positifs ou négatifs au HLB.

2.3.4.2. Diagnostic du virus de la tristeza (CTV)

On utilise 0,15g de tissu broyé dans 3ml de tampon commercial (composition propriétaire) pour réaliser un diagnostic CTV grâce à des CTV ImmunoStrip® développés par Agdia®. L'ensemble des échantillons s'est révélé positif au CTV.

2.3.5. Classification par *machine learning* sous R

Le nombre de lots de feuilles positives varie en fonction du seuil de C_T que l'on se fixe. Avec un $C_T = 32$, on a 8 lots positifs sur les 42 lots échantillonnés. Avec un $C_T = 36$, on a 16 lots positifs. On effectue une analyse soit par lot, en considérant les lots d'un même arbre comme indépendants, soit par arbre, en considérant tous les lots d'un même arbre comme positifs si au moins un des trois lots est positif.

Une classification non supervisée par *Support Vector Machine* (SVM) a été effectuée sur le logiciel R grâce au package *e101*. La variable à expliquer est le statut HLB des arbres et les variables explicatives sont les valeurs de réflectance des feuilles pour chaque longueur d'onde comprise entre 350 et 2500 nm et les variables phénotypiques (flush, chlorose, chancre citrique). Il y a donc 2154 variables explicatives.

Le jeu de données a été divisé en deux jeux de données servant respectivement à l'apprentissage et à la validation du modèle selon un ratio 75/25. Une validation croisée (K-fold Cross Validation) a été réalisée avec 10 répétitions afin de limiter le risque de surapprentissage du modèle. Les performances moyennes du modèle ont été évaluées pour différentes valeurs de C_T seuil et en conservant l'ensemble des modalités des variables phénotypiques ou en les transformant en variables binaires.

L'importance de chaque variable dans la classification par SVM a été déterminée grâce à la fonction *Importance* du package *Rminer* (Cortez & Embrechts, 2013).

3. Résultats

3.1. Prédiction du parcellaire réunionnais

La prédiction de la nature des parcelles par un modèle SVM (cf figure 11) présente une accuracy de 65%, c'est-à-dire que 65% des prédictions sont justes. Cependant, seulement 3% des parcelles d'agrumes sont effectivement classées en « agrumes ». Parmi les parcelles classées en « agrumes », 86% sont bien des agrumes mais cela ne compense pas la très faible sensibilité du modèle. A côté de ces performances moyennes, le modèle prédit très bien les parcelles « non-agrume » (100% de spécificité).

Evaluation de la prediction de la nature des parcelles			
Modele SVM			
accuracy	sensitivity	specificity	precision
0.65	0.03	1	0.86

Figure 11 Evaluation de la prédiction de la nature des parcelles par Support Vector Machine

L'emploi d'une validation croisée à 10 répétitions améliore légèrement la précision (accuracy) mais la prédiction des parcelles d'agrumes n'est pas satisfaisante.

La prédiction de la nature des parcelles par un modèle Random Forest est bien meilleure (cf figure 12), 82% des parcelles sont bien classées. La prédiction des parcelles « agrumes » est bien meilleure qu'avec un SVM, ici 75 % des parcelles d'agrumes sont bien classées (sensitivity) et parmi les parcelles prédites en agrumes, 75% sont réellement des agrumes. La prédiction des « non-agrumes » reste la meilleure car c'est la classe majoritaire, 86% des « non-agrumes » sont prédites comme telles. Cela nous permet de construire un parcellaire agrumicole prédit sous Qgis (cf figure 14).

Evaluation de la prediction de la nature des parcelles			
Modele Random Forest			
accuracy	sensitivity	specificity	precision
0.82	0.75	0.86	0.75

Figure 12 Evaluation de la prédiction de la nature des parcelles par Random Forest

Le poids des variables explicatives dans la classification de la nature des parcelles n'est pas le même entre les deux méthodes(cf figure 13). Parmi les variables les plus importantes, on retrouve les densités de vergers d'agrumes dans un rayon de 100 m (densite_100_Verger.agrume), 300 m et 900 m, la nature des sols (STIPA), l'altitude (alti) et la température maximale journalière (T_max_mean). L'exposition a également un rôle prépondérant. L'indice de connectivité (IC) entre les parcelles d'agrumes est également important.

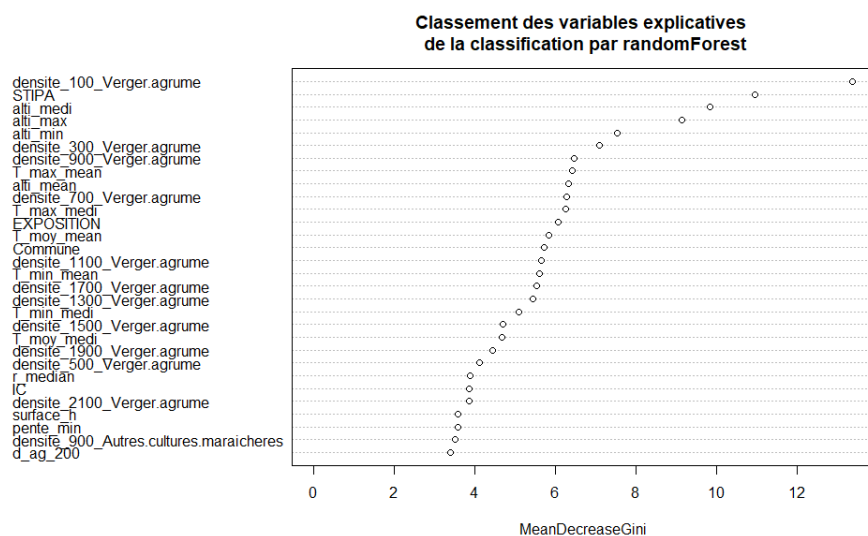


Figure 13 Classement de l'importance des variables explicatives dans la prédiction de la nature des parcelles par Random Forest

L'emploi d'une validation croisée à 10 répétitions amène aux mêmes résultats.

La prédiction de la nature des parcelles par Random Forest nous permet d'améliorer la connaissance du parcellaire agrumicole réunionnais. Le parcellaire agrumicole « prédit » pourra être utilisé par la suite pour faire tourner des modèles épidémiologiques du HLB.

Cartographie du parcellaire agrumicole réunionnais

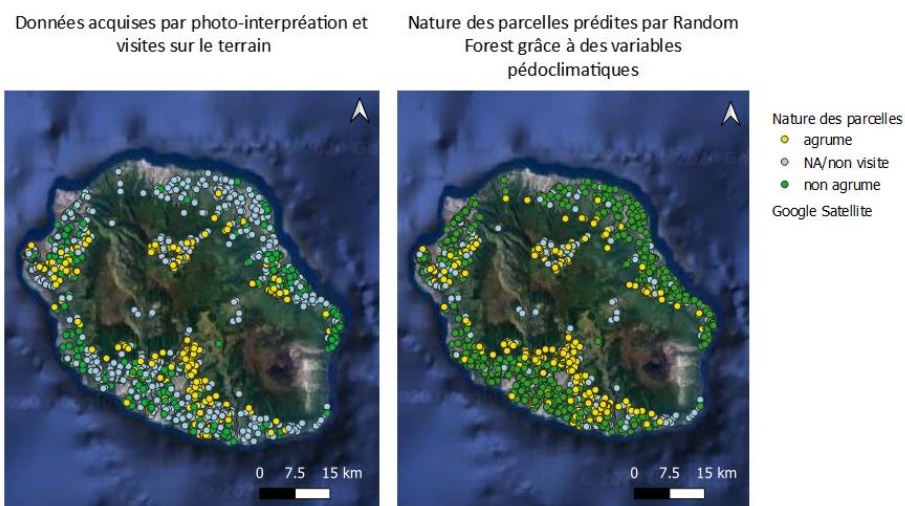


Figure 14 Cartographie du parcellaire agrumicole réunionnais

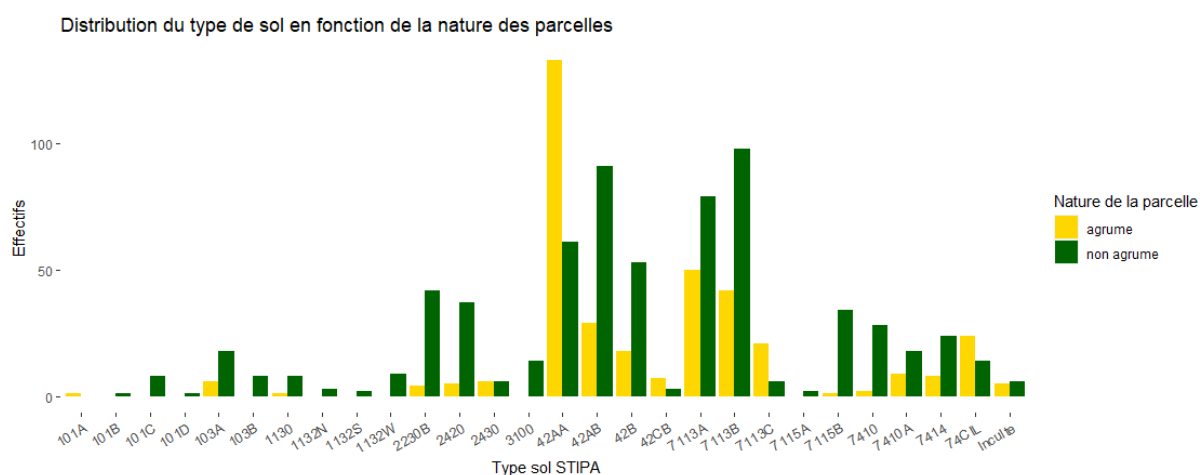


Figure 15 Distribution du type de sol en fonction de la nature des parcelles

On a vu précédemment que le type de sol (variable STIPA) avait beaucoup d'importance dans la classification de la nature des parcelles. La distribution des types de sol est différente en fonction de la nature des parcelles (cf figure 15). Les sols au code STIPA 42AA sont largement majoritaires chez les parcelles d'agrumes. Ils correspondent à des sols andiques non perhydratés (Pouzet, 2002). Les sols bruns andiques, codes STPA 7113A, 7113B (Pouzet, 2002), sont plutôt majoritaires chez les parcelles « non agrume ».

3.2. Cartographier le risque de présence du HLB

On prédit maintenant l'état sanitaire des parcelles (statut HLB) grâce à des variables pédoclimatiques et de caractérisation de la structure du paysage en utilisant un modèle Random Forest avec une validation croisée à 10 folds (cf figure 15). 82% des prédictions sont justes (cf accuracy). Le modèle peine à prédire les parcelles qui sont positives : seulement 51% des parcelles réellement positives sont prédites comme telles (cf sensitivity) et parmi les parcelles prédites positives, 68% le sont effectivement (precision). Par contre, la prédiction est bien meilleure pour les parcelles négatives : 92% de parcelles négatives sont prédites comme telles (cf specificity).

Evaluation de la prediction de l'état sanitaire des parcelles

Modele Random Forest, CV=10

accuracy	sensitivity	specificity	precision
0.82	0.51	0.92	0.68

Figure 16 Evaluation de la prédiction de l'état sanitaire des parcelles par Random Forest

Parmi les variables les plus explicatives (cf figure 17), on retrouve des variables liées aux bassins de production d'agrumes (Commune, type de sol STIPA) et des variables bioclimatiques (altitude, température minimale journalière) qui peuvent être corrélées à la localisation des parcelles d'agrumes ou bien à la niche écologique des psylles, vecteurs de la maladie du HLB ou les préférences thermiques de la bactérie. Les densités en vergers

d'agrumes dans un rayon variant de 1,1 à 2,1 km ont également un rôle majeur dans la classification.

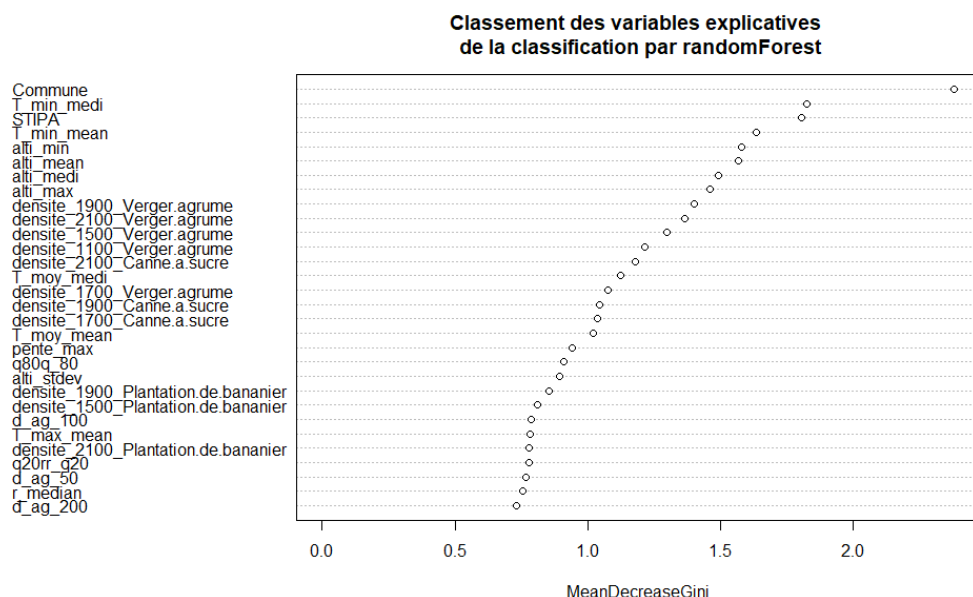


Figure 17 Classement de l'importance des variables explicatives dans la prédiction du statut HLB des parcelles

La prédiction de l'état sanitaire des parcelles nous permet de cartographier le risque de présence de la maladie (cf figure 18). On retrouve deux foyers de maladie, à Saint-Benoit au Nord-Est et à Petite-Île au Sud. Il n'y a pas parcelle prédite comme positive ailleurs mais le nombre de parcelles dont le statut n'est pas prédit à cause d'un manque de données reste considérable.

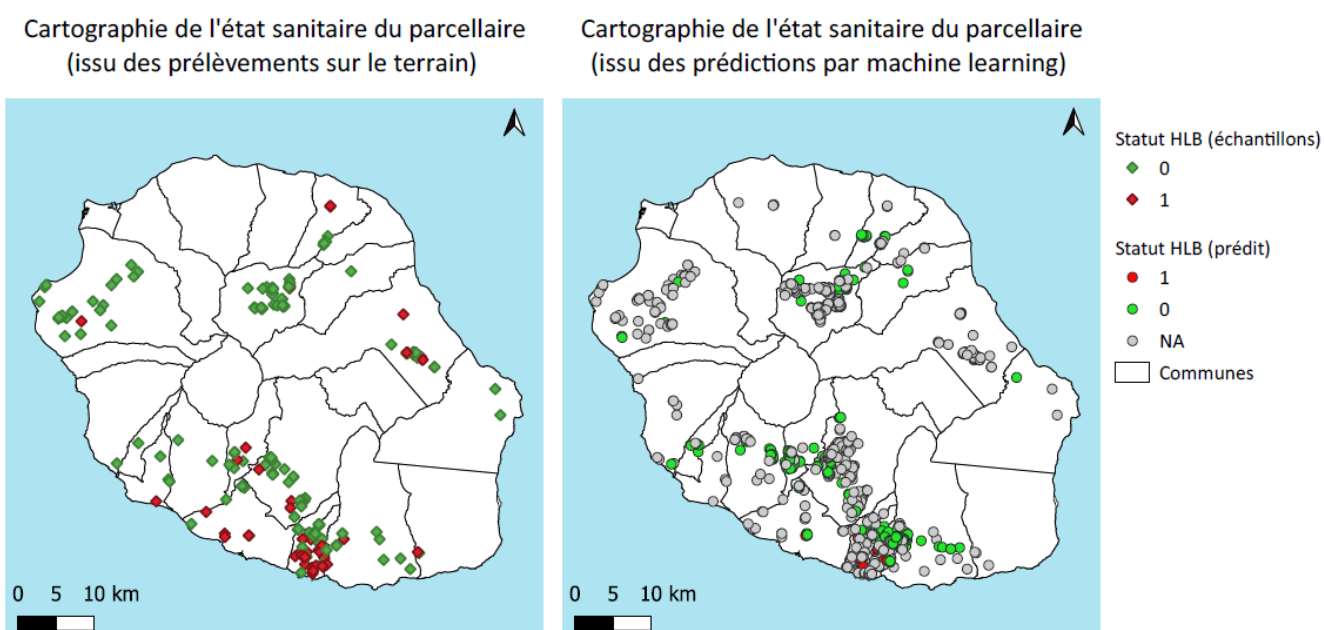


Figure 18 Cartographie de l'état sanitaire du parcellaire agrumicole réunionnais

3.3. Signature spectrale de la maladie du HLB

3.3.1. Spectres proche-infrarouge

On dispose de 2520 spectres proche infra-rouge (PIR) au total. Pour chaque longueur d'onde, on calcule l'inverse de la réflectance moyenne en fonction du statut HLB de l'arbre (cf figure 19).

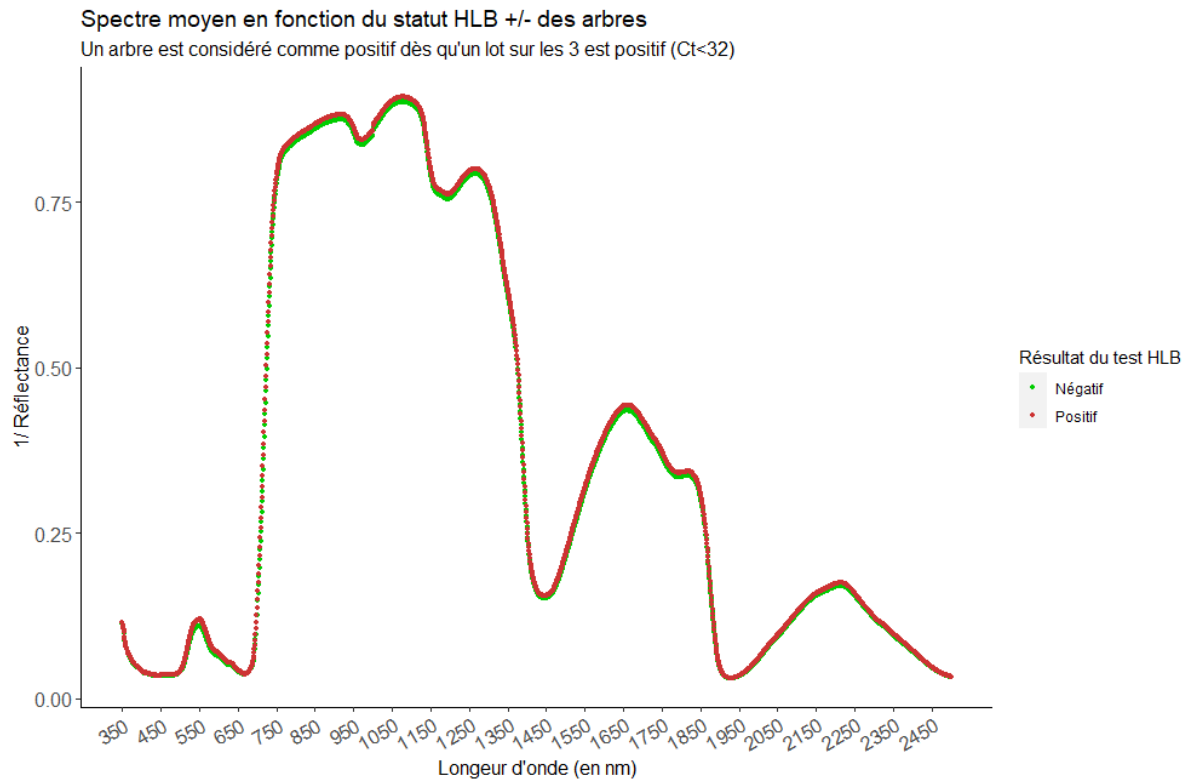


Figure 19 Spectre IR moyen en fonction du statut HLB des arbres

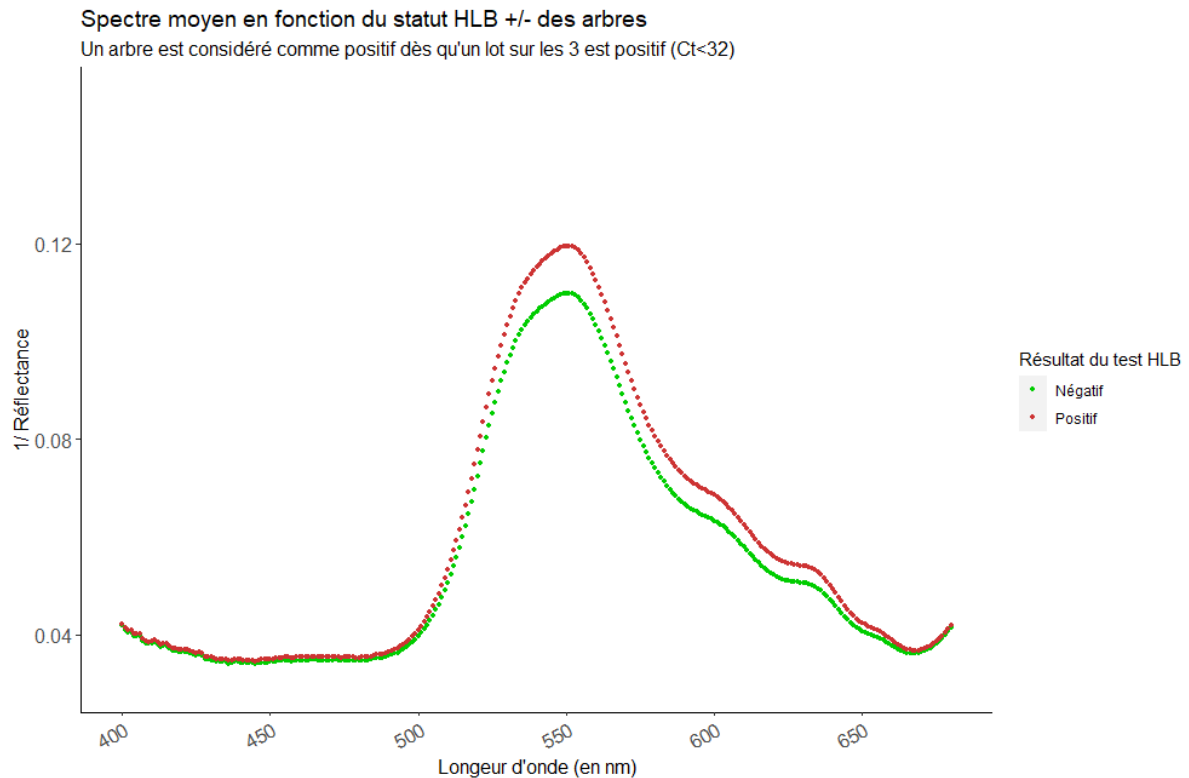


Figure 20 Spectre IR moyen en fonction du statut HLB des arbres (zoom sur zone 400-600 nm)

3.3.2. Prédiction du statut infecté des arbres par Spectroscopie Proche Infra-Rouge (SPIR)

On évalue la prédiction des modèles en fonction de l'AUC (Area Under the Curve). L'AUC correspond à l'aire sous la courbe ROC (Sensibilité en fonction de 1-Spécificité, cf figure 20). En maximisant l'AUC, on maximise le taux de vrais positifs et on minimise le taux de faux positifs.

Ici, l'AUC est maximale pour un $C_T = 32$ et en ne considérant pas les lots comme indépendants. L'accuracy n'est pas maximale (75.9%) mais on détecte à la fois bien les positifs (sensitivity = 76,5%) et les négatifs (specificity=74,7%).

En prenant un $C_T = 36$ pour les mêmes conditions, la précision chute à 69,4% car soit le nombre de faux positifs en qPCR augmente, soit le nombre de faux négatifs en SPIR augmente auquel cas la SPIR peine à détecter des échantillons faiblement positifs. Le C_T seuil de 32 est donc à privilégier.

Si l'on veut détecter au mieux les positifs, il faut maximiser la sensitivity. En prenant un $C_T=32$, des lots indépendants et un poids de 1 pour les variables phénotypiques, entre 98% et 99% des positifs sont prédits comme tels (cf deux dernières lignes figure 22). Cependant, la prédiction des arbres sains est mauvaise car environ 20% seulement des arbres sains sont bien classés (cf sensitivity).

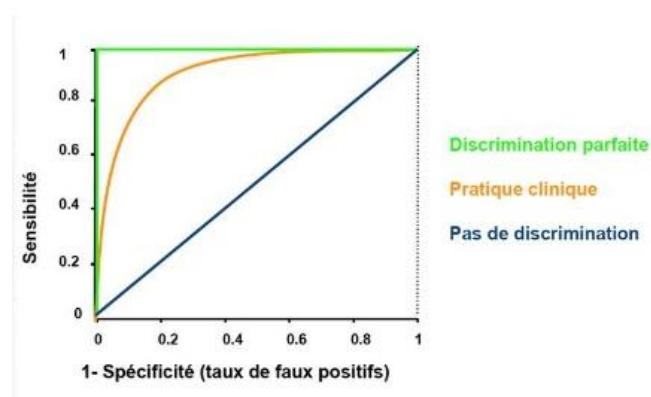


Figure 21 Interprétation d'une courbe ROC

Evaluation de la prediction du statut HLB des spectres proches infra-rouge											
Modele SVM, kernel linéaire											
Accuracy	Precision	Specificity	Error_rate	Sensitivity	Kappa	AUC	CTseuil	lots_indépendants	variables_pheno	poids_var_pheno	
0.759	0.845	0.747	0.241	0.765	0.517	0.756	32	non (arbre)	catégories	inverse	
0.754	0.694	0.746	0.246	0.765	0.509	0.756	36	non (arbre)	catégories	inverse	
0.755	0.698	0.755	0.245	0.756	0.510	0.755	36	non (arbre)	binaires	inverse	
0.758	0.841	0.739	0.242	0.769	0.516	0.754	32	non (arbre)	binaires	inverse	
0.752	0.722	0.801	0.248	0.688	0.505	0.744	36	non (arbre)	catégories	1	
0.766	0.919	0.708	0.234	0.780	0.532	0.744	32	oui	binaires	inverse	
0.740	0.703	0.783	0.260	0.683	0.480	0.733	36	non (arbre)	binaires	1	
0.762	0.913	0.683	0.238	0.781	0.525	0.732	32	oui	catégories	inverse	
0.768	0.781	0.551	0.232	0.889	0.537	0.720	32	non (arbre)	binaires	1	
0.720	0.804	0.714	0.280	0.724	0.440	0.719	36	oui	catégories	inverse	
0.765	0.778	0.544	0.235	0.887	0.529	0.715	32	non (arbre)	catégories	1	
0.709	0.801	0.715	0.291	0.705	0.418	0.710	36	oui	binaires	inverse	
0.719	0.746	0.543	0.281	0.826	0.437	0.685	36	oui	binaires	1	
0.720	0.738	0.510	0.280	0.850	0.440	0.680	36	oui	catégories	1	
0.850	0.851	0.267	0.150	0.988	0.701	0.627	32	oui	catégories	1	
0.841	0.841	0.203	0.159	0.992	0.683	0.597	32	oui	binaires	1	

Figure 22 Evaluation de la prédiction du statut HLB des spectres proche infra-rouge par SVM

3.3.3. Identification des longueurs d'onde caractéristiques de la maladie

Grâce à la fonction *Importance* du package *Rminer* (Cortez & Embrechts, 2013), on calcule l'importance de chaque variable dans la classification du statut HLB des spectres par SVM. Cela nous permet d'identifier les longueurs d'onde qui sont les plus caractéristiques de la signature

spectrale du HLB (cf figure 23). Les longueurs d'onde situées entre 350 et 650 nm et entre 1900 et 2000 nm sont les plus importantes. Celles situées entre 750 et 1300 nm sont d'importance moyenne.

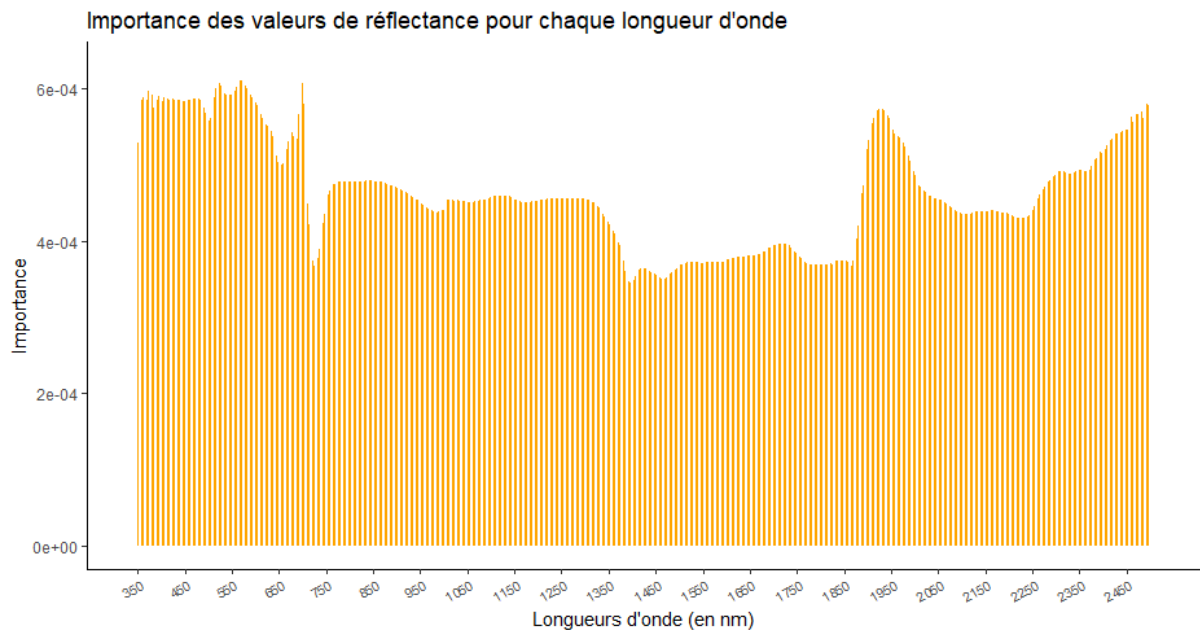


Figure 23 Importance de chaque variable (longueur d'onde) dans la prédiction du statut HLB par SVM

Les variables phénotypiques *chancre*, *chlorose* et *flush* ont presque deux fois moins d'importance que n'importe quelle longueur d'onde du spectre PIR (cf figure 24).

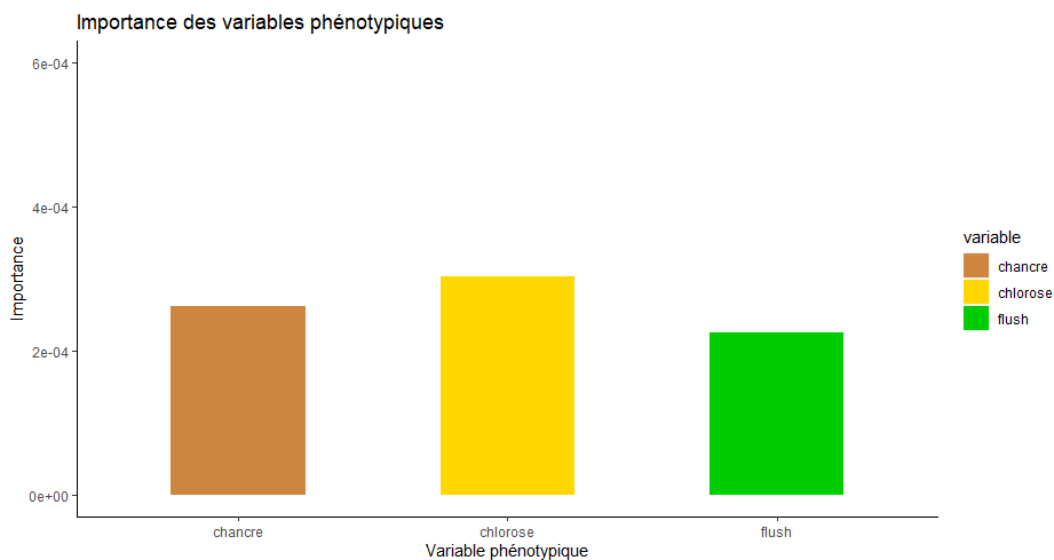


Figure 24 Importance des variables phénotypiques dans la prédiction du statut HLB par SVM

La prédiction du statut HLB des feuilles seulement grâce aux variables phénotypiques conduit à une moins bonne prédiction que lorsque l'on combine données de réflectance et variables

phénotypiques. L'AUC du modèle est de 50%, ce qui équivaut à une classification des individus au hasard (cf figure 25).

La justesse de la prédiction (cf accuracy) est dans le meilleur des cas de 81% mais il faut prendre en compte le fait que le jeu de données est déséquilibré et présente beaucoup plus d'individus négatifs que positifs. Ici, le modèle prédit très mal les individus positifs, parmi les prédits positifs, aucun ne le sont réellement (cf sensitivity). Dans les 3 premiers cas (lignes 1 à 3 de la figure 25), le modèle classe en fait tous les individus en tant que négatifs. Dans le dernier cas (ligne 4, figure 25), tous les individus sont classés en positifs.

Evaluation de la prédiction du statut HLB							
variables phénotypiques uniquement							
Accuracy	Precision	Specificity	Error_rate	Sensitivity	auc	CTseuil	lots_indépendants
0.810	NaN	1	0.190	0	0.5	32	oui
0.619	NaN	1	0.381	0	0.5	36	oui
0.643	NaN	1	0.357	0	0.5	32	non (arbre)
0.571	0.571	0	0.429	1	0.5	36	non (arbre)

Figure 25 Evaluation de la prédiction du statut HLB grâce aux variables phénotypiques seulement

4. Discussion

4.1. Prédiction du parcellaire agrumicole

En épidémiologie du paysage, les méthodes qualifiées de « spatialement explicites » nécessitent une connaissance précise de la localisation des hôtes potentiels de la maladie étudiée, dans notre cas des parcelles d'agrumes, et doivent donc concilier avec un manque de données ou de ressources disponibles. La photo-interprétation d'images satellites permet d'identifier des parcelles susceptibles d'être des agrumes mais elle manque de justesse pour effectuer des simulations par des modèles épidémiologiques sur le parcellaire photo-interprété. La prédiction du parcellaire agrumicole grâce à des méthodes de *machine learning* apparaît donc comme une solution prometteuse car on réduit considérablement l'effort d'échantillonnage en se focalisant seulement sur la construction d'une base de données d'apprentissage du modèle. En théorie, les SVM sont plus adaptés à des problèmes de classification à deux classes alors que les Random forest à des problèmes multi classes (Statnikov et al., 2008). On s'attendait donc à obtenir une meilleure classification par SVM mais ce n'est pas le cas. En effet, de manière générale, le modèle Random Forest ressort comme l'un des meilleurs classificateurs (Fernandez-Delgado et al., 2014).

Grâce au Random Forest, la prédiction des parcelles non-agrumes est bien meilleure que celles des parcelles d'agrumes (spécificité de 86% et sensibilité de 75%). La prédiction de la nature des parcelles peut donc être utile pour restreindre l'effort de terrain en n'effectuant des vérifications terrain seulement sur les parcelles prédites en agrumes. Le déséquilibre des effectifs des deux classes peut expliquer en partie la meilleure prédiction des parcelles non-agrumes car le classificateur est biaisé et favorise les prédictions vers la classe majoritaire (Kubat, 2015).

Deux solutions existent pour faire face à ce problème de déséquilibre de classes : effectuer un sous échantillonnage aléatoire du jeu de données pour se retrouver avec le même nombre d'individus dans chaque classe ou accorder un poids plus important aux individus de la classe minoritaire (Burez & Van den Poel, 2009).

Néanmoins, la nature de certaines parcelles ne peut être prédite car leurs valeurs pour certaines variables pédoclimatiques ne sont pas présentes dans la base d'apprentissage du modèle. Il est donc important de bien penser l'échantillonnage afin d'établir un modèle prédictif sur l'ensemble du territoire étudié.

4.2. Prédiction de l'état sanitaire des parcelles et cartographie du risque HLB

La prédiction de l'état sanitaire des parcelles permet de repérer les zones où la maladie a la plus grande probabilité de se trouver en relation avec les variables pédoclimatiques, afin d'améliorer la surveillance en se concentrant sur les zones à plus haut risque (Stephen Parnell et al., 2017). La prédiction de l'état sanitaire pourra être utilisée pour identifier les bassins de production indemnes et ceux à haut risque dans le but d'accompagner les institutions dans la prise de mesures de gestion de l'épidémie de HLB ainsi que les agriculteurs souhaitant se lancer dans la production d'agrumes.

De la même manière que pour la prédiction de la nature des parcelles, on fait face à un déséquilibre de classes (171 parcelles négatives contre 55 positives) ce qui porte préjudice à la bonne prédiction du statut des parcelles (Kubat, 2015). A l'avenir, l'analyse des échantillons prélevés qui n'ont pas encore été testés par qPCR permettra d'enrichir la base de données d'apprentissage et d'améliorer la prédiction. La prédiction du statut infecté pourra donc être utilisée pour orienter l'effort d'échantillonnage vers les parcelles les plus susceptibles d'être positives.

Enfin, une des limites de cette prédiction de l'état sanitaire des parcelles est qu'elle ne prend pas en compte la dynamique temporelle de la maladie. Pour cela, la date d'échantillonnage devrait apparaître dans les variables explicatives.

Concernant les variables les plus importantes dans la classification du statut des parcelles, on retrouve des variables liées aux bassins de production d'agrumes (Commune, type de sol STIPA) et des variables bioclimatiques (altitude, température minimale journalière) qui peuvent être corrélées à la localisation des parcelles d'agrumes, à la niche écologique des psylles, vecteurs de la maladie du HLB, et/ou aux préférences thermiques de la bactérie qui demeurent assez mal connues. Les densités en vergers d'agrumes dans un rayon variant de 1,1 à 2,1 km ont également un rôle majeur dans la classification et pourraient témoigner des capacités de dispersion de *Diaphorina citri* qui s'élevaient jusqu'à 2,4 km (Stelinski, 2019).

4.3. Prédiction de l'état sanitaire des arbres par SPIR

L'utilisation de données de SPIR couplées à des variables phénotypiques est prometteuse pour établir un diagnostic des arbres atteints du HLB. Avec seulement une mesure de réflectance, on peut prédire la présence de *Ca. Liberibacter asiaticus* dans une feuille avec un justesse (accuracy) de 75,9%, tout en optimisant la détection des échantillons positifs et des négatifs. D'autres études parviennent à une justesse de prédiction du statut des feuilles de l'ordre de 80% en prédisant sur la moyenne de 5 à 6 spectres PIR (Sankaran et al., 2010; Sankaran & Ehsani, 2011). Il serait intéressant de reprendre la prédiction du statut des feuilles en moyennant les valeurs de réflectance de nos 6 spectres PIR collectés, on s'attend à une nette amélioration de la prédiction.

En fixant un C_T seuil à 32 plutôt qu'à 36, on maximise le taux de vrais positifs parmi les échantillons prédits positifs (cf precision). Les échantillons sont donc plus facilement séparés en deux classes. On peut supposer qu'il y a des échantillons réellement négatifs présentant un C_T compris entre 32 et 36 (contaminations au laboratoire) ou que la méthode spectrale manque de capacité de discrimination chez des échantillons faiblement contaminés. Lors des expériences de qPCR, les témoins négatifs n'ont pas montré de signal tardif d'amplification.

Les longueurs d'onde les plus caractéristiques de la signature spectrale de la maladie sont à mettre en lien avec des différences dans la composition chimique des feuilles saines et malades. Le pic d'absorbance (et donc de l'inverse de la réflectance également) entre 900 et 1050 nm est attribué à la présence de glucides dans les feuilles sous forme d'amidon (Dupuy et al., 1997; Hawkins et al., 2010; Kačuráková & Wilson, 2001; Sankaran et al., 2010). Le HLB

cause en effet la formation de bouchons de callose dans les vaisseaux du phloème, ce qui empêche la distribution des photo-assimilés vers les racines, d'où le stockage sous forme d'amidon au niveau des feuilles. Les longueurs d'onde comprises entre 570 et 650 nm sont attribuées à la présence d'eau. Dans notre modèle, ces longueurs d'onde sont importantes dans la prédiction du statut HLB. Il n'a pas été démontré qu'il existait une différence de teneur en eau entre les feuilles saines et malades (Sankaran et al., 2010). La prise de mesures de SPIR en blocs lors de l'expérience nous permet d'écarter tout biais lié à un dessèchement des feuilles au cours de la prise de mesures.

Il serait intéressant d'établir différents modèles prédictifs du statut infecté des arbres en fonction des caractères phénotypiques des feuilles étudiées. Néanmoins, on ne dispose pas d'assez de données pour calibrer des modèles sur des sous-jeux de données.

Lorsque l'on considère tous les lots d'un même arbre comme indépendants (on peut avoir à la fois des lots positifs et négatifs sur un même arbre), on augmente la justesse de la prédiction en améliorant la détection des positifs (precision, sensibility) mais en diminuant celle des négatifs (specificity). Certains lots peuvent être positifs et négatifs sur un même arbre, cela témoigne de la distribution hétérogène de la bactérie au sein des rameaux de l'arbre. Cette hétérogénéité a été montrée dans plusieurs études (Li et al., 2009; Louzada et al., 2016; Hilf & Luo, 2018).

La relativement bonne prédiction du statut des spectres lorsque l'ensemble des lots du même arbre sont considérés comme positifs dès lors qu'au moins l'un d'entre eux est positif nous laisse penser qu'il existe une signature spectrale systémique de la maladie. La présence de la maladie au sein de rameaux infectés affecte la signature spectrale des feuilles sur les rameaux sains.

La SPIR est donc une méthode prometteuse pour diagnostiquer les arbres atteints du HLB. Des méthodes similaires non destructives sont actuellement en cours de développement, telles que la spectroscopie Raman qui permet d'établir un diagnostic plus précoce et plus juste qu'un diagnostic par qPCR (Sanchez et al., 2020) en étudiant les variations d'énergie de photons inélastiques traversant un échantillon (Bumrah & Sharma, 2016).

5. Conclusion

Face à la réémergence actuelle du HLB à la Réunion, des mesures de gestion efficaces doivent être prises rapidement. La construction de modèles épidémiologiques permettra de simuler différents scénarios de gestion et d'en évaluer l'efficacité. La cartographie du risque de présence de la maladie permettra d'accentuer l'effort de surveillance sur les foyers de production les plus à risque.

Cependant, les modèles épidémiologiques requièrent la connaissance exacte du parcellaire agrumicole, c'est-à-dire des sous-populations d'hôtes du HLB, car le pathogène est spécialiste et le paysage agrumicole est fortement morcelé. La prédiction de la nature des vergers préalablement repérés par photo-interprétation d'images satellites, par des méthodes de machine learning en utilisant des variables pédoclimatiques apporte une solution au manque de ressources disponibles et permet de cartographier un parcellaire « prédit ».

La cartographie du risque de présence de la maladie est également essentielle pour adapter les mesures de gestion aux zones à plus haut risque et aux zones indemnes. Cependant, un échantillonnage et un diagnostic de l'ensemble des parcelles d'agrumes n'est pas envisageable au regard des ressources disponibles. Ainsi, prédire le statut HLB des parcelles d'agrumes par machine learning est une solution prometteuse bien que la base de données d'apprentissage doive être étoffée pour garantir une prédiction plus juste.

Enfin, il est nécessaire de réduire le coût et le délai du diagnostic HLB afin de réaliser un échantillonnage plus rapide et plus précis. L'objectif étant de déterminer exactement quels arbres sont malades du HLB au sein d'une parcelle pour accompagner au mieux les agriculteurs. Nos travaux soulignent le potentiel d'un diagnostic HLB par SPIR, à la fois rapide, juste et peu onéreux qui pourrait être couplé à une confirmation moléculaire à l'aide de méthodes comme la LAMP (loop-mediated isothermal amplication) plus rapide et portable que la qPCR. Plusieurs méthodes publiées sont actuellement en cours d'évaluation par l'ANSES.

6. Bibliographie

- Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., & Daszak, P. (2004). Emerging infectious diseases of plants : Pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, 19(10), 535-544. <https://doi.org/10.1016/j.tree.2004.07.021>
- Aubert, B., Grisoni, M., Villemin, M., & Rossolin, G. (1996). A Case Study of Huanglongbing (Greening) Control in Reunion. *International Organization of Citrus Virologists Conference Proceedings (1957-2010)*, 13(13). <https://escholarship.org/uc/item/8024t7fj>
- Bertaux, C. (2014). *Mise en place de la spectroscopie proche infrarouge selon les approches Quality By Design et Lean Six Sigma pour le pilotage en ligne de l'humidité d'une poudre* [Université Bordeaux 2]. <https://dumas.ccsd.cnrs.fr/dumas-01109550/document>
- Bové, J. (2006). Huanglongbing : A destructive, newly emerging, century-old disease of citrus. *Journal of Plant Pathology*, 88. <https://doi.org/10.4454/jpp.v88i1.828>
- Bové, J. M. (2014). Huanglongbing or yellow shoot, a disease of Gondwanan origin : Will it destroy citrus worldwide? *Phytoparasitica*, 42(5), 579-583. <https://doi.org/10.1007/s12600-014-0415-4>
- Brasier, C. M. (2008). The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathology*, 57(5), 792-808. <https://doi.org/10.1111/j.1365-3059.2008.01886.x>
- Bumbrah, G. S., & Sharma, R. M. (2016). Raman spectroscopy – Basic principle, instrumentation and selected applications for the characterization of drugs of abuse. *Egyptian Journal of Forensic Sciences*, 6(3), 209-215. <https://doi.org/10.1016/j.ejfs.2015.06.001>
- Burez, J., & Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. *Expert Systems with Applications*, 36(3, Part 1), 4626-4636. <https://doi.org/10.1016/j.eswa.2008.05.027>
- Carlsson-Granér, U., & Thrall, P. H. (2002). The spatial distribution of plant populations, disease dynamics and evolution of resistance. *Oikos*, 97(1), 97-110. Scopus. <https://doi.org/10.1034/j.1600-0706.2002.970110.x>
- Cifuentes-Arenas, J. C., Beattie, G. A. C., Peña, L., & Lopes, S. A. (2019). *Murraya paniculata* and *Swinglea glutinosa* as Short-Term Transient Hosts of ' *Candidatus Liberibacter asiaticus*' and Implications for the Spread of Huanglongbing. *Phytopathology*®, 109(12), 2064-2073. <https://doi.org/10.1094/PHYTO-06-19-0216-R>
- Cirad AWARE. (2020). <https://aware.cirad.fr/>
- Cook, A. r, Gibson, G. j, Gottwald, T. r, & Gilligan, C. a. (2008). Constructing the effect of alternative intervention strategies on historic epidemics. *Journal of The Royal Society Interface*, 5(27), 1203-1213. <https://doi.org/10.1098/rsif.2008.0030>
- Cortez, P., & Embrechts, M. J. (2013). Using sensitivity analysis and visualization techniques to open black box data mining models. *Information Sciences*, 225, 1-17. <https://doi.org/10.1016/j.ins.2012.10.039>
- Dupuy, S. (2019). *Reunion Island—2018, learning spatial database*.
- Dupuy, Wojciechowski, Ta, C. D., Huvenne, J. P., & Legrand, P. (1997). Mid-infrared spectroscopy and chemometrics in corn starch classification. *Journal of Molecular Structure*, 410-411, 551-554. [https://doi.org/10.1016/S0022-2860\(96\)09517-8](https://doi.org/10.1016/S0022-2860(96)09517-8)

- Fernandez-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). *Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?* 15, 3133-3181.
- Gilland, B. (2002). World population and food supply : Can food production keep pace with population growth in the next half-century? *Food Policy*, 27(1), 47-63.
[https://doi.org/10.1016/S0306-9192\(02\)00002-7](https://doi.org/10.1016/S0306-9192(02)00002-7)
- Gottwald, T. R. (1989). Preliminary Analysis of Citrus Greening (Huanglungbin) Epidemics in the People's Republic of China and French Reunion Island. *Phytopathology*, 79(6), 687.
<https://doi.org/10.1094/Phyto-79-687>
- Graziosi, I., Minato, N., Alvarez, E., Ngo, D. T., Hoat, T. X., Aye, T. M., Manuel Pardo, J., Wongtiem, P., & Wyckhuys, K. A. G. (2016). Emerging pests and diseases of South-east Asian cassava : A comprehensive evaluation of geographic priorities, management options and research needs. *Pest Management Science*, 72(6), 1071-1089. <https://doi.org/10.1002/ps.4250>
- Grenfell, B., & Harwood, J. (1997). (Meta)population dynamics of infectious diseases. *Trends in Ecology & Evolution*, 12(10), 395-399. [https://doi.org/10.1016/S0169-5347\(97\)01174-9](https://doi.org/10.1016/S0169-5347(97)01174-9)
- Gullino, M. L., Leroux, P., & Smith, C. M. (2000). Uses and challenges of novel compounds for plant disease control. *Crop Protection*, 19(1), 1-11. [https://doi.org/10.1016/S0261-2194\(99\)00095-2](https://doi.org/10.1016/S0261-2194(99)00095-2)
- Haight, R. G., Homans, F. R., Horie, T., Mehta, S. V., Smith, D. J., & Venette, R. C. (2011). Assessing the Cost of an Invasive Forest Pathogen : A Case Study with Oak Wilt. *Environmental Management*, 47(3), 506-517. <https://doi.org/10.1007/s00267-011-9624-5>
- Halbert, S. E., & Manjunath, K. L. (2004). Asian citrus psyllids (Sternorrhyncha : Psyllidae) and greening disease of citrus: A literature review and assessment of risk in Florida. *Florida Entomologist*, 87(3), 330-353. [https://doi.org/10.1653/0015-4040\(2004\)087\[0330:ACPSPA\]2.0.CO;2](https://doi.org/10.1653/0015-4040(2004)087[0330:ACPSPA]2.0.CO;2)
- Hanski, I. (2001). Spatially realistic theory of metapopulation ecology. *Naturwissenschaften*, 88(9), 372-381. <https://doi.org/10.1007/s001140100246>
- Hawkins, S. A., Park, B., Poole, G. H., Gottwald, T., Windham, W. R., & Lawrence, K. C. (2010). Detection of Citrus Huanglongbing by Fourier Transform Infrared—Attenuated Total Reflection Spectroscopy: *Applied Spectroscopy*.
<https://doi.org/10.1366/000370210790572043>
- Hilf, M. E., & Luo, W. (2018). Dynamics of « Candidatus Liberibacter asiaticus » Colonization of New Growth of Citrus. *Phytopathology*, 108(10), 1165-1171. <https://doi.org/10.1094/PHYTO-12-17-0408-R>
- Jones, D. R., & Baker, R. H. A. (2007). Introductions of non-native plant pathogens into Great Britain, 1970-2004. *Plant Pathology*, 56(5), 891-910. <https://doi.org/10.1111/j.1365-3059.2007.01619.x>
- Kačuráková, M., & Wilson, R. H. (2001). Developments in mid-infrared FT-IR spectroscopy of selected carbohydrates. *Carbohydrate Polymers*, 44(4), 291-303. [https://doi.org/10.1016/S0144-8617\(00\)00245-9](https://doi.org/10.1016/S0144-8617(00)00245-9)
- Kubat, M. (2015). *An Introduction to Machine Learning*. Springer International Publishing.
<https://doi.org/10.1007/978-3-319-20010-1>
- Legg, J., Somado, E. A., Barker, I., Beach, L., Ceballos, H., Cuellar, W., Elkhoury, W., Gerling, D., Helsen, J., Hershey, C., Jarvis, A., Kulakow, P., Kumar, L., Lorenzen, J., Lynam, J., McMahon, M., Maruthi, G., Miano, D., Mtunda, K., ... Fauquet, C. (2014). A global alliance declaring war

- on cassava viruses in Africa. *Food Security*, 6(2), 231-248. <https://doi.org/10.1007/s12571-014-0340-x>
- Li, W., Hartung, J. S., & Levy, L. (2006). Quantitative real-time PCR for detection and identification of *Candidatus Liberibacter* species associated with citrus huanglongbing. *Journal of Microbiological Methods*, 66(1), 104-115. <https://doi.org/10.1016/j.mimet.2005.10.018>
- Li, W., Levy, L., & Hartung, J. S. (2009). Quantitative distribution of « *Candidatus Liberibacter asiaticus* » in citrus plants with citrus huanglongbing. *Phytopathology*, 99(2), 139-144. <https://doi.org/10.1094/PHYTO-99-2-0139>
- Louzada, E. S., Vazquez, O. E., Braswell, W. E., Yanev, G., Devanaboina, M., & Kunta, M. (2016). Distribution of « *Candidatus Liberibacter asiaticus* » Above and Below Ground in Texas Citrus. *Phytopathology*, 106(7), 702-709. <https://doi.org/10.1094/PHYTO-01-16-0004-R>
- Ma, Z. H., & Michailides, T. J. (2005). Advances in understanding molecular mechanisms of fungicide resistance and molecular detection of resistant genotypes in phytopathogenic fungi. *Crop Protection*, 24(10), 853-863. <https://doi.org/10.1016/j.cropro.2005.01.011>
- Martinetti, D., & Soubeyrand, S. (2018). Identifying Lookouts for Epidemio-Surveillance : Application to the Emergence of *Xylella fastidiosa* in France. *Phytopathology*, 109(2), 265-276. <https://doi.org/10.1094/PHYTO-07-18-0237-FI>
- Meentemeyer, R., Haas, S., & Václavík, T. (2012). Landscape Epidemiology of Emerging Infectious Diseases in Natural and Human-Altered Ecosystems. *Annual review of phytopathology*, 50, 379-402. <https://doi.org/10.1146/annurev-phyto-081211-172938>
- Munir, S., He, P., Wu, Y., He, P., Khan, S., Huang, M., Cui, W., He, P., & He, Y. (2018). Huanglongbing Control : Perhaps the End of the Beginning. *Microbial Ecology*, 76(1), 192-204. <https://doi.org/10.1007/s00248-017-1123-7>
- Nnadi Nnaemeka, E., Nimzing, L., Mark Ojogba, O., & Onyedibe Kenneth, I. (2011). Landscape epidemiology : An emerging perspective in the mapping and modelling of disease and disease risk factors. *Asian Pacific Journal of Tropical Disease*, 1(3), 247-250. [https://doi.org/10.1016/S2222-1808\(11\)60041-8](https://doi.org/10.1016/S2222-1808(11)60041-8)
- Ostfeld, R. S., Glass, G. E., & Keesing, F. (2005). Spatial epidemiology : An emerging (or re-emerging) discipline. *Trends in Ecology & Evolution*, 20(6), 328-336. <https://doi.org/10.1016/j.tree.2005.03.009>
- Papaix, J., Goyeau, H., Du Cheyron, P., Monod, H., & Lannou, C. (2011). Influence of cultivated landscape composition on variety resistance:an assessment based on wheat leaf rust epidemics. *New Phytologist*, 191(4), 1095-1107. <https://doi.org/10.1111/j.1469-8137.2011.03764.x>
- Parnell, S., Gottwald, T. R., Riley, T., & van den Bosch, F. (2014). A generic risk-based surveying method for invading plant pathogens. *Ecological Applications: A Publication of the Ecological Society of America*, 24(4), 779-790. <https://doi.org/10.1890/13-0704.1>
- Parnell, Stephen, van den Bosch, F., Gottwald, T., & Gilligan, C. A. (2017). Surveillance to Inform Control of Emerging Plant Diseases : An Epidemiological Perspective. *Annual Review of Phytopathology*, 55(1), 591-610. <https://doi.org/10.1146/annurev-phyto-080516-035334>
- Pautasso, M., Doering, T. F., Garbelotto, M., Pellis, L., & Jeger, M. J. (2012). Impacts of climate change on plant diseases-opinions and trends. *European Journal of Plant Pathology*, 133(1), 295-313. <https://doi.org/10.1007/s10658-012-9936-1>
- Plantegenest, M., Le May, C., & Fabre, F. (2007). Landscape epidemiology of plant diseases. *Journal of The Royal Society Interface*, 4(16), 963-972. <https://doi.org/10.1098/rsif.2007.1114>

- Pouzet, D. (2002). *Caractérisation chimique des sols réunionnais*.
- QGIS.org. (2020). *QGIS Geographic Information System. Open Source Geospatial Foundation Project. Version 3.10.4 Coruna*. <http://qgis.org>
- R Development Core Team. (2005). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0*. <http://www.R-project.org>
- Rimbaud, L., Bruchou, C., Dallot, S., Pleydell, D. R. J., Jacquot, E., Soubeyrand, S., & Thébaud, G. (2018). Using sensitivity analysis to identify key factors for the propagation of a plant epidemic. *Royal Society Open Science*, 5(1). <https://doi.org/10.1098/rsos.171435>
- Rimbaud, L., Dallot, S., Gottwald, T., Decroocq, V., Jacquot, E., Soubeyrand, S., & Thébaud, G. (2015). Sharka Epidemiology and Worldwide Management Strategies : Learning Lessons to Optimize Disease Control in Perennial Plants. *Annual Review of Phytopathology*, 53(1), 357-378. <https://doi.org/10.1146/annurev-phyto-080614-120140>
- Sanchez, L., Pant, S., Mandadi, K., & Kurouski, D. (2020). Raman Spectroscopy vs Quantitative Polymerase Chain Reaction In Early Stage Huanglongbing Diagnostics. *Scientific Reports*, 10(1), 10101. <https://doi.org/10.1038/s41598-020-67148-6>
- Sankaran, S., & Ehsani, R. (2011). Visible-near infrared spectroscopy based citrus greening detection : Evaluation of spectral feature extraction techniques. *Crop Protection*, 30(11), 1508-1513. <https://doi.org/10.1016/j.cropro.2011.07.005>
- Sankaran, S., Ehsani, R., & Etxeberria, E. (2010). Mid-infrared spectroscopy for detection of Huanglongbing (greening) in citrus leaves. *Talanta*, 83(2), 574-581. <https://doi.org/10.1016/j.talanta.2010.10.008>
- Sechler, A., Schuenzel, E. L., Cooke, P., Donnua, S., Thaveechai, N., Postnikova, E., Stone, A. L., Schneider, W. L., Damsteegt, V. D., & Schaad, N. W. (2009). Cultivation of 'Candidatus Liberibacter asiaticus', 'Ca. L. africanus', and 'Ca. L. americanus' Associated with Huanglongbing. *Phytopathology*®, 99(5), 480-486. <https://doi.org/10.1094/PHYTO-99-5-0480>
- Spraying diseased citrus orchards with antibiotics could backfire. (2019). *Nature*, 567(7748), 283-283. <https://doi.org/10.1038/d41586-019-00875-7>
- Statnikov, A., Wang, L., & Aliferis, C. F. (2008). A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinformatics*, 9, 319. <https://doi.org/10.1186/1471-2105-9-319>
- Stelinski, L. L. (2019). Ecological Aspects of the Vector-Borne Bacterial Disease, Citrus Greening (Huanglongbing) : Dispersal and Host Use by Asian Citrus Psyllid, Diaphorina Citri Kuwayama. *Insects*, 10(7), 208. <https://doi.org/10.3390/insects10070208>
- Taylor, P., & Reeder, R. (2020). Antibiotic use on crops in low and middle-income countries based on recommendations made by agricultural advisors. *CABI Agriculture and Bioscience*, 1(1), 1. <https://doi.org/10.1186/s43170-020-00001-y>
- Wang, N., & Trivedi, P. (2013). Citrus huanglongbing : A newly relevant disease presents unprecedented challenges. *Phytopathology*, 103(7), 652-665. <https://doi.org/10.1094/PHYTO-12-12-0331-RVW>
- Wilson, C., & Tisdell, C. (2001). Why farmers continue to use pesticides despite environmental, health and sustainability costs. *Ecological Economics*, 39(3), 449-462. [https://doi.org/10.1016/S0921-8009\(01\)00238-5](https://doi.org/10.1016/S0921-8009(01)00238-5)

Wyenandt, C. A., McGrath, M. T., Everts, K. L., Rideout, S. L., Gugino, B. K., & Kleczewski, N. (2018).

Fungicide Resistance Management Guidelines for Cucurbit Downy and Powdery Mildew

Control in the Mid-Atlantic and Northeast Regions of the United States in 2018. *Plant Health*

Progress, 19(1), 34-36. <https://doi.org/10.1094/PHP-12-17-0077-BR>