Structure of agricultural landscapes and epidemic risk, a demo-genetic approach.

Julien Papaïx

INRA - Ecole doctorale ABIES - AgroParisTech

Thèse

pour obtenir le grade de docteur délivré par

L'INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT (AgroParisTech)

Spécialité : Biologie des populations et écologie

présentée et soutenue publiquement

le 26 septembre 2011

par

Julien Papaïx

Structure du paysage agricole et risque épidémique, une approche démo-génétique

---

Thèse effectuée au sein de l'unité mixte INRA-AgroParisTech BIOlogie et GEstion du Risque en agriculture (UMR 1290 BIOGER, F-78850 Thiverval Grignon) et de l'unité INRA Mathématiques et Informatiques Appliquées de Jouy-en-Josas (UR 341 MIAJ, F-78352 Jouy-en-Josas).

Thèse encadrée par Christian Lannou (DR BIOGER) et Hervé Monod (DR MIAJ).

Présentée devant le jury composé de:

M. Frank van den Bosch		
Principal Investigator	Rothamsted Research	*Rapporteur*
M. Sylvain Gandon		
Directeur de recherche	CEFE - CNRS	*Rapporteur*
Mme. Elisabeth Fournier		
Chargée de recherche	BGPI - INRA	*Examinateur*
M. Philippe Gate		
Directeur scientifique	Arvalis - Institut du végétal	*Examinateur*
M. Jean-Jacques Daudin		
Professeur	AgroParisTech	*Représentant ED*

Structure du paysage agricole et risque épidémique, une approche démo-génétique

L'intensification de l'agriculture a amélioré de façon considérable la production alimentaire ces dernières cinquante années mais elle s'est accompagnée d'un impact croissant sur l'environnement. En particulier, la modernisation de l'agriculture a impliqué une simplification de la structure des paysages agricoles rendant nos agro-écosystèmes plus sensibles au risque épidémique. L'utilisation de la diversité génétique des cultures est une solution prometteuse pour réduire le risque d'occurrence et de propagation des maladies des cultures. Elle nécessite cependant une gestion collective des espaces agricoles. En conséquence, l'échelle d'étude ne doit plus se focaliser sur la parcelle mais sur le paysage. Dans cette thèse, nous nous interessons aux processus se déroulant à l'échelle du paysage et au rôle de la diversité des plantes cultivées pour le contrôle des épidémies. Nous avons identifié trois questions: comment les populations pathogènes se propagent-elles dans un paysage d'hôtes hétérogène ? Comment les différents génotypes composant la population pathogène entrent-ils en compétition au sein d'une population hôte diversifiée ? et, à plus long terme, comment les populations pathogènes évoluent-elles en réponse à la structure des populations hôtes ? Chacune de ces questions a été approfondie grâce à l'analyse de données obtenues en condition de production mais aussi par des approches théoriques. Nous avons montré que la composition et la structure spatiale des populations hôtes influence fortement la population pathogène. Cependant, les recommandations que peut fournir ce travail pour gérer la diversité génétique dépendent de l'objectif visé.

Evolution | Epidémiologie du paysage | Hétérogénéité spatiale | Spécialisation

Landscape structure and epidemic risk, a demo-genetic approach

Agriculture intensification has improved food production impressively in the past 50 years but this came with an increasing impact on the environment. In particular, modern agriculture has led to the simplification of the environmental structure over vast areas. As a consequence, agro-ecosystems are particularly susceptible to epidemics. The increase of crop genetic diversity is a promising way for reducing the risk of occurrence and development of diseases in crops but the technical and organisational conditions required to manage the genetic resources at this scale have not been established yet. This will require shifting the scale of crop protection investigations from the field to the agricultural landscape. In this PhD thesis we focus on landscape-scale processes and on the potential role of functional diversity in cultivated landscapes to better control plant diseases. We identified three questions: how does a pathogen population spread over a heterogeneous host landscape? How do pathogen genotypes compete in a diversified host population? And, in a longer term, how do pathogen populations evolve in response to host landscape structure? Each of these questions is investigated through the analysis of real data and the development of theoretical approaches. We demonstrate that the composition and the spatial structure of the host landscape greatly influence the pathogen population dynamics and evolution. The recommendations that this work may provide in order to practically manage the genetic resources will depend on the desired aim and will request further collaborative work with the professional operators.

Evolution | Landscape epidemiology | Spatial heterogeneity | Specialisation

Remerciements

Je tenais tout d'abord à remercier Sylvain Gandon et Frank van den Bosch pour avoir pris le temps de lire et d'évaluer mon travail. Je remercie aussi Elisabeth Fournier, Philippe Gate et Jean-Jacques Daudin d'avoir accepté de participer à mon jury de thèse. Je souhaitais remercier également les membres de mon comité de thèse pour m'avoir suivi pendant ces trois ans: Philippe Du Cheyron, Jean-Jacques Daudin, Etienne Klein, Manuel Plantegenest mais aussi Frédérique Angevin, Marie-Hélène Jeuffroy, Claire Lavigne, Ivan Sache et Samuel Soubeyrand.

Cette thèse chemine entre l'épidémiologie végétale, les statistiques et la modélisation. Elle n'a pu se faire sans de nombreuses discussions avec des personnes venant d'horizons divers et variés. Je les en remercie. Je tiens à remercier tout particulièrement Christian et Hervé pour leur patience, leurs conseils et pour m'avoir accompagné pendant ces trois années de recherches. Je remercie également Henriette, Suzanne et Olivier pour m'avoir autant aidé. Je suis aussi reconnaissant envers l'ensemble des personnes des deux unités (BIOGER et MIAJ) qui m'ont accueilli, fait rire, soutenu et appris énormément de choses. Je n'oublie pas non plus la master boulet team du CEFE et les vinoc' de l'Agro qui sont toujours par là !

Cette thèse a été réalisée dans le cadre d'un contrat ASC financé par l'INRA avec la participation des départements SPE et MIA.

Tenez-le bien, Applefish !!

Manu Larcenet,
Le savant et la femme (Soyons fous 1, 2005)

Contents

List of Tables

List of Figures

Eléments de contexte

L'intensification de l'agriculture a amélioré de façon considérable la production alimentaire au cours de ces dernières cinquantes années (Food and Agriculture Organization, 2009). Cependant, ce gain de productivité s'est accompagné d'un impact croissant sur l'environnement (Vitousek et al., 1997; Tilman et al., 2002). En particulier, la modernisation de l'agriculture a impliqué une simplification de la structure des paysages agricoles sur de grands espaces en remplaçant la diversité naturelle par un faible nombre de plantes cultivées et d'animaux domestiques (Altieri, 1999). Cette trop forte homogénéité a rendu nos agro-écosystèmes particulièrement sensibles au risque épidémique (Pautasso et al., 2005; Stukenbrock and McDonald, 2008; Meehan et al., 2011). Pour contrecarrer cette sensibilité, l'approche dominante, et la plus simple, est l'utilisation de produit phytosanitaires et elle vise essentiellement les échelles de la parcelle et du cycle cultural (des rotations courtes sont aussi pratiquées). Cependant, la trop forte dépendance de nos agro-écosystèmes à l'utilisation de pesticides est aujourd'hui remise en question par la société (Grenelle, 2008) mais aussi par une partie de la filière agricole. En effet, elle impose de fortes contraintes sur les biocénoses des espaces cultivés, des espaces adjacents et des ressources naturelles. De plus, le recours aux biocides pour le contrôle des parasites implique un coût non négligeable pour les agriculteurs. A titre d'exemple, l'utilisation des fongicides sur blé représente environ 70 € par hectare ce qui, au total, peut représenter une dépense de 300 millions d'euros par an à l'échelle de la France (Arvalis). Enfin, cette volonté de changer d'approche pour la protection des cultures se reflète dans une réglementation de plus en plus stricte.

Les approches alternatives qui permettraient la transition vers une agriculture écologiquement intensive, sans pour autant augmenter de façon critique les risques de perte de récolte liés à la pression parasitaire, doivent mobiliser des leviers agronomiques efficaces (Tilman, 1999; Meynard et al., 2002). Parmi ceux-ci, l'utilisation de la diversité génétique inter- mais aussi intra-spécifique semble une solution prometteuse pour réduire le risque d'occurrence et de propagation des maladies des cultures. Cependant, le remplacement d'une solution pesticide par une solution génétique nécessite encore plus une gestion collective de cette ressource. En effet, de nombreux exemples de contournements de résistances variétales dus à leur popularité existent (Wolfe and Schwarzbach, 1978; Samborski, 1985; Bayles

et al., 2000). Comme toute ressource finie, les résistances variétales sont sujettes à la *tragédie des communs* (Hardin, 1968): l'utilisation non raisonnée de la ressource amène à sa disparition. Il est donc nécessaire de définir au niveau technique mais aussi organisationnel de nouvelles stratégies de gestion de notre système agricole. L'utilisation collective de la ressource génétique nécessite des changements drastiques à la fois des itinéraires culturaux et de l'organisation de la filière dans son ensemble (Butault et al., 2010). Il existe bien sûr des exemples d'études visant à produire des conseils aux agriculteurs de façon globale (Gladders et al., 2006), mais elles restent de l'ordre de conseils généraux : les études épidémiologiques doivent changer d'échelle pour se placer à celle du paysage afin d'introduire cet élément de façon explicite dans les modèles pour améliorer notre compréhension des épidémies ainsi que leur prédiction (Ostfeld et al., 2005).

L'étude des interactions entre les structures paysagères et les processus écologiques relève de l'écologie du paysage (Turner, 2005). L'idée de coupler ces approches avec l'étude de dynamiques épidémiques date des années 1930 et a pris le nom d'épidémiologie du paysage (Ostfeld et al., 2005). Les approches d'épidémiologie du paysage ont été relativement bien développées en épidémiologie humaine, animale ou encore pour la gestion des parasites en agriculture (Langlois et al., 2001; Allan et al., 2003; Bianchi et al., 2006; Tildesley et al., 2010). En pathologie végétale, si l'idée de définir des stratégies de gestion des variétés à l'échelle de paysages agricoles est dans l'air depuis la fin des années 1970 (Zadoks and Kampmeijer, 1977; Mundt and Brophy, 1988), le potentiel de l'épidémiologie du paysage pour en explorer de nouvelles voies est encore trop peu exploité par les pathologistes (Plantegenest et al., 2007). Dans ce travail de thèse nous nous intéressons justement aux processus ayant cours à l'échelle du paysage dans l'idée de ré-introduire de la diversité fonctionelle au sein de nos cultures pour mieux contrôler les maladies des plantes.

Dans son acception la plus large, la *diversité* décrit le nombre d'entités (une entité pouvant aller du génotype à l'écosystème), leur abondance et leur distribution (Haccou et al., 2005). Pour comprendre le lien entre la diversité et une propriété particulière d'un écosystème, la diversité doit être considérée du point de vue du trait fonctionnel par rapport à la propriété étudiée. Ainsi, la *diversité fonctionnelle* en épidémiologie aura trait à la diversité des hôtes jouant un rôle différent vis-à-vis d'un pathogène ou d'une souche de pathogène. Pautasso et al. (2005) ont classé les écosystèmes en fonction de leur diversité et de leur sensibilité aux maladies. Ils ont identifié quatre cas: (i) les monocultures chanceuses (écosystèmes peu diversifiés et peu sensibles), (ii) les écosystèmes sensibles et peu diversifiés, (iii) les écosystème sensibles et diversifiés et (iv) les écosystèmes gouvernés par l'hypothèse d'assurance (écosystèmes diversifiés et peu sensibles). Dans ce travail nous soutiendrons la thèse selon laquelle les sytèmes agricoles doivent effectuer la transition de (ii) vers (iv). L'hypothèse d'assurance lie le fonctionnement d'un écosystème à sa diversité fonctionnelle (Yachi and Loreau, 1999). Cette hypothèse avance le fait que la diversité fonctionnelle

assure l'intégrité d'un écosystème contre les variations temporelles ou spatiales d'une variable biotique ou abiotique de l'environnement (Loreau et al., 2001, 2003a). Appliquer cette hypothèse à la pathologie végétale revient à observer que toutes les plantes sont sensibles à au moins un pathogène, mais que toutes les plantes ne sont pas sensibles à tous les pathogènes (Pautasso et al., 2005) : une population hôte diversifiée assure mieux son intégrité face au développement des maladies.

Quels mécanismes procurent une assurance aux populations hôtes diversifiées ? Au cours de ce travail de thèse, nous avons identifié trois mécanismes majeurs. Premièrement, l'effet de dilution : l'augmentation d'hôtes sub-optimaux pour le pathogène dans la population hôte dilue l'effet de l'hôte principal en augmentant la proportion de propagules du pathogène qui seront retenues par des hôtes incompétents. Deuxièmement, la compétition entre plusieurs souches d'un pathogène adaptées de façon différentielle à l'hôte peut diminuer la prévalence de la maladie au sein de la population hôte. Troisièmement, la structure de la population hôte influence l'évolution de la population pathogène vers des souches plus ou moins agressives et spécialisées. Dans le premier chapitre de ce document nous présentons plus en détail les bases théoriques de ces mécanismes. Nous expliquons aussi comment les différents chapitres (de 2 à 6) peuvent être vus au travers de ce filtre.

Ce travail de thèse s'appuie à la fois sur l'analyse de données récoltées en condition de production et sur des approches théoriques afin d'étudier les différents mécanismes sous-jacents à la définition de stratégies de gestion durable du risque épidémique dans nos paysages agricoles. Dans le chapitre 2, nous analysons trois jeux de données récoltés à l'échelle de la France dans le but d'étudier le lien entre la composition du paysage en blé et sa sensibilité à la rouille brune. Dans ce chapitre nous démontrons le lien entre la composition du paysage en blé, la structure des populations de rouille et les niveaux de maladie observés au champ. L'étude de processus écologiques complexes à de grandes échelles de temps et d'espace rend la collecte de données difficile et les modèles de simulation peuvent alors jouer un rôle important autant dans la formalisation et l'inférence théorique que dans la prédiction ou la structuration d'échantillonnages futurs (Peck, 2004, 2008). Les chapitres 3, 4 et 5 sont tous trois basés sur un modèle de simulation de la dynamique épidémique à l'échelle d'un paysage agricole. Le premier de ces chapitres décrit un cadre de modélisation commun liant une description détaillée du paysage à des modèles de type métapopulation. Dans les deux autres chapitres nous étudions respectivement l'influence de la structure du paysage sur la dynamique épidémique et sur la dynamique de la composition génétique de la population pathogène. Enfin, le chapitre 6 se base sur un modèle de dynamique adaptative pour étudier l'évolution à long terme d'une population dans un milieu spatialement hétérogène.

Part I

Diversity, agriculture and epidemics

Chapter 1

Host diversity and epidemic risk

Contents

1.1 Introduction

Agriculture intensification has improved food production impressively in the past 50 years (Food and Agriculture Organization, 2009) but it came with an increasing impact on the environment (Vitousek et al., 1997; Tilman et al., 2002). In particular, modern agriculture has implied the simplification of the environment structure over vast areas, 'replacing nature's diversity with a small number of cultivated plants and domesticated animals' (Altieri, 1999). As a consequence, agro-ecosystems are particularly susceptible to the epidemic risk (Pautasso et al., 2005; Stukenbrock and McDonald, 2008; Meehan et al., 2011). The dominant approach for pest control, which is also the simplest, focuses on the use of chemical treatments. However, the reliance on pesticides is challenged both by society (Grenelle, 2008) and part of the agricultural sector. Indeed, such an approach imposes strong pressure on biotic communities of cultivated areas, adjacent spaces and natural resources. In addition, it implies important costs for farmers. As an example, almost €70 per wheat hectare are spent for fungicides in France, which represents €300 millions per year (Arvalis). This course change echoes an increasing hardening of the crop protection legislative framework (see figure 1.1 for an illustration of the European legislative framework for crop protection).

Alternative approaches must be developed in order to support the transition towards an ecologically intensive agriculture, without significantly increasing the risk of pest pressure. Such approaches must mobilise a combination of effective agronomic levers (Tilman, 1999; Meynard et al., 2002). Among these, the increase of crop genetic diversity is a promising way for reducing the risk of occurrence and development of diseases in crops. Diversification strategies have already been studied at the within field and crop cycle scales. However, the replacement of a pesticide by a genetic solution requires to manage the genetic resources carefully and collectively. Indeed, many examples show a resistance breakdown when resistances are widely used (Wolfe and Schwarzbach, 1978; Samborski, 1985; Bayles et al., 2000): like any finite resource, variety resistances are subject to the 'tragedy of the commons' (Hardin, 1968), a non-rational use of the resource leads to its extinction. It is therefore necessary to draw new technical and organisational conditions to sustainably manage this resource in a productive system because its use at the territory scale requires significant changes both in farming systems and in the organisation of the involved actors (Butault et al., 2010). Previous studies have attempted to produce technical advice for farmers that includes rotations in time and space of different resistance genes, pyramiding strategies or the use of multilines or mixtures of cultivars (for example Gladders et al., 2006). Nevertheless existing advice remains too global. Explicit landscape approaches should be better incorporated in epidemiological studies in order to improve our understanding and prediction of disease risk (Ostfeld et al., 2005).

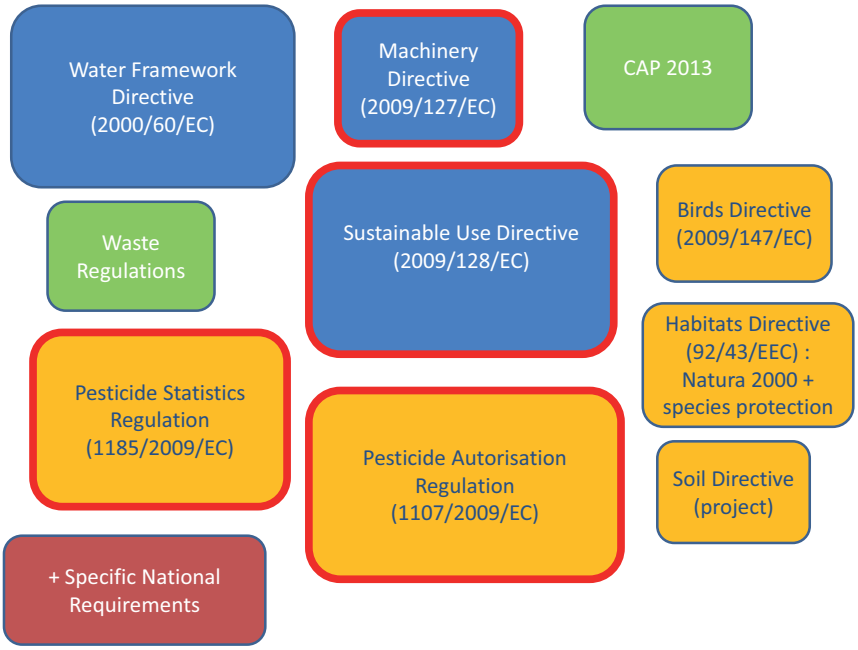


Figure 1.1: The crop protection regulation framework has recently evolved with the adoption by the European Parliament in 2009 of four correlated texts. Regulation 1107/2009 sets the principles of pesticide registration across Europe and replaces Directive 91/414/EC which was in force since 1993. In addition to the principles of the risk assessment of each molecule and formulation, which already existed, the new regulation sets the basis of a common evaluation within European member states and encourages the substitution of the most dangerous molecules. Directive 2009/128/EC concerns the sustainable use of pesticides across Europe. It concerns several aspects of the European agriculture in the aim of reducing the risks of pesticide use : definition of national plans for risk reduction (in France : Ecophyto 2018), improvement of the efficacy of the application machinery, training of farmers and retailers, information of the general public, development of alternatives to chemicals. Directive 2009/127 on application machinery modifies directive 2006/42 which already existed, and was written in order to harmonise the machinery requirements among the European Member States. This measure contributes to obtain the same application conditions in every country, and is consistent with the idea of super-national risk assessment as required by regulation 1107/2009 on pesticide authorisation. Regulation 1185/2009 completes the new pesticide framework by forcing each Member State to produce consistent, harmonized and quantified data on pesticide sales and utilisations every year. This kind of data is essential to assess the European pesticide policies and build relevant indicators.

In addition to these four main texts, several other European laws have an impact on pesticide use. First of all, the Water Framework Directive 2000/60/EC which requires the Member States to obtain a good chemical and biological state for all the surface water and groundwater bodies by 2015. This directive includes a list of chemicals whose emissions in water have to be reduced, and some of these chemicals are pesticides. The Common Agricultural Policy, which is to evolve in 2013, has a strong impact on pesticide use, by fixing the criteria of allocation grant to European farmers. Other texts concerning wildlife protection (Birds and Habitats Directive), or environment protection (Soil Directive project) have an indirect but not negligible impact on pesticide use. Finally, pesticides and pesticide containers are considered as dangerous waste and are therefore concerned by the legislation on waste. National additional requirements exist and differ from a Member State to another.

The study of interactions between spatial patterns of a landscape and ecological processes is at the heart of landscape ecology (Turner, 2005). Association of landscape ecology and epidemiology dates back from the thirties and was logically named landscape epidemiology (Ostfeld et al., 2005). The landscape epidemiology approach is relatively well developed for human and animal epidemiology as well as for pest control in agriculture (Langlois et al., 2001; Allan et al., 2003; Bianchi et al., 2006; Tildesley et al., 2010). In plant pathology, the idea of designing variety management strategies at the scale of cultivated landscapes has been around for a long time (Zadoks and Kampmeijer, 1977; Mundt and Brophy, 1988) but the potential of landscape epidemiology for exploring new strategies of host resistance management still remains largely underexploited by plant pathologists (Plantegenest et al., 2007). In this PhD thesis we focus on landscape-scale processes and on the potential role of functional diversity in crops to better control plant-diseases.

In its broader definition, *diversity* refers to the number of entities (from genotypes to ecosystems), their abundance and distribution (Haccou et al., 2005). In order to understand the link between diversity and a particular ecosystem property we must consider diversity with respect to the functional traits susceptible to influence that ecosystem property. Thus, *functional diversity* will refer in our case to hosts that act differently with respect to a pathogen or to a particular pathogen strain. Pautasso et al. (2005) classified ecosystems according to their susceptibility and diversity (figure 1.2). It identified four cases: (i) 'lucky monoculture' (low diversity and low susceptibility), (ii) susceptibility without diversity, (iii) susceptibility with diversity, and (iv) ecosystems governed by the insurance hypothesis (high diversity and low susceptibility). In this PhD thesis we argue that agro-ecosystems must evolve from (ii) to (iv). The insurance hypothesis links ecosystem functioning and species diversity (Yachi and Loreau, 1999). It suggests that functional diversity insures ecosystem functioning against temporal and/or spatial fluctuations of a biotic or abiotic variable in the environment (Loreau et al., 2001, 2003a). Applying this hypothesis to plant disease results in observing that 'most plants are susceptible to more than one pathogen, but not all plants are susceptible to all pathogens' (Pautasso et al., 2005): a diversified host population better insures its integrity against disease development.

What are the mechanisms that could provide an insurance account for functional diversity of host populations? In this work we identified three main mechanisms. First, the dilution effect insures that the increase of non-efficient hosts will dilute the effect of principal host. Indeed, non-efficient hosts increase the proportion of pathogen propagules coming from incompetent hosts. Second, competition between differentially adapted pathogen strains could decrease infection prevalence in the host population. Third, host population structure influences pathogen evolution towards more effective specialised strains. In the following of this chapter we will describe in more details these three mechanisms from a theoretical point of view. We will also underline how the different chapters (from 2 to 6) can be classified

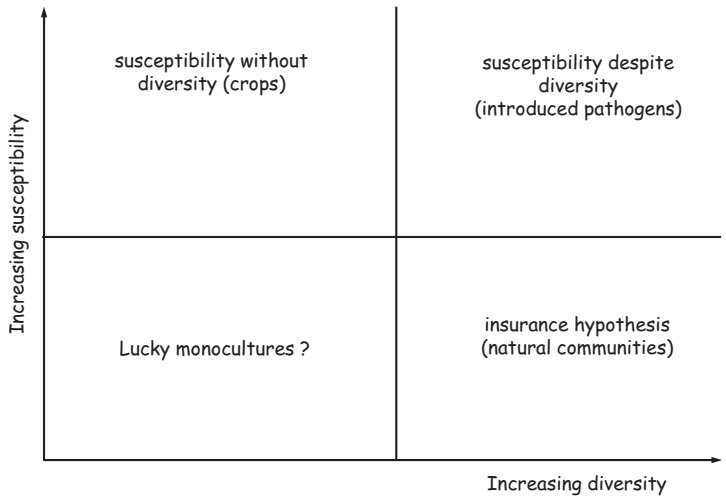


Figure 1.2: Classification of ecosystems according to their susceptibility and diversity. From Pautasso et al. (2005)

according to these mechanisms.

This PhD thesis relies on both the analysis of real data and on theoretical approaches in order to investigate several mechanisms underlying the design of sustainable strategies for disease control in agricultural landscapes. We first begin (chapter 2) by analysing three large scale datasets in order to track a potential relationship between the varietal wheat landscape and their susceptibility to leaf rust. Chapter 2 demonstrates the relationship between wheat landscape composition, leaf rust population genetic structure and observed disease level scored in field conditions. The study of complex ecological processes at large scales of time and space makes data collection difficult. That is why theoretical approaches such as simulation or mathematical models can play an important role for theoretical formalisation and inference as well as in the structuring of future sampling (Peck, 2004, 2008). Chapters 3, 4 and 5 are all based on a simulation model for pathogen population dynamics on an agricultural landscape. Chapter 3 describes a general modelling framework that links a detailed characterisation of the landscape to classical metapopulation models. In chapters 4 and 5, we study respectively the role of landscape structure on the pathogen population dynamics and on the genetic composition of the pathogen population. Finally, in chapter 6 we used an adaptive dynamics approach in order to study the long term evolution of pathogen populations in spatially heterogeneous environments.

1.2 Dilution effect

The increase in species richness can reduce disease risk through a lot of mechanisms (Mundt, 2002; Keesing et al., 2006). However, the term *dilution effect* will be use restrictively in this work and apply only to 'situations in which a pathogen can be acquired by a variety of hosts but is transmitted efficiently by only one or a few' (Keesing et al., 2006, box 2). Lively (2010) proposed a simple example in order to explain how the dilution effect acts in a heterogeneous host population. Consider a host population composed of several genotypes. Suppose that each strain of the pathogen population can develop on one host genotype only, *i.e.* interaction between host and pathogen is governed by the 'gene-for-gene' model (Flor, 1971). Without considering spatial structure, the spread of infection requires that (Lively, 2010):

$$g_i N > \frac{1}{1 - e^{-\lambda}},$$

where g_i is the frequency of host genotype i, N is the total number of individuals in the host population and λ is the mean number of matching spores that contact each host. This simple relationship shows that increasing diversity (*i.e.* decreasing g_i) without increasing host density (N) may reduce the disease risk by increasing the number of alternative hosts which are incompetent reservoirs. As species or genotypes are added to the host community, these alternative hosts dilute infections by increasing the proportion of pathogen propagules that infect incompetent hosts (Ostfeld and Keesing, 2000). The expected result is a lower infection prevalence in the host population.

However this basic framework lacks two important aspects. First, it does not take into account quantitative interactions between plant and pathogen (Pariaud et al., 2009a), which are of increasing importance due to their better sustainability (McDonald, 2010). In such a case, in addition to the binary response of the gene-for-gene model, a continuous response has to be considered to describe the host-pathogen interaction. The components of the variety mixture then appear more or less conducive to the disease and the results could be changed. Second, the non spatial context is very limiting. In fact, when dispersal is limited, host types exchange parasite propagules according to their spatial organisation. Empirical evidence of such an effect is abundant in the phytopathological literature, in particular dealing with variety mixtures, but focus on the intra-field scale (for a review see Mundt, 2002).

In this PhD thesis, the dilution effect was studied by using a spatially explicit model for pathogen population dynamics in an agricultural landscape. We investigated how the landscape composition and the spatial deployment of varieties affect the disease spread. We considered several conductivities for the alternative host and several abilities of the pathogen to disperse. The modelling framework is described in chapter 3. Chapter 3 also provides an analysis of thresholds for pathogen invasion in heterogeneous landscapes. The epidemic

spread in a heterogeneous host population is studied more thoroughly in chapter 4.

1.3 Competition and coexistence between pathogen strains

In agricultural pathosystems, pathogen strains adapt differentially to host varieties regarding quantitative traits (rewied by Pariaud et al., 2009a). As a consequence, the composition of the pathogen population that lives on a given host determines the amount of disease that will be observed. The landscape varietal composition shapes the genetic structure of pathogen populations encountered on a particular host variety, which explains the relationship between the composition of the host population and its susceptibility to disease. How pathogen strains compete and coexist in an heterogeneous host population is therefore of prime interest for disease control.

In community ecology, the causes of species diversity maintenance is a crucial question. This has led to the current debate between niche theory and neutral perspectives (Whitfield, 2002; Gewin, 2006). While the first assumes partitioning of resource use, the second is based on stochastic processes only. More recently Clark et al. (2007) propose a third explanation in order to resolve the *biodiversity paradox*. It is based on the roles of deterministic biological process *vs* stochastic approximations in dynamic population models. We briefly review these theories before explaining how we address the question of pathogen strains coexistence in this work. The aim is not to be exhaustive but to provide some theoretical basis on the mechanisms that enhance coexistence between species.

1.3.1 Ecological niche

Hutchinson (1957) defined the fundamental niche as 'a n-dimensional hypervolume [...], every point in which corresponds to a state of the environment which could permit the species [...] to exist indefinitely'. From this definition, Pulliam (2000) identified three contexts that diverge from this idealised point of view (Figure 1.3). Firstly, competition with a better competitor could reduce the fundamental niche to the realised niche (Figure 1.3 B). This is due to the fact that when several species are present in a given environment, they are not only affected by the environment but they also change it and interact directly with one another. Secondly and thirdly, adding dispersal to the niche concept changes crucially the relationship between fundamental niche and species distribution: source-sink dynamics will extend the area of occurrence (Figure 1.3 C) whereas dispersal limitation and frequent local extinctions may reduce the species occurrence inside of its fundamental niche area (Figure 1.3 D).

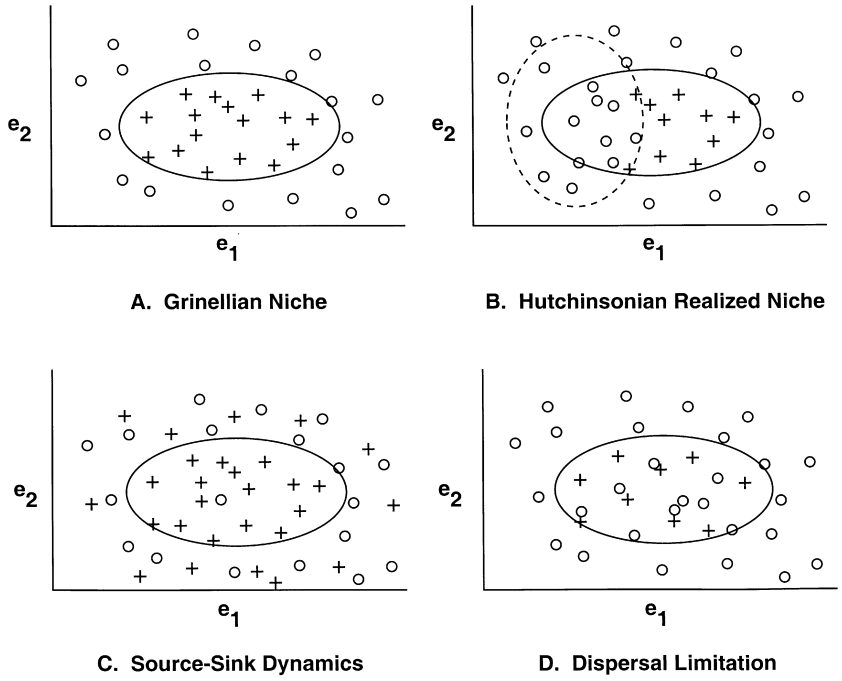


Figure 1 Four views of the relationship between niche and species distribution. In each diagram, the solid oval refers to the fundamental niche or the combination of environmental factors (e_1 and e_2) for which the species has a finite rate of increase (λ) greater than or equal to 1.0. The "pluses" indicate the presence of the species in a patch of habitat characterized by particular values of e_1 and e_2, and the "zeroes" similarly indicate the absence of the species in a patch of habitat. According to the Grinnellian niche concept (A), a species occurs everywhere that conditions are suitable and nowhere else. Hutchinson's realized niche concept (B) postulates that a species will be absent for those portions of the niche space that are utilized by a dominant competitor. According to source–sink theory (C), a species may commonly occur in sink habitat where λ is less than 1.0. Metapopulation dynamics and dispersal limitation (D) posit that species are frequently absent from suitable habitat because of frequent local extinctions and the time required to recolonize suitable patches.

Figure 1.3: Four views of the relationship between niche and species distribution. From Pulliam (2000)

The basic principles of coexistence between species exploiting the same resource are classically exposed using the Lokta-Volterra model. In terms of absolute competition coefficients, the two-species Lotka-Volterra competition equations can be written as follows (Chesson, 2000):

$$\frac{1}{N_i} \cdot \frac{dN_i}{dt} = r_i\Big(1 - \alpha_{ii}N_i - \alpha_{ij}N_j\Big), \quad i \in \{1,2\},\ i \in \{1,2\} \text{ and } j \neq i \tag{1.1}$$

where r_i is the *per capita* growth rates and N_i is the abundance of species i. α_{ii} and α_{ij} are, respectively, absolute intraspecific and interspecific competition coefficients. Assuming equation (1.1), species 1 and 2 coexist if they cannot competitively exclude one another, *i.e.* if $\alpha_{11} > \alpha_{21}$ and $\alpha_{22} > \alpha_{12}$. That means that species coexist if they limit themselves more than they limit the other species (Chesson, 2000; Adler et al., 2007).

Usually, coexistence among species is characterised by the 'long-term low-density growth rate', $\bar{r_i}$, which measures the capacity of species i to recover from low density. Coexistence between several species requires that each species is able to increase from low density, *i.e.* that $\bar{r_i} > 0$. Immersing the basic Lokta-Volterra model (equation (1.1)) into the niche framework can be done by incorporating a measure of resource-use overlap among species as proposed by Chesson (1990). In such a model and under particular assumptions, coexistence in multi-species communities depends on two additive terms (Chesson, 2000):

$$\bar{r_i} \approx b_i(k_i - \bar{k}) + \frac{b_i(1-\rho)D}{n-1}. \tag{1.2}$$

Parameters are the number of species in the system (n), the measures of fitness of individual species (k_i), the average fitness of the competitors of species i ($\bar{k}$), the rates at which the *per capita* growth rates decline as resources decline in abundance (b_i), the niche overlap (ρ) and a positive constant (D). The first term in equation (1.2) compares the fitness of species i to the mean fitness of the community: it is positive for the best competitors and negative for the poorest competitors. The second term renders the niche differentiation. If there is no niche partitioning, *i.e.* overlap is total ($\rho = 1$), then it is equal to 0. It is called by Chesson (2000) a 'stabilising' term since it buffer differences in fitness. In fact, this second term is always positive and thus it can counterbalance negative fitness differences. Without this second term, the best competitor would exclude all the other species. As a consequence, stable coexistence is an interplay between 'equalising mechanisms', which tend to limit the differences between fitness, and 'stabilising mechanisms' that make intraspecific effects more negative than interspecific effects (Adler et al., 2007, figure 2 and box 1). It is important to note that in the niche framework, there is no stable coexistence possible without the stabilising term. In other words, stable coexistence, even if mean fitness are very close, requires niche differences, which means that 'species must differ in terms of how they respond to and/or affect the environment, including resources, they share with other species' (Clark, 2009, box 1).

How does spatial structure of the resource act on the equalising and stabilising terms? This question is obviously dependent on the ability of the organism to disperse in the environment. Both in isolated or well mixed communities the resource spatial structure has no effect on the results presented before. On the contrary, at intermediate dispersal rates, local species diversity is enhanced. An intermediate dispersal rate may prevent local competitive exclusion by the local dominant species while being low enough to prevent exclusion by the global dominant one (Loreau et al., 2003a). Several classifications of coexistence mechanisms are possible (Melbourne et al., 2007). Chesson (2000) identified four such mechanisms in spatial environments: fluctuation-independent mechanisms, spatial relative non-linearity, spatial storage effect and spatial covariance of local population growth with local population density (Chesson, 2000; Gravel et al., 2011).

Fluctuation-independent mechanisms group all mechanisms that are not due to variation of the resource (*e.g.*, mean fitness differences or classic niche partitioning). Spatial relative non-linearity measures how far from a linear response the response of a species to resource variation is. The more variable the environment, the more disadvantaged the species with high non-linearity response is. Thus, if the species with the highest non-linearity is also the species that, in average, have the best fitness, non-linearity will tend to decrease the importance of differences between fitness (Gravel et al., 2011). That is why this mechanism could be split into an equalising term - that reflects the relative fitness taking into account of the environment fluctuations in space - and a possible stabilising term due to the non-linearity amplitude between species. The spatial storage effect is a stabilising mechanism. In this case, a species takes more advantage of spatial locations where its fitness is better than the other species than spatial locations where the environment is unfavourable for it. Fitness-density covariance is a purely spatial mechanism. As an example, competitors produce more propagules in favourable conditions and, due to a limited dispersal, propagules are locally retained. As a consequence competitors suffer proportionately stronger competition under favourable environments, which leads to more negative intraspecific effects than interspecific effects (Gravel et al., 2011).

1.3.2 Neutral theory

On the contrary to niche theory, the neutral theory of biodiversity (Hubbell, 2001; Bell, 2001) assumed that species are ecologically equivalent (Nee, 2005; Kopp, 2010), that is, 'all individuals had identical demographic characteristics, regardless of the species they belong to' (Bell, 2001). It then predicts species abundance as a dynamic equilibrium involving speciation, dispersal and extinction (Hubbell, 2003). In such a context, abundances of all species fluctuate at random because only demographic stochasticity operates. This is called 'ecological drift' by Hubbell (2003) and is analogous to the well known 'genetic

drift' (Nee, 2005). A historical review that links the Hubbell's neutral theory to the concept of neutrality in population genetics (Kimura, 1968) is proposed by Chave (2004) (see also Nee, 2005; Gewin, 2006).

The basic neutral model refers to the island biogeography model whereby a local community is linked via dispersal to an external pool. It involves only five parameters: the immigration probability, the birth and death probabilities, the number of individuals in the community and the total number of species in the external pool (Bell, 2001). Neutral theory leads to the definition of a biodiversity number which controls several characteristics of biodiversity patterns (Hubbell, 2003) and speciation (de Aguiar et al., 2009; Kopp, 2010).

Even if niche differentiation is not questioned, Leibold and McPeek (2006) propose three reasons for considering neutral perspectives. The first is that neutral models could be seen as a null hypothesis. Second, neutral theory has proved to produce realistic patterns of biodiversity particularly in highly diversified communities and for rare species (Volkov et al., 2003; Chave, 2004; Halley and Iwasa, 2011). Hence, invoking the parsimony principle neutral models could be retained because they involve less parameters. Finally, neutral models may enlarge the niche view to complex dynamics of equivalent species in interaction with the rest of the community.

1.3.3 Between zero and complete uncertainty, a 'Laplacian' vision of the biodiversity paradox

According to Clark et al. (2007), the biodiversity paradox comes with the fact that stable coexistence in models demands precise conditions on resource use (see the niche section) that are difficult to identify in the real world. The neutral theory of biodiversity is an alternative view that needs no functional differentiation between species to explain biodiversity patterns. The link between niche and neutral theories was made in several ways. Adler et al. (2007) put the neutral theory within the Chesson (2000) framework (see also the niche section), which leads to see niche and neutral theory as the extrema of a continuum. Niche constitutes the case when differences between fitness are high and coexistence is performed due to strong stabilising mechanisms whereas neutrality refers 'to weak stabilisation operating on species of similar fitness'. The idea of a continuum is also described by Gewin (2006). Chave (2004) mentioned that both theories are complementary. In fact niche theory focuses on small numbers of species with relative simple interactions while neutral theory focuses on highly diversified communities when the role of stochasticity is unavoidable.

Clark et al. (2007) link both theories via a continuum but 'the continuum in models is one of knowledge, not cause. The two types of models in the debate are special cases, low-dimensional trade-offs and neutrality representing zero uncertainty and complete uncertainty,

respectively'. This view is discussed by Clark (2009). He argues that stochasticity is misinterpreted, which leads to consider stochasticity as 'an actual force'. Gravel et al. (2011, box 1) summarise Clark's view well. Consider the following population growth model

$$\frac{dN_i(x,t)}{dt} = f(\text{covariates}, \text{parameters}) + \text{error}, \tag{1.3}$$

where the first term is deterministic and accounts for our knowledge on the processes that act on the population dynamics. In Clark (2009), the second term (error) represents only our uncertainty due to modelling. While the basic niche view assumed that all processes are known (*i.e.* their is no error term), the neutral theory assumed that there is the error term only. With this mixed approach, patterns of diversity are explained always by niche differences among species but, due to their high dimensionality, these differences are difficult to identify (Clark et al., 2010; Ricklefs, 2011).

This conceptual framework seems particularly interesting and makes niche differences a central mechanism for explaining coexistence between species. Moreover, it recalls the extreme high-dimensionality of natural systems. Nevertheless, I see some ambiguities in the Clark (2009) article. In fact, pushing his reasoning to the extreme could lead to another debate on the existence of stochasticity in Nature. Fisher's works (Fisher Box, 1978) show that the second term in equation (1.3) might be split into two terms: a modelling error and an incompressible 'natural' error. However, Clark (2009) does not allow for the latter. In fact he wrote

'First, there is no evidence for stochasticity in nature at observable scales. Stochasticity is an attribute of models. When an individual gives birth or dies 'at random', that event results from real processes.'

And latter,

'The fact that there might be a practical limit to how much we know does not mean that there exists some residual inherent stochasticity.'

With these sentences it seems that Clark does not consider stochasticity as a possible force in the real world, which reminds the deterministic vision of Laplace. Potentially, it might lead to a misunderstanding of community dynamics especially for rare species in which demographic stochasticity could be extremely influential.

1.3.4 Retained approaches

The study of coexistence of several pathogen strains and the role of the spatial organisation of varieties on this coexistence were studied in this PhD thesis using two

approaches. First, the relative roles of environment composition and pathogen strain fitness on the competition of several pathogen strains was studied in the specific example of wheat leaf rust. This work was based on a 10-year survey of wheat leaf rust populations at the scale of France. We constructed a hierarchical model in a Bayesian framework in order to infer the relationship between the wheat varietal landscape composition, the composition of leaf rust populations and observed symptoms on the main varieties. This work in presented in chapter 2.

Based on these results we explored further how habitat spatial structure may influence the coexistence between pathogen strains, using the simulation framework presented in chapter 3. We studied the composition of a pathogen population that spreads on an agricultural landscape composed of two varieties. Three pathogen strains were defined: two symmetric specialists and one generalist for which the landscape is homogeneous. In chapter 5, we investigated in which condition a new strain may establish a population into a resident population and, in which conditions coexistence between strains is stable. According to the theoretical framework depicted before, the model assumed low dimensional niche differences and stochasticity in the population dynamics. We concentrated our study on spatial mechanisms that promote coexistence.

1.4 Long term evolution of pathogens

Strong natural selection exerted by anthropogenic disturbances accelerates evolutionary changes in pathogens (Palumbi, 2001). Integration of evolutionary principles into epidemiological models was thus unavoidable in order to better understand the rapid evolution of pathogens in response to changes in their host ecology (Galvani, 2003). Among global changes that affect pathogen evolution, the increase in connectivity between hosts is particularly influential (Boots and Sasaki, 2000). It is assumed that, together with the increase of host homogeneity (Galvani, 2003), the high conductivity of agro-ecosystems has led to the selection of highly specialised and virulent[1] pathogens (Stukenbrock and McDonald, 2008).

Usually, description of host population dynamics is based on a set of ordinary differential equations that describes the three states of the host population as healthy, infected, and recovered. From this model, non-spatial studies on virulence evolution compute the basic reproductive number (Hethcote, 2000), R_0, that determines the ability of the pathogen to invade the host population. The evolutionarily stable pathogen strain is defined by the set of parameters that maximise R_0. Pathogen-host models generally involve two parameters

[1] In classical epidemiological studies, virulence is defined as the induced mortality of hosts due to pathogen infection.

that describe the ability of the pathogen population to grow: pathogen's transmission rate and induced host mortality (virulence). Simple non-spatial models predict that, without explicit trade-off between transmission and virulence, pathogen evolves to become avirulent and have infinitely fast transmission (but see Alizon and van Baalen, 2005).

An alternative method to R_0 maximisation is adaptive dynamics (Geritz et al., 1998). Due to several limits of R_0 maximisation, spatial studies use an adaptive dynamics approach (Messinger and Ostling, 2009). Adaptive dynamics was introduced by Hofbauer and Sigmund (1990) and Nowak and Sigmund (1990). It aims to study the evolutionary outputs of a resident population that is exposed to repeated invasion by mutants. It is based on two fundamental ideas: the population remains monomorphic - because demographic equilibrium is reached before a new mutant arrives - and initial growth rate (invasion fitness) of the mutant determine the output of the competition between mutant and resident. More recently, Day and Proulx (2004) proposed to use techniques from theoretical population genetics in order to study the evolution of virulence. This method has several advantages with respect to adaptive dynamics, in particular the epidemiological dynamics equilibrium assumption can be relaxed (Day and Gandon, 2007).

In this PhD work, we used an adaptive dynamics approach in order to study the evolution of specialisation in spatially heterogeneous landscapes (chapter 6). The role of habitat spatial structure was addressed with different perspectives in spatially explicit adaptive dynamics model. Some studies suppose that the habitat changes gradually with space for its altitude or temperature (Doebeli and Dieckmann, 2003; Champagnat and Méléard, 2007) and others introduce explicit patches in the environment either in a continuous environment (Débarre and Gandon, 2010) or in a metapopulation framework (Meszéna et al., 1997; Parvinen and Egas, 2004). In chapter 6, we develop a metapopulation model in order to investigate how the landscape structure influences the dynamics of adaptation using both analytical and simulation studies.

1.5 Use of functional diversity in agriculture

The diversity for resistance genes has mainly been considered at two different scales so far: plants and fields. At the scale of the plant, pyramiding strategies consist in combining several resistance genes in the same variety. It is then supposed that the gene combination is more difficult for the pathogen to overcome than an isolated resistance gene (but see Mundt, 1990). Quantitative resistance is also sometimes considered as a resistance diversification strategy since it relies on combinations of QTLs. Pyramiding major resistance genes or combining QTLs for quantitative resistance require a major effort in plant breeding.

At the scale of the field, two approaches exist: multilines and cultivar mixtures. The idea

of the multiline dates back from the end of the fifties. Here, on the contrary to pyramiding, resistance genes are distributed in different lineages that are grown as a mixture on the same plot. The interest is to obtain a composite variety with homogeneous individuals for agronomic characters but diversified for disease resistance factors. This principle has been applied in practise but in a limited way (Finckh et al., 2000). Indeed, selection and maintenance over time of multiline cultivars represent a considerable effort.

Cultivar mixtures, as for multilines, are based on the idea of increasing functional diversity for resistance to diseases. The principle is simply to grow in a mixture several cultivars with different resistance genes. However, these cultivars must be compatible regarding their agronomic behaviour and complementary according to their resistance to disease (which implies a good knowledge of the pathogen population). Cultivar mixtures generally present excellent yield stability with respect to their components grown alone, due to compensation effects between the different cultivars in the case of biotic or abiotic stress (Finckh et al., 2000). The efficiency of cultivar mixtures with respect to disease control depends on several epidemiological parameters (Mundt, 2002). In particular, it is expected to decline for higher autoinfection / alloinfection ratio, which is determined by both pathogen and host physical and epidemiological characteristics.

In this PhD thesis, we consider the scale of agricultural landscapes. It was observed empirically that the varieties used over large areas tend to become highly susceptible and, conversely, that highly susceptible varieties could become more resistant when their frequency in the landscape decreases. Inspired by such observations we dissect the mechanisms involved in the diversity / susceptibility relationship at the regional scale with the aim of designing sustainable strategies for disease control in agricultural landscapes.

Chapter 2

Influence of cultivated landscape composition on variety resistance: an assessment based on the wheat leaf rust epidemics

This chapter is based on an article published in New Phytologist by Julien Papaïx, Henriette Goyeau, Philippe Du Cheyron, Hervé Monod and Christian Lannou

Contents

Foreword to chapter 2

Any modelling study have to rely with some observed reality. Our work begins with observations on the wheat leaf rust pathosystem (*Puccinia triticina*/*Triticum aestivum*). The variety *Soissons* was highly popular in the 90's and cover up to 40% of the wheat surfaces at the scale of France. Due to its increasing susceptibility (as well as other factors) *Soissons* became less and less popular and only represented 3% of wheat acreage in 2008. These last years, technical institutes and agricultural cooperative have noted that *Soissons* was in pass to appear resistant to leaf rust in field conditions, up to be envisaged it in rotations for organic agriculture.

Based on this observation, we investigated further the dynamics of leaf rust at the scale of France on the mainly grown wheat varieties. For this purpose, we jointly analysed three large data-sets describing the wheat leaf rust pathosystem: (i) the frequencies of the most frequently grown varieties in France; (ii) a ten-year population survey of *P. triticina* on each of these varieties; and (iii) the assessment of the resistance level of these varieties in multi-local trials during the same period of time. In order to link and analyse these data-sets, we developed a hierarchical model within a Bayesian framework.

We show that among all compatible pathotypes, some were preferentially associated with a variety, that the frequency of a pathotype on a variety was affected by the landscape varietal composition, and that the observed resistance level of a variety was linked to the frequency of the most aggressive pathotypes among all compatible pathotypes. As a consequence, the landscape varietal composition was found to influence the resistance level (as measured in the field) of the most frequently grown wheat varieties by altering the structure of pathogen populations. In addition we illustrated that the quantitative aspects of the host-pathogen relationship have to be considered in addition to the major resistance/virulence factors in landscape epidemiology approaches.

2.1 Introduction

In modern agriculture, structure simplification and genetic uniformity of cultivated landscapes facilitate the spread of epidemics and the genetic evolution of pathogens towards a higher virulence (Oerke and Dehne, 2004; Stukenbrock and McDonald, 2008; Margosian et al., 2009). Nevertheless, both experimental and theoretical approaches support the idea that increasing functional diversity based on resistance factors would make agricultural systems less susceptible to diseases (Altieri, 1999; Zhu et al., 2000; Mundt, 2002; Bianchi et al., 2006; Garrett et al., 2009). More generally, a relationship between functional diversity and susceptibility to diseases had been demonstrated in ecological systems (Pautasso et al., 2005; Keesing et al., 2006) and this relationship can be considered at different spatial scales, one of them being the landscape (Gilligan, 2008). Studies in animal and human epidemiology (Keeling, 1999; Tildesley et al., 2010) have shown that landscape structure and connectivity may greatly influence pathogen invasion rates: the simplification of agricultural landscape structure and composition, along with the decline of non-crop habitat, has led to a decrease in natural pest control (Bianchi et al., 2006).

In plant pathology, the idea of designing variety management strategies at the scale of cultivated landscapes has been around for a long time (Zadoks and Kampmeijer, 1977) and was tested more than 20 years ago by Mundt and Brophy (1988) with a simulation model. More recently, several authors have explored the potential of large-scale approaches to optimise the deployment of host resistance. With an approach based on the metapopulation theory (Hanski, 1998), Parnell et al. (2006) suggest that the spread of pathogen strains that are resistant to fungicides would be more effectively controlled with a landscape-scale approach; Margosian et al. (2009) assessed the connectivity, with regard to pathogen transmission, of the four main crops in the USA; Skelsey et al. (2009) developed a spatio-temporal model of the potato late blight pathosystem that will make it possible to evaluate spatial deployment of host resistance in large growing areas. Nevertheless, the potential of landscape epidemiology for exploring new strategies of host resistance management still remains largely underexploited by plant pathologists (Plantegenest et al., 2007). A likely reason for this is the difficulty to obtain and analyse experimental data at a large geographic scale. A main objective of this study is therefore to contribute to filling this gap.

The development of simulation models to design variety deployment strategies in agricultural landscapes requires the identification of the effects that take place at large scales on host and pathogen populations. Most available datasets on pathogen population structure and host resistance are based on qualitative host-pathogen interactions, as described by the gene-for-gene model (Flor, 1971). It is well known that interactions between major resistance genes and avirulence genes shape pathogen population structure at large scales in cropping systems (Wolfe and Schwarzbach, 1978; Hovmøller et al., 1993; Rouxel et al., 2003; Goyeau

et al., 2006; Barrès et al., 2008) and that, reciprocally, invasions of new virulent strains render ineffective the corresponding resistance genes in the crops. Nevertheless, although the qualitative host-pathogen interactions are necessary to describe pathogen populations and to explain the observed resistance level of host varieties, they are not sufficient. Quantitative interactions determined by pathogen aggressiveness (see Pariaud et al., 2009a, for a review) and host quantitative resistance (Brun et al., 2010; Marcel et al., 2008) can play a major role as well in shaping pathogen populations (Thrall and Burdon, 2003; Miller et al., 1998; Pariaud et al., 2009b). In such a case, in addition to the binary response of the gene-for-gene model, a continuous response has to be considered to describe the host-pathogen interaction.

In this paper, we test the hypothesis that landscape composition (in terms of host variety frequencies) has an impact on the changes in the observed resistance level of the main varieties grown by the farmers, at the French national scale. For this, we consider both qualitative and quantitative information on the interactions between the pathogen and its host. Data analysis focuses on a wheat (*Triticum aestivum*) foliar disease, leaf rust, caused by *Puccinia triticina*, a basidiomycete fungus (Bolton et al., 2008). This pathosystem has been studied in depth. As a result, we had access to three datasets to carry out our analysis, related to: (i) the frequencies of the most frequently grown varieties in France; (ii) a ten-year population survey of *P. triticina* on each of these varieties; and (iii) the assessment of the resistance level of these varieties in multi-local trials during the same period of time. In order to link and analyse these datasets, we developed a hierarchical model within a Bayesian framework.

2.2 Material and methods

2.2.1 Data description

2.2.1.1 Wheat varieties

The French institute, *FranceAgriMer*, publishes annual statistics on the most frequently sown wheat varieties in France (ONIGC, 2008). This dataset records the frequency, with respect to the total French wheat acreage, of the ten most frequently sown varieties each year. From 1999 to 2008, 30 varieties were recorded, representing between 53.8% and 76% of the wheat acreage each year.

The INRA Grignon laboratory routinely identifies the major resistance genes present in the cultivated varieties. Based on this information (Goyeau et al., 2006), the recorded varieties were classified into five main groups. The first group contained the varieties bearing the resistance gene $Lr13$, and the second group the varieties with $Lr14a$. Groups 3, 4 and 5 contained varieties with the following combinations of resistance genes: $Lr10 + Lr13$, $Lr13 + Lr37$ and $Lr10 + Lr13 + Lr37$, respectively. Since the frequency of the first group

Table 2.1: Frequencies of the studied wheat varieties from 1999 to 2008 in France.

Variety	Year									
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Apache	NA	NA	0.08	0.17	0.24	0.23	0.23	0.14	0.11	0.11
Caphorn	0	NA	NA	NA	0.03	0.08	0.08	0.13	0.13	0.12
Charger	0.04	0.089	0.11	0.10	0.08	0.07	0.06	0.05	NA	NA
Isengrain	0.084	0.114	0.14	0.14	0.09	0.06	0.07	0.04	0.034	0.03
Orvantis	0	NA	NA	0.04	0.06	0.06	0.06	0.04	0.032	0.032
Soissons	0.153	0.108	0.11	0.09	0.07	0.05	0.05	0.04	0.038	0.03
Trémie	0.113	0.076	0.06	0.04	0.03	0.02	NA	NA	NA	NA
Total frequency	0.39	0.387	0.50	0.58	0.60	0.57	0.55	0.44	0.344	0.322

was low and declining over the considered period, we limited our study to the last four groups.

Among these groups, we focused on the most representative varieties (in terms of frequency). Seven of them were chosen from groups 2, 3, 4 and 5: *Isengrain* and *Soissons* (group 2), *Charger* and *Trémie* (group 3), *Apache* (group 4) and *Caphorn* and *Orvantis* (group 5). These varieties represented between 32.2% and 60% of the French wheat acreage, depending on the year (table 2.1). The other varieties, present at low frequencies, are considered together in the model as a background host population. Because only the first ten varieties are recorded each year in the wheat survey, the frequency of those varieties in the cultivated landscape was not always known over the whole period. *Apache* was recorded from 2001 to 2008, *Caphorn* from 2003 to 2008, *Charger* from 1999 to 2006, *Isengrain* from 1999 to 2008, *Orvantis* from 2002 to 2008, *Soissons* from 1999 to 2008, and *Trémie* from 1999 to 2004.

2.2.1.2 *Puccinia triticina* population

Isolates of *Puccinia triticina* are collected each year from a network of unsprayed nurseries in 64 different locations throughout the country. A pustule is sampled at each site for each variety and the isolate is increased for pathotype determination. A detailed description of the French leaf rust survey is given in Goyeau et al. (2006). A pathotype, or phenotype for qualitative virulence (Gilmour, 1973), is attributed to each isolate collected. The pathotype indicates whether or not the isolate is able to overcome the major resistance genes.

A total of 2521 isolates were sampled on 124 varieties over the period considered (1999–2008). The varieties *Soissons*, *Isengrain*, *Charger* and *Apache* were sampled from 1999 to 2008, *Trémie* from 1999 to 2007, *Orvantis* from 2001 to 2008, and *Caphorn* from 2003 to

Table 2.2: Number of collected isolates on each variety. Data from the French leaf rust survey.

Variety	Year									
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Apache	3	4	19	31	41	35	51	47	71	43
Caphorn	NA	NA	NA	NA	13	23	44	45	64	41
Charger	2	9	1	18	27	17	26	26	24	8
Isengrain	48	50	61	52	43	30	23	13	22	10
Orvantis	NA	NA	18	24	32	23	46	39	48	16
Soissons	53	58	63	47	45	30	46	37	44	31
Trémie	32	36	55	20	11	22	26	25	10	NA
Sample length	166	187	234	229	224	194	306	306	417	238

Table 2.3: Number of collected isolates for each pathotype. Data from the French leaf rust survey.

Pathotype	Year									
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
006*xxx*	0	0	0	3	4	22	50	37	69	23
016206	10	6	12	28	41	18	10	8	0	0
073100	36	76	98	67	51	31	28	31	15	9
077317	6	4	12	14	40	20	24	12	6	0
106314	0	0	0	0	2	14	38	87	109	104

2008 (table 2.2). Over the whole period, 196 different pathotypes were identified, among which only a few had a high frequency (figure 2.1, table 2.3). We focused on the most frequent pathotypes, referred to as 006*xxx*, 016206, 073100, 077317 and 106314 (see Goyeau et al., 2006). The first pathotype (006*xxx*) aggregates three individual pathotypes (006106, 006504 and 006506) with very close phenotypic expressions and identical micro-satellite profiles. The low-frequency pathotypes are not ignored by the model but are grouped together and considered as a background pathogen population.

2.2.1.3 Disease scoring

The French technical institute, *Arvalis Institut du Végétal*, carries out annual trials to evaluate the resistance level of the main wheat varieties under field conditions, in order to produce technical advice for farmers. Trials are distributed over 40 different locations throughout the wheat-growing areas in France. They consist of complete blocks containing

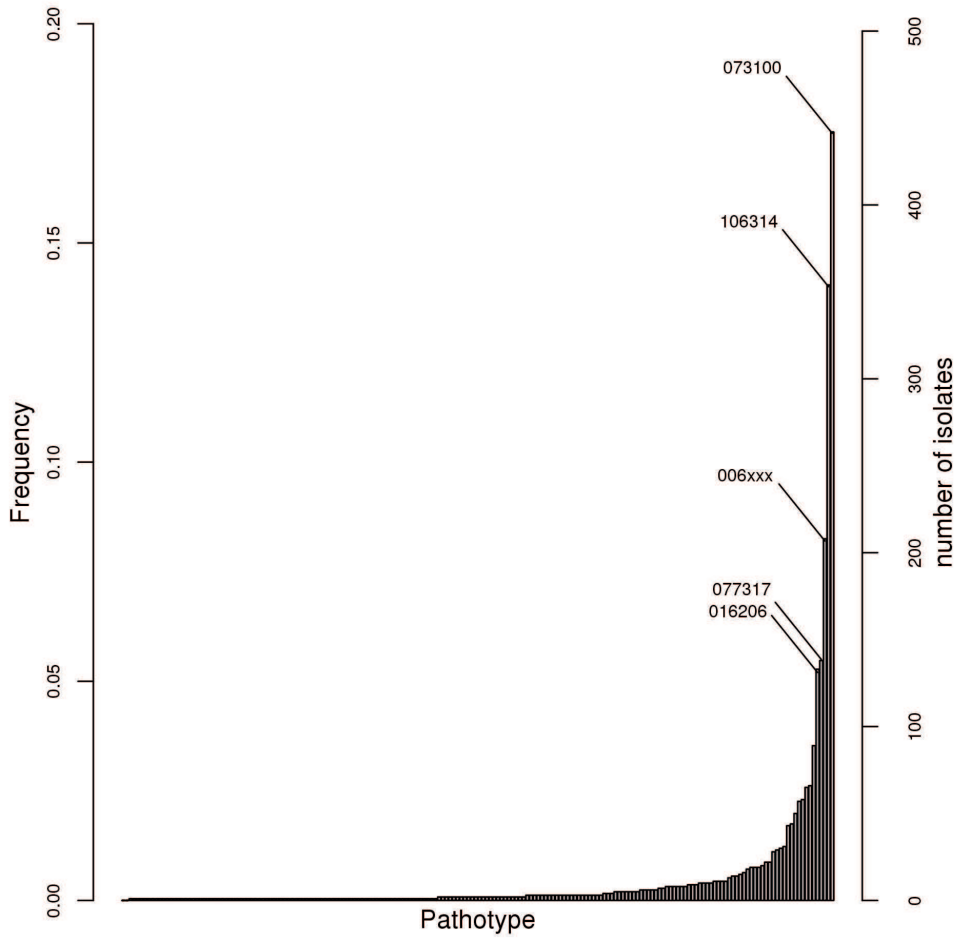


Figure 2.1: Ordered frequencies and corresponding number of isolates of the pathotypes identified in the Puccinia triticina survey over the considered period (1999-2008). The pathotypes that were specifically considered in this work are indicated.

$2 \times 12 m^2$ plots (ten rows) sown with several varieties. The sowing date is chosen as an average of the optimal sowing dates of all the varieties. Sowing density and nitrogen fertilisation are determined according to local practises, based on soil type, expected yield, etc. The plots are not sprayed with pesticides. The variety resistance level is directly evaluated from the disease scoring as observed resistance level $= 1 -$ disease score, where the disease score is the proportion of diseased leaf surface. In this paper, 'observed resistance' is defined as the resistance level of a variety as it is scored in the field (the term 'field resistance' is sometimes used in the literature) and represents the variety resistance as it is perceived by farmers and advisors. This 'observed resistance' is the result of environmental conditions, the genetic resistance factors of the varieties and the genetic composition of the pathogen population.

Scoring of the leaf rust symptoms is performed once a year in May or June. The *Apache* variety was scored from 1999 to 2008, *Caphorn* from 2001 to 2008, *Charger* from 1999 to 2007, *Isengrain* from 1999 to 2005, *Orvantis* from 2000 to 2008, *Soissons* from 1999 to 2008, and *Trémie* from 1999 to 2004 (figure 2.2).

Disease scoring yields both qualitative information (presence/absence of disease, usually referred to as 'incidence' in the phytopathological literature) and quantitative information (level of observed disease, usually referred to as 'severity'– figure 2.2). Absence of disease results either from the absence of the pathogen itself or, more likely in the case of leaf rust, from incompatibility of the pathotypes that were present at the scoring time with the considered variety. As classically observed, disease incidence on a variety was significantly linked to the proportion of virulent pathotypes in the pathogen population (figure 2.2). In order to account for such effects, qualitative virulence is introduced in the statistical model but, as stated before, we focus on the quantitative aspects of the host-pathogen relationship when exploring the datasets.

2.2.2 Statistical modelling

We constructed a statistical model in order to jointly analyse the three large-scale datasets describing the wheat leaf rust pathosystem. The model aims to look for the existing correlations between the wheat variety frequencies and the *P. triticina* population composition, in the one hand, and between the *P. triticina* population composition and the observed disease severity on the main wheat varieties, on the other hand. The need for a convenient and flexible framework to combine information from several parallel data sources led us to develop a state-space model (SSM). In a SSM, the datasets are first described by observation variables that constitute the observation process layer (figure 2.3). The observation variables are then linked to each other via unobserved hidden variables, classically referred to as 'latent variables' that constitute the system process layer. The model thus consisted of two sets of equations. The state equations described the statistical

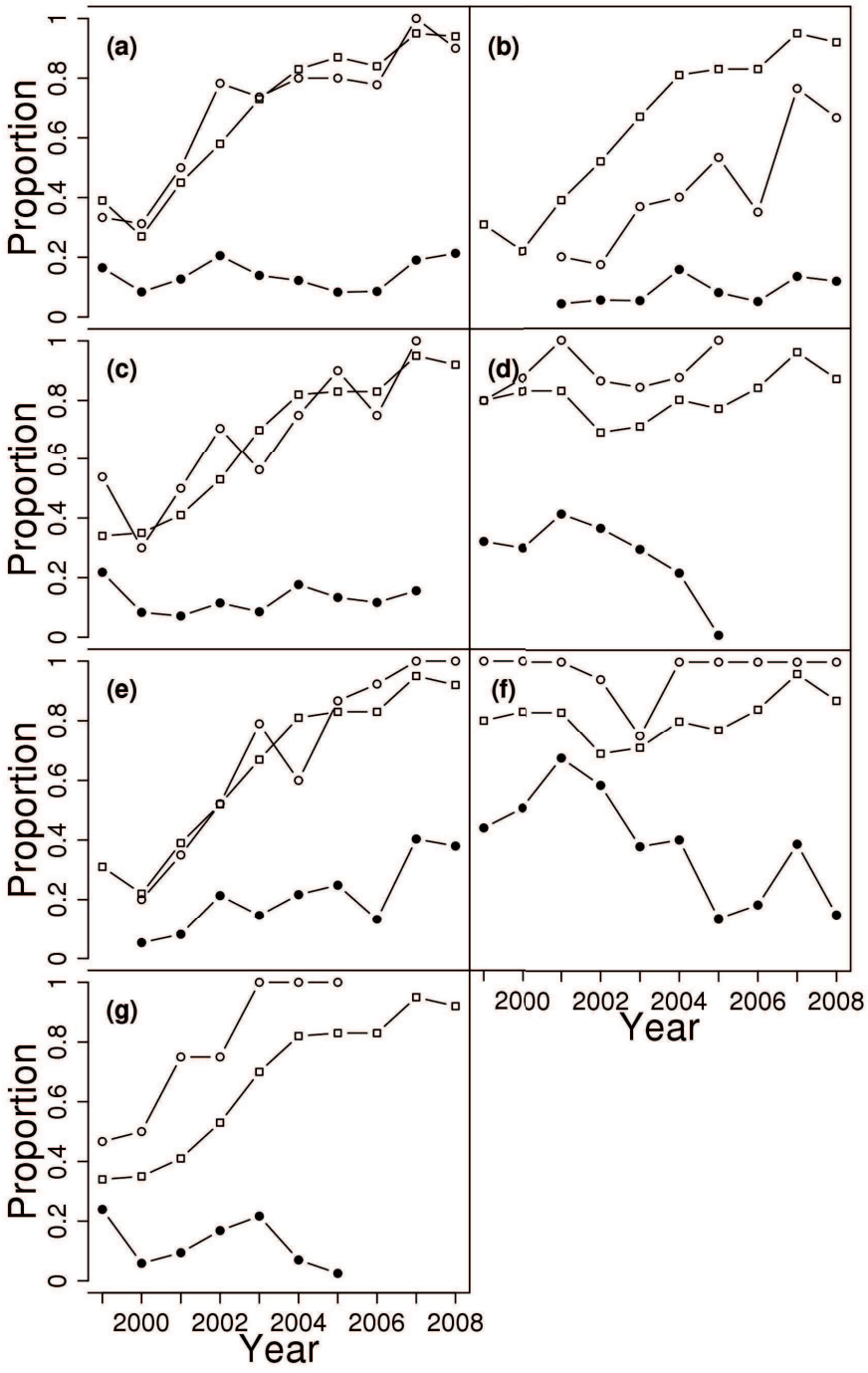


Figure 2.2: Evolution across all observation sites of the *Arvalis* survey of leaf rust incidence (open circles) and disease scores (closed circles), along with the frequency of pathotypes (open squares) that are virulent for varieties *Apache* (a), *Caphorn* (b), *Charger* (c), *Isengrain* (d), *Orvantis* (e), *Soissons* (f), and *Trémie* (g). The relationship between disease scores and time was tested by a GLM with beta distributed errors. The slope was significantly greater than 0 for *Orvantis* and smaller than 0 for *Soissons*, with a 0.001 threshold.

links between the latent variables. The observation equations linked the latent variables and the observation variables. Parameters appear both in the system and observation process layers.

With regard to the biological system, the model was broken down into two sub-models (figure 2.3). The population composition sub-model was centred on the *P. triticina* population composition and on its dependence on wheat variety frequencies. The disease severity sub-model was dedicated to the influence of *P. triticina* population composition on the observed leaf rust severity for each wheat variety.

We first present the main characteristics and assumptions of the model. We then describe the population composition sub-model and the disease severity sub-model, how they are connected and how the parameters can be interpreted in biological terms. Finally, we define the criteria that were used to interpret the results. Additional technical information is given in the appendix 1 of supporting information and table 2.4 provides a summary of terms and parameters definitions.

2.2.2.1 Model characteristics and assumptions

The host landscape was considered as a set of $V = 7$ varieties with year-dependent frequencies and a background landscape with low but unknown variety frequencies. The pathogen population was composed of $P = 5 + 1$ pathotypes: the five individual pathotypes presented in the data description section, plus a generic pathotype that included all the other pathotypes (figure 2.1). The pathotype distribution was assumed to be homogeneous at the scale of the country (Goyeau et al., 2006). No explicit dependence from one year to the next one was included (*i.e.* there is no time dependence in the model). Since data were collected at the end of the epidemic season, the effect of the current year on the pathogen population structure was assumed to be predominant over the potential effect of previous years. Even though the analysis was done at the French national scale, we considered $r = 7$ climatic regions: southeast, southwest, centre-south, centre, northeast, north and northwest of France, in order to account for the influence of climatic conditions on disease development in the disease severity sub-model.

2.2.2.2 Population composition sub-model

In the leaf rust survey, year t, $N_{v,t}$ pustules were sampled on variety v and each one was classified as one of the P pathotypes defined above. Let $Y_{v,t} = (y_{v,t,1}, \cdots, y_{v,t,P})$ be the categorical random variable whose elements $y_{v,t,p}$ denote the numbers of leaf rust pustules sampled on variety v, year t, and assigned to pathotype p. Let $\pi_{v,t,p}$ denote the proportions of pathotype p on variety v, year t. Assuming that the probability for pathotype p to be

Table 2.4: Definitions of the main terms and parameters used in this study.

Terms	Symbols	Description/biological interpretation
Pathotype	p	Phenotype for qualitative virulence.
Disease score	X	Proportion of diseased leaf surface.
Observed resistance level	–	1 – disease score (results from genetic resistance factors of the varieties, genetic composition of the pathogen population and the environmental conditions).
Incidence	–	Presence/absence of disease (qualitative information resulting from the disease scores).
Severity	–	Level of observed disease (quantitative information resulting from the disease scores).
Aggressiveness	–	Quantitative component of pathogenicity.
Virulence	δ	Capacity of a pathotype to overcome a major resistance gene.
Basic affinity	α	Indication of a pathotype aggressiveness on a variety at the landscape scale.
–	β	Response of the pathogen to changes in the landscape composition.
–	a^0	Susceptibility of a variety confronted to the global leaf rust population.
–	a^1	Year effect.
–	a^2	Climatic region effect.
–	b	Effect of the proportion of a pathotype on the observed disease on a variety.
–	C_1	For a given variety, minimal value of the posterior probability that a particular parameter for a pathotype is greater than the same parameter for another pathotype.
–	C_2	For a given pathotype, minimal value of the posterior probability that a particular parameter for a variety is greater than the same parameter for another variety.

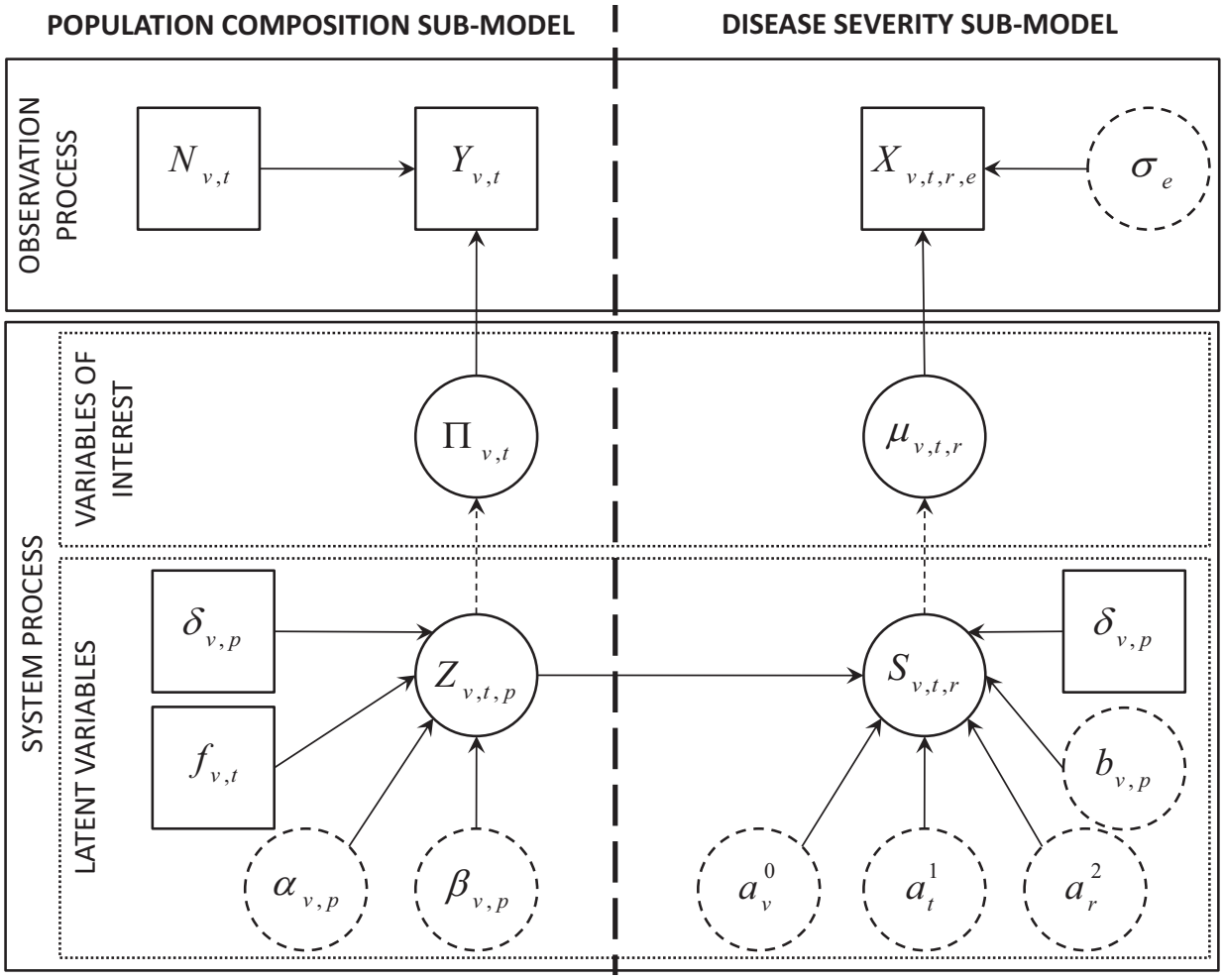


Figure 2.3: Graphical representation of the state-space model (p: pathotype; v: variety; t: year; r: climatic region; e: trial in Arvalis survey). The system process is composed of the latent variables (solid circles), $Z_{v,t,p}$ and $S_{v,t,r}$, depending on unknown parameters (dashed circles) and covariables (solid squares) via stochastic links (solid arrows). From these latent variables, the unobserved variables of interest (solid circles), $\Pi_{v,t}$ (the vector of pathotype proportions on variety v year t) and $\mu_{v,t,r}$ (the mean proportion of diseased leaf area of variety v, year t in region r), are deduced via a deterministic link (dashed arrows). The observed data (solid square) are the visible part of the system process. In the population composition sub-model, $Y_{v,t}$ classify the leaf rust pustules sampled on variety v, year t, for the p pathotypes. In the disease severity sub-model, $X_{v,t,r,e}$ denote the disease score attributed to variety v on trial e, year t and region r in the *Arvalis* survey. The parameter σ_e (dashed circle) represents the within trial variability. The other parameters are: $\phi_{v,t}$ (variety frequencies), $\delta_{v,p}$ (variety–pathotype compatibility), $\alpha_{v,p}$ (basic affinity), $\beta_{v,p}$ (sensitivity to changes in variety frequencies), a_v^0 (variety effect), a_t^1 (year effect), a_r^2 (region effect) and $b_{v,p}$ (sensitivity to changes in pathotype frequencies).

sampled only depends on its proportion, $Y_{v,t}$ has a multinomial distribution with parameters $N_{v,t}$ and $\Pi_{v,t} = (\pi_{v,t,1}, \cdots, \pi_{v,t,P})$. This led to the observation equation that describes the *P. triticina* population composition:

$$Y_{v,t}|N_{v,t}, \Pi_{v,t} \sim \text{Multinomial}\left(N_{v,t}, \Pi_{v,t}\right).$$

In the state part of this sub-model (figure 2.3), the proportions $\pi_{v,t,p}$ were associated to a latent variable $Z_{v,t,p}$ that represents the relative population size of pathotype p on variety v year t. Note that $Z_{v,t,p}$ is not the actual size of the pathotype population but a scale variable instead. $\pi_{v,t,p}$ and $Z_{v,t,p}$ were linked through the state equations (see the appendix 1 of supporting information for more details):

$$\mathrm{E}[Z_{v,t,p}] = \bar{Z}_{v,t,p} = \alpha_{v,p} + \sum_{i=1}^{V} \delta_{i,p}\beta_{i,p}\phi_{i,t} \qquad (\alpha_{v,p}, \beta_{v,p}) \geq 0, \tag{2.1a}$$

$$Z_{v,t,p}|\bar{Z}_{v,t,p} \sim \text{Gamma}\left(\bar{Z}_{v,t,p}, 1\right), \tag{2.1b}$$

$$\pi_{v,t,p} = \frac{Z_{v,t,p}}{Z_{v,t,1} + \cdots + Z_{v,t,P}}. \tag{2.1c}$$

In equation (2.1a), $\alpha_{v,p}$ represents a basic affinity between pathotype p and variety v, and the second term accounts for the effect of the landscape composition, variable $\phi_{v,t}$ denoting the frequency of variety v during year t, and parameter $\beta_{v,p}$ denoting the sensitivity of pathotype p to the frequency of variety v. Prior information on the gene-for-gene relationship was integrated through the binary parameter $\delta_{v,p}$ defined by:

$$\delta_{v,t} = \begin{cases} 1 & \text{if } p \text{ and } v \text{ are compatible,} \\ \\ 0 & \text{otherwise.} \end{cases}$$

Note that in equation (2.1a), there can be a non-zero basic affinity even if the variety and the pathotype are incompatible ($\delta_{v,p} = 0$). This is to take the situation in which an incompatible pathotype is nonetheless able to produce a few pustules on a resistant variety into account, as sometimes occurs in field epidemics (Samborski, 1985).

In this statistical model, the links that were established between each component are descriptive and the associated coefficients should not be directly assimilated to biological parameters of the plant-pathogen relationship. Interpretation of equation (2.1a) can however be made as follows. The basic affinity $\alpha_{v,p}$ of pathotype p for variety v is the part of $\mathrm{E}[Z]$ that is not accounted for by the frequencies in the landscape of the seven varieties considered in the analysis. A high $\alpha_{v,p}$ value means that pathotype p was always well represented on variety v during the period studied, regardless of the landscape composition. Therefore, for a compatible pathotype, this parameter provides an indication on the pathotype aggressiveness on a variety, based on the size of its population on that variety and relative to the other

pathotypes. Parameter $\beta_{v,p}$ quantifies the response of the pathogen to changes in the landscape composition. A simple analogy can be made here with a linear regression, with α being analogous to the intercept and β to the slope.

2.2.2.3 Disease severity sub-model

Let $X_{v,t,r,e}$ denote the disease score attributed in the *Arvalis* survey to variety v on trial e, year t and in region r. Its expectation $\mu_{v,t,r}$ was assumed to reflect the mean rust severity on variety v that year in that region. The score $X_{v,t,r,e}$ varied in the $[0, 1]$ interval but the actual range of notation differed between trials. Consequently the scores were assumed to follow a Beta distribution with mean value, $\mu_{v,t,r}$, and scale parameter σ_e depending on the trial. This led to the observation equation that describes the disease scoring dataset:

$$X_{v,t,r,e}|\mu_{v,t,r}, \sigma_e \sim \text{Beta}\left(\mu_{v,t,r}, \sigma_e\right).$$

The expected disease scores $\mu_{v,t,r}$ were associated to two latent variables $S_{v,t,r}$ and $S'_{v,t,r}$ through the state equations (see the appendix 1 of supporting information for more details):

$$\text{E}[S_{v,t,r}] = \bar{S}_{v,t,r} = a^0_v + a^1_t + a^2_r + \sum_{j=1}^{P-1} \delta_{v,j} b_{v,j} \pi_{v,t,i} \qquad \left(a^0_v, a^1_t, a^2_r, b_{v,p}\right) \geq 0, \tag{2.2a}$$

$$S_{v,t,r}|\bar{S}_{v,t,r} \sim \text{Gamma}\left(\bar{S}_{v,t,r}, 1\right), \tag{2.2b}$$

$$\mu_{v,t,r} = \frac{S_{v,t,r}}{S_{v,t,r} + S'_{v,t,r}}. \tag{2.2c}$$

Since disease scores are defined in the scoring procedure as the observed proportion of diseased leaf area, $\mu_{v,t,r} = \text{E}[X_{v,t,r,e}]$ can be identified to the mean proportion of leaf area that was diseased for variety v during year t in region r. In equations (2.2), $S_{v,t,r}$ represents the diseased leaf area of a variety v, year t, in region r and $S'_{v,t,r}$ denotes the healthy leaf area. As for Z, S and S' should not be considered as actual areas but as scale variables instead. In equation (2.2a), parameters a^0, a^1 and a^2 represent the variety, year and region main effects, respectively, and define a basal disease pressure. In particular, a^0 can be seen as the basic susceptibility of a variety, confronted to the global leaf rust population over the entire period considered. The last term of the equation accounts for the effect of the pathogen population composition, where δ still denotes 0-1 qualitative virulence. Here again, b would be analogous to a slope, quantifying the link between the proportion of a particular pathotype and the observed disease on a variety, whereas $a^0 + a^1 + a^2$ could be assimilated to an intercept. The summation in equation (2.2a) is performed over the $(P - 1)$ individual pathotypes considered earlier. The index P identifies the generic pathotype that groups all the other pathotypes that were present during the period. By construction, their joint effect is included into the basal disease pressure.

2.2.2.4 Index and criterion definitions

In the model defined above, the interactions between pathotypes and varieties are quantified through the parameters $\theta_{v,p}$, with θ in $\{\alpha, \beta, b\}$. In order to synthesise the resulting information, two sets of criteria varying in $[0, 1]$ were defined:

$$C_1(\theta; v, p) = \min_{p'}\{\text{Prob}(\theta_{v,p} > \theta_{v,p'})\},$$
$$C_2(\theta; v, p) = \min_{v'}\{\text{Prob}(\theta_{v,p} > \theta_{v',p})\}.$$

Where Prob $(\bullet)$ denotes posterior probabilities. In the first case ($C_1(\theta; v, p)$), the variety v is fixed and $\theta_{v,p}$ values are compared among pathotypes. In the second case ($C_2(\theta; v, p)$), the pathotype p is fixed and $\theta_{v,p}$ values are compared among varieties. With these definitions, large $C_1(\theta; v, p)$ values identify the pathotypes that are most strongly linked to variety v with respect to parameter θ, whereas, in a symmetrical way, large $C_2(\theta; v, p)$ values identify the varieties to which pathotype p is the most strongly linked with respect to parameter θ.

In order to understand how the pathogen population responds to host frequencies in the cultivated landscape, it is worthwhile to compare the effect of the basic affinity and the effect of the landscape composition, as defined in equation (2.1a). The relative weights of both types of effects were defined by

$$W(\alpha; v, t, p) = \frac{\alpha_{v,t,p}}{\alpha_{v,p} + \sum_{i=1}^{V} \delta_{i,p}\beta_{i,p}\phi_{i,t}},$$

and

$$W(\beta, i; v, t, p) = \frac{\delta_{i,p}\beta_{i,p}\phi_{i,t}}{\alpha_{v,p} + \sum_{i=1}^{V} \delta_{i,p}\beta_{i,p}\phi_{i,t}}.$$

2.2.3 Bayesian implementation

Inference on the parameters was performed by Bayesian statistical methods, resulting in a joint posterior distribution (Gelman et al., 2004). This posterior distribution was computed via a Markov Chain Monte Carlo (MCMC) method using Jags software (Plummer, 2010).

2.2.3.1 Prior densities

Only the δ parameters were considered to be known *a priori*. For all the other parameters no *a priori* information was assumed to be available and non-informative prior densities were used: α, β, a^0, a^1, a^2 and b were defined as being uniformly distributed on $[0, 10000]$ and the trial variances σ were defined as being uniformly distributed on $[0, 1]$. It was systematically

verified that the upper bound of the uniform prior was large enough to have no influence on the posterior densities.

For identification reasons, we fixed $\bar{Z}_{v,t,P} = 1, \forall(v,t)$ and $S'_{v,t,r} = 5, \forall(v,t,r)$. This choice was made after checking the data-fitting and performing sensitivity analyses. In particular, it was verified that variation around the chosen values had little impact on the criteria defined above.

2.2.3.2 MCMC convergence and mixing

Three MCMC-chains of 125,000 iterations were computed. Convergence was assessed using the Gelman and Rubin statistic ($\hat{R}$) which compares the within to the between variability of chains started at different and dispersed initial values (Gelman et al., 2004). Burn-in was set to 25,000 and thinning every 100 iterations resulted in acceptable mixing and convergence ($\hat{R} < 1.1$, for all the parameters).

2.2.3.3 Data fitting

To assess the fit of the model to the data, we used an approach known as posterior predictive checking, which is a Bayesian counterpart of the classical tests for goodness-of-fit (Gelman et al., 2004). The idea is to generate replicated data from the posterior distribution of the parameters. If the model fits the data well, then the replicated data should be similar to the observed data. For both pathotype proportion and disease score the posterior distributions of the replicated data showed an adequate fit of the model (figure 2.4).

2.3 Results

2.3.1 Overview

In a first step, the model is used to explore the relationship between the host and the pathogen populations. Two aspects are considered: the basic affinity between pathotypes and varieties, described by parameters α, and the response of the pathogen to changes in the landscape composition, described by parameters β. Posterior densities of parameters α and β are given in figures 2.5 and 2.6 but more synthetic information is provided by criteria $C_1(\theta; v, p)$ and $C_2(\theta; v, p)$, with $\theta = \alpha$ or β, which will be used to discuss how the pathogen population responds to changes in the host population (tables 2.5 and 2.6).

In a second step, we will attempt to link the observed resistance in the field to the composition of the pathogen population. Criterion $C_1(b; v, p)$ (table 2.5) is used to describe

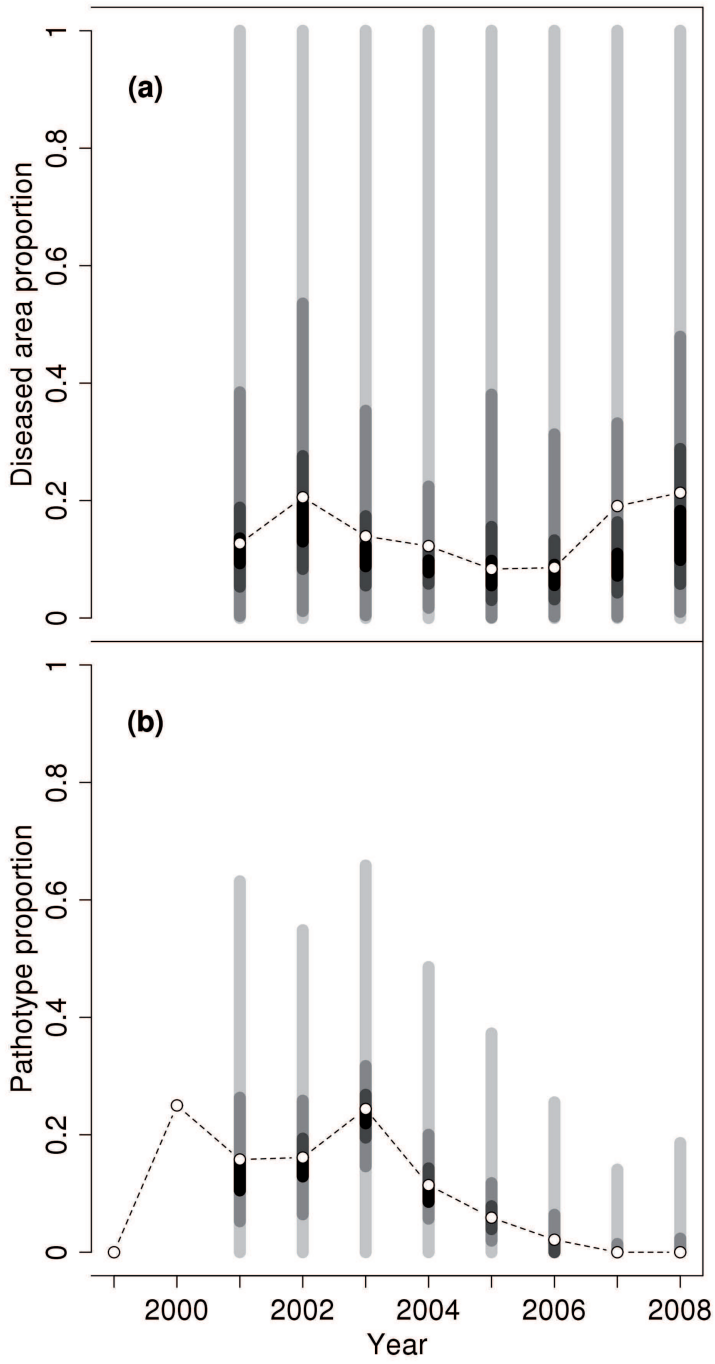


Figure 2.4: Posterior densities of replicated data (grey bars), generated from the posterior distribution of the parameters, and observed data (open circles). Example of the mean disease score on variety *Apache* (a) and the proportion of pathotype 016206 (b). Grey bars are median centred quantile intervals, from the darkest to the lightest: $[0.45, 0.55]$, $[0.35, 0.65]$, $[0.15, 0.85]$ and $[0, 1]$). The model does not incorporate time-dependency between years.

Table 2.5: Values of the C_1 criterion for each parameter and each pathotype-variety pair. This criterion compares the parameter values among pathotypes for a given variety. Large $C_1(\theta; v, p)$ (with θ in $\{\alpha, \beta, b\}$) values identify the pathotypes that are most strongly linked to variety v according to parameter θ.

	Apache			*Caphorn*			*Charger*			*Isengrain*		
	α	β	b	α	β	b	α	β	b	α	β	b
006xxx	0.40	0.21	0.36	0.19	0.29	0.32	0.23	0.16	0.53	0.00	0.11	-
016206	0.45	0.38	0.55	0.15	0.00	0.28	0.64	0.57	0.47	0.00	-	-
073100	0.55	-	-	0.09	-	-	0.24	-	-	1.00	0.73	-
077317	0.37	0.62	0.45	0.06	0.00	0.68	0.36	0.43	0.45	0.00	0.27	-
106314	0.25	0.13	0.31	0.81	0.71	0.10	0.28	0.12	0.38	0.00	0.09	-
	Orvantis			*Soissons*			*Trémie*					
	α	β	b	α	β	b	α	β	b			
006xxx	0.78	0.26	0.20	0.00	0.14	0.04	0.03	0.18	-			
016206	0.22	0.40	0.07	0.00	-	-	0.51	0.59	-			
073100	0.10	-	-	1.00	0.70	0.95	0.49	-	-			
077317	0.15	0.60	0.16	0.00	0.30	0.05	0.08	0.41	-			
106314	0.08	0.16	0.80	0.00	0.11	0.00	0.03	0.16	-			

α: basic affinity for a pathotype to a variety; β: sensitivity of pathotypes to changes in variety frequencies; b: sensitivity of disease severity on a variety to the pathotype proportions.

how the disease level observed on a variety can be linked to the pathogen population composition. Posterior densities of the parameters are given in figure 2.8.

2.3.2 Variety frequencies and pathogen population composition

2.3.2.1 Basic affinity (α)

The C_1 criterion (table 2.5) indicates the dominance of a pathotype on a variety. Pathotypes 106314 and 006xxx were dominant on the two most recent varieties, *Caphorn* and *Orvantis* (released in 2000), respectively ($C_1(\alpha, Caphorn, 106314) = 0.81$ and $C_1(\alpha, Orvantis, 006xxx) = 0.78$). These two pathotypes are also the most recent ones in the French *P. triticina* population (no isolate of these pathotypes were found before 2003 and 2002, respectively). Pathotype 073100 was clearly dominant on varieties *Isengrain* and *Soissons*. Two pathotypes, 016206 and 073100, were dominant on *Trémie*. Pathotype 016206 was moderately dominant on *Charger*. No pathotype was found to be specifically dominant on *Apache*. Note that the *Trémie*-073100 association was unexpected here since, according to its qualitative virulence pattern, this interaction should be incompatible (Goyeau et al., 2006). Nevertheless, infection of *Trémie* with that pathotype often produces

Table 2.6: Values of the C_2 criterion for parameter α and β for each pathotype-variety pair. This criterion compares the parameter values among pathotypes for a given variety. Large $C_2(\theta; v, p)$ (with θ in $\{\alpha, \beta\}$) values identify the pathotypes that are most strongly linked to variety v according to parameter θ.

	006*xxx*		016206		073100		077317		106314	
	α	β	α	β	α	β	α	β	α	β
Apache	0.17	0.00	0.19	0.19	0.00	-	0.44	0.20	0.06	0.00
Caphorn	0.35	0.95	0.29	0.14	0.00	-	0.37	0.15	0.94	1.00
Charger	0.09	0.00	0.25	0.35	0.00	-	0.40	0.20	0.06	0.00
Isengrain	0.05	0.00	0.11	-	0.62	0.52	0.56	0.11	0.03	0.00
Orvantis	0.65	0.05	0.23	0.64	0.00	-	0.44	0.21	0.03	0.00
Soissons	0.04	0.00	0.01	-	0.38	0.48	0.44	0.13	0.02	0.00
Trémie	0.05	0.00	0.71	0.36	0.00	-	0.33	0.21	0.02	0.00

α: basic affinity for a pathotype to a variety; β: sensitivity of pathotypes to changes in variety frequencies.

a few viable pustules. Since the frequency of 073100 in the pathogen population was high, and the disease levels (total number of pustules) on *Trémie* were low, it was logical to detect a link between *Trémie* and 073100.

Criterion C_2 (table 2.6) makes it possible to determine the preference of a pathotype for one or several of the seven varieties considered. Pathotype 006*xxx* was found preferentially on *Orvantis* but was also present on *Apache* and *Caphorn*. These three varieties are those that share the *Lr*37 resistance gene. On the contrary, pathotype 106314 was only related to *Caphorn* ($C_2(\alpha, Caphorn, 106314) = 0.94$). Pathotype 073100 was highly abundant on *Soissons* and *Isengrain* and was never related to the other varieties ($C_2 = 0$). Pathotype 016206 was related to *Trémie* and was infrequently found on *Isengrain* and *Soissons*. Finally, pathotype 077317 was identified as a generalist, with no preference for any of the varieties.

2.3.2.2 Response of pathogen to changes in landscape composition (β)

Changes in the frequencies of *Isengrain* and *Soissons* in the cultivated landscape strongly affected the proportion of pathotype 073100 in the pathogen population (table 2.5). Changes in the frequencies of *Apache*, *Charger*, *Orvantis* and *Trémie* mainly affected 016206 and 077317 (table 2.5 and figure 2.6) but 016206 was more especially affected by *Orvantis* (table 2.6). As for α, pathotype 077317 was not specifically sensitive to any of these varieties (table 2.6). Changes in the frequency of *Caphorn* mainly affected pathotypes 006*xxx* and 106314 (table 2.5 and figure 2.6) and, among the considered varieties, *Caphorn* was the most influential variety on both pathotypes (table 2.6). Data analysis also suggested that changes

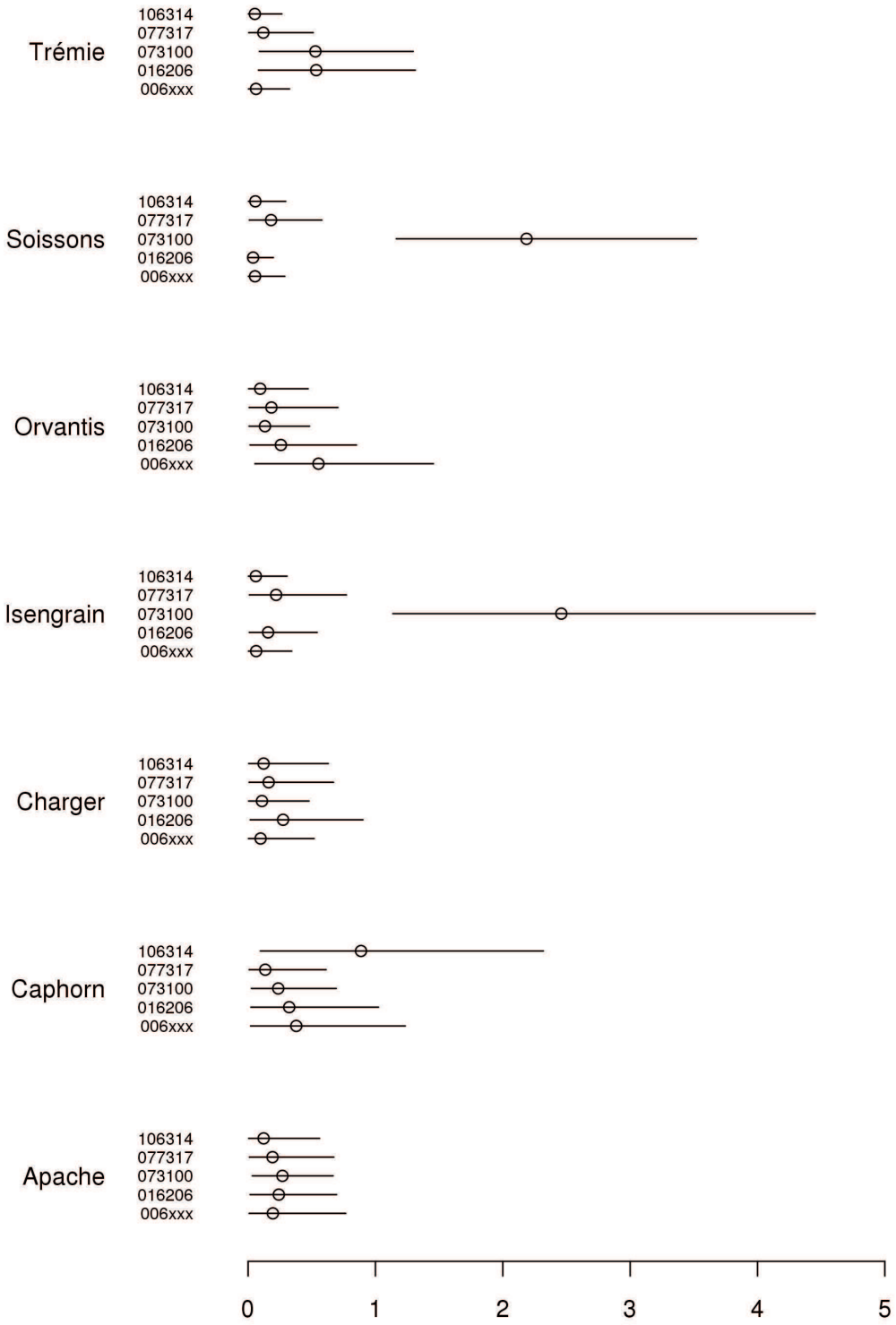


Figure 2.5: Posterior densities of parameter α. The median (open circle) and the 95% credibility interval (horizontal solid line) are indicated for each pathotype-variety pair.

in *Orvantis* influenced pathotype 006xxx (figure 2.6). Note that, although pathotype 006xxx presented a strong affinity α for *Orvantis*, it did not appear very sensitive to changes in the proportion of *Orvantis* in the host population (table 2.6, parameter β). A simple reason for that is that *Orvantis*'s frequency did not vary much between 2002 and 2008 (table 2.1).

2.3.2.3 Relative effect of basic affinity and changes in variety frequencies (W)

Figure 2.7 shows the respective weights of the basic affinity $W(\alpha; v, t, p)$ and of the changes in the variety frequencies $W(\beta, i; v, t, p)$ on the population size of a pathotype on a variety, as estimated by Z (equation (2.1)). Overall, figure 2.7 suggests that the frequency of a pathotype on a variety is the highest when the landscape composition is the most favourable, i.e. when $W(\alpha; v, t, p)$ is minimal.

The presence of 073100 on *Soissons* was mainly explained by its very high affinity for this variety (figure 2.7a). On the contrary, the frequency of 106314 on the same variety was mainly related to the frequency of *Caphorn* in the landscape (figure 2.7c). The case of 077317 is more complex: changes in the landscape composition influenced this pathotype through a combination of varieties that varied over time: mainly *Trémie* and *Soissons* in the beginning of the period, then *Charger* and finally *Apache* and *Orvantis* (figure 2.7b).

In the right column of figure 2.7, we examine the situation of pathotype 106314. The frequency of this pathotype on *Caphorn* was explained by both its affinity (α) and the increasing frequency of *Caphorn* in the landscape, with weights of comparable magnitude (figure 2.7e). This suggests an amplification effect, with the increase in *Caphorn* in the landscape (table 2.1) resulting in a higher frequency of 106314 on this variety. On *Soissons* and *Apache* (figure 2.7c and d), 106314 increased from 0% in 2003 to 35% and 39%, respectively, in 2008 and this increase was accounted for by the frequency of *Caphorn* in the landscape. On *Orvantis*, pathotype 106314 also increased in frequency, probably for the same reasons, but it only reached 13% of the pathogen population sampled on this variety at the end of the period (figure 2.7f). This might be due to a competition with 006xxx, which had a high affinity for *Orvantis* with respect to the other pathotypes (table 2.5).

2.3.3 Disease dynamics

2.3.3.1 Effect of pathogen population composition on observed resistance (b)

For *Isengrain* and *Trémie*, very large credibility intervals were obtained for parameter b (figure 2.8), probably because pathotypes 006xxx and 106314 were observed for only one or two years on these varieties, which made the data difficult to exploit. On the other varieties, three different kinds of responses were obtained. For *Apache*, *Caphorn* and *Charger*, the

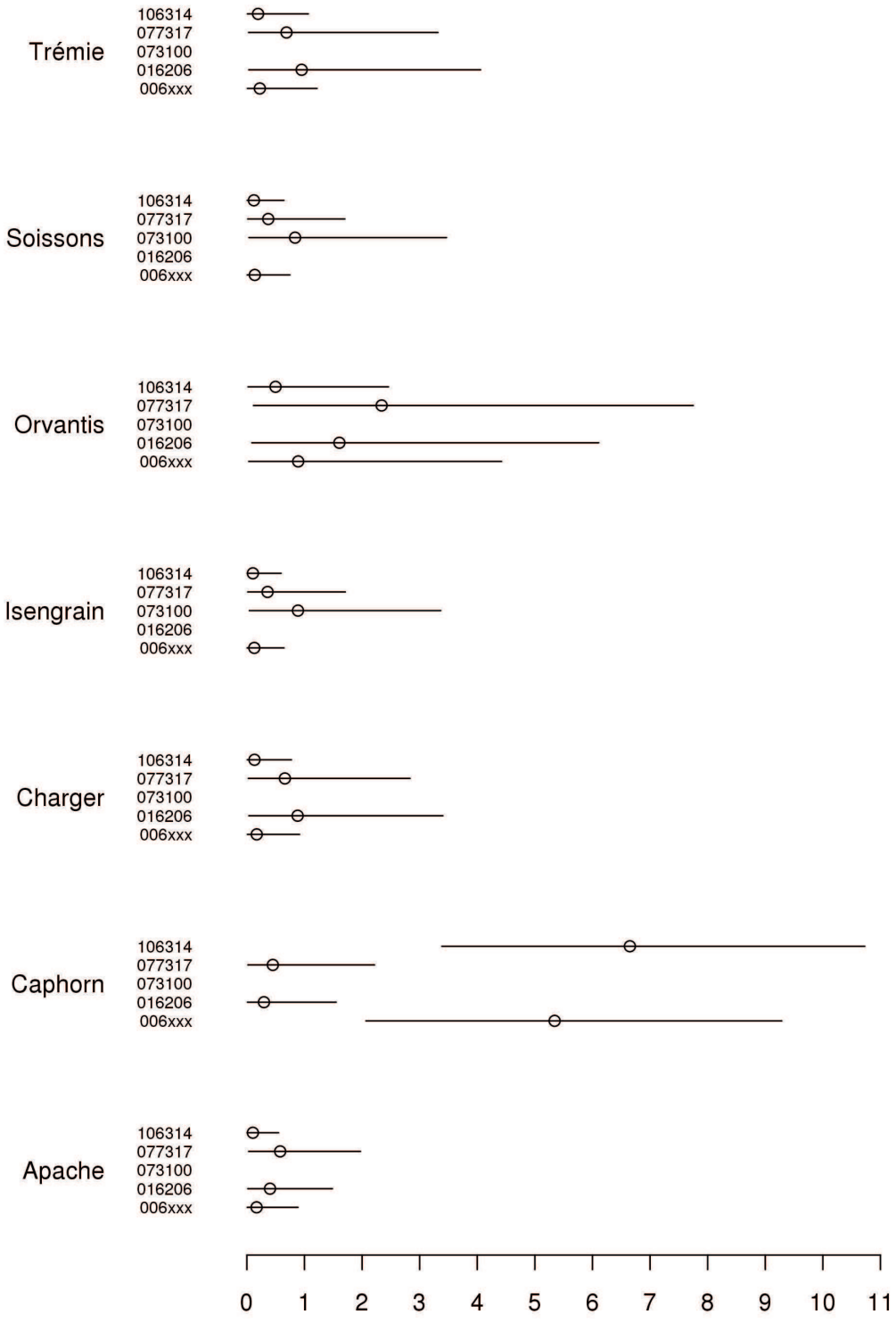


Figure 2.6: Posterior densities of parameter β. The median (open circle) and the 95% credibility interval (horizontal solid line) are indicated for each pathotype-variety pair.

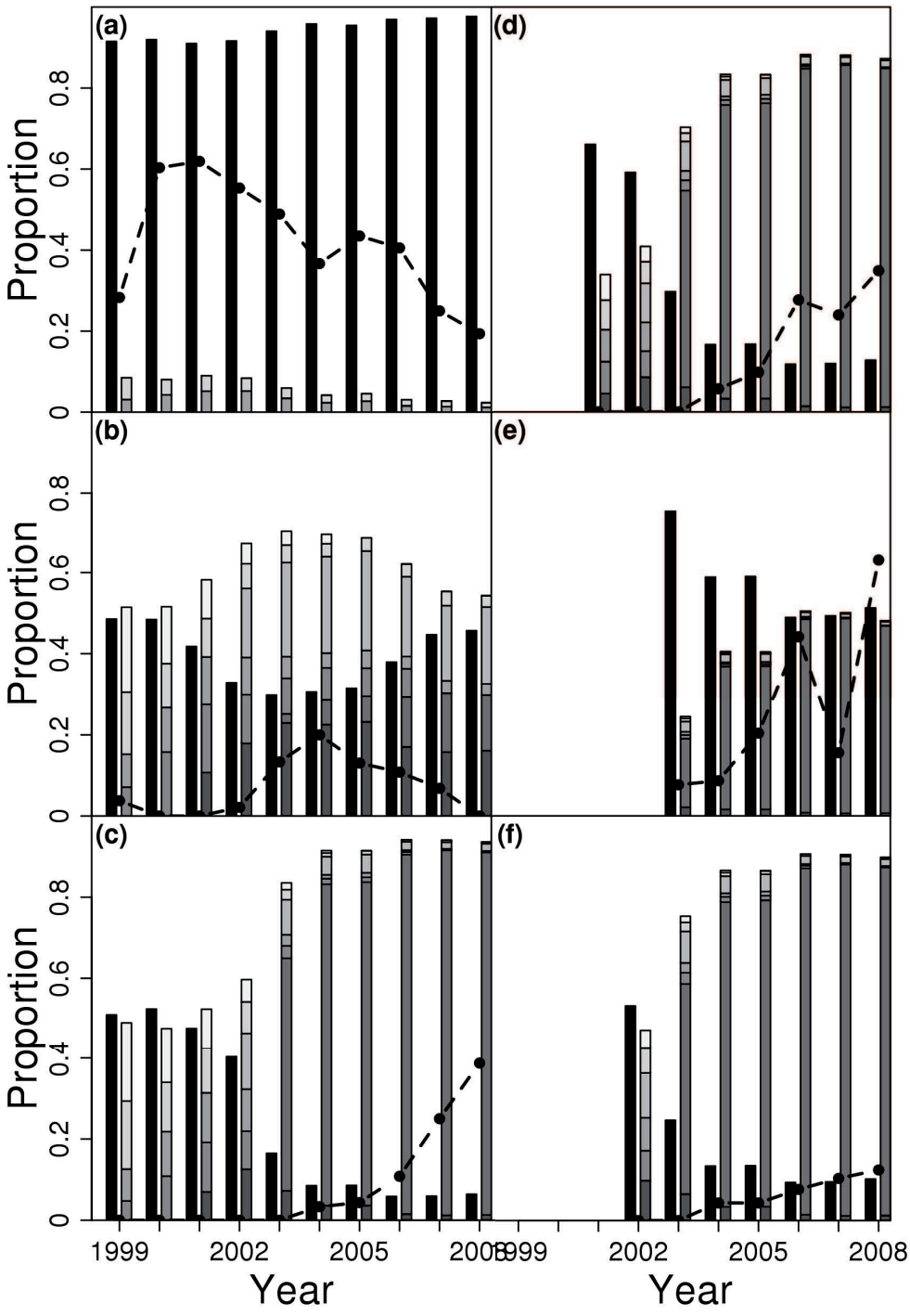


Figure 2.7: Observed proportion of pathotype p on variety v (closed circles) along with associated model-based indices (bars), for selected (p, v) pairs. Pathotypes 073100 (a), 077317 (b) and 106314 (c) on variety *Soissons*; pathotype 106314 on varieties *Apache* (d), *Caphorn* (e) and *Orvantis* (f). Weights of basic affinity, $W(\alpha; v, p)$: black bars and changes in landscape composition, $W(\beta, i; v, t, p)$: grey bars, on the considered pathotype frequency. Grey bars indicate varieties *Apache*, *Caphorn*, *Charger*, *Isengrain*, *Orvantis*, *Soissons* and *Trémie* from the darkest to the lightest.

disease scoring yielded constant values (figure 2.2) and it seems difficult to link the observed resistance levels to a specific pathotype. Nevertheless, the disease scores on *Apache* could be related to the proportions of pathotypes 077317 and 016206 and the disease scores on *Caphorn* could be related to 077317 (table 2.5 and figure 2.8). It is interesting to note that the b value for the pair *Caphorn*-106314 was very low, even though the frequency of 106314 was strongly linked to *Caphorn* (tables 2.5 and 2.6). This suggests that *Caphorn*, when grown extensively, influenced the pathotype frequencies in the pathogen population without being affected by severe epidemics itself. The second kind of response is that of *Orvantis*, for which the observed resistance level was decreasing (so that the disease scores kept increasing - figure 2.2). This was correlated to the increase in 106314 in the pathogen population ($C_1(b; Orvantis, 106314) = 0.80$). The last response type is that of *Soissons*, on which the observed resistance increased (and disease scores decreased - figure 2.2). This could be linked to the decrease in the frequency of 073100 in the pathogen population ($C_1(b; Soissons, 73100) = 0.95$).

2.3.3.2 Variety, year and region effects

The variety effect, a^0, can be used as a criterion to rank the varieties according to their susceptibility to leaf rust, taking both qualitative and quantitative pathogenicity into account. As expected, *Isengrain* and *Soissons* had the highest a^0 values and *Caphorn* the lowest (figure 2.9a). The year effect, a^1 (figure 2.9b), accounted for the *P. triticina* population breakdown in 2003 and subsequent changes in the following years (Goyeau et al., 2006). The region effect, a^2 (figure 2.9c), was consistent with the known behavior of the pathogen, notably that northern and southeastern France are not favourable to wheat leaf rust.

2.4 Discussion

2.4.1 Overview

In this article, we developed a statistical model in order to jointly analyse three large-scale datasets describing the wheat leaf rust pathosystem. Many published papers establish a relationship between the frequency of resistance genes in the host population and the evolution of the pathogen population structure in terms of pathotypes, based on qualitative virulence factors (*e.g.* Hovmøller et al., 1993; Goyeau et al., 2006; Kolmer, 2002). The originality of the present approach was to account for the quantitative aspects of the host-pathogen relationship and to relate host and pathogen genotype frequencies to observed disease severity values. The analysis demonstrated that the landscape varietal composition influences the observed resistance level of the most frequently grown wheat varieties by

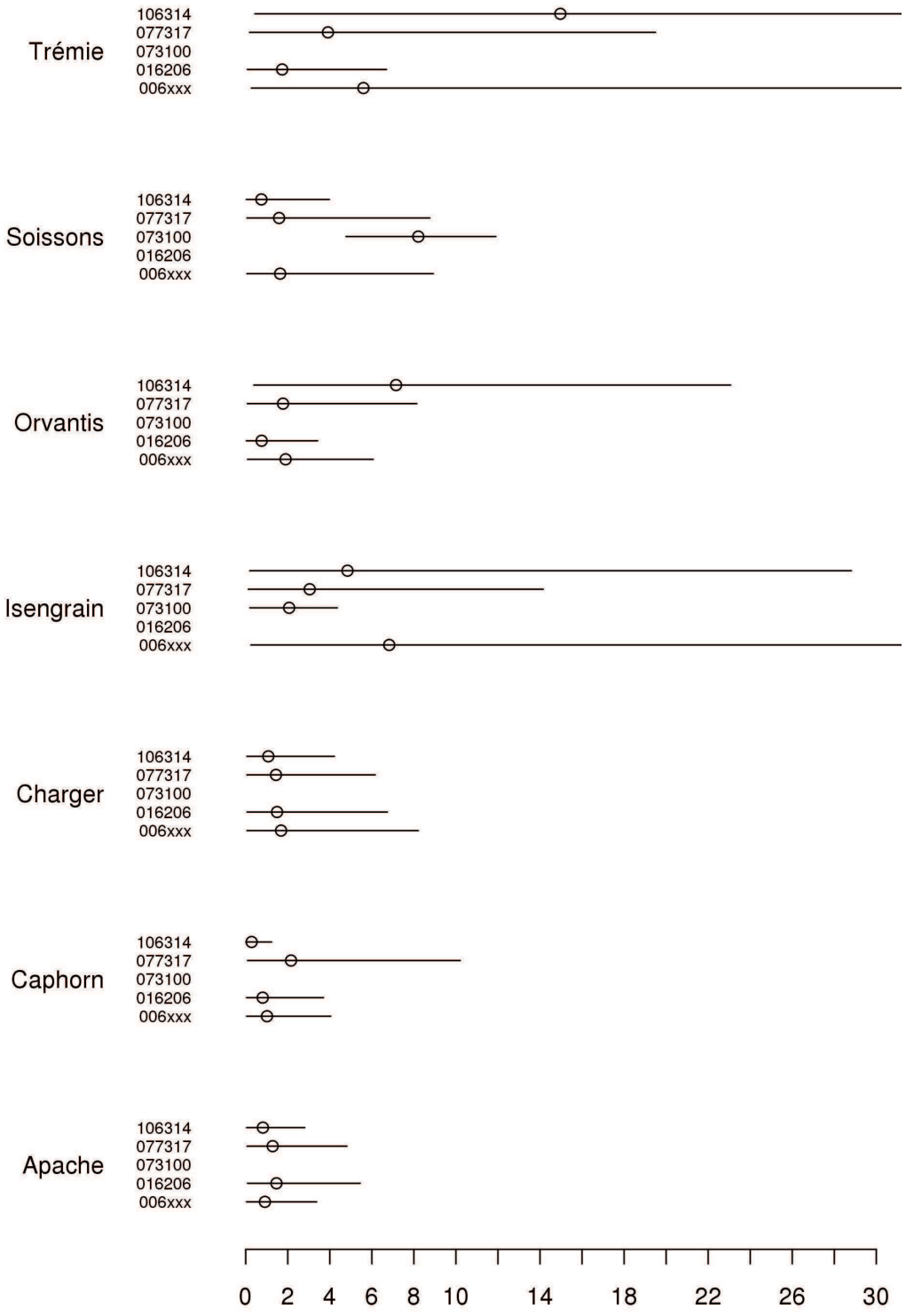


Figure 2.8: Posterior densities of parameter b. The median (open circle) and the 95% credibility interval (horizontal solid line) are indicated for each pathotype-variety pair.

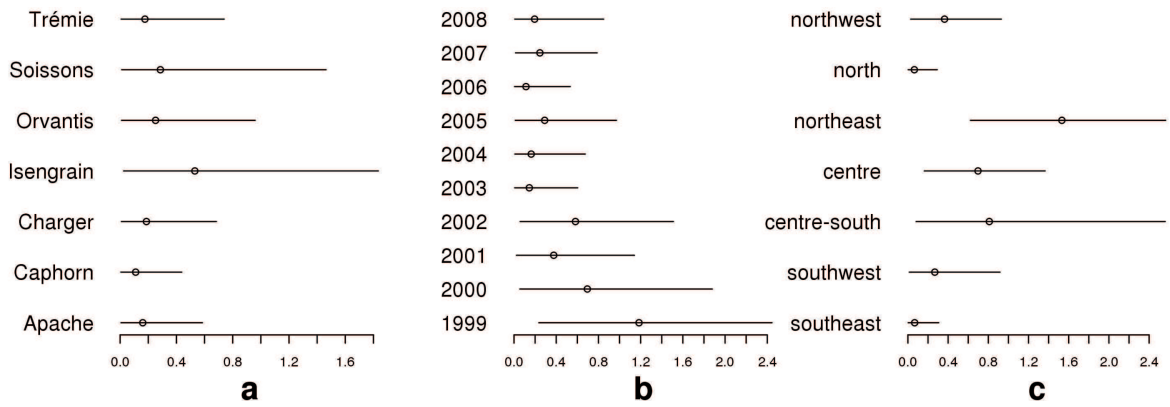


Figure 2.9: Posterior densities of parameter a^0 (a), a^1 (b) and a^2 (c). The median (open circle) and the 95% credibility interval (horizontal solid line) are indicated for each pathotype-variety pair.

altering the structure of the pathogen population. Another conclusion of the study is that quantitative effects (linked either to host quantitative resistance or pathogen aggressiveness) played a major role in shaping the leaf rust population structure in France over the past ten years.

2.4.2 Results interpretation

The analysis of the link between the pathogen population composition (in terms of pathotypes) and the landscape varietal composition revealed preferential associations between varieties and pathotypes that were not accounted for by the known compatibility relationships based on avirulence-resistance interactions. The strength of these associations is quantified in the model by parameter $\alpha_{v,p}$ (see equation (2.1a)), which can be interpreted as an indication of the aggressiveness level of a pathotype on a variety. Two compatible (virulent) pathotypes may have very different $\alpha_{v,p}$ values on a variety, as in the case of 073100 ($\alpha_{Soissons,73100} = 2.2$, $IC_{95\%} = [1.2, 3.5]$) and 106314 ($\alpha_{Soissons,106314} = 0.057$, $IC_{95\%} = [0.0023, 0.30]$) on *Soissons*. This means that, although 106314 is fully compatible with *Soissons* according to the gene-for-gene relationship, it exhibits a low aggressiveness on *Soissons* in the field, whereas 073100 appears as a very aggressive pathotype on that variety. More generally, criterion C_1 indicates that 073100 was largely dominant over all other compatible pathotypes on *Soissons* and on *Isengrain* and, reciprocally, criterion C_2 indicates that 073100 was mainly found on these two varieties. This pathotype thus appears as an aggressive specialist. On the opposite, according to the same criteria, pathotype 077317 was characterised as a generalist, with no preference for any of the considered varieties.

It then appears that quantitative effects that can be linked to the host quantitative

resistance and the pathogen aggressiveness level had a major effect on the leaf rust population structure in France between 1999 and 2008. Such interactions should be taken into account for designing varietal allocation strategies but they are generally not documented. In a recent study, Skelsey et al. (2009) measured aggressiveness traits (Pariaud et al., 2009b) of two *Phytophthora infestans* isolates on five potato varieties in order to parametrise a spatiotemporal model of potato late blight epidemics. This approach made it possible to obtain specific values for different quantitative parameters underlying the host-pathogen relationship but was restricted, for practical reasons, to a set of isolates and varieties that was not necessarily representative of the field populations. Our approach can be seen as complementary since it did not provide specific values of aggressiveness traits but globally identified the associations between pathotypes and varieties at the landscape scale. Note however that *P. triticina* has a clonal population structure (Goyeau et al., 2007), which makes easier the establishment of associations between pathotypes and host genotypes.

Even though the basic affinity (in terms of our model) between a pathotype and a variety strongly influenced the pathogen population structure, its effect was modulated by that of the other varieties present in the landscape. Indeed, the analysis of the link between the pathogen population composition and the varietal frequencies in the cultivated landscape showed that a variety, by increasing the population size of a pathotype, may significantly influence the composition of the pathogen population on other varieties. This effect is quantified in the model by parameter β (see equation (2.1a)). For example, the variety *Caphorn* is expected to have a strong influence on the presence of pathotype 106314 on the other varieties. This can be seen in figure 2.7, where the presence of 106314 on *Soissons*, *Apache* and *Orvantis* is mainly accounted for by the frequency of *Caphorn* in the landscape. The population size of a pathotype on a variety thus resulted from its aggressiveness level on that variety as well as from inoculum produced by other varieties, and the model was able to characterise both effects.

Another major conclusion of our analysis is that the observed resistance level of a variety could be linked to the composition of the pathogen population, which itself depended on the landscape composition (as seen above). The link between the expected disease severity on a variety and the pathotype frequencies is quantified in the model by parameter b (see equation (2.2a)). It is usual to observe a resistance breakdown when a major resistance gene is overcome by a new pathotype. In the case of quantitative resistance, a gradual decline is expected when the frequency of aggressive individuals increases in the pathogen population. A comparison of both situations can be found in Mundt (2002).

Some varieties maintained a fairly constant observed level of resistance over the period considered (figure 2.2). For example, the fact that no pathotype presented a marked affinity for *Apache* or was linked to its observed resistance level is consistent with a high level of

quantitative resistance in this variety. The situation was different for *Orvantis*. This variety was released in 2000 and it exhibited a decreasing level of observed resistance until the end of the period studied (figure 2.2). This decrease could be attributed to pathotype 106314 and, to a lesser extent, 006xxx and 077317. Both 106314 and 006xxx were sensitive to changes in *Caphorn*'s frequency and it seems that the influence of *Caphorn* on the pathogen population (see figure 2.7c, d and f) contributed to the decline in *Orvantis*' resistance.

Resistance breakdowns are frequent in crops, but it is less common to observe an increase in the observed resistance level of a variety. This is, however, what happened here for *Soissons*. The disease scores obtained on *Soissons* were strongly linked to the frequency of 073100 in the pathogen population (see parameter b in table 2.5) and, therefore, the relative decline in 073100 accounts for the increase in *Soissons*' observed resistance level. After 2002, 073100 was partly replaced by 077317 and 106314, both of which produce susceptible-type lesions on *Soissons* (Goyeau et al., 2006) but appear to be much less aggressive than 073100 in the field (table 2.5) and, as a result, probably do not cause severe epidemics on that variety. The increase in the frequency of 077317 and 106314 on *Soissons* can be attributed to the influence of other varieties (figure 2.7b and c). The decrease of 073100 on *Soissons* followed the decline in *Soissons*' frequency after the mid-90s and was due to the fact that this pathotype was not compatible with other varieties, except *Isengrain*. In recent years, *Soissons* has been gradually rated as more and more resistant by extension services. We were able to establish here that this increased resistance did not result from a global decrease of the virulent population but was linked to the frequency of a single highly aggressive pathotype.

2.4.3 Limits

In order to build and estimate the model, it was assumed that the pathotype frequencies were independent between years and that the geographical distribution of the pathotypes was homogeneous at the scale of the country. Given that we worked at the country scale and with data collected at the end of the epidemic season, it is reasonable to assume that the effect of the current year on the pathogen population structure was predominant over the potential effect of previous years. A preliminary exploration of the datasets supported this hypothesis. The assumption of a homogeneous pathotype distribution at the scale of the country was consistent with the population structures described for leaf rust (Goyeau et al., 2006, 2007), and the fact that rust spores are dispersed over large distances (Park et al., 2000). We also assumed that the five main pathotypes considered in the study were pre-existing and homogeneously distributed at the country scale. Theoretical models predict that local structures in the host population could be crucial in the invasion dynamics of pathogens (Keeling, 1999; Park et al., 2001; Débarre et al., 2009). A possible improvement of the analysis would be to test whether weakening the assumption of complete mixing of

the population at the national level would alter the results.

Another limit of the study was due to boundary effects in the datasets. We worked on a sequence of ten consecutive years, which was long enough to capture major changes in the host and pathogen populations. Nevertheless, the beginning of the period may have been influenced by what happened before 1999. Moreover, the model cannot estimate the parameters for varieties that were introduced too late towards the end of the period since this estimation requires a certain amount of information. In particular, a decrease in the landscape representativeness occurred in 2007 and 2008 due to the rise of a new variety, *Sankara*, which represented 8 to 9% of the wheat landscape those last two years. This variety bears the same resistance genes as *Caphorn* and *Orvantis*. It is susceptible to pathotype 106314 and high disease score values were recorded on *Sankara* in the field severity assessments. It is therefore likely that this variety played a role in the multiplication of pathotype 106314, and it is possible that the large influence on 106314 attributed to *Caphorn* was partly overestimated if the effect of *Sankara* was confused with that of *Caphorn*. Based on the existing dataset, the specific effect of *Sankara* cannot yet be properly assessed by the model because the introduction of this variety in the system is too recent.

2.4.4 Conclusion

The approach developed here provided documented situations and information that can be used with landscape epidemiology models for designing variety management strategies. It also made it possible to identify major effects that have to be taken into account in the simulation of large-scale epidemics. Based on parameters α and β, interaction groups can be defined that account for both qualitative (gene-for-gene) and quantitative host-pathogen interactions in a landscape: pathotype 073100 was clearly related to *Isengrain* and *Soissons*; 006xxx was influenced by *Caphorn* and *Orvantis* and 106314 by *Caphorn*; 016206 was influenced by *Trémie* with a moderate influence of *Charger*; 077317 appeared as a generalist and was not linked to a specific variety. This pattern can be linked to ecological specialisation (Devictor et al., 2010). In that field of research, species specialisation indices are commonly calculated at the landscape scale, *e.g.*, to understand the impact of human activities on the structure of natural communities (Clavero and Brotons, 2010). The present study suggests that such specialisation indices could be relevant in plant epidemiology to identify pathotype preferences in a heterogeneous host population and to better understand and predict pathogen population dynamics over the years in cultivated landscapes.

Part II

Population dynamics in agricultural landscapes

Foreword to part 2

As shown in chapter 2, landscape composition has a great influence on the genetic structure of wheat leaf rust populations as they can be observed in the fields. In order to build the model and estimate its parameters, it was assumed that the geographical distribution of the pathotypes and the wheat varieties were homogeneous at the scale of the country. This is quite a reasonable approximation for wheat leaf rust, because it has a large dispersal range. However, when dispersal is limited, spatial structures of the host population are crucial for the spread of epidemics and they determine the strength of competition between pathogen strains. In this second part, we focus on the role of host spatial structures on the demo-genetic dynamics of the pathogen population. For that purpose, we developed and analysed a simulation model of epidemics spread in agricultural landscapes.

In chapter 3, we present the modelling framework, which is situated at the interface between landscape ecology and metapopulation modelling. It is based on three components: the representation of agricultural landscapes, the computation of dispersal rates from an individual dispersal function and the life cycle of the organism under investigation. It is also based on the constant search for complementarity between analytic and simulation-based approaches. A study on epidemics thresholds illustrates this continuum between mathematics and simulations.

Chapter 4 addresses the relationship between spatial structure of host populations and epidemics spread over the landscape. We consider here that the pathogen population has no genetic structure, *i.e.* the pathogen population is composed by one strain only. The landscape is composed of two varieties whose deployment is controlled through their proportions and spatial aggregation.

Chapter 5 addresses the relationship between spatial structure of host populations and the pathogen population composition. In this chapter the pathogen population is composed by three strains that live in a two-host landscape. Two of the pathogen strains are considered as specialists, *i.e.* they have a better fitness on one of the hosts. The third strain is a generalist that perceives the landscape as homogeneous.

Chapter 3

Integrated modelling of population dynamics in an agricultural landscape

This chapter is based on an article project by Julien Papaïx, Katarzyna Adamczyk, Annie Bouvier, Suzanne Touzeau, Kiên Kiêu, Christian Lannou and Hervé Monod

Contents

3.1 Introduction

As advocated by Tilman (1999), consideration of the principles governing ecosystems will provide precious insights to face the challenge of agriculture to improve productivity and sustainability while decreasing its environmental impact. It will require shifting the scale of crop protection investigations and strategies from the field to the agricultural landscape (Plantegenest et al., 2007). Thus the concepts of landscape ecology, which focuses on interactions between spatial patterns and ecological processes (Burel and Baudry, 2004; Turner, 2005), will be essential to build a new paradigm for agriculture with consequences on crop protection strategies.

Modelling processes that occur at the landscape scale is gaining more and more attention from theoretical ecologists as well as agricultural managers. With (2002) studies how spatial pattern, such as habitat fragmentation or resource distribution, affects the various stages of the invasion process at the landscape scale. Loreau et al. (2003b) propose the meta-ecosystem concept as a theoretical framework to study the dynamics and functioning of ecosystems from local to global scales. In the agricultural context, several models were developed in order to study the risk of gene escape from genetically modified crops (for example Colbach et al., 2001a,b; Angevin et al., 2008) and to better control plant diseases through variety management at the landscape scale (see for example the SIPPOM model for phoma stem canker of Lô-Pelzer et al. (2010)).

Among these models, three approaches for spatial ecology can be identified (Hanski, 1998). First, theoretical ecologists usually describe space as a simple continuum or as a discrete lattice, or even as a 1D-environment. On the opposite, landscape ecologists consider a very high level of details of the landscape physical structure. While the first approach lacks applicability for managers, the second one lacks a sound theoretical basis (Seppelt et al., 2009). The third approach is the metapopulation one (Levins, 1969) and Hanski (1998) introduced it as an intermediate between the first two.

In a model based on a metapopulation structure, the landscape is considered as a set of habitat patches, linked by dispersal and immersed in an unsuitable environment with respect to the organism under investigation. A strong theoretical background has been developed in such a context about population persistence and species coexistence (Tuljapurkar and Caswell, 1997; Caswell, 2001). However, classical approaches based on a given metapopulation structure may prove too schematic to cape with agricultural landscapes. Indeed, landscape structure needs to be addressed more explicitly. Moreover, one should be able to vary its main features in order to better study their influence on the agro-ecological processes. The landscape must thus be seen as another model entity, made of a patron around which variability can be added.

In this article we propose a modelling framework in order to fill the gap between landscape ecology and metapopulation modelling that takes advantage of theoretical results developed in the metapopulation context while considering much more realistic landscapes. In the following, we first expose how we represent an agricultural landscape. Then, we explain how to model dispersal in order to couple such landscapes with classical matrix models for population dynamics. Finally we propose a case study in order to show how theoretical analysis and simulation approaches can be matched in a coherent way.

3.2 Modelling framework

3.2.1 Environmental heterogeneity of the agricultural landscape

Following (Turner and Gardner, 1991; Forman, 1995), landscape is defined here as a 'spatially and/or temporally heterogeneous area' at any scale relevant to the ecological process or organism under investigation. Note that on the contrary to Gustafson (1998) we make a difference between spatial heterogeneity and temporal heterogeneity by identifying three cases (figure 3.1): heterogeneity can be either purely spatial or purely temporal or spatio-temporal.

This section deals with purely spatial heterogeneity. We first define heterogeneity and how it is quantified. Then, we describe a landscape simulator which involves two steps (Le Ber et al., 2009): the simulation of the physical landscape structure (*i.e.* the field pattern) and of the qualitative landscape structure (*i.e.* spatial repartition of varieties).

3.2.1.1 Quantification of spatial heterogeneity

Ecologists distinguish the environmental (or abiotic) heterogeneity from the biotic heterogeneity (Melbourne et al., 2007). Environmental heterogeneity deals with variation in the physical environment, whereas biotic heterogeneity deals with variation in the occurrence and abundance of organisms. We focus on environmental heterogeneity, including that due to the host varieties.

Li and Reynolds (1994, 1995) based their definition of spatial heterogeneity on two components: the factor of interest and its variability in space. The factor of interest could be either continuous (like temperature) or discrete or categorical (like habitat types). The same authors and then Gustafson (1998) proposed to quantify heterogeneity through non-spatial components (composition of the environment) and spatial components (configuration of the environment). For a categorical factor, the non-spatial component includes the categories present and their quantities. Indices such as the number of categories, their

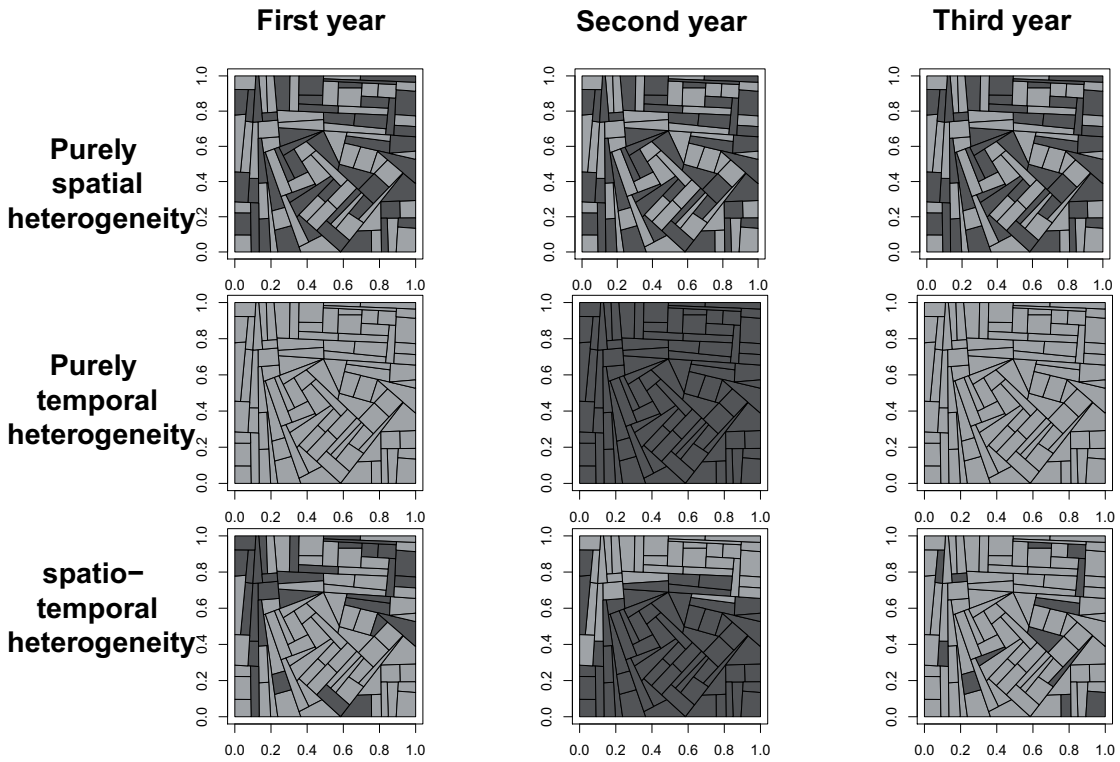


Figure 3.1: Purely spatial heterogeneity (top line) deals with landscapes that do not present temporal variations in crops. In landscape where heterogeneity is purely temporal (middle line), crops change each year but a given year only one crop is cultivated. Finally, several crops could be cultivated a given year and the landscape composition and structure could change every year what leads to spatio-temporal heterogeneous landscapes.

proportions, dominance and differentiation levels provide a quantification of this component (O'Neill, 1988; Riitters et al., 1995). The spatial component relates to how categories are dispatched in space. Several indices were developed in order to quantify the category connectivity based on the geometry of aggregates or their aggregation level, taking or not into account the differentiation level between categories (O'Neill, 1988; Riitters et al., 1995; He et al., 2000; Ahlqvist and Shortridge, 2010).

These indices are called Landscape Pattern Indices (LPI) or landscape pattern metrics in landscape ecology. Much bibliography has been dedicated to devising new LPIs and to avoid some pitfalls (Tischendorf, 2001; Li and Wu, 2004), due to the high correlation between LPI (Gustafson and Parker, 1992; Fortin et al., 2003) and the fact that they quantify a complex combination of several heterogeneity components (Li and Reynolds, 1994; Peng et al., 2010).

In an agricultural landscape, a major source of heterogeneity is varietal diversity and this is a categorical factor. In addition, varieties are dispatched among a discrete number of fields that have well-defined geometrical shapes. The framework we present just below relies on appropriate algorithms and on relevant metrics to control the results when generating agricultural landscapes. A first series of metrics characterises the field pattern. They include non-spatial indices (number of fields, average area, area variance, shape indices) and possibly spatial indices related to area or shape aggregation for example. A second series of metrics relates to variety allocation and includes again non-spatial components (number of categories, proportions) and spatial components such as the level of aggregation between fields with the same variety.

3.2.1.2 Field pattern simulation

This section is essentially based on a unpublished work realised by Katarzyna Adamczyk and Kiên Kiêu from INRA-MIAJ laboratory.

A field pattern may be seen as a polygonal meshing of a plane (Okabe et al., 1992). Several methods are currently used such as Voronoi tessellation or repetition of a particular polygon. In the first case, space is partitioned using a set of points, or seeds. For each seed, a polygon is defined as the part of space closer to that seed than to any other seed. In the second case, one or several polygons are repeated in order to make a partition of space. Le Ber et al. (2006) and Adamczyk et al. (2007) compared these approaches in the idea to generate field patterns. The Voronoi tessellation method produces convex polygons with a high number of vertices whereas the second method produces highly regular paving. Thus, neither method was relevant for modelling agricultural landscapes.

In a polygonal tessellation, vertices are defined as the meeting points of two or more

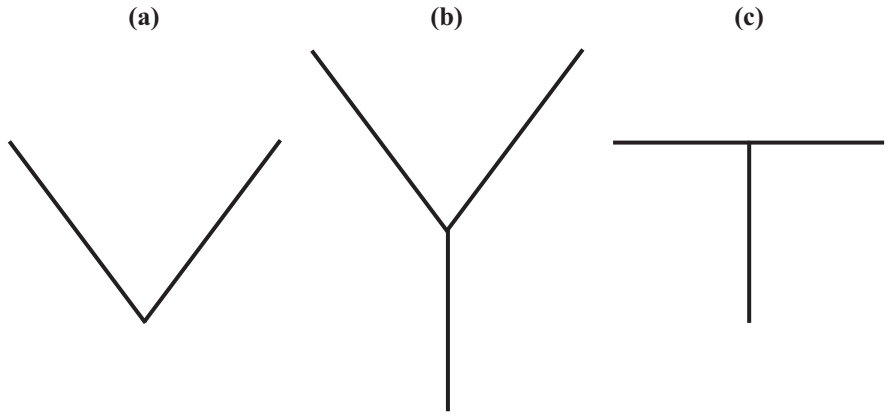


Figure 3.2: V, Y and T vertices. **(a)**: V vertex (*i.e.* degree 2 vertex). **(b)** and **(c)**: two kinds of degree 3 vertices, Y vertex **(b)** and T vertex **(c)**

segments. The degree of a vertex is its number of incident segments and it is equal to 2 or more. A degree 2 vertex is named a V vertex (figure 3.2). A degree 3 vertex could be either in Y or in T (figure 3.2). A T vertex has two of its incident segments aligned and, among all the existing tessellations, they seem to be the most relevant for modelling field patterns. Arak and Surgailis (1989) and Arak et al. (1993) proposed a T-tessellation model that defined a probability distribution on the T-tessellations space by using an energy function. The limit of the Arak's model is that it produces highly irregular tessellations.

Based on Arak's model, a T-tessellation algorithm was constructed by inserting in the energy function terms that make it possible to control key LPIs of the field pattern. A Metropolis-Hasting algorithm was used in order to explore the T-tessellation space using three geometrical operations to generate a new tessellation (figure 3.3). The split operation consists in splitting one polygon into two polygons. The merge operation consists in suppressing a non-blocking segment. Finally, during the flip operation a blocking segment is extended and one of the aligned segments is suppressed.

In practise, field pattern was controlled through three LPIs: the number of fields, the field area variability and the square like form of fields. The first two LPIs are biologically important because they control the frequency of fields with very extreme (small or high) area. Between two different agricultural landscapes, field size could vary greatly but, within the same agricultural landscape fields are generally close to rectangular and homogeneous according to their area. The three LPIs described above allow to quantify how much the simulated landscape is far from this situation. In order to generate field patterns with contrasted LPIs, three terms were added to the energy function: the number of segments, the sum of square polygons areas and the sum of the acute angles. Figure 3.4 shows landscapes for which one or several indices are controlled. Moreover, the stochasticity of the algorithm

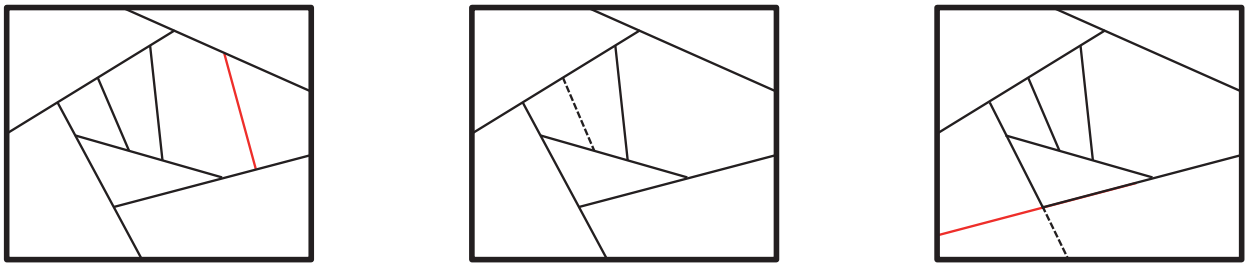


Figure 3.3: Elementary operations of the T-tessellation algorithm. From the left to the right: split, merge and flip. Black solid lines: unmodified segments. Black dashed lines: removed segments. Red solid lines: added segments.

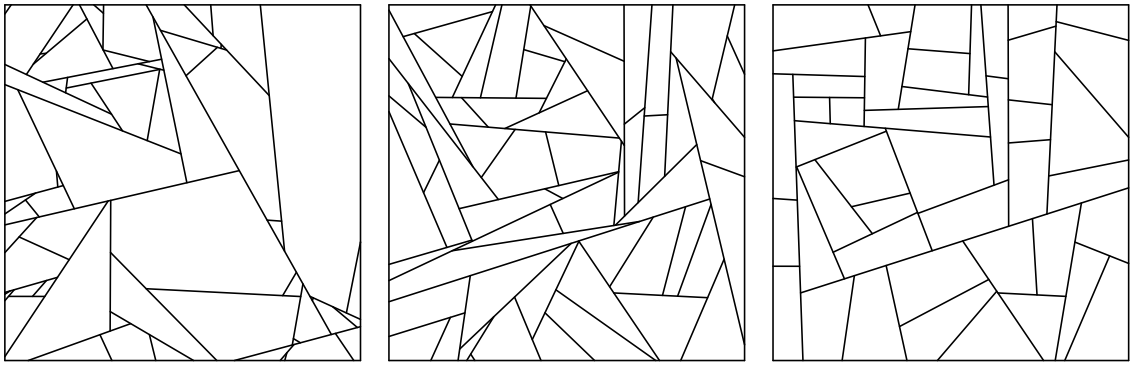


Figure 3.4: Simulated field patterns with controlled LPIs. From the left to the right: the number of fields is controlled, number of fields and variability are controlled, the three LPIs (number of fields, area variability and square-like form) are controlled.

makes it possible to produce several field patterns having the same LPIs (figure 3.5).

3.2.1.3 Variety allocation

The landscape qualitative structure was described by two non-spatial indices, the total number of varieties and their respective acreage proportions, and one spatial index, the aggregation level of varieties.

The aggregation level of variety k, AI_k, is defined as the mean proportion, over the landscape, of neighbour fields that share the same variety:

$$AI_k = \frac{\sum_{i,v(i)=k} \nu_{ik}}{\sum_{i,v(i)=k} N_i},$$

where $v(i)$ is the variety present on field i, ν_{ik} is the number of fields that share variety k in the neighbourhood of field i and N_i is the total number of fields in the neighbourhood of

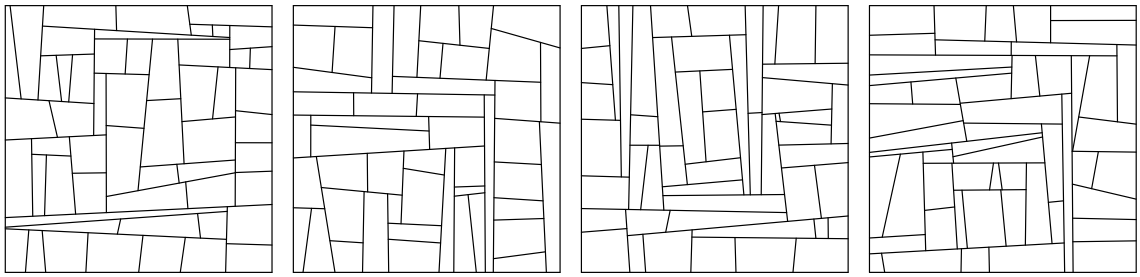


Figure 3.5: Examples of field patterns having the same LPIs. Four field patterns of around 50 fields are presented. They were simulated by four independents MCMC chains starting with four distinct initial values.

field i. This aggregation index is highly correlated with variety proportions. In the case of regular tessellation, the aggregation index of He et al. (2000), AI'_k, can be used for limiting the correlation between aggregation and proportion.

The allocation of varieties to fields was performed by constrained optimisation using a simulated annealing algorithm (Kirkpatrick et al., 1983). The energy function incorporated a first term that controls variety proportions and a second one that controls variety aggregations. Some examples of variety allocations are given in figure 3.6. Note that it also exists advanced software that simulates crop allocation over years (see Castellazzi et al. (2010) for an example on the *LandSFACTS* software).

3.2.2 Ecological heterogeneity and dispersal

3.2.2.1 Management scale and dispersal range

Fields are the management units but the population dynamics and dispersal may occur at a smaller scale. As an example Soubeyrand et al. (2007) studied the spread of yellow rust of wheat (*Puccinia striiformis* on *Triticum aestivum*) and estimated that spores were deposited up to 150m from their source. In the following the term 'field' still denotes the management unit. In addition, we define the patch as a finer geographical unit homogeneous both with respect to dispersal and to habitat type. The patch scale can be considered as quite realistic or it can be considered as a discrete approximation to continuous space. The important point is that the population is assumed to be perfectly mixed within each patch.

Patches can be generated by using 2D-meshing methods among which two are exposed here. The first one is the Delaunay tessellation (Persson and Strang, 2004). The Delaunay tessellation consists in paving each field with triangles whose circumscribed circle does not contain any vertex of another triangle. The second method consists in defining a regular tessellation whose elementary polygon area is adapted to the dispersal function and to define

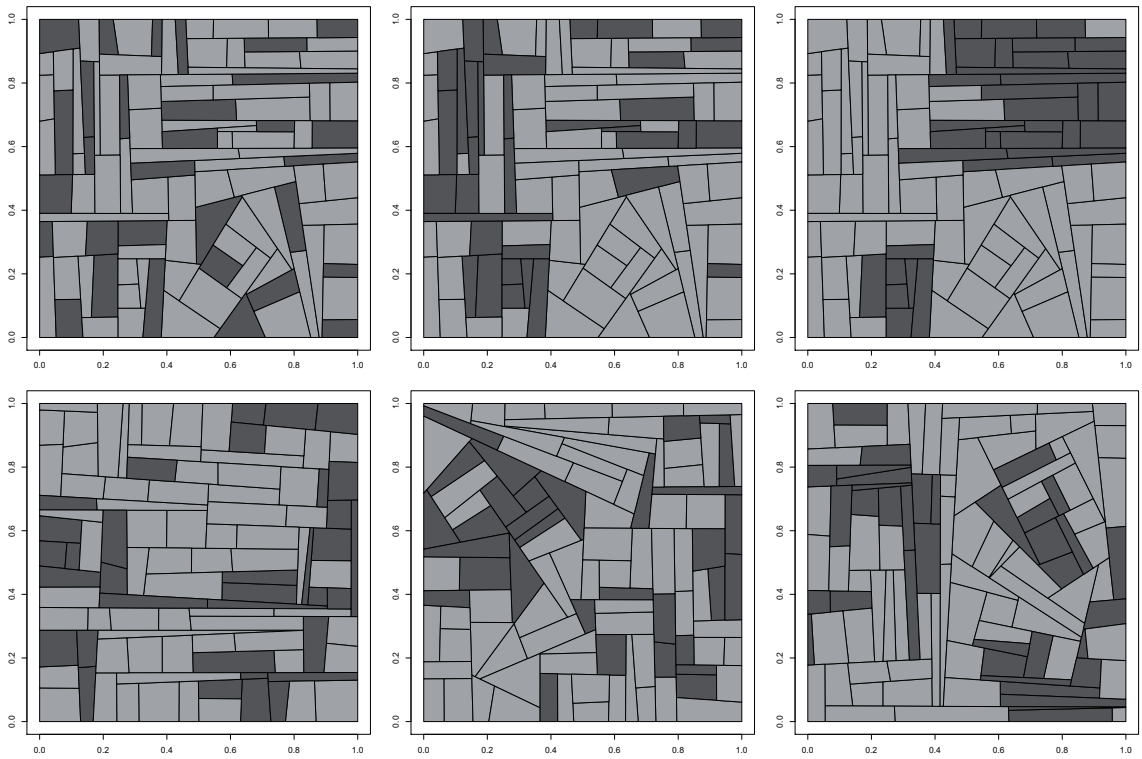


Figure 3.6: Examples of simulated landscape structures with two varieties (light (70%) and dark (30%) grey) dispatched among 150 fields (around). The top line shows an increasing variety aggregation level from left to right but the field pattern remains unchanged. The bottom line shows three field patterns with the same level of variety aggregation.

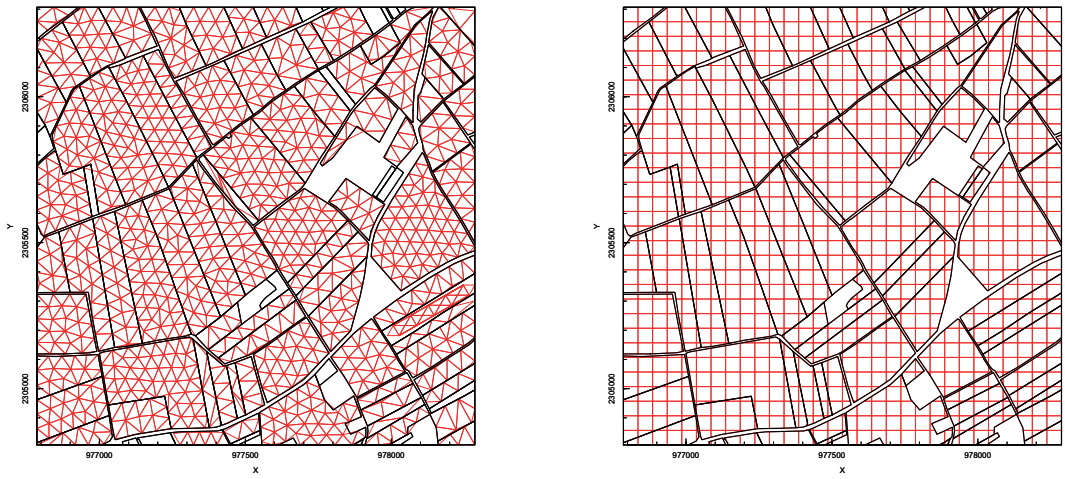


Figure 3.7: Real field patterns drawn together with patches. In the left panel patches are determined using the Delaunay methods whereas in the right panel intersection with a regular tessellation is used.

patches as the intersection between each field and each elementary polygon. Both methods are illustrated in figure 3.7. The Delaunay tessellation becomes problematic for non-convex fields due to the difficulty to fill the field in its entirety (figure 3.8). Nevertheless, it gives patches that have roughly the same size, which is not the case with the second method (figure 3.8). Besides, the Delaunay tessellation is longer to compute than the intersection with a regular tessellation.

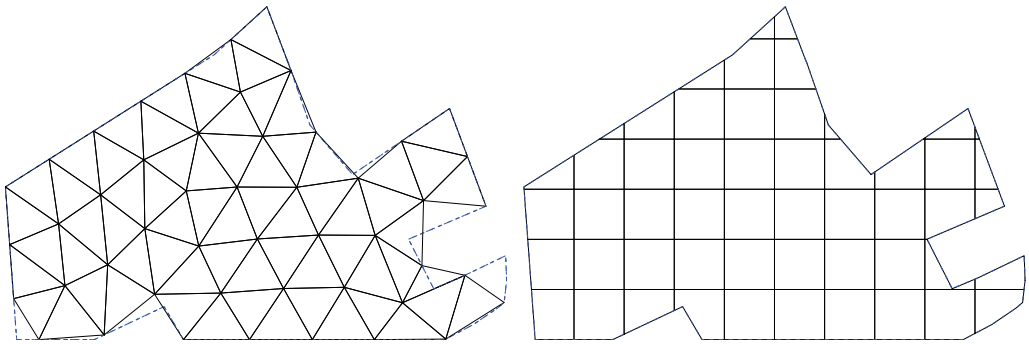


Figure 3.8: The left panel shows Delaunay tessellation on a non-convex field, the right panel shows the same field intercepted with a regular tessellation. Blue dashed line: field, black solid line, patches.

3.2.2.2 Individual dispersal function

Dispersal is the spatio-temporal component of the spread process. Most spatially explicit models for population dynamics consider migration as a reaction-diffusion model (Débarre et al., 2009; Sapoukhina et al., 2010). Another more realistic way for modelling dispersal at the landscape scale is the use of an individual dispersal function (IDF). An IDF is defined as the probability density of the deposit position, (x, y), of a propagule emitted from a punctual source in $(0, 0)$. In the case of air-borne dispersal, IDFs are classified in three categories. Empirical models are made of a parametric probability density function based on the observed data. Quasi-mechanistic models take into account the major atmospheric mechanisms. Mechanistic models describe the physical movement of a particle in the atmosphere with parameters that have a physical meaning.

In this work, we use empirical models because they involve less parameters and take less time to compute. Several empirical functions are classically used (Tufto et al., 1997). They differ by the weight that they give to far dispersal events. This is an important characteristic to take into account because the form of the distribution tail determines the population spread (Neubert and Caswell, 2000; Mundt et al., 2009). In fact, Neubert and Caswell (2000) have found that when dispersal contains both long- and short-distance components, it is the long-distance component that governs the invasion speed, even when long-distance dispersal is rare. In addition, such empirical models may also involve anisotropies in distance or density (Soubeyrand et al., 2007).

3.2.2.3 Dispersal rates computation

Let $g(z, z')$ denote the individual dispersal function between spatial points z and z'. The dispersal rates between patches is obtained by integration according to the formula

$$m_{ij} = \int_{\mathcal{A}} \int_{\mathcal{B}} g(|| \, z - z' \, ||) \, dz' \, dz. \tag{3.1}$$

The integration is performed between pairs of points that belong to the area $\mathcal{A}$ and $\mathcal{B}$ of patches i and j, respectively. Note that when using equation (3.1), the implicit assumption is that the population mixes perfectly in each patch. By computing m_{ij} over all pairs of patches, we obtain now a dispersal matrix that could be used in classical matrix population models.

The dispersal rates m_{ij} in equation (3.1) were computed using the CaliFloPP algorithm (Bouvier et al., 2009). This algorithm computes the integral of a point-wise dispersal function between any pair of source and target polygons. Bouvier et al. (2009) showed that

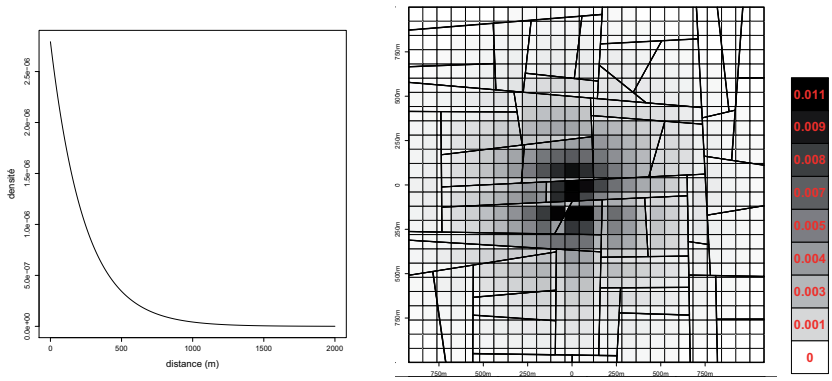


Figure 3.9: Dispersal function (left panel) and dispersal rates (right panel). The following dispersal function was used: $g(|| z - z' ||) = \frac{2\pi}{m_0^2} \exp\left(\frac{2\pi}{m_0} || z - z' ||\right)$, where $|| z - z' ||$ is the euclidean distance between points z and z', where m_0 is the dispersal range ($m_0 = 150$m). Dispersal rates from the patch situated in $(0, 0)$ were computed from the equation (3.1) using the CaliFloPP algorithm.

the computation of particle fluxes between any pair of polygons can be done by computing the integral of the individual dispersal function on the intersection between polygon i and the translation of polygon j by a displacement vector, $\vec{u}$, for all $\vec{u}$. The CaliFloPP algorithm uses first algorithmic geometry tools in order to decompose, triangulate and intercept polygons then numerical integration tools in order to compute the particle flux. Figure 3.9 displays both the individual dispersal function $g(.)$ and the dispersal rates m_{ij}.

When using integration with CaliFloPP, patch size has an influence on the spread speed computed at the landscape scale, because the assumption of patch homogeneity implies an additional dispersal within the patch, which is an artifact. This technical issue was not studied in detail during the PhD, but it was checked that the patch size was sufficiently small to avoid strong biases. As an illustration, figure 3.10 shows the effect of patch size on the epidemic spread given an exponential dispersal function of range 150m. It shows that, given this dispersal function, the influence of patch size is small provided the polygons of the regular tessellations are smaller than 200m $\times$ 200m.

3.2.3 Sensitivity analysis at the landscape scale

In this PhD, modelling is not an end in itself, but a means to acquire better understanding of complex spatio-temporal processes. To reach this objective, modelling must be completed by sensitivity analyses, in order to explore how the model output behaves when parameters or input variables of interest vary (Ginot et al., 2006).

Local sensitivity analysis (or elasticity analysis) is based on the computation of the model derivatives with respect to model inputs of interest (Caswell, 2001). However, this derivative-

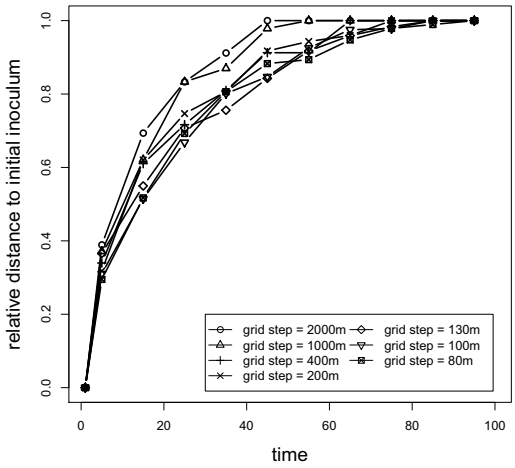


Figure 3.10: The evolution over time of the maximal distance from the inoculum to the epidemic front plotted for several values of the regular tessellation step. Epidemics were simulated using the model describe in section 3.3.2. Each points results from the median computed over 20 simulations.

based approach is unwarranted when the inputs vary in large intervals or in discrete sets, especially when the model is highly non-linear. Consequently, sensitivity analysis in this thesis is based on the global approach advocated by Saltelli et al. (2000, 2008). In addition, it is adapted to include the landscape itself among the factors of interest.

3.2.3.1 Global sensitivity analysis

In global sensitivity analysis, the influence of several input factors is studied simultaneously. Here an input factor designates any model parameter or input variable whose influence on the model output is of interest. Each input factor is given on a discrete or continuous domain of variation, called the uncertainty domain. Several methods are then available to quantify the influence of the factors (Monod et al., 2006; Iooss, 2011).

The variance-based methods are based on a decomposition of the model output variance presented in Sobol (1993) or Saltelli (2000). When the input factors vary independently in their uncertainty domains, the Sobol' decomposition leads to a unique decomposition of the output variance:

$$V_G = V_1 + \ldots + V_n + V_{1,2} + \ldots + V_{n-1,n} + \ldots\ldots + V_{1,\ldots,n}, \tag{3.2}$$

where V_G is the global variance, V_i is the first order effect (or main effect) of input factor i, $V_{i,j}$ is the interaction of order one between factors i and j, and the rest of the decomposition

consists of higher order interactions. When the model is stochastic, an additional term must be added to the decomposition to reflect the average output variance when the input factors are fixed.

The Sobol' sensitivity indices are calculated as the ratios between the variance terms and the global variance. For example, $SI_i = V_i/V_G$ denotes the first order sensitivity index of factor i and $SI_{i,j} = V_{i,j}/V_G$ denotes the interaction sensitivity index between factors i and j. Note that the sensitivity indices belong to the $(0, 1)$ interval, that they sum to one, and that a high index indicates a high influence on the model output. The total sensitivity index of factor i is defined as the sum of all first order and interaction sensitivity indices that involve factor i. It also belongs to $(0, 1)$ and gives a global measure of the influence of factor i.

3.2.3.2 Computations

In practise, the sensitivity indices cannot be calculated exactly and they must be estimated after running simulations. This requires to perform a computer experiment by (i) defining the input factors and their uncertainty domains, (ii) choosing the most appropriate method and the number N of simulations to run, (iii) drawing a sample of size N in the input domain according to the chosen method, (iv) running the N simulations according to the sampling output, (v) calculating the estimated sensitivity indices according to the chosen method.

The sensitivity analyses in the sequel of this PhD include between two and four input factors with either discrete or continuous uncertainty domains. They are usually based on a replicated complete factorial design. Details are given below and inside the chapters.

3.2.3.3 Landscape as a complex input factor

Landscape is characterised by its physical structure (the field pattern, here) and its qualitative structure (the varieties and their allocation). This is a complex input factor for sensitivity analyses and so it requires specific attention. Problems and possible solutions associated with complex input factors are discussed by Iooss and Ribatet (2009). Agricultural landscape structure and composition are integrated explicitly in the sensitivity analyses performed in Viaud et al. (2008), Lavigne et al. (2008) and Colbach et al. (2009). However these studies cope with deterministic models and with factorial designs that involve a small number of levels per factor. In this PhD, we apply the method to stochastic models, as in Lurette et al. (2009). In addition, we allow for a refined exploration of the effects of continuous landscape features, by using more levels per factor and applying the metamodelling technique based on polynomial chaos expansion of model output (Sudret, 2008).

In section 3.2.1, landscapes were described by landscape pattern indices (LPIs) such as the number of fields or the aggregation level of the varieties. LPIs did not define a unique landscape, as stressed by Li and Wu (2004). But they allowed to decompose the landscape description into a part that can be measured and controlled through quantitative variables, and a part that can be considered as residual variability when the quantitative variables are given. Two main questions then arise when we study the influence of landscape on an ecological process. First, how much is the process influenced by the LPIs of interest? Second, how robust are the sensitivity analysis results to landscape residual variability? To answer such questions, we include the LPIs of interest among the input factors of the sensitivity analyses and we perform replications.

More precisely, the LPIs of interest are considered as input factors that can be controlled at the sampling stage of the sensitivity analysis (figure 3.11). For example, the method presented in section 3.2.1 allows to control the number of fields and some of their shape characteristics when generating the field pattern. It is also possible to control the area occupied by different varieties and the degree of aggregation when allocating varieties to fields. Consequently, we consider each landscape as a random realisation around a 'patron' determined by specific values of the LPIs of interest. The sensitivity analyses then integrate the landscape as follows:

1. (main sampling design) the LPIs of interest are fully integrated among the input factors when constructing the sampling design and analysing the results;

2. (landscape replications) for each combination of the LPIs in the sampling design, several distinct landscapes are generated;

3. (model replications) for each element of the sampling design and each landscape replication, several simulations are performed.

This hierarchical structure of the simulations allows to calculate the sensitivity indices associated with all input factors, including the LPIs. It also allows to assess the output variability that can be attributed to landscape residual variability (through landscape replications) and to other sources of model stochasticity (through model replications). The key idea is to decompose the landscape description into a part that can be measured and controlled through quantitative variables, and a part that is considered as residual variability when the quantitative variables are fixed.

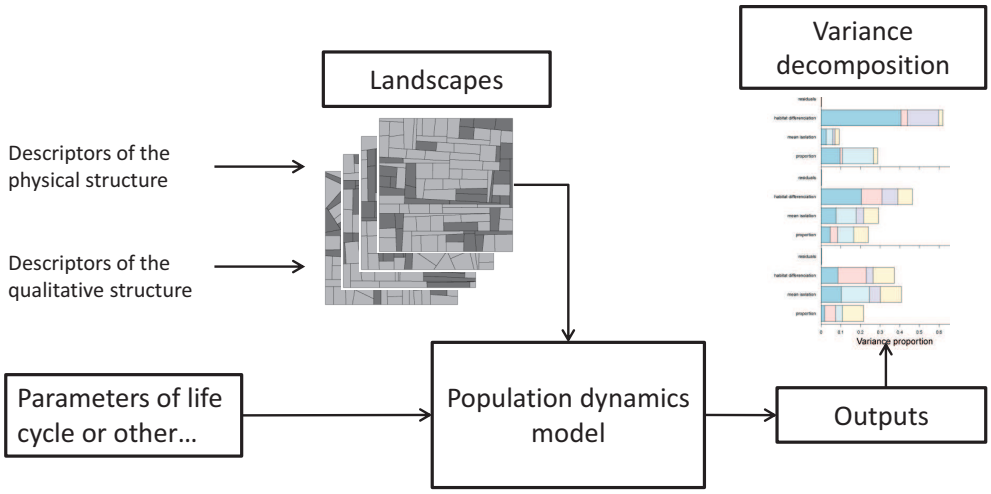


Figure 3.11: Schematic representation of a sensitivity analysis with a landscape as an input factor. Landscape descriptors allow to define a variety of landscapes which are used as an input of the model. The sensitivity of model outputs to landscape descriptors and other parameters could then be done through variance decomposition.

3.3 Spatial thresholds for invasion and simulations of epidemics in agricultural landscapes

In this section, we used the approach developed before and we provide a study that aimed to compare epidemics simulations to thresholds for pathogen invasion at the landscape scale. Classically, the basic reproductive number, R_0, is used in order to predict the occurrence of epidemics (Gilligan and van den Bosch, 2008). R_0 gives the number of potential new infections during the average duration of the infectious period of a single infected host. For the epidemics to occur, the basic reproductive number must be greater than 1.

We developed here a population dynamics model for a pathogen foliar fungus on an agricultural landscape. We used both a deterministic version and a stochastic simulation model. The first one made it possible to define thresholds for invasion at the landscape scale that took or not into account the initial location of inoculum. We show here how spatial structure of landscape heterogeneity influence the epidemics spread and compare analytical to simulation based results.

3.3.1 Landscapes

Let us consider a $2000 \times 2000\text{m}^2$ agricultural landscape composed of around 150 fields. Two wheat varieties are cultivated in equal proportions (50% of the total acreage for

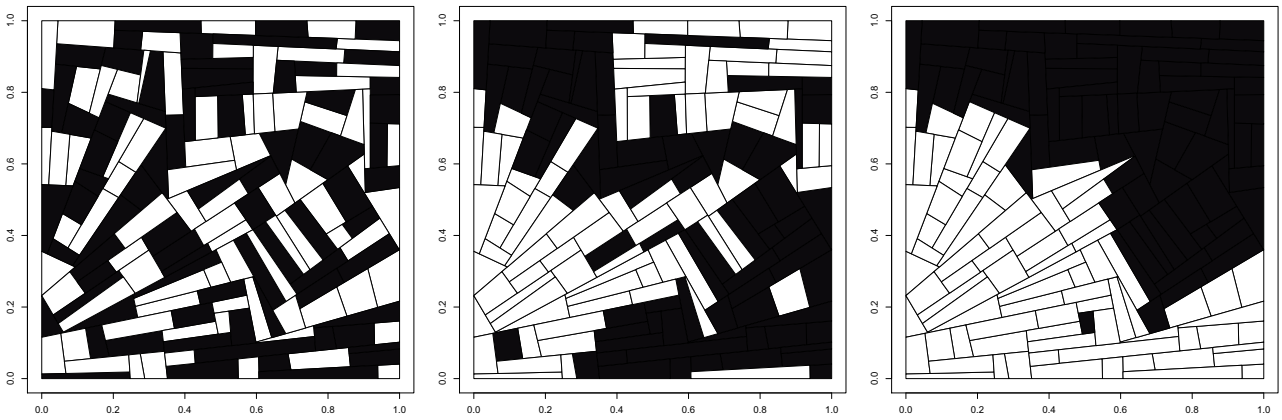


Figure 3.12: Two varieties are dispatched in equal proportions but with an increasing aggregation level from the left to the right (mixed, mosaic and grouped).

each variety) but with three different allocation strategies: mixed, mosaic and grouped (figure 3.12). In the first allocation strategy, $AI = 30\%$ of field neighbours share the same variety. The aggregation index AI is equal to 70% for the mosaic strategy and to 90% for the grouped one. In this example, two fields are considered as neighbours if the shorter distance between edges is lesser than 50m. Five field patterns were generated and for each of them two realisations were simulated for each allocation strategy. Thus, for each allocation strategy, ten different landscape structures were generated. Fields were subdivided into patches by intercepting field patterns with a regular grid of $100 \times 100\text{m}^2$ squares.

3.3.2 Model

The model is a stochastic and discrete time HLIR model (Madden et al., 2008) with static infection sites and spores as a propagule state. It describes the evolution of the infection sites in each of the following states: healthy sites (H), contaminated sites (S'), latent lesions (L), infectious lesions (I) and removed sites (R). The carrying capacity of patch i, K_i, is fixed and proportional to the area of patch i. We consider a pathogen population composed by one strain only and only the infection efficiency was considered variety-dependent.

Let $v(i)$ be the host type of patch i. In this patch, a healthy site on patch i is contaminated by a spore with probability $\pi\left(\frac{H_i(t-1)}{K_i}\right)$ (see supporting information, appendix 2). Spores that do not succeed to contaminate an infection site are removed. A contaminated site becomes infected with probability $e_{v(i)}$, the infection efficiency. When infected, the site remains latent during τ days before producing spores. After T days of sporulation, the site is removed. An infectious site produces r spores per day. Spores remain in patch i with probability m_{ii} and

disperse to patch j with probability m_{ij}. The complete description of the model is given in supporting information (appendix 2). Table 3.1 gives the transitions between each state.

The function $\pi(.)$ can take different forms that reflect heterogeneity among sites. When $\pi\left(\frac{H_i(t-1)}{K_i}\right) = \frac{H_i(t-1)}{K_i}$, all the sites have the same probability to be contaminated. On an individual plant, however, not all sites are equally accessible since the plant cover is heterogeneous. Here we considered that the last infection sites are more difficult to reach. In order to render this effect, we used the following sigmoid function for $\pi(.)$:

$$\pi(x) = 1 - \frac{\exp\Big(-\kappa x^{\sigma}\Big) - \exp(-\kappa)}{1 - \exp(-\kappa)},$$

where parameters $\kappa > 0$ and $\sigma > 0$ determined the sigmoid form. This allows to simulate smoother and probably more realistic epidemic dynamics.

The deterministic version of the model describe above is displayed by figure 3.13 and is synthesised by the following system of equations:

$$\begin{cases} \frac{dH_j}{dt} &= -e_{v(j)} \cdot \pi_j \cdot \sum_{i=1}^{N} r \cdot m_{ij} \cdot I_i, \\ \frac{dL_j}{dt} &= e_{v(j)} \cdot \pi_j \sum_{i=1}^{N} r \cdot m_{ij} \cdot I_i - \frac{1}{\tau} \cdot L_j, \\ \frac{dI_j}{dt} &= \frac{1}{\tau} \cdot L_j - \frac{1}{T} \cdot I_j, \\ \frac{dR_j}{dt} &= \frac{1}{T} \cdot I_j. \end{cases} \tag{3.3}$$

Table 3.1: Transitions of the stochastic model on a given patch i.

Description	Notation	Transition
Spore cloud	$S_i(t)$	$r \sum_{j=1}^{N} I_j(t-1) m_{ji}$
Contaminated sites	$S'_i(t)$	$\min\left(\text{Bin}\left(H_i(t-1), \pi_i(t)\right), S_i(t)\right)$
New latent lesions	$H_i(t-1) \rightarrow L_{i,p}(t)$	$\text{Bin}\left(S'_i(t), e_{v(i)}\right)$
New infectious lesions	$L_i(t-1) \rightarrow I_i(t)$	$\text{Bin}\left(L_i(t-1), 1 - e^{-\frac{1}{\tau}}\right)$
New removed sites	$I_i(t-1) \rightarrow R_i(t)$	$\text{Bin}\left(I_i(t-1), 1 - e^{-\frac{1}{T}}\right)$

3.3.3 Invasion thresholds

From the model defined by the system of equations 3.3, two invasion thresholds can be defined. The first invasion threshold, $R_0^{(1)}$, is based on the next generation matrix (van den Driessche and Watmough, 2002; Fulford et al., 2002). The next generation matrix gives

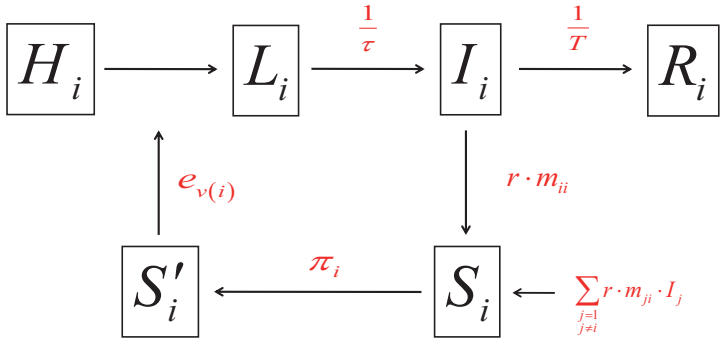


Figure 3.13: Graphical representation of the population dynamics model. H_i: number of healthy sites in patch i, L_i: number of latent sites in patch i, I_i: number of infectious sites in patch i, R_i: number of removed sites in patch i, S_i: number of spores in patch i, S'_i: number of contaminated sites in patch i. τ: latency period, T: infectious period, r: spore production per lesion per day, m: dispersal rates, $\pi(.)$: contamination function, e: infection efficiency.

the number of infection sites in each state and for each patch at time t from the same quantities at time $t-1$. $R_0^{(1)}$ is thus given by the dominant eigenvalue of the matrix whose elements in line i and column j are $Tre_{v(i)}m_{ji}$ (see supporting information, appendix 3). In the general case, there is no explicit formulation of $R_0^{(1)}$ but it can be computed numerically. $R_0^{(1)}$ supposed that epidemics start with an infinitesimal infectious lesion dispatched among all patches.

The second invasion threshold, $R_0^{(2)}$, takes into account the spatial location of the inoculum. $R_0^{(2)}$ was defined using an heuristic proposed by Park et al. (2001): a single infectious lesion in patch i gives birth to $rTe_{v(i)}m_{ii}$ potential new infectious lesions in the same patch i and to $rTe_{v(j)}m_{ij}$ potential new infectious lesions in the patch j. Thus, a single infectious lesion give potentially birth to a total of $R_0^{(2)} = rT\sum_{j=1}^{P} e_{v(j)}m_{ij}$ new infectious lesions.

3.3.4 Simulations context

Epidemics were simulated over 1000 time steps using the stochastic version of the model (Table 3.1). The number of spores produced by a single lesion during one day was set to $r = 2$. Lesions were supposed to remain latent during $\tau = 5$ days before producing spores during $T = 10$ days. On variety V_1, the infection efficiency was fixed to $e_{V_1} = 0.04$ that corresponds to a non spatial basic reproductive number equal to $2 \times 10 \times 0.04 = 0.8$. On variety V_2, the infection efficiency varied from $e_{V_2} = 0.02$ to $e_{V_2} = 0.2$, giving a non spatial reproductive number varying from 0.4 to 4. Epidemics were initialised by 1 infectious lesion

in one V_1 patch chosen at random. Model stochasticity was taken into account by simulating two epidemics with two different starting points on each of the 10 landscape structures of one allocation strategy.

The following isotropic and exponential individual dispersal function was used:

$$g(|| z - z' ||) = \frac{2\pi}{m_0^2} \exp\left(\frac{2\pi}{m_0} || z - z' ||\right).$$

where $||z - z'||$ is the distance between z and z', and m_0 is a range parameter. Dispersal rates among patches were computed using CaliFloPP (see section 3.2.2.3). We consider a short ($m_0 = 150$m) and long ($m_0 = 1500$m)dispersal range.

3.3.5 Results

3.3.5.1 Comparison of invasion thresholds

Figure 3.14 compares predictions by the two invasion thresholds. First, a clear difference appear between $R_0^{(1)}$ and $R_0^{(2)}$ particularly when dispersal is limited. This difference is visible not only in the values of $R_0^{(1)}$ and $R_0^{(2)}$ but also in their predictions of pathogen invasion. For example, $R_0^{(2)}$ decreases when aggregating varieties and it predicts less and less epidemics whereas $R_0^{(1)}$ increases and its predictions remain unchanged. This was due to the fact that we inoculate a V_1 field for which the non spatial basic reproductive number was lesser that one. So, by grouping varieties the epidemics are less probable because the pathogen encounter more often the host for which it is non efficient. Finally, $R_0^{(2)}$ was found highly variable between simulations which is consistent with a determinant role of local spatial structures in the epidemics viability. Differences between $R_0^{(1)}$ and $R_0^{(2)}$ decrease when increasing dispersal range. They are roughly equal and predict epidemics in the same way for mixed strategies. This could be explain by the fact that a mixed strategy along with a long dispersal range context corresponds to the more non-spatial case. Thus, the initial position of inoculum does not import.

3.3.5.2 Comparison between invasion thresholds and simulated epidemics

Figure 3.14 allows to compare invasion thresholds to simulated epidemics. When dispersal range was low, local structures played a leading role and simulated epidemics were globally consistent with $R_0^{(2)}$ predictions. On the contrary, $R_0^{(1)}$ was far from simulations results due to the fact that it does not consider the initial position of inoculum. This is particularly the case when grouped strategies were used. When dispersal range was large, $R_0^{(2)}$ predicted simulation results correctly for grouped strategies but not in the case of mixed ones. Moreover, while $R_0^{(1)}$ and the observed probability of an epidemics to occur increased by aggregating varieties, $R_0^{(2)}$ decreased by aggregating varieties.

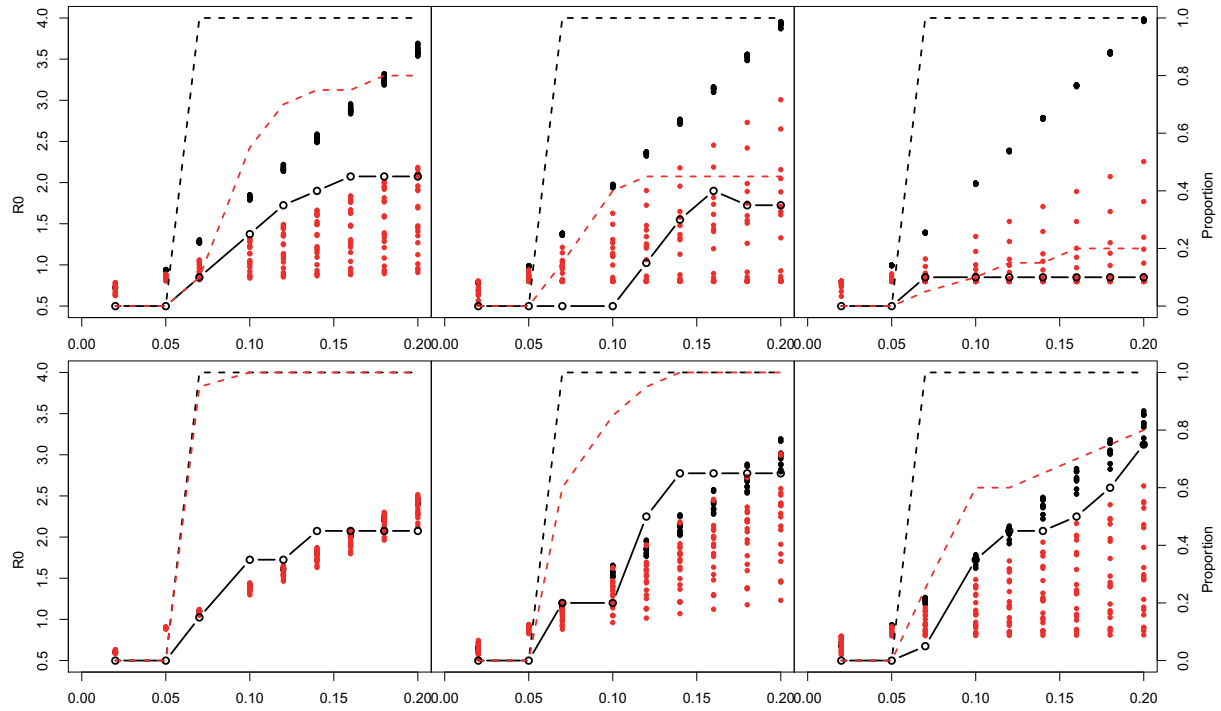


Figure 3.14: Each panel draws $R_0^{(1)}$ (black closed circles) and $R_0^{(2)}$ (red closed circles) values along with their respective predictions (black and red dashed lines, respectively) of pathogen invasion and the pathogen invasion observed on simulated epidemics (open circles) against e_{V_2}. Top line: short dispersal range (m_0 = 150m), bottom line: large dispersal range (m_0 = 1500m). Left column: mixed variety allocation strategy, middle column: mosaic variety allocation strategy, right column: mosaic variety allocation strategy. For each e_{V_2} value, 10 landscape structures × 2 inoculum locations were available.

3.3.5.3 Conclusions on the results

Classically, invasion thresholds are computed using the next generation matrix. This method takes into account the spatial organisation of hosts but not the spatial structure of the pathogen population. The rarity hypothesis from which such invasion criteria are based is difficult to characterise in the context of spatially structured populations: few individuals could be dispatched in several patches or in one patch only. Using the heuristic of Park et al. (2001) (but see Massol et al., 2009) we show here that initial spatial location of inoculum is crucial for predicting pathogen invasion at the landscape scale. The effect of local spatial structures is decreased by increasing pathogen dispersal.

We also compared prediction by invasion thresholds with simulated epidemics. The basic reproductive number that takes into account the spatial location of the inoculum ($R_0^{(2)}$) was more consistent with the simulation study, which reinforces the importance of taking into account of spatial structures in the pathogen population. Moreover, both criteria predicted more epidemics than we observed on simulations with the stochastic model. That is consistent with the fact that stochasticity at the early stage of population growth is an important force to take into account.

In this study, only one pathogen strain was present. Results presented here suggests that spatial structure could be of prime importance when several strains are in competition. In particular for studies on mutant invasion when abundance between pathogen strains is strongly unbalanced. In fact, in such a context stochasticity could play a leading role due to the possibility of foundation effects (see chapter 5).

3.4 Discussion

We have presented an integrated modelling framework at the interface between metapopulation-based models and landscape ecology approaches (figure 3.15). This is an integrated framework because it makes a coherent link between field scale and a finer scale associated with ecological processes. What does interface mean here? Considering more complex landscape structures makes numerical investigation and simulation based methods unavoidable. However, it is necessary to ensure coherence with analytical solutions when they are available. The first step of the model consisted in defining the environment and the habitat distribution. The agricultural context has naturally led to consider space as both continuous and structured by 2D-objects, fields. Then, we used a particular algorithm, CaliFloPP (Bouvier et al., 2009), in order to compute patch-to-patch dispersal rates from an individual dispersal function (IDF). Considering an IDF allows to model migration in a much more realistic way. Finally, from these dispersal rates, any kind of metapopulation

model can be immersed into a realistic landscape.

The CaliFloPP algorithm performs the integration of an individual dispersal function between any pair of polygons. Respectively to raster approaches it insures non biased dispersal rates for three reasons. First, raster methods suppose that dispersal occurs from barycentre to barycentre. However, due to the high non-linearity for short distances of the IDF, dispersal rates can be biased. Second, by integrating the IDF, the geometrical form of polygons is taken into account, which is not the case otherwise. Third, if fields are not contiguous but separated by a gap or a road, this is automatically taken into account.

Field patterns are obtained by using a T-tessellation algorithm that controls the polygon number, their form and their area variability. This approach was chosen in order to obtain field patterns sharing common characteristics but all different from each other. As we explained, such landscape replicates are of prime interest for testing the robustness of results or to study the variability of a model output between a set of landscapes. However, some structures in the landscape can not be generated. With the development of geographical tools another approach would be to search in landscape data-bases landscapes that correspond to the desired LPI.

In plant pathology, the scale at which control strategy is defined has proved crucial for its effectiveness (*e.g.* Gilligan et al., 2007). More generally, the effect of spatial scale is an important issue in landscape ecology (Ricklefs, 1987). The metacommunity concept (Leibold et al., 2004) appears interesting in order to operate the transition into a regional scale model. A metacommunity is defined as a set of local community that are linked by dispersal. In this context, the transition to larger scales would be done by considering a set of landscapes interconnected via dispersal. This hierarchical concept could appear useful to identify the role of spatial heterogeneity at different scales on pathogen invasion at the regional scale (Melbourne et al., 2007).

The issue of data-based inference must also be considered. The modelling framework exposed here could help to find the most influential features in the landscape on the population dynamics and spread in order to guide further statistical investigation or set up experiments (Fortin et al., 2003). In addition, the increasing development of statistical inference when likelihood is not available (Beaumont, 2010; Hartig et al., 2011) offers an interesting framework for parameter estimation at the landscape scale.

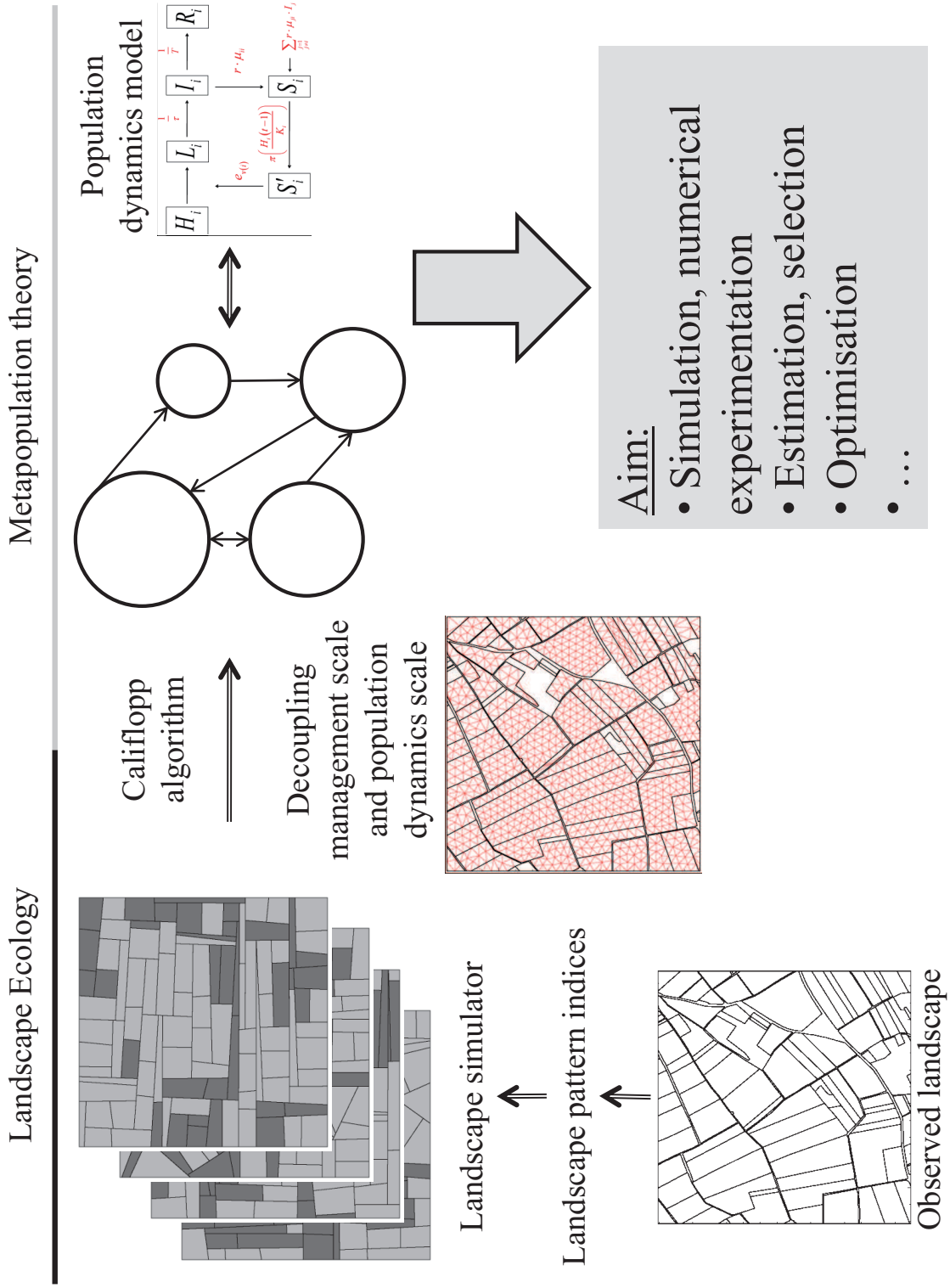


Figure 3.15: Modelling framework.

Chapter 4

Population spread at the landscape scale: How could spatial heterogeneity decrease disease severity?

This chapter is based on a article project by Julien Papaïx, Suzanne Touzeau, Hervé Monod and Christian Lannou

Contents

4.1 Introduction

The strong genetic uniformity of agricultural systems, whatever the scale considered, facilitates the spread of diseases. The threat to agricultural production posed by plant pathogens is such that intensive production relies on massive pesticide inputs. Plant breeding has been a considerable success in terms of yield improvement but has not been able to durably reduce the disease risk, mainly because pathogen have easily adapted to the resistant varieties. Considerable efforts are made in two directions: producing genotypes with a more durable resistance to pathogens and developing intelligent strategies, mainly based on the idea of functional diversity, for using these varieties.

It is now well established that increasing host diversity reduces the rate of disease transmission, and this defines the term 'functional diversity' in the context of plant pathology. This relationship between host diversity and disease transmission is mainly accounted for by 'dilution effects' (Keesing et al., 2006) even though other mechanisms are involved that can be of importance in crops (Chin and Wolfe, 1984; Lannou et al., 2005). Dilution effects are based on the fact that, in a diversified host community or population, even though all individuals are susceptible to one or several pathogen genotypes, each host individual is resistant to a fraction of the pathogen population. Moreover, a host genotype represents a habitat to which the different compatible pathogen genotypes are more or less well adapted. Thus, increasing genetic diversity (for resistance characters) in a host population results in reductions in the pathogen transmission rate that can considerably impact the final epidemic severity (Burdon and Chilvers, 1982; Mundt, 2002). In natural ecosystems, dilution effects participate to the notion of 'insurance hypothesis' that links diversity and resistance to fluctuations of biotic or abiotic variables in the environment (Mitchell et al., 2002; Pautasso et al., 2005).

In plant pathology, the idea of disease management based on mixtures of susceptible and resistant hosts has led to numerous studies and applications (Lannou et al., 1994; Garrett and Mundt, 1999; Zhu et al., 2000; Mundt, 2002). The effectiveness of such a control strategy has been demonstrated for foliar diseases (Browning and Frey, 1969), first without considerations of the host and pathogen spatial distribution. Then, in a series of experimental and simulation studies, Chris Mundt and coll. showed how the host genotypes spatial arrangement and the pathogen distribution largely account for the effectiveness of the mixture strategy (Mundt and Leonard, 1985, 1986; Mundt et al., 1986) and he introduced the notion of genotype unit area (GUA, contiguous area occupied by a single host genotype). More recently, Sapoukhina et al. (2010) showed with a spatially explicit model that, when long dispersal events occur, host mixture hamper disease spread only when susceptible units are separated by a minimal distance whereas for short-distance dispersal of the pathogen, random mixtures are the most efficient. But in a practical situation, the mixture layout

results more of a compromise between theoretical consideration and farming practises (Zhu et al., 2000).

In the context of variety mixtures, the notion of functional diversity only applies at the field scale. It is however strongly suggested that host diversification at a larger scale could be of great interest for reducing epidemic severity in agricultural systems (Mundt and Brophy, 1988; Zhu et al., 2000). Few experimental data are available at such scales (but see Papaïx et al., 2011), which probably explains the low interest of plant pathologists for landscape epidemiology approaches so far (Plantegenest et al., 2007). At the scale of a large production area, the host unit is the field (whether it is sown with a single genotype or a mixture) and the expected dilution effects rely on the landscape connectivity, relative to the pathogen dispersal. In general, habitat connectivity is crucial for the spread of spatially structured populations (Söndgerath and Schröder, 2002; Condeso and Meentemeyer, 2007). At the landscape scale, the connectivity degree measures whether the landscape structure facilitates or impedes movements among habitat patches (Taylo et al., 1993). It depends on four components: (1) the proportion and (2) the aggregation level of habitat types, (3) the ability of the organism to develop on each habitat and (4) to disperse among them (With et al., 1997). Several models in the ecological literature study the effect of habitat connectivity on the spread of an organism at the landscape scale and conclude that habitat structure interacts strongly with the dispersal ability of the organism (Hale et al., 2001; With, 2002). In agricultural landscapes, most of the space is devoted to a few host species or varieties, which makes the connectivity of pathogen habitats very high (Margosian et al., 2009). In this context, a general objective of landscape epidemiology is to evaluate whether this connectivity can be reduced with a significant impact on epidemic severity.

Landscape epidemiology approaches are still scarce in the phytopathological literature. Parnell et al. (2009), for Asiatic Citrus Canker, and Skelsey et al. (2010), for potato late blight, are two recent examples. Parnell et al. (2006) study the spread of fungicide resistant pathogen strains and show that its ability to invade depends on an interaction between the proportion of suitable habitat in the landscape and the intra-field reproductive capacity of the pathogen. Nevertheless, they do not consider explicitly the pathogen dispersal capacities and the spatial distribution of sprayed fields. Skelsey et al. (2010) shows that clustering potato cultivation in some parts of a region enhanced the spread within such a cluster while it delayed spread from one cluster. However they do not specifically consider the effect of differences in the resistance levels of the varieties on the epidemic spread.

In this article, we study how the four above cited components of habitat connectivity determine the spreading capacity of a pathogen population at the landscape scale. To this aim, we developed a spatially explicit model representing the population dynamics of an air-borne foliar disease in an agricultural landscape by coupling tessellation methods (Okabe

et al., 1992) and classical matrix population models (Caswell, 2001). We first present the modelling approach and the analyses performed on the model. Then the results of the simulation study are shown and discussed.

4.2 Materials and methods

4.2.1 Population dynamics model

We describe here an air-borne plant parasite spreading model on a spatially explicit agricultural landscape. The landscape is considered as a set of fields represented by polygons, each containing a susceptible or a resistant variety. Field patterns are obtained with tessellation methods. Varieties of the host plant are allocated to the field with a control of the variety proportions and aggregation levels. In order to describe the pathogen dispersal at a realistic scale, the fields are subdivided into smaller units named 'intra-field patches' and the pathogen disperses among patches, whether intra-field or between-fields. The pathogen population dynamics is simulated with a matrix population model structured by the patches. The dispersal rates among patches are computed from an individual dispersal function.

4.2.1.1 Host environment

The host environment is a $2000 \times 2000\text{m}^2$ region composed of about 150 fields. The field sizes and shapes are defined with a tessellation method that consisted in meshing the square region while controlling both the number of polygons and their geometrical form (see chapter 3). The tessellation algorithm is based on Markov Chain Monte-Carlo (MCMC) methods. Thanks to the stochasticity of the algorithm, several field patterns with the same spatial characteristics (number of polygons and their form) can be generated (figure 4.1).

Since pathogens usually disperse at a smaller scale than the whole field, at the temporal scale considered, each field is divided into patches by intercepting a regular grid of $100\times100\text{m}^2$ squares with the whole simulated area. The patch is then the spatial unit of the system and the pathogen population is supposed to be perfectly mixed at the patch level. The pathogen disperses among the patches and this simulates both intra- and between-field dispersal.

Each field contains either a susceptible (SV) or a resistant (RV) variety (all patches of the same field contain the same variety). The spatial arrangement of SV and RV among fields is described by their proportions and their aggregation level, at the landscape scale. The variety proportions are calculated from their respective total acreage in the landscape. For each field of the landscape, the proportion of neighbour fields (i.e. fields distant from less than 50m) containing the same variety is computed and the average value of this variable

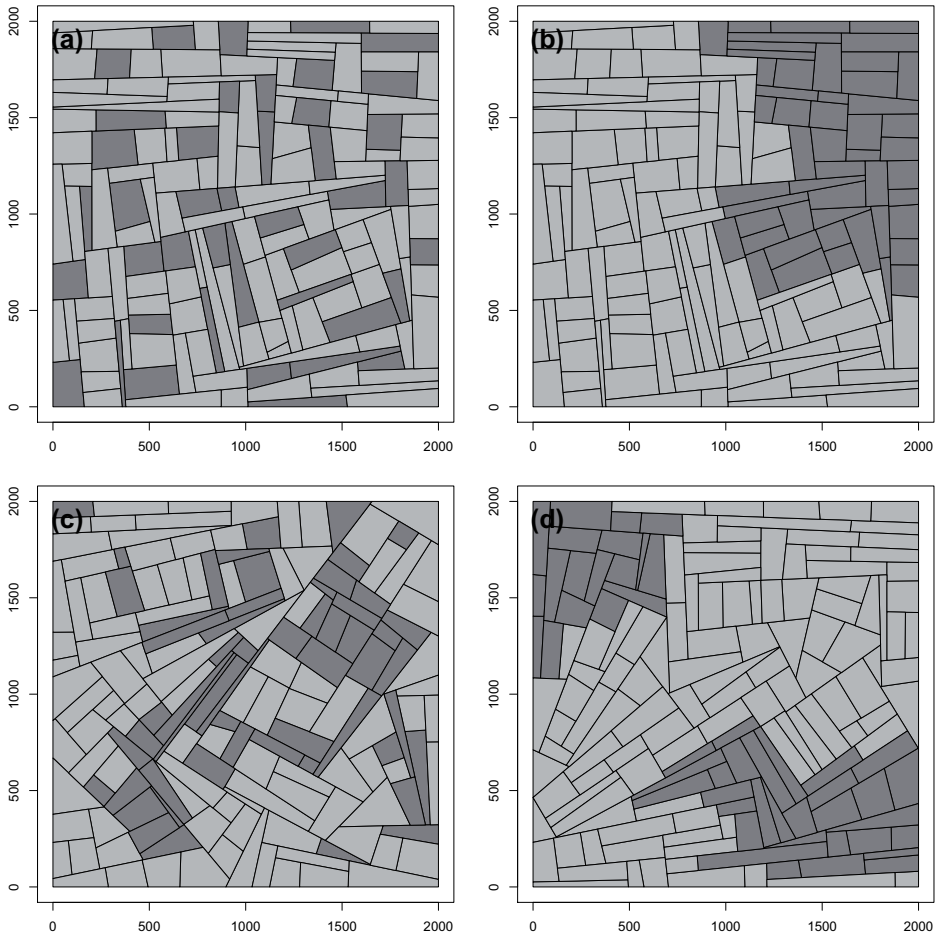


Figure 4.1: Example of landscapes with different variety allocations: a, 30% of RV and mixed strategy; b and d, 30% of RV and grouped strategies; c, 30% of RV and mosaic strategy. Panels a and b have the same field pattern, differing from c and from d.

defines the aggregation level of the varieties. SV and RV are dispatched among fields using a simulated annealing algorithm (Kirkpatrick et al., 1983). It is a stochastic algorithm that performs constrained optimisation and makes it possible to generate pseudo-random variety allocation replicates while controlling their proportion and aggregation level (figure 4.1). Finally, the landscape structure is defined by the field pattern (number of fields and their form), the variety proportions and the variety aggregation level.

4.2.1.2 Within-patch local dynamics

The model of local pathogen dynamics is a stochastic and discrete time HLIR model (Madden et al., 2008) in which spores are the propagule state. Since we assumed that the local pathogen population is perfectly mixed at the patch level, each patch is considered as

a set of infection sites with no spatial structure. The carrying capacity (*i.e.* the number of infection sites) of patch i, K_i, is fixed and proportional to the area of the patch. The model describes the evolution of the number of infection sites in each of the following states: healthy sites (H), contaminated sites (S'), latent lesions (L), infectious lesions (I) and removed sites (R). Figure 4.2 shows a graphical representation of the model.

Let $v(i)$ be the host type (variety) encountered on patch i. In this patch, a contaminated site becomes infected with a probability $e_{v(i)}$ (infection efficiency). When infected, the site remains latent during $\tau_{v(i)}$ days (latent period) before producing new spores. After $T_{v(i)}$ days of sporulation (infectious period), the site is removed. An infectious lesion produces $r_{v(i)}$ spores per day. The spores remain in their patch of origin i with a probability m_{ii} (autoinfection) or disperse on another patch (alloinfection) j with a probability m_{ij}. In patch i, a spore is deposited on a healthy site (and the site becomes contaminated) with the probability $\pi\Big(\frac{H_i(t-1)}{K_i}\Big)$, an increasing function of the proportion of healthy sites in the patch. Spores that do not succeed in infecting a site are removed. The complete description of the model is given in supporting information (appendix 2) and table 4.1 gives the transitions between each state.

The function $\pi(.)$ can take different forms that reflect a certain heterogeneity among sites. When $\pi\Big(\frac{H_i(t-1)}{K_i}\Big) = \frac{H_i(t-1)}{K_i}$, all the sites in a patch have the same probability to be contaminated. On an individual plant, however, not all sites are equally accessible to the spores because the host plant has a certain physical structure. Here we consider that the last available infection sites are more difficult to reach by using the following sigmoid function for $\pi(.)$:

$$\pi(x) = 1 - \frac{\exp\Big(-\kappa x^{\sigma}\Big) - \exp(-\kappa)}{1 - \exp(-\kappa)}$$

Parameters $\kappa > 0$ and $\sigma \geq 0$ determine the sigmoid form. This allows simulating smoother and probably more realistic epidemic dynamics.

Based on its local life cycle, the pathogen ability to develop on each variety can be described by its basic reproductive number, R_0. According to Madden et al. (2008) the basic reproductive number on SV (respectively RV) is $R_0^{(SV)} = e_{SV} r_{SV} T_{SV}$ (respectively $R_0^{(RV)} = e_{RV} r_{RV} T_{RV}$). Reciprocally, the resistance level of a variety is characterised by the pathogen R_0 on that variety.

4.2.1.3 Dispersal modelling

Soubeyrand et al. (2007) estimated the intra-field dispersal of yellow rust (a wheat disease caused by the fungal pathogen *Puccinia striiformis*) using an exponential individual

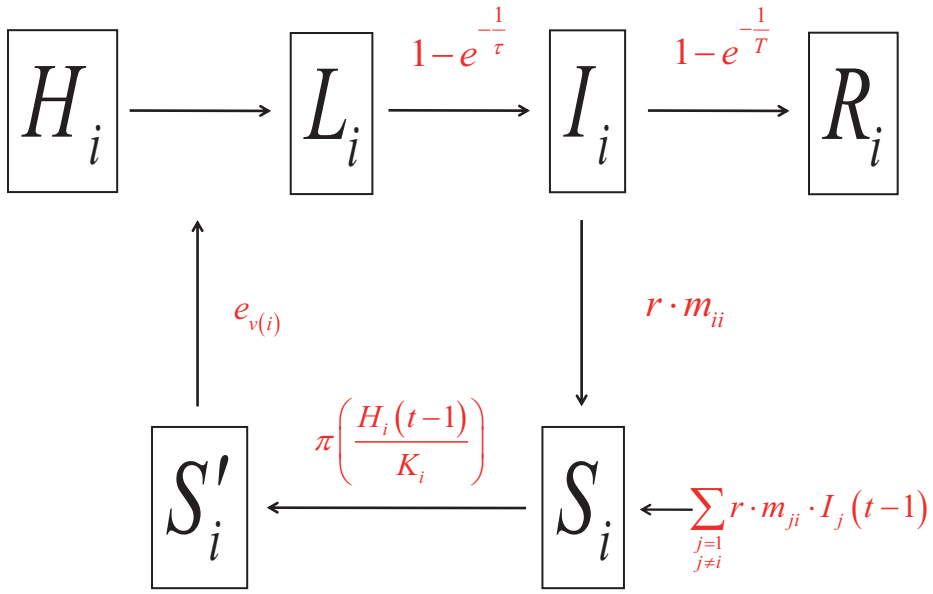


Figure 4.2: Graphical representation of the population dynamics model. H_i: number of healthy sites in patch i, L_i: number of latent sites in patch i, I_i: number of infectious sites in patch i, R_i: number of removed sites in patch i, S_i: number of spores in patch i, S'_i: number of contaminated sites in patch i. τ: latency period, T: infectious period, r: spore production per lesion per day, m: dispersal rates, $\pi(.)$: contamination function, e: infection efficiency.

Table 4.1: Transitions of the stochastic model.

Description	Notation	Transition
Spore cloud	$S_i(t)$	$\sum_{j=1}^{P} rI_j(t-1)\mu_{ji}$
Site contamination	$S'_i(t)$	$\min\left(\text{Bin}\left(H_i(t-1), \pi_i(t)\right), S_i(t)\right)$
Infection of healthy sites	$H_i(t-1) \to L_{i,p}(t)$	$\text{Bin}\left(S'_i(t), e_{v(i)}\right)$
New infectious lesions	$L_i(t-1) \to I_i(t)$	$\text{Bin}\left(L_i(t-1), 1-e^{-\frac{1}{\tau}}\right)$
New removed lesions	$I_i(t-1) \to R_i(t)$	$\text{Bin}\left(I_i(t-1), 1-e^{-\frac{1}{T}}\right)$

dispersal function that considers two kinds of anisotropy, in density and distance. Here we use a simplified version of this function, assuming that spore dispersal is isotropic. More specifically, the proportion of spores emitted from a given source point z and arriving at a given reception point z' is given by the individual dispersal function:

$$g(|| z - z' ||) = \frac{2\pi}{m_0^2} \exp\left(\frac{2\pi}{m_0} || z - z' ||\right),$$

where $||z - z'||$ is the distance between z and z', and m_0 is a range parameter. For a more realistic description of the spore dispersal, the amount of spores dispersed from patch i to patch j can be calculated as:

$$m_{ij} = \int_{\mathcal{A}_i} \int_{\mathcal{A}_j} g(|| z - z' ||)\, dz'\, dz. \tag{4.1}$$

The integration is performed between all pairs of points that belong to the areas $\mathcal{A}_i$ and $\mathcal{A}_j$ of patches i and j, respectively. This was done with the CaliFloPP algorithm (Bouvier et al., 2009), which computes the integral of a pointwise dispersal function between any pair of source and target polygons (here the patches).

4.2.2 Model analysis

4.2.2.1 Numerical experimentation

Five field patterns were constructed (see chapter 3). On each of them, five proportions of the susceptible variety and three levels of variety aggregation were simulated. For each variety allocation, *i.e.* proportion × aggregation combination, two replicates were realised, leading to a total of 5 field patterns × 5 variety proportions × 3 aggregation levels × 2 variety allocation replicates = 150 landscape structures. The RV proportions used were 10, 30, 50, 70 and 90%. The first aggregation level corresponded to a mixed strategy, the third to a grouped strategy and the second was an intermediate case which will be named mosaic strategy (figure 4.1).

$R_0^{(SV)}$ was considered constant and fixed to 2 while $R_0^{(RV)}$ varied from 0 to 1.8 by 0.2 in order to account for several host resistance levels. All the life-cycle parameters potentially depend on both the host and the pathogen genotype. Here the pathogen population was composed of a single pathogen genotype whose infection efficiency only depended on the variety. Hence, to obtain the above $R_0^{(RV)}$ values, parameters r and T were fixed to 2 spores per day and 10 days, respectively, for both varieties. The latency time, τ, was fixed to 5 days for both variety. The parameters κ and σ of function $\pi(.)$ were both set to 6.

The epidemics started from 100 infectious lesions located on a patch within a susceptible field chosen at random. Two replicate simulations, with two different initial inoculum

locations, were made for each landscape structure (150) × infection efficiency (10) combination. Moreover, for each of these combinations, a short range (m_0 = 150m) and a long range (m_0 = 1500m) dispersal ability were used. Epidemics were simulated over 2000 time steps. This led to a total of 6000 simulations: 2 replicates × 150 landscape structures × 10 infection efficiencies (i.e. resistance levels of RV) × 2 dispersal ranges.

4.2.2.2 Outputs

First, the global dynamics of the epidemic was measured by computing the integrated relative green surface over the epidemic season (GS, van den Bosch and Gilligan (2003)) for the whole landscape (GS_{tot}), SV (GS_{SV}) and RV (GS_{RV}). The integrated relative green surface was defined as the sum over the epidemic season and over all patches of a given variety of the ratio between the number of healthy sites in a patch and its carrying capacity. Note that in the case of complete resistance, *i.e.* $R_0^{(RV)} = 0$, only GS_{SV} was calculated.

From GS_{SV}, we defined γ as the gain, in term of integrated green surface, provided by the introduction of RV:

$$GS_{SV} = GS_{ref}(1+\gamma), \qquad 0 \leq \gamma \leq \frac{GS_{max} - GS_{ref}}{GS_{ref}},$$

Where GS_{ref} is the integrated green surface of reference and is defined as the integrated green surface when only the susceptible variety was present in the landscape, averaged over 10 epidemic simulations. It was equal to 793 and 725 when dispersal range was 150m and 1500m, respectively. GS_{max} is the maximal integrated green surface, *i.e.* the integrated green surface without disease ($GS_{max} = 2000$).

For each simulated epidemic and at each time step, the maximal distance between the initial inoculum position and the current position of the most distant infectious lesions (for all directions) was also computed in order to evaluate the spatial spread of the pathogen (figure 4.3). At each time, this distance was expressed relatively to the maximal distance reached at the end of the epidemic in order to limit the influence of the initial inoculum location.

Finally, the consequences of the changes in $R_0^{(RV)}$ on GS_{RV} were analysed for each combination of variety proportions and aggregation levels. For that, GS_{RV} was plotted against $R_0^{(RV)}$ and the curves were fitted to the following logistic function, defined by three parameters:

$$GS_{RV} = GS_{max} - \frac{GS_{max} - GS_{min}}{1 + \exp\Big(- a(R_0^{(RV)} - b)\Big)}. \tag{4.2}$$

In equation (4.2), GS_{max} still denotes the maximal integrated green surface ($GS_{max} = 2000$); GS_{min} is GS_{RV} obtained when $R_0^{(RV)} = 1.8$; a is the rate of GS_{RV} decrease due to the increase

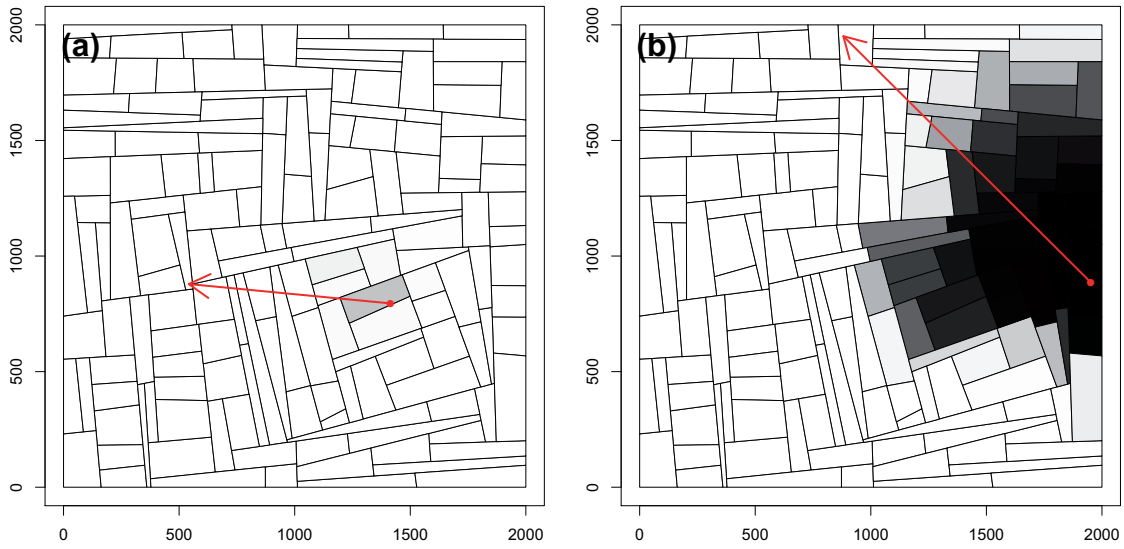


Figure 4.3: Spatial expansion of the epidemic at time $t = 300$ after inoculation. Landscape structures used are those of figure 4.1a and b for panel a and b, respectively. Diseased sites ($L + I + R$) are represented in a grey scale according to their density. Intra-field patches are not apparent. The red dot marks the inoculated patch and the red arrow the maximal distance in all directions to the current position of infectious lesions.

of $R_0^{(RV)}$; b is the $R_0^{(RV)}$ value that corresponds to $GS_{RV} = 50\%$ of GS_{max}. Parameters GS_{min}, a and b were estimated using least squares.

4.2.2.3 Sensitivity analysis

We studied the sensitivity of GS_{tot}, GS_{SV} and GS_{RV} to $R_0^{(RV)}$, to the RV proportion and to the aggregation level of the varieties by computing global sensitivity indices (Saltelli et al., 2008). Global sensitivity indices were defined using the Sobol's decomposition of the model variance (Sobol, 1993). Global sensitivity index for parameter p_1 (SI_{p_1}) represents the part of the model variance explained by p_1:

$$SI_{p_1} = \frac{V_{p_1} + \sum_{p_2} V_{p_1 p_2} + \sum_{p_2 p_3} V_{p_1 p_2 p_3}}{V}$$

where V_{p_1} is the variance due to parameter p_1 alone, $V_{p_1 p_2}$ is the variance due to p_1 in interaction with parameter p_2 and $V_{p_1 p_2 p_3}$ is the variance due to p_1 in interaction with p_2 and p_3. Variances explained by each parameter were estimated using a third order chaos polynomial and a classical linear regression (Sudret, 2008).

4.3 Results

4.3.1 Effect of *RV* introduction

The relative gain in green surface (γ) due to the introduction of *RV* in the landscape varied from 0 to 1.40 when the dispersal range was low and from 0 to 1.75 when the dispersal range was large, depending on the *RV* proportion, on the *RV* resistance level and on the variety aggregation level (figure 4.4). When $m_0 = 1500$m, γ sometimes reaches its maximal value ($\gamma = 1.75$), which means that the epidemic stopped and the pathogen went extinct at the first time.

As expected, the resistant variety protected the susceptible variety but also itself (figure 4.5): when the resistant variety proportion was greater in the landscape, the disease severity was lower on the susceptible and the resistant varieties, and thus their integrated green surface was greater. Reducing the aggregation level of the varieties resulted in a decrease of the disease severity on *SV* (figure 4.5, mixed strategy *vs.* grouped strategy). When varieties were mixed, however, the resistant variety received more spores because of its spatial proximity with the susceptible variety and this resulted in an increase of the disease severity on *RV* (figure 4.5). The interplay between the disease reduction on *SV* and of disease increase on *RV* for lower aggregation levels led, in certain cases, to the fact that grouped strategies were more efficient than mixed strategies.

Similar results were obtained with a large dispersal range. In that case, the integrated green surface tended to be smaller but the effect of the proportion of *RV* on the disease severity on *SV* was greater. When the *RV* proportion was high, the integrated green surface remained always high when mixed strategies were used (figure 4.5b). This was due to the fact that the spores dispersed from their source field had a high probability to encounter unfavourable conditions on a resistant field (see also figure 4.6d). When the dispersal range was low ($m_0 = 150$m), autoinfection at the field level was equal to 54% while it dropped at 7% for $m_0 = 1500$m.

4.3.2 Spatial expansion of the epidemic

Figure 4.6 shows the distance between the initial inoculum location and the epidemic front over time. For a limited dispersal range and when *RV* resistance level and *RV* proportion were high, the epidemic expanded faster for grouped strategies (figure 4.6a and b). When the varieties were grouped, the initial progression rate did not vary much with the *RV* resistance level while it changed for the mosaic strategy and, to a higher extent, for the mixture strategy (figure 4.6a, b and c). This indicates that the mixture strategy allowed slowing down the spatial progression of the disease on the landscape.

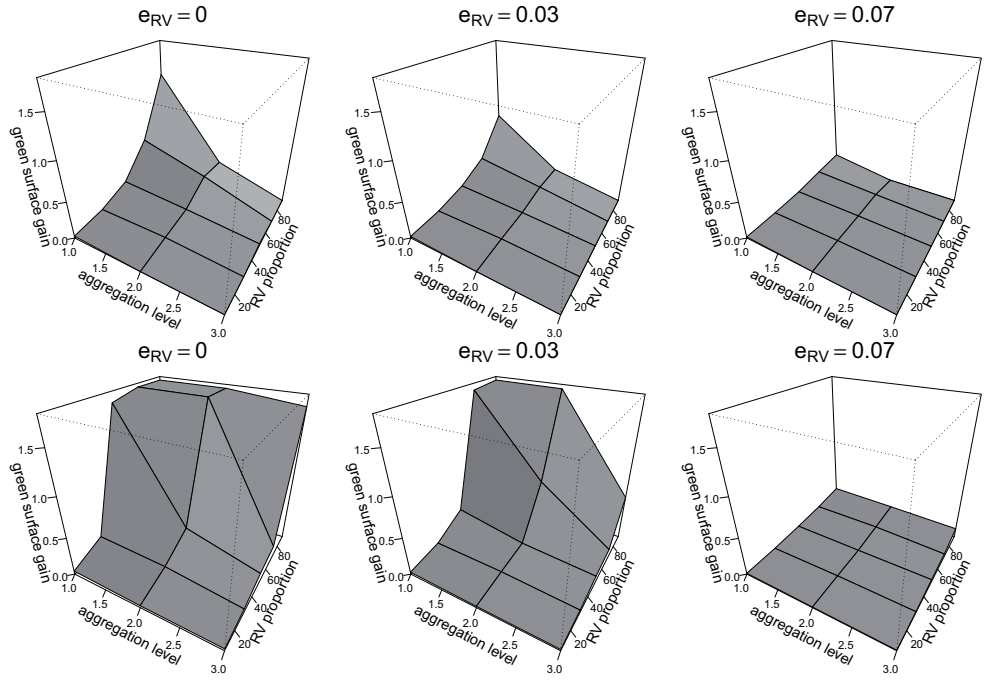


Figure 4.4: Gain in term of integrated green surface, provided by the introduction of RV (γ). Top line: $m_0 = 150$m, bottom line: $m_0 = 1500$m. Left column: $e_{RV} = 0$, meddle column: $e_{RV} = 0.03$ and right column: $e_{RV} = 0.07$. x-axis: variety aggregation level, 1=mixed, 2=mosaic, 3=grouped; y-axis, RV proportion in percent; z-axis: γ.

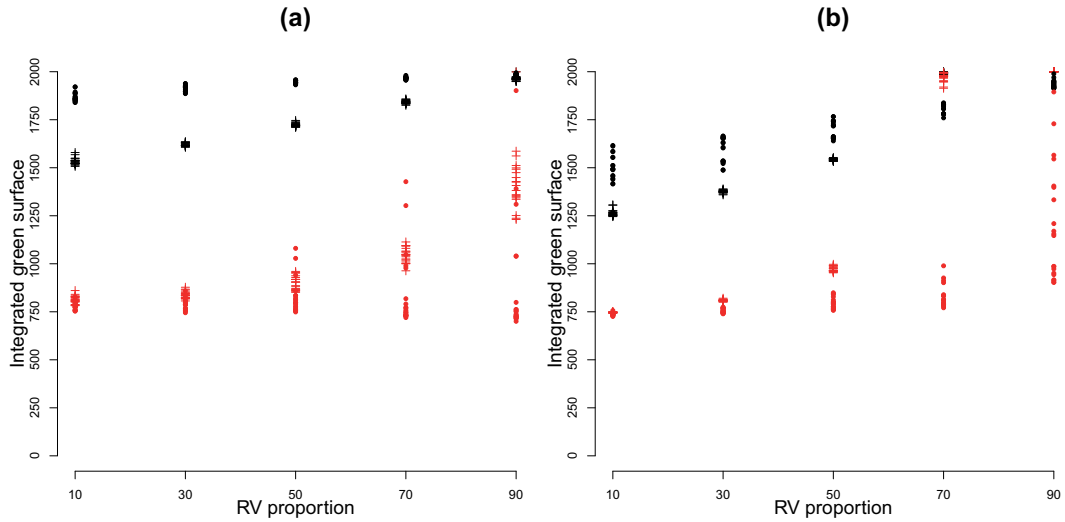


Figure 4.5: Integrated green surface (GS) for each variety according to RV proportion. a: $e_{RV} = 0.03$ and $\mu_0 = 150m$, b: $e_{RV} = 0.03$ and $\mu_0 = 1500m$. Red: SV, black: RV. Dot: grouped strategy for variety allocation, +: mixed strategy for variety allocation.

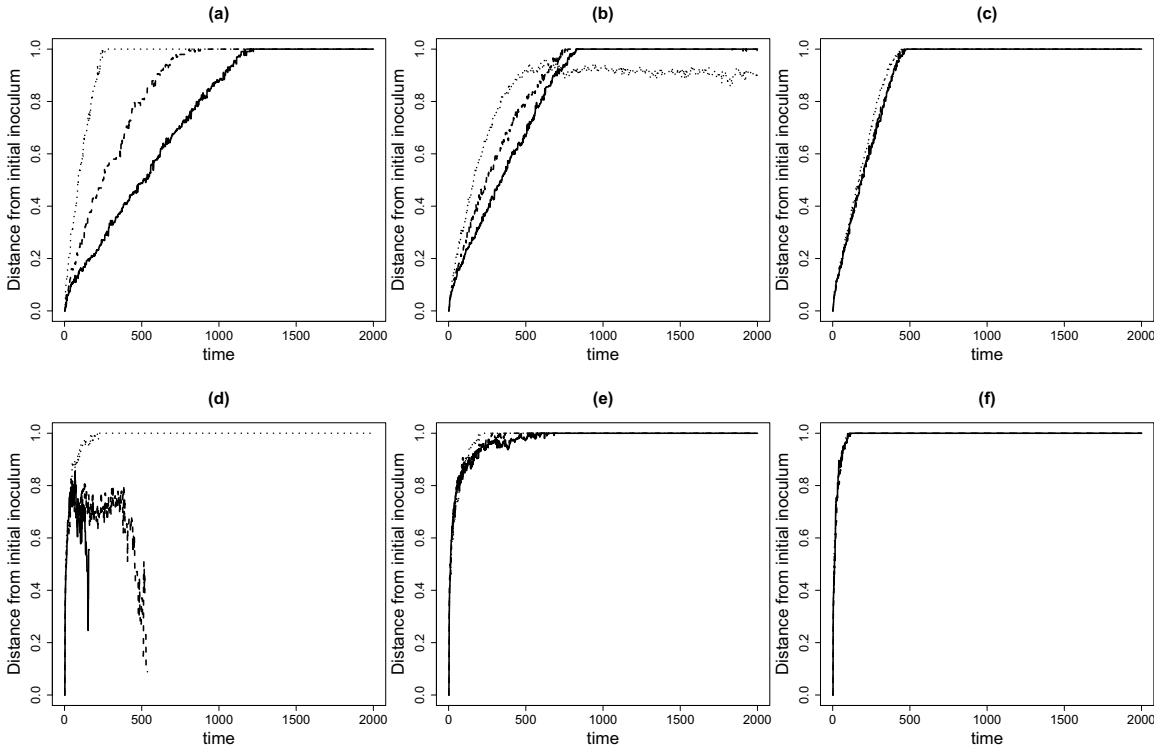


Figure 4.6: Spatial expansion of the pathogen population as a function of time. Top line: $\mu_0 = 150m$, bottom line: $\mu_0 = 1500m$. Left column: $R_0^{(RV)} = 0$ and 70% of RV, middle column: $R_0^{(RV)} = 0.06$ and 70% of RV, right column: $R_0^{(RV)} = 0.06$ and 30% of RV. Solid line: mixed strategy for variety allocation, dashed line: mosaic strategy for variety allocation and dotted line: grouped strategy for variety allocation.

With a higher dispersal range, the pathogen spread more easily over the landscape and no difference in the spatial progression rate between variety allocation strategies was observed at the beginning of epidemics (figure 4.6d, e and f). In addition, the pathogen went extinct for mixed strategies in some cases (figure 4.6d).

4.3.3 Sensitivity of integrated green surface to landscape structure

In this section we analyse the sensitivity of the integrated green surfaces to the landscape structure variables, globally (table 4.2) or for each $R_0^{(RV)}$ separately (figure 4.7).

4.3.3.1 Global sensitivity analysis

When the resistant variety was completely resistant ($R_0^{(RV)} = 0$) and the dispersal range was low, the variety aggregation explained 28.2% of the integrated green surface variability,

Table 4.2: Sensitivity indices of GS_{tot}, GS_{RV} and GS_{SV} to RV proportion, variety aggregation level and RV resistance level.

	dispersal range	GS_{tot} $R_0^{(RV)} > 0$	GS_{RV} $R_0^{(RV)} > 0$	GS_{SV} $R_0^{(RV)} > 0$	GS_{SV} $R_0^{(RV)} = 0$
RV proportion	150 m	41.1	3.0	32.0	45.4
	1500 m	41.0	7.9	42.0	64.0
Aggregation level	150 m	0.2	1.4	18.1	28.2
	1500 m	1.0	0.5	11.6	23.6
RV resistance level	150 m	54.9	90.9	26.2	-
	1500 m	53.9	88.1	38.1	-
Residuals	150 m	3.8	4.5	23.8	26.4
	1500 m	4.0	3.5	8.4	12.4

and the proportion of RV more than 45%. For a high dispersal range, the effects of the RV proportion were enhanced while these of the variety aggregation decreased (table 4.2).

In the case of quantitative resistance ($R_0^{(RV)} > 0$), the resistance level of RV was the most influential parameter on GS_{tot} (table 4.2) but the RV proportion also showed a major effect. On the contrary, the variety aggregation level was not influential. Changing the dispersal range gave roughly the same results.

The sensitivity analysis of the integrated green surface on each variety (GS_{SV} and GS_{RV}) exhibited very different results according to the variety considered. The disease severity on the resistant variety was essentially explained by its own resistance level while the effects of the three parameters were more balanced on the susceptible variety. The RV proportion and resistance level were the most influential parameters on GS_{SV} variance, but the variety aggregation level also had a noticeable effect. When the dispersal range was high, the effect of RV proportion and RV resistance level increased whereas the effect of variety aggregation decreased. The fact that GS_{SV} was sensitive to the variety aggregation level but not GS_{tot} is consistent with the fact that SV and RV behaved in an opposite way when the variety aggregation level changed (see section 4.3.1).

The high level of the residual variance on GS_{SV} denoted the interaction between local spatial structures (local aggregates of a variety and shape of these aggregates) and the position of initial inoculum. The residual variance decreased when dispersal range increased (table 4.2), which is consistent with an effect of local spatial structures.

4.3.3.2 Sensitivity analysis for each $R_0^{(RV)}$ separately

In order to investigate in more details the effect of landscape structure, we computed the sensitivity indices of GS_{SV} and GS_{RV} for each value of $R_0^{(RV)}$ separately (figure 4.7). When $R_0^{(RV)}$ was lower than 1, the proportion of resistant variety and the varieties aggregation level explained 56% and 42%, respectively, of the GS_{RV} variability, in average. When $R_0^{(RV)}$ was greater than 1, the part of the GS_{RV} variability explained by the variety aggregation level dropped to 4% and the residual variability reached 80% of the variance. At the same time, the effect of the variety proportion first increased and then decreased to 14%. These observations are in accordance with the classical threshold effect of the basic reproductive number, which was obtained here at the landscape scale for RV. When $R_0^{(RV)}$ was greater than 1, RV became a source of spore production, which limited the dilution effects (see the discussion section for more details).

In the case of SV, a progressive decrease of the effect of the RV proportion was observed when $R_0^{(RV)}$ increased. When $R_0^{(RV)}$ was lower than 1, the RV proportion had the highest effect on GS_{SV} variability but when $R_0^{(RV)}$ was greater than 1 variety aggregation became more influential. As for the global analysis above, the proportion of residual variance remained high.

When the dispersal range was high, the threshold effect disappeared and, for both varieties, the RV proportion accounted for most of the variance in GS_{SV} and GS_{RV}. But the variety aggregation level still explained 20% of the variance of GS_{SV} for high RV resistance levels.

This analysis must be modulated with regards to the total GS variance. For RV, the total variance of GS_{RV} reached a maximum just before $R_0^{(RV)} = 1$. After this threshold the landscape structure little impacted GS_{RV}. For SV, the total variance continuously decreased when $R_0^{(RV)}$ increased and it became very low when $R_0^{(RV)}$ exceeded one. This mean that for $R_0^{(RV)} > 1$, the landscape structure did not reduce much the epidemics on the susceptible variety.

4.3.4 Effect of resistance level

Figure 4.8 shows the integrated green surface of the resistant variety (GS_{RV}) as a function of $R_0^{(RV)}$. The curves in figure 4.8 can be characterised by three parameters (see equation (4.2)): GS_{min}, a, and b. Their values are shown in figure 4.9 for each combination of the RV proportions and variety aggregation levels.

Variations of GS_{min}, (figure 4.9a and d) were very small for all landscape structures. This means that when $R_0^{(RV)}$ was high (*i.e.* the resistance of RV was low), RV was equally

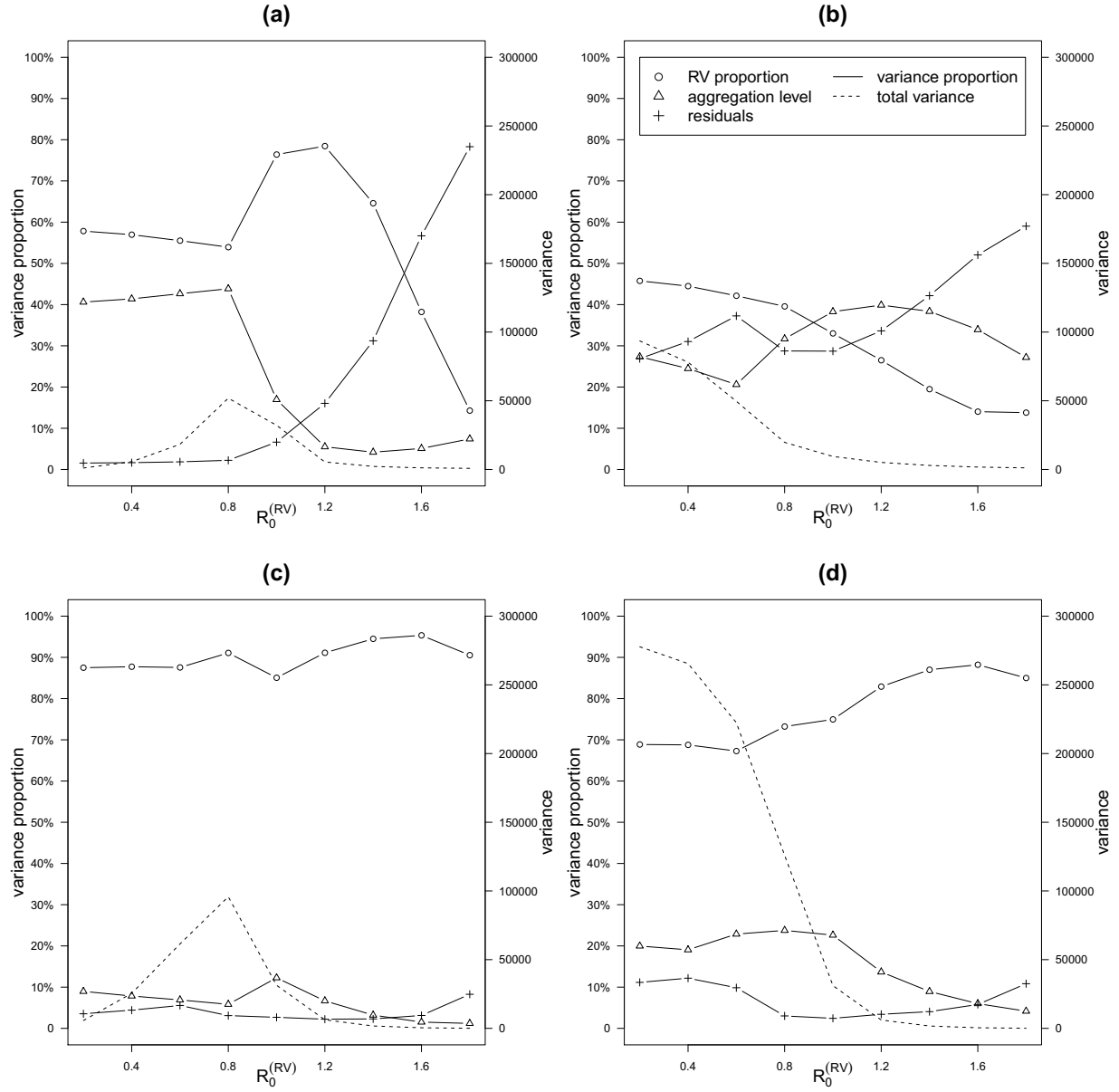


Figure 4.7: Sensitivity indices for each $R_0^{(RV)}$ separately. For each RV resistance value, the proportion of variance accounted for by the proportion of RV, the aggregation level and the residuals are shown. Left column, GS_{RV}; right column, GS_{sV}; top line, $\mu_0 = 150m$; bottom line, $\mu_0 = 1500m$.

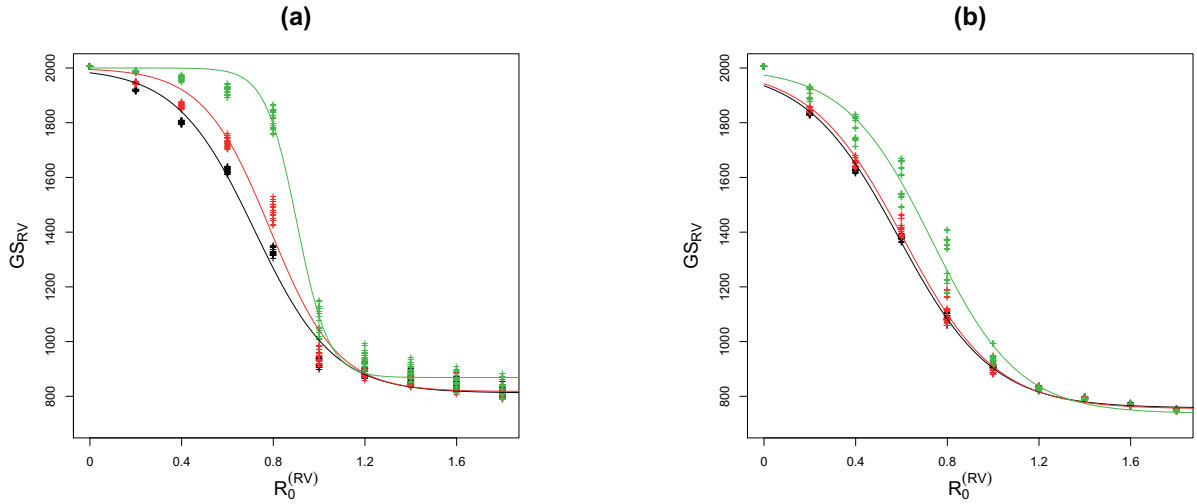


Figure 4.8: Fitted values (solid lines) and observation of GS_{RV} (+) as a function of $R_0^{(RV)}$. Black: mixed strategy for variety allocation, red: mosaic strategy for variety allocation and green: grouped strategy for variety allocation. a, 30% of RV and $m_0 = 150$m; b, 30% of RV and $m_0 = 1500$m.

diseased over the epidemic season, whatever the deployment strategy chosen.

The shape of the decline of GS_{RV} between low and high value of $R_0^{(RV)}$ depend on the variety aggregation level and the dispersal range (figure 4.8). Analysis of parameter a (figure 4.9b) shows that decreasing the aggregation level made the transition of GS_{RV} from high to low values (when $R_0^{(RV)}$ increased) more progressive. When the varieties were grouped, a sharp decrease in the integrated green surface (high values of a) occurred when $R_0^{(RV)}$ increased. On the contrary, for mixed and mosaic strategies, the transition from high to low GS_{RV} was more progressive. When mixed or mosaic strategies were used, increasing the proportion of resistant variety accelerated the rate of GS_{RV} decrease for increasing $R_0^{(RV)}$. Increasing the dispersal range made the transition of GS_{RV} between high to low values more progressive in all cases and reduced the differences between variety deployment strategies (figure 4.9e and b).

The $R_0^{(RV)}$ value that corresponds to $GS_{RV} = 50\%$ of GS_{max} (parameter b) is around $R_0^{(RV)} \approx 1$, depending on variety proportions and aggregation level as well as on the dispersal range. When the proportion as well as the aggregation level of the resistant variety increased, b increased (figure 4.9c and f). In addition b values were higher for large dispersal range (figure 4.9f). This means that, when RV received lower spores, the GS_{RV} decrease occurred for higher $R_0^{(RV)}$ values, *i.e.* for lower resistance level of RV.

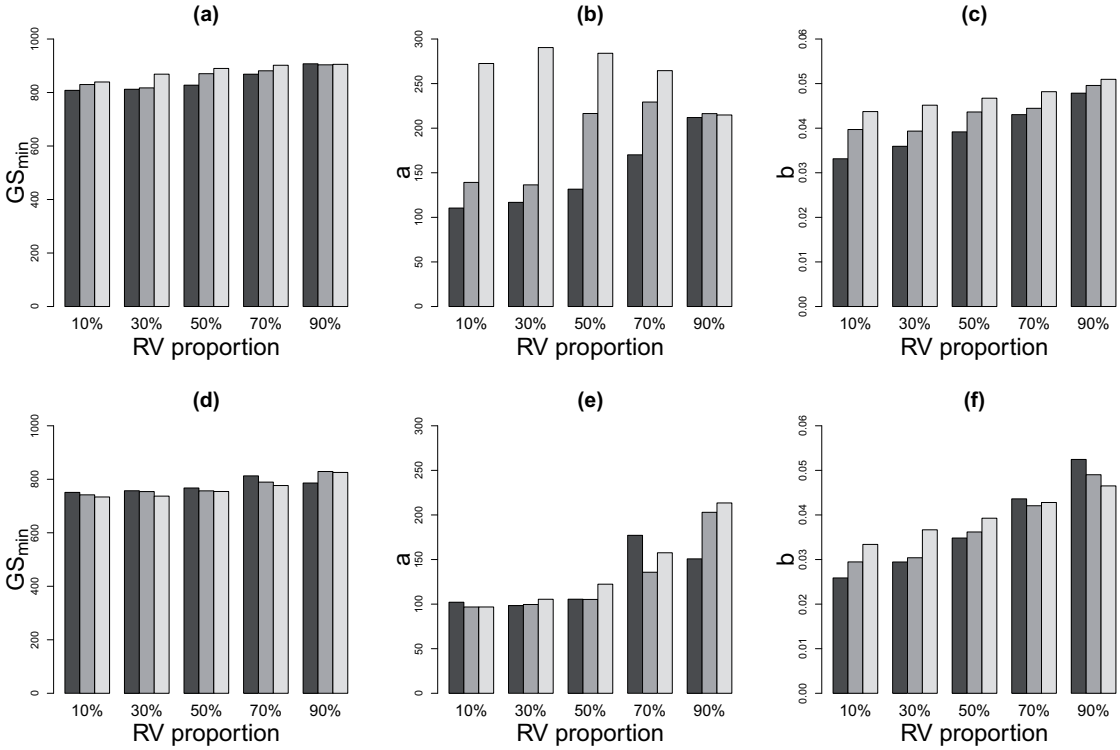


Figure 4.9: Estimation of parameters GS_{min} (a and d), a (b and e) and b (c and f) of equation (4.2) plotted against the proportion of the resistant variety. The level of variety aggregation (mixed, mosaic and grouped) is increasing from the darkest to the lightest grey. Top line: $\mu_0 = 150m$, bottom line: $\mu_0 = 1500m$.

4.4 Discussion

How does habitat connectivity influence the spread of a population? We addressed this question in the context of agricultural landscapes to evaluate the benefits of functional diversity for epidemic control. In this context, the host varieties were considered as habitats for the pathogen. The four components describing habitat connectivity were then the varietal composition of the landscape, the spatial aggregation of the varieties, the ability for the pathogen to develop on each variety (variety resistance level), and the dispersal ability of the pathogen. These four components had a strong impact on the epidemic development at the landscape scale and they show strong interactions.

The first conclusion of the study is that the introduction of a resistant variety in the landscape has an effect in reducing the epidemic development on the susceptible variety. This effect depended on the variety proportions and their spatial arrangement. The greatest effect on GS_{SV} was however that of the RV proportion, especially when dispersal range was high. This is not surprising and means that dilution effects at the landscape scale strongly depends on the level of resistance that are available. In addition, the resistance level of RV is also crucial: when $R_0^{(RV)} > 1$, RV becomes a source of spores and its effects on SV decrease drastically.

We considered the case of a complete resistance ($R_0^{(RV)} = 0$) as well as a quantitative resistance ($R_0^{(RV)} > 0$). In this latter case, the resistant variety was itself affected by the pathogen. Not surprisingly, the resistance level of RV was the most influential parameter for determining its level of disease. Spatial arrangement of RV plays opposite roles with respect to the susceptible variety. While mixed strategy decreases landscape connectivity and protects SV, it exposes RV to a greater amount of spores. As a consequence, RV was more diseased in mixed rather than grouped variety allocation strategies. The balance between the gain on SV and the loss on RV can make grouped strategies for variety allocation more resistant to disease.

A threshold effect was found on the integrated green surface of the resistant variety and was related to the local basic reproductive number $R_0^{(RV)}$. This can be explained by a source-sink dynamics. When $R_0^{(RV)}$ is lower than 1 local population growth rates are positive on SV (the source), negative on RV (the sink) and the spore dispersal generates a source-sink dynamics. A simple model for a single-species source-sink system is exposed by (Holt, 1993). It assumes a saturated source patch in which juveniles are forced to emigrate to a sink patch where, in the absence of this immigration, the population would decline. Holt showed that the size of the sink population is directly proportional to the immigration from the source, scaled by the growth rate of the population in the sink patch. In our context, when the RV proportion is fixed, immigration from source (SV) to sink (RV) can be enhanced by

increasing the average spatial proximity (going from grouped to mixed landscapes) between RV and SV or by increasing the spore dispersal range. Our results are thus in line with this simple prediction since in mixed landscapes RV exhibited a higher level of disease, as well as when dispersal range was large. In addition, RV exhibited a higher level of disease when $R_0^{(RV)}$ increased up to 1 (increase of the local growth rate). The difference of behaviour between mixed and grouped landscapes when the spore dispersal range was short could be explained by the fact that immigration from SV to RV was too low in grouped landscapes for maintaining a viable population on RV. Then, the pathogen could develop on RV only when $R_0^{(RV)}$ was greater than 1. In mixed landscapes, immigration from SV to RV allowed the pathogen to maintain on RV, which explains that a progressive decrease in GS_{RV} was observed instead of a threshold effect.

The data analysis presented in chapter 2 suggests that such source-sink dynamics operate in the field. *Trémie* is a wheat variety with a high level of resistance to leaf rust. It was found related to a particular leaf rust genotype (pathotype 073100) even though it is resistant to 073100. Infection of *Trémie* with that pathotype in controlled conditions however produces a few viable pustules. Since the frequency of 073100 in the global pathogen population was high, due to its increase by variety *Soissons*, it can be assumed that the inoculum provided by *Soissons* accounted for the presence of 073100 on *Trémie*. Pathotype 073100 did not develop epidemics on *Trémie* because of a very low R_0 value but was constantly reintroduced on that variety as external inoculum. This illustrates with a real situation the theoretical results presented here.

When resistant varieties are introduced in agrosystems, they usually show a decreasing resistance level over time, due to rapid adaptation of pathogens (Johnson, 1961; Stukenbrock and McDonald, 2008). One important issue related to the re-introduction of functional diversity in agriculture is that of the preservation of the efficiency of new resistant varieties. In this work we supposed that the resistance level was constant during the epidemics, *i.e.* that the pathogen population could not adapt to the resistant variety by increasing its basic reproductive number. However, some of the results could be seen from an evolutionary point of view. Figure 4.8 shows the integrated green surface of the resistant variety according to its resistance level. Suppose that the adaptation to the pathogen population results in an erosion of the resistance level of RV (and thus an increase in $R_0^{(RV)}$). This erosion could take the form of a brutal breakdown in aggregated landscapes or be more progressive in mixed landscapes, having a lower impact on the final yield and leaving more time to the farmer to adjust.

In this work the pathogen population was composed by a single genotype. However, epidemics on agricultural landscape are usually caused by more diversified pathogen populations. A recent study (Papaïx et al. (2011) and chapter 2) based on a large database

analysis coupling pathogen frequency data and disease observation data, showed that the genetic structure of the pathogen population present on a particular variety is partly due to the reciprocal influence of the different varieties composing the landscape. Moreover, the observed disease on a variety was linked to the landscape composition, through the pathogen population structure. A perspective would be to consider a structured pathogen population with more genotypes and study how the landscape structure shapes this pathogen population and determines the distribution of the pathogen genotypes on each variety.

Chapter 5

Mutant establishment and coexistence between pathogen genotypes in a heterogeneous host landscape

This chapter is based on an article project by Julien Papaïx, Suzanne Touzeau, Hervé Monod and Christian Lannou

Contents

5.1 Introduction

Selection for quantitative traits influences pathogen evolution in agricultural systems and can result in differential adaptation to host varieties (Pariaud et al., 2009a). In a large scale study on wheat leaf rust Papaïx et al. (2011) demonstrated that such differential adaptation of pathogen genotypes to host varieties was largely accounted for by the relationship between the composition of the host population and its susceptibility to disease. It also revealed that the pathogen population included genotypes (pathotypes) that could be defined as specialists and others with a larger host panel, that could be qualified of generalists. The landscape varietal composition influences the observed resistance level on the most frequently grown wheat varieties by altering the structure of the pathogen population. Which pathogen genotypes will develop on which variety? This is a crucial question in plant epidemiology for understanding the link between the host population structure and its susceptibility to disease. In the present work, we develop a theoretical approach to investigate the conditions of emergence and establishment of a mutant pathogen with generalist features in an agricultural landscape, then we determine the conditions of co-existence between specialists and generalists pathotypes.

In all communities, the first condition for stable coexistence is success in establishment of a population (With, 2002). Establishment success in homogeneously mixed host-pathogen systems is determined by the pathogen basic reproductive number. When two pathogens are competing for a single host, the pathogen with the highest basic reproductive number will invade the resident population. However, when dispersal is limited, the pathogen population will typically progress in a susceptible host population as a travelling wavefront (Mundt et al., 2009). In such conditions, a mutant genotype will be overwhelmed if it appears in the already diseased region and it will become more likely established if it arises closer to the wavefront (Wei and Krone, 2005). This effect is known as the 'surfing effect' in population genetics (Excoffier and Ray, 2008). The effects of landscape structure on the establishment of a mutant have been rarely investigated in the literature. To our knowledge, the most advanced work is that of Burton and Travis (2008). These authors constructed a model for simulating the expansion range of haploid individual in a rectangular lattice. Heterogeneity was introduced by separating the left and the right side of the lattice either by an unsuitable block or by considering that the two areas of suitable habitat were connected by a narrow corridor. They found that landscape structure, together with the spatial location of mutant introduction, have a considerable influence on the mutant survival probability and on the population dynamics of the mutants. In particular the landscape structure could favour deleterious mutations that would normally disappear.

When all types (species or genotypes) are well established in the community, the question of stable coexistence can be addressed. A classical mechanism for coexistence is niche

partitioning (Chesson, 2000). This requires that types differ in their use of resources but not necessarily that these resources fluctuate in space (homogeneous environment). Environment heterogeneity makes however easier a stable coexistence because it promotes mechanisms that are not possible in homogeneous environments (Melbourne et al., 2007). Chesson (2000) identified three mechanisms that depend on the variation of resources in space: storage effect, relative non-linearity and fitness-density covariance. More recently, Débarre and Lenormand (2011) added another mechanism: 'habitat boundary polymorphism'. This new mechanism requires both habitat heterogeneity and distance-limited dispersal, which creates maladaptation at habitat edges and favour the maintenance of more generalist genotypes. Débarre and Lenormand (2011) found that this mechanism acts in very different environment: continuous or stepping stone one-dimensional structures as well as circular or flower-shaped two-dimensional structures. However, their framework is not readily applicable to the specific situation of agricultural landscapes.

We propose here to study the conditions for establishment and stable coexistence of pathogen genotypes in agricultural landscapes with a simulation model. The model is stochastic and the life-cycle is based on an air-borne plant-pathogen foliar fungus. The landscape is represented as a set of fields on which two varieties are cultivated with a controlled proportion and spatial organisation. We consider also several dispersal abilities and several specialisation costs for the pathogen. We consider two case-studies: first, the pathogen population is composed of two specialised genotypes that can mutate towards a generalist genotype. The host population size is fixed and we examine the conditions for establishment of a generalist population during an epidemic. Second, the three genotypes are considered to pre-exist in the pathogen population, there is no new mutation, and the host is able to grow. We then study the equilibrium of the pathogen population composition.

5.2 Materials and methods

5.2.1 The model

5.2.1.1 The landscape

We used the framework described in chapter 3 in order to generate 2000×2000 km^2 agricultural landscapes composed of around 150 fields. Two varieties, V_1 and V_2 were deployed with controlled spatial arrangements defined by their proportions and aggregation level. Remind that this framework requires two different scales: the management scale (fields) and the dispersal scale (intra-field patches) at which the pathogen population is assumed to be perfectly mixed.

5.2.1.2 Population dynamics model

Based on the asexual life cycle of *Puccinia triticina* (supporting information, appendix 2), we developed an HLIR model with spores as a propagule state (Madden et al., 2008). The environment is composed of N patches and the pathogen population of $P = 3$ genotypes, P_1, P_2 and P_3. P_1 and P_2 are both specialists of varieties V_1 and V_2, respectively. P_3 is a generalist for which the landscape is homogeneous. Each patch was considered as a set of infection sites with no spatial positions. The model describes the dynamics of the number of infection sites in each of the following states: healthy sites (H), latent lesions (L), infectious lesions (I) and removed sites (R). For host patch j we have:

- host dynamics

$$\frac{dH_j}{dt} = -\sum_{p=1}^{P}\Big(e_{P_p,v(j)}\cdot\pi_j\cdot\sum_{i=1}^{N} r_{P_p,v(i)}\cdot m_{ij}\cdot I_i(P_p)\Big) + \delta H_j(1-\frac{H_j}{C_j}), \tag{5.1}$$

- dynamics of genotype P_3

$$\begin{cases} \dfrac{dL_j(P_3)}{dt} = e_{P_3,v(j)}\cdot\pi_j\displaystyle\sum_{i=1}^{N} m_{ij}\Big(r_{P_3,v(i)}\cdot I_i(P_3) + \sum_{p=1}^{P-1} r_{P_p,v(i)}\cdot\mu\cdot I_i(P_p)\Big) - \frac{1}{\tau_{P_3,v(j)}}\cdot L_j(P_3), \\ \dfrac{dI_j(P_3)}{dt} = \dfrac{1}{\tau_{P_3,v(j)}}\cdot L_j(P_3) - \dfrac{1}{T_{P_3,v(j)}}\cdot I_j(P_3), \\ \dfrac{dR_j(P_3)}{dt} = \dfrac{1}{T_{P_3,v(j)}}\cdot I_j(P_3), \end{cases} \tag{5.2}$$

- dynamics of genotypes P_p, for $p = 1, 2$

$$\begin{cases} \dfrac{dL_j(P_p)}{dt} = e_{P_p,v(j)}\cdot\pi_j\displaystyle\sum_{i=1}^{N} r_{P_p,v(i)}\cdot m_{ij}(1-\mu)\cdot I_i(P_p) - \frac{1}{\tau_{P_p,v(j)}}\cdot L_j(P_p), \\ \dfrac{dI_j(P_p)}{dt} = \dfrac{1}{\tau_{P_p,v(j)}}\cdot L_j(P_p) - \dfrac{1}{T_{P_p,v(j)}}\cdot I_j(P_p), \\ \dfrac{dR_j(P_p)}{dt} = \dfrac{1}{T_{P_p,v(j)}}\cdot I_j(P_p), \end{cases} \tag{5.3}$$

where $v(j)$ indicates the variety cultivated in patch j, $e_{P_p,v(j)}$ is the infection efficiency of genotype P_p on variety $v(j)$, π_j is the probability for a spore to be deposited on a free

infection site in patch j, $r_{P_p,v(j)}$ is the per day number of spores produced by an infectious lesion of genotype P_p on variety $v(j)$, m_{ij} is the proportion of spores released from patch i that are deposited in patch j, $\tau_{P_p,v(j)}$ and $T_{P_p,v(j)}$ are the latent period and infectious period of genotype P_p latent lesion on variety $v(j)$, respectively. We assumed that P_3 can appear by mutation from the other genotypes present in the pathogen population at a mutation rate μ. The host population grows at rate δ. The carrying capacity for the host in patch j, C_j , is proportional to the area of patch j.

In order to take into account demographic stochasticity, which could be very important in the establishment phase of a population, we developed a stochastic version of the model. The dynamics of infection sites in each state (healthy, latent, infectious and removed) and for each patch is described by the following sequence of events. First, the composition of the spore cloud arriving in patch j is determined. Then the sites to be contaminated are determined and dispatched among genotypes. Finally, new latent, infectious and removed lesions are calculated. Each of these steps is described in details in supporting information (appendix 2).

5.2.2 Numerical experiments

5.2.2.1 Landscapes and dispersal

Landscapes of around 150 fields were simulated. Using the framework of chapter 3 we generated 5 field patterns. Two situation were studied: with balanced (50% of both varieties) and unbalanced (30% of V_1 and 70% of V_2) variety proportions. For each situation, three variety deployment strategies with an increasing aggregation level were considered: mixed, mosaic and grouped.

The landscape is defined as the field pattern (physical structure) along with the varieties (qualitative structure). For each field pattern $\times$ variety proportions $\times$ variety aggregation levels, two landscape replicates were generated by changing variety allocation in space. In addition, for each of these replicates, two independent epidemic simulations were performed. Thus, for each variety proportion $\times$ variety aggregation level combination, 5 field patterns $\times$ 2 landscape replicates $\times$ 2 epidemic replicates $=$ 20 epidemics were simulated.

Based on Soubeyrand et al. (2007), the individual dispersal function was assumed to decrease exponentially with distance. More precisely, the proportion of spores emitted from a source point z and arriving at a reception point z' was given by the individual dispersal function:

$$g(|| z - z' ||) = \frac{2\pi}{m_0^2} \exp\left(\frac{2\pi}{m_0} || z - z' || \right).$$

where $||z - z'||$ is the distance between z and z', and m_0 is a range parameter. Dispersal rates among patches were computed according to the method described in chapter 3.

Since the effect of host spatial structure can be counterbalanced by the dispersal capacity of the pathogen (see chapter 4), we studied epidemics for both short range ($m_0 = 150$m) and long range ($m_0 = 1500$m) dispersal conditions.

5.2.2.2 Life-cycle parameters

Pathogen genotypes were described by their non-spatial basic reproductive number, $R_0^{P_p,V_v} = e_{P_p,V_v} r_{P_p,V_v} T_{P_p,V_v}$, on each variety ($v = 1, 2$; $p = 1, 2, 3$; see Figure 5.1). The genotype P_3 was considered as a generalist since it had the same fitness on each variety, *i.e.*, $R_0^{P_3,V_1} = R_0^{P_3,V_2}$. On the contrary, P_1 and P_2 were both specialist genotypes on variety V_1 and V_2, respectively, in a symmetric way: $R_0^{P_1,V_1} = R_0^{P_2,V_2}$ and $R_0^{P_1,V_2} = R_0^{P_2,V_1}$.

All life-cycle parameters could potentially depend on both the variety and the pathogen genotype. Here we fixed τ, T and r respectively to 5 days, 10 days and 2 spores per days for both varieties and the three genotypes. The different values of R_0 were obtained by varying the infection efficiency. Infection efficiency of the generalist genotype, P_3 was fixed to 0.1, leading to $R_0^{P_3,V_1} = R_0^{P_3,V_2} = 2$. For the specialist genotypes, e_{P_1,V_1} and e_{P_2,V_2} were fixed to 0.11, leading to $R_0^{P_1,V_1} = R_0^{P_2,V_2} = 2.2$ and e_{P_1,V_2} and e_{P_2,V_1} varied in $\{0.05, 0.07, 0.08, 0.09\}$, which resulted in $R_0^{P_1,V_2}$ and $R_0^{P_2,V_1}$ varying in $\{1, 1.4, 1.6, 1.8\}$.

In the following, for each specialist, the variety with the highest R_0 will be referred to as the susceptible variety whereas the other one will be referred to as the resistant variety. With regards to P_1, V_1 is the susceptible variety and V_2 is the resistant variety, and it is the opposite for P_2. With regards to P_3, both V_1 and V_2 are susceptible (see figure 5.1).

5.2.3 Case studies

Invasion process of a resident population by a new genotype can be split into three phases: mutant introduction by long dispersal events or by mutation, establishment and spread of mutant population, and takeover or coexistence with the resident genotype(s). We study in this chapter the establishment phase (section 5.3) and the conditions in which genotypes are able to coexist in the landscape (section 5.4).

In a first step (section 5.3), the generalist genotype P_3 was continually introduced by mutation of specialist genotypes P_1 and P_2 (at rate $\mu = 10^{-5}$). Because foundation effects could be of prime importance in the establishment phase, we assumed no growth of the host ($\delta = 0$) and we started the epidemic with 20 infectious lesions (10 of P_1 and 10 of P_2) on a randomly chosen V_1 patch, and 20 infectious lesions (10 of P_1 and 10 of P_2) on a randomly chosen V_2 patch.

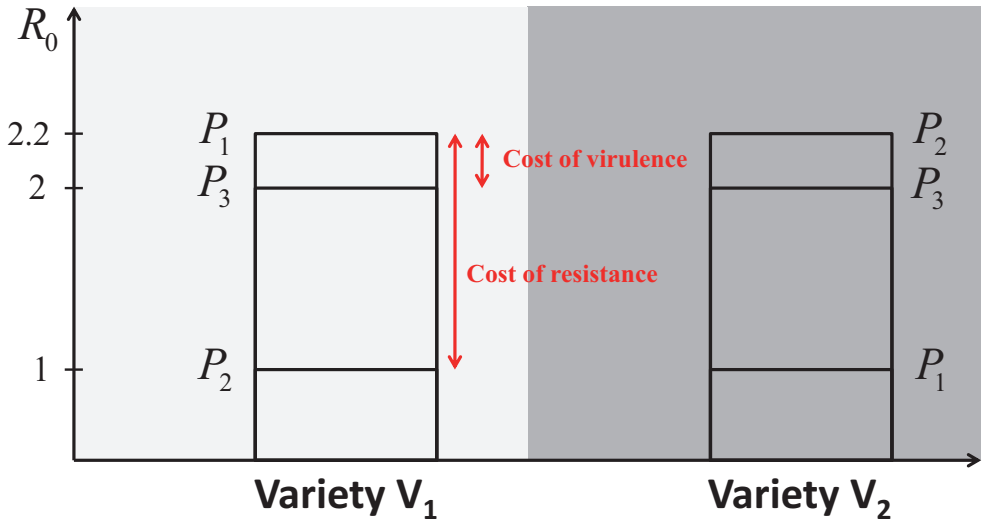


Figure 5.1: Basic reproductive number of each pathogen genotype on each variety. P_1 and P_2 are specialists on variety V_1 and V_2, respectively. P_3 is a generalist genotype. V_1 is susceptible to P_1 and P_3; V_2 is susceptible to P_2 and P_3.

In a second step (section 5.4), we assumed that all genotypes pre-existed and we studied the coexistence between genotypes, We set the mutation rate to 0, and we studied the equilibrium of the pathogen population composition in a growing host population ($\delta = 0.1$). Epidemics started with $30 + 30 + 30$ infectious lesions of genotypes P_1, P_2 and P_3 in all patches.

5.3 Establishment of mutant population

The ability of the mutant P_3 to establish a population was evaluated (i) by computing its relative abundance (population size) at the scale of the landscape and (ii) by establishing its spatial distribution among fields, in order to check whether it was diffuse across the landscape or whether localised populations emerged. In this latter case, we investigated in more details how host spatial structures allowed local emergence of the mutant. Relative abundance of P_3 at the field-scale was computed by averaging its relative abundance in intra-field patches.

5.3.1 Global population size

Figure 5.2 shows the relative population size of the mutant genotype P_3 at the landscape scale. It appears clearly that it was very difficult for the mutant to establish a population as soon as the basic reproductive number of the specialists on the resistant host was greater than 1. It also appears that the global relative abundance of P_3 varied greatly among landscapes.

In all cases, it was more difficult for the mutant to establish a population in grouped strategies. Mixed strategies were the most favourable environments for the mutant to get established when the variety proportions were unbalanced, whereas it was the mosaics when the variety proportions were balanced. Increasing the dispersal range made the establishment of P_3 easier for mixed and mosaic strategies.

When the dispersal range was large, V_1 proportion = 70% and $R_0^{P_2,V_1} = 1$ (figure 5.2c), P_2 was not able to establish a population for mixed and mosaic strategies, which left more room for P_3.

5.3.2 Intra field population size

Intra-field relative abundance of P_3 was computed only for $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$, *i.e.* when P_3 was able to establish a population. Figure 5.3 displays the distribution of the intra-field abundance of P_3 and shows contrasted situations.

When the spore dispersal range is large (figure 5.3g to l), the population distribution tails are short, which indicates a diffuse population across the landscape: P_3 is present in all or most fields but in very low proportions. This is consistent with the fact that long range dispersal mixed populations and did not lead to spatial structuring. Note that figure 5.3i and l confirm that, for the grouped strategy of variety allocation and long spore dispersal range, the mutant never established a population (*cf* figure 5.2c and d).

On the contrary, for the low dispersal range, several distribution patterns can be observed, depending on the variety allocation strategy (figure 5.3a to f). Mixed strategies exhibited shorter distribution tails than mosaic and grouped strategies, which indicates that mixed strategies resulted in more diffuse P_3 populations (Figure 5.4).

The spatial patterns of P_3 abundance differed in the mosaic strategy, for which 58% (figure 5.3b) to 52% (figure 5.3e) of the fields contained a P_3 population, and in the grouped strategy, for which only 20% of the fields contained a P_3 population. For the grouped strategy, most of fields were mutant free and the emergence of localised P_3 populations resulted from particular landscape structures (see section 5.3.3).

5.3.3 Local emergence of mutant populations

The high variability among simulations (figure 5.2) reflects a strong interaction between the landscape structure and the localisation in space of initial inoculum. Below we develop two examples.

Figure 5.5 illustrates the interaction between landscape spatial structure and initial inoculum position. In the first landscape (figure 5.5a and b), both initial inoculum sources

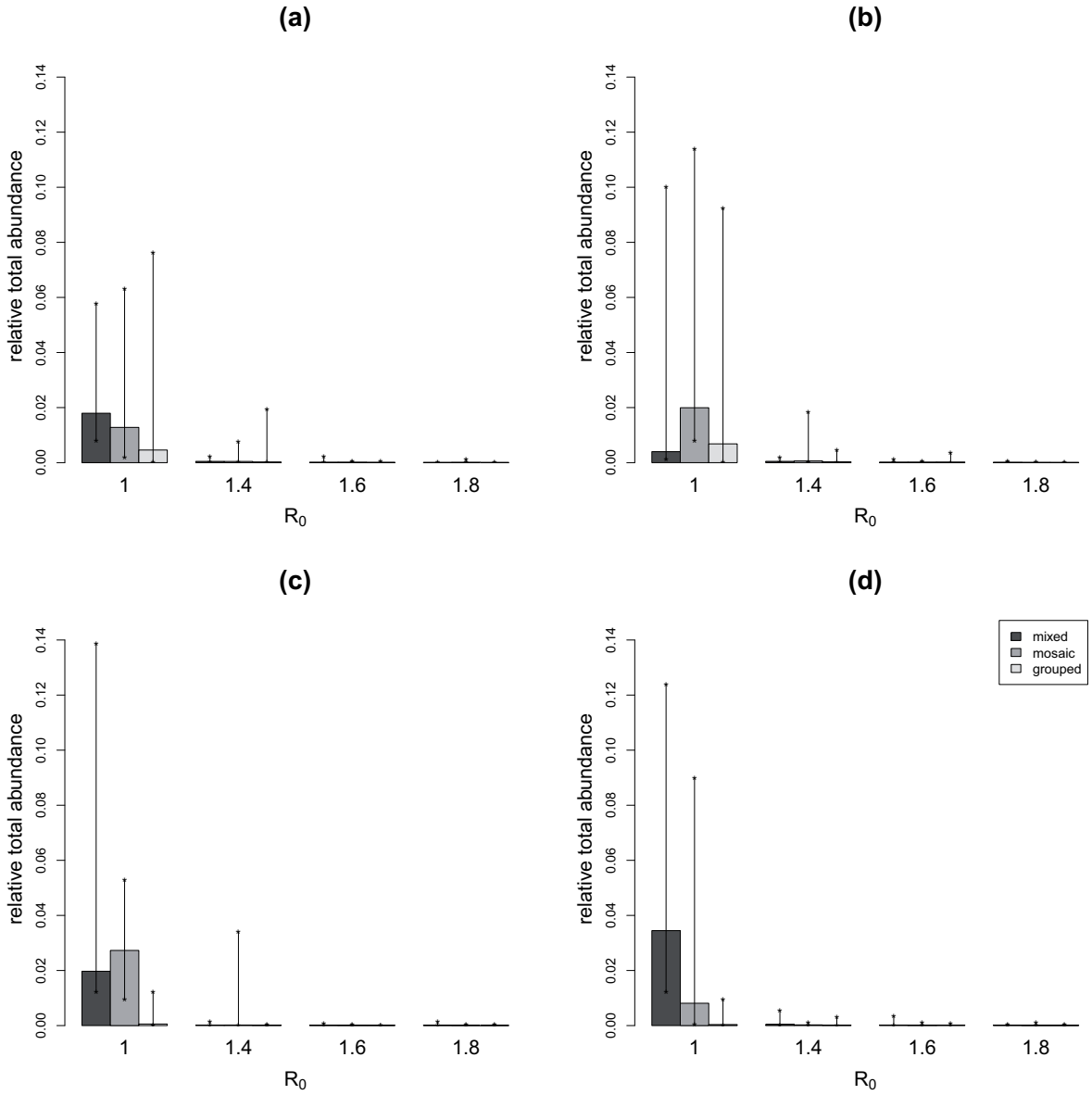


Figure 5.2: Relative abundances of genotype P_3 at the landscape scale plotted against the basic reproductive number of specialists on the resistant host and for different variety allocation strategies: a and b, short dispersal range ($m_0 = 150$m); c and d, large dispersal range ($m_0 = 1500$m); a and c, unbalanced variety proportions ($30\%-70\%$); b and d, balanced variety proportions ($50\% - 50\%$). Grey bars indicate the median over the 20 simulations and the black solid line the $[2.5\%, 97.5\%]$ quantile interval. From the darkest to the lightest grey: mixed, mosaic and grouped variety allocation strategies.

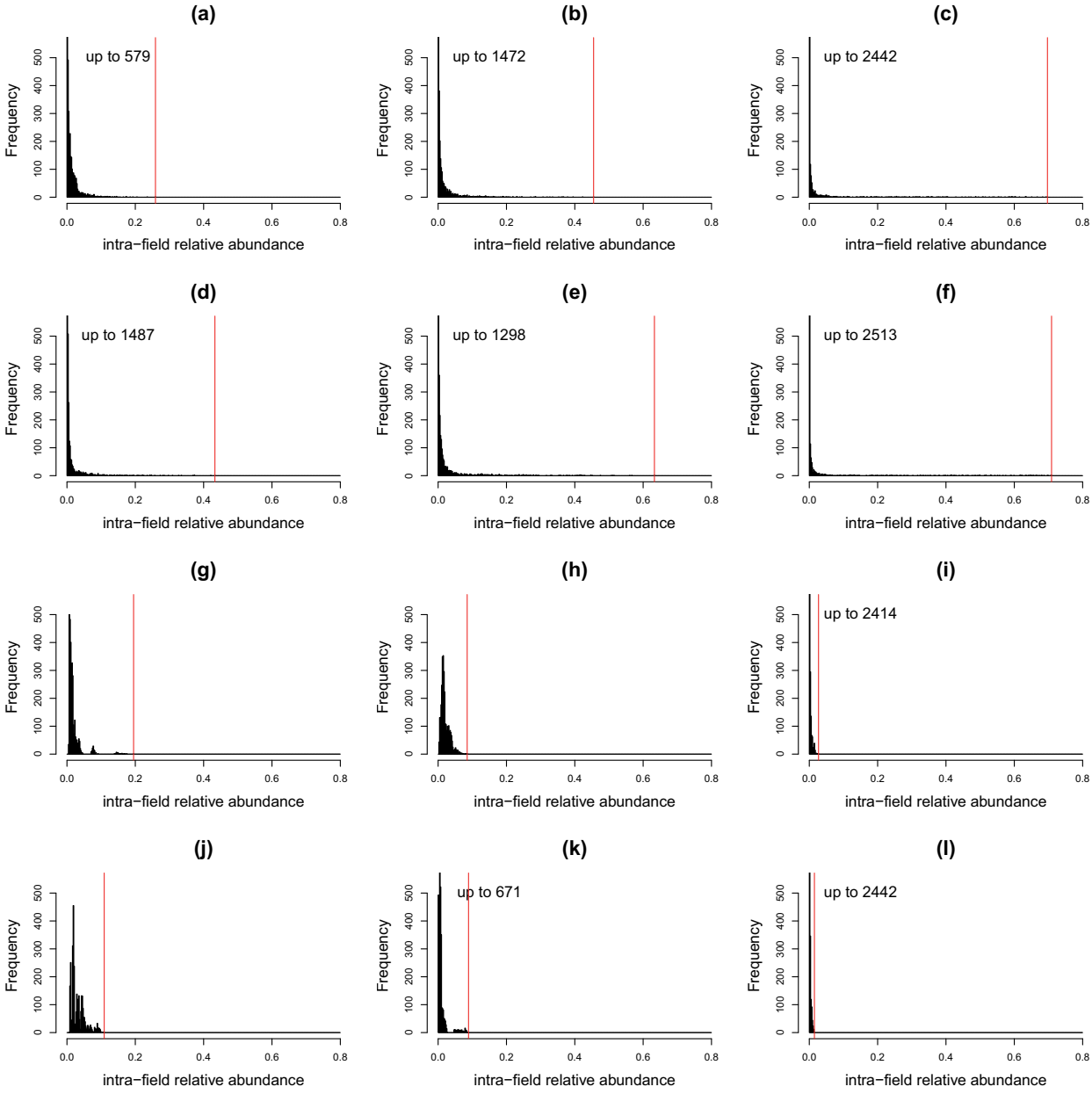


Figure 5.3: Histograms of intra-field relative abundance of P_3 for 3080 fields corresponding the combinations of variety proportions × variety aggregation levels (around 150 fields × 5 field patterns × 2 variety allocation replicates × 2 epidemic replicates = 3080 fields). Left column, mixed strategy; middle column, mosaic strategy and right column, grouped strategy. a,b,c, unbalanced variety proportions ($30\% - 70\%$) and short dispersal range ($m_0 = 150$m); d,e,f, balanced variety proportions ($50\% - 50\%$) and short dispersal range ($m_0 = 150$m); g,h,i, unbalanced variety proportions ($30\% - 70\%$) and large dispersal range ($m_0 = 1500$m); j,k,l, balanced variety proportions ($50\% - 50\%$) and large dispersal range ($m_0 = 1500$m). The red solid line indicates the upper bound of the histogram. Other parameters: $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$. For technical reasons, the y-axis is truncated but the highest y value is indicated on each graph when necessary.

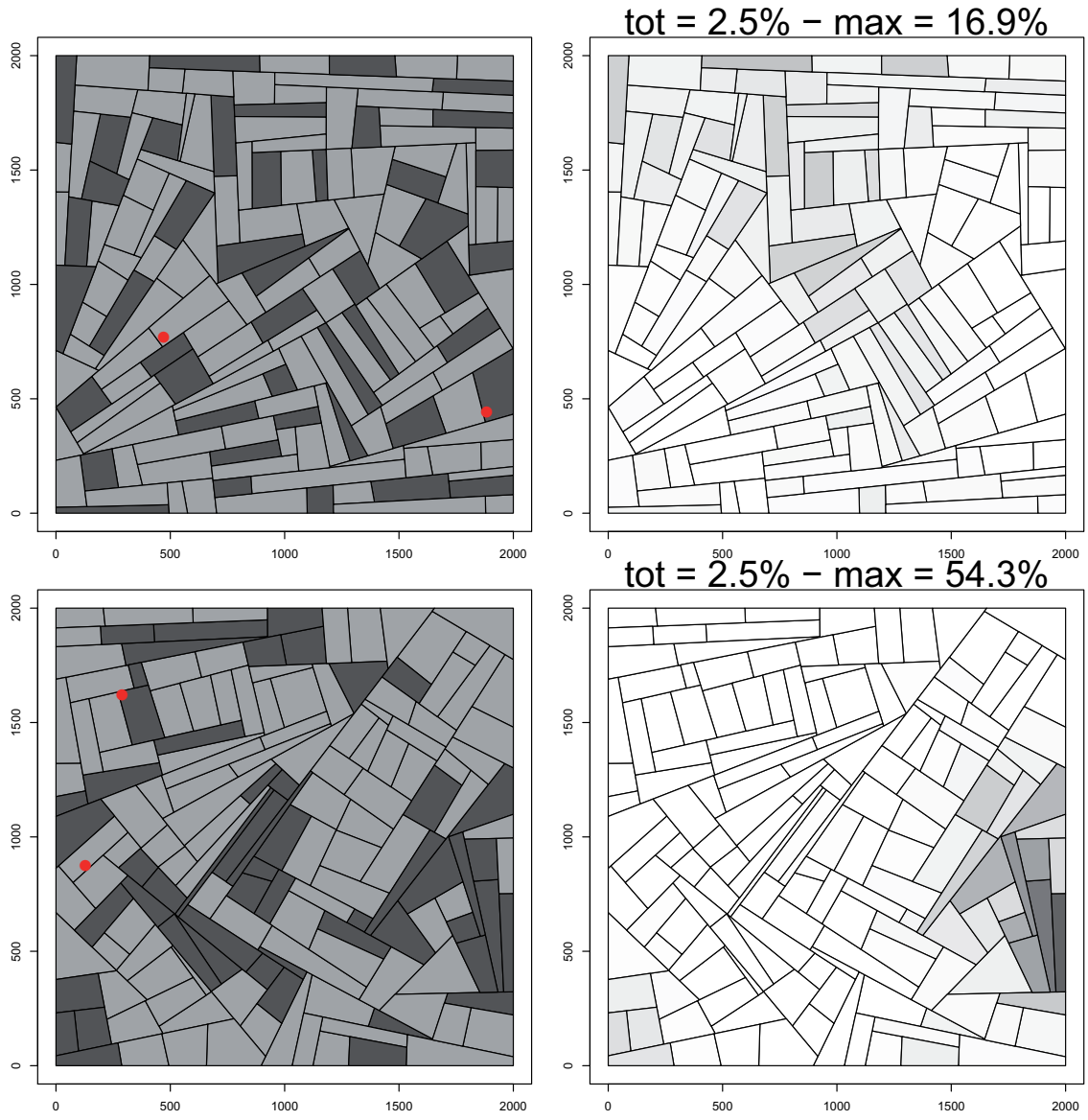


Figure 5.4: Relative abundance of genotype P_3 at the field scale landscapes with mixed (top line) *vs* mosaic (bottom line) variety allocation strategy. The left column shows the landscape structure. Light grey: V_1 (70%), dark grey: V_2 (30%), red point: initial position of inoculum. The right column shows the frequency of P_3. Grey scale: frequency of P_3 from 0 (white) to 0.55 (black), tot: relative abundance value at the landscape scale, max: maximum of the intra-field relative abundance values. Other parameters: $m_0 = 150$m, $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$. This is an example of a simulation.

are relatively central and the two variety aggregates do not present any particular structure. In such a pattern, both specialist pathogens rapidly spread on their respective susceptible varieties and P_3 could not establish a population. On the contrary, in the second landscape (figure 5.5c and d), the position of initial inoculum relative to the geometric form of the V_2 aggregate resulted in constraints for the spread of P_1, which allowed P_3 to locally establish a population.

Figure 5.6 illustrates the same kind of interaction with a completely different landscape structure, corresponding to the mixed strategy. When both P_1 and P_2 epidemics started from positions that are relatively close in space, both specialists spread concomitantly over the landscape and did not leave room for the development of a P_3 population. In that case, P_3 was present in the global population (2.1%) but locally at low frequency (11.7% at most). When the initial position of one of the specialists was central in the landscape whereas the other specialist started its own epidemic from a peripheral position, this produced a spatio-temporal lag in the spread of P_1 and P_2, giving the generalist the opportunity to develop a population on several fields of the landscape.

More generally, it was shown (figure 5.2b) that the mosaic strategy was more favourable for P_3 when V_1 and V_2 were at the same proportion in the landscape. Indeed, in such landscape structures, natural barriers and particular geometric forms increased the probability of successful foundation effects for the mutant.

5.4 Coexistence among pathogen genotypes

In this section, we suppose that the three genotypes P_1, P_2 and P_3 are initially present in the pathogen population and we study how the landscape structure influences the coexistence between those genotypes. All patches were inoculated with the same quantity of each pathogen genotype and the host was continuously growing. In a first step, we computed the equilibrium of infectious lesions at the landscape scale. Then we computed the frequency of each genotype (infectious lesions) at the field scale by averaging the values of the intra-field patches. This allowed to establish distribution maps of the genotypes at the landscape scale.

5.4.1 Stable coexistence

Figure 5.7 shows the output of the competition between P_1, P_2 and P_3 at the landscape scale. For the balanced variety proportions, we observed three situations: the generalist alone, coexistence between the specialists and the generalist and both specialists without the generalist. For unbalanced variety proportions (70% of V_2), two other situations were observed: the specialist of V_2 together with the generalist and the specialist of V_2 alone.

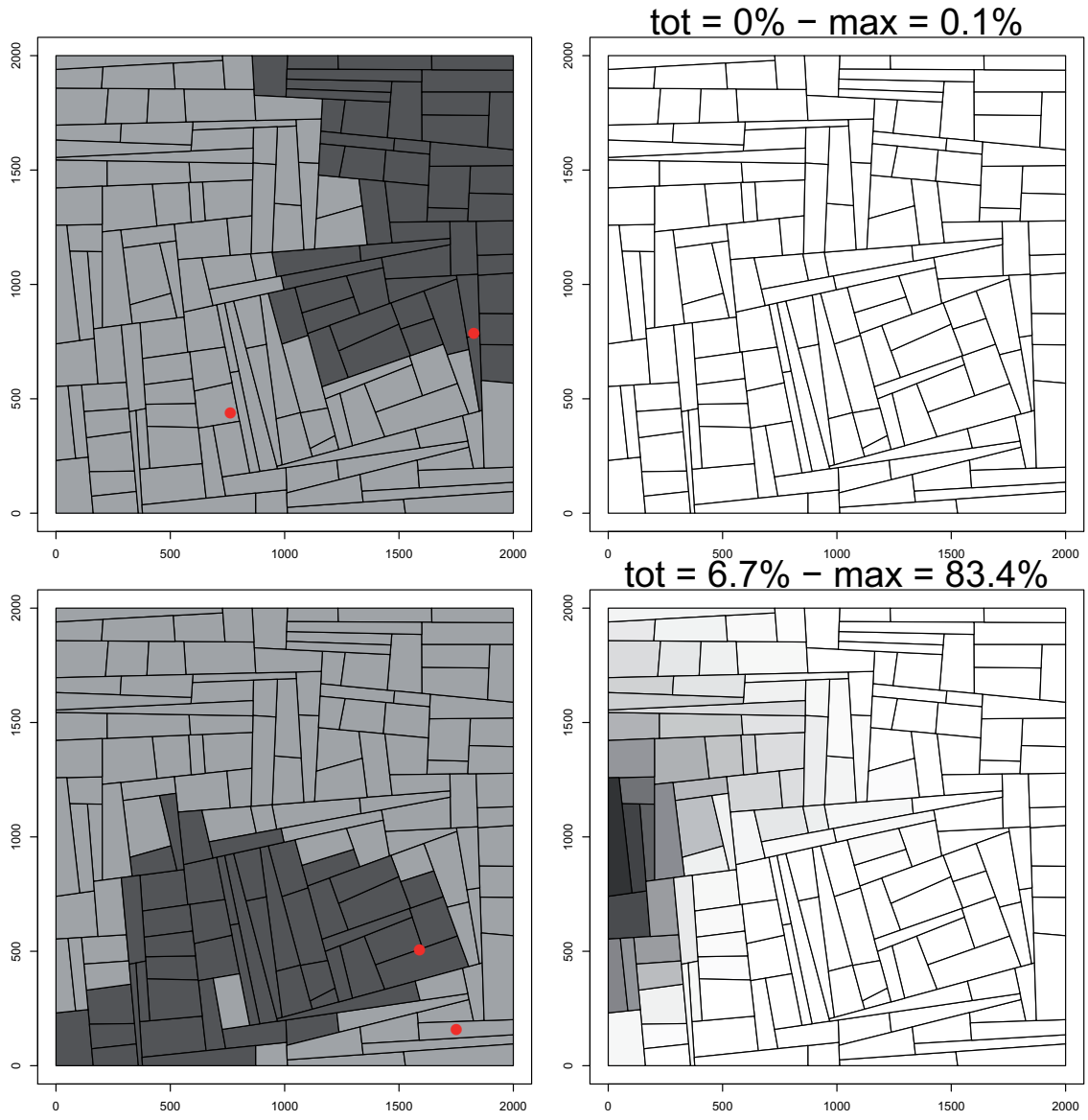


Figure 5.5: Relative abundance of genotype P_3 at the field scale in two landscapes with a grouped variety allocation strategy. The left column shows the landscape structure. Light grey: V_1 (70%), dark grey: V_2 (30%), red point: initial position of inoculum. The right column shows the frequency of P_3. Grey scale: frequency of P_3 from 0 (white) to 0.70 (black), tot: relative abundance value at the landscape scale; max: maximum of the intra-field relative abundance values. Other parameters: $m_0 = 150$m, $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$. This is an example of a simulation.

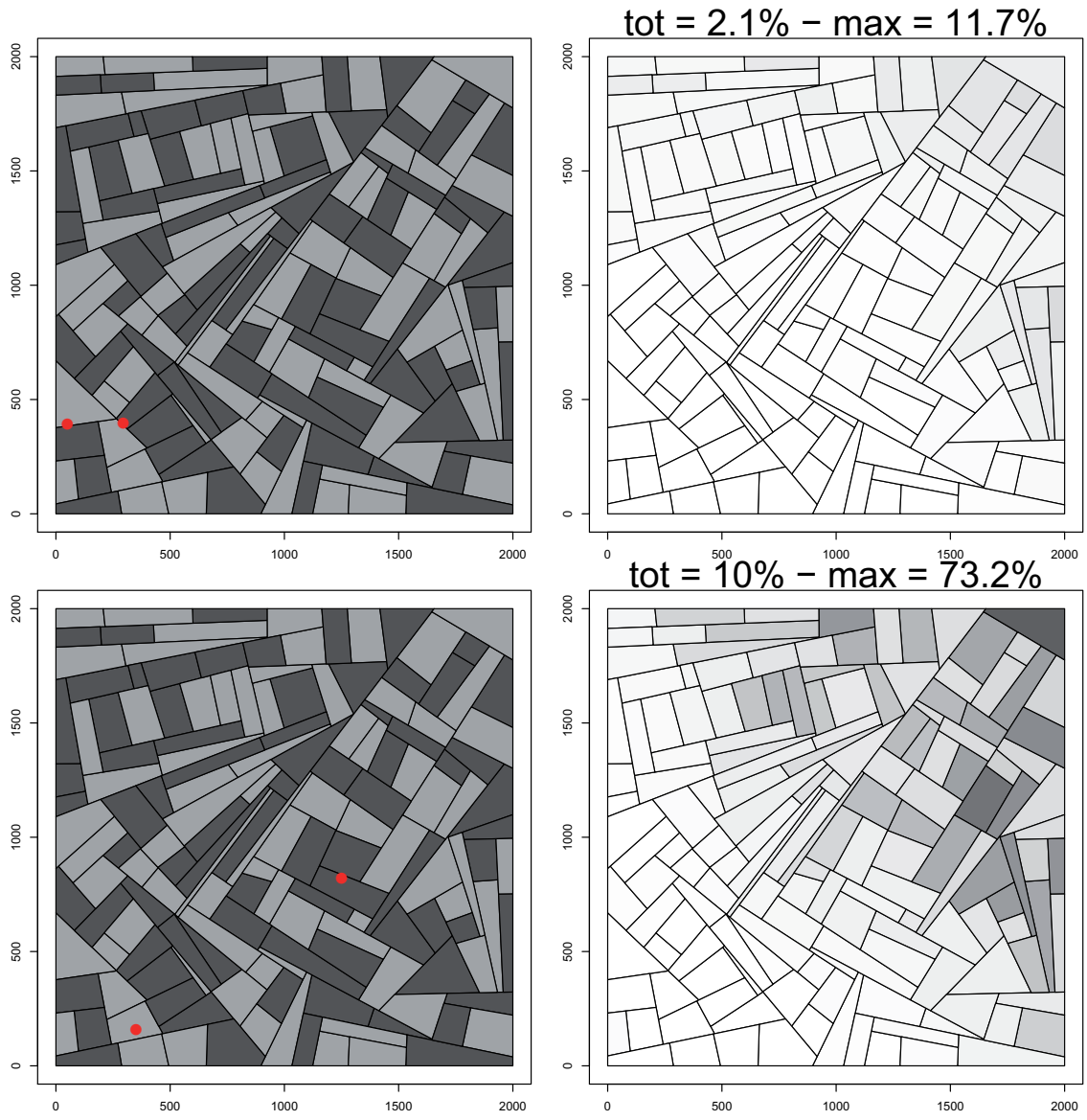


Figure 5.6: Relative abundance of genotype P_3 at the field scale in two landscapes with mixed variety allocation strategy. The left column shows the landscape structure. Light grey: V_1 (70%), dark grey: V_2 (30%); red point: initial position of inoculum. The right column shows the frequency of P_3. Grey scale: frequency of P_3 from 0 (white) to 0.70 (black), tot: relative abundance value at the landscape scale, max: maximum of the intra-field relative abundance values. Other parameters: $m_0 = 150$m, $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$. This is an example of a simulation.

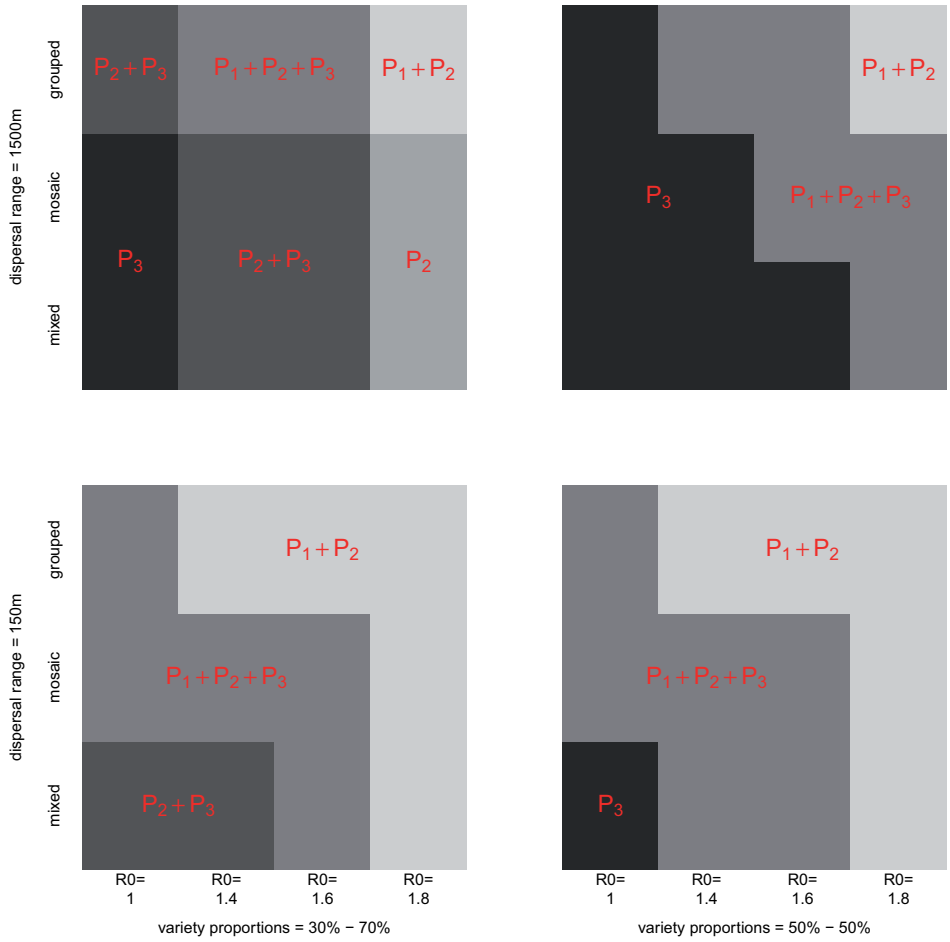


Figure 5.7: Conditions for coexistence between pathogen genotypes P_1, P_2 and P_3. Top line, large dispersal range ($m_0 = 1500$m); bottom line, short dispersal range ($m_0 = 150$m); left column, unbalanced variety proportions ($30\% - 70\%$); right column, balanced variety proportions ($50\% - 50\%$). y-axes, variety allocation strategies; x-axes, $R_0^{P_1,V_2} = R_0^{P_2,V_1}$.

Overall, the generalist genotype was favoured by mixing varieties, by decreasing the ability of specialists to reproduce on the resistant host and by increasing the pathogen dispersal range.

When the pathogen dispersal range was large, the specialists excluded the generalist for high variety aggregation levels or when their reproduction rate on the resistant host was high. When the dispersal range was low, changes in variety proportion did not change greatly the competition output, excepted for the mixed strategy. For equal variety proportions, the specialists presented a symmetrical behaviour and co-existed or not with the generalist. For unbalanced variety proportions, one of the specialists could be excluded.

5.4.2 Spatial distribution at population equilibrium

Figure 5.8 illustrates the spatial patterns of pathogen genotype distributions at population equilibrium, for three variety allocation strategies. For the mixed strategy (figure 5.8, left column), P_1 did not persist, P_2 was present (with only a few exceptions) in all V_2 fields but was more abundant in local aggregates of V_2, and the generalist P_3 developed in all fields. For the mosaic strategy, (figure 5.8, central column), P_1 was now able to persist in V_1 aggregates, P_2 became more abundant, and P_3 was concentrated in the most heterogeneous zones, showing an apparent spatial structure. Finally, for the grouped strategy (figure 5.8, right column), both specialists were highly abundant in the aggregate of their respective susceptible host and totally absent in the other fields, whereas the generalist genotype was restricted to the borders of the variety aggregates.

These examples can be generalised: figure 5.9 shows that the specialist abundance is highly related to the number of neighbour fields with an identical variety. On the contrary, the generalist tended to be more abundant in heterogeneous zones. This relationship was stronger for more aggregated variety distributions.

The spatial distribution of the pathogen population was however greatly influenced by its ability to disperse. Figure 5.10 presents the case of a grouped strategy with a fairly high multiplication rate of P_1 and P_2 on their resistant resistant variety ($R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1.8$). Here, the generalist did not persist. For the low dispersal range, each specialist was more or less limited to its susceptible variety (figure 5.10d and f). By increasing the dispersal range, P_1 and P_2 became more mixed (Figure 5.10a and c).

Figure 5.11 compares mixed and grouped strategies when the pathogen dispersal range is large and variety proportions are the same for V_1 and V_2. Here, the three genotypes coexisted. When the varieties were mixed, there was no spatial structure in the pathogen population. When the varieties were grouped, the specialists were restricted to their respective susceptible host and the generalist was evenly distributed over the whole landscape. This last situation can be explained by a habitat boundary polymorphism (Débarre and Lenormand, 2011), the high dispersal ability leading to a generalised distribution of P_3. On the contrary, the first situation is closer to a non-spatial landscape structure (due to the high dispersal range and the low aggregation level of varieties). As a consequence, basic reproductive number of genotypes can be averaged at the landscape scale. We obtained a mean basic reproductive number identical for all the three genotypes ($R_0 = 2$). We thus have three equivalent genotypes evolving in a quasi-neutral model.

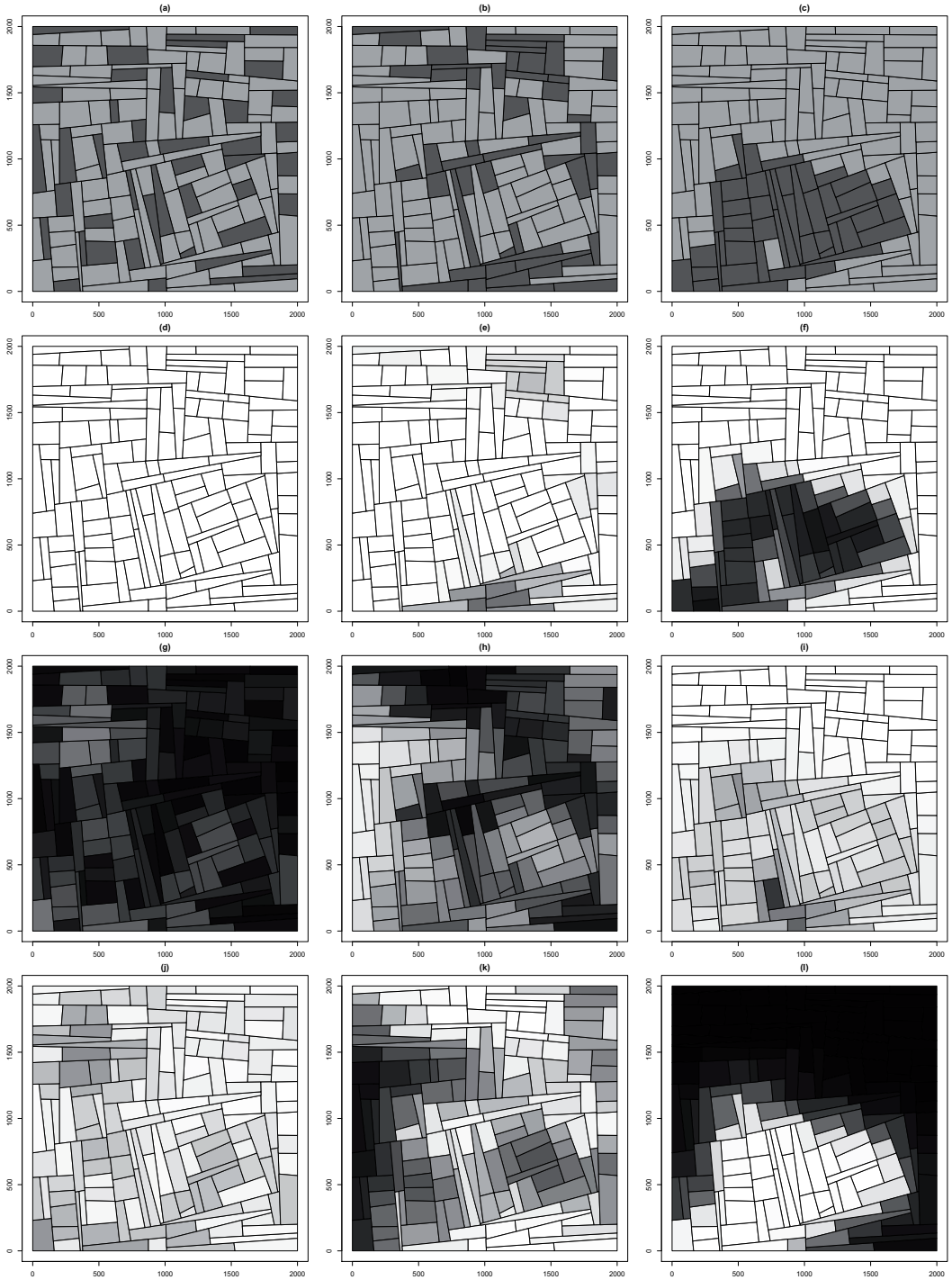
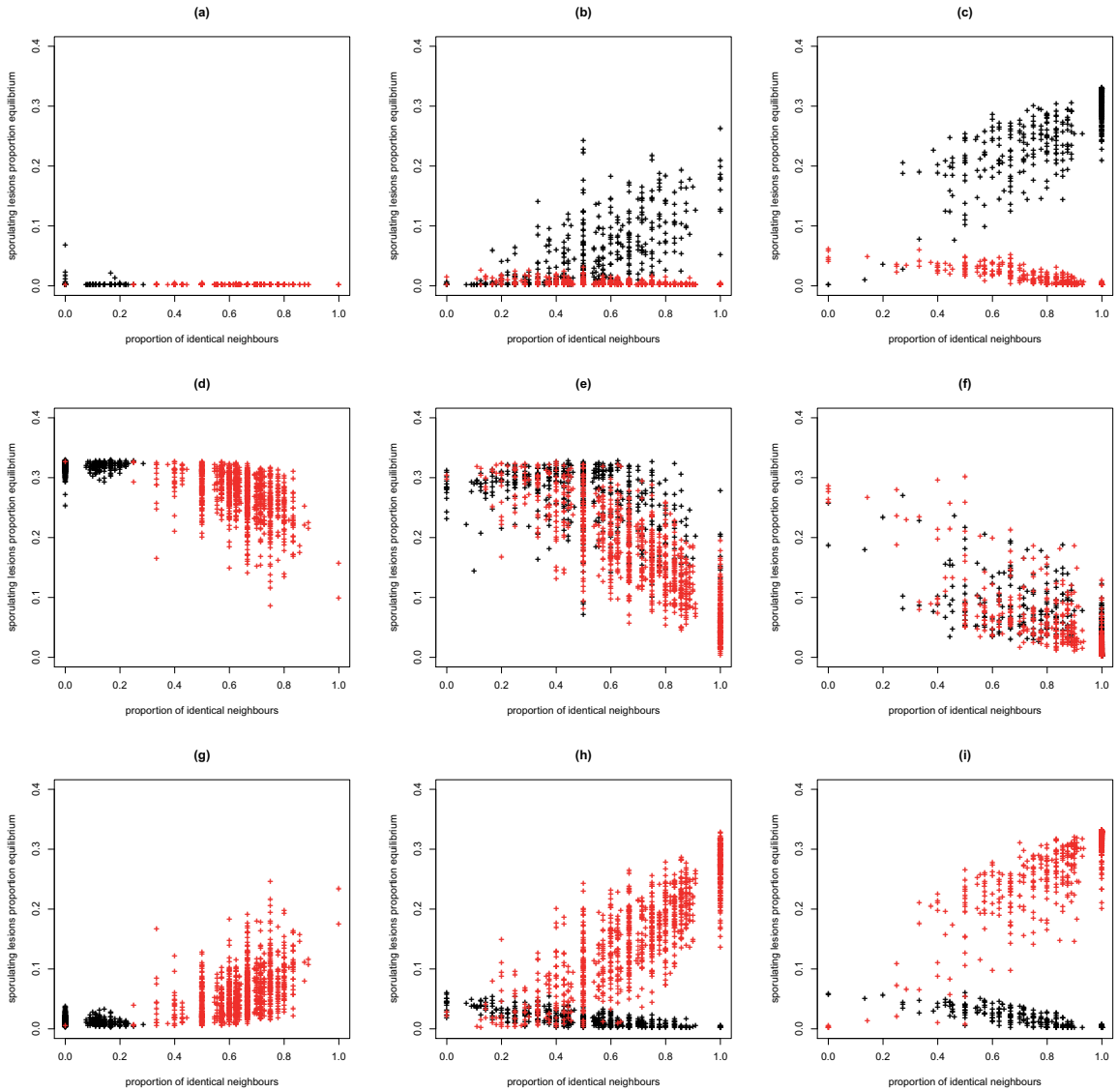


Figure 5.8: Coexistence of two specialists and a generalist pathogen in landscapes with three variety allocation strategies: mixed (left column), mosaic (central column) and grouped (right column). Top line: landscape structures (dark grey: V_1, 30%; light grey: V_2, 70%). Relative abundance of each pathogen genotype at the population equilibrium: d,e,f: P_1 (specialist of V_1); g,h,i: P_3 (generalist) ; j,k,l : P_2 (specialist of V_2). The grey scale indicates the genotype frequency, from 0 (white) to 0.35 (black). Other parameters are: $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$ and $m_0 = 150$m. This is an example of a simulation.



proportion of identical neighbours

Figure 5.9: The frequencies of infectious lesions in each of the 3080 fields (around 150 fields × 5 field patterns × 2 variety allocation replicates × 2 epidemic replicates = 3080 fields) at population equilibrium plotted against the proportion of neighbour fields that share the same variety. The three genotypes are P_1 (a, b, c), P_3 (d, e, f) and P_2 (g, h, i). The relative abundance of the different genotypes is in black for V_1 fields and in red for V_2 fields. Left column: mixed variety allocation strategy, middle column: mosaic variety allocation strategy, right column: grouped variety allocation strategy. Other parameters: V_1 proportion = 30%, V_2 proportion = 70%, $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$, $m_0 = 150$m.

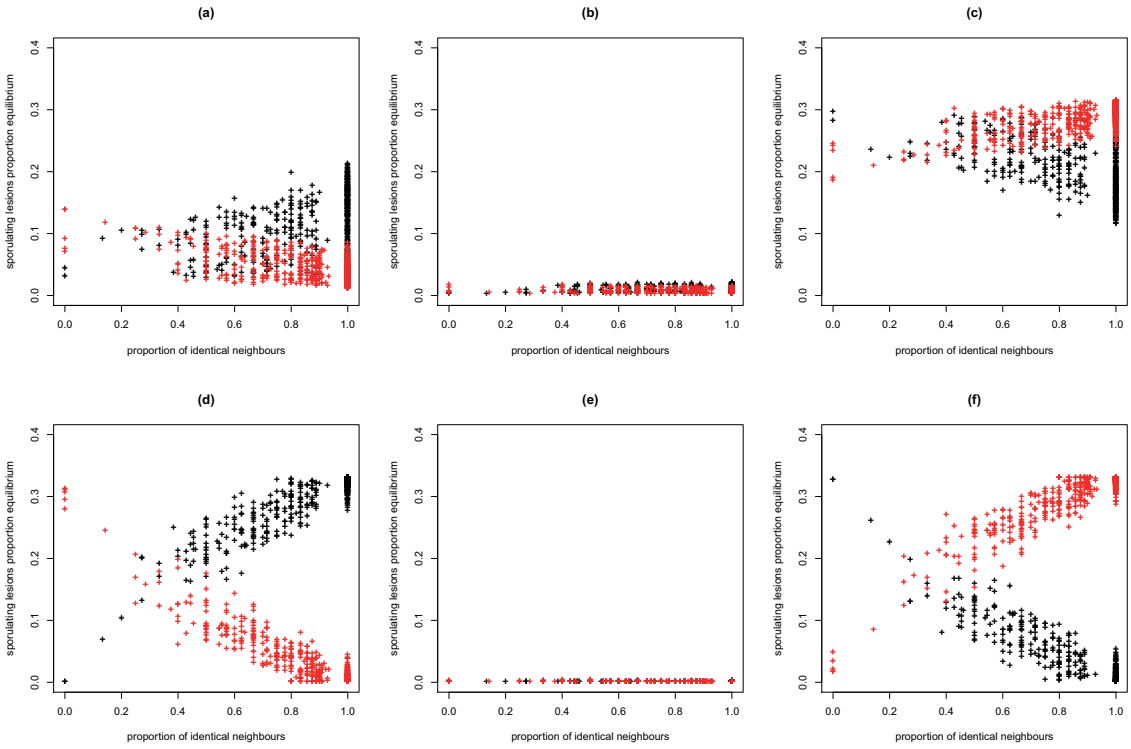


Figure 5.10: The frequencies of infectious lesions in each of the 3080 fields at population equilibrium plotted against the proportion of neighbour fields that share the same variety. The three genotypes are P_1 (a and d), P_3 (b and e) and P_2 (c and f). The relative abundance of the different genotypes is in black for V_1 fields and in red for V_2 fields. Top line: large dispersal range ($m_0 = 1500$m), bottom line: short dispersal range ($m_0 = 150$m). Other parameters: V_1 proportion $= 30\%$, V_2 proportion $= 70\%$, $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1$, $m_0 = 150$m.

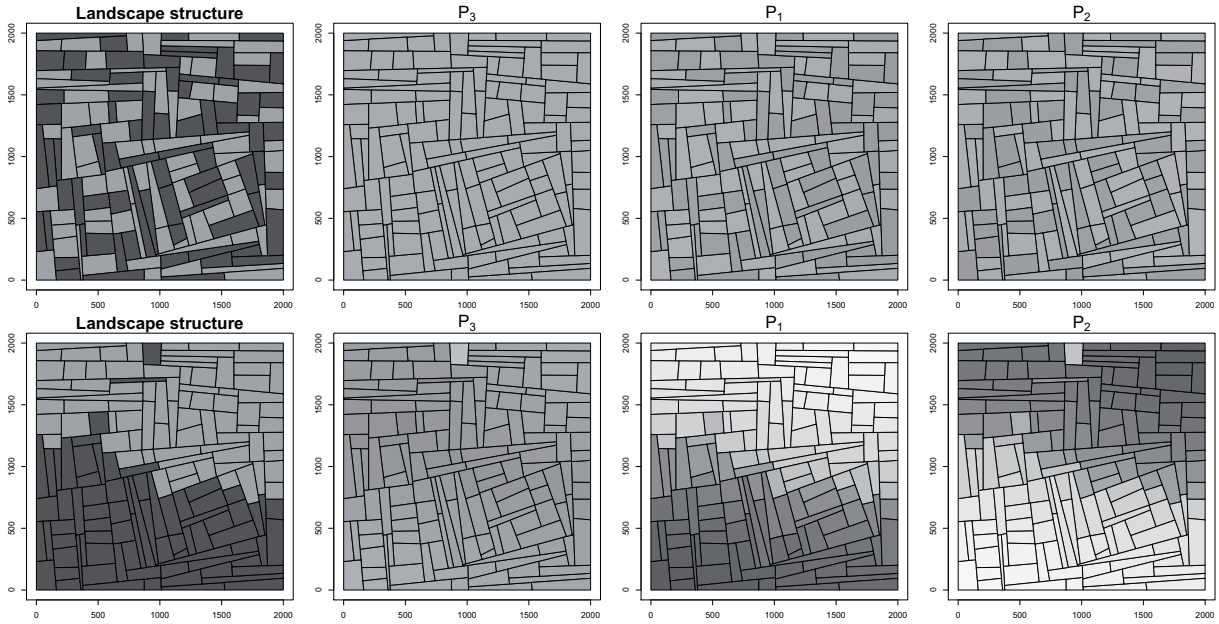


Figure 5.11: Spatial structure of the pathogen population when dispersal range is large ($m_0 = 1500$m). Left column shows the landscape structures. Dark grey: V_1 (50%), light grey: V_2 (50%). The other columns show the frequency of each pathogen genotype at the equilibrium. The grey scale indicates the frequency from 0 (white) to 0.35 (black). Top line: $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1.8$ and mixed strategy, bottom line: $R_0^{P_1,V_2} = R_0^{P_2,V_1} = 1.4$ and grouped strategy. This is an example of a simulation.

5.5 Discussion

During both the population establishment phase and the coexistence at population equilibrium, landscape structure had a strong effect on the mutant population. Highly fragmented landscapes were found favourable to the generalist genotype. When the variety aggregation level was high (mosaic and grouped landscapes), the generalist genotype had more difficulties establishing a population but the geometry and size of the variety aggregates interacted with the initial position of inoculums to produce local populations. Such effect could give opportunities for the generalist to locally build a population. Another output of the study was that coexistence between the three genotypes occurred rather easily.

During the establishment phase, two kinds of spatial structures were obtained for the mutant population. Highly fragmented landscapes led to diffuse P_3 populations, whereas higher aggregation levels of the varieties led to the establishment of local populations. This can be of practical interest for disease management since localised populations can be controlled more easily by crop rotations or local fungicide applications, while diffuse populations seem more difficult to handle. However, the establishment of strong

local populations of the generalist genotype make it less susceptible to demographic or environmental stochasticity. The high variability of the size of the mutant population was also discussed by Burton and Travis (2008). They found substantial variation in the fate of mutations depending on where they arise in space during the population expansion, in particular when corridor structures were present in the landscape. Corridor structures tended to slow down the resident population and so allowed the mutant population to get locally established through foundation effects. Our model also produced similar effects. On the contrary to Burton and Travis (2008), we considered a resident population composed by two genotypes each one specialised on one variety. In addition to the interaction between landscape structure and initial position of inoculums, there was also a strong effect of the relative position of initial inoculum of each genotype.

Hallatschek and Nelson (2009) developed and analysed a model that described the population dynamics and the genetic segregation on expanding microbial colonies. They proposed several relationship between the observed genetic patterns and the selective advantage / disadvantage of mutations. As an example, they showed that beneficial mutations give rise to sectors with an opening angle that depends on the selective advantage of the mutants. However, they did not consider heterogeneous environments. In our case, we were able to provide some relationship between landscape structure and observed pathogen genetic patterns. Nevertheless, the description of the structure of the pathogen genetic pattern and the shape of the variety aggregates needs to go further in order to extract as much information as possible on which landscape structure (in interaction with the initial position of inoculum) gives rise to which spatial pattern of the mutant population.

After the establishment phase (at the population equilibrium), the coexistence among pathogen genotypes was highly stable between landscape structures, as long as the variety proportions and aggregation level and the spore dispersal range did not vary. In accordance with Débarre and Lenormand (2011), we found that habitats boundaries provide opportunities for the generalist genotype to persist in the landscape. In our study, however, the 'habitat boundary polymorphism' acted within a more flexible framework, including complex landscape structures, an explicit life-cycle and stochasticity in the pathogen life events.

Papaïx et al. (2011) (chapter 2) describes situations of pathogen genotype coexistence and the consequences on observed disease level in the field. In chapter 2, we investigated the dynamics of leaf rust at the scale of France on the mainly grown wheat varieties. In particular, we were able to compare the effect of the local fitness (basic affinity in chapter 2) and the effect of the landscape composition on the proportion of a pathogen genotype sampled on a given wheat variety. We showed that a variety, by increasing the population size of a pathotype may significantly influence the composition of the pathogen population on other

varieties. A striking example is that of variety *Soissons* and its specialist pathotype (073100). In the 1990's, *Soissons* was highly popular and covered up to 40% of the wheat acreage in France. It selected a highly specialised and damageable pathotype (073100), which is in good accordance with the prediction of our model in the case of aggregated landscapes (section 5.4.1). At the beginning of the wheat leaf rust survey, the pathogen population sampled on *Soissons* was composed up to 60% of 073100. It is likely that the decrease of *Soissons* frequency resulted in a more fragmented distribution of that variety. Here again, the observed changes in the pathogen population meet the model predictions: the pathogen population on *Soissons* was progressively invaded by maladapted pathotypes originating from other varieties (remind that 073100 was not able to infect the recent varieties). The relative decline in 073100 accounted for the increase in Soissons' observed resistance level.

Habitat boundaries are highly diversified zones. In agro-ecosystems, particular boundaries are those between production and natural components of agricultural landscapes. It is recognised that the agro-ecological interface significantly influences disease dynamics and evolution of plant pathogen (Burdon and Thrall, 2008). Agro-ecosystems provide a range of situations including both agricultural crops and wild reservoirs. Our study highlights the importance of borders as a diversity reservoir contributes to answering the question of the effect of the agro-ecological interface on evolutionary changes in plant pathogens.

Part III

Long term evolution of pathogen populations

Foreword to part 3

In the part before, and especially in chapter 5, we have shown that the landscape structure has a key role for determining which strains win the competition and thus which strains compose the pathogen population. However, we have considered fixed phenotypes for the pathogen strains and the considered time scale was comparable to the pathogen life-span.

At a longer time scale, evolution of pathogen populations results from a succession of new strains apparition (called mutant strains), which will compete with the resident pathogen population. In an idealized world, the evolution process take more time than the demographic one. As a consequence, mutant strains appear once the pathogen population composition has reached its equilibrium. In such conditions, the pathogen population evolves gradually up to a generalist strain, which phenotype depends of the landscape composition. This generalist strain can be stable. In this case mutant strains are always eliminated after the competition period and the pathogen population remains monomorphic. But, the generalist strain may also be unstable: the population splits then in various specialised populations.

In this part, we opt for such a framework in order to study the evolution of pathogen specialisation in heterogeneous environments. We generalise the soft selection model of Levene (1953) to any number of patches interconnected with any set of pairwise dispersal values. We then analyse this model via both analytical and simulation approaches.

Chapter 6

Evolution of specialisation in spatial metapopulations

This chapter is based on a article project by Julien Papaïx, Oliver David, Christian Lannou and Hervé Monod

Contents

6.1 Introduction

Plant diseases have been a threat for crop production since the origin of agriculture (Stukenbrock and McDonald, 2008; Zadoks, 1982). From field to landscape, agro-ecosystems offer more conducive environments for the emergence and spread of pathogens, compared to natural ecosystems. In a crop field, plants are grown at a high density with identical or very similar genotypes grouped in the same place. Practises such as fertilisation and irrigation provide stable and favourable conditions for many diseases. At a larger scale, the shift from a complex and diversified natural environment to much more simplified and genetically uniform agrosystems over vast areas (Robinson and Sutherland, 2002) has made easier the adaptation of plant pathogens to their hosts (Johnson, 1981). The progresses in crop management in modern agriculture have ensured sufficient yields but present the major drawback to facilitate the occurrence and spread of highly specialised, and thus highly damaging, plant pathogens. In a period when European agriculture is attempting to massively reduce the use of chemical pesticides, the development of control strategies that hamper the evolution towards more specialised pathogens is a major challenge in crop protection. After a short review of pathogen specialisation in agricultural systems we present our approach, which took a more theoretical point of view.

In agricultural systems, anthropogenic disturbances have accelerated adaptive phenotypic changes (Palumbi, 2001; Hendry et al., 2008) leading to the emergence of new pathogen species (*e.g.*Stukenbrock et al., 2007; Gibbs et al., 2008) or to the differentiation and subsequent specialisation of pathogen sub-populations (*e.g.* Munkacsi et al., 2008; Gladieux et al., 2010). A well documented example is that of the rice-infecting lineage of *Magnaporthe oryzae*, a fungal pathogen causing devastating epidemics in crops. *M. oryzae* emerged between 5000 and 7000 years BP following a host shift from *Setaria millet* to rice (Couch et al., 2005). More recently (in 1989), a host shift of the same pathogen from rice to wheat is assumed to have caused the emergence of wheat blast in Brazil. Such major events have been described for several crop pathogens (for a review, see Stukenbrock and McDonald, 2008) and are probably unavoidable.

Although more discrete, pathogen specialisation to cultivated varieties is much more common and causes the most popular varieties to become increasingly susceptible to different diseases. Johnson (1961) used the term 'Man-guided' to describe how wheat selection for resistance during the 20th century, as well as growing practises, have shaped the *Puccinia* populations, a group of pathogens causing the 'rust' disease on wheat. Breeding for resistance in cultivated plants has largely relied on the exploitation of the 'gene-for-gene' system (Flor, 1971). According to this system, avirulence genes, in the pathogen, are matched by resistance genes, in the host. These qualitative resistance genes confer immunity to the plant, which explains their popularity in breeding programs. However they are easily overcome by

the pathogen after mutation or deletion of the avirulence gene and, in the recent past, the release of resistant varieties has usually led to the adaptation of pathogen populations through the accumulation of qualitative pathogenicity factors. A very demonstrative example can be found by comparing the French populations of *Melampsora larici-populina*, a pathogen causing poplar rust, on wild and cultivated host populations (Gérard et al., 2006).

The shape of pathogen populations driven by the 'gene-for-gene' system has been largely documented in agricultural systems (Wolfe and Schwarzbach, 1978; Hovmøller et al., 1993; Rouxel et al., 2003; Goyeau et al., 2006; Barrès et al., 2008) but is not sufficient for explaining the high specialisation level that is observed in crop pathogens. A recent study (Papaïx et al., 2011) based on a large database analysis coupling pathogen frequency data and disease observation data showed that pathogen specialisation in crops is largely accounted for by quantitative pathogenicity. This data exploration, together with a fitness study (Pariaud et al., 2009a), revealed in particular that a *Puccinia triticina* lineage was responsible of major epidemics on the widely grown wheat variety *Soissons*. When the frequency of that variety decreased, the pathogen lineage decreased as well and the observed disease levels became progressively lower. Yet, most of the leaf rust lineages present in France at the same period were still able to infect *Soissons* (*i.e.* were compatible according to the 'gene-for-gene' system). But these maladapted lineages only developed mild epidemics on that variety. This example illustrates well the fact that widely grown varieties may select specialised pathogen lineages and suggests that deployment strategies could be designed to prevent or limit the development of these lineages.

As underlined by Thrall et al. (2010), the application of evolutionary principles could provide relevant tools to maintain agriculture productivity while reducing environmental impacts. Such an application is illustrated by Hendry et al. (2011) as follow. The mismatch between the current phenotype and an optimal phenotype gives the adaptation level of a population to a given environment. A population with high adaptation level will exhibit high abundance while a large mismatch will limit population sizes. Thus, for pest control in agriculture, we wish to enlarge this mismatch, or keep it as large as possible, in order to limit pathogen adaptation to crop varieties. In a heterogeneous environment, migration counterbalances local adaptation leading to a decrease in the mean fitness of local populations (Lenormand, 2002). By altering the structure of agricultural landscapes, selection could be manipulated in order to influence pathogen adaptation. The role of spatial heterogeneity within agricultural landscapes on adaptation and specialisation of pathogen is thus of practical interest. These questions have been the subject of numerous theoretical studies in evolutionary biology. However, as we will show, existing theoretical approaches lack in the consideration of spatial perspectives. They must be extended to a framework with more flexible assumptions on the environment structure in order to match with the agricultural problematic. In this work we adopted the adaptive dynamics framework which is a powerful

approach for the study of ecological adaptation (Maynard Smith and Price, 1973; Geritz et al., 1998; Geritz and Gyllenberg, 2005; Waxman and Gavrilets, 2005). In the following of this introduction we briefly review spatial models for for the dynamics of adaptation in heterogeneous environment then present our work.

Spatial structures in the environment have been introduced with different perspectives in adaptive dynamics models. Some studies suppose that the habitat changes gradually with space, for example, for altitude or temperature (Doebeli and Dieckmann, 2003; Champagnat and Méléard, 2007). Doebeli and Dieckmann (2003) showed that the gradient strength, together with the moving distance of individuals, strongly influences the evolutionary outputs. A different approach, more suitable to agricultural systems, is to introduce explicit patches in the environment. Débarre and Gandon (2010) developed a soft selection model for a population evolving in a one-dimensional space divided into two different habitats. They showed that habitat differentiation and proportion determine the evolution into two specialists or a single generalist. To go beyond unidimensional environments, other authors have considered metapopulation structures (Hanski, 1998), in which a network of local populations were interconnected by dispersal. However, these studies are most of the time based on a spatially implicit description of the environment (Meszéna et al., 1997; Parvinen and Egas, 2004). Recently, Hanski et al. (2011) proposed an eco-evolutionary dynamics model for a spatially explicit metapopulation inhabiting a finite network of patches, and they studied the scale at which the population was adapted. Depending on gene flow and demo-genetic parameters they found that adaptation may be local, at the network scale or may lead to a mosaic specialisation. They did not, however, specifically address the question of the effect of habitat spatial structures on adaptation.

How do pathogens adapt to a heterogeneous environment? And how does landscape structure determine host specificity of pathogens? In this work, we propose to address these questions from a theoretical point of view based on adaptive dynamics. A spatially explicit metapopulation model is developed to investigate how selection, dispersal range, habitat proportion and habitat spatial structure interplay to influence the evolution of specialisation at the local and landscape scales. In the proposed model, the landscape is decomposed into a network of patches and individuals are described by one phenotypic trait that controls their fitness in each habitat. Both analytical and simulation studies are used in order to determine the evolutionarily stable phenotypes, the evolutionary speed and the evolution of the phenotypic variance. In section 2, the model is describe and an invasion analysis is performed to obtain general analytical results. Then, the role of landscape structure on the stability of the singular strategy is investigated by considering two types network structure. In section 3, the main analytical results are detailed assuming a hierarchical highly symmetric network structure. In section 4, they are applied numerically to spatially explicit lattice. In section 5, a simulation experiment is performed to study the dynamics of specialisation when

the population splits into two specialised phenotypes. Finally in the discussion section, we discuss the practical consequences of our findings for the design of sustainable strategies for disease control in agricultural landscapes.

6.2 Theoretical framework

This paper is based on a discrete-time deterministic population dynamics model for a metapopulation composed of several phenotypes that develop on a finite number of patches. We consider dispersion as a passive process only, *i.e.* there is no habitat choice. This model generalises the soft selection model of Levene (1953) to any number of patches interconnected with any set of pairwise dispersal values.

After presenting the model, we carry out an invasion analysis on this generic model (section 6.2.2). Conditions are given for a new mutant to grow and replace the resident phenotype assuming that the metapopulation is monomorphic. These results extend classical ones in adaptive dynamics to any metapopulation structure.

6.2.1 Model

6.2.1.1 Spatial heterogeneity

Consider a metapopulation that develops in a spatially heterogeneous environment consisting of a network of P patches. Two kinds of spatial heterogeneity, physical and biological, are involved in our model.

The physical heterogeneity acts on each individual on the same way, whatever its phenotype. Two components account for the physical heterogeneity, the carrying capacity of each patch and dispersal. Each patch j carries a finite and constant number K_j of individuals. We define the relative carrying capacity of patch j as $\overline{K}_j = K_j/K_T$, where $K_T = \sum_{j=1}^{P} K_j$ denotes the total carrying capacity of the environment. Dispersal is heterogeneous so that a propagule dispersed from a patch is deposited in another patch according to a specified dispersal distribution that is not necessarily uniform. The dispersal rate from patch j' to j is denoted by $m_{j'j}$. The number of propagules received by patch j is quantified by the input connection of patch j:

$$m_{+j} = \sum_{j'=1}^{P} m_{j'j}\, \overline{K}_{j'}.$$

In other words, the more propagules patch j receives, the more connected it is. This model is more general than most models in the literature since no assumption is needed on dispersal

rates. As a consequence, our model deals with any metapopulation structure with any set of pairwise dispersal values. In particular, dispersal rates can be computed from an individual dispersal function, which gives a unified framework to handle classical metapopulation as well as spatially explicit models.

Besides, several habitats are present in the environment and represent the biological nature of spatial heterogeneity. The environment is composed of H different habitats (or niches) with one habitat per patch. The habitat of patch j is denoted by $h(j)$. Each habitat k is in proportion $\pi_k = \sum_{j,h(j)=k}^{P} \overline{K}_j$ in the environment.

6.2.1.2 Description of individuals

Individuals are assumed to be haploid and are classified with respect to their phenotype. They reproduce asexually with non-overlapping generations and the progeny of an individual generally have the same phenotype than that parent. Phenotype i is characterised by the value, or strategy, x_i of a continuous trait x. The population size of phenotype i in patch j at time t is denoted by $n_{ij}(t)$. As the total population size is constant, it satisfies

$$\sum_{i=1}^{I(t)} \sum_{j=1}^{P} n_{ij}(t) = K_T,$$

where $I(t)$ is the number of phenotypes present in the metapopulation at time t.

Given its trait value x_i and the habitat encountered in patch j the survival probability of phenotype i in patch j is proportional to $f_{h(j)}(x_i)$. In this paper, the function $f_k(.)$ is assumed to be Gaussian with habitat-specific values of the optimal trait β_k and equal spread σ around the optimal trait (Geritz et al., 1998), so that

$$f_k(x) = \exp \frac{-(x-\beta_k)^2}{2\sigma^2}.$$

Differences between optimal traits β_k for the different habitats generate a trade-off, that depends on σ, between survival functions on the habitats: adaptation to a particular habitat causes maladaptation on the others. In particular, consider two habitats 1 and 2 with opposite values of the optimal trait. Then the survival functions of strategy x are $f_1(x) = \exp\left(\frac{-(x-\delta)^2}{2\sigma^2}\right)$ and $f_2(x) = \exp\left(\frac{-(x+\delta)^2}{2\sigma^2}\right)$ with $\delta = \beta_1 = -\beta_2$. The associated trade-off function, $u(\cdot)$, between survival functions in habitats 1 and 2 is defined by $f_2(x) = u\Big(f_1(x)\Big)$. As shown in Débarre and Gandon (2010), it satisfies

$$u(y) = \exp\left[-2\left(\frac{\delta}{\sigma} - \sqrt{-\frac{1}{2}\ln y}\right)^2\right]. \tag{6.1}$$

The differentiation between the two habitats is quantified by δ/σ. When $\delta/\sigma < 1$, the trade-off is weak. When $\delta/\sigma > 1$, the trade-off is strong. When δ/σ is close to one, the trade-off is very sensitive to the phenotype value x (see Débarre and Gandon, 2010, Fig.1c).

6.2.1.3 Metapopulation dynamics

The demography of the metapopulation is modelled using deterministic discrete-time equations. The model is based on the life cycle of individuals that involves the following sequence of events: reproduction, dispersal, selection and regulation. Reproduction rates are assumed to be constant among habitats and phenotypes, so we only present in detail the dispersal, selection and regulation phases.

During dispersal, a proportion $m_{j'j}$ of propagules produced in patch j' is deposited on patch j. So the number of individuals of phenotype i that is deposited on patch j is equal to

$$\sum_{j'=1}^{P} m_{j'j} n_{ij'}(t).$$

In each patch, new individuals are subject to a selection process with a survival probability of phenotype i in patch j proportional to $f_{h(j)}(x_i)$. This frequency-independent selection is followed by a non-selective competition for space. Thus, in patch j, a fraction

$$\frac{\left(\sum_{j'=1}^{P} m_{j'j}\, n_{ij'}(t)\right) f_{h(j)}(x_i)}{\sum_{i'=1}^{I}\left(\left(\sum_{j'=1}^{P} m_{j'j}\, n_{i'j'}(t)\right) f_{h(j)}(x_{i'})\right)}$$

of space is allocated to strategy x_i.

After the selection phase, local regulation for space makes the population size of phenotype i on patch j at time $t+1$, equal to

$$n_{ij}(t+1) \;=\; K_j \frac{\left(\sum_{j'=1}^{P} m_{j'j}\, n_{ij'}(t)\right) f_{h(j)}(x_i)}{\sum_{i'=1}^{I}\left(\left(\sum_{j'=1}^{P} m_{j'j}\, n_{i'j'}(t)\right) f_{h(j)}(x_{i'})\right)}. \tag{6.2}$$

6.2.2 Results on the evolution of a monomorphic metapopulation

To study the long-term phenotypic evolution of the population analytically, we consider a simplified and standard framework in which the metapopulation is monomorphic and evolves through episodic mutations of small amplitude (Geritz et al., 1998). Applying the theory of invasion analysis (Geritz et al., 1998) to the setting presented in section 6.2.1.3, we derive

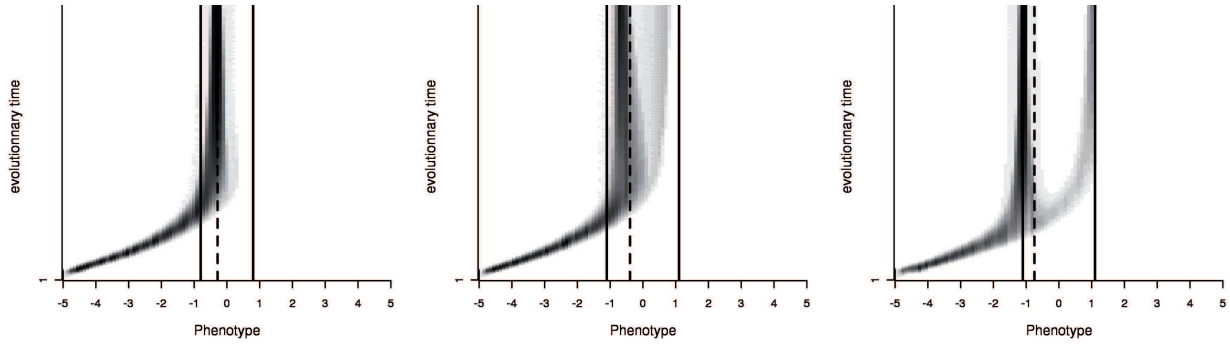


Figure 6.1: Examples of evolutionary trajectories simulated from equation (6.2), see section 6.5. From the left to the right the branching criterion (equation (6.9)) is equal to 0.6, 1.07 and 7.5. When it is lesser than one, the generalist strategy is stable (first panel). Otherwise, the population splits into two populations of specialists (middle and right panels). Dashed line: theoretical singular strategy ($x^\star$), solid lines: habitat optima (β_1 and β_2). Other parameters are: left panel: $\delta/\sigma = 0.8, m_s = 75\%, \pi = 0.32, \text{AI} = 0.1$; middle panel: $\delta/\sigma = 1.1, m_s = 75\%, \pi = 0.32, \text{AI} = 0.1$; right panel: $\delta/\sigma = 1.1, m_s = 15\%, \pi = 0.16, \text{AI} = 0.7$.

the 'evolutionarily singular strategy' ($x^\star$) and we characterise its stability. If the singular strategy is stable, the population remains monomorphic. Otherwise there is a branching point and the population becomes polymorphic, *i.e.* the environment selects for specialist phenotypes. Figure 6.1 gives some examples of evolutionary trajectories in both cases.

As we will show, a lot of results in the literature can be viewed as particular cases in our framework. The potential of such a framework is then demonstrated in two particular metapopulation structures in sections 6.3 and 6.4.

6.2.2.1 Short-term dynamics

The resident population is assumed to be monomorphic with trait value x_1. At a given time t_1, an individual produces a mutant with trait value x_2. At a time t_2 shortly after t_1 ($t_2 > t_1$, $t_2 - t_1$ small), the mutant is rare enough to consider that the resident population remains monomorphic ($n_{2j}(t) << n_{1j}(t)$ for all patches j and for $t_1 \leq t \leq t_2$). Under this approximation and according to equation (6.2), the number of mutants in patch j at time $t+1$ is equal to

$$n_{2j}(t+1) = \frac{\left(\sum_{j'=1}^{P} m_{j'j}\, n_{2j'}(t)\right) f_{h(j)}(x_2)}{\left(\sum_{j'=1}^{P} m_{j'j}\, K_{j'}\right) f_{h(j)}(x_1)} K_j.$$

Let $N_i(t) = \left(n_{i1}(t), \cdots, n_{iP}(t)\right)'$ denote the vector of the population sizes $n_{ij}(t)$ of phenotype i on patches $j = 1, \cdots, P$. The vector of mutant population sizes on each patch satisfy the

matrix equation $N_2(t+1) = A(x_1, x_2)\, N_2(t)$, where $A(x_1, x_2)$ is the $P \times P$ matrix which element in row j and column j' is equal to

$$[A(x_1, x_2)]_{jj'} = \frac{m_{j'j}\,\overline{K}_j}{m_{+j}} \frac{f_{h(j)}(x_2)}{f_{h(j)}(x_1)}. \tag{6.3}$$

Provided that $A(x_1, x_2)$ is an irreducible and primitive matrix, the growth of phenotype x_2 in the period after t_1 is approximately given by

$$N_2(t) \approx (\lambda^{(1)}(x_1, x_2))^{t-t_1}\, r^{(1)}(x_1, x_2)\, l^{(1)}(x_1, x_2)' N_2(t_1), \tag{6.4}$$

where $\lambda^{(1)}(x_1, x_2)$ is the dominant eigenvalue of $A(x_1, x_2)$, and $l^{(1)}(x_1, x_2)$ and $r^{(1)}(x_1, x_2)$ are its associated left and right eigenvectors, respectively (Caswell, 2001, chapter 4). Note that $A(x_1, x_2)$ is irreducible if and only if there is a dispersal path from each patch to every other patch, *i.e.* there are no disconnected subsets of patches or traps in the landscape, with respect to dispersal (Caswell, 2001, chapter 4).

6.2.2.2 Medium-term evolution

The mutant's fate is determined by its invasion fitness function $s(x_1, .)$ defined by $s(x_1, x_2) = \ln(\lambda^{(1)}(x_1, x_2))$ (Durinx et al., 2008, see also equation (6.4)). If $s(x_1, x_2) \leq 0$, the mutant becomes extinct shortly after t_1. If $s(x_1, x_2) > 0$, then the mutant either becomes extinct shortly after t_1 by genetic drift, or it increases in frequency in the population. Because our model dynamics are deterministic, only the latter alternative occurs.

Mutation steps are assumed to be small. Thus the mutant and resident phenotypes have close trait values x_1 and x_2. In this case, a stable dimorphism of x_1 and x_2 is not possible (Champagnat et al., 2006). A mutant with positive invasion fitness will eventually replace the resident and will generate a new monomorphic resident population with trait value x_2. It is assumed that mutations occur sufficiently infrequently so that the population becomes monomorphic before a new mutant appears.

6.2.2.3 Singular strategy

At a longer time scale, the resident population evolves gradually through a succession of mutant invasions towards a phenotype with trait value $x^\star$ called the 'singular strategy' (figure 6.1). As mutation steps are small, the first order approximation of the invasion fitness $s(x_1, x_2)$ is

$$s(x_1, x_2) \approx s(x_1, x_1) + \left.\frac{\partial s(x_1, x_2)}{\partial x_2}\right|_{x_2 = x_1} . (x_2 - x_1). \tag{6.5}$$

$\left.\frac{\partial s(x_1,x_2)}{\partial x_2}\right|_{x_2=x_1}$ defines the local fitness gradient (Geritz et al., 1998). Its sign determine the direction of selection: if $\left.\frac{\partial s(x_1,x_2)}{\partial x_2}\right|_{x_2=x_1} > 0$, then only mutants with $x_2 > x_1$ can invade, whereas if $\left.\frac{\partial s(x_1,x_2)}{\partial x_2}\right|_{x_2=x_1} < 0$, then this is only possible for mutants with $x_2 < x_1$. Thus, the population evolves until it reaches, the trait value $x^\star$ for which the local fitness gradient is zero, *i.e.* $\left.\frac{\partial s(x^\star,x_2)}{\partial x_2}\right|_{x_2=x^\star} = 0$. We show in the supporting information (appendix 4) that a monomorphic population has a unique singular strategy which trait value $x^\star$ is equal to

$$x^\star = \sum_{j=1}^{P} \frac{\overline{K}_j \left[l^{(1)}(x^\star,x^\star)\right]_j}{\sum_{j'=1}^{P} \overline{K}_{j'} \left[l^{(1)}(x^\star,x^\star)\right]_{j'}} \beta_{h(j)}. \tag{6.6}$$

Note that in our case, the singular strategy is always reachable by gradual evolution (convergence stability, see the appendix 4 of supporting information).

When dispersal rates are symmetric, *i.e.* $m_{j_1 j_2} = m_{j_2 j_1}$, the j^{th} component of the eigenvector $l^{(1)}(x^\star,x^\star)$ is equal to the input connection of patch j, m_{+j}. It follows that the singular strategy is then equal to

$$x^\star = \frac{1}{\sum_{j'=1}^{P} \overline{K}_{j'} m_{+j'}} \sum_{k=1}^{H} \Big(\sum_{j,h(j)=k}^{P} \overline{K}_j m_{+j} \Big) \beta_k.$$

In other words, the singular strategy is a weighted average of the habitat phenotypic optima β_k, which weights are an increasing function of the relative carrying capacities and input connections of the patches in the habitat. The singular strategy corresponds to a generalist phenotype since it represents a balanced strategy with respect to the habitat frequencies in the environment.

If all the patches have the same carrying capacity and the same input connection, the singular strategy is equal to

$$x^\star = \frac{1}{P} \sum_{j=1}^{P} \beta_{h(j)} = \sum_{k=1}^{H} \pi_k \beta_k. \tag{6.7}$$

In this case, the weights are equal to the habitat proportions and the spatial allocation of the different habitats has no impact on the singular strategy (see also Débarre and Gandon, 2010, Equation 17)).

6.2.2.4 Branching criterion

The second order approximation of the invasion fitness $s(x_1, x_2)$ is

$$s(x_1, x_2) \approx s(x_1, x_1) + \left.\frac{\partial s(x_1, x_2)}{\partial x_2}\right|_{x_2=x_1} . (x_2 - x_1) + \left.\frac{\partial^2 s(x_1, x_2)}{\partial x_2^2}\right|_{x_2=x_1} . \frac{1}{2}(x_2 - x_1)^2. \quad (6.8)$$

The stability of $x^\star$ is determined by the third term of Equation (6.8). If $\left.\frac{\partial^2 s(x^\star, x_2)}{\partial x_2^2}\right|_{x_2=x^\star} > 0$, then $x^\star$ is a branching point. On the contrary, if $\left.\frac{\partial^2 s(x^\star, x_2)}{\partial x_2^2}\right|_{x_2=x^\star} < 0$, then $x^\star$ is an 'evolutionarily stable strategy' (ESS, figure 6.1).

In supporting information (appendix 4) we show that the singular strategy $x^\star$ is evolutionarily stable if and only if

$$l^{(1)}(x^\star, x^\star)' \Delta(x^\star) \left(\mathrm{Id}_P + 2 \sum_{j=2}^{P} \frac{\lambda^{(j)}(x^\star, x^\star) r^{(j)}(x^\star, x^\star) l^{(j)}(x^\star, x^\star)'}{\lambda^{(1)}(x^\star, x^\star) - \lambda^{(j)}(x^\star, x^\star)} \right) \Delta(x^\star) r^{(1)}(x^\star, x^\star) < 1. \quad (6.9)$$

In this equation, Id_P is the identity matrix of order P, $\Delta(x)$ is the diagonal matrix whose diagonal elements are given by $[\Delta(x)]_{jj} = -\frac{x - \beta_{h(j)}}{\sigma}$ and prime denotes transposition. The eigenvalues of $A(x^\star, x^\star)$ in decreasing order are denoted by $\lambda^{(j)}(x^\star, x^\star)$, for $j = 1, \cdots, P$; the corresponding left and right normalised eigenvectors are denoted by $l^{(j)}(x^\star, x^\star)$ and $r^{(j)}(x^\star, x^\star)$. By 'normalised eigenvectors', we mean that $l^{(j)}(x^\star, x^\star)' r^{(j)}(x^\star, x^\star) = 1$, $l^{(j)}(x^\star, x^\star)' r^{(j')}(x^\star, x^\star) = 0$ for $j \neq j'$.

Equation (6.9) is too complex to be interpreted directly. However it opens the way to a better understanding of how environment heterogeneity and structure influence the stability of the singular strategy in specific cases. This will be illustrated in the next sections with hierarchical and spatially explicit metapopulations. In the following, the branching criterion will be defined as the quantity on the left-hand side of equation (6.9). It depends on the patch structure and on the habitat phenotypic optima.

6.2.2.5 Evolutionary speed

The model representing the population dynamics is described by equation (6.2). To obtain the first results on the singular strategy and its stability we have considered that the population remained monomorphic. In order to study the evolutionary speed up to the singular strategy two other assumptions must be made. First, a mechanism of mutant apparition must be provided as well as a description of the mutant trait value. This

mechanism is assumed to be stochastic and to alter the phenotype trait with a small variance. Second, when rare, the mutant population is assumed to be subject to demographic stochasticity.

The evolutionary speed of the phenotype trait x in a monomorphic population is approximately given by the canonical equation of adaptive dynamics (Champagnat et al., 2006; Dieckmann and Law, 1996). This equation is based on asymptotics with three nested time scales. We intuitively describe the principles here and give the technical details in supporting information (appendix 4).

At the finest time scale, time is discrete and the dynamics model of section (6.2.2.1) applies to each new generation. At medium time scale, the number of generations per time unit tends to infinity. Time appears as continuous but phenotype trait evolution still appears as discrete with a series of monomorphic resident metapopulations each identified by its unique phenotype. Each mutation is assumed to alter the phenotype trait with a small variance $\gamma^2(x)$ and to occur according to a continuous-time Poisson process with rate $\theta(x)$. In addition, the stochasticity that affects the demography of the mutant when it is still rare is taken into account through a parameter τ^2 that quantifies the variability of the offspring distribution of an individual (see supporting information, appendix 4). At large time scale, a large number of small mutations occur at each time unit, so that the phenotype trait evolution appears as a continuous and derivable process. The evolution speed $\dot{x}$ is defined as the derivative of the resident phenotype trait x with respect to time at this larger scale.

In our case, the canonical equation of adaptive dynamics can be written as (see the appendix 4 of supporting information for details)

$$\dot{x} = -\frac{\gamma^2(x)\theta(x)}{\sigma^2\tau^2(x)}(x - x^\star). \tag{6.10}$$

When $\gamma^2(x)$, $\theta(x)$ and $\tau^2(x)$ do not depend on x, the solution to this equation satisfies:

$$\frac{x(t) - x^\star}{x(0) - x^\star} = \exp\Big(-\frac{\gamma^2\theta}{\sigma^2\tau^2}\,t\Big). \tag{6.11}$$

Thus the time taken by the trait x to reach a given percentage of $x(0) - x^\star$, *i.e.* the distance between a given starting point ($x(0)$) and the singular strategy ($x^\star$), is equal to:

$$t = -\frac{\sigma^2\tau^2}{\gamma^2\theta}\ln\Big(\frac{x(t) - x^\star}{x(0) - x^\star}\Big). \tag{6.12}$$

In equation (6.12), σ^2 (the variance of the survival function), γ^2 (the variance of the mutation distribution) and θ (the mutation probability) do not depend on the landscape. On the contrary the variability of the offspring distribution of an individual, τ^2, depends on the spatial structure of the environment since it involves the eigenvectors of the A matrix describe

in equation (6.3). However, when all patches have the same carrying capacity and dispersal rates are symmetric, with basic assumptions on the demographic stochasticity, τ^2 does not depend of the landscape structure (see supporting information, appendix 4). In this case the evolutionary speed up to the singular strategy only depends of the distance up to $x^\star$ but not of the spatial repartition of habitats.

In the following two sections we applied these general results in two particular cases: section 6.3 consider a spatially implicit metapopulation with a hierarchical structure and section 6.4 consider a spatially explicit lattice with dispersal rates computed from an individual dispersal function.

6.3 Hierarchical metapopulations

6.3.1 Assumptions on the environment

Consider a hierarchical network composed by p_1 groups of p_2 patches with the same size, so that $P = p_1 p_2$ (figure 6.2). The questions of interest are how this group structure may change evolutionary outcomes and whether mixing habitats within groups may limit specialisation. Thanks to this framework, we will show how classical results could be found again and extended in order to study the role of spatial structures in the environment. In particular we will show that the branching criterion splits into a non-spatial term that depends on the fitness function and on the global proportion of habitats and a spatial term that reflects habitat allocation to groups.

Let the three dispersal rates m_0, m_1 and m_2 be defined in the following way: m_0 is the proportion of propagules issued from a patch that are deposited in a patch of another group, $m_0 + m_1$ is the proportion of propagules that are deposited in another patch of the same group, $m_0 + m_1 + m_2$ is the proportion of propagules that remain in its patch of origin. This parametrisation leads to the dispersal matrix

$$M = m_0 \, J_{P,P} + m_1 \, \mathrm{Id}_{p_1} \otimes J_{p_2,p_2} + m_2 \, \mathrm{Id}_P,$$

where $J_{P,P}$ is the $P \times P$ matrix of 1 and $\mathrm{Id}_{p_1} \otimes J_{p_2,p_2}$ denotes the block-diagonal matrix with diagonal matrices J_{p_2,p_2}. More generally, $\otimes$ denotes the tensor product between matrices.

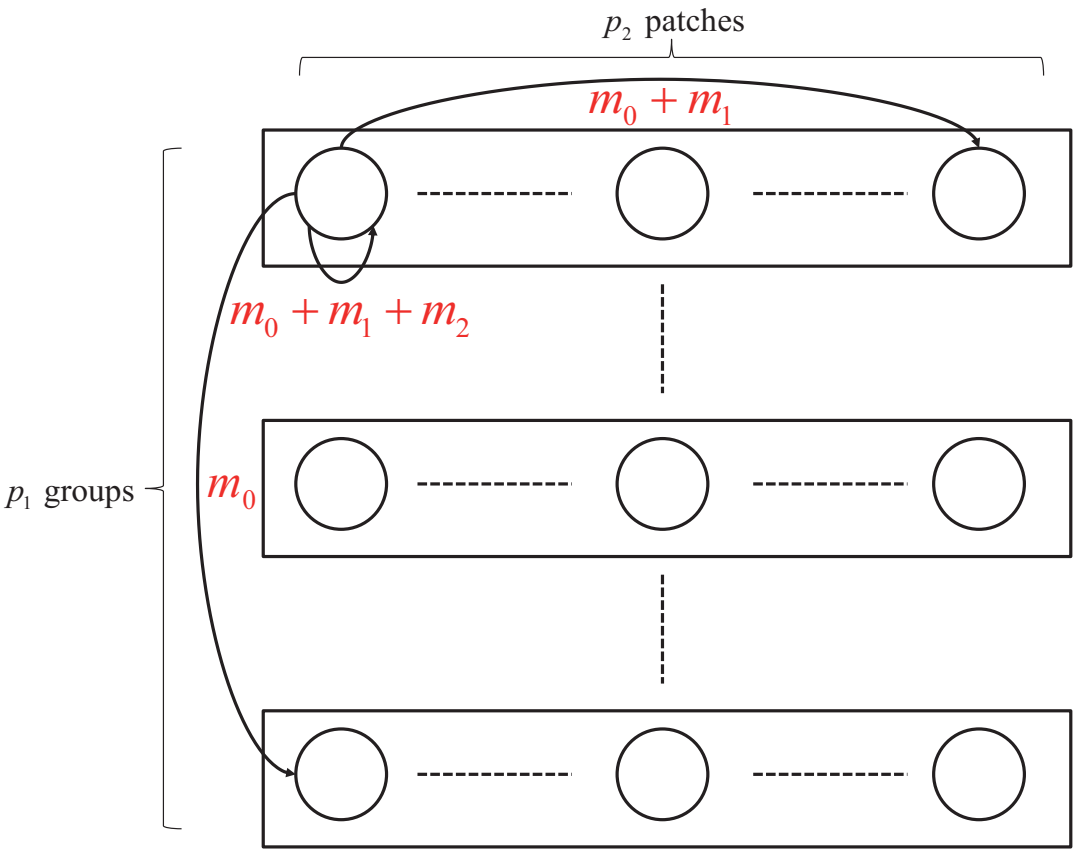


Figure 6.2: The hierarchical metapopulation structure. $P = p_1 \times p_2$ patches are dispatched among p_1 groups of p_2 patches. m_0: dispersal rate between patches that belongs to different groups, m_0+m_1: dispersal rate between patches that belongs to the same group, m_0+m_1+m_2: intra-patch dispersal rate.

6.3.2 Conditions for evolutionarily stable strategy

When dispersal is homogeneous ($m_0 \neq 0$ and $m_1 = m_2 = 0$), the condition for evolutionarily stability in equation (6.9) becomes:

$$\sum_{k=1}^{H} \pi_k \frac{(x^\star - \beta_k)^2}{\sigma^2} < 1 \tag{6.13}$$

(see the appendix 4 of supporting information for more details). This result for non-spatial environments has already been established by Geritz et al. (1998, Appendix 2). If there are two habitats in proportions π and $1-\pi$ with $\beta_1 = \delta$ and $\beta_2 = -\delta$, equation (6.13) becomes:

$$\pi(1-\pi) < \frac{\sigma^2}{4\delta^2}. \tag{6.14}$$

This simple relation shows that specialisation is facilitated when the trade-off is strong (δ/σ is large). Moreover, because of the homogeneous dispersal, only the composition (π) of the environment is involved. Equation (6.14) also shows that both habitats must exist in sufficient proportions for specialists to emerge. Otherwise, evolution leads to a single generalist poorly adapted to the habitat with the lowest proportion (equation (6.7)).

The first hierarchical level consists in distinguishing dispersal between patches from dispersal within patches ($m_0 \neq 0$, $m_2 \neq 0$ and $m_1 = 0$). In this case dispersal is hindered so that a propagule is more likely to stay in its patch of origin than to move to another patch ($m_2 \neq 0$) but there is no group structure ($m_1 = 0$), the condition for evolutionary stability in equation (6.9) becomes:

$$(1+2\xi)\sum_{k=1}^{H} \pi_k \frac{(x^\star - \beta_k)^2}{\sigma^2} < 1,$$

(see the appendix 4 of supporting information for more details), where $\xi = \dfrac{m_2}{1-m_2} = \dfrac{m_2}{p_1 p_2 m_0}$. The coefficient ξ provides a measure of the patch isolation. Adding heterogeneity in dispersal makes specialisation easier by multiplying the left-hand side of equation (6.13) by $(1+2\xi)$.

A second hierarchical level can be added by distinguishing dispersal between groups from dispersal within groups and within patches ($m_0 \neq 0$, $m_1 \neq 0$ and $m_2 \neq 0$). In this case, the condition for evolutionary stability in equation (6.9) becomes:

$$(1+2\xi)\sum_{k=1}^{H} \pi_k \frac{(x^\star - \beta_k)^2}{\sigma^2} + 2\nu \sum_{k=1}^{H}\sum_{k'=1}^{H} v_{kk'} \frac{(x^\star - \beta_k)(x^\star - \beta_{k'})}{\sigma^2} < 1, \tag{6.15}$$

(see the appendix 4 of supporting information for more details) where $\xi = \dfrac{m_2}{1-m_2} = \dfrac{m_2}{p_2 m_1 + p_1 p_2 m_0}$,

$\nu = (1+\xi)\dfrac{m_1}{p_2 m_0}$ and $v_{kk'} = \dfrac{1}{p_1}\displaystyle\sum_{g=1}^{p_1}(\pi_{kg} - \pi_k)(\pi_{k'g} - \pi_{k'})$, with π_{kg} the frequency of habitat k in group g. As before, the parameter ξ provides a measure of the patch isolation, while the quantity $\dfrac{m_1}{p_1 m_0}$ that appears in the definition of ν measures the group isolation. The presence of a group structure adds the second term of the left-hand side of equation (6.15) to the branching condition. This term is always positive, so that the presence of a group structure makes specialisation easier. This effect is greater for stronger isolation of groups and patches since ν is an increasing function of $\dfrac{m_1}{p_1 m_0}$ and of ξ.

The quantity $v_{kk'}$ in equation (6.15) can be interpreted as the covariance between the within-group frequencies of habitats k and k'. Thus, this term reflects the effect of the allocation of habitats to patches. When all the groups share the same habitat proportions, it is equal to zero. On the contrary when groups are not balanced with respect to habitat $|v_{kk'}|$ increases until each group contains one habitat only. In addition, the weight of the second term in equation (6.15) increases with $|v_{kk'}|$, *i.e.*, the effect of the group structure is greater when allocation of habitats to groups is more unbalanced. Finally, habitat allocation and proportion interact strongly since the range of variation of $|v_{kk'}|$ decreases when a habitat is increasingly present in the environment. This means that when the proportion of a particular habitat increases in the environment, the effect of the habitat allocation decreases.

6.3.2.1 Optimality of habitat and mixture strategies

Consider a situation when the within-group habitat frequencies π_{kg} can be controlled and when all the other variables are fixed, that is π_k, β_k, σ, m_0, m_1 and m_2 are fixed. Then the first term of the left-hand side of equation (6.15) is fixed and the left-hand side of equation (6.15) is minimal when its second term is zero. This occurs when $\pi_{k1} = \ldots = \pi_{kp_1} = \pi_k$ for $k = 1, \ldots, H$, since the covariances $v_{kk'}$ are then equal to zero. Thus allocating habitats to patches so that all the groups have the same habitat composition is optimal to limit specialisation among all the allocations, for fixed global frequencies π_k. We show here that mixture strategies are optimal for damping pathogen adaptation.

6.4 Lattice networks

In this section, we still study the conditions for branching of a monomorphic metapopulation for which equation (6.9) applies. However, we now consider a spatial rather than a hierarchical structure for the network in order to relax the strong symmetry properties that were imposed on the environment in section 6.3. Space is now explicit and dispersal rates are computed from an individual dispersal function. As a consequence, we will rely

on numerical computation to study the conditions for branching. Based on the numerical results, a global sensitivity analysis (Ginot et al., 2006; Saltelli et al., 2008) is performed to quantify the influence of the landscape pattern on the branching criterion defined in section 6.2.2.4.

6.4.1 Methods

6.4.1.1 Environment

We considered an arbitrary square area $\mathcal{E}$ partitioned into a regular lattice of size 30×30, resulting in $P = 900$ contiguous square patches. The landscape pattern was determined by the allocation of two habitats ($H = 2$) in proportion $1 - \pi$ and π to the 900 patches. Without loss of generality, we only considered patterns that satisfied $1 - \pi \geq \pi > 0$. To avoid border effects, $\mathcal{E}$ was considered as a torus in the calculations of dispersal rates and habitat aggregation (see below).

Two landscape pattern indices (LPI) were defined in order to characterise the different possible patterns: *(i)* the proportion of habitat 2 (π) and *(ii)* an aggregation index of habitat 2, AI (He et al., 2000), varying between 0, when two habitat 2 patches are never neighbours in the landscape, and 1, when habitat 2 is as patchy as possible (figure 6.3, first line).

6.4.1.2 Dispersal

The dispersal was assumed to be isotropic and to decrease exponentially with distance. More precisely, the proportion of propagules dispersed from a given source point z and arriving at a given reception point z' was given by the individual dispersal function

$$g(|| z - z' ||) = \frac{2\pi}{m_s^2} \exp\left(\frac{2\pi}{m_s} || z - z' ||\right),$$

where $|| z - z' ||$ is the distance between z and z', and m_s is a range parameter (Soubeyrand et al., 2009, Appendix C). The between-patch dispersal proportions were deduced by integration according to the formula

$$m_{jj'} = \int_{\mathcal{A}} \int_{\mathcal{A}'} g(|| z - z' ||)\, dz'\, dz. \tag{6.16}$$

When using equation (6.16), the integration is performed between pairs of points that belong to the area $\mathcal{A}$ and $\mathcal{A}'$ of patches j and j', respectively. The implicit assumption is that the population mixes perfectly in each patch. The dispersal rates $m_{jj'}$ in equation (6.16) were computed using the CaliFloPP algorithm (Bouvier et al., 2009) on the 900 patches. This algorithm computes the integral of a point-wise dispersal function between any pair of source and target polygons.

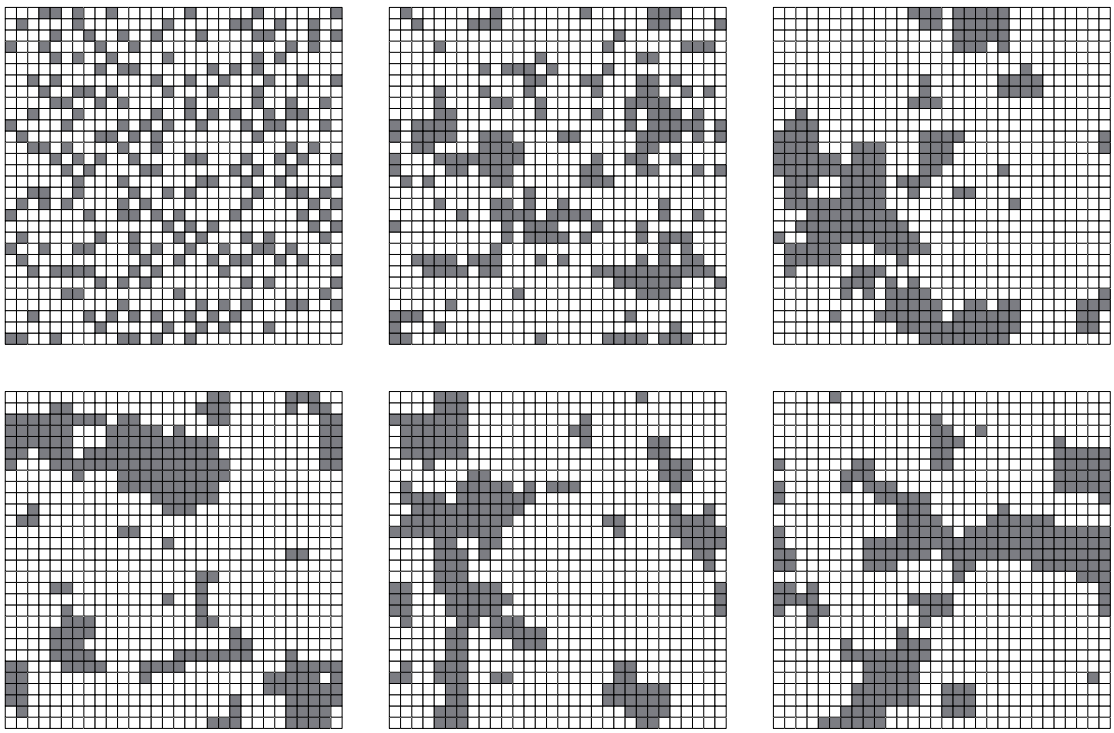


Figure 6.3: Repartition by simulated annealing of habitats in the lattice environment. Habitat 1 (white) and habitat 2 (grey) represent $1 - \pi = 75\%$ and $\pi = 25\%$, respectively. In the top row aggregation index (AI) is increasing: left panel, AI $= 0.1$; middle panel, AI $= 0.5$; right panel: AI $= 0.8$. The bottom row displays three pseudo-random allocations with AI $= 0.8$.

6.4.1.3 Experimental design

Three input factors were taken into account in the sensitivity analysis: *(i)* the LPI π; *(ii)* the LPI AI; *(iii)* the habitat differentiation δ/σ. In addition, the dispersal range was investigated by considering three different cases: m_s equal to 15%, 37.5% and 75% of the environment length.

The π index varied at ten levels in the range $(0; 0.5)$ and the AI index by 0.1 in the range $(0.1; 0.8)$. Landscape patterns were randomly generated for 75 combinations of π and AI with up to 30 replicates per combination (see below and figure 6.3). The remaining five combinations were impossible to generate by our algorithm. The habitat differentiation δ/σ was varied in the range $(0.4; 2)$.

A simulated annealing algorithm (Kirkpatrick et al., 1983) was used to generate landscapes with controlled values of the LPI π and AI according to the experimental design described before. Under these constraints, pseudo-random landscape patterns were generated by dispatching both habitats 1 and 2 among the 900 patches (figure 6.3, last line).

6.4.1.4 Sensitivity analysis

In order to quantify the influence of the input factors (π, AI and δ/σ), global sensitivity analysis was applied to the branching criterion defined in equation (6.9) and denoted by here Y. Calculation of the sensitivity indices was based on the Sobol' decomposition (Sobol, 1993) as described e.g. in Saltelli et al. (2000). According to this decomposition and assuming that the input factors vary independently, the variance of Y reads:

$$V = V_\pi + V_{\mathrm{AI}} + V_{\delta/\sigma} + V_{\pi,\mathrm{AI}} + V_{\pi,\delta/\sigma} + V_{\mathrm{AI},\delta/\sigma} + V_{\pi,\mathrm{AI},\delta/\sigma},$$

where V_π (respectively V_{AI} and $V_{\delta/\sigma}$) is the variance associated with the main effect of factor π (respectively AI and δ/σ), $V_{\pi,\mathrm{AI}}$ (respectively $V_{\pi,\delta/\sigma}$ and $V_{\mathrm{AI},\delta/\sigma}$) is the additional variance due to the interaction between factor π and AI (respectively between factor π and δ/σ and between factor AI and δ/σ), $V_{\pi,\mathrm{AI},\delta/\sigma}$ is the variance due to the interaction between the three factors, π, AI and δ/σ.

The sensitivity indices for factor v, $v \in \{\pi, \mathrm{AI}, \delta/\sigma\}$, are defined by:

- main effect indices: $\mathrm{SI}_v^{(1)} = \frac{V_v}{V}$,
- interaction indices: $\mathrm{SI}_v^{(2)} = \frac{\sum_{w \neq v} V_{vw}}{V}$,
- triple interaction index: $\mathrm{SI}^{(3)} = \frac{V_{\pi,\mathrm{AI},\delta/\sigma}}{V}$,
- total indices: $\mathrm{TSI}_v = \mathrm{SI}_v^{(1)} + \mathrm{SI}_v^{(2)} + \mathrm{SI}^{(3)}$.

They vary between 0 (no effect) and 1 (maximum effect).

In practise, the sensitivity indices cannot be calculated exactly and they must be estimated after running simulations. To estimate the indices, we applied the metamodelling technique proposed by Sudret (2008). First a polynomial-chaos metamodel of order five was fitted to the simulation results by linear regression. Then the estimated sensitivity indices were calculated using *ad hoc formulae* based on the estimated parameters of the metamodel.

6.4.2 Numerical experiment outputs

Figure 6.9 maps the stability of the singular strategy (*i.e.* the generalist strategy) depending on proportion (π) and aggregation (AI) of habitat 2 in several cases of habitat differentiation and dispersal range. Remind that $0 < \pi \leq 0.5$ since we only consider cases where $1 - \pi \geq \pi > 0$.

The variability of the branching criterion was much larger when $m_s = 15\%$ ($V = 495$) than when $m_s = 75\%$ ($V = 4.5$). However, in both cases, branching occurred for some combination of the factors. As expected, an increase in proportion and aggregation of habitat 2 made branching easier (figure 6.9). When the dispersal range (m_s) decreased, as when habitat differentiation (δ/σ) increased, branching occurred for lower habitat 2 proportions and levels of aggregation. When the trade-off was weak ($\delta/\sigma < 1$) and the dispersal range was large, the singular strategy was stable except for high levels of aggregation and large proportions of habitat 2. On the contrary, when the trade-off was strong ($\delta/\sigma > 1$) or the dispersal range was low, specialists were selected at high level of aggregation (respectively proportion) of habitat 2, whatever the proportion (respectively aggregation) of habitat 2 (figure 6.9).

The sensitivity analysis of the branching criterion (figure 6.5) identified habitat differentiation (δ/σ) as the most influential factor, especially when the dispersal range was large ($\mathrm{TSI}_{\frac{\delta}{\sigma}} = 0.62$ when $m_s = 75\%$ and $\mathrm{TSI}_{\frac{\delta}{\sigma}} = 0.37$ when $m_s = 15\%$). The influence of environment composition was also strong when the dispersal range was large, as quantified by the sensitivity indices of π ($\mathrm{TSI}_\pi = 0.29$ when $m_s = 75\%$ and $\mathrm{TSI}_\pi = 0.22$ when $m_s = 15\%$). On the contrary, spatial structures had the strongest effect when the dispersal range was small ($\mathrm{TSI}_{\mathrm{AI}} = 0.41$ when $m_s = 15\%$ and $\mathrm{TSI}_{\mathrm{AI}} = 0.09$ when $m_s = 75\%$).

An important part of the variance of the branching criterion was explained by interactions between factors, especially when the dispersal range was low (figure 6.5). When the dispersal range was large, interactions between the proportion of habitat 2 and differentiation among habitats were strong. When the dispersal range was low, interactions between the differentiation among habitats and the aggregation of habitat 2 were strong. The strength

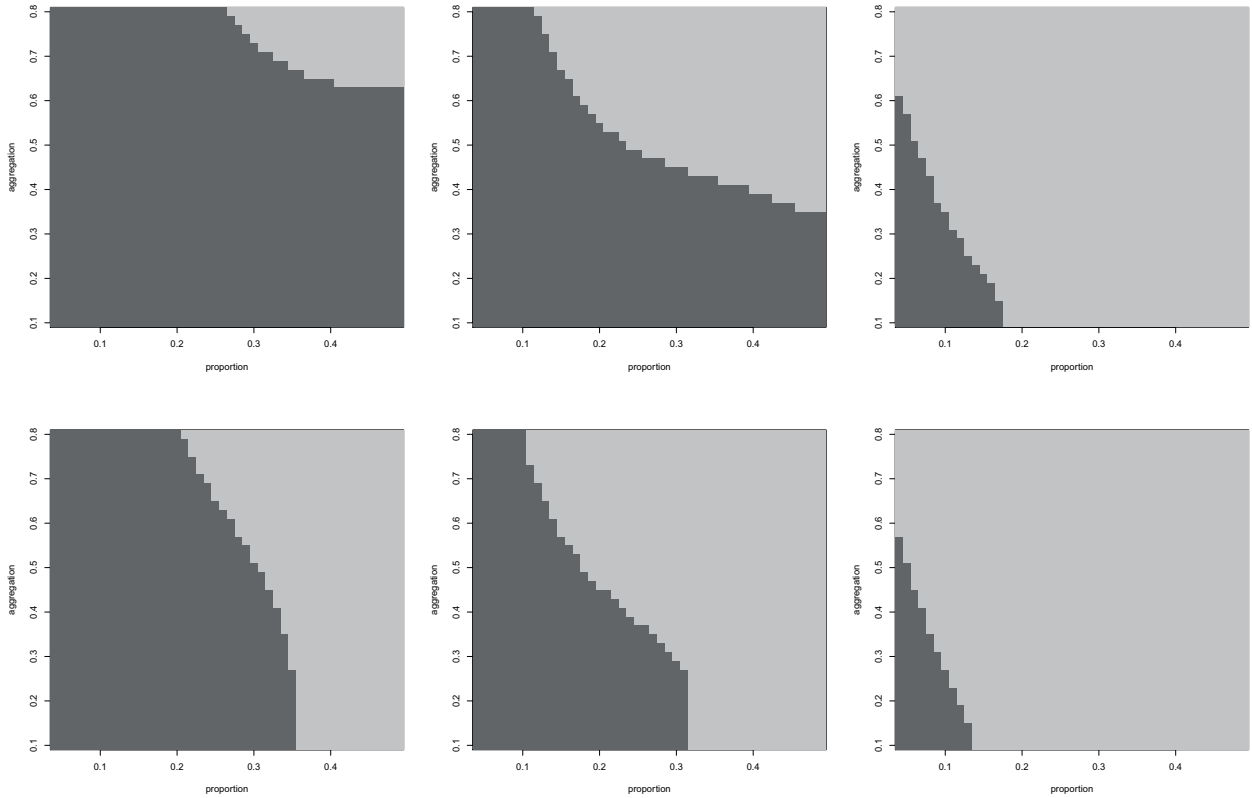


Figure 6.4: Stability of the generalist strategy against habitat 2 proportion and aggregation level. The generalist strategy is stable when the branching criterion (equation (6.9)) is lower than 1 (dark grey) and unstable when it is greater than 1 (light grey). In the top row the trade-off is weak ($\delta/\sigma = 0.84$), in the bottom row the trade-off is strong ($\delta/\sigma = 1.16$). From left column to right column dispersal range is decreasing: left column, $m_s = 75\%$; middle column, $m_s = 37.5\%$; right column, $m_s = 15\%$

of the third order interactions, when dispersal range was low, indicated complex interactions between the three factors.

6.5 What occurs after branching? A simulation study of specialist evolution

The main assumption in order obtain analytical results was the monomorphic population framework. As a consequence, it was not possible to study the evolution of specialists when the generalist strategy was not stable. In this section, we relaxed the monomorphic assumption and we studied the effects of spatial heterogeneity on the specialist strategies. In particular we investigated their degree of specialisation, the speed at which they evolved

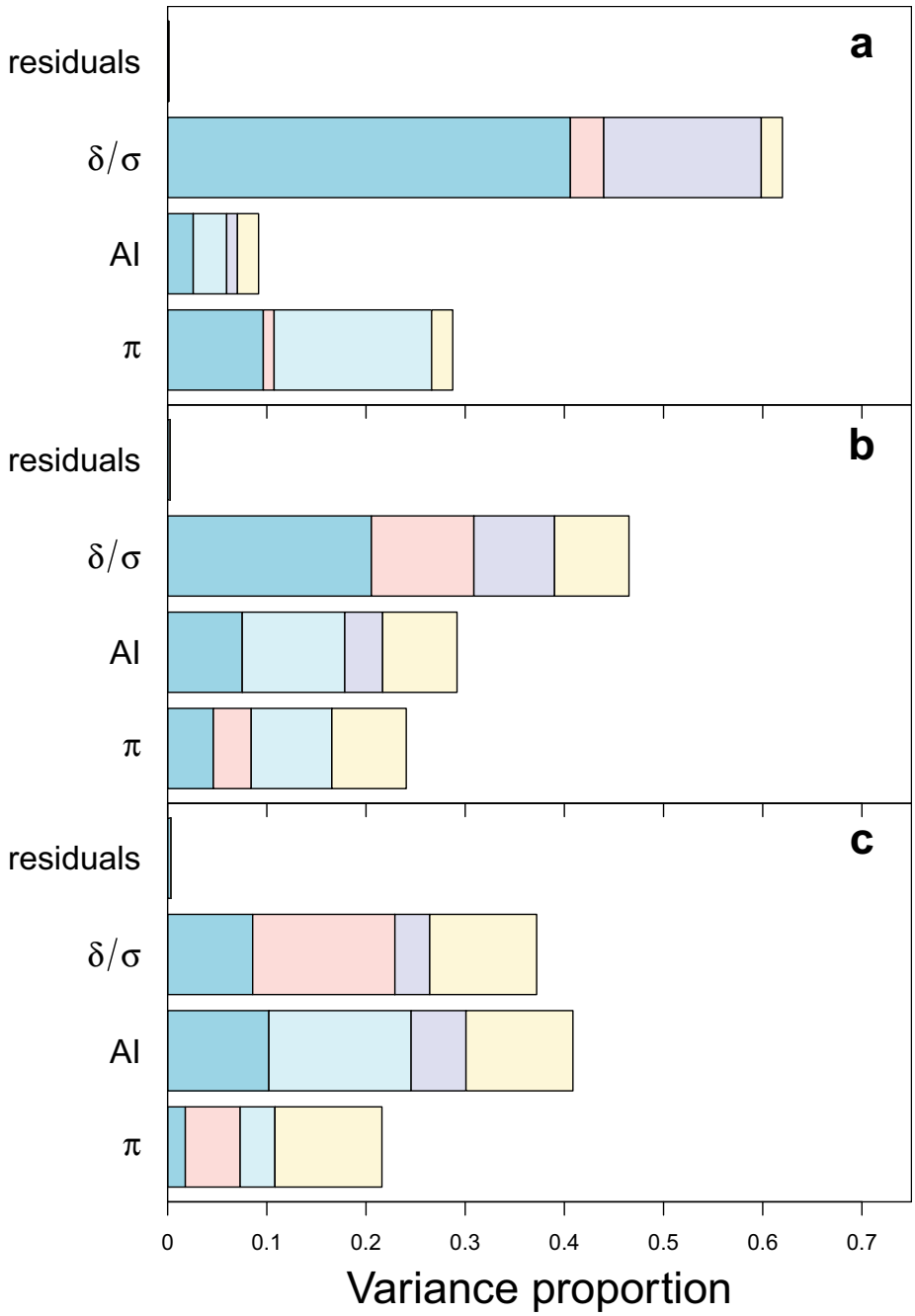


Figure 6.5: Sensitivity analysis performed on the branching criterion (equation (6.9)). The 3 factors analysed are: δ/σ, the habitat differentiation; AI, the habitat 2 aggregation level and π the habitat 2 proportion. From top to bottom dispersal range decrease: a, $m_s = 75\%$, b, $m_s = 37.5\%$ and c, $m_s = 15\%$. The first part of bars (blue) correspond to main indices (effect of the factor alone, $SI_v^{(1)}$) and full bars correspond to total indices (TSI_v). Pink, interaction with AI; purple, interaction with π; light blue, interaction with δ/σ; yellow, triple interaction.

and their spatial distribution.

We developed here a simulation model based on equation (6.2) representing the evolution of a metapopulation that developed on a lattice environment. The environment framework is that of section 6.4. We considered three case studies. The first allowed to check the agreement between theoretical predictions (section 6.2.2) and simulation results on the singular strategy and on the branching conditions. The second dealt with a large dispersal range and the third with low dispersal range. In these last cases, the other parameters (π and δ/σ) were fixed to values that favoured specialisation when habitat aggregation varied.

6.5.1 Methods

6.5.1.1 Case studies

In each case study there were two habitats. The dispersal range m_s, the habitat differentiation δ/σ and the proportion of habitat 2 (π) were fixed, while the aggregation index (AI) varied by 0.1 in the range $(0.1; 0.8)$.

In the first case study, the parameters were fixed at $m_s = 75\%$, $\delta/\sigma = 0.9$ and $\pi = 0.32$. Parameters were chosen in order to observe both stable and unstable conditions for the generalist when habitat aggregation varied.

In the last two case studies, the parameters m_s, δ/σ and π were chosen in order to observe a branching situation whatever the aggregation index. In the first one (situation A), the parameters were fixed at $m_s = 75\%$, $\delta/\sigma = 1.1$ and $\pi = 0.32$. In this case the large dispersal range limited the influence of the spatial structure and the branching criterion had a low variability when the aggregation index varied. In the second one (situation B), the parameters were fixed at $m_s = 15\%$, $\delta/\sigma = 1.1$ and $\pi = 0.16$. This set of parameters allowed to cover a broad range of branching criterion values when AI varied (figure 6.6).

For each case study and each AI value, 15 replicates were conducted each one on 15 different landscape patterns chosen at random as described in sections 6.4.1.1 and 6.4.1.3.

6.5.1.2 Model implementation

The assumptions on the environment and on the dispersal function were the same as in section 6.4. The patch carrying capacity was set to $K_j = 100$ for all patches. The continuous trait x varied between -5 and 5 and this range was discretised into 101 phenotypes. The simulations started with a monomorphic population with trait value $x = -5$. At each time step, mutants were generated from an existing phenotype provided its local population size was over a threshold equal to 1. New mutants were generated according to a Gaussian

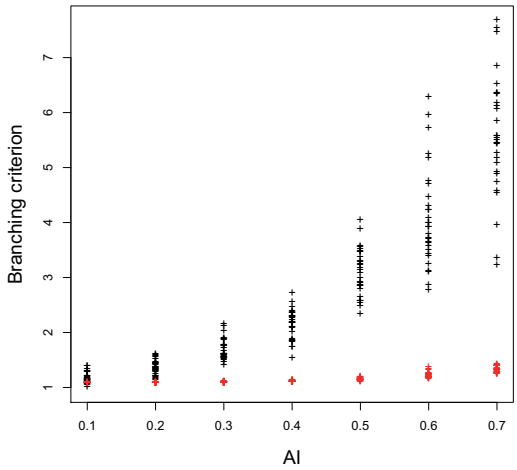


Figure 6.6: Branching criterion value against aggregation levels of habitat 2 (AI). Red, situation A ($m_s = 75\%$, $\delta/\sigma = 1.1$ and $\pi = 0.32$); black, situation B ($m_s = 15\%$, $\delta/\sigma = 1.1$ and $\pi = 0.16$).

perturbation centred on the pre-existing trait value and with variance γ^2 equal to 0.01. There was no genetic drift. The multitype metapopulation dynamics followed equation (6.2) and was simulated over 2000 time steps.

6.5.1.3 Post-simulation computations

Before branching, we estimated the singular strategy ($\hat{x}^\star$) and the time taken to reach it (T_{branch}) on the simulations. After branching, the metapopulation was considered as a mixture of the two populations of specialists. Simulations were then compared through three measures: *(i)* the trait value of specialists, *(ii)* the time taken to reach them (T_{ESS}) and *(iii)* the within population phenotypic variance. In addition, the level of adaptation of local populations was computed at the end of the evolution, in each patch.

More technically, at each time step the metapopulation composition was summarised by the means and standard deviation of a two-component Gaussian mixture that was fitted by the EM algorithm (Meng and Rubin, 1993) implemented in the function *normalmixEM* of the R package *mixtools* (Young et al., 2010, Version 0.4.4.). The time to reach the singular strategy (T_{branch}) was defined as the first time step where the EM algorithm identified two distinct populations. The estimated singular strategy ($\hat{x}^\star$) was then defined as the mean metapopulation phenotype at $t = T_{branch} - 1$. Trait values of specialists were estimated by the means of the two-component Gaussian mixture at the end of evolution. The time to reach the specialists (T_{ESS}) was defined as the first time step when the means of the

two-component Gaussian mixture were contained in the 5% interval around the trait value of the specialist strategies.

Finally, in patch j, the mean phenotype was calculated as a weighted average of the existing phenotypes, the weights being the proportions of each phenotype in the local population:

$$\bar{x}_j(t) = \frac{\sum_{i=1}^{I} x_i n_{ij}(t)}{K_j}.$$

Then we defined the level of local adaptation as the mismatch between the mean phenotype of patch j ($\bar{x}_j(t)$) and the optimal phenotype in patch j ($\beta_{h(j)}$): $|\bar{x}_j(t) - \beta_{h(j)}|$.

6.5.2 Simulation results

6.5.2.1 Comparing theoretical predictions and simulation results

Despite the monomorphic assumption, simulation results were highly consistent with analytical predictions. When the branching criterion was greater than 1, a branching point was observed in 90.2% of the simulations (figure 6.7). Reciprocally, no branching was observed when the branching criterion was lower than 1. Moreover, the estimated singular strategy $\hat{x}^\star$ was in line with the theoretical singular strategy $x^\star$: $\hat{x}^\star$ values were close to $x^\star$ and $\hat{x}^\star$ was independent of the aggregation of habitat 2 (figure 6.7).

Figure 6.8 shows the time to reach the singular strategy. According to section 6.2.2.5, the time to reach the singular strategy could potentially dependent on the aggregation index. This was not the case for the situation A. However, in the situation B with stronger spatial effects (due to a lower dispersal range), the singular strategy was reached faster when habitats were grouped. This tendency was significant but differences in the median of T_{branch} between isolated habitats and grouped habitats varied between 39 and 49 only (figure 6.8).

6.5.2.2 Effect of habitat aggregation on specialist strategies

Habitat aggregation level highly affected the specialist populations. First, figure 6.9 shows that specialists were more adapted to the habitats when the habitats were more grouped (AI is higher). Second, the within-population variability was smaller when habitats were aggregated (figure 6.10). This means that, when habitats were more grouped, the specialist populations had a mean phenotype more adapted and were more homogeneous. Third, when habitat were more grouped specialist populations evolved more rapidly to their stable strategies (figure 6.11). Thus, specialists evolved faster and became more adapted when habitats were more grouped.

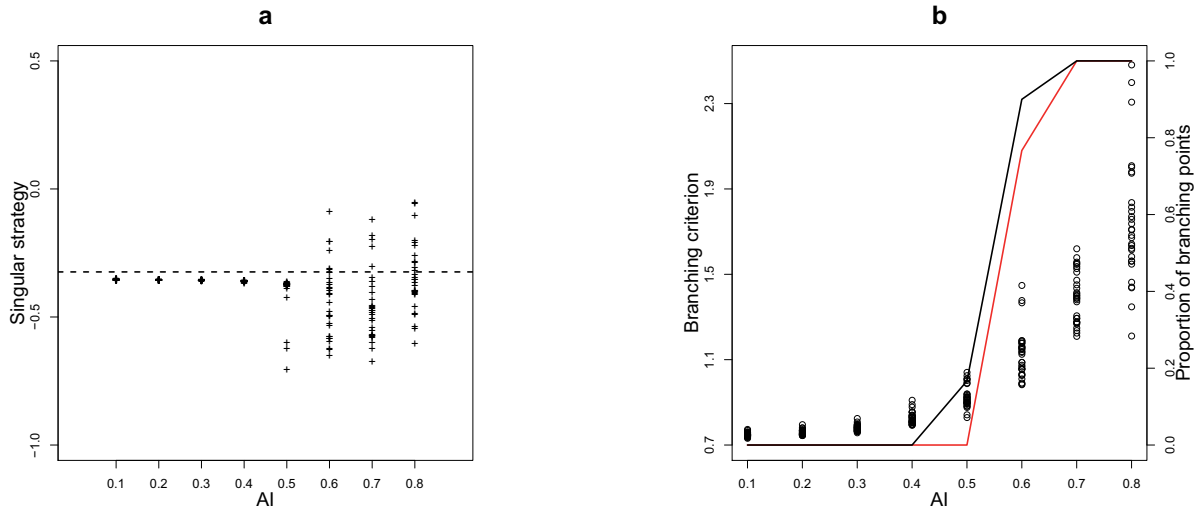


Figure 6.7: Comparison between analytical and simulation results. a, the estimated singular strategy $\hat{x^\star}$ (+) plotted along with the theoretical singular strategy $x^\star$ (dashed line) against the aggregation level of habitat 2 (AI). b, the branching criterion value (open circle and left Y-axis) plotted along with the predicted (black solid line and right Y-axis) and observed on simulations (red solid line and right Y-axis) proportions of branching points, against the aggregation level of habitat 2. Other parameters are: $m_s = 75\%$, $\delta/\sigma = 0.9$ and $\pi = 0.32$.

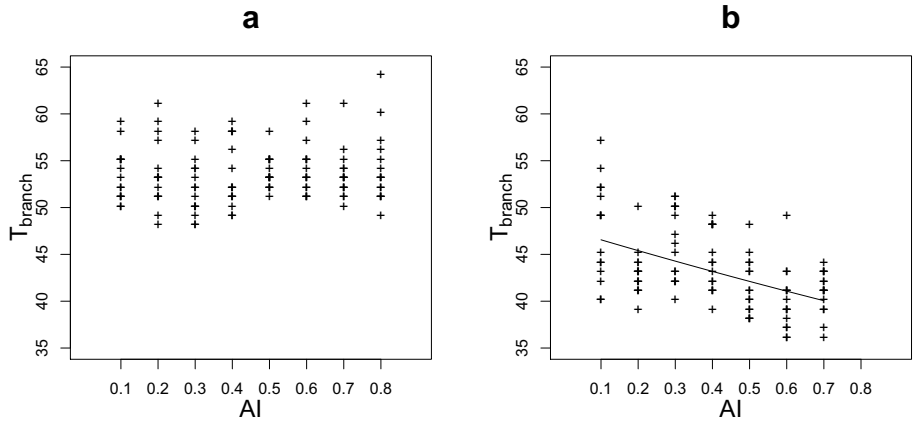


Figure 6.8: The time to reach the singular strategy (T_{branch}) plotted against the aggregation level of habitat 2 (AI) in the case of situation A (a) and of situation B (b). The relationship between T_{branch} and AI was tested by a GLM with Poisson distributed errors. The solid line indicates that the slope was significantly greater than 0 with a 0.001 threshold. Other parameters are: a, $m_s = 75\%$, $\delta/\sigma = 1.1$ and $\pi = 0.32$; b, $m_s = 15\%$, $\delta/\sigma = 1.1$ and $\pi = 0.16$.

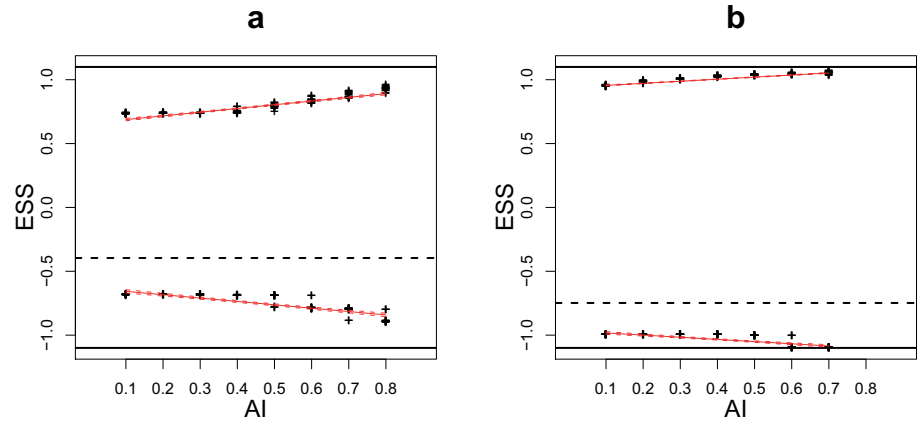


Figure 6.9: Specialists trait value at the equilibrium (ESS) plotted against the aggregation level of habitat 2 (AI) in the case of situation A (a) and of situation B (b). The relationship between ESS values and AI was tested by classical linear regression. Red lines indicates that the slope was significantly greater than 0 with a 0.001 threshold. Other parameters are: a, $m_s = 75\%$, $\delta/\sigma = 1.1$ and $\pi = 0.32$; b, $m_s = 15\%$, $\delta/\sigma = 1.1$ and $\pi = 0.16$.

Specialist populations obtained in situation A appeared globally less adapted, more variable and with a slower evolutionary speed than those obtained in situation B. This was due to the fact that the branching criterion was always higher in situation B with respect to situation A (figure 6.6). However, while the range of variation of the branching criterion was greater in situation B, effects of habitat aggregation level were larger in the situation A. This means that the value of the branching criterion was crucial especially when it was near the branching threshold of 1.

The monitoring of the within population phenotypic variance revealed a first diversification phase during which the phenotypic variance increased then a selection phase during which the phenotypic variance decreased until it was stable (figure 6.10). The maximum phenotypic variance coincided with the branching point.

Figure 6.12 shows the pattern of adaptation across the environment in a particular example. Populations living at the centre of habitat aggregates showed a high local adaptation level whereas populations inhabiting isolated patches or edge patches were poorly adapted. These differences in local adaptation were not explained by the persistence of a generalist in the edges but rather by the coexistence of the two specialists in the same patch (not shown) due to the migration-selection balance (Mouquet and Loreau, 2003). This was observed in all the simulations.

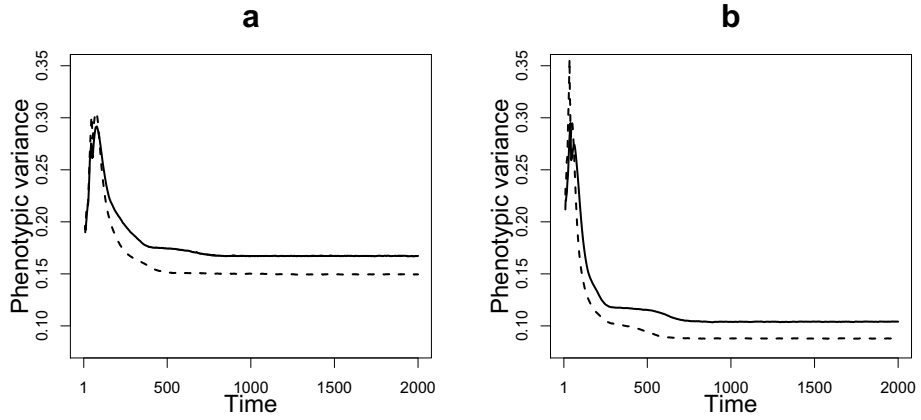


Figure 6.10: Evolution of the within populations phenotypic variance for AI = 0.1 (solid line) and for AI = 0.7 (dashed line). Other parameters are: situation A (a), $m_s = 75\%$, $\delta/\sigma = 1.1$ and $\pi = 0.32$; situation B (b), $m_s = 15\%$, $\delta/\sigma = 1.1$ and $\pi = 0.16$.

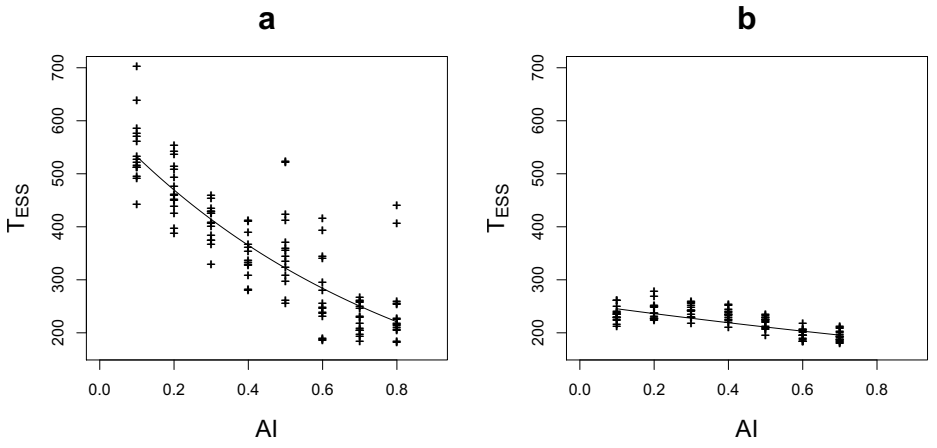


Figure 6.11: The time to reach the ESSs (T_{ESS}) plotted against the aggregation level of habitat 2 (AI) in the case of situation A (a) and of situation B (b). The relationship between T_{ESS} and AI was tested by a GLM with Poisson distributed errors. The solid line indicates that the slope was significantly greater than 0 with a 0.001 threshold. Other parameters are: a, $m_s = 75\%$, $\delta/\sigma = 1.1$ and $\pi = 0.32$; b, $m_s = 15\%$, $\delta/\sigma = 1.1$ and $\pi = 0.16$.

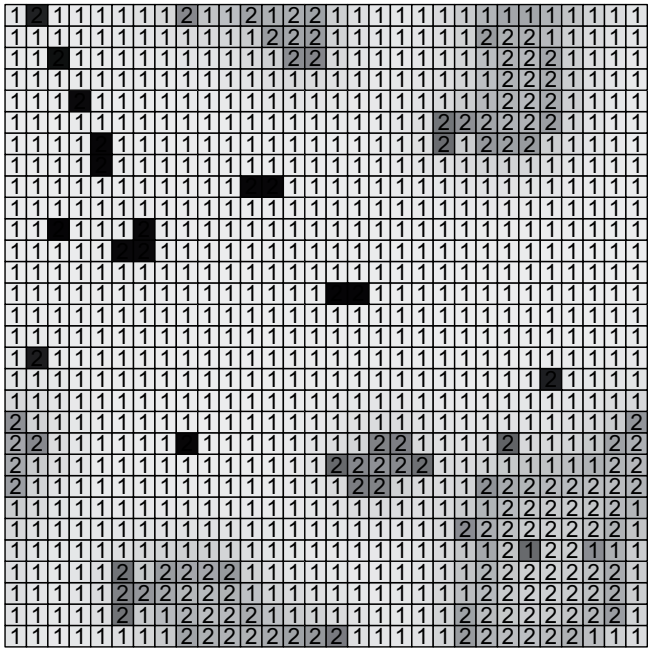


Figure 6.12: An example of the level of local adaptation: the more adapted the local population, the lightest the grey. Patches marked with 1 (respectively 2) shared the habitat 1 (respectively the habitat 2). In this case, specialisation was predicted by the analytical branching criterion. Parameters are $m_s = 15\%$, $d = 1.1$, $\pi = 0.16$ and AI $= 0.7$ (situation B). This is an example of a simulation.

6.6 Discussion

With regard to agricultural consequences of pathogen specialisation, we identified two questions related to applied evolutionary biology: how do pathogens adapt to a heterogeneous environment? And how does landscape structure determine host specificity of pathogens? In this article, we addressed these questions from a theoretical point of view. We studied the consequences of spatial heterogeneity on the dynamics of adaptation of a population that lives on a finite network of habitat patches interconnected via passive dispersal. By an adaptive dynamics approach, we developed a model that described the phenotypic changes occurring in a metapopulation under soft selection. Our model provided a more general framework than most models in the literature since no assumption was needed on dispersal rates: classical as well as spatially explicit metapopulation models can be handled. By analysing the model through analytical as well as simulation methods, we were able to generalise classical results: whatever dispersal rates, (i) we defined the singular strategy, (ii) we characterised its stability and (iii) we provided the evolutionary speed. We also (iv) studied the effects of spatial heterogeneity on the specialist strategies.

The singular strategy corresponds to a generalist phenotype since it represents a balanced strategy with respect to the habitat frequencies and connectivity. When dispersal rates are symmetric and patches have the same input connection, the singular strategy is a function of habitat proportion and optima only: it is independent of the spatial distribution of habitats in the environment. This classical result can be also found in Geritz et al. (1998) or Débarre and Gandon (2010). However, it represents only a particular case since, in equation (6.6), spatial configuration can influence the singular strategy: small and/or isolated patches have a smaller weights on the generalist phenotype. Although this had been pointed out in the case of the dynamics and persistence of a metapopulation by Ovaskainen and Kanski (2003) and discussed by Hanski et al. (2011) with evolutionary perspectives, the role of individual habitat patches was not yet demonstrated with an adaptive dynamics approach.

The stability of the singular strategy determines if the population remains monomorphic with a generalist phenotype or if it splits into two specialised populations. It is characterised by the value of the branching criterion. The first models that investigated the respective roles of environment heterogeneity and dispersal rate on the stability of the singular strategy involved two patches with different habitats and linked by migration (*e.g.* Meszéna et al., 1997). In this case, limiting migration between patches helped specialisation. In spatial environments with given habitat proportions, migration between both habitats can be reduced by hindering dispersal or by aggregating patches that share the same habitat. As expected, the model developed here predicts that both effects are crucial in determining if the metapopulation remains monomorphic or if it splits into two specialists. As it was underlined by Débarre and Gandon (2010), spatial and two-patch models give qualitatively the same results. However, we were able here to identify in several dispersal contexts the respective roles of habitat differentiation, habitat frequencies and spatial distribution. Considering a metapopulation that develops in a hierarchical network of habitat patches, we showed that the branching criterion splits into a non-spatial term that depends on the fitness function and on the global proportion of habitats and a spatial term that reflects habitat allocation to groups. We also provided a sensitivity analysis on the branching criterion of a spatially explicit metapopulation inhabiting a lattice network. In this case, we showed that for short dispersal range the spatial distribution of habitats is the most influential factor on branching. However, its effects decrease when dispersal range increases and for long dispersal ranges, habitat proportions and habitat differentiation are the most influential factors.

The evolutionary speed of a monomorphic population is approximately given by the canonical equation of adaptive dynamics. In our case, provided some assumptions on both demographic stochasticity and dispersal rates, we found that the time taken by the population to reach its generalist phenotype is independent of the spatial organisation of habitats. However, using a simulation approach, we found that this is not the case when dispersal range was short, due to the stronger effects of the habitat spatial structure. More

efforts must be done in order to characterise the role of demographic stochasticity and habitat spatial structure in the evolutionary speed of a monomorphic population.

What happens after branching is scarcely treated in the literature. Meszéna et al. (1997) demonstrated that adaptation of coexisting phenotypes to habitats is stronger for higher patch differences or inverse migration rates. Geritz et al. (1998) obtained similar results with three habitats and random migration. Thanks to a simulation approach, we have showed that the spatial arrangement of habitats impacts both the mean phenotype of specialist populations and their phenotypic variance: coexisting phenotypes were found more adapted in grouped habitats. In addition, a clear effect of habitat aggregation on specialisation speed was also observed in all the simulations: specialists evolved faster when habitats are more grouped. As a consequence, specialists adapt faster and better when habitats are grouped.

Coexistence among specialist and generalist genotypes in heterogeneous landscapes was investigated in chapter 5 using a demographic approach. We used the term 'demographic' because life-cycle was described explicitly and because genotypes were pre-defined. Both approaches lead to comparable conclusions: grouped allocation strategies as well as short dispersal range favour specialist genotypes whereas mixed allocation strategies and large dispersal range favour generalist genotypes. However, one important conclusion differs between the two approaches: according to chapter 5 three genotypes can coexist in a two-habitat landscape. This was never the case in the present study. We underline here an important difference between gradual evolution of a population and competition between pre-existing genotypes. In the first case, the population evolves first towards a generalist genotype, which can be stable (the population remains monomorphic) or not. It then splits into two coexisting specialists. When genotypes pre-exist and cannot evolve through mutation, we only observe the output of a competition. How much is it stable to mutations? This remains an open question.

The adaptive dynamics framework could appear as a little far from the applied question of the effect of agricultural landscape spatial heterogeneity on adaptation and specialisation of pathogens. However, some of our results can be of practical interest for designing sustainable strategies for disease control in agricultural landscapes. Our results suggest that highly homogeneous landscapes select for highly specialised pathogens but also that these pathogens are selected more rapidly. Cultivar mixtures, at the field or the landscape scale, appear then as unavoidable in order to increase varietal resistance durability for three reasons: (i) they hamper pathogen specialisation and, in the case when specialists are unavoidable, (ii) they slow down specialist evolution and (iii) limit their adaptation level. The particular case of the hierarchical metapopulation can be interesting from the agricultural point of view. The environment may be a field where a mixture of varieties is grown and where subplots (patches) receive different varieties. Alternatively, the environment may be an agrosystem

where patches are fields and groups are local production basins. We show here that mixture strategies are optimal for damping pathogen adaptation. More generally, landscapes that increase the durability of resistant varieties could be designed by defining a criterion and optimising this criterion. The branching criterion we used was derived from an invasion study and characterises the stability of the singular strategy. Our simulations showed, however, that this criterion seems to be a good predictor of the global evolutionary trajectory: the higher the branching criterion value, the more favourable the landscape for specialisation. It would be interesting to design and test variety allocation strategies in more realistic agricultural landscapes using this criterion along with optimisation methods.

Adaptive dynamics allows studying the long-term evolution of a population. This framework can be used to model the evolution of pathogen with short life cycles such as bacteri or fungi over one or several years. In our study, we assumed that the landscape structure did not change over time. However, crop rotations in time impact the evolution of pathogen populations (van den Berg et al., 2010) and could be a good control strategy for disease (Xu, 2011). An improvement of our study would be to study how this temporal heterogeneity modifies the results (see the appendix 5 of supporting information).

Part IV

Discussion

Chapter 7

General discussion

Contents

7.1 Introduction

The transition towards an ecologically intensive agriculture needs to mobilise an association of effective agronomic levers. Among these, crop genetic diversity is a promising way. More than any finite resource, variety resistances are subject to 'the tragedy of the commons' (Hardin, 1968) and thus require the design of collective strategies for their durable management. This PhD work aims at shifting from local to landscape scales in order to improve our understanding and prediction of disease risk and to provide an eco-evolutionary support for new organisational strategies.

Three ecological mechanisms that link host genetic diversity and susceptibility of host populations to disease were identified: dilution effect, competition among pathogen strains and long term evolution of pathogen populations. These mechanisms were investigated in agricultural landscapes using both statistical tools for the analysis of real data and theoretical approaches, including simulation models and mathematical analysis. The aim of this section is not to discuss particular results again but rather to give insight into the complementarity of approaches and to provide some ideas to go further.

7.2 Draw me a resistant landscape

The relationship between host diversity and disease susceptibility raises three questions: how does a pathogen population spread over a heterogeneous host landscape? How do pathogen genotypes compete in a diversified host population? And, at a longer term, how do pathogen populations evolve in response to host population structure? In a consistent way, all studies in the previous chapters showed that the composition and the spatial structure of the host population influence greatly the pathogen population. However, the recommendations that this work could provide will depend on the desired aim.

Suppose that a new variety with a new quantitative resistance is now available. How should it be introduced in the varietal landscape? The first objective could be to protect this variety against disease. At short time scale the better strategy would be to cultivate this variety over large and non fragmented areas. In this way we minimise the exchange of pathogen propagules among varieties and preserve the integrity of the new one. However, at longer time scales, this is the best way to favour the emergence of highly specialised and damageable pathogen strains. In such a context we could expect a brutal breakdown of the variety resistance. The second objective could be to protect the other varieties with the new one. In this case mixed strategies are the more adapted ones because they decrease landscape connectivity. This strategy cannot be the most productive one at short time scale because the new resistant variety receives more pathogen propagules and will exhibit more

disease. Nevertheless, in such a context, pathogen populations will evolve more slowly and towards less aggressive strains: to protect old varieties in the short term implies to protect the new one in a longer term.

Consider now the varieties that are available in a given year. The question is how to deploy these varieties in order to have a resistant landscape. The leaf rust study showed that pathogen populations are composed by a collection of strains with a variable specialisation degree, from highly specialised strains (that could be highly damageable) to generalist ones. A dominant variety, both at the global scale or at the local scale, will enhance its own specialist, leading to a decrease in its resistance level and greater disease severity. On the contrary, mixed strategies will favour generalist strains and will intensify competition among specialists. As a consequence, non frequent varieties, both globally and locally, will be less exposed to specialist pathogens and this will preserve their integrity. In purely spatial terms, variety diversity must then be enhanced both at local and global scale in order to avoid the dominance of a particular variety and to promote competition among pathogen strains. The temporal scale must not be forgotten. Temporal succession of varieties could lead to the same results. Varying spatial and temporal composition of the varietal landscape offers new dimensions: highly popular and susceptible varieties could become 'resistant' when they decrease in frequency, as in the case of variety *Soissons* (see chapter 2).

7.3 Modelling approaches in agro-ecological studies

Modelling approaches in agro-ecological studies may aim at investigating mechanisms from a basic point of view or at providing predictive models for managers. This PhD thesis dealt with the first objective. We based our work on three complementary modelling approaches: statistics, simulations and mathematics. One vision of the relationship between each of them could be a continuum in the description of observations. Mathematics analysis requires drastic hypotheses in order to put the problem in a tractable form. Simulation-based methods allow to relax some hypotheses but still consider an idealised world. Finally, statistical analyses are directly confronted to real data-sets. Exchanges between these approaches are useful in both ways. As an example, simulation approaches can give information on the robustness of some necessary hypothesis. Moreover, with the increase of likelihood-free inference methods, the border between simulation and statistical models is less and less clear-cut.

Another important point that I would like to discuss is the role of data. Any modelling approach must be based, more or less, on observed phenomena. All types of data should be considered: from precise experimental design in highly controlled situations to large temporal and spatial scale datasets that could come from diversified sources. Even if they

are not obtained in controlled conditions, many large scale data-sets allow to confront ecological theories to the real world. There are a lot of 'dormant' data-sets, such as in naturalistic associations, agricultural technical institutes, state agencies and others that would be interesting to see better exploited.

7.4 Perspectives

An important point in the approach of this PhD work is that we studied quantitative traits for pathogenicity rather than qualitative characters. However, we only considered discrete phenotypes and ignored variance (but see the simulation study of chapter 6). A further step would require to focus on the population variability in pathogenicity both in theoretical models and in statistical analyses, where the split into homogeneous pathotypes is sometimes contestable. In this vein, we always assumed clonal reproduction: adding sexual recombination will give more insights to our findings.

Another point that will be exciting is to widen our work to natural systems and to investigate further how spatial structure of the host and pathogen populations interact with co-evolving patterns between host resistance and pathogen infectivity. In the agricultural part this echoes with variety rotations. Here we essentially worked on spatial heterogeneity (but see chapter 2 and the appendix 5 of supporting information) but adding a temporal dimension would be of prime interest in order to design management strategies of variety resistances.

Finally, some agricultural cooperatives and technical institutes are working on pesticide reduction by reasoning the varietal choice. I think that this represents a great opportunity to acquire data-sets at the regional scale in order to confront models such as presented in this work to the real world.

7.5 Conclusion

I would like to conclude this work by underlying that collective changes must be operated in our agricultural systems. In fact, the opposition between natural and agricultural environments is not tenable anymore: we must reintegrate our agricultural systems into the natural environment. Such a change requires a collective management of agricultural spaces. As a consequence, the practical design of new management strategies must be done together with social and economical perspectives. Analysis and modelling of farmer's decisions with respect to land uses is an important field of research. These studies aim at assessing the impact of public policy on agricultural activity, or at determining the optimal location of

cropping systems based on both economic and environmental criteria. However, coupling agronomic models with decision models has been so far underexploited and is usually limited to the farm scale. The scale of interest must now be widened and landscape epidemiology approaches must integrate the operators and their potentially different strategies.

Bibliography

K. Adamczyk, F. Angevin, N. Colbach, C. Lavigne, F. Le Ber, and J.F. Mari. Genexp, un logiciel simulateur de paysages agricoles pour l'étude de la diffusion de transgènes. *Géomatique*, 17:469–487, 2007.

P.B. Adler, J. HilleRisLambers, and J.M. Levine. A niche for neutrality. *Ecology Letters*, 10:95–104, 2007.

O. Ahlqvist and A. Shortridge. Spatial and semantic dimensions of landscape heterogeneity. *Landscape Ecology*, 25:573–590, 2010.

S. Alizon and M. van Baalen. Emergence of a convex trade-off between transmission and virulence. *The American Naturalist*, 165:E155–E167, 2005.

B. Allan, F. Keesing, and R. Ostfeld. Effect of forest fragmentation on lyme disease risk. *Conservation Biology*, 17:267–272, 2003.

M.A. Altieri. The ecological role of biodiversity in agroecosystems. *Agriculture Ecosystems and Environment*, 74:19–31, 1999.

F. Angevin, E.K. Klein, C. Choimet, A. Gauffreteau, C. Lavigne, A. Messean, and J.M. Meynard. Modelling impacts of cropping systems and climate on maize cross-pollination in agricultural landscapes: the MAPOD model. *European Journal of Agronomy*, 28:471–484, 2008.

T. Arak and D. Surgailis. Markov fields with polygonal realizations. *Probability Theory and Related Fields*, 80:543–579, 1989.

T. Arak, P. Clifford, and D. Surgailis. Point-based polygonal models for random graphs. *Advances in Applied Probability*, 25:348–372, 1993.

B. Barrès, F. Halkett, C. Dutech, A. Andrieux, J. Pinon, and P. Frey. Genetic structure of the poplar rust fungus *Melampsora larici-populina*: evidence for isolation by distance in europe and recent founder effects overseas. *Infection, Genetics and Evolution*, 8:577–587, 2008.

R.A. Bayles, K. Flath, M.S. Hovmø ller, and C. de Vallavieille-Pope. Breakdown of the Yr17 resistance to yellow rust of wheat in northern europe. *Agronomie*, 20:805–811, 2000.

M.A. Beaumont. Approximate bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution and Systematic*, 41:379–406, 2010.

G. Bell. Neutral macroecology. *Science*, 293:24–132418, 2001.

F.J.J.A. Bianchi, C.J.H. Booij, and T. Tscharntke. Sustainable pest regulation in agricultural landscapes: a review on landscape composition, biodiversity and natural pest control. *Proceedings of the Royal Society B-Biological Sciences*, 273:1715–1727, 2006.

M.D. Bolton, J.A. Kolmer, and D.F. Garvin. Wheat leaf rust caused by *Puccinia triticina*. *Molecular Plant Pathology*, 9:563–575, 2008.

M. Boots and A. Sasaki. The evolutionary dynamics of local infection and global reproduction in host-parasite interactions. *Ecology Letters*, 3:181–185, 2000.

A. Bouvier, K. Kieu, K. Adamczyk, and H. Monod. Computation of the integrated flow of particles between polygons. *Environmental Modelling and Software*, 24:843–849, 2009.

J.A. Browning and K.J. Frey. Multiline cultivars as a means of disease control. *Annual Review of Phytopathology*, 14:355–382, 1969.

H. Brun, A.M. Chèvre, B.D.L. Fitt, S. Powers, A.L. Besnard, M. Ermel, V. Huteau, B. Marquer, F. Eber, M. Renard, and D. Andrivon. Quantitative resistance increases the durability of qualitative resistance to *Leptosphaeria maculans* in *Brassica napus*. *New Phytologist*, 185:285–299, 2010.

J.J. Burdon and G.A. Chilvers. Host density as a factor in plant-disease ecology. *Annual Review of Phytopathology*, 20:143–166, 1982.

J.J. Burdon and P.H. Thrall. Pathogen evolution across the agro-ecological interface: implications for disease management. *Evolutionary Applications*, 1:57–65, 2008.

F. Burel and J. Baudry. *Landscape Ecology: Concepts, Methods and Applications*. Science Publishers, Inc., Plymouth (UK), 2004.

O.J. Burton and J.M.J. Travis. Landscape structure and boundary effects determine the fate of mutations occurring during range expansions. *Heredity*, 101:329–340, 2008.

J.P. Butault, C.A. Dedryver, C. Gary, L. Guichard, F. Jacquet, J.M. Meynard, P. Nicot, M. Pitrat, R. Reau, B. Sauphanor, I. Savini, and T. Volay. *Ecophyto R&D. Quelles voies pour réduire l'usage des pesticides? Synthèse du rapport d'étude*. INRA Editeur, Paris (Fr), 2010.

M.S. Castellazzi, J. Matthews, F. Angevin, C. Sausse, G.A. Wood, P.J. Burgess, Brown I., K.F. Conrad, and J.N. Perry. Simulation scenarios of spatio-temporal arrangement of crops at the landscape scale. *Environmental Modelling and Software*, 25:1881–1889, 2010.

H. Caswell. *Matrix Population Models*. Sinauer Associates Inc. Publishers, Sunderland (USA), 2001.

N. Champagnat and S. Méléard. Invasion and adaptive evolution for individual-based spatially structured populations. *Journal of Mathematical Biology*, 55:147–188, 2007.

N. Champagnat, R. Ferrière, and S. Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical Population Biology*, 69:297–321, 2006.

J. Chave. Neutral theory and community ecology. *Ecology Letters*, 7:241–253, 2004.

P. Chesson. MacArthur's consumer-resource model. *Theoretical Population Biology*, 37: 26–38, 1990.

P. Chesson. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematic*, 31:343–366, 2000.

K.M. Chin and M.S. Wolfe. The spread of *Erysiphe graminis f.sp. hordei* in mixtures of barley varieties. *Plant Pathology*, 33:89–100, 1984.

J.S. Clark. Beyond neutral science. *Trends in Ecology and Evolution*, 24:8–15, 2009.

J.S. Clark, M. Dietze, S. Chakraborty, P.K. Agarwal, I. Ibanez, S. LaDeau, and M. Wolosin. Resolving the biodiversity paradox. *Ecology Letters*, 10:647662, 2007.

J.S. Clark, D. Bell, C. Chu, B. Courbaud, M. Dietse, M. Hersh, J. HilleRisLambers, I. Ibanez, S. LaDeau, S. McMahon, J. Metcalf, J. Mohan, E. Moran, L. Pangle, S. Pearson, C. Salk, Z. Shen, D. Valle, and P. Wyckoff. High-dimensional coexistence based on individual variation: a synthesis of evidence. *Ecological Monographs*, 80:569–608, 2010.

M. Clavero and L. Brotons. Functional homogenization of bird communities along habitat gradients: accounting for niche multidimensionality. *Global Ecology and Biogeography*, 19: 684–696, 2010.

N. Colbach, C. Clermont-Dauphin, and J.M. Meynard. Genesys: a model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers - i. temporal evolution of a population of rapeseed volunteers in a field. *Agriculture Ecosystems and Environment*, 83:235–253, 2001a.

N. Colbach, C. Clermont-Dauphin, and J.M. Meynard. Genesys: a model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers - ii. genetic exchanges among volunteer and cropped populations in a small region. *Agriculture Ecosystems and Environment*, 83:255–270, 2001b.

N. Colbach, H. Monod, and C. Lavigne. A simulation study of the medium-term effects of field patterns on cross-pollination rates in oilseed rape (*Brassica napus L.*). *Ecological Modelling*, 220:662–672, 2009.

T.E. Condeso and R.K. Meentemeyer. Effects of landscape heterogeneity on the emerging forest disease sudden oak death. *Journal of Ecology*, 95:364–375, 2007.

B.C. Couch, I. Fudal, M.H. Lebrun, D. Tharreau, B. Valent, P. van Kim, J.L. Notteghem, and L.M. Kohn. Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics*, 170:613–630, 2005.

T. Day and S. Gandon. Applying population-genetic models in theoretical evolutionary epidemiology. *Ecology Letters*, 10:876–888, 2007.

T. Day and S.R. Proulx. A general theory for the evolutionary dynamics of virulence. *The American Naturalist*, 163:E40–E63, 2004.

M.A.M. de Aguiar, M. Baranger, E.M. Baptestini, L. Kaufman, and Y. Bar-Yam. Global patterns of speciation and diversity. *Nature*, 460:384–388, 2009.

F. Débarre and S. Gandon. Evolution of specialization in a spatially continuous environment. *Journal of Evolutionary Biology*, 23:1090–1099, 2010.

F. Débarre and T. Lenormand. Distance-limited dispersal promotes coexistence at habitat boundaries: reconsidering the competitive exclusion principle. *Ecology Letters*, 14:260–266, 2011.

F. Débarre, T. Lenormand, and S. Gandon. Evolutionary epidemiology of drug-resistance in space. *PLoS Computational Biology*, 5, 2009.

V. Devictor, J. Clavel, R. Julliard, S. Lavergne, D. Mouillot, W. Thuiller, P. Venail, S. Villeger, and N. Mouquet. Defining and measuring ecological specialization. *Journal of Applied Ecology*, 47:15–25, 2010.

U. Dieckmann and R. Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34:579–612, 1996.

M. Doebeli and U. Dieckmann. Speciation along environmental gradients. *Nature*, 421: 259–264, 2003.

M. Durinx, J.A.J. Metz, and G. Meszena. Adaptive dynamics for physiologically structured population models. *Journal of Mathematical Biology*, 56:673–742, 2008.

L. Excoffier and N. Ray. Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology and Evolution*, 23:347–351, 2008.

M.R. Finckh, E.S. Gacek, H. Goyeau, C. Lannou, U. Merz, C.C. Mundt, L. Munk, J. Nadziak, A.C. Newton, C. de Vallavieille-Pope, and M.S. Wolfe. Cereal variety and species mixtures in practice, with emphasis on disease resistance. *Agronomie*, 20:813–837, 2000.

J. Fisher Box. *R.A. Fisher, the Life of a Scientist*. Wiley, New York, (USA), 1978.

H. Flor. Current status of the gene-for-gene concept. *Annual Review of Phytopathology*, 9: 275–296, 1971.

Food and Agriculture Organization. *FAOSTAT*. Food and Agriculture Organization of the United Nations, Rome (It), 2009.

R.T.T. Forman. Some general principles of landscape and regional ecology. *Landscape Ecology*, 10:133–142, 1995.

M.J. Fortin, B. Boots, F. Csillag, and T.K. Remmel. On the role of spatial stochastic models in understanding landscape indices in ecology. *OIKOS*, 102:203–212, 2003.

G.R. Fulford, M.G. Roberts, and J.A.P. Heesterbeek. The metapopulation dynamics of an infectious disease: tuberculosis in possums. *Theoretical Population Biology*, 61:15–29, 2002.

A.P. Galvani. Epidemiology meets evolutionary ecology. *Trends in Ecology and Evolution*, 18:132–139, 2003.

K.A. Garrett and C.C. Mundt. Epidemiology in mixed host populations. *Phytopathology*, 89:984–990, 1999.

K.A. Garrett, L.N. Zúñiga, E. Roncal, G.A. Forbes, C.C. Mundt, Z. Su, and R.J. Nelson. Intraspecific functional diversity in hosts and its effect on disease risk across a climatic gradient. *Ecological Applications*, 19:1868–1883, 2009.

A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis - Second Edition*. Chapman & Hall, New York, USA, 2004.

P.R. Gérard, C. Husson, J. Pinon, and P. Frey. Comparison of genetic and virulence diversity of *Melampsora larici-populina* populations on wild and cultivated poplar and influence of the alternate host. *Phytopathology*, 96:1027–1036, 2006.

S.A.H. Geritz and M. Gyllenberg. Seven answers from adaptive dynamics. *Journal of Evolutionary Biology*, 18:1174–1177, 2005.

S.A.H. Geritz, E. Kisdi, G. Meszena, and J.A.J. Metz. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, 12:35–57, 1998.

V. Gewin. Beyond neutrality - ecology finds its niche. *PLoS Biology*, 4:1306–1310, 2006.

A.J. Gibbs, K. Ohshima, M.J. Phillips, and M.J. Gibbs. The prehistory of potyviruses: their initial radiation was during the dawn of agriculture. *PLoS ONE*, 3:e2523, 2008.

C.A. Gilligan. Sustainable agriculture and plant diseases: an epidemiological perspective. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 363:741–759, 2008.

C.A. Gilligan and F. van den Bosch. Epidemiological models for invasion and persistence of pathogens. *Annual Review of Phytopathology*, 46:2008, 2008.

C.A. Gilligan, J.E. Truscott, and A.J. Stacey. Impact of scale on the effectiveness of disease control strategies for epidemics with cryptic infection in a dynamical landscape: an example for a crop disease. *Journal of the Royal Society Interface*, 4:925–934, 2007.

J. Gilmour. Octal notation for designating physiologic races of plant pathogens. *Nature*, 242:620, 1973.

V. Ginot, S. Gaba, R. Beaudouin, F. Aries, and H. Monod. Combined use of local and ANOVA-based global sensitivity analyses for the investigation of a stochastic dynamic model: application to the case study of an individual-based model of a fish population. *Ecological modelling*, 193:479–491, 2006.

P. Gladders, N. Evans, S. Marcroft, and X. Pinochet. Dissemination of information about management strategies and changes in farming practices for the exploitation of resistance to *Leptosphaeria maculans* (phoma stem canker) in oilseed rape cultivars. *European Journal of Plant Pathology*, 114:117–126, 2006.

P. Gladieux, X.G. Zhang, I. Roldan-Ruiz, V. Caffier, T. Leroy, M. Devaux, S. Van Glabeke, E. Coart, and B. Le Cam. Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Molecular Ecology*, 19:658–674, 2010.

H. Goyeau, R. Park, B. Schaeffer, and C. Lannou. Distribution of pathotypes with regard to host cultivars in french wheat leaf rust populations. *Phytopathology*, 96:264–273, 2006.

H. Goyeau, F. Halkett, M.F. Zapater, J. Carlier, and C. Lannou. Clonality and host selection in the wheat pathogenic fungus *Puccinia triticina*. *Fungal Genetics and Biology*, 44:474–483, 2007.

D. Gravel, T. Bell, C. Barbera, T. Bouvier, T. Pommier, P. Venail, and N. Mouquet. Experimental niche evolution alters the strength of the diversity-productivity relationship. *Nature*, 469:89–92, 2011.

Grenelle. Le grenelle de l'environnement, table ronde - récapitulatif. Technical report, Ministère de l'écologie, Ministère de l'écologie du développement et de l'aménagement durables - Paris, France, 2008.

E.J. Gustafson. Quantifying landscape spatial pattern: what is the state of the art? *Ecosystems*, 1:143–156, 1998.

E.J. Gustafson and G.R. Parker. Relationships between landcover proportion and indices of landscape spatial pattern. *Landscape Ecology*, 7:101–110, 1992.

P. Haccou, P. Jagers, and V.A. Vatutin. *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, Cambridge (GB), 2005.

M.L. Hale, P.W. Lurz, M.D. Shirley, S. Rushton, R.M. Fuller, and K. Wolff. Impact of landscape management on the genetic structure of red squirrel populations. *Science*, 293: 2246–2248, 2001.

O. Hallatschek and D.R. Nelson. Life at the front of an expanding population. *Evolution*, 64:193–206, 2009.

J.M. Halley and Y. Iwasa. Neutral theory as a predictor of avifaunal extinctions after habitat loss. *PNAS*, 108:2316–2321, 2011.

I. Hanski. Metapopulation dynamics. *Nature*, 396:41–49, 1998.

I. Hanski, T. Mononen, and O. Ovaskainen. Eco-evolutionary metapopulation dynamics and the spatial scale of adaptation. *The American Naturalist*, 177:29–43, 2011.

G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

F. Hartig, J.M. Calabrese, B. Reineking, T. Wiegand, and A. Huth. Statistical inference for stochastic simulation models - theory and application. *Ecology Letters*, doi: 10.1111/j.1461-0248.2011.01640.x, 2011.

H.S. He, B.E. DeZonia, and D.J. Mladenoff. An aggregation index (AI) to quantify spatial patterns of landscapes. *Landscape Ecology*, 15:591–601, 2000.

A.P. Hendry, T.J. Farrugia, and M.T. Kinnison. Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology*, 17:20–29, 2008.

A.P. Hendry, M.T. Kinnison, M. Heino, T. Day, T.B. Smith, G. Fitt, C.T. Bergstrom, J. Oakeshott, P.S. Jørgensen, M.P. Zalucki, G. Gilchrist, S. Southerton, A. Sih, S. Strauss, R.F. Denison, and S.P. Carroll. Evolutionary principles and their practical application. *Evolutionary Applications*, 4:159–183, 2011.

H.W. Hethcote. The mathematics of infectious diseases. *SIAM review*, 42:599–653, 2000.

J. Hofbauer and K. Sigmund. Adaptive dynamics and evolutionary stability. *Applied Mathematics Letters*, 3:75–79, 1990.

R.D. Holt. *Species Diversity in Ecological Communities*, chapter 'Ecology at the mesoscale: the influence of regional processes on local communities', pages 77–88. R. Ricklefs and D. Schluter, eds University of Chicago Press, Chicago, IL, 1993.

M. Hovmøller, L. Munk, and H. Ostergard. Observed and predicted changes in virulence gene-frequencies at 11 loci in a local barley powdery mildew population. *Phytopathology*, 83:689–689, 1993.

S.P. Hubbell. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, New Jersey, (USA), 2001.

S.P. Hubbell. Modes of speciation and the lifespans of species under neutrality: a response to the comment of Robert E. Ricklefs. *OIKOS*, 100:193–199, 2003.

G.E. Hutchinson. Concluding remarks. *Cold Spring Harbor Symp. Quantitative Biol.*, 22: 415–427, 1957.

B. Iooss. Revue sur l'analyse de sensibilité globale de modèles numériques. *Journal de la Société Française de Statistique*, 152:3–25, 2011.

B. Iooss and M. Ribatet. Global sensitivity analysis of computer models with functional inputs. *Reliability Engineering and System Safety*, 94:1194–1204, 2009.

R. Johnson. Durable resistance - definition of, genetic-control, and attainment in plant-breeding. *Phytopathology*, 71:567–568, 1981.

T. Johnson. Man-guided evolution in plant rusts. *Science*, 10:357–362, 1961.

M. Keeling. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society B-Biological Sciences*, 266:859–867, 1999.

F. Keesing, R. Holt, and R. Ostfeld. Effects of species diversity on disease risk. *Ecology Letters*, 9:485–498, 2006.

M. Kimura. Evolutionary rate at the molecular level. *Nature*, 217:624–626, 1968.

S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

J.A. Kolmer. Virulence phenotypes of *Puccinia triticina* in the south atlantic states in 1999. *Plant Disease*, 86:288–291, 2002.

M. Kopp. Speciation and the neutral theory of biodiversity. *Bioessays*, 32:564–570, 2010.

J. Langlois, L. Fahrig, G. Merriam, and H. Artsob. Landscape structure influences continental distribution of hantavirus in deer mice. *Landscape Ecology*, 16:255–266, 2001.

C. Lannou, C. de Vallavieille-Pope, C. Biass, and H. Goyeau. The efficacy of mixtures of susceptible and resistant hosts to 2 wheat rusts of different lesion size - controlled experiments and computerized simulations. *Journal of Phytopathology*, 140:227–237, 1994.

C. Lannou, P. Hubert, and C. Gimeno. Competition and interactions among stripe rust pathotypes in wheat cultivar mixtures. *Plant Pathology*, 54:699–712, 2005.

C. Lavigne, E.K. Klein, J.F. Mari, F. Le Ber, K. Adamczyk, H. Monod, and F. Angevin. How do genetically modified (GM) crops contribute to background levels of GM pollen in an agricultural landscape? *Journal of Applied Ecology*, 45:1104–1113, 2008.

F. Le Ber, C. Lavigne, J.F. Mari, K. Adamczyk, and F. Angevin. Genexp, un logiciel pour simuler des paysages agricoles en vue de l'étude de la diffusion de transgènes. *SAGEO*, pages 1–12, 2006.

F. Le Ber, C. Lavigne, K. Adamczyk, F. Angevin, N. Colbach, J.F. Mari, and H. Monod. Neutral modelling of agricultural landscapes by tessellation methods - Application for gene flow simulation. *Ecological Modelling*, 220:3536–3545, 2009.

M.A. Leibold and M.A. McPeek. Coexistence of the niche and neutral perspectives in community ecology. *Ecology*, 87:1399–1410, 2006.

M.A. Leibold, M. Holyoak, N. Mouquet, P. Amarasekare, J.M. Chase, M.F. Hoopes, R.D. Holt, J.B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, 7:601–613, 2004.

T. Lenormand. Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, 17:183–189, 2002.

H. Levene. Genetic equilibrium when more than one niche is available. *The American Naturalist*, 87:331–333, 1953.

R. Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15:237–240, 1969.

H.B. Li and J.F. Reynolds. A simulation experiment to quantify spatial heterogeneity in categorical maps. *Ecology*, 75:2446–2455, 1994.

H.B. Li and J.F. Reynolds. On definition and quantification of heterogeneity. *OIKOS*, 73: 280–284, 1995.

H.B. Li and J. Wu. Use and misuse of landscape indices. *Landscape Ecology*, 19:389–399, 2004.

C.M. Lively. The effect of host genetic diversity on disease spread. *The American Naturalist*, 175:E149–E152, 2010.

E. Lô-Pelzer, L. Bousset, M.H. Jeuffroy, M.U. Salam, X. Pinochet, M. Boillot, and J.N. Aubertot. SIPPOM-WOSR: a simulator for integrated pathogen population management of phoma stem canker on winter oilseed rape I. description of the model. *Field Crops Research*, 118:73–81, 2010.

M. Loreau, S. Naeem, P. Inchausti, J. Bengtsson, J.P. Grime, A. Hector, D.U. Hooper, M.A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D.A. Wardle. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, 294:804–808, 2001.

M. Loreau, N. Mouquet, and A. Gonzalez. Biodiversity as spatial insurance in heterogeneous landscapes. *PNAS*, 100:12765–12770, 2003a.

M. Loreau, N. Mouquet, and R.D. Holt. Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. *Ecology Letters*, 6:673–679, 2003b.

A. Lurette, S. Touzeau, M. Lamboni, and H. Monod. Sensitivity analysis to identify key parameters influencing salmonella infection dynamics in a pig batch. *Journal of Theoretical Biology*, 258:43–52, 2009.

L.V. Madden, G. Hughes, and F. van den Bosch. *The Study of Plant Disease Epidemics*. APS PRESS, St. Paul (USA), 2008.

T.C. Marcel, B. Gorguet, M.T. Ta, Z. Kohutova, A. Vels, and R.E. Niks. Isolate specificity of quantitative trait loci for partial resistance of barley to *Puccinia hordei* confirmed in mapping populations and near-isogenic lines. *New Phytologist*, 177:743–755, 2008.

M. Margosian, K. Garrett, J. Shawm Hutchinson, and K. With. Connectivity of the American agricultural landscape: assessing the national risk of crop pest and disease spread. *Bioscience*, 59:141–151, 2009.

F. Massol, V. Calcagno, and J. Massol. The metapopulation fitness criterion: proof and perspectives. *Theoretical Population Biology*, 75:183–200, 2009.

J. Maynard Smith and G.R. Price. The logic of animal conflict. *Nature*, 246:15–18, 1973.

B.A. McDonald. How can we achieve durable disease resistance in agricultural ecosystems? *New Phytologist*, 185:5–7, 2010.

T.D. Meehan, B.P. Werling, D.A. Landis, and C. Gratton. Agricultural landscape simplification and insecticide use in the midwestern united states. *PNAS*, 108:11500–11505, 2011.

B.A. Melbourne, H.V. Cornell, K.F. Davies, C.J. Dugaw, S. Elmendorf, A.L. Freestone, R.J. Hall, S. Harrison, A. Hastings, M. Holland, M. Holyoak, J. Lambrinos, K. Moore, and H. Yokomizo. Invasion in a heterogeneous world: resistance, coexistence or hostile takeover? *Ecology Letters*, 10:77–94, 2007.

X.L. Meng and D.B. Rubin. Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika*, 80:267–278, 1993.

S. Messinger and A. Ostling. The consequences of spatial structure for the evolution of pathogen transmission rate and virulence. *The American Naturalist*, 174:441–454, 2009.

G. Meszéna, I. Czibula, and S.A.H. Geritz. Adaptative dynamics in a 2-patch environment: a toy model for allopatric and parapatric speciation. *Journal of Biological Systems*, 5: 265–284, 1997.

J.M. Meynard, T. Doré, and P. Lucas. Agronomic approach: cropping systems and plant diseases. *Comptes Rendus Biologies*, 326:37–46, 2002.

J.S. Miller, D.A. Johnson, and P.B. Hamm. Aggressiveness of isolates of *Phytophthora infestans* from the columbia basin of washington and oregon. *Phytopathology*, 88:190–197, 1998.

C.E. Mitchell, D. Tilman, and J.V. Groth. Effects of grassland plant species diversity, abundance, and composition on foliar fungal disease. *Ecology*, 83:1713–1726, 2002.

H. Monod, C. Naud, and D. Makowski. Uncertainty and sensitivity analysis for crop models. In D. Wallach, D. Makowski, and J.W. Jones, editors, *Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Applications*, chapter 4, pages 55–100. Elsevier, 2006.

N. Mouquet and M. Loreau. Community patterns in source-sink metacommunities. *The American Naturalist*, 162:544–557, 2003.

C.C. Mundt. Probability of mutation to multiple virulence and durability of resistance gene pyramids. *Phytopathology*, 80:221–223, 1990.

C.C. Mundt. Use of multiline cultivars and cultivar mixtures for disease management. *Annual Review of Phytopathology*, 40:381–410, 2002.

C.C. Mundt and L. Brophy. Influence of number of host genotype units on the effectiveness of host mixtures for disease control: a modeling approach. *Phytopathology*, 78:1087–1094, 1988.

C.C. Mundt and K.J. Leonard. Effect of host genotype unit area on epidemic development of crown rust following focal and general inoculations of mixtures of immune and susceptible oat plants. *Phytopathology*, 75:1141–1145, 1985.

C.C. Mundt and K.J. Leonard. Effect of host genotype unit area on development of focal epidemics of bean rust and common maize rust in mixtures of resistant and susceptible plants. *Phytopathology*, 76:895–900, 1986.

C.C. Mundt, K.J. Leonard, W.M. Thal, and J.H. Fulton. Computerized simulation of crown rust epidemics in mixtures of immune and susceptible oat plants with different genotype unit areas and spatial distributions of initial disease. *Phytopathology*, 86:590–598, 1986.

C.C. Mundt, K.E. Sackett, L.D. Wallace, C. Cowger, and J.P. Dudley. Long-distance dispersal and accelerating waves of disease: empirical relationships. *The American Naturalist*, 173:456–466, 2009.

A.B. Munkacsi, S. Stoxen, and G. May. *Ustilago maydis* populations tracked maize through domestication and cultivation in the Americas. *Proceedings of the Royal Society B-Biological Sciences*, 275:1037–1046, 2008.

S. Nee. The neutral theory of biodiversity: do the numbers add up? *Functional Ecology*, 19: 173–176, 2005.

M.G. Neubert and H. Caswell. Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. *Ecology*, 81:1613–1628, 2000.

M. Nowak and K. Sigmund. The evolution of stochastic strategies in the prisoner's dilemma. *Acta Applicandae Mathematicae*, 20:247–265, 1990.

E. Oerke and H. Dehne. Safeguarding production - losses in major crops and the role of crop protection. *Crop Protection*, 23:275–285, 2004.

A. Okabe, B. Boots, and K. Sugihara. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley, New York, (USA), 1992.

R. O'Neill. Indices of landscape pattern. *Landscape Ecology*, 1:153–162, 1988.

ONIGC. Répartition des variétés - récoltes 1999-2008 (blé tendre, orge, blé dur,triticale). Technical report, France Agrimer, Paris, France, 2008.

R.S. Ostfeld and F. Keesing. Biodiversity and disease risk: the case of lyme disease. *Conservation Biology*, 14:722–728, 2000.

R.S. Ostfeld, G.E. Glass, and F. Keesing. Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology and Evolution*, 20:328–336, 2005.

O. Ovaskainen and I. Kanski. How much does an individual habitat fragment contribute to metapopulation dynamics and persistence? *Theoretical Population Biology*, 64:481–495, 2003.

S.R. Palumbi. Evolution - humans as the world's greatest evolutionary force. *Science*, 293: 1786–1790, 2001.

J. Papaïx, H. Goyeau, P. du Cheyron, H. Monod, and C. Lannou. Influence of cultivated landscape composition on variety resistance: an assesment based on the wheat leaf rust epidemics. *New phytologist*, 191:1095–1107, 2011.

B. Pariaud. *Agressivité de Puccinia triticina (agent de la rouille brune du blé) et son adaptation à l'hôte*. PhD thesis, Université Paris-Sud 11, 2008.

B. Pariaud, V. Ravigné, F. Halkett, H. Goyeau, J. Carlier, and C. Lannou. Aggressiveness and its role in the adaptation of plant pathogens. *Plant Pathology*, 58:409–424, 2009a.

B. Pariaud, C. Robert, H. Goyeau, and C. Lannou. Aggressiveness components and adaptation to a host cultivar in wheat leaf rust. *Phytopathology*, 99:869–878, 2009b.

A. Park, S. Gubbins, and C.A. Gilligan. Invasion and persistence of plant parasites in a spatially structured host population. *OIKOS*, 94:162–174, 2001.

R.F. Park, A. Jahoor, and F.G. Felsenstein. Population structure of *Puccinia recondita* in Western Europe during 1995, as assessed by variability in pathogenicity and molecular markers. *Journal of Phytopathology*, 148:169–179, 2000.

S. Parnell, F. van den Bosch, and C.A. Gilligan. Large-scale fungicide spray heterogeneity and the regional spread of resistant pathogen strains. *Phytopathology*, 96:549–555, 2006.

S. Parnell, T.R. Gottwald, F. van den Bosch, and C.A. Gilligan. Optimal strategies for the eradication of asiatic citrus canker in heterogeneous host landscapes. *Phytopathology*, 99: 1370–1376, 2009.

K. Parvinen and M. Egas. Dispersal and the evolution of specialisation in a two-habitat type metapopulation. *Theoretical Population Biology*, 66:233–248, 2004.

M. Pautasso, O. Holdenrieder, and J. Stenlid. Susceptibility to fungal pathogens of forests differing in tree diversity. *Ecological Studies*, 176:263–289, 2005.

S.L. Peck. Simulation as experiment: a philosophical reassessment for biological modeling. *Trends in Ecology and Evolution*, 19:530–534, 2004.

S.L. Peck. The hermeneutics of ecological simulation. *Biology and Philosophy*, 23:383–402, 2008.

J. Peng, Y. Wang, Y. Zhang, J. Wu, W. Li, and Y. Li. Evaluating the effectiveness of landscape metrics in quantifying spatial patterns. *Ecological Indicators*, 10:217–223, 2010.

P.O. Persson and G. Strang. A simple mesh generator in matlab. *SIAM Review*, 46:329–345, 2004.

M. Plantegenest, C. Le May, and F. Fabre. Landscape epidemiology of plant diseases. *Journal of the Royal Society Interface*, 4:963–972, 2007.

M. Plummer. Jags version 2.1.0 Manual. Technical report, accessed 11 March 2011, URL: http://www-fis.iarc.fr/ martyn/software/jags/, 2010.

H.R. Pulliam. On the relationship between niche and distribution. *Ecology Letters*, 3:349–361, 2000.

R.E. Ricklefs. Community diversity: relative roles of local and regional processes. *Science*, 235:167–171, 1987.

R.E. Ricklefs. Applying a regional community concept to forest birds of Eastern North America. *PNAS*, 108:2300–2305, 2011.

K.H. Riitters, R.V. O'Neill, C.T. Hunsaker, J.D. Wickham, D.H. Yankee, S.P. Timmins, K.B. Jones, and B.L. Jackson. A factor analysis of landscape pattern and structure metrics. *Landscape Ecology*, 10:23–39, 1995.

R.A. Robinson and W.J. Sutherland. Post-war changes in arable farming and biodiversity in great britain. *Journal of Applied Ecology*, 39:157–176, 2002.

T. Rouxel, A. Penaud, X. Pinochet, H. Brun, L. Gout, R. Delourme, J. Schmit, and M. Balesdent. A 10-year survey of populations of *Leptosphaeria maculans* in France indicates a rapid adaptation towards the *rlm1* resistance gene of oilseed rape. *European Journal of Plant Pathology*, 109:871–881, 2003.

A. Saltelli. Variance-based methods. In A. Saltelli, K. Chan, and E.M. Scott, editors, *Sensitivity Analysis*, Probability and Statistics, chapter 8. Wiley & Sons, 2000.

A. Saltelli, S. Tarantola, and F. Campolongo. Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15:377–395, 2000.

A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis. The Primer*. Wiley, Chichester (UK), 2008.

D.J. Samborski. *Wheat Leaf Rust*. Roelfs, A.P. and Bushnall, W.R., Orlando, USA, 1985.

N. Sapoukhina, Y. Tyutyunov, I. Sache, and R. Arditi. Spatially mixed crops to control the stratified dispersal of airborne fungal diseases. *Ecological Modelling*, 221:2793–2800, 2010.

R. Seppelt, F. Müller, B. Schröder, and M. Volk. Challenges of simulating complex environmental systems at the landscape scale: a controversial dialogue between two cups of espresso. *Ecological Modelling*, 220:3481–3489, 2009.

P. Skelsey, G.J.T. Kessel, W.A.H. Rossing, and W. van der Werf. Parameterization and evaluation of a spatiotemporal model of the potato late blight pathosystem. *Phytopathology*, 99:290–300, 2009.

P. Skelsey, W.A.H. Rossing, G.J.T. Kessel, and W. van der Werf. Invasion of *Phytophthora infestans* at the landscape level: how do spatial scale and weather modulate the consequences of spatial heterogeneity in host resistance? *Phytopathology*, 100:1146–1161, 2010.

I.M. Sobol. Sensitivity estimates for non linear mathematical models. *Mathematical Modelling and Computational Experiments*, 1:407–414, 1993.

D. Söndgerath and B. Schröder. Population dynamics and habitat connectivity affecting the spatial spread of populations - a simulation study. *Landscape Ecology*, 17:57–70, 2002.

S. Soubeyrand, J. Enjalbert, A. Sanchez, and I. Sache. Anisotropy, in density and in distance, of the dispersal of yellow rust of wheat: experiments in large field plots and estimation. *Phytopathology*, 97:1315–1324, 2007.

S. Soubeyrand, A.L. Laine, I. Hanski, and A. Penttinen. Spatiotemporal structure of host-pathogen interactions in a metapopulation. *The American Naturalist*, 174:308–320, 2009.

E. Stukenbrock and B.A. McDonald. The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, 46:75–100, 2008.

E.H. Stukenbrock, S. Banke, M. Javan-Nikkhah, and B.A. McDonald. Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Molecular Biology and Evolution*, 24:398–411, 2007.

B. Sudret. Global sensitivity analysis unsing polynomials chaos expansions. *Reliability Engineering & System Safety*, 93:964–979, 2008.

P.D. Taylo, L. Fahrig, K. Henein, and G. Merriam. Connectivity is a vital element of landscape structure. *OIKOS*, 68:571–573, 1993.

P.H. Thrall and J.J. Burdon. Evolution of virulence in a plant host-pathogen metapopulation. *Science*, 299:1735–1737, 2003.

P.H. Thrall, J.D. Bever, and J.J. Burdon. Evolutionary change in agriculture: the past, present and future. *Evolutionary Applications*, 3:405–408, 2010.

M. Tildesley, T. House, M. Bruhn, R. Curry, M. O'Neil, J. Allpress, G. Smith, and M. Keeling. Impact of spatial clustering on disease transmission and optimal control. *PNAS*, 107:1041 – 1046, 2010.

D. Tilman. Global environmental impacts of agricultural expansion: the need for sustainable and efficient practices. *PNAS*, 96:5995–6000, 1999.

D. Tilman, K.G. Cassman, P.A. Matson, R. Naylor, and S. Polasky. Agricultural sustainability and intensive production practices. *Nature*, 418:671–677, 2002.

L. Tischendorf. Can landscape indices predict ecological processes consistently? *Landscape Ecology*, 16:235–254, 2001.

J. Tufto, S. Engen, and K. Hindar. Stochastic dispersal processes in plant populations. *Theoretical Population Biology*, 52:16–26, 1997.

S. Tuljapurkar and H. Caswell. *Structured-Population Models in Marine, Terrestrial, and Freshwater Systems*. Chapman & Hall, New York (USA), 1997.

M.G. Turner. Landscape ecology: what is the state of the science? *Annual Review of Ecology Evolution and Systematics*, 36:319–344, 2005.

M.G. Turner and R.H. Gardner. *Quantitative Methods in Landscape Ecology: an Introduction*. Springer-Verlag, New York (USA), 1991.

F. van den Berg, N. Bacaer, J.A.J. Metz, C. Lannou, and F. van den Bosch. Periodic host absence can select for higher or lower parasite transmission rates. *Evolutionary Ecology*, 2010.

F. van den Bosch and C.A. Gilligan. Measures of durability of resistance. *Phytopathology*, 93:616–625, 2003.

P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180:29–48, 2002.

V. Viaud, H. Monod, C. Lavigne, F. Angevin, and K. Adamczyk. Spatial sensitivity of maize gene-flow to landscape pattern: a simulation approach. *Landscape Ecology*, 23:1067–1079, 2008.

P.M. Vitousek, H.A. Mooney, J. Lubchenco, and J.M. Melillo. Human domination of earth's ecosystems. *Science*, 277:494–499, 1997.

I. Volkov, J.R. Banavar, S.P. Hubbell, and A. Maritan. Neutral theory and relative species abundance in ecology. *Nature*, 424:1035–1037, 2003.

D. Waxman and S. Gavrilets. 20 questions on adaptive dynamics. *Journal of Evolutionary Biology*, 18:1139–1154, 2005.

W. Wei and S.M. Krone. Spatial invasion by a mutant pathogen. *Journal of Theoretical Biology*, 236:335–348, 2005.

J. Whitfield. Neutrality versus the niche. *Nature*, 417:480–481, 2002.

K.A. With. The landscape ecology of invasive spread. *Conservation Biology*, 16:1192–1203, 2002.

K.A. With, R.H. Gardner, and M.G. Turner. Landscape connectivity and population distributions in heterogeneous environments. *OIKOS*, 78:151–169, 1997.

M. Wolfe and E. Schwarzbach. Patterns of race changes in powdery mildews. *Annual Reviews of Phytopathology*, 16:159–180, 1978.

X.M. Xu. A simulation study on managing plant diseases by systematically altering spatial positions of cultivar mixture components between seasons. *Plant Pathology*, 2011.

S. Yachi and M. Loreau. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *PNAS*, 96:1463–1468, 1999.

D. Young, T. Benaglia, D. Chauveau, R. Elmore, T. Hettmansperger, D. Hunter, H. Thomas, and F. Xuan. Tools for analyzing finite mixture models. Technical report, R CRAN, 2010, Version 0.4.4.

J.C. Zadoks. Cereal rusts, dogs and stars in antiquity. *Garcia de Orta, Ser. Est. Agron. Lisboa*, 9:13–20, 1982.

J.C. Zadoks and P. Kampmeijer. The role of crop populations and their deployment, illustrated by means of a simulator, epimul 76. *Annals of the New York Academy of Sciences*, 287:164–190, 1977.

Y. Zhu, H. Chen, J. Fan, Y. Wang, Y. Li, J. Chen, J. Fan, S. Yang, L. Hu, H. Leung, T. Mew, P. Teng, Z. Wang, and C.C. Mundt. Genetic diversity and disease control in rice. *Nature*, 406:718–722, 2000.

Supporting information

Appendix 1 :
Technical information related to the bayesian hierarchical model for wheat leaf rust epidemics

Appendix 2 :
Technical information related to the stochastic simulation model of a plant pathogen foliar fungus dynamics on an agricultural landscape

Appendix 3 :
Computation of the basic reproductive number

Appendix 4 :
Supporting information relative the chapter 6, Evolution of specialization in spatially heterogeneous environments

Appendix 5 :
Hétérogénéités temporelle et spatio-temporelle dans un modèle de dynamique adaptative

Appendix 6 :
PhD defence slides

Appendix 1

Technical information related to the bayesian hierarchical model for wheat leaf rust epidemics

This appendix gives technical details related to model construction. We first explain how the population composition sub-model was constructed and we then detail the disease severity sub-model.

In order to make MCMC convergence easier, we defined a conjugated model for the population composition sub-model. Thus, because $Y_{v,t} = \left(y_{v,t,1}, \cdots, y_{v,t,P}\right)$ has a multinomial distribution with parameters $N_{v,t}$ and $\Pi_{v,t} = \left(\pi_{v,t,1}, \cdots, \pi_{v,t,P}\right)$, we supposed that $\Pi_{v,t}$ had a Dirichlet distribution. A way to simulate a Dirichlet distribution was to generate P variables $Z_{v,t,p}$ as $Z_{v,t,p} \sim \text{Gamma}\left(\bar{Z}_{v,t,p}, 1\right)$. $\pi_{v,t,P}$ was then defined as $\frac{Z_{v,t,p}}{\sum_{i=1}^{P} Z_{v,t,p}}$. Note that Z should not be considered as the actual size of the pathotype population but as a scale variable instead. Since $\text{E}[Z_{v,t,p}] = \bar{Z}_{v,t,p}$, it was possible to link the $\pi_{v,t,p}$ to wheat variety frequencies assuming that the expectation of $Z_{v,t,p}$ is influenced by the landscape composition through the regression equation

$$\text{E}[Z_{v,t,p}] = \bar{Z}_{v,t,p} = \alpha_{v,p} + \sum_{i=1}^{V} \delta_{i,p}\, \beta_{i,p}\, \phi_{i,t}.$$

Based on the modelling framework of the population composition sub-model, we constructed a coherent model for the disease severity sub-model. Let $X_{v,t,r,e}$ be the disease score attributed to variety v year t on trial e in region r in the Arvalis survey, $X_{v,t,r,e}|\mu_{v,t,r}, \sigma_e \sim \text{Beta}\left(\mu_{v,t,r}, \sigma_e\right)$. The Beta distribution is parameterised by its mean value $\mu_{v,t,r}$ and scale parameter σ_e. As for $\pi_{v,t,p}$, we introduced two scale variables $S_{v,t,r}$ and $S'_{v,t,r}$ that represent the diseased leaf area and the healthy leaf area of a variety v, respectively, during year t in the climatic region r. Since disease scores are defined in the scoring procedure as the observed proportion of diseased leaf area, $\mu_{v,t,r} = \text{E}[X_{v,t,r,e}]$ can be identified with the mean proportion of leaf area that was diseased for variety v during year t in region r. Thus, $\mu_{v,t,r} = \frac{S_{v,t,r}}{S_{v,t,r} + S'_{v,t,r}}$. To keep the model globally coherent, we assumed that $S_{v,t,r} \sim \text{Gamma}\left(\bar{S}_{v,t,r}, 1\right)$. Since different pathotypes may produce different levels of disease severity on a given variety, the expectation of $S_{v,t,r}$ depends on the composition of the pathogen population through the regression equation

$$\text{E}[S_{v,t,r}] = \bar{S}_{v,t,r} = a_v^0 + a_t^1 + a_r^2 + \sum_{j=1}^{P-1} \delta_{v,j}\, b_{v,j}\, \pi_{v,t,j}.$$

For identifiability reasons, $S'_{v,t,r}$ was fixed *a priori*.

We provide below the JAGS script to fit the model to the three datasets on wheat leaf rust.

```
model
{
### a priori ###
for (v in 1:V)
  {
    a0[v] ~ dunif(0,10000)
    b1[v] ~ dunif(0,10000)
    b2[v] ~ dunif(0,10000)
    b3[v] ~ dunif(0,10000)
    b4[v] ~ dunif(0,10000)
    b5[v] ~ dunif(0,10000)

    for (p in 1:(P-1))
      {
        alpha[v,p] ~ dunif(0,10000)
      }
  }

for (t in 1:T)
  {
    a1[t] ~ dunif(0,10000)
  }

for (r in 1:R)
  {
    a2[r] ~ dunif(0,10000)
  }

for (p in 1:(P-1))
  {
    beta1[p]~dunif(0,10000)
    beta2[p]~dunif(0,10000)
    beta3[p]~dunif(0,10000)
    beta4[p]~dunif(0,10000)
```

```
    beta5[p]~dunif(0,10000)
    beta6[p]~dunif(0,10000)
    beta7[p]~dunif(0,10000)
  }

for (e in 1:E)
  {
    sig[e] ~ dunif(0,1)
  }

### meta model  ###

# Sbar model
for (v in 1:V)
  {
    for (t in prem[v]:der[v])
      {
        for (r in 1:R)
          {
            Sbar[v,t,r] <- a0[v]+ a1[t] + a2[r]
                           + b1[v]*Prop[v,t,1]*DELTA1[v]
                           + b2[v]*Prop[v,t,2]*DELTA2[v]
                           + b3[v]*Prop[v,t,3]*DELTA3[v]
                           + b4[v]*Prop[v,t,4]*DELTA4[v]
                           + b5[v]*Prop[v,t,5]*DELTA5[v]
          }
      }
  }

# Zbar model
for (v in 1:V)
  {
    for (t in prem[v]:der[v])
      {
        Z[v,t,P] <- 1
        for (p in 1:(P-1))
          {
            Zbar[v,t,p] <- alpha[v,p]
```

```
                            + beta1[p]*phi[1,t]*delta1[p]
                            + beta2[p]*phi[2,t]*delta2[p]
                            + beta3[p]*phi[3,t]*delta3[p]
                            + beta4[p]*phi[4,t]*delta4[p]
                            + beta5[p]*phi[5,t]*delta5[p]
                            + beta6[p]*phi[6,t]*delta6[p]
                            + beta7[p]*phi[7,t]*delta7[p]
          }
      }
  }

# mu and pi computation
for (v in 1:V)
  {
    for (t in prem[v]:der[v])
      {
        for (r in 1:R)
          {
            S[v,t,r] ~ dgamma(Sbar[v,t,r],1)
            mu[v,t,r] <- S[v,t,r]/(S[v,t,r]+5)
          }
        for (p in 1:P)
          {
            Z[v,t,p] ~ dgamma(Zbar[v,t,p],1)
            pi[v,t,p] <- Z[v,t,p]/sum(Z[v,t,])
          }
      }
  }

### likelihood ###

# composition of leaf rust population
for (v in 1:V)
  {
    for (t in prem[v]:der[v])
      {
        Y[v,t,1:P] ~ dmulti(pi[v,t,],N[v,t])
        Yrep[v,t,1:P] ~ dmulti(pi[v,t,],N[v,t]) #prediction
```

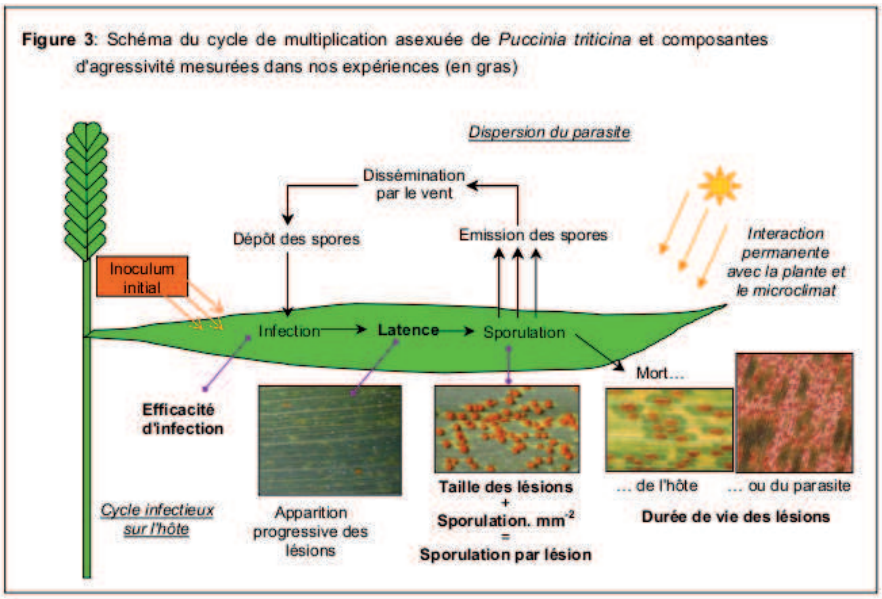
```
      }
  }
# Yrep vectorisation
for (i in 1:NYrep)
  {
    vecYrep[i] <- Yrep[varieteY[i],anneeY[i],pathoY[i]]
  }

# beta parametrization
for (v in 1:V)
  {
    for (t in prem[v]:der[v])
      {
        for (r in 1:R)
          {
            for (e in 1:E)
              {
                p1[v,t,r,e] <- ((1-(sig[e]*sig[e]))*(1-mu[v,t,r]))/(sig[e]*sig[e])
                p2[v,t,r,e] <- ((1-(sig[e]*sig[e]))*mu[v,t,r])/(sig[e]*sig[e])
              }
          }
      }
  }
# disease scores
for(i in 1:NX)
  {
    X[i] ~ dbeta(p1[variete[i],annee[i],region[i],essai[i]],
                 p2[variete[i],annee[i],region[i],essai[i]])
    Xrep[i] ~ dbeta(p1[variete[i],annee[i],region[i],essai[i]],
                    p2[variete[i],annee[i],region[i],essai[i]])   #prediction
  }
```

Appendix 2

Technical information related to the stochastic simulation model of a plant pathogen foliar fungus dynamics on an agricultural landscape

The model is based on the biology of a plant pathogen foliar fungus like wheat leaf rust (*Puccinia triticina* on *Triticum aestivum*). *Puccinia triticina* is an air-borne basidiomycete fungus strictly biotrophe (Bolton et al., 2008). Wheat leaf rust epidemics result from the seccession of asexual life cycles (Figure 1 and Pariaud, 2008). Once deposited on the host and germinated, spores that have penetrated in the leaf give latent lesions and develop structures for host explotation resources. Then the lesions produce spores and die.



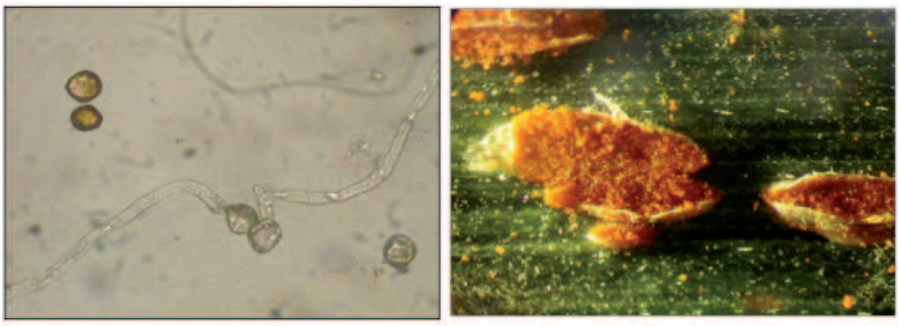


Photo 1: Spores germées et non germées de *Puccinia triticina* (microscope optique x640)

Photo 2: Lésions de *Puccinia triticina* (loupe binoculaire, x40)

Figure 1: *Puccinia triticina* life cycle. From Pariaud (2008)

A host individual could be seen as a set of infection sites in each of which just one lesion could develop. Thus, in the following we do not consider host individuals but total infection sites present in a patch. (Pariaud, 2008) showed also that the total number of emitted spores per surface unit of sporulating tissue just exhibited a small linear decrease when sporulating lesions density increases. So, we do not consider active competition between lesions. We describe below each step of the stochastic model in its more general version.

1. Computation of the spore cloud composition arriving in patch j:

$$S_{j,p}(t) = \sum_{i=1}^{m} Multi\Big(r_{v(i),p} I_{i,p}(t-1), \Big[m_{i1}, \cdots, m_{im}\Big]\Big)_j.$$

2. Determination of contaminated sites:

   (a) infection sites accessibility in patch j:

$$\pi_j(t) = f\left(\frac{H_j(t-1)}{K_j}\right), \quad f(x) = \xi \cdot \left(1 - \frac{e^{-\kappa \cdot x^{\sigma}} - e^{-\kappa}}{1 - e^{-\kappa}}\right) \ \forall x \in [0,1].$$

   Where ξ is the threshold parameter and κ and σ are two form parameters ($\kappa > 0$ and $\sigma > 0$). Figure 2 gives some examples for $f(.)$.

   (b) Accessible infection sites:

$$H'_j(t) \sim Bin\left(H_j(t-1), \pi_j(t)\right).$$

   (c) Repartition of accessible infection sites among pathogen strains:

$$H''_j(t) \sim Multi\left(H'_j(t), \left[\frac{S_{j,1}(t)}{\sum_{p=1}^{P} S_{j,p}(t)}, \cdots, \frac{S_{j,P}(t)}{\sum_{p=1}^{P} S_{j,p}(t)}\right]\right).$$

   (d) Regulation by disponible spores:

$$S'_{j,p}(t) = min\left(H''_{j,p}(t), S_{j,p}(t)\right).$$

3. Infected infection sites:

$$(H_j(t-1) \to L_{j,p}(t)) \sim Bin\left(S'_{j,p}(t), e_{v(i),p}\right).$$

Latent and sporulating lesions are determined as follow::

$$(L_{j,p}(t-1) \to I_{j,p}(t)) \sim Bin\left(L_{j,p}(t-1), 1 - e^{-\frac{1}{\tau_{v(j),p}}}\right),$$

$$(I_{j,p}(t-1) \to R_{j,p}(t)) \sim Bin\left(I_{j,p}(t-1), 1 - e^{-\frac{1}{T_{v(j),p}}}\right).$$

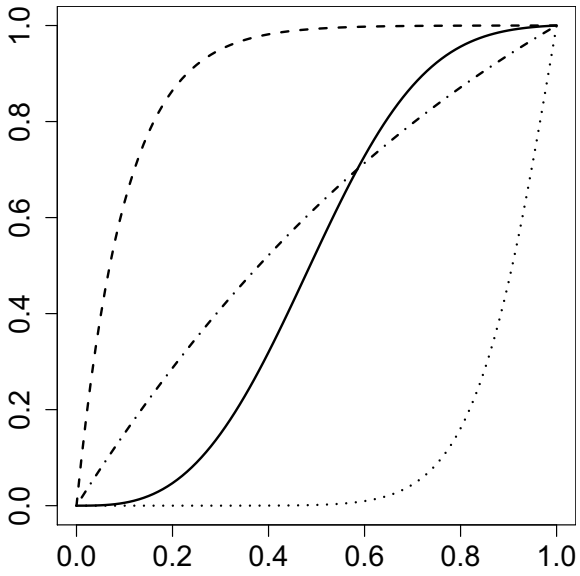


Figure 2: π function for several parameters κ and σ. Solid line: $\kappa = 6$, $\sigma = 6$; dashed line: $\kappa = 10$, $\sigma = 1$; dotted line: $\kappa = 1$, $\sigma = 10$; dash-dot line: $\kappa = 1$, $\sigma = 1$.

Appendix 3

Computation of the basic reproductive number

The deterministic version of the model described in chapter 3 is synthesised for patch i ($i = 1, ..., N$) by the following system of equations:

$$\begin{cases} \frac{dH_i}{dt} &= -e_{v(i)} \cdot \pi_i \cdot \sum_{j=1}^{N} r \cdot m_{ji} \cdot I_j, \\ \frac{dL_i}{dt} &= e_{v(i)} \cdot \pi_i \sum_{j=1}^{N} r \cdot m_{ji} \cdot I_j - \frac{1}{\tau} \cdot L_i, \\ \frac{dI_i}{dt} &= \frac{1}{\tau} \cdot L_i - \frac{1}{T} \cdot I_i, \\ \frac{dR_i}{dt} &= \frac{1}{T} \cdot I_i, \end{cases} \tag{1}$$

where $v(i)$ denotes the variety present in patch i, e_u is the infection efficiency of the pathogen on variety u, π_i is the probability for a spore to encounter a free infection site in patch i, r is the number of spores produced per day by a sporulating lesion, m_{ji} is the proportion of spores emitted by patch j that land in patch i, τ is the latency duration and T is the duration of the infectious period.

With the next-generation matrix method (van den Driessche and Watmough, 2002), we only need to consider the dynamics of the infected states, to compute the basic reproductive number, R_0. Considering all patches, it follows from equation (1):

$$\begin{cases} \frac{dL_1}{dt} &= e_{v(1)} \cdot \pi_1 \sum_{j=1}^{N} r \cdot m_{j1} \cdot I_j - \frac{1}{\tau} \cdot L_1, \\ \vdots & \\ \frac{dL_N}{dt} &= e_{v(N)} \cdot \pi_N \sum_{j=1}^{N} r \cdot m_{jN} \cdot I_j - \frac{1}{\tau} \cdot L_N, \\ \frac{dI_1}{dt} &= \frac{1}{\tau} \cdot L_1 - \frac{1}{T} \cdot I_1. \\ \vdots & \\ \frac{dI_N}{dt} &= \frac{1}{\tau} \cdot L_N - \frac{1}{T} \cdot I_N. \end{cases} \tag{2}$$

The dynamics of the infected states can be decomposed into new infections and transfers between compartments:

$$\begin{pmatrix} \dot{L} \\ \dot{I} \end{pmatrix} = \mathcal{F} + \mathcal{V},$$

where

$$\mathcal{F} = \begin{pmatrix} e_{v(1)}\pi_1 r \sum_{j=1}^{N} m_{j1} I_j \\ \vdots \\ e_{v(N)}\pi_N r \sum_{j=1}^{N} m_{jN} I_j \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } \mathcal{V} = \begin{pmatrix} -\frac{1}{\tau} L_1 \\ \vdots \\ -\frac{1}{\tau} L_N \\ \frac{1}{\tau} L_1 - \frac{1}{T} I_1 \\ \vdots \\ \frac{1}{\tau} L_N - \frac{1}{T} I_N \end{pmatrix}.$$

The next generation matrix K is then given by $K = -FV^{-1}$, where F is the Jacobian matrix of new infections and V is the Jacobian matrix of transfers between compartments, evaluated at the disease-free equilibrium (DFE). The basic reproductive number is defined as the spectral radius, *i.e.* the greatest eigenvalue in magnitude, $R_0 = \rho\left(-FV^{-1}\right)$. If $R_0 > 1$, then the DFE is unstable whereas, if $R_0 < 1$, the DFE is locally stable.

In our case, the DFE is, for all i, $L_i^\star = I_i^\star = 0$ and $H_i^\star = K_i$, where K_i is the carrying capacity of patch i. In addition, we assume that at the DFE, $\pi_i = 1$. We have thus:

$$F = \begin{array}{c} \\ L_1 \\ \vdots \\ L_N \\ I_1 \\ \vdots \\ I_N \end{array} \begin{array}{c} \begin{array}{cccccc} \partial L_1 & \cdots & \partial L_N & \partial I_1 & \cdots & \partial I_N \end{array} \\ \left(\begin{array}{ccc|ccc} & & & & & \\ & 0 & & & re_{v(i)}m_{ji} & \\ & & & & & \\ \hline & & & & & \\ & 0 & & & 0 & \\ & & & & & \end{array}\right) \end{array} \text{ and } V = \begin{array}{c} \\ L_1 \\ \vdots \\ L_N \\ I_1 \\ \vdots \\ I_N \end{array} \begin{array}{c} \begin{array}{cccccc} \partial L_1 & \cdots & \partial L_N & \partial I_1 & \cdots & \partial I_N \end{array} \\ \left(\begin{array}{ccc|ccc} \ddots & & 0 & & & \\ & -\frac{1}{\tau} & & & 0 & \\ 0 & & \ddots & & & \\ \hline \ddots & & 0 & \ddots & & 0 \\ & \frac{1}{\tau} & & & -\frac{1}{T} & \\ 0 & & \ddots & 0 & & \ddots \end{array}\right) \end{array}.$$

The next generation matrix is then equal to:

$$K = -FV^{-1} = \left(\begin{array}{c|c} Tre_{v(i)}m_{ji} & Tre_{v(i)}m_{ji} \\ \hline 0 & 0 \end{array}\right).$$

The eigenvalues of K are 0, with multiplicity N, and the eigenvalue of the $N \times N$ submatrix $A = \left(Tre_{v(i)}m_{ji}\right)$. R_0 is thus equal to the dominant eigenvalue of A. It is computed numerically according to the landscape structure.

Appendix 4

Supporting information relative the chapter 6: Evolution of specialization in spatially heterogeneous environments

In these appendix, it is assumed that the matrix $A(x_1, x_2)$ is primitive (or positive regular) that is there exists an integer k such that $A^k(x_1, x_2)$ has all entries strictly positive for any x_1 and any x_2 in the neighbourhood of x_1. Then $A(x_1, x_2)$ has a dominant eigenvalue $\lambda_1(x_1, x_2)$ that is positive, striclty greater in magnitude than the other eigenvalues and simple (Caswell, 2001, chapter 4). $r_1(x_1, x_2)$ and $l_1(x_1, x_2)$ denote the right and left eigenvectors of the matrix $A(x_1, x_2)$ associated with the dominant eigenvalue. For simplicity all the eigenvalues and the eigenvectors of $A(x_1, x_2)$ are assumed to be real in this appendix.

Evolutionary outcomes

Singular strategy

We consider a mutant with trait x_2 in a monomorphic population with trait x_1. The fitness gradient $D(x_1, x_2)$ is equal to (Caswell, 2001, Chapter 9, eqn. 9.10):

$$D(x_1,x_2)=\frac{\partial \ln(\lambda_1(x_1,x_2))}{\partial x_2}=\frac{1}{\lambda_1(x_1,x_2)}\cdot\frac{\partial\lambda_1(x_1,x_2)}{\partial x_2}=\frac{1}{\lambda_1(x_1,x_2)}\cdot\frac{l_1'(x_1,x_2)\frac{\partial A(x_1,x_2)}{\partial x_2}r_1(x_1,x_2)}{l_1'(x_1,x_2)r_1(x_1,x_2)},$$

where prime denotes transposition. Since $[A(x_1,x_2)]_{jj'}=\frac{\mu_{j'j}\bar{K}_j}{\mu_{+j}}\cdot\frac{f_{h(j)}(x_2)}{f_{h(j)}(x_1)}$,

$$\left[\frac{\partial A(x_1,x_2)}{\partial y}\right]_{jj'}=-\frac{\mu_{j'j}\bar{K}_j}{\mu_{+j}}\cdot\frac{f_{h(j)}(x_2)}{f_{h(j)}(x_1)}\cdot\frac{x_2-\beta_{h(j)}}{\sigma^2}.$$

So,

$$l_1'(x_1,x_2)\frac{\partial A(x_1,x_2)}{\partial x_2}r_1(x_1,x_2)=l_1'(x_1,x_2)\Delta(x_2)A(x_1,x_2)r_1(x_1,x_2)$$

where $\Delta(x_2)$ denotes the diagonal matrix whose diagonal elements are $[\Delta(x_2)]_{jj}=-\frac{x_2-\beta_{h(j)}}{\sigma^2}$. Because $r_1(x_1,x_2)$ is the right eigenvector of the matrix $A(x_1,x_2)$ associated with the dominant eigenvalue $\lambda_1(x_1,x_2)$ it follows that, $l_1'(x_1,x_2)\frac{\partial A(x_1,x_2)}{\partial x_2}r_1(x_1,x_2)=l_1'(x_1,x_2)\Delta(x_2)\lambda_1(x_1,x_2)r_1(x_1,x_2)$ and

$$D(x_1,x_2)=\frac{l_1'(x_1,x_2)\Delta(x_2)r_1(x_1,x_2)}{l_1'(x_1,x_2)r_1(x_1,x_2)}.$$

$[A(x_1,x_2)]_{jj'}=\frac{\mu_{j'j}\bar{K}_j}{\mu_{+j}}$ so its dominant eigen value is 1. Since $[r_1(x_1,x_1)]_j=\bar{K}_j$ the fitness gradient taken at $x_2=x_1$, $D(x_1,x_1)$ is equal to

$$\begin{aligned}D(x_1,x_1)=&-\frac{\sum_j[l_1(x_1,x_1)]_j\bar{K}_j\frac{x_1-\beta_{h(j)}}{\sigma^2}}{\sum_j[l_1(x_1,x_1)]_j\bar{K}_j}=-\frac{x_1-x^\star}{\sigma^2},\\ x^\star=&\frac{\sum_j\bar{K}_j[l_1(x_1,x_1)]_j\beta_{h(j)}}{\sum_j\bar{K}_j[l_1(x_1,x_1)]_j}=\sum_j\frac{\bar{K}_j[l_1(x_1,x_1)]_j}{\sum_{j'}\bar{K}_{j'}[l_1(x_1,x_1)]_{j'}}\beta_{h(j)}.\end{aligned}\tag{3}$$

A strategy is singular iff the fitness gradient $D(x_1, x_1)$ is zero. Thus $x^\star$ is the singular strategy.

Stability of the singular strategy

From Equation 3, $dD(x_1, x_1)/dx = -1/\sigma^2 < 0$, thus the singular strategy $x^\star$ is always convergence stable (Geritz et al., 1998).

In order to know if the strategy $x^\star$ is evolutionary stable, we calculate $\partial^2 s(x^\star, x_2)/\partial x_2^2 \mid_{x_2=x^\star}$. We have:

$$\begin{aligned}
\frac{\partial^2 s(x^\star, x_2)}{\partial x_2^2} \mid_{x^\star} &= \frac{\partial^2 \ln(\lambda_1(x^\star, x_2))}{\partial x_2^2} \mid_{x^\star} \\
&= \left[\frac{1}{\lambda_1(x^\star, x_2)} \frac{\partial^2 \lambda_1(x^\star, x_2)}{\partial x_2^2} - \frac{1}{\lambda_1(x^\star, x_2)^2} (\frac{\partial \lambda_1(x^\star, x_2)}{\partial x_2})^2\right] \mid_{x^\star} \\
&= \left[\frac{1}{\lambda_1(x^\star, x_2)} \frac{\partial^2 \lambda_1(x^\star, x_2)}{\partial x_2^2} - D(x^\star, x_2)^2\right] \mid_{x^\star} \\
&= \frac{\partial^2 \lambda_1(x^\star, x_2)}{\partial x_2^2} \mid_{x^\star},
\end{aligned}$$

since $\lambda_1(x^\star, x^\star) = 1$ and $D(x^\star, x^\star) = 0$. Using that:

$$\frac{\partial \lambda_1(x^\star, x_2)}{\partial x_2} = \frac{l_1'(x^\star, x_2) \frac{\partial A(x^\star, x_2)}{\partial x_2} r_1(x^\star, x_2)}{l_1'(x^\star, x_2) r_1(x^\star, x_2)}$$

we have:

$$\begin{aligned}
\frac{\partial^2 s(x^\star, x_2)}{\partial x_2^2} \mid_{x^\star} = &\left[\frac{\partial \frac{1}{l_1'(x^\star, x_2) r_1(x^\star, x_2)}}{\partial x_2} l_1'(x^\star, x_2) \frac{\partial A(x^\star, x_2)}{\partial x_2} r_1(x^\star, x_2)\right. \\
&+ \frac{1}{l_1'(x^\star, x_2) r_1(x^\star, x_2)} \{\frac{\partial l_1'(x^\star, x_2)}{\partial x_2} \frac{\partial A(x^\star, x_2)}{\partial x_2} r_1(x^\star, x_2) \\
&\left. + l_1'(x^\star, x_2) \frac{\partial^2 A(x^\star, x_2)}{\partial^2 x_2} r_1(x^\star, x_2) + l_1'(x^\star, x_2) \frac{\partial A(x^\star, x_2)}{\partial x_2} \frac{\partial r_1(x^\star, x_2)}{\partial x_2}\}\right] \mid_{x^\star} \\
= &\frac{1}{l_1'(x^\star, x^\star) r_1(x^\star, x^\star)} \left[\frac{\partial l_1'(x^\star, x_2)}{\partial x_2} \frac{\partial A(x^\star, x_2)}{\partial x_2} r_1(x^\star, x_2)\right. \\
&+ l_1'(x^\star, x_2) \frac{\partial^2 A(x^\star, x_2)}{\partial^2 x_2} r_1(x^\star, x_2) \\
&\left. + l_1'(x^\star, x_2) \frac{\partial A(x^\star, x_2)}{\partial x_2} \frac{\partial r_1(x^\star, x_2)}{\partial x_2}\right] \mid_{x^\star}
\end{aligned}$$

as $l_1'(x^\star, x_2) \frac{\partial A(x^\star, x_2)}{\partial x_2} r_1(x^\star, x_2)$ is zero at $x^\star$.

First term computation The first eigenvalue of $A(x^\star, x_2)$ is $\lambda_1(x^\star, x_2)$. Let $\lambda_k(x^\star, x_2)$ be the k^{th} eigenvalue of $A(x^\star, x_2)$ $(k \geq 2)$ and $l_k(x^\star, x_2)$ and $r_k(x^\star, x_2)$ be respectively, the left

and right eigenvectors associated with $\lambda_k(x^\star, x_2)$. It is assumed that the eigenvectors have been scaled so that $l_k'(x^\star, x_2) r_k(x^\star, x_2) = 1$, $l_k'(x^\star, x_2) r_{k'}(x^\star, x_2) = 0$ for $k \neq k'$. According to Caswell (2001, chapter 9, eqn. 9.132),

$$\begin{aligned}
\frac{\partial l_1(x^\star, x_2)}{\partial x_2} &= \sum_{k,k\neq 1} \frac{\left\langle \frac{\partial A'(x^\star, x_2)}{\partial x_2} l_1(x^\star, x_2), r_k(x^\star, x_2) \right\rangle}{\lambda_1(x^\star, x_2) - \lambda_k(x^\star, x_2)} l_k(x^\star, x_2) \\
&= \sum_{k,k\neq 1} \frac{r_k(x^\star, x_2)' \left(\Delta(x_2) A(x^\star, x_2)\right)' l_1(x^\star, x_2)}{\lambda_1(x^\star, x_2) - \lambda_k(x^\star, x_2)} l_k(x^\star, x_2) \\
&= \sum_{k,k\neq 1} \frac{r_k(x^\star, x_2)' A'(x^\star, x_2) \Delta'(x_2) l_1(x^\star, x_2)}{\lambda_1(x^\star, x_2) - \lambda_k(x^\star, x_2)} l_k(x^\star, x_2) \\
&= \sum_{k,k\neq 1} \frac{\lambda_k(x^\star, x_2) r_k(x^\star, x_2)' \Delta(x_2) l_1(x^\star, x_2)}{\lambda_1(x^\star, x_2) - \lambda_k(x^\star, x_2)} l_k(x^\star, x_2) \\
&\qquad \Delta'(x_2) = \Delta(x_2) \text{ because } \Delta \text{ is diagonal.}
\end{aligned}$$

Thus,

$$\frac{\partial l_1'(x^\star, x_2)}{\partial x_2} \frac{\partial A(x^\star, x_2)}{\partial x_2} r_1(x^\star, x_2)$$

$$\begin{aligned}
&= \left(\sum_{k,k\neq 1} \frac{\lambda_k(x^\star,x_2) r_k(x^\star,x_2)' \Delta(x_2) l_1(x^\star,x_2)}{\lambda_1(x^\star,x_2) - \lambda_k(x^\star,x_2)} l_k(x^\star, x_2) \right)' \Delta(x_2) A(x^\star, x_2) r_1(x^\star, x_2) \\
&= \lambda_1(x^\star, x_2) \left(\sum_{k,k\neq 1} \frac{\lambda_k(x^\star,x_2) r_k(x^\star,x_2)' \Delta(x_2) l_1(x^\star,x_2)}{\lambda_1(x^\star,x_2) - \lambda_k(x^\star,x_2)} l_k(x^\star, x_2) \right)' \Delta(x_2) r_1(x^\star, x_2) \\
&= \lambda_1(x^\star, x_2) \left(\sum_{k,k\neq 1} \frac{\lambda_k(x^\star,x_2) l_k(x^\star,x_2)' l_1'(x^\star,x_2) \Delta(x_2) r_k(x^\star,x_2)}{\lambda_1(x^\star,x_2) - \lambda_k(x^\star,x_2)} \right) \Delta(x_2) r_1(x^\star, x_2) \\
&= \lambda_1(x^\star, x_2) l_1'(x^\star, x_2) \Delta(x_2) \left(\sum_{k,k\neq 1} \frac{\lambda_k(x^\star,x_2) r_k(x^\star,x_2) l_k(x^\star,x_2)'}{\lambda_1(x^\star,x_2) - \lambda_k(x^\star,x_2)} \right) \Delta(x_2) r_1(x^\star, x_2).
\end{aligned}$$

Second term computation Let $diag(\frac{1}{\sigma^2})$ be the diagonal matrix whose diagonal elements are $\left[diag(\frac{1}{\sigma^2})\right]_{jj} = \frac{1}{\sigma^2}$. Since $\frac{\partial A(x^\star, x_2)}{\partial x_2} = \Delta(x_2) A(x^\star, x_2)$, we have:

$$\begin{aligned}
\frac{\partial^2 A(x^\star, x_2)}{\partial^2 x_2} &= \frac{d\Delta(x_2)}{dx_2} A(x^\star, x_2) + \Delta(x_2) \frac{\partial A(x^\star, x_2)}{\partial x_2} \\
&= \left[\Delta(x_2)^2 - diag(\frac{1}{\sigma^2}) \right] A(x^\star, x_2)
\end{aligned}$$

So that,

$$\begin{aligned}
l_1'(x^\star, x_2) \frac{\partial^2 A(x^\star, x_2)}{\partial^2 x_2} r_1(x^\star, x_2) &= l_1'(x^\star, x_2) \left[\Delta(x_2)^2 - diag(\frac{1}{\sigma^2}) \right] A(x^\star, x_2) r_1(x^\star, x_2) \\
&= \lambda_1(x^\star, x_2) l_1'(x^\star, x_2) \left[\Delta(x_2)^2 - diag(\frac{1}{\sigma^2}) \right] r_1(x^\star, x_2).
\end{aligned}$$

Third term computation As for the first term, according to Caswell (2001, chapter 9, eqn. 9.131),

$$\begin{aligned}\frac{\partial r_1(x^\star,x_2)}{\partial x_2} &= \sum_{k,k\neq 1}\frac{\left\langle \frac{\partial A(x^\star,x_2)}{\partial x_2} r_1(x^\star,x_2), l_k(x^\star,x_2)\right\rangle}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)} r_k(x^\star,x_2)\\ &= \sum_{k,k\neq 1}\frac{l_k(x^\star,x_2)'\Delta(x_2)A(x^\star,x_2)r_1(x^\star,x_2)}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)} r_k(x^\star,x_2)\\ &= \sum_{k,k\neq 1}\frac{\lambda_1(x^\star,x_2)l_k(x^\star,x_2)'\Delta(x_2)r_1(x^\star,x_2)}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)} r_k(x^\star,x_2).\end{aligned}$$

Thus,

$$l_1'(x^\star,x_2)\frac{\partial A(x^\star,x_2)}{\partial x_2}\frac{\partial r_1(x^\star,x_2)}{\partial x_2}$$

$$\begin{aligned}&= l_1'(x^\star,x_2)\Delta(x_2)A(x^\star,x_2)\textstyle\sum_{k,k\neq 1}\frac{\lambda_1(x^\star,x_2)l_k(x^\star,x_2)'\Delta(x_2)r_1(x^\star,x_2)}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)} r_k(x^\star,x_2)\\ &= \lambda_1(x^\star,x_2)l_1'(x^\star,x_2)\Delta(x_2)\textstyle\sum_{k,k\neq 1} A(x^\star,x_2)r_k(x^\star,x_2)\frac{l_k(x^\star,x_2)'\Delta(x_2)r_1(x^\star,x_2)}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)}\\ &= \lambda_1(x^\star,x_2)l_1'(x^\star,x_2)\Delta(x_2)\left(\textstyle\sum_{k,k\neq 1}\frac{\lambda_k(x^\star,x_2)r_k(x^\star,x_2)l_k(x^\star,x_2)'}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)}\right)\Delta(x_2)r_1(x^\star,x_2).\end{aligned}$$

Finally we have,

$$\frac{\partial^2 s(x^\star,x_2)}{\partial x_2^2}\,|_{x^\star}$$

$$\begin{aligned}= \frac{1}{l_1'(x^\star,x^\star)r_1(x^\star,x^\star)}&[\lambda_1(x^\star,x_2)l_1'(x^\star,x_2)\Delta(x_2)\left(\textstyle\sum_{k,k\neq 1}\frac{\lambda_k(x^\star,x_2)r_k(x^\star,x_2)l_k(x^\star,x_2)'}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)}\right)\Delta(x_2)r_1(x^\star,x_2)\\ &+\lambda_1(x^\star,x_2)l_1'(x^\star,x_2)\left[\Delta(x_2)^2-diag(\tfrac{1}{\sigma^2})\right]r_1(x^\star,x_2)\\ &+\lambda_1(x^\star,x_2)l_1'(x^\star,x_2)\Delta(x_2)\left(\textstyle\sum_{k,k\neq 1}\frac{\lambda_k(x^\star,x_2)r_k(x^\star,x_2)l_k(x^\star,x_2)'}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)}\right)\Delta(x_2)r_1(x^\star,x_2)]\,|_{x^\star}\\ = \frac{1}{l_1'(x^\star,x^\star)r_1(x^\star,x^\star)}&\times\\ \Big[l_1'(x^\star,x_2)\Delta(x_2)&\left(2\textstyle\sum_{k,k\neq 1}\frac{\lambda_k(x^\star,x_2)r_k(x^\star,x_2)l_k(x^\star,x_2)'}{\lambda_1(x^\star,x_2)-\lambda_k(x^\star,x_2)}+Id_P\right)\Delta(x_2)r_1(x^\star,x_2)\Big]\,|_{x^\star}-\tfrac{1}{\sigma^2},\end{aligned}$$

since $\lambda_1(x^\star,x^\star)=1$. **$\Delta$ in this appendix is equal to $\frac{1}{\sigma}\Delta$ in the main text.**

Canonical equation

We consider the evolution of the quantitative trait x when the population is monomorphic before x reaches $x^\star$. The probability that a new individual is mutant is $u_N\theta(x)$. Population size N is assumed to be large and mutations are assumed to be rare with $Nu_N\to 0$ when

$N \to \infty$ (Champagnat et al., 2006). The timescale is changed to a continuous one in which one unit is equal to $1/(Nu_N)$ generations. In this new timescale, mutations occur according to a continuous-time Poisson process with rate $\theta(x)$.

Mutational steps are assumed to be small so that the trait of a mutant is denoted by $x_2 = x + \varepsilon z$ where ε is small. The quantity z is assumed to be distributed according to the mutation distribution $M(x, z)$ that is assumed to be symmetrical with respect to 0 and hence centered. The timescale is changed a second time so that one unit in the new timescale is equal to ε^{-2} units in the previous timescale. Then the trait of the resident follows the equation (Champagnat et al., 2006; Durinx et al., 2008):

$$\dot{x} = \gamma^2(x)\theta(x)\frac{\partial p_x(x_2)}{\partial x_2} \mid_x /2,$$

where $p_x(x_2)$ is the mutant's survival probability, $\dot{x} = dx/dt$ and $\gamma^2(x) = \int z^2 M(x, z)dz$ is the variance of the mutation distribution. The mutant can survive only if $s_x(x_2) > 0$. When $s(x, x_2) > 0$, the extinction of the mutant population mainly occurs at the begining of the population growth when the population is still small. The demography of this population is modelled with a slightly supercritical multi-type branching process with a reproduction mean matrix equal to $A(x, x_2)$ (Champagnat et al., 2006; Durinx et al., 2008; Haccou et al., 2005, section 5.6). Then the mutant's survival probability is found to be approximately equal to:

$$p_x(x_2) \approx 2\ln(\lambda_x(x_2))/\tau^2.$$

In this equation the parameter τ^2 is equal to:

$$\tau^2 = \sum_i [r_x(x)]_i \mathrm{Var}[\sum_j [l_x(x)]_j \xi_{ij}],$$

where ξ_{ij} is the random number of descendants in patch j of a resident in patch i and it is assumed that the eigenvectors have been scaled so that $l'_x(x)r_x(x) = \sum_i [r_x(x)]_i = 1$. It follows that the canonical equation is equal to:

$$\dot{x} = -\gamma^2(x)\theta(x)(x - x^\star)/(\sigma^2\tau^2).$$

Only τ^2 can depend on the landscape structure through ξ_{ij}. A simple way in order to introduce stochasticity in the mutant demography consists in supposing, from equation (6.3) of the main text, that:

$$\xi_{ij} \sim \mathrm{B}\Big(K_j, \frac{m_{ij}}{m_{+j}K_T}\Big).$$

In this case, the spatial repartition of habitat does not affect ξ_{ij} because the fitness functions are not involved.

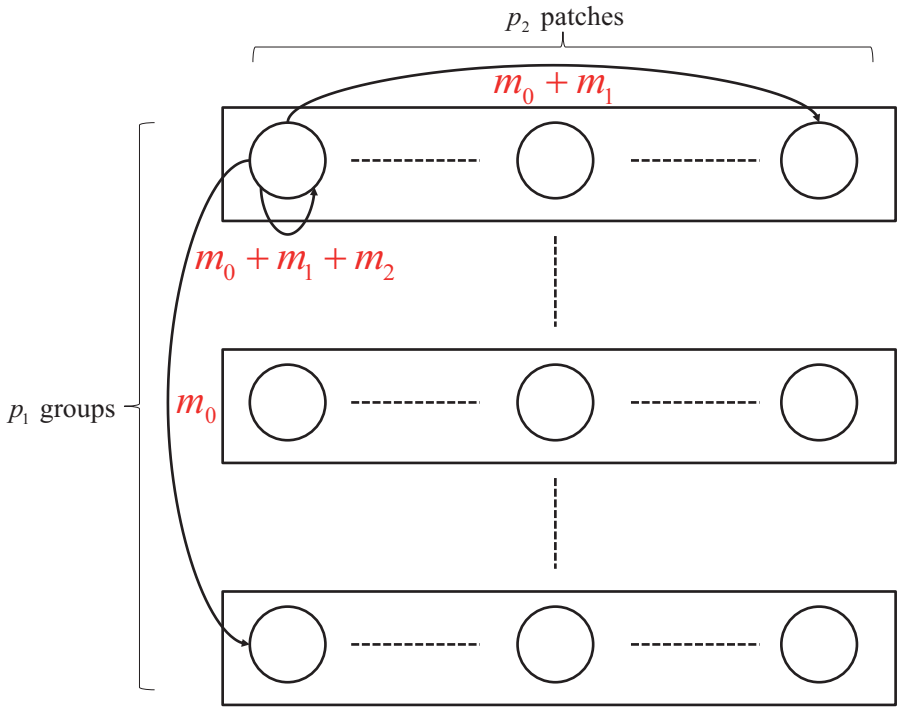


Figure 3: The hierarchical metapopulation structure. $p_1 \times p_2$ patches are dispatched among p_1 groups of p_2 patches. m_0 is the dispersal rate between patches that belongs to different groups. $m_0 + m_1$ is the dispersal rate between patches that belongs to the same group. $m_0 + m_1 + m_2$ is the intra-patch dispersal rate. Intra patch dispersal is greater than intra-group dispersal which is itself greater than inter-group dispersal.

Hierarchical metapopulation

Let us consider a metapopulation composed of p_1 groups of p_2 patches (Figure 3).

Three dispersal rates, m_0, m_1 and m_2, are defined as:

- $m_0 + m_1 + m_2$ is the probability that a propagule remains in its patch of origin,
- $m_0 + m_1$ is the probability that a propagule lands in another given patch of the same group,
- m_0 is the probability that a propagule lands in a given patch of another group.

Since all the patches have the same relative size $\bar{K}_j$ and dispersal is symmetric, $A(x, x) = M$. M is the dispersal matrix and is equal to:

$$M = m_0 \, J_{P,P} + m_1 \, Idp_1 \otimes J_{p_2,p_2} + m_2 \, IdP,$$

where $J_{P,P}$ is the $P \times P$ matrix of 1 and $Idp_1 \otimes J_{p_2,p_2}$ denotes the block-diagonal matrix with diagonal matrices J_{p_2,p_2}. More generally, $\otimes$ denotes the tensor product between matrices.

As $A(x,x)$ is symmetric it can be decomposed as $A(x,x) = \sum_{i=1}^{P} \lambda_i(x,x) r_i(x,x) r_i'(x,x)$ where $\lambda_i(x,x)$ and $r_i(x,x)$ denote the i^{th} eigenvalue and the i^{th} scaled eigenvector respectively. Thus, following this decomposition we have

$$A(x,x) = (m_2 + p_2 m_1 + p_1 p_2 m_0) S_1 + (m_2 + p_2 m_1) S_2 + m_2 S_3,$$

where $S_1 = \frac{1}{p_1} J_{p_1} \otimes \frac{1}{p_2} J_{p_2}$, $S_2 = (I_{p_1} - \frac{1}{p_1} J_{p_1}) \otimes \frac{1}{p_2} J_{p_2}$, $S_3 = I_{p_1} \otimes (I_{p_2} - \frac{1}{p_2} J_{p_2})$, I_{p_2} and I_{p_1} are the indentity matrices of size p_2 and p_1, J_{p_2} and J_{p_1} are the square matrices of ones of size p_1 and p_2. Note that $\lambda_1(x,x) = m_2 + p_2 m_1 + p_1 p_2 m_0 = 1$, $r_1(x,x) = 1_{p_2 p_1} / \sqrt{p_1 p_2}$ where $1_{p_1 p_2}$ is the vector of ones of length $p_1 p_2$. The second-order derivative of fitness can be written as:

$$\frac{\partial^2 s(x^\star, x_2)}{\partial x_2^2} \mid_{x^\star}$$

$$\begin{aligned} &= \left[l_1'(x^\star, x_2) \Delta(x_2) \left(2 \sum_{k,k \neq 1} \frac{\lambda_k(x^\star,x_2) r_k(x^\star,x_2) l_k(x^\star,x_2)'}{\lambda_1(x^\star,x_2) - \lambda_k(x^\star,x_2)} + I_{p_1 p_2} \right) \Delta(x_2) r_1(x^\star, x_2) \right] \mid_{x^\star} - \frac{1}{\sigma^2} \\ &= \left[l_1'(x^\star, x_2) \Delta(x_2) \left(2\rho S_2 + 2\xi S_3 + I_{p_1 p_2} \right) \Delta(x_2) r_1(x^\star, x_2) \right] \mid_{x^\star} - \frac{1}{\sigma^2}, \end{aligned}$$

where $\rho = \dfrac{m_2 + p_2 m_1}{p_1 p_2 m_0}$ and $\xi = \dfrac{m_2}{p_2 m_1 + p_1 p_2 m_0}$. Moreover,

$$\Delta(x_2) r_1(x^\star, x_2) = (l_1'(x^\star, x_2) \Delta(x_2))' = \frac{1}{\sqrt{p_1 p_2}} X \Theta(x_2),$$

where X is the $p_1 p_2 \times H$ matrix with elements equal to 0 or 1 (where H is the total number of different habitats),

$$X = \begin{bmatrix} \frac{X_1}{X_2} \\ \hline \vdots \\ \hline X_{p_1} \end{bmatrix},$$

where X_g is the design matrix of habitats in group g. $\Theta(x_2)$ is the vector of length H equal to

$$\Theta(x_2) = \begin{bmatrix} \vdots \\ -\frac{x_2 - \beta_{h(i)}}{\sigma^2} \\ \vdots \end{bmatrix}.$$

Let us define the vector $R = \sum_{g=1}^{p_1} R_g$ of the total number of patches bearing each habitat (R_g is vector of the number of patches bearing each habitat in group g). Moreover, we define the vector of the mean number of patches bearing each habitat as $\bar{R} = \frac{1}{p_1} R$. We have $\rho = \xi + \nu$, where $\nu = \dfrac{m_1}{p_1 m_0} \left(\dfrac{m_2}{p_2 m_1 + p_1 p_2 m_0} + 1 \right) \geq 0$. It follows that:

$$\frac{\partial^2 s(x^\star, x_2)}{\partial x_2^2} \mid_{x^\star}$$

$$
\begin{aligned}
&= \tfrac{2}{p_1p_2}\Theta'(x_2)\left[(\xi+\nu)X'S_2X+\xi X'S_3X\right]\Theta(x_2)\mid_{x^\star}+\Theta'(x_2)X'I_{p_1p_2}X\Theta(x_2)/(p_1p_2)\mid_{x^\star}-\tfrac{1}{\sigma^2}\\
&= \tfrac{2}{p_1p_2}\Theta'(x_2)\left[\xi X'(S_2+S_3)X+\nu X'S_2X\right]\Theta(x_2)\mid_{x^\star}+\Theta'(x_2)B\Theta(x_2)/(p_1p_2)\mid_{x^\star}-\tfrac{1}{\sigma^2}
\end{aligned}
$$

where B is the diagonal matrix with diagonal elements given by R. As

$$
\begin{aligned}
S_2+S_3 &= (I_{p_1}-\frac{1}{p_1}J_{p_1})\otimes\frac{1}{p_2}J_{p_2}+I_{p_1}\otimes(I_{p_2}-\frac{1}{p_2}J_{p_2})\\
&= I_{p_1}\otimes\frac{1}{p_2}J_{p_2}-\frac{1}{p_1}J_{p_1}\otimes\frac{1}{p_2}J_{p_2}+I_{p_1}\otimes I_{p_2}-I_{p_1}\otimes\frac{1}{p_2}J_{p_2}\\
&= I_{p_1p_2}-\frac{1}{p_1p_2}J_{p_1p_2}
\end{aligned}
$$

and $\Theta'(x^\star)X'J_{p_1p_2}X\Theta(x^\star)=0$ we have:

$$
\frac{\partial^2 s(x^\star,x_2)}{\partial x_2^2}\mid_{x^\star}
$$

$$
\begin{aligned}
&= \tfrac{2}{p_1p_2}\Theta'(x_2)\left[\xi X'(I_{p_1p_2}-\tfrac{1}{p_1p_2}J_{p_1p_2})X+\nu C\right]\Theta(x_2)\mid_{x^\star}+\Theta'(x_2)B\Theta(x_2)/(p_1p_2)\mid_{x^\star}-\tfrac{1}{\sigma^2}\\
&= \tfrac{1}{p_1p_2}\Theta'(x_2)\left[(1+2\xi)B+2\nu C\right]\Theta(x_2)\mid_{x^\star}-\tfrac{1}{\sigma^2}
\end{aligned}
$$

where $C=X'S_2X=\frac{1}{p_2}\sum_{g=1}^{p_1}(R_g-\bar{R})(R_g-\bar{R})'$. It follows that $x^\star$ is an ESS if:

$$
(1+2\xi)\sum_{k=1}^{H}\pi_k\frac{(x^\star-\beta_k)^2}{\sigma^2}+2\nu\sum_{k=1}^{H}\sum_{k'=1}^{H}v_{kk'}\frac{(x^\star-\beta_k)(x^\star-\beta_k')}{\sigma^2}<1,
$$

where $v_{kk'}=\sum_{g=1}^{p_1}(\pi_{kg}-\pi_k)(\pi_{k'g}-\pi_{k'})/p_1$ and π_{kg} is the frequency of habitat k in group g. When dispersal is homogeneous ($m_2=m_1=0$), the condition for evolutionary stability becomes:

$$
\sum_{k=1}^{H}\pi_k\frac{(x^\star-\beta_k)^2}{\sigma^2}<1.
$$

When dispersal is hindered so that a propagule is more likely to stay in its patch than to move to another patch ($m_2\neq 0$) and when there is no group structure ($m_1=0$), the condition for evolutionary stability becomes:

$$
(1+2\xi)\sum_{k=1}^{H}\pi_k\frac{(x^\star-\beta_k)^2}{\sigma^2}<1.
$$

Appendix 5

Hétérogénéités temporelle et spatio-temporelle dans un modèle de dynamique adaptative

This work was realized by Florian Claeys (ENS Ulm) during an internship suppervising by Julien Papaïx, Hervé Monod and Olivier David.

École Normale Supérieure
Département de Biologie

Hétérogénéités temporelle et spatio-temporelle dans un modèle de dynamique adaptative

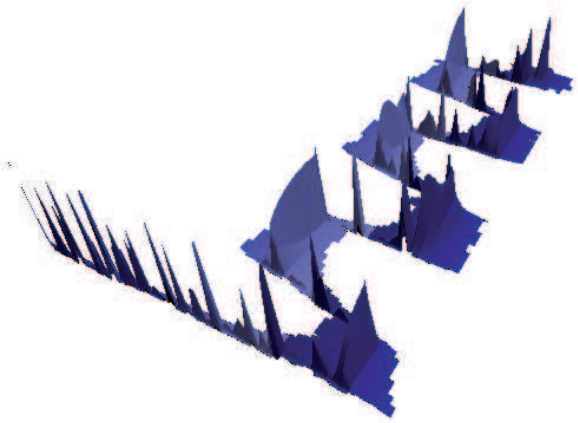
Florian Claeys

21 juillet 2011

Rapport de Stage de Master 1
7 mars 2011 – 7 juillet 2011
Centre INRA de Jouy–en–Josas
Unité de Mathématiques et d'Informatique Appliquées

sous la direction d'Hervé Monod[1] et de Julien Papaïx[1][2]

(1) INRA, Mathématiques et Informatique Appliquées, Domaine de Vilvert, F-78352 Jouy-en-Josas, France
(2) BIOlogie et GEstion du Risque en agriculture, avenue Lucion Brétigère, F-78850 Thiverval-Grignon, France



Résumé

Les pratiques modernes de l'agriculture et notamment la céréaliculture intensive, ont permis d'augmenter considérablement la production de nourriture pour une population humaine en forte croissance. Cependant, ces agroécosystèmes sont des zones à fort risque épidémique : leur grande homogénéité favorise la sélection de pathogènes spécialisés et agressifs. Une stratégie pour éviter cette spécialisation est de réintroduire de l'hétérogénéité dans ces milieux. Ainsi, en modifiant les structures spatio-temporelles de l'environnement, il est possible de manipuler les pressions de sélection et d'influencer l'adaptation des pathogènes.
La dynamique adaptative est une théorie reposant sur un puissant ensemble de techniques d'étude de l'évolution phénotypique et en particulier des spécialisations. Elle réutilise notamment le concept de fréquence-dépendance de la théorie des jeux, associé aux notions de valeur sélective et de mutation, pour mettre en évidence des stratégies évolutives optimales. Certaines de ces dernières, dites stratégies de convergence stable, peuvent conduire à un branchement évolutif à partir duquel une population initialement monomorphe devient dimorphique ou à une stratégie évolutivement stable, qui une fois établie, empêche tout mutant de se fixer. Les caractéristiques environnementales vont fortement influencer le niveau d'adaptation et le comportement évolutif des populations.
Dans une perspective de modélisation à l'échelle du paysage agricole, un modèle de dynamique de métapopulation couplé à un modèle de dynamique du paysage a été développé permettant d'étudier l'influence de différents paramètres de l'hétérogénéité sur la dynamique de l'adaptation. Indépendamment de l'hétérogénéité spatiale, l'hétérogénéité temporelle peut, elle-aussi, influencer l'adaptation des pathogènes. Ainsi, sur un paysage agricole monovariétal, des changements variétaux saisonniers rapprochés dans le temps permettent de favoriser les stratégies généralistes aux stratégies spécialistes et de limiter la valeur sélective des pathogènes spécialisés. Pour aborder l'hétérogénéité spatio-temporelle, la dynamique de paysage a consisté en la variation saisonnière des proportions de deux variétés sur un paysage agricole parcellaire. La comparaison des observations de ces simulations aux résultats théoriques déduits de l'influence de l'hétérogénéité spatiale montre un ont été réalisées, où les proportions de deux variétés sur un paysage agricole varient à chaque saison. La comparaison des observations aux résultats théoriques issus de l'étude de l'hétérogénéité spatiale montre que l'ajout de la dimension temporelle peut favoriser une évolution vers le dimorphisme, pour des paysages avec des habitats très différenciés et très agrégés dans le temps et/ou dans l'espace, ou au contraire, favoriser une évolution vers le monomorphisme pour des habitats peu différenciés et/ou soumis à des changements spatio-temporels fréquents.

Mots-clés Dynamique adaptative, hétérogénéité, spécialisation, pathogènes, branchement évolutif

1 Introduction

Face à l'accroissement démographique de la population mondiale et à l'augmentation des besoins en nourriture de l'élevage, les productions agricoles devront considérablement être augmentées pour assurer la sécurité alimentaire de 2050. L'extension des zones cultivées ou le recours accru aux pesticides et aux engrais ne sauraient constituer des réponses pertinentes à ce défi, car c'est une agriculture durable qui doit être promue alliant les hauts rendements aux faibles impacts environnementaux. Un des éléments primordiaux de cette « révolution doublement verte » (Griffon 2006) est l'amélioration de la gestion des épidémies végétales.

Les pratiques de lutte contre les épidémies végétales ne sont encore que faiblement efficaces. L'efficacité est définie comme le pourcentage de pertes de production agricole évitées et est calculée à partir des pertes actuelles observées et des pertes potentielles estimées (Oerke *et al.* 1994). Selon Oerke et Dehne (2004), en matière de lutte contre les pathogènes bactériens et fongiques, l'efficacité atteindrait 33,8 % et tombe à 12,9 % contre les virus. Les raisons de ce manque d'efficacité sont à chercher dans la nature même des agroécosystèmes modernes (Stukenbrock et McDonald 2008). Face à l'hétérogénéité, la complexité et la diversité des écosystèmes naturels, les agroécosystèmes modernes sont des systèmes simplifiés à l'extrême, cumulant de nombreux facteurs de risque épidémique : la généralisation de vastes étendues de monocultures génétiquement homogènes (Robinson et Sutherland 2002) a facilité l'adaptation des pathogènes à leurs hôtes (Stukenbrock et McDonald 2008). Par exemple, depuis la moitié du XX^{e}siècle, la diffusion successive de nouvelles variétés de blés, chacune possédant un nouveau gène de résistance a entraîné une pression de sélection graduelle ayant pour principale conséquence une évolution « guidée par la main de l'homme » des populations de rouille noire du blé (*Puccinia germinis*), où de nouvelles populations virulentes apparaissaient en réponse à l'introduction de nouveaux gènes de résistances (Johnson 1961).

L'homme, à travers la modification de l'environnement, influence l'évolution des pathogènes. Le contrôle des épidémies végétales doit donc inclure la gestion de l'évolution des pathogènes. Comme dans toute intégration de principes évolutifs dans un domaine appliqué, le concept-clé est celui de décalage entre le phénotype actuel des organismes et le phénotype optimal (Hendry *et al.* 2011). Si ce décalage est faible, les populations sont bien adaptées à leur environnement et en conséquence possèderont des abondances élevées. Inversement, un décalage fort signifie une mauvaise adaptation, et peut entraîner un déclin des populations. En épidémiologie végétale, l'objectif est d'accroître le plus possible ce décalage, et de diminuer le niveau d'adaptation des pathogènes, ou de ralentir les spécialisations (Papaïx *et al.* 2011b).

Une des stratégies développées dans cet objectif est d'introduire de l'hétérogénéité écologique dans les exploitations agricoles. En effet l'hétérogénéité, en générant une diversité des habitats variétaux va jouer sur des compromis écologiques liés à l'utilisation des ressources : une capacité accrue à exploiter un habitat entraîne une moindre capacité d'en exloiter un autre (Kassen 2002, Gravel *et al.* 2011). L'objectif de cette étude est d'aborder les questions de recherche portant sur l'influence de l'hétérogénéité temporelle sur la dynamique évolutive des pathogènes et plus précisément, l'influence de la fréquence des changements variétaux saisonniers sur la spécialisation des pathogènes. L'influence de l'hétérogénéité spatio-temporelle sera abordée dans le cas de changements saisonniers de la proportion de chaque variété. Les résultats présentés proviennent de l'analyse d'un modèle de dynamique adaptative, une théorie permettant d'étudier l'évolution de traits continus soumis à une sélection fréquence-dépendance sous des scénarios écologiques donnés.

Après quelques éléments de définition sur la notion d'hétérogénéité écologique, quelques principes fondamentaux de la théorie de la dynamique adaptative seront expliqués. La description du modèle employé s'attachera à détailler l'articulation entre dynamique du paysage et dynamique de populations. Les résultats concernant l'influence des hétérogénéités temporelle et spatio-temporelle sur la spécialisation seront présentés, puis discutés.

2 Quelques éléments du cadre scientifique d'étude

2.1 L'hétérogénéité, aspects spatiauxs et temporels

Le mot «hétérogénéité » provient de deux radicaux grec, ετερος « autre, différence » et γένος « origine » et est employé dans le langage courant pour désigner ce qui est de différentes natures ou pour qualifier une distribution répartie inégalement ou un mélange dans lequel on peut distinguer au moins deux constituants à l'œil nu (Académie Française 1932-1935).
Que ce soit « la constitution de plusieurs éléments différents les uns des autres » (Art 1995), « le fait d'avoir une structure ou une composition non uniforme » (Lincoln *et al.* 1998) ou encore, « la complexité résultant de l'interaction entre la distribution spatiale des contraintes environnementales et les réponses différentielles des organismes à ces contraintes » (Milne 1991), la plupart des définitions de l'hétérogénéité en écologie insistent sur les discontinuités spatiales et temporelles sans vraiment répondre à la question de la nature de l'hétérogénéité. Là où Kolasa *et al.* (1991) distingue une douzaine de composantes de l'hétérogénéité, certains ont opté pour une définition opérationnelle.
L'hétérogénéité désignerait ainsi la complexité ou la variabilité d'une propriété du système étudié dans le temps et/ou dans l'espace (Li et Reynolds 1994; 1995), selon que l'on se réfère à des descripteurs qualitatifs et catégoriques (complexité) ou quantitatifs et numériques (variabilité). Cette définition opérationnelle permet de distinguer une hétérogénéité structurelle d'une hétérogénéité fonctionnelle, selon que l'on tienne compte ou non des effets de la complexité et de la variabilité des structures sur les fonctions écologiques. Autre avantage de cette définition, elle permet d'asseoir une base conceptuelle à une hétérogénéité que l'on déclinerait en hétérogénéités spatiale, temporelle et spatio-temporelle (1).

La notion d'hétérogénéité est fortement dépendante de l'échelle d'étude à commencer par le grain, défini comme la résolution la plus fine des données et l'envergure, définie comme la durée totale ou l'aire maximale de définition de la propriété étudiée. Le grain et l'envergure sont les premiers facteurs affectant l'hétérogénéité. Il existe deux approches pour quantifier l'hétérogénéité : la première, directe, passe par la mesure de la complexité (composition et configuration) et de la variabilité (tendance, corrélation, anisotropie). Wiens (2000) reprend cette vision de l'hétérogénéité spatiale en en distinguant quatre variances - agrégation, motifs, composition, localisation -, dans une vision de l'espace de plus en plus explicite.

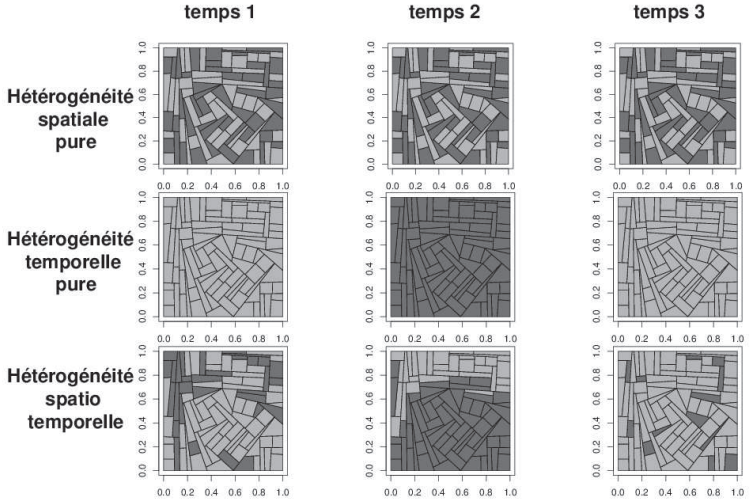


FIGURE 1 – Hétérogénéités spatiale, temporelle et spatio-temporelle d'un paysage agricole

De la même manière que pour l'hétérogénéité spatiale, l'hétérogénéité temporelle peut être décrite sous plusieurs aspects : l'échelle considérée (à associer à la longévité des habitats), le contraste, la fréquence de changement, l'agrégation temporelle et la prédictabilité au cours du temps (Stuefer 1996). L'espace et le temps peuvent être substitués l'un à l'autre dans de nombreuses situations écologiques : la coexistence des espèces en fonction de la séparation des ressources (Giller 1984), les stratégies de recherche de nourriture, les successions écologiques (Pickett et Likens 1989). Un paysage peut être discrétisé dans le temps de manière semblable à la discrétisation spatiale en parcelles ou *patchs* (Turner *et al.* 2001). Les interactions entre « parcelles temporelles » deviennent des décalages temporels ou des héritages, avec des effets frontières ou des effets barrières (Strayer *et al.* 2003). Une différence fondamentale subsiste tout de même : l'espace permet des interactions directionnelles de type « aller » et « retour », ce qui est impossible dans le temps Wiens (2000).
L'échelle, du grain et de l'envergure de la variabilité temporelle déterminent grandement l'hétérogénéité temporelle : une hétérogénéité à petite échelle ne pourra générer que des réponses phénotypiques rapides essentiellement physiologiques (plasticité physiologique de Hutchings et de Kroon (1994)), tandis que que les phénomènes évolutifs résulteront de variations temporelles à grande échelle. Des considérations théoriques et des simulations

mathématiques ont montré que l'hétérogénéité temporelle pouvait contraindre fortement les réponses adaptatives des organismes (Oborny 1994) : la nature adaptative de la réponse de croissance à un habitat donné dépend largement de la prédictabilité temporelle, définie à partir de l'entropie de Shannon en théorie de l'information (Juhász-Nagy et Podani 1983). Pour des ressources pulsatiles, ces réponses peuvent être des stratégies d' « attente les bras croisés » , par exemple) (Hutchings et de Kroon 1994).

2.2 Dynamique adaptative, éléments fondamentaux et singularités évolutives

Les différentes composantes spatiales et temporelles de l'hétérogénéité, en contraignant les réponses adaptatives des organismes offrent autant de moyens de contrôler l'adaptation des pathogènes en jouant sur les structures spatiales et temporelles des paysages agricoles. La manipulation de l'hétérogénéité entre dans le cadre plus large de l'intégration dans les agrosystèmes de principes écologiques et évolutifs régissant les écosystèmes naturels, élément stratégique pour relever les défis de l'agriculture de demain (Tilman 1999, Thrall *et al.* 2011). La dynamique adaptative permet d'étudier le devenir évolutif d'une population et, en complémentarité des nouvelles techniques de séquençages à haut-débit, permettant d'analyser rapidement la composition génétique des communautés de pathogènes (Garrett *et al.* 2006), s'inscrit dans une approche intégrative, « éco-évo-génomique », de la gestion des structures spatio-temporelles du paysage agricole.

Fondamentaux de la dynamique adaptative La dynamique adaptative est un ensemble de techniques introduit par Hofbauer et Sigmund (1990) et Nowak et Sigmund (1990) pour comprendre les conséquences évolutives des variations phénotypiques introduites par de légères mutations dans une population dite résidente. Cette approche, liant à la dynamique de populations une dynamique évolutive, s'appuie sur le concept de fréquence-dépendance, issu de la théorie des jeux. La sélection correspond donc au modèle de « sélection douce » (Levene 1953, Wallace 1975).
La dynamique adaptative repose sur deux idées fondamentales (Brännström et Festenberg 2006). Selon la première, la population résidente, monomorphe, est supposée être dans un équilibre dynamique lorsque les mutants apparaissent. Selon la seconde, le devenir des mutants, soit l'invasion et le remplacement de la population résidente, soit le non-maintien et la disparition, peut être déduit du taux de croissance initiale, appelée exposant d'invasion (Diekmann 2003).
Dans la plupart des modèles de dynamique adaptative, et cette étude ne fait pas exception, la valeur sélective des individus, définie comme le taux de croissance à long-terme d'un phénotype dans un environnement donné (Metz *et al.* 1992), est caractérisée par un seul trait phénotypique continu. La valeur sélective $s_x(x)$ de la population résidente, supposée monomorphe de phénotype x, est nulle pour tout x : la population résidente est supposée en équilibre dynamique (Geritz *et al.* 1998).
Dans cette population résidente, des mutants vont apparaître. Du fait de sa rareté, le mutant ne va pas influer immédiatement sur le phénotype de la population résidente, de sorte que la valeur sélective $s_x(y)$ d'un mutant de phénotype y ne va dépendre que de y et de x. Le signe de $s_x(y)$ indique la capacité du mutant à envahir ou non la population résidente. La fréquence-dépendance de la valeur sélective implique que le paysage d'aptitude, correspondant à l'ensemble des phénotypes possibles, à leurs degrés de similitudes et à leur valeur sélective respective, est un paysage mouvant, changeant à chaque invasion.
Le gradient de sélection $\left.\frac{\partial s_x(y)}{\partial y}\right|_{y=x}$ détermine les phénotypes mutants pouvant envahir la stratégie singulière x : si le gradient est positif (respectivement négatif), des mutants avec une valeur de trait supérieure à x (respectivement inférieure à x) pourront remplacer la population résidente.

Singularités évolutives Une stratégie singulière correspond à la valeur phénotypique $x^\star$ pour laquelle le gradient de sélection est nul, signifiant pour un mutant que le paysage d'aptitude est localement plat (Brännström et Festenberg 2006). La stabilité de convergence et la stabilité évolutive des stratégies singulières déterminent le devenir évolutif des populations (Geritz *et al.* 1998). Les stratégies évolutivement stables (ESS) sont des stratégies singulières telles que, si elles sont adoptées par la plupart des individus d'une population, aucune stratégie mutante ne pourra être envahissante (Smith et Price 1973). Les ESS correspondent à des pièges évolutifs, où aucun changement évolutif n'est possible sans changement de l'environnement. Mathématiquement, une ESS vérifie :

$$\left.\frac{\partial^2 s_x(y)}{\partial y^2}\right|_{x=y=x^\star} < 0$$

Une singularité évolutive stable par convergence (CSS) se comporte comme un attracteur (Christiansen 1991) : toute population, quelque soit son phénotype, va évoluer vers cette singularité. Mathématiquement, une CSS vérifie :

$$\left.\frac{\partial^2 s_y(x)}{\partial x^2}\right|_{x=y=x^\star} > \left.\frac{\partial^2 s_x(y)}{\partial y^2}\right|_{x=y=x^\star}$$

Certaines singularités peuvent être stables par convergence et évolutivement instables. Localement, la sélection est disruptive et peut correspondre à un point de branchement évolutif (EBP), où une population monomorphe devient dimorphique. Si $s_{x^\star}(m)$ et $s_m(x^\star)$ sont tous les deux positifs, la population va pouvoir se diviser en deux sous-populations et le dimorphisme sera stabilisé. Mathématiquement cette condition peut s'écrire :

$$\left.\frac{\partial^2 s_x(y)}{\partial y^2}\right|_{x=y=x^\star} > -\left.\frac{\partial^2 s_y(x)}{\partial x^2}\right|_{x=y=x^\star}$$

3 Description du modèle employé

Le modèle employé est basé sur le modèle de Papaïx *et al.* (2011a), dont il reprend les éléments d'hétérogénéité spatiale. Il s'agit d'un modèle de dynamique adaptive modélisant l'évolution d'un trait phénotypique d'une métapopulation répartie sur un paysage composé d'un nombre fini de parcelles monovariétales. Dans ce modèle, la dispersion est considérée comme un processus passif : il n'y a pas de choix de l'habitat par les pathogènes.

3.1 Métapopulation, paysage et phénotypes

La population modélisée est supposée être constituée d'individus haploïdes se reproduisant de manière asexuée et sans chevauchement des générations. La valeur sélective de ces individus est caractérisée par un trait phénotypique, de valeur x comprise entre -5 et 5. Cette gamme phénotypique est subdivisée en I intervalles, correspondant à autant de phénotypes.
Cette métapopulation occupe un espace subdivisé en P parcelles, correspondant à des unités élémentaires spatiales. Pour éviter les effets de bords, cet espace est un tore. Chaque unité est caractérisée par la présence d'un seul type d'habitat pouvant varier dans le temps, correspondant à une variété et assimilé à une niche écologique, et par sa capacité de charge $K_{j,j\in[\![1,P]\!]}$, constante dans le temps.
Le modèle calcule les effectifs de chaque phénotype des populations de chaque parcelle et à chaque pas de temps jusqu'à un temps final T. On note $n(i,j,t)$ la taille de la population de phénotype i occupant la parcelle j au temps t et $h(j,t)$ l'habitat de la parcelle j au temps t.
La taille totale de la population est constante :

$$\forall t \in \{1,T\}, \sum_{i=1}^{I}\sum_{j=1}^{P} n(i,j,t) = \sum_{j=1}^{P} K_j$$

On définit la taille relative de la parcelle j par $K_j^* = \frac{K_j}{\sum_{j=1}^{P} K_j}$

3.2 Dynamique de paysage

3.2.1 Habitats et paysage spatio-temporel

L'ensemble des habitats $h(j,t)_{j\in[\![1,P]\!]t\in[\![1,T]\!]}$ constitue un paysage spatio-temporel caractérisé en chaque instant par les proportions moyennes et l'agrégation des habitat. L'agrégation de l'habitat h varie entre 0, lorsqu'il n'y a aucune paire de parcelles voisines avec ce même habitat h, et 1, lorsque le nombre de paires de parcelles voisines recevant l'habitat h est maximal, compte tenu des proportions entre habitats (He *et al.* 2000).
Soit H le nombre d'habitats différents possibles. À chaque pas de temps, le paysage spatial peut être caractérisé par le nombre d'habitats, la proportion et l'agrégation de chaque habitat.
La présente étude ne porte que sur des paysages à deux types d'habitats, d'indices 1 et 2. Le paysage spatio-temporel, noté $\mathcal{P}_{ET}$, est caractérisé par $\pi(t) \in [0,1]$, la proportion de parcelles avec l'habitat 1 au temps t et $\alpha(t) \in [0,1]$ l'agrégation de l'habitat minoritaire au temps t.

3.2.2 Création de paysages spatiaux paramétrés

Les dimensions spatiales des paysages spatio-temporels sont au nombre de deux et forment une aire modélisée par un treillis régulier de 900 parcelles carrées et contiguës (Papaïx *et al.* 2011a). Pour éviter les effets de bords dans le calcul des taux de dispersion et d'agrégation des habitats, cette aire est considérée comme étant la surface d'un tore (Griffith 1983). En répartissant les habitats 1 et 2 sur les 900 parcelles, et en contrôlant leur proportion et leur agrégation à l'aide d'un algorithme de recuit simulé (Kirkpatrick *et al.* 1983), une banque de paysages spatiaux a pu être constituée.

3.2.3 Stratégie d'étude de la dimension temporelle des paysages

La dimension temporelle est discrétisée en saisons. Une saison correspond à un intervalle de temps pendant lequel les composantes spatiales du paysage restent inchangées. Au sein d'une simulation, il y a S saisons de durée constante égale à $t_S = \frac{T}{S}$. L'ensemble des paysages spatio-temporels peut s'écrire $\mathcal{P}_{ET}(\pi(s), \alpha(s), d, t_S)$ où $s \in [\![1, S]\!]$ est la s–ième saison, $\pi(s)$ la proportion moyenne, $\alpha(s)$ l'agrégation de l'habitat 1 pendant la saison s et d la différenciation des habitats (section §3.3.3).
Dans le cas de l'hétérogénéité temporelle pure, le paysage est spatialement homogène : à un temps donné, toutes les parcelles possèdent le même habitat de sorte que le paysage est assimilé à une seule parcelle. Chaque saison voit le basculement d'un habitat à l'autre. On a donc

$$\forall s \in [\![1, S]\!], \alpha(s) = 1$$

$$\forall s \in [\![1, S]\!], \pi(s) = 1$$

L'ensemble des paysages temporels, considéré comme un sous-ensemble de $\mathcal{P}_{ET}$, est alors noté $\mathcal{P}_T(d, t_s)$ avec d, la différenciation des habitats et t_s la durée des saisons.

3.3 Cycle de vie

Le cycle de vie est constitué de quatre étapes : après une première phase de reproduction avec mutation, s'ensuit une phase de dispersion, puis une phase de sélection avant une phase de régulation.

3.3.1 Reproduction avec mutation

Pour chaque sous-population, le taux de reproduction est supposé constant quelque soit le phénotype ou les caractéristiques de la parcelle occupée. Pour une population d'un phénotype donné sur une parcelle donnée, un mutant apparaît dès que l'effectif de cette population franchit un certain seuil. En notant $x_r, r \in [\![1, I]\!]$ la valeur de trait de la population d'où provient le mutant, $x_m, m \in [\![1, I]\!]$ la valeur du trait du mutant provient de la catégorisation du trait phénotypique x'_m. $x'_m \in \mathbb{R}$ suit une loi normale de moyenne x_r et de variance $0,1$:

$$x'_m \sim N(x_r; 0, 1)$$

3.3.2 Dispersion

Durant la phase de dispersion, chaque parcelle reçoit une certaine proportion des propagules produites par les autres parcelles. En notant $m_{j'j}$ la proportion de propagules produites par la parcelle j' et arrivant dans la parcelle j, et $p(i, j, t)$ la quantité de propagules de phénotypes x_i arrivant sur la parcelle j au temps t, on a :

$$\forall j \in [\![1, P]\!], \forall i \in [\![1, I]\!], \forall t \in [\![1, T]\!], \qquad p(i, j, t) = \sum_{j'=1}^{P} m_{j'j}(t) n(i, j', t).$$

La dispersion est supposée isotropique et décroissante exponentiellement avec la distance. Par ailleurs, on suppose que la population se mélange parfaitement dans chaque parcelle. On peut alors écrire :

$$m_{j'j} = \int_{\mathcal{A}_1} \int_{\mathcal{A}_2} \frac{2\pi}{m_s^2} \exp\left(\frac{2\pi}{m_s^2} |z_{j'} - z_j|\right) \mathrm{d}z_{j'} \cdot \mathrm{d}z_j$$

avec $\mathcal{A}_1$ et $\mathcal{A}_2$ l'ensemble des points des parcelles j' et j respetivement, $|z_{j'} - z_j|$ la distance entre les points $z_{j'}$ de la parcelle j' et z_j de la parcelle j et m_s est la distance de dispersion moyenne. Notre étude se limite au cas où $m_s = 15\%$ de la longueur du côté du treillis.
L'ensemble des $m_{j'j}$ a été calculé en utilisant l'algorithme Califlopp (Bouvier *et al.* 2009), lequel permet de calculer entre un polygone source et un polygone cible des intégrales de fonctions de dispersion point à point.

3.3.3 Sélection

À chaque pas de temps, les individus de chacune des parcelles sont sélectionnés selon leur phénotype et le type d'habitat présent, mais indépendamment des effectifs. Le taux de survie des individus de phénotype i occupant l'habitat $h(j,t)$ de la parcelle j au temps t est égal à $f(x_i, h(j,t))$ où f est définie par :

$$f(x,h) = \exp\left(-\frac{(x-\beta_h)^2}{2\sigma^2}\right)$$

avec, β_h la valeur de trait optimale de l'habitat h et σ^2 la variance de trait autour de cette valeur optimale (Geritz *et al.* 1998). Des différences de valeur optimale entre les habitats génèrent un compromis écologique entre les survies sur chaque habitat, d'autant plus fort que les habitats sont différenciés : l'adaptation à un habitat entraîne obligatoirement une mal-adaptation aux autres habitats. Dans le cas de deux habitats d'indices 1 et 2 aux valeurs optimales opposées $\delta = \beta_1 = -\beta_2$, les taux de survie sont $f(x,1) = e^{-\frac{(x-\delta)^2}{2\sigma^2}}$ et $f(x,2) = e^{-\frac{(x+\delta)^2}{2\sigma^2}}$. Le compromis écologique entre les taux de survie est la fonction $\mathcal{C}(y), y \in [0,1]$ vérifiant :

$$\forall x, f(x,2) = \mathcal{C}\left(f(x,1)\right)$$

D'après Débarre et Gandon (2010), $u(y)$ peut s'écrire :

$$\mathcal{C}(y) = \exp\left[-2\left(\frac{\delta}{\sigma} - \sqrt{-\frac{1}{2}\ln y}\right)^2\right].$$

La force du compromis est déterminée par la différenciation des habitats $d = \frac{\delta}{\sigma}$. Pour $d < 1$, le compromis est faible ; pour $d > 1$, le compromis est fort. Dans le cas limite $d = 1$, le compromis est très sensible aux valeurs de trait x.

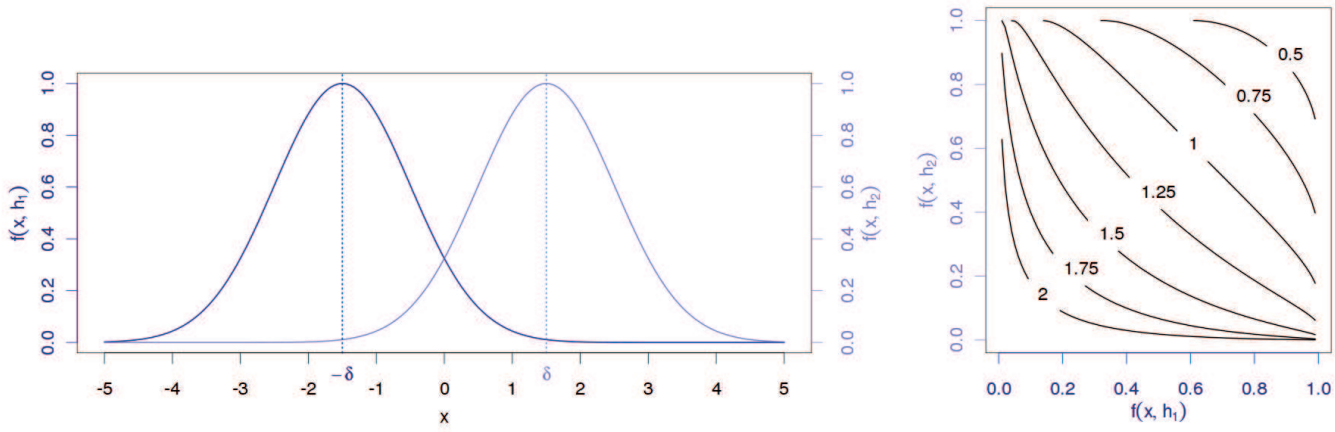


FIGURE 2 – Taux de survie et compromis. Pour chaque habitat h, la valeur sélective suit une gaussienne centrée sur sa valeur sélective optimale β_h, définissant ainsi un compromis écologique entre les survies sur chaque habitat. Dans le cas de deux habitats avec $\delta = \beta_1 = -\beta_2$, la différenciation des habitats permet de moduler la force du compromis écologique des survies (Débarre et Gandon 2010).

3.3.4 Régulation

Afin de respecter la condition d'une taille totale de population constante au cours du temps, une phase de régulation est appliquée aux différentes sous-populations.

$$\forall\{k,j,t\} \in [\![1,I]\!] \times [\![1,P]\!] \times [\![1,T-1]\!], \qquad n(i,j,t+1) = K_j \times \frac{p(k,j,t)f(k,h(j,t))}{\sum_{i=1}^{I} p(i,j,t)f(x_i,h(j,t))}$$

.

3.4 Critère de branchement

Supposons qu'au temps t_m un individu mutant de phénotype x_m apparaisse dans une population résidente de phénotype x_r. Peu de temps après son apparition, au temps t, cette population continue d'être supposée monomorphique. Le nombre de mutants de la parcelle j au temps $t+1$, $n(m,j,t+1)$ est égal à :

$$n(m,j,t+1) = K_j \frac{\sum_{j'=1}^{P} m_{j'j} n(r,j',t)}{\sum_{j'=1}^{P} m_{j'j} K_{j'})} \frac{f\left(x_m, h(j)\right)}{f\left(x_r, h(j)\right)}$$

En notant $N_i(t)$ le vecteur d'élément générique correspondant à la taille de population du phénotype i sur la parcelle j, la taille de population mutante vérifie l'égalité matricielle (Caswell 2001) :

$$N_m(t+1) = A(x_r, x_m) N_m(t)$$

où $A(x_1, x_2) \in \mathcal{M}_{P,P}(\mathbb{R})$ est la matrice d'élément générique :

$$[A(x_1,x_2)]_{j,j'} = \frac{m_{j'j} K_j}{\sum_{j'=1}^{P} m_{j'j} K_{j'}} \frac{f\left(x_2, h(j)\right)}{f\left(x_1, h(j)\right)}$$

En reprenant les notations introduites précédemment dans la présentation de la théorie de la dynamique adaptative, la stratégie singulière $x^\star$ est une stratégie évolutive stable si et seulement si :

$${}^t g^{(1)}\left(x^\star, x^\star\right) \cdot \Delta(x^\star) \cdot \left(\mathrm{Id}_P + 2 \sum_{j=2}^{P} \frac{\lambda^{(j)}\left(x^\star, x^\star\right) \cdot d^{(j)}\left(x^\star, x^\star\right) \cdot \ {}^t g^{(j)}\left(x^\star, x^\star\right)}{\lambda^{(1)}\left(x^\star, x^\star\right) - \lambda^{(j)}\left(x^\star, x^\star\right)} \right) \cdot \Delta(x^\star) \cdot r^{(j)}\left(x^\star, x^\star\right) < 1.$$

Dans cette équation, Id_P est la matrice identité d'ordre P, $\Delta(x^\star)$ est la matrice diagonale d'élément générique $[\Delta(x)]_i = -\frac{x-\beta_{h_i}}{\sigma}$. Les valeurs propres de $A\left(x^\star, x^\star\right)$ sont notés, dans l'ordre croissant $\lambda^{(j)}\left(x^\star, x^\star\right), j \in [\![1, P]\!]$. Les vecteurs propres normalisés gauche et droite correspondant à ces valeurs propres sont notés $g^{(j)}\left(x^\star, x^\star\right), j \in [\![1, P]\!]$ et $d^{(j)}\left(x^\star, x^\star\right), j \in [\![1, P]\!]$, respectivement. Par « vecteur normalisé », on entend que :

$${}^t g^{(j)}\left(x^\star, x^\star\right) \cdot r^{(j)}\left(x^\star, x^\star\right) = 1 \quad \text{et} \quad \forall j \neq j', \ {}^t g^{(j)}\left(x^\star, x^\star\right) \cdot r^{(j)}\left(x^\star, x^\star\right) = 0$$

La partie gauche de cette inéquation correspond à un critère de branchement. Celui-ci dépend de la structure parcellaire et des phénotypes optimaux des habitats : au-delà de 1, ce critère indique que les paramètres spatiaux permettent un branchement évolutif ; en-deçà, la stratégie singulière est une ESS.
La comparaisons des prédictions théoriques déduites de ce critère de branchement aux résultats provenant des résultats de simulation portant sur l'hétérogénéité spatiale montrent une très bonne concordance : plus de 90 % des simulations montrent un branchement évolutif lorsque le critère de branchement est supérieur à 1 (Papaïx *et al.* 2011a).

4 Résultats

4.1 Hétérogénéité temporelle

Le jeu de simulations contient 400 simulations et explore les deux paramètres utilisés pour la définition des paysages temporels. La durée des saisons t_S varie entre 25 et 500 pas de temps. Enfin, $\sigma^2 = 1$ de sorte que $d = \delta$: la différenciation des habitats est assimilée à la valeur absolue des stratégies optimales. La différenciation des habitats varie entre $0,2$ et 2.

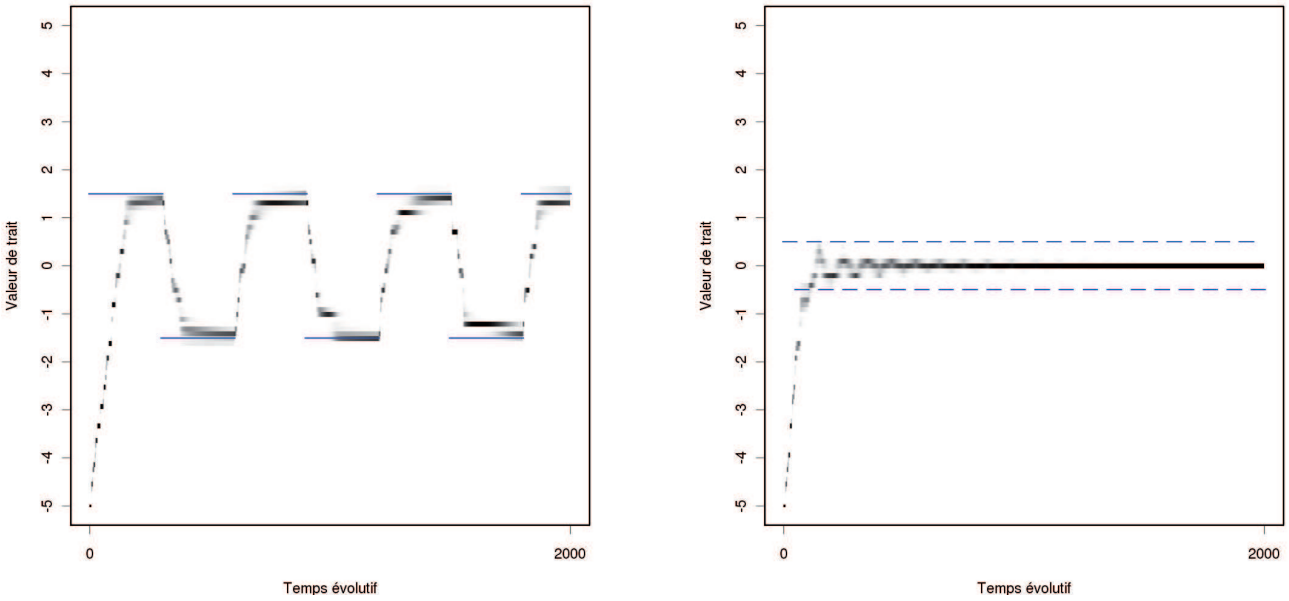


FIGURE 3 – Deux stratégies évolutives : à gauche, une stratégie spécialiste ($\delta = 1.5$, $t_S = 300$) ; à droite une stratégie généraliste ($\delta = 0.5$, $t_S = 50$). Avec, en ordonnée la valeur de trait phénotypique, en abscisse, le temps évolutif et en bleu, la valeur de trait optimale δ au cours du temps.

4.1.1 Comportements évolutifs observés

La valeur de trait moyenne de la population varie entre δ et $-\delta$ en fonction des changements saisonniers d'habitats : le trait de la population évolue vers δ (respectivement $-\delta$) lorsque le paysage contient l'habitat 1 (respectivement 2). On obtient ainsi des oscillations, de période égale à la durée des saisons.
Deux comportements évolutifs peuvent être distingués (figure 3) : dans le premier cas, les oscillations se maintiennent tout au long de la simulation, ce qui correspond à une population qui se spécialise à chaque saison ; dans le second cas, les oscillations s'amortissent jusqu'à s'éteindre, ce qui est la marque d'une population adoptant une stratégie généraliste.
L'étude de l'influence de l'hétérogénéité temporelle sur ces deux stratégies passe par l'analyse des phénotypes présents en fin de simulation. Plus exactement, on s'intéresse à la valeur absolue moyenne du phénotype de la population. La symétrie des traits optimaux des habitats implique qu'une population adoptant la stratégie spécialiste verra, saison après saison, son phénotype moyen évoluer vers des valeurs opposées (δ pendant les saisons où l'habitat est occupé par la variété 1 et $-\delta$ pour les saisons où l'habitat est occupé par la variété 2). La valeur absolue varie entre 0 et δ : une population adoptant une stratégie généraliste verra son phénotype moyen converger vers 0, tandis qu'une population adoptant une stratégie spécialiste optimale verra son phénotype moyen converger vers δ.
Prendre la valeur absolue moyenne du phénotype en fin de simulation est un outil commode pour apprécier le niveau de spécialisation de cette population, indépendamment du signe du trait optimal et tout en restant sensible aux stratégies généralistes.

4.1.2 Étude du phénotype final

On appellera « phénotype final » la valeur absolue moyenne du phénotype vers laquelle tend la population en fin de saison, que l'on notera ϕ_f. Celui-ci ne peut excéder la valeur de la stratégie optimale : $\phi_f = 0$ correspond à la stratégie généraliste, $\phi_f = \delta$ correspond à la stratégie optimale.
La représentation graphique bidimensionnelle du phénotype final en fonction de la différenciation des habitats et de la durée des saisons, montre que l'évolution vers une spécialisation ou au contraire l'émergence d'un généraliste dépend fortement de ces paramètres (figure 4.1.1).
La figure 4.1.1 montre une distinction entre les deux stratégies évolutives généraliste et spécialiste en fonction des paramètres δ et t_s. Globalement, la stratégie généraliste est d'autant plus favorisée que la durée des saisons est courte et que la différenciation des habitats est faible. Inversement, plus la différenciation des habitats est élevée et plus la durée des saisons est longue, plus le spécialisme est fort avec des niveaux d'adaptation élevés.

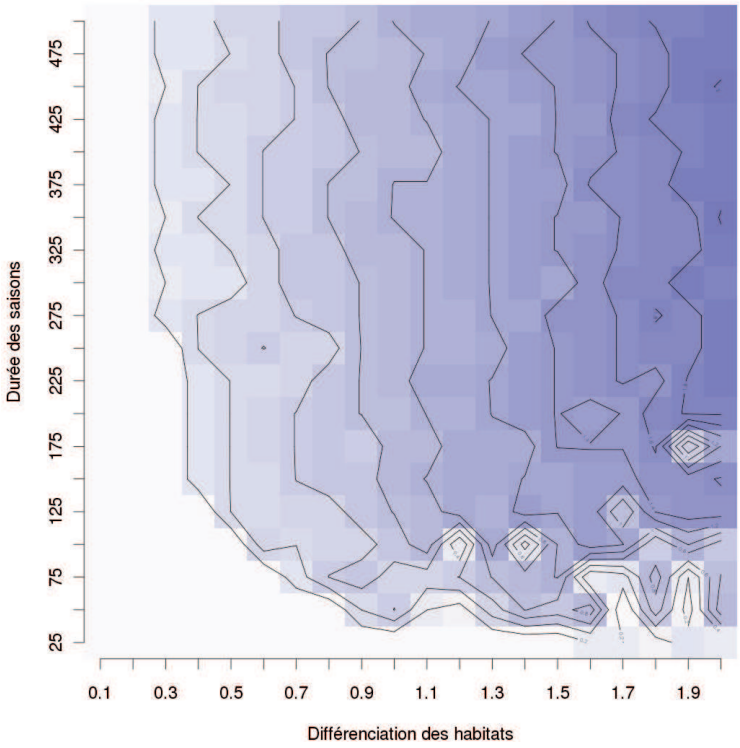


FIGURE 4 – Représentation graphique du phénotype final en fonction de la durée des saisons (en ordonnée) et de la différenciation des habitats (en abscisse). La valeur du phénotype final est proportionnelle à l'intensité de la couleur grise, avec en blanc la stratégie généraliste.

4.1.3 Influence de la différenciation des habitats sur le phénotype final

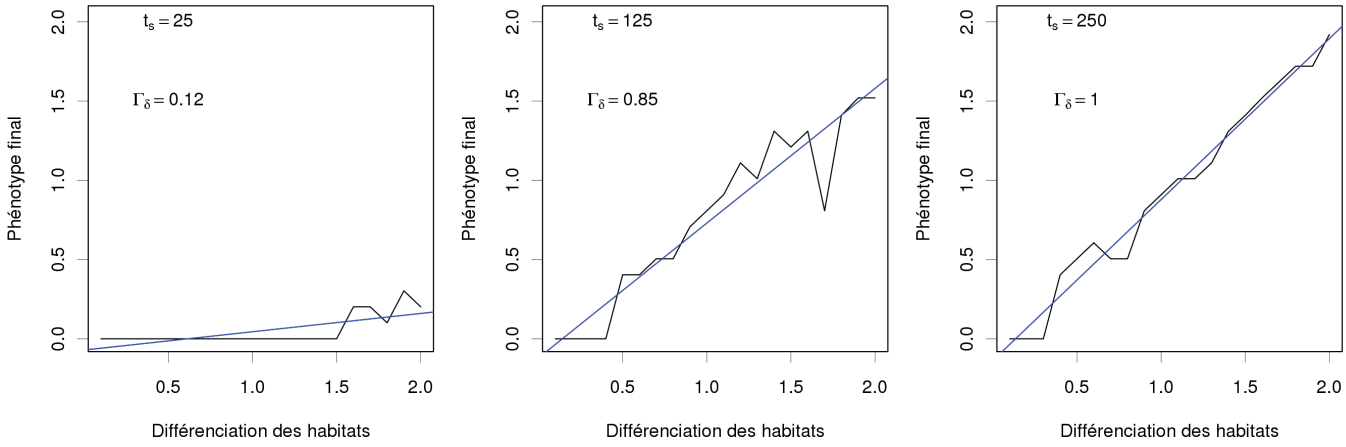


FIGURE 5 – ϕ_f en fonction de δ pour trois durées de saison fixées : de gauche à droite, $t_s = 25$, $t_s = 125$,$t_s = 250$. En bleu, figurent les coefficients de corrélation, corrélation dont la *p–value* est inférieure à 0.001 pour chaque cas.

Pour chaque valeur de t_s supérieure ou égale à 25 pas de temps, il existe une corrélation positive entre la différenciation des habitats et le phénotype final (figure 5). Ces corrélations sont significatives, d'après les résultats de tests de coefficient de corrélation linéaire de Pearson, avec un risque de première erreur inférieure à $0,01$. Pour les valeurs de t_s inférieures à 25, le phénotype final correspond, dans tous les cas, à un généraliste : pour des durées de saisons courtes, ou, autrement dit, des fréquences élevées de changements variétaux, la spécialisation est impossible.
La relation de proportionnalité entre ϕ_f et δ dépend de la durée des saisons t_S : lorsque la durée des saisons est élevée, le coefficient de proportionnalité Γ_δ tend vers 1 (figures 5 et 4.1.3).

$$\lim_{\substack{t_s \to T \\ T \to \infty}} \Gamma_\delta = 1$$

La fréquence des changements variétaux est un aspect de l'hétérogénéité temporelle qui va limiter la spécialisation des populations. Au delà d'une certaine durée de saisons, le phénotype final atteint la stratégie optimale.

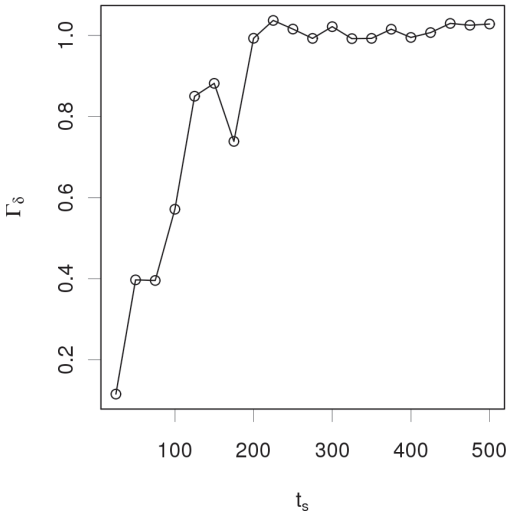


FIGURE 6 – Γ_δ en fonction de t_s. Pour des saisons de durée supérieure à 200, la différenciation des habitats détermine complètement le phénotype final.

4.1.4 Influence de la durée des saisons sur le phénotype final

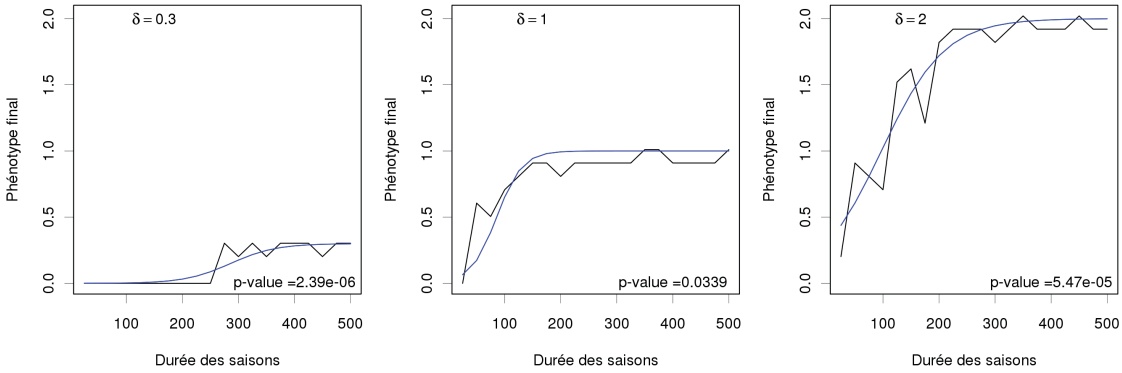


FIGURE 7 – $\phi_f = f(t_s)$ pour trois valeurs de différenciation des habitats : de gauche à droite, $\delta = 0,3$, $\delta = 1$,$\delta = 2$. En bleu, figurent les courbes de corrélation selon un modèle logistique de Verhulst. La *p–value*, indiquée pour chaque cas, provient d'un test de corrélation linéaire effectué entre $\ln\left(\frac{\delta}{\phi_f} - 1\right)$ et t_s.

Pour $\delta < 0,3$, les habitats ne sont pas suffisamment différenciés pour permettre des spécialisations : le phénotype final est nul, indiquant une évolution vers un généralisme. Pour $\delta \geq 0,3$, on peut observer une spécialisation. Cette spécialisation atteint la stratégie optimale ($\phi_f = \delta$) pour des durées de saisons suffisamment longue, ce qui justifie l'emploi pour l'interprétation de la corrélation entre la durée des saisons et le phénotype final du modèle logistique de Verhulst suivant, avec $\rho > 0$, $\tau > 0$ et $\|\epsilon_1\| \ll 1$:

$$\phi_f(\delta, t_s) = \frac{\delta}{1 + e^{-\rho \cdot (t_s - \tau}} + \epsilon_1$$

où τ correspond à la durée de saison nécessaire pour que le phénotype final atteigne la moitié du phénotype optimal ; ρ est un paramètre de croissance permettant de rendre compte de la rapidité à laquelle l'influence de δ sur ϕ_f va s'effectuer et $\|\epsilon_1\|$ est lié à l'erreur du modèle statistique. On vérifie aisément que cette équation vérifie la condition de proportionnalité entre ϕ_f et δ, à valeur de t_s fixée, observée précédemment : pour des

durées de saisons longues, le phénotype final correspond au phénotype optimal.
Une transformation logarithmique permet de linéariser cette relation :

$$\ln\left(\frac{\delta}{\phi_f} - 1\right) = \rho \cdot \tau - \rho \cdot t_s + \epsilon_2$$

avec $\|\epsilon_2\| \ll 1$, l'erreur statistique. Un test de corrélation linéaire de Pearson, effectué sur la forme linéaire du modèle de Verhulst employé, confirme la validité du modèle pour chaque cas.
Le modèle logistique dépend de la différenciation des habitats (figure 7). Si le paramètre ρ variant entre $0,015$ et $0,08$, pour une valeur moyenne de $0,039$, semble peu dépendant de δ, il n'en va pas de même pour τ.

La relation entre τ et δ n'est pas monotone et on observe une inversion de tendance pour $\delta \approx 1$ (figure 4.1.4). Pour $\delta \leq 1$, plus δ est faible, plus τ est grand signifiant que sous des habitats faiblement différenciés, le phénotype final est très sensible aux fréquences des changements d'habitat. Une certaine durée de saisons, d'autant plus élevée que δ est faible, est nécessaire pour que ϕ_f atteigne la stratégie optimale. Autrement dit, plus les habitats sont différenciés, plus les populations peuvent atteindre un haut niveau d'adaptation dans un paysage temporel hétérogène.
Cette tendance n'est pas immuable et pour $d \geq 1$, la relation entre τ et δ devient positive. Une explication possible est que plus δ est élevé, plus la distance phénotypique à parcourir à chaque fin de saison entre les deux valeurs optimales est grande : 2δ. La vitesse évolutive peut dépendre de la différenciation des habitats et de leurs valeurs optimales, de sorte que l'évolution peut aller « plus vite et plus loin » : une population peut mettre moins de temps à parcourir une plus grande distance phénotypique (Papaïx *et al.* 2011a). Ici, ce n'est pas le cas : pour des valeurs élevées de δ, le temps nécessaire à passer d'une stratégie optimale à la stratégie optimale opposée est d'autant plus important que la distance entre ces deux stratégies est grande, offrant ainsi une sensibilité accrue à l'hétérogénéité temporelle.

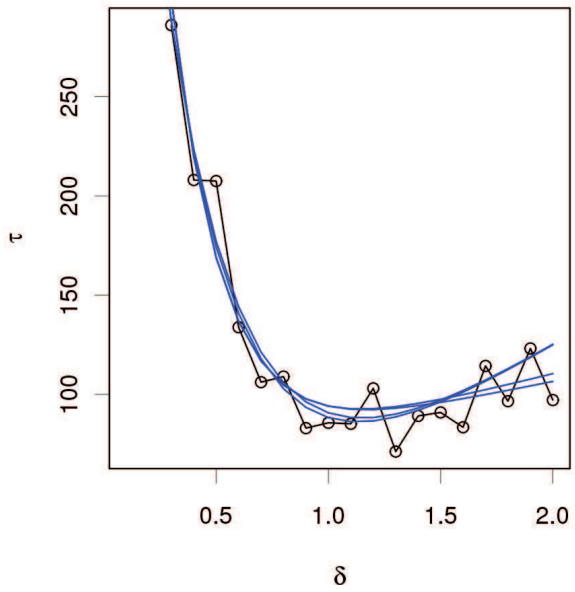


FIGURE 8 – τ en fonction de δ. Les courbes en bleu sont des courbes d'interprétation provient de l'ajustement d'une fonction du type $\tau = a \cdot e^{-b\tau} + c \cdot \tau + \delta$, par la méthode des moindres carrés.

4.2 Hétérogénéité spatio-temporelle

En reprenant la notation introduite dans la description du modèle proposé où $\mathcal{P}_{ET}(\pi(s), \alpha(s), d, t_S)$ désigne l'ensemble des paysages spatio-temporels, notre étude de l'hétérogénéité spatio-temporelle s'est limitée aux paysages vérifiant :

$$\forall s \in [\![1, S]\!], \qquad \alpha(s) = \alpha$$

$$\forall s \in [\![1, S]\!], \qquad \pi(s) = \begin{cases} \pi_1 & \text{si} \quad s \equiv 1 \pmod 2 \\ \pi_2 & \text{si} \quad s \equiv 0 \pmod 2 \end{cases}$$

$$\pi_1 \geq \pi_2$$

Autrement dit, la dynamique de paysage consiste en une alternance saisonnière de deux proportions d'habitat 1, à agrégation et différenciation d'habitats fixées. Comme précédemment, la différenciation des habitats est assimilée à la valeur absolue des stratégies optimales : $\sigma^2 = 1$.
Le jeu de simulations contient 540 simulations et explore les différents paramètres de la manière suivante :

- les valeurs possibles d'agrégation de l'habitat minoritaire sont $0,2$; $0,4$ et $0,7$;
- les valeurs possibles de la différenciation des habitats sont $0,8$; 1 et $1,2$, illustrant ainsi les trois types de compromis écologiques
- les durées des saisons peuvent être de 20, 50, 100 ou 200
- les proportions en habitat 1 peuvent être $0,04$; $0,25$; $0,49$; $0,75$ et $0,96$.

4.2.1 Trois comportements évolutifs

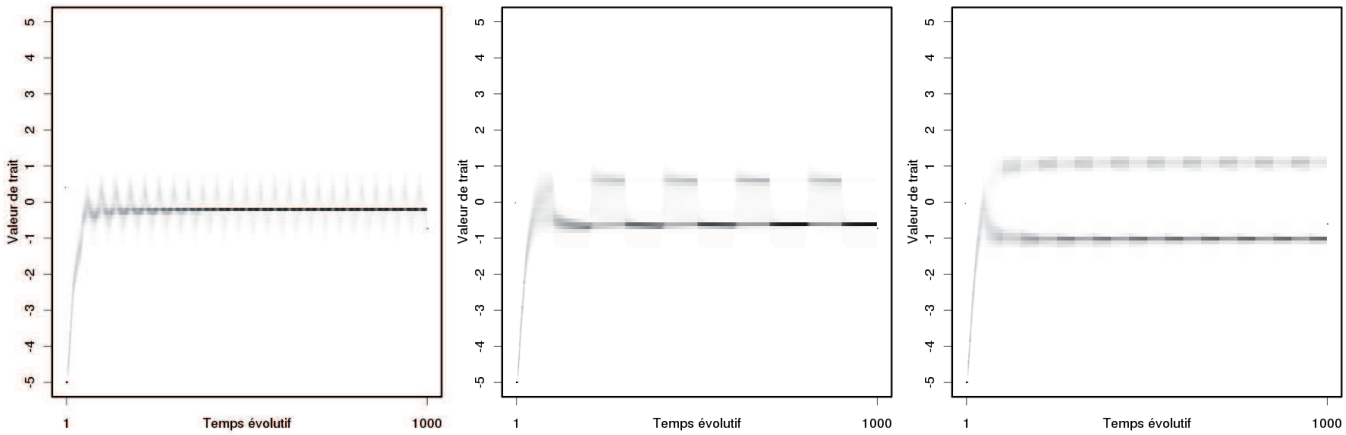


FIGURE 9 – Trois comportements évolutifs : à gauche, un monomorphisme à toutes les saisons ($\delta = 0,8$; $t_s = 20$; $\alpha = 0,2$;$\pi_1 = 0,96$; $\pi_2 = 0,75$) ; au milieu, alternance entre des saisons de monomorphisme et des saisons de dimorphismes ($\delta = 0,8$; $t_s = 100$; $\alpha = 0,7$; $\pi_1 = 0,49$; $\pi_2 = 0,04$) ; à droite, un dimorphisme à toutes les saisons ($\delta = 1,2$; $t_s = 50$; $\alpha = 0,2$; $\pi_1 = 0,49$; $\pi_2 = 0,25$). Avec, en ordonnée la valeur de trait phénotypique, en abscisse, le temps évolutif.

Trois comportements évolutifs peuvent être distingués (figure 9). Dans le premier cas, la population reste monomorphe et on observe les deux stratégies mises en évidence précédemment : le phénotype moyen de la population suit des oscillations de période égale à la durée des saisons et évolue soit vers une stratégie généraliste, soit vers une stratégie spécialiste. Dans le second cas, on observe l'alternance saisonnière entre une population monomorphe et une population dimorphique. Dans le troisième cas, la population est dimorphique à toutes les saisons (exceptée la première saison).
Ces deux derniers comportements évolutifs permettent de mettre en évidence des branchements évolutifs, rendus possibles par la présence des deux habitats à toutes les saisons. Ces branchements évolutifs conduisent à la formation de deux sous-populations développant deux stratégies de spécialisation. Les caractéristiques spatiales de l'environnement semblent déterminer la stabilité du dimorphisme engendré. Par ailleurs, le monomorphisme est généralement associé à l'adoption d'une stratégie généraliste tandis que le dimorphisme correspond plutôt à populations bien spécialisées sur l'un ou l'autre des habitats. L'alternance entre les deux est bien souvent un cas intermédiaire avec des populations moins spécialisées pendant les saisons de dimorphisme.

4.2.2 Comparaison des comportements observés aux résultats théoriques

La figure 10 montre le comportement évolutif pour chaque simulation (graphiques de gauche) et le comportement théorique déduit des critères de branchement spatiaux appliqués à chaque saison (graphiques de droite). La caractérisation des différents comportements évolutifs s'appuie sur une analyse des phénotypes dominants de la population pour chaque pas de temps de la seconde moitié du temps de simulation. Un phénotype dominant correspond au phénotype moyen d'un groupe de classes phénotypiques successives et dont les effectifs représentent plus de 2 % de la population totale. Selon le nombre de phénotypes dominants ainsi caractérisés, 1 ou 2, la population est considérée monomorphique ou dimorphique.
L'hétérogénéité spatio-temporelle est un objet d'étude beaucoup plus complexe que l'hétérogénéité temporelle, et ce, notamment en raison de la multiplicité des dimensions. Le nombre de simulations n'est pas suffisant pour mettre en évidence, de manière aussi précise que précédemment, les influences des différents paramètres sur le comportement évolutif des populations mais certains effets semblent cependant, pouvoir être avancés. La comparaison des observations des simulations aux résultats théoriques de l'hétérogénéité spatiale montrent des différences dues à la dimension temporelle de l'hétérogénéité spatio-temporelle.
La différenciation des habitats semble être le principal facteur influençant le comportement évolutif : plus la différenciation est élevée, plus le dimorphisme est favorisé. L'effet de la durée des saisons est moins net, mais il semble que plus les durées de saison sont courtes, plus le monomorphisme est favorisé. Ces observations sont tout à fait cohérentes avec les résultats issus de l'étude de l'hétérogénéité temporelle.

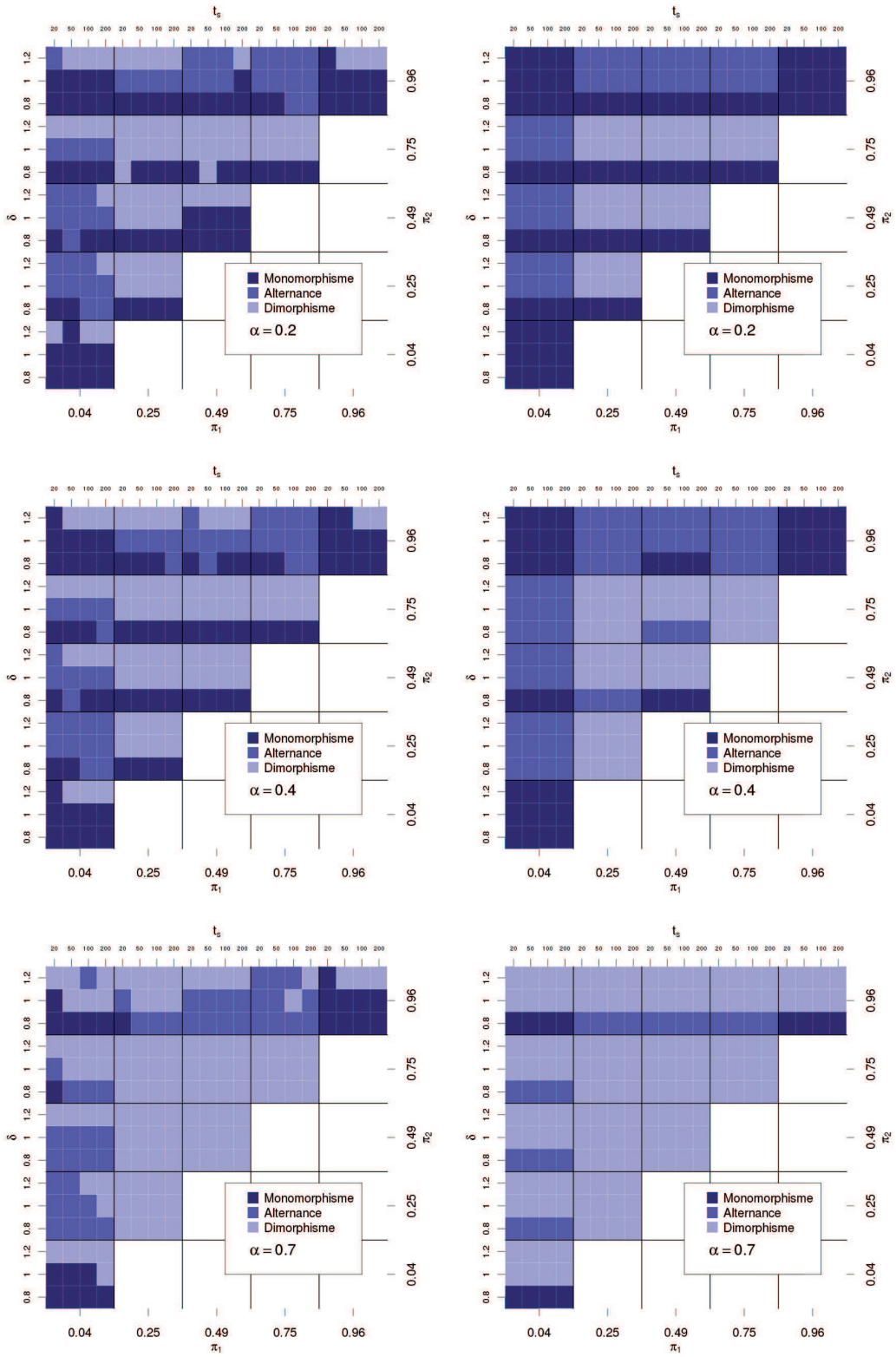


FIGURE 10 – Comportements évolutifs théoriques (à droite) et observés (à gauche). De haut à bas, $\alpha(s) = 0,2$; $\alpha(s) = 0,4$ et $\alpha(s) = 0,7$. Avec π_1 en abscisse, π_2 en ordonnée, t_s en seconde abscisse et δ en seconde ordonnée. En bleu foncé, les populations sont monomorphes à toutes les saisons ; en bleu clair, les populations sont dimorphiques à toutes les saisons ; en bleu moyen, les populations alternent entre des saisons de monomorphisme et des saisons de dimorphisme.

Par rapport aux résultats théoriques issus du critère spatial de branchement, l'effet de la différenciation des habitats semble amplifié en hétérogénéité spatio-temporelle. Par exemple, les simulations du type ($\alpha = 0,4$; $\pi_1 = 0,04$; $\pi_2 = 0,75$) montrent un effet très net de ce paramètre, en comparaison aux résultats déduits du critère spatial de branchement qui prédisent une alternance entre monomorphisme et dimorphisme quel que soit la valeur de la différenciation.
Concernant les paramètres spatiaux, plus l'agrégation des habitats est élevée plus le dimorphisme semble favorisé. Ces résultats sont cohérents avec ceux provenant de l'étude de l'hétérogénéité spatiale (Papaïx *et al.* 2011a) : plus les habitats sont agrégés, plus les spécialistes ont un niveau d'adaptation élevé, et les populations dimorphiques étant généralement des populations spécialisées, cet effet n'est pas surprenant. L'effet de la proportion des habitats semble moins déterminant que l'agrégation, excepté pour les proportions extrêmes qui favorisent nettement le monomorphisme. Enfin, plus l'amplitude des changements de proportion d'habitats est importante, plus le monomorphisme est favorisé.
Globalement, les résultats d'hétérogénéité spatio-temporelle concordent avec ceux de l'hétérogénéité temporelle : des changements fréquents de paysage favorisent des populations monomorphes et généralistes, là où des habitats bien agrégés dans l'espace et dans le temps, favorisent des populations dimorphiques de spécialistes. Qui plus est, les simulations dont les variations de proportions d'habitats impliquent des valeurs extrêmes, $0,04$ ou $0,96$, semblent plus sensibles à l'hétérogénéité spatio-temporelle. De même, les simulations dont les changements saisonniers sont très rapprochés ($t_s = 20$) montrent une plus grande sensibilité aux paramètres spatiaux. Inversement, sous de fortes agrégations, que l'on peut considérer comme une moindre hétérogénéité spatiale, et des valeurs moyennes de proportions d'habitats, la différenciation et la durée des saisons n'ont plus d'effet. Il y aurait donc des phénomènes d'amplification des effets de l'hétérogénéité : plus les populations sont soumises à une forte hétérogénéité sous certains aspects, plus elles seront sensibles à d'autres aspects de l'hétérogénéité.

5 Discussion

L'hétérogénéité temporelle, tout comme l'hétérogénéité spatiale, peut influencer l'adaption des pathogènes. Un modèle de dynamique adaptative, décrivant les changements phénotypique au cours du temps d'une population de pathogènes soumis à des changements périodiques d'habitat, permet de montrer que la spécialisation d'une population peut dépendre fortement des caractéristiques paysagères temporelle qui lui sont imposées. En cela, les résultats présentés confirment et illustrent sur un aspect temporel la supériorité qu'ont les stratégies généralistes dans l'exploitation de l'hétérogénéité écologique (Loreau *et al.* 2001, Gravel *et al.* 2011).
En particulier, la durée des saisons, correspondant au temps de présence d'un type d'habitat dans l'environnement et que l'on peut assimiler à une sorte d'« agrégation temporelle », est un paramètre déterminant pour l'évolution de pathogènes. Plus la durée des saisons est courte, plus le niveau d'adaptation des pathogènes sera bas, jusqu'à pouvoir favoriser l'adoption d'une stratégie généraliste. Le niveau d'adaptation est défini par rapport à la stratégie optimale, déterminée par la différenciation des habitats. Les influences de ces deux paramètres, durée des saisons et différenciation des habitats, sont croisées et complexes : la sensibilité à la durée des saisons (ou à la fréquence des changements saisonniers) est d'autant plus forte que la différenciation des habitats est faible. Ce n'est plus vrai pour des habitats très différenciés, où la distance phénotypique à parcourir entre deux optima de spécialisation devient suffisamment importante pour que le temps nécessaire à la parcourir soit allongé, et en conséquence, entraîner une sensibilité accrue aux perturbations des changements saisonniers.
Ces résultats illustrent l'influence complexe que peut jouer un aspect de l'hétérogénéité temporelle sur un compromis écologique, ici la spécialisation entre deux habitats. D'autres compromis écologiques peuvent être influencés de la sorte. À titre d'exemple, l'absence périodique de l'hôte peut entraîner chez l'hôte une augmentation du taux de transmission, selon un compromis entre transmission et virulence, ou sa diminution, selon un compromis entre transmission et survie saisonnière (van den Berg *et al.* 2010). L'influence de l'hétérogénéité temporelle sur des compromis décisifs dans l'adoption de stratégies écologiques par les pathogènes, et de leur devenir évolutif, est un argument fort pour intégrer de manière explicite les aspects temporels dans les études sur les liens entre hétérogénéité environnementale et évolution des pathogènes.
Les résultats provenant de l'étude de l'hétérogénéité spatio-temporelle montrent en effet qu'il peut exister des phénomènes d'amplification entre différents aspects : l'effet de la fréquence des changements saisonniers, ou de la différenciation des habitats est d'autant plus fort qu'il s'accompagne d'une faible agrégation et de saisons avec des proportions d'habitat extrêmes. Inversement, les effets de certains aspects de l'hétérogénéité peuvent être contrariés par des facteurs d'homogénéité : l'agrégation des habitats peut inhiber complètement l'effet de la différenciation des habitats, par exemple.

L'idée d'appliquer une hétérogénéité spatio-temporelle à un espace agricole pour mieux réguler les épidémies végétales et plus freiner l'apparition de spécialisations, par rapport à la seule hétérogénéité spatiale ou temporelle, est illustrée par la rotation de mélanges de culture. Cette pratique agricole est un outil majeur du maintien des stratégies agricoles modernes conciliant productivité agricole, réduction des maladies et préservation de l'environnement (Oddino *et al.* 2008, Li *et al.* 2009). Elle repose sur l'emploi d'un mélange de cultures, pour forcer les pathogènes à survivre dans un environnement d'hôte spatialement hétérogène, et une rotation dans le temps de ces cultures, pour forcer les pathogènes à survivre dans un environnement d'hôte temporellement hétérogène (Xu 2011).
Mais au-delà des ces principes généraux, la compréhension des effets des aspects spatiaux, temporels et spatio-temporels de l'hétérogénéité sur l'écologie et l'évolution des pathogènes nécessite un effort de recherche important, passant par l'amélioration des modèles actuellement développés et l'exploration de leurs hypothèses. Le modèle ici étudié est un modèle de dynamique adaptative dont les limites sont liées à la simplification des caractéristiques des populations et des paysages spatio-temporels. L'étude de l'influence de l'hétérogénéité s'est basée sur la caractérisation de phénotypes dominant. Les résultats ici montrés s'appuient sur la caractérisation de phénotypes dominants. Cette caractérisation ne doit pas faire oublier qu'en réalité l'ensemble des classes phénotypiques entre les deux stratégies optimales est représenté, mais dans des effectifs très faibles. Ce polymorphisme marginal peut provenir de la présence des deux habitats à chaque saison qui permet de maintenir de faibles effectifs spécialisés sur un habitat parfois très minoritaire au cours d'une saison, ou simplement provenir d'effets de retard, liés aux paramètres démographiques. Ce polymorphisme à bas niveau permet d'expliquer l'absence de branchements évolutifs ultérieurs dans les simulations dont le comportement évolutif est une alternance entre dimorphisme et monomorphisme : les effectifs, certes très faibles, sont cependant suffisants pour entraîner immédiatement une nouvelle spécialisation (figure 9, graphique du milieu). L'étude de ce polymorphisme à bas niveau, l'élargissement des hypothèses d'haploïdie, d'asexualité, de générations non chevauchantes, l'étude de l'adaptation sur plusieurs traits sont autant de perspectives pour développer les aspects populationnels du modèle. De manière semblable, l'intégration d'autres aspects de l'hétérogénéité et l'augmentation du nombre d'habitats sont les premiers développements envisageables des aspects environnementaux du modèle proposé. L'objectif dans lequel s'inscrit ce projet est l'élaboration d'un modèle intégrant l'ensemble des aspects de l'hétérogénéité spatio-temporelle. Un tel modèle constituerait un outil à même de pouvoir étudier l'ensemble des interactions et des effets des différentes composantes de l'hétérogénéité écologique, préalable indispensable à l'élaboration de stratégies agricoles de gestion de la spécialisation des pathogènes.

6 Conclusion

Dans un contexte où la législation européenne cherche à contrôler l'emploi des produits phytosanitaires afin de garantir un niveau élevé de protection de l'environnement, et dans le même temps de préserver la compétitivité de l'agriculture communautaire (Parlement européen 2009), la manipulation des structures spatio-temporelles des paysages semble une voie prometteuse pour la gestion de l'évolution des pathogènes et du risque épidémique en milieu agricole. Le modèle proposé s'appuie sur la théorie de la dynamique adaptative pour montrer l'influence de la fréquence des changements de cultures sur la spécialisation des pathogènes. Cet aspect de l'hétérogénéité peut contraindre fortement le niveau d'adaptation des pathogènes sur des habitats pourtant monovariétaux à un temps donné, et potentiellement réduire d'autant leur agressivité. Au niveau de l'hétérogénéité spatio-temporelle, le devenir évolutif d'une population est influencée à la fois par des aspects temporels et des aspects spatiaux, avec des phénomènes d'amplification ou d'inhibition. L'étude du rôle évolutif de l'hétérogénéité s'inscrit dans le courant de la biologie évolutive appliquée, et défend une meilleure intégration des principes évolutifs et écologiques dans les techniques modernes agricoles. À l'heure où les impacts environnementaux de l'agriculture et sa vulnérabilité face aux épidémies végétales constituent des critères de plus en plus valorisés à côté des traditionnelles maximisations de la production et du rendement, l'un des principaux enseignements de ce courant est de rapprocher le fonctionnement des systèmes cultivés à celui des écosystèmes. La réintroduction de l'hétérogénéité écologique dans les agrosystèmes constitue donc une priorité en politique agricole, à laquelle doit contribuer la recherche agronomique.

Références

Académie Française : *Dictionnaire de l'Académie Française, 8^e édition*. Hachette, Paris (France), 1932-1935.

H. W. Art : *The Dictionary of Ecology and Environmental Science*. Henry Holt & Company, New York City (États-Unis d'Amérique), 1995.

A. Bouvier, K. Kieu, K. Adamczyk et H. Monod : Computation of the integrated flow of particles between polygons. *Environmental Modelling and Software*, 24:843–849, 2009.

Å. Brännström et N. V. Festenberg : The hitchhiker's guide to adaptive dynamics, 2006. URL `http://adtoolkit.sourceforge.net/adintro.pdf`.

H. Caswell : *Matrix population models*. Sinauer Associates Inc. Publishers, Sunderland (États-Unis d'Amérique), 2001.

N. Champagnat, R. Ferrière et S. Méléard : Unifying evolutionary dynamics : from individual stochastic processes to macroscopic models. *Theoretical Population Biology*, 69:297–321, 2006.

F. B Christiansen : On conditions for evolutionary stability for a continuously varying character. *The American Naturalist*, 138:37–50, 1991.

F. Débarre et S. Gandon : Evolution of specialization in a spatially continuous environment. *Journal of Evolutionnary Biology*, 23:1090–1099, 2010.

U. Dieckmann et R. Law : The dynamical theory of coevolution : a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34:579–612, 1996.

O. Diekmann : A beginners guide to adaptive dynamics. *Banach Center Publications*, 63:47–86, 2003.

K. A. Garrett, S. H. Hulbert, J. E. Leach et S. E Travers : Ecological genomics and epidemiology. *European journal of plant pathology*, 115:35–51, 2006.

S. Geritz, E. Kisdi, G. Meszena et J. Metz : Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, 12:35–57, 1998.

P. S Giller : *Community structure and the niche*. Chapman and Hall, New York City (États-Unis d'Amérique), 1984.

D. Gravel, T. Bell, C. Barbera, T. Bouvier, T. Pommier, P. Venail et N. Mouquet : Experimental niche evolution alters the strength of the diversity-productivity relationship. *Nature*, 469:89–92, 2011.

D.A. Griffith : The boundary value problem in spatial statistical analysis. *Journal of Regional Science*, 23:377–387, 1983.

Michel Griffon : *Nourrir la planète : pour une révolution doublement verte*. Odile Jacob, Paris (France), 2006.

H. S. He, B. E. DeZonia et D.J. Mladenoff : An aggregation index (AI) to quantify spatial patterns of landscapes. *Landscape Ecology*, 15:591–601, 2000.

A. P. Hendry, M. T. Kinnison, M. Heino, T. Day, T. B. Smith, G. Fitt, C. T. Bergstrom, J. Oakeshott, P., S. Jørgensen, M. P Zalucki, G. Gilchrist, S. Southerton, A. Sih, S. Strauss et R. F. Denison S.P. Carroll : Evolutionary principles and their practical application. *Evolutionary Applications*, 4:159–183, 2011.

J. Hofbauer et K. Sigmund : Adaptive dynamics and evolutionary stability. *Applied Mathematics Letters*, 3:75–79, 1990.

M. J. Hutchings et H. de Kroon : Foraging in plants : the role of morphological plasticity in resource acquisition. *Advances in Ecological Research*, 25:159–238, 1994.

T. Johnson : Man-guided evolution in plant rusts. *Science*, 10:357–362, 1961.

P. Juhász-Nagy et J. Podani : Information theory methods for the study of spatial processes and succession. *Plant Ecology*, 51(3):129–140, 1983.

R. Kassen : The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of Evolutionary Biology*, 15:173–190, 2002.

S. Kirkpatrick, C. D. Gelatt et M.P. Vecchi : Optimization by simulated annealing. *Science*, 220:671–680, 1983.

J. Kolasa, S. T. Pickett et T. F. H. Allen : *Ecological Heterogeneity*. Springer-Verlag, New York City (États-Unis d'Amérique), 1991.

H. Levene : Genetic equilibrium when more than one niche is available. *American Naturalist*, 87:331–333, 1953.

C. Li, X. He, S. Zhu, H. Zhou, Y. Wang, Y. Li, J. Yang, J. Fan, J. Yang, G. Wang, Y. Long, J. Xu, Y. Tang, G. Zhao, J. Yang, L. Liu, Y. Sun, Y. Xie, H. Wang et Y. Zhu : Crop diversity for yield increase. *PLoS One*, 4:e8049, 2009.

H. Li et J.F. Reynolds : On definition and quantification heterogeneity. *Oikos*, 73:280–284, 1995.

H. B. Li et J.F. Reynolds : A simulation experiment to quantify spatial heterogeneity in categorical maps. *Ecology*, 75:2446–2455, 1994.

Roger J. Lincoln, Geoffrey Allan Boxshall et P. F. Clark : *A dictionary of ecology, evolution, and systematics*. Cambridge University Press, Cambridge, Royaume-Uni, 1998.

M. Loreau, S. Naeem, P. Inchausti, J. Bengtsson, JP Grime, A. Hector, DU Hooper, MA Huston, D. Raffaelli, B. Schmid *et al.* : Biodiversity and ecosystem functioning : current knowledge and future challenges. *Science*, 294:804–808, 2001.

J. A. J. Metz, R. M. Nisbet et S. A. H. Geritz : How should we define 'fitness' for general ecological scenarios ? *Trends in Ecology & Evolution*, 7:198–202, 1992.

B. T Milne : Heterogeneity as a multiscale characteristic of landscapes. *Ecological Studies*, 86:69–84, 1991.

M. Nowak et K. Sigmund : The evolution of stochastic strategies in the prisoner's dilemma. *Acta Applicandae Mathematicae*, 20:247–265, 1990.

B. Oborny : Growth rules in clonal plants and environmental Predictability–A simulation study. *Journal of Ecology*, 82:341–351, 1994.

C. M Oddino, A. D Marinelli, M. Zuza et G. J March : Influence of crop rotation and tillage on incidence of brown root rot of peanut caused by fusarium solani in argentina. *Canadian Journal of Plant Pathology*, 30:575–580, 2008.

E. Oerke et H. Dehne : Safeguarding production - losses in major crops and the role of crop protection. *Crop Protection*, 23:275–285, 2004.

E.C. Oerke, H.W. Dehne, F. Schönbeck et A. Weber : *Crop production and crop protection : estimated losses in major food and cash crops.* Elsevier Science Publishers, Amsterdam (Pays-Bas) & New York City (États-Unis d'Amérique), 1994.

J. Papaïx, O. David, C. Lannou et H. Monod : The dynamics of adaptation in spatially heterogenous environment. *en préparation*, 2011a.

J. Papaïx, H. Goyeau, P. Du Cheyron et C. Lannou : Influence of cultivated landscape composition on variety resistance : an assessment based on wheat leaf rust epidemics. *New Phytologist*, 2011b.

Parlement européen : Règlement (CE) n° 1107/2009 du Parlement européen et du Conseil du 21 octobre 2009 concernant la mise sur le marché des produits phytopharmaceutiques et abrogeant les directives 79/117/CEE et 91/414/CEE du Conseil. *Journal officiel*, n° L 309 (24/11/2009):1 – 5, 2009.

S. T. A. Pickett et G. E. Likens : Space-for-time substitution as an alternative to long-term studies. *In Long-Term Studies in Ecology - Approaches and Alternatives*, pages 110–135. Springer-Verlag, Berlin (Allemagne), 1989.

R. A. Robinson et W.J. Sutherland : Post-war changes in arable farming and biodiversity in Great Britain. *Journal of Applied Ecology*, 39:157–176, 2002.

J. Maynard Smith et G. R Price : The logic of animal conflict. *Nature*, 246:15–18, 1973.

D. L Strayer, H. A Ewing et S. Bigelow : What kind of spatial and temporal details are required in models of heterogeneous systems ? *Oikos*, 102:654–662, 2003.

J. F Stuefer : Potential and limitations of current concepts regarding the response of clonal plants to environmental heterogeneity. *Plant Ecology*, 127:55–70, 1996.

E. Stukenbrock et B. McDonald : The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, 46:75–100, 2008.

P. H Thrall, J. G Oakeshott, G. Fitt, S. Southerton, J. J Burdon, A. Sheppard, R. J Russell, M. Zalucki, M. Heino et R. Ford Denison : Evolution in agriculture : the application of evolutionary approaches to the management of biotic interactions in agro-ecosystems. *Evolutionary Applications*, 4:200–215, 2011.

D. Tilman : Global environmental impacts of agricultural expansion : the need for sustainable and efficient practices. *Proceedings of the National Academy of Sciences of the United States of America*, 96:5995–6000, 1999.

M. G Turner, R. H Gardner et R. V O'Neill : *Landscape Ecology in Theory and Practice : Pattern and Process*. Springer Verlag, Berlin (Allemagne), 2001.

F. van den Berg, N. Bacaër, J. A.J Metz, C. Lannou et F. van den Bosch : Periodic host absence can select for higher or lower parasite transmission rates. *Evolutionary Ecology*, 25:121–137, 2010.

B. Wallace : Hard and soft selection revisited. *Evolution*, 29:465–473, 1975.

J. A. Wiens : Ecological heterogeneity : an ontogeny of concepts and approaches. *In The ecological consequences of environmental heterogeneity : the 40th symposium of the British Ecological Society, held at the University of Sussex, 23-25 March 1999*, 2000.

X. M Xu : A simulation study on managing plant diseases by systematically altering spatial positions of cultivar mixture components between seasons. *Plant Pathology*, 60, 2011.

Premiers développements des aspects analytiques

L'établissement d'une trame analytique rigoureuse à même d'expliquer l'ensemble des observations est une tâche très ardue. Cependant, dans le cas extrême d'un changement saisonnier à chaque pas de temps évolutif, il peut facilement être montré que la stratégie généraliste est une stratégie évolutive stable (ESS). Par ailleurs, dans le cas de changements saisonniers impliquant des durées de saison suffisamment longues, une application de l'équation canonique de la dynamique adaptative permet de rendre compte des oscillations du phénotype de la population telles qu'observées dans le graphique gauche de la figure 3.
Les simulations où le phénotype subit des oscillations amorties (par exemple, graphique droite de la figure 3) correspondent à des cas intermédiaires. Cet amortissement est sans doute lié à la cinétique de la dynamique de la distribution phénotypique de la population. Le polymorphisme des populations soumises à ces oscillations amorties n'est plus compatible avec l'hypothèse de monomorphisme et en cela réside l'origine de leur difficulté analytique.

Caractérisation des stratégies singulières

Soit une population résidente x dans laquelle apparaît un mutant y. Dans un environnement constant dans le temps, le devenir de ce mutant est déterminé par son coefficient d'invasion $s_x^{(h)}(y)$ avec h l'habitat présent lors de l'apparition du mutant. Ce coefficient est le rapport des taux de survie du mutant et de la population résidente :

$$s_x^{(h)}(y) = \frac{f(y,h)}{f(x,h)} = \frac{e^{-\frac{(y-\beta_h)^2}{2\sigma^2}}}{e^{-\frac{(x-\beta_h)^2}{2\sigma^2}}}$$

Cas de saisons longues Au cours d'une saison, le paysage est invariant dans le temps. Dans le cas de saisons longues, ce paysage permet l'évolution d'une population vers la stratégie optimale.

$$\log s_x^{(h)}(y) = -\frac{1}{2\sigma^2}\left((y-\beta_h)^2 - (x-\beta_h)^2\right)$$

$$\log s_x^{(h)}(y) = -\frac{1}{2\sigma^2}(y-x)(y+x-2\beta_h)$$

Le gradient de sélection est égal à :

$$\Upsilon_x^{(h)}(y) = \frac{\partial s_x^{(h)}(y)}{\partial y} = -\frac{\beta_h - y}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}(y-x)(y+x-2\beta_h)\right)$$

$\Upsilon_{\beta_h}^{(h)}(\beta_h) = 0$ donc $x^\star = \beta_h$ est une stratégie singulière.
Le calcul des dérivées secondes du coefficient d'invasion permettent de déterminer la convergence et la stabilité des stratégies singulières (Geritz *et al.* 1998, Brännström et Festenberg 2006). Le calul de la dérivée seconde du coefficient d'invasion par rapport à x^2 donne :

$$\frac{\partial^2 s_x^{(h)}(y)}{\partial x^2} = \frac{\partial}{\partial x}\left[\frac{x-\beta_h}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}(y-x)(y+x-2\beta_h)\right)\right]$$

$$\frac{\partial^2 s_x^{(h)}(y)}{\partial x^2} = \frac{1}{\sigma^2}\left(1 + \frac{(x-\beta_h)^2}{\sigma^2}\right)\exp\left(-\frac{1}{2\sigma^2}(y-x)(y+x-2\beta_h)\right)$$

Le calcul de la dérivée seconde du coefficient d'invasion par rapport à y^2 donne :

$$\frac{\partial^2 s_x^{(h)}(y)}{\partial y^2} = \frac{\partial}{\partial y}\left[\frac{\beta_h - y}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}(y-x)(y+x-2\beta_h)\right)\right]$$

$$\frac{\partial^2 s_x^{(h)}(y)}{\partial y^2} = \frac{1}{\sigma^2}\left(-1 + \frac{(\beta_h - y)^2}{\sigma^2}\right)\exp\left(-\frac{1}{2\sigma^2}(y-x)(y+x-2\beta_h)\right)$$

En la singularité singularité, nous avons :

$$\left.\frac{\partial^2 s_x^{(h)}(y)}{\partial x^2}\right|_{x=y=x^\star} = \frac{1}{\sigma^2} \qquad \text{et} \qquad \left.\frac{\partial^2 s_x^{(h)}(y)}{\partial y^2}\right|_{x=y=x^\star} = -\frac{1}{\sigma^2}$$

$$\left.\frac{\partial^2 s_x^{(h)}(y)}{\partial x^2}\right|_{x=y=x^\star} > \left.\frac{\partial^2 s_x^{(h)}(y)}{\partial y^2}\right|_{x=y=x^\star}$$

Donc $x^\star = \beta_h$ est une stratégie évolutive convergente (CSS).
Par ailleurs,

$$\left.\frac{\partial^2 s_x^{(h)}(y)}{\partial y^2}\right|_{x=y=x^\star} = -\frac{1}{\sigma^2} < 0.$$

Donc $x^\star = \beta_h$ est une stratégie évolutive stable (ESS).

Cas des saisons très courtes Dans un environnement marqué par des changements saisonniers d'habitat, le devenir du mutant est déterminé par le produit de ses coefficients d'invasion propores à chaque habitat. Dans le cas extrême d'un changement saisonnier à chaque cycle, on peut écrire :

$$\varsigma_x(y) = s_x^{(1)}(y) \cdot s_x^{(2)}(y)$$

$\varsigma(x, y)$ est indépendant de δ :

$$\log \varsigma_x(y) = -\frac{1}{2\sigma^2}\left((y+\delta)^2 - (x+\delta)^2 + (y-\delta)^2 - (x-\delta)^2\right)$$

$$\log \varsigma_x(y) = -\frac{1}{2\sigma^2}\left((y-x)(x+y+2\delta) + (y-x)(x+y-2\delta)\right)$$

$$\log \varsigma_x(y) = \frac{x^2-y^2}{\sigma^2}$$

Le gradient de sélection est égal à :

$$\Upsilon_x(y) = \frac{\partial \varsigma_x(y)}{\partial y} = -\frac{2y}{\sigma^2} e^{\frac{x^2-y^2}{\sigma^2}}$$

Pour $x = y$, le gradient de sélection devient :

$$\Upsilon_x(x) = -\frac{2x}{\sigma^2}$$

$\Upsilon_0(0) = 0$ donc $x^\star = 0$ est une stratégie singulière.
Le calcul de la dérivée seconde par rapport à x^2 donne :

$$\frac{\partial^2 \varsigma_x(y)}{\partial x^2} = \frac{\partial}{\partial x}\left[\frac{2x}{\sigma^2} e^{\frac{x^2-y^2}{\sigma^2}}\right]$$
$$\frac{\partial^2 \varsigma_x(y)}{\partial x^2} = \frac{2}{\sigma^2}\left(1 + \frac{2x^2}{\sigma^2}\right) e^{\frac{x^2-y^2}{\sigma^2}}$$

Le calcul de la dérivée seconde par rapport à y^2 donne :

$$\frac{\partial^2 \varsigma_x(y)}{\partial y^2} = \frac{\partial}{\partial y}\left[-\frac{2y}{\sigma^2} e^{\frac{x^2-y^2}{\sigma^2}}\right]$$
$$\frac{\partial^2 \varsigma_x(y)}{\partial y^2} = \frac{2}{\sigma^2}\left(-1 + \frac{2y^2}{\sigma^2}\right) e^{\frac{x^2-y^2}{\sigma^2}}$$

En la singularité singularité, nous avons :

$$\left.\frac{\partial^2 \varsigma_x(y)}{\partial x^2}\right|_{x=y=x^\star} = \frac{2}{\sigma^2} \qquad \text{et} \qquad \left.\frac{\partial^2 \varsigma_x(y)}{\partial y^2}\right|_{x=y=x^\star} = -\frac{2}{\sigma^2}$$

$$\left.\frac{\partial^2 \varsigma_x(y)}{\partial x^2}\right|_{x=y=x^\star} > \left.\frac{\partial^2 \varsigma_x(y)}{\partial y^2}\right|_{x=y=x^\star}$$

Donc $x^\star = 0$ est une stratégie évolutive convergente (CSS).
Par ailleurs,

$$\left.\frac{\partial^2 \varsigma_x(y)}{\partial y^2}\right|_{x=y=x^\star} = -\frac{2}{\sigma^2} < 0.$$

Donc $x^\star = 0$ est une stratégie évolutive stable (ESS).

Équation canonique de la dynamique adaptative

La vitesse évolutive du trait phénotypique x dans une population monomorphe peut être décrite par l'équation canonique de la dynamique adaptative (Champagnat *et al.* 2006, Dieckmann et Law 1996) :

$$\dot{x} = \frac{\gamma^2(x)\theta(x)}{2}\, \partial p_x(y)\partial y|_x \tag{1}$$

Dans cette équation, $p_x(y)$ est la probabilité de survie du mutant de phénotype y dans une population résidente x. L'échelle de temps est modifiée de sorte que les mutations apparaissent selon un processus de Poisson à temps continu de taux $\theta(x)$. Chaque mutation modifie le trait phénotypique selon une variance $\gamma^2(x)$.
D'après Papaïx *et al.* (2011a), la stochasticité affectant la démographie des individus mutants rares peut être prise en compte par un paramètre ν^2 qui quantifie la variabiltié de la distibution des descendants d'un individu, de sorte que l'équation canonique peut être réécrite sous la forme :

$$\dot{x} = -\frac{\gamma^2(x)\theta(x)}{\sigma^2\nu^2}(x - x^\star)$$

Avec σ^2 la variance des courbes de survie. Si $\gamma^2(x)$ et $\theta(x)$ ne dépendent pas de x, cas dans lequel nous nous plaçons, les solutions de cette équation vérifient :

$$\frac{x(t) - x^\star(t)}{x_0(t) - x^\star(t)} = \exp\left(-\frac{\gamma^2\theta}{\sigma^2\nu^2}t\right) \tag{2}$$

avec $x^\star(t)$ la stratégie optimale au temps t et $x_0(t)$ la valeur du phénotype de la population au début de la saison correspondant au temps t.
La figure 11 représente l'estimation de $\frac{\gamma^2\theta}{\sigma^2\nu^2}$ pour chaque simulation. Pour des différenciations d'habitat supérieures à $0,5$, ce paramètre semble constant et indépendant de δ et de t_s. Pour des différenciations d'habitat inférieures à $0,5$, ce paramètre est plus faible (Ajouter Graphique « Influence de δ sur τ »). Cet écart peut s'expliquer par un éloignement du comportement des populations simulées par rapport à l'équation de la dynamique adaptative. En effet, les changements de saison entraînent chez des populations soumis à des habitats faiblement différenciés une modification plus sigmoïde qu'exponentielle du phénotype.
Dans la suite des calculs, nous considérerons $\frac{1}{\tau} = \frac{\gamma^2\theta}{\sigma^2\nu^2}$ constant et indépendant de δ et t_s.

$$x(t) = (-1)^s\delta + (x_0(t) - (-1)^s\delta)\, e^{-\frac{t}{\tau}}$$

avec $s \in [\![1, S]\!]$ la saison correspondant au temps t.
En notant $x_s(\delta, t_s)$ le phénotype à la fin de la saison s d'une population soumise à des saisons de durée t_s et à des habitats de différenciation δ et $x_0(\delta, t_s)$ la valeur initiale du phénotype de cette population, on a la relation de récurrence suivante :

$$\forall s \in [\![1, S]\!], x_s(\delta, t_s) = (-1)^s\delta + (x_{s-1}(\delta, t_s) - (-1)^s\delta)\, e^{-\frac{t_s}{\tau}} \tag{3}$$

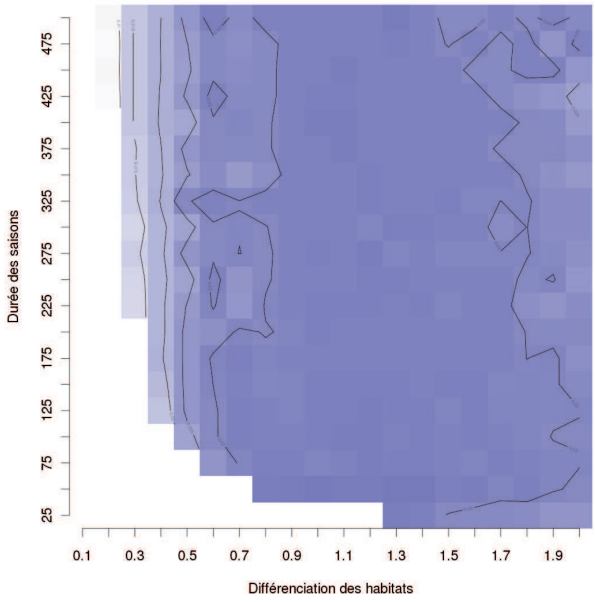


FIGURE 11 – Représentation de la valeur estimée de $\frac{\gamma^2\theta}{\sigma^2\nu^2}$ en fonction de la différenciation des habitats et de la durée des saisons. Pour une simulation donnée, la valeur représentée correspond à la moyenne des estimations réalisées par la méthode des moindres carrés à chaque changement saisonnier. Ont été exclues les simulations pour lesquelles le phénotype final était nul.

Par définition, la valeur absolue de cette suite tend vers le phénotype final :

$$\lim_{s\to\infty} |x_s(\delta, t_s)| = \phi_f(\delta, t_s)$$

Une expression analytique de ϕ_f en fonction de δ et t_s peut être déduite de l'équation 3 en supposant que $x_{s-1} = -x_s$:

$$x_s = (-1)^s\delta + (x_{s-1} - (-1)^s\delta)\, e^{-\frac{t_s}{\tau}}$$

$$x_s = (-1)^s\delta + (-x_s - (-1)^s\delta)\, e^{-\frac{t_s}{\tau}}$$

$$x_s = \frac{1 - e^{-\frac{t_s}{\tau}}}{1 + e^{-\frac{t_s}{\tau}}}(-1)^s\delta$$

$$x_s = (-1)^s\delta \tanh\left(\frac{t_s}{2\tau}\right)$$

D'où une expression du phénotype final :

$$\phi_f = \delta \tanh\left(\frac{t_s}{2\tau}\right) \tag{4}$$

Le graphique de gauche de la figure 12 est une représentation de $\phi_f(\delta, t_s)$ en fonction de δ et t_s, déduit de ces considérations analytiques. On peut observer une remarquable concordance entre les résultats analytiques et les observations de simulations.

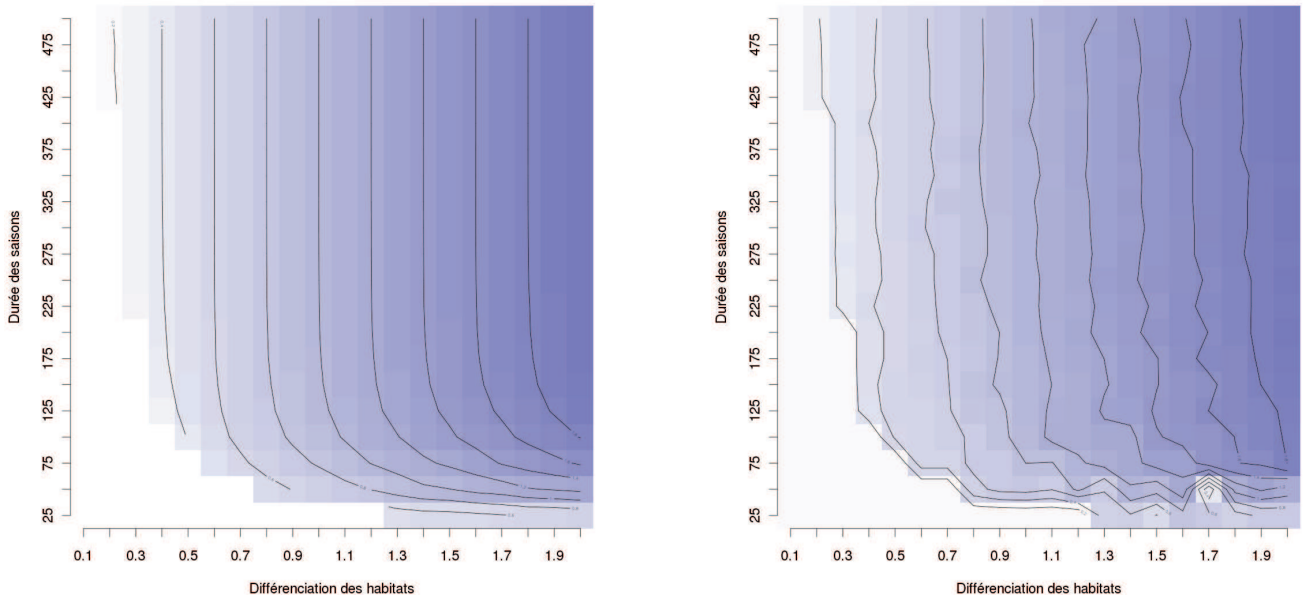


FIGURE 12 – Comparaison du phénotype final provenant de la résolution de l'équation 4 (à gauche) au phénotype final observé dans les simulations (à droite).

Appendix 6

PhD defence slides

These slides were presented during the defence of my PhD on Monday Septembre 26th 2011 at 2:30 pm - amphitheatre Tisserand, AgroParisTech 16 rue Claude Bernard 75005 PARIS.

Landscape structure and epidemic risk, a demo-genetic approach

Julien Papaïx

Supervised by Christian Lannou and Hervé Monod

Contents

Agricultural context

- Agro-ecosystems are susceptible to the epidemic risk.
- Alternative approaches, such as the use of genetic diversity, should be developed.
- Broader spatial scales should be considered and collective strategies should be designed.

Agro-ecosystems should evolve from low to high diversity (insurance hypothesis).

Ecological questions

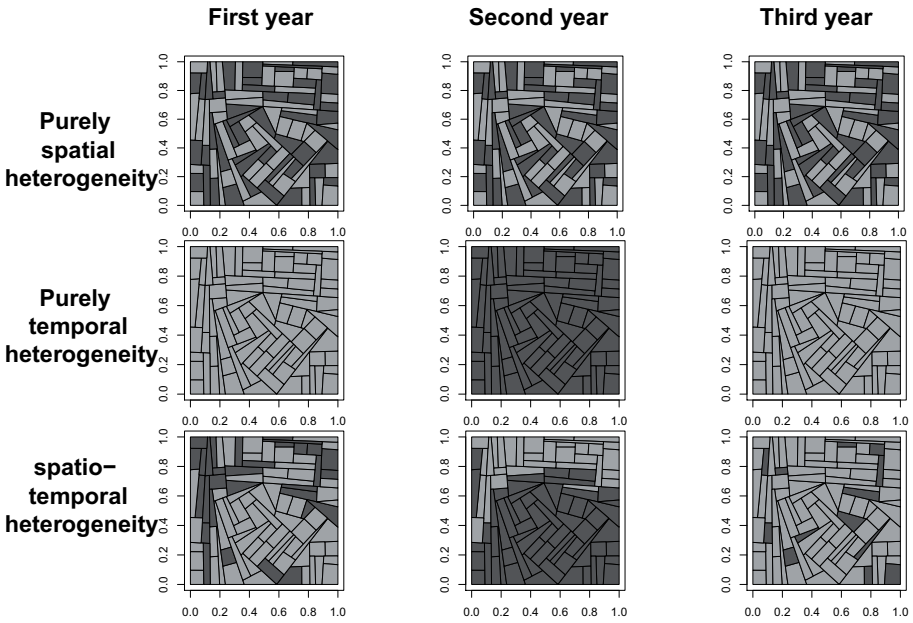
- How does a population spread over an heterogeneous landscape ?
- How do genotypes (or species) compete in spatially heterogeneous environments ?
- How does the spatial repartition of habitats influence the long term evolution of populations ?

Requires the interplay between three disciplines:

- epidemiology,
- landscape ecology,
- evolutionary biology.

Spatial heterogeneity

Environmental heterogeneity



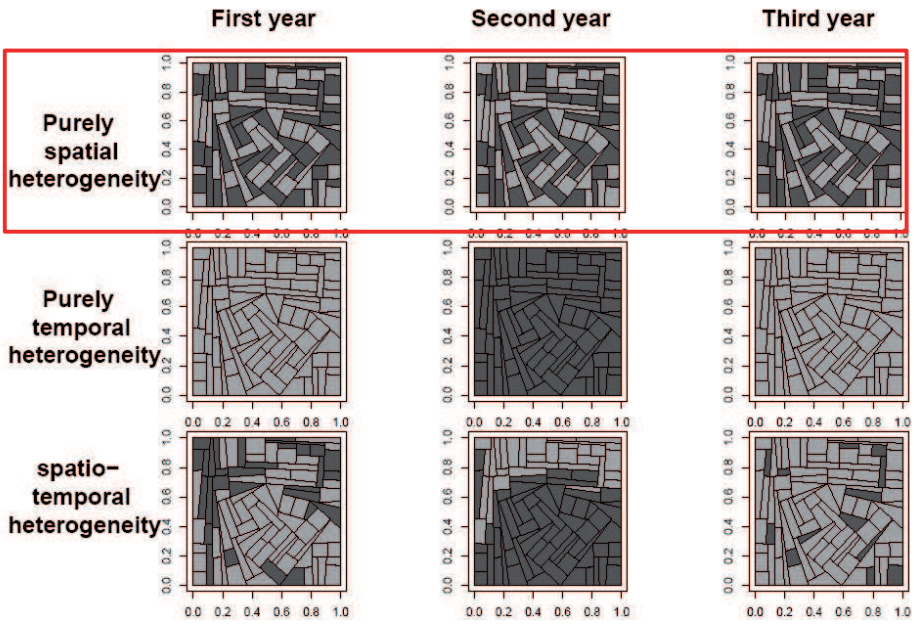
Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

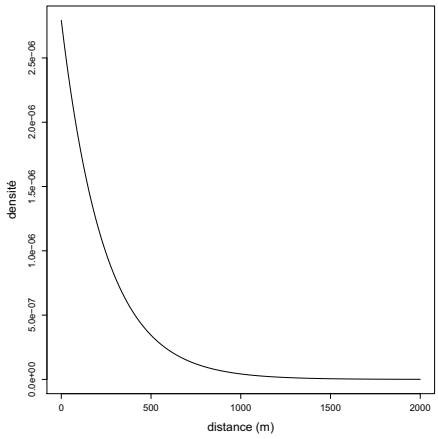
Spatial heterogeneity

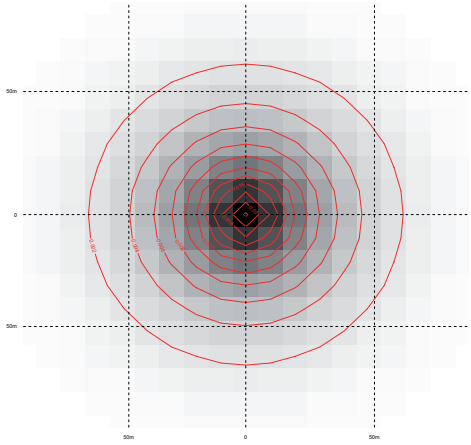
Environmental heterogeneity



Spatial heterogeneity

Dispersal limitation





Contents

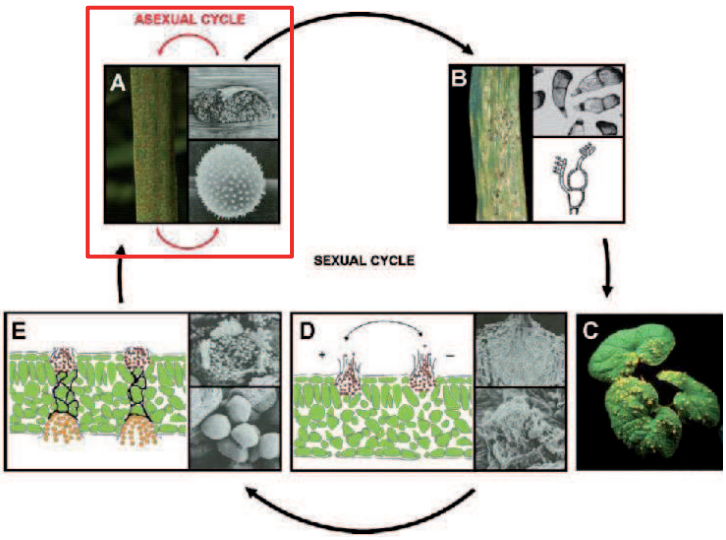
Contents

The wheat leaf rust pathosystem

Leaf rust (from Bolton 2008)

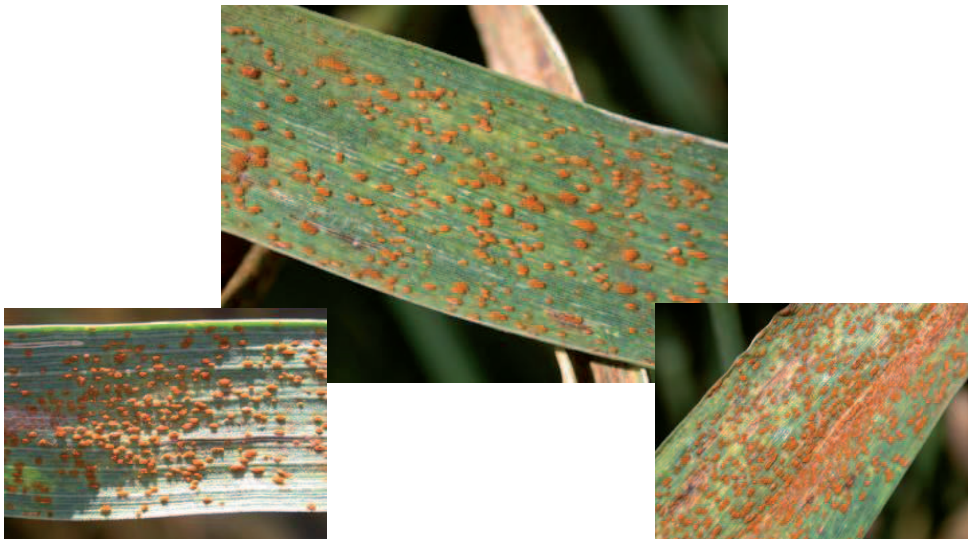
Triticum aestivum

Thalictrum speciosissimum

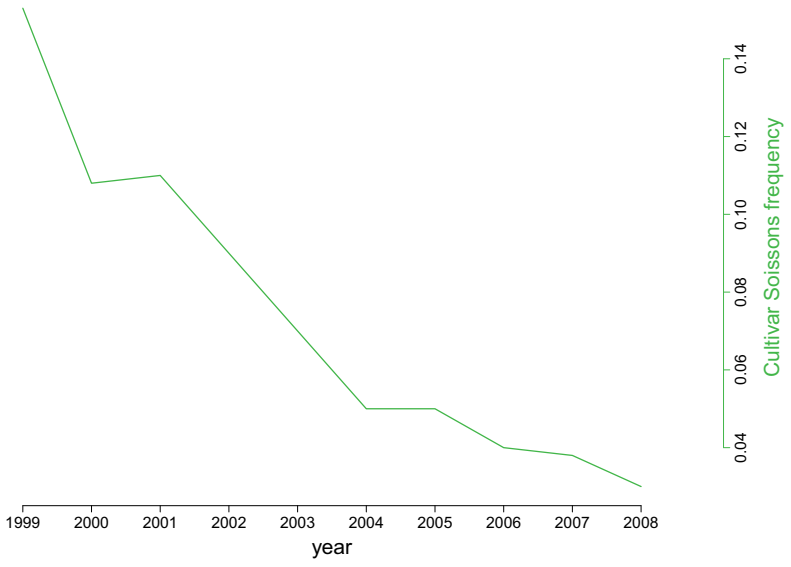


The wheat leaf rust pathosystem

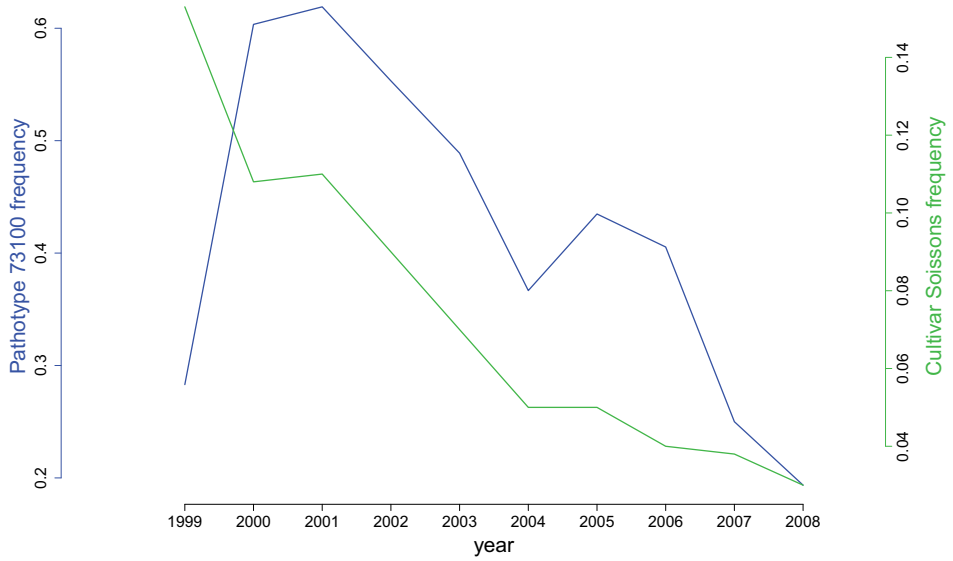
Leaf rust (from Pariaud 2008)



The Soissons-073100 case



The Soissons-073100 case

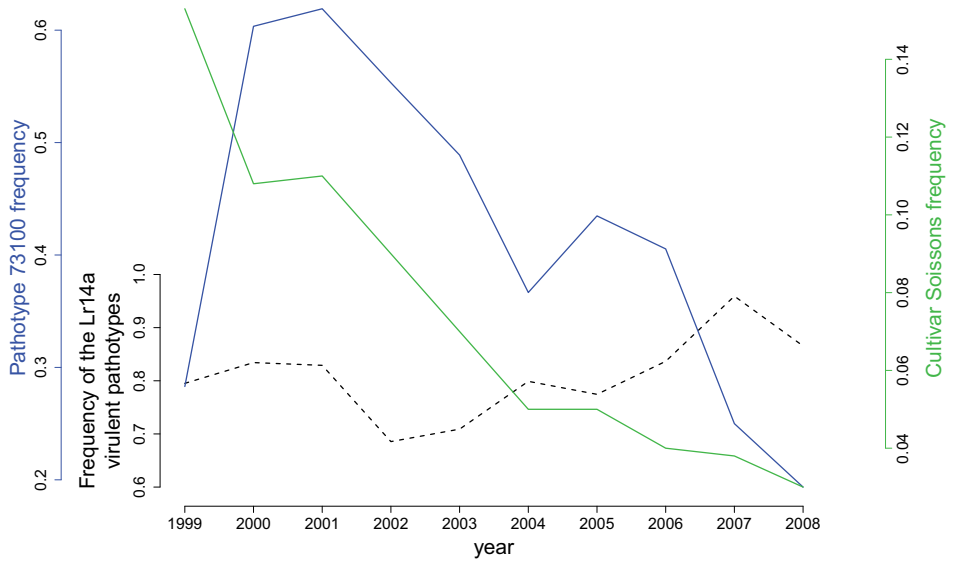


Landscape structure and epidemic risk

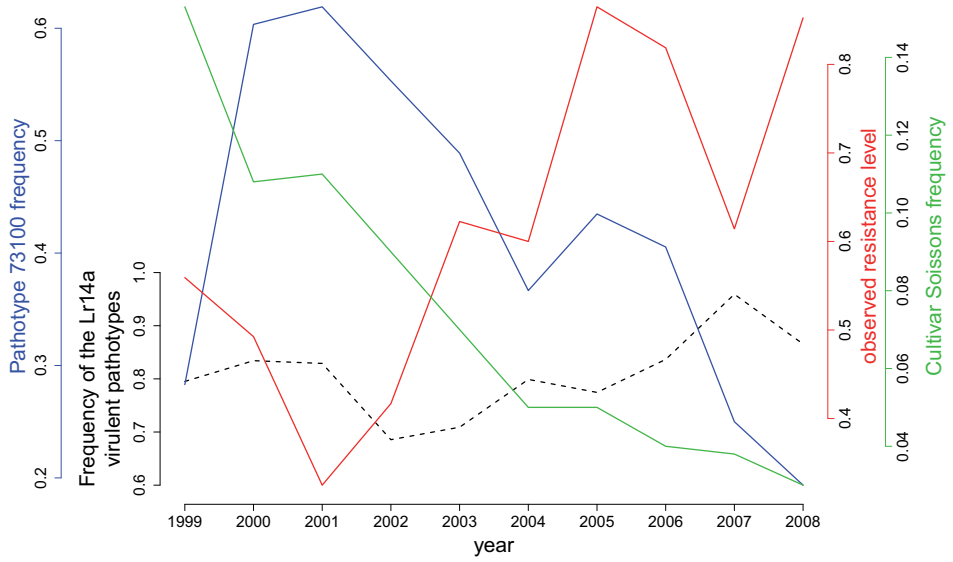
Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

The Soissons-073100 case



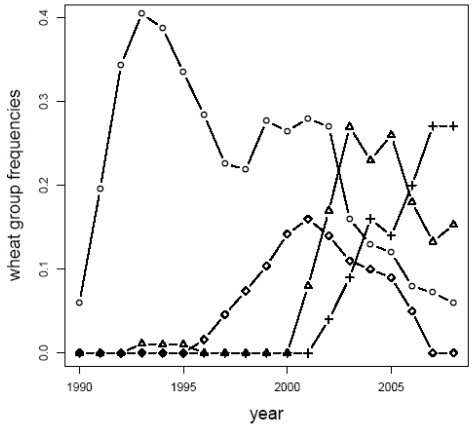
The Soissons-073100 case



Datasets

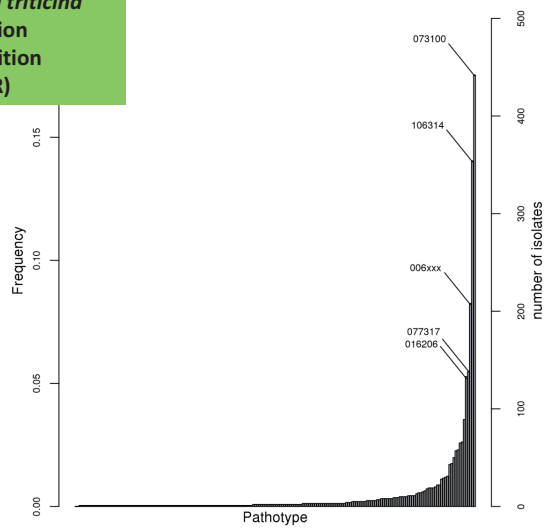
Papaïx et al., 2011, New Phytologist

Wheat landscape composition (FranceAgriMer)



Datasets

***Puccinia triticina* population composition (BIOGER)**



Wheat landscape composition (FranceAgriMer)

Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents

Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest

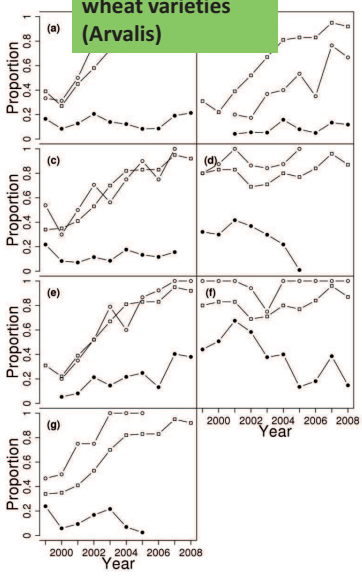
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population

General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

Datasets

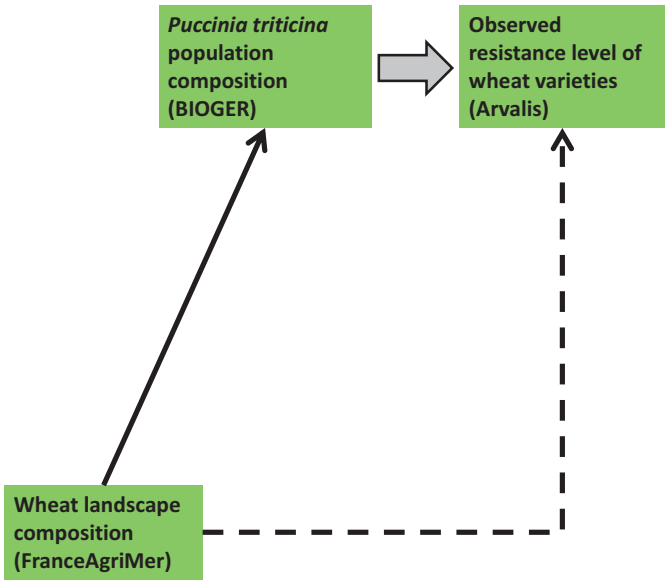
***Puccinia triticina* population composition (BIOGER)**

Observed resistance level of wheat varieties (Arvalis)

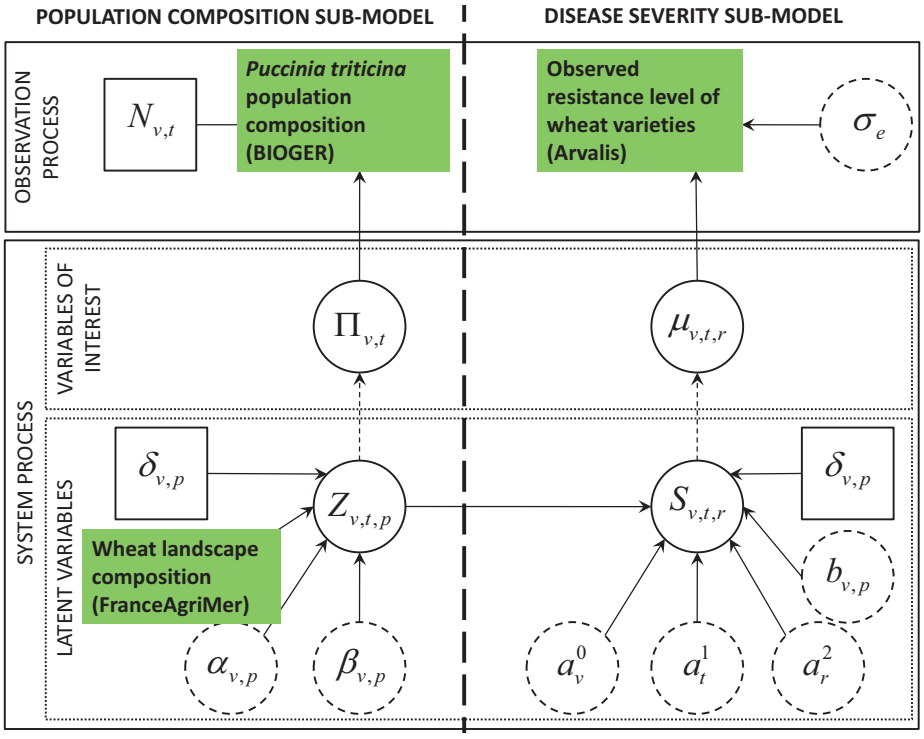


Wheat landscape composition (FranceAgriMer)

Datasets

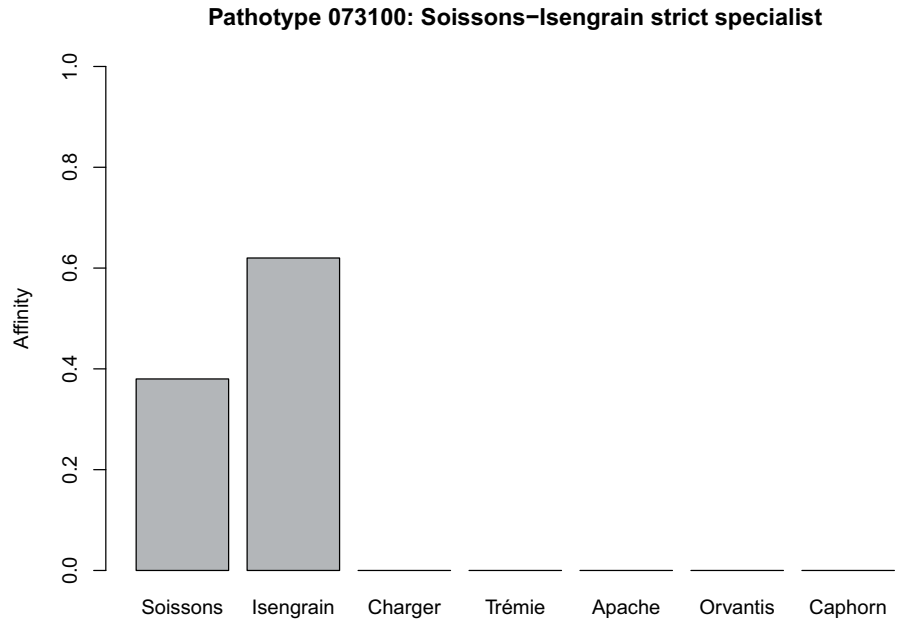


Model



Results

Pattern of ecological specialisation



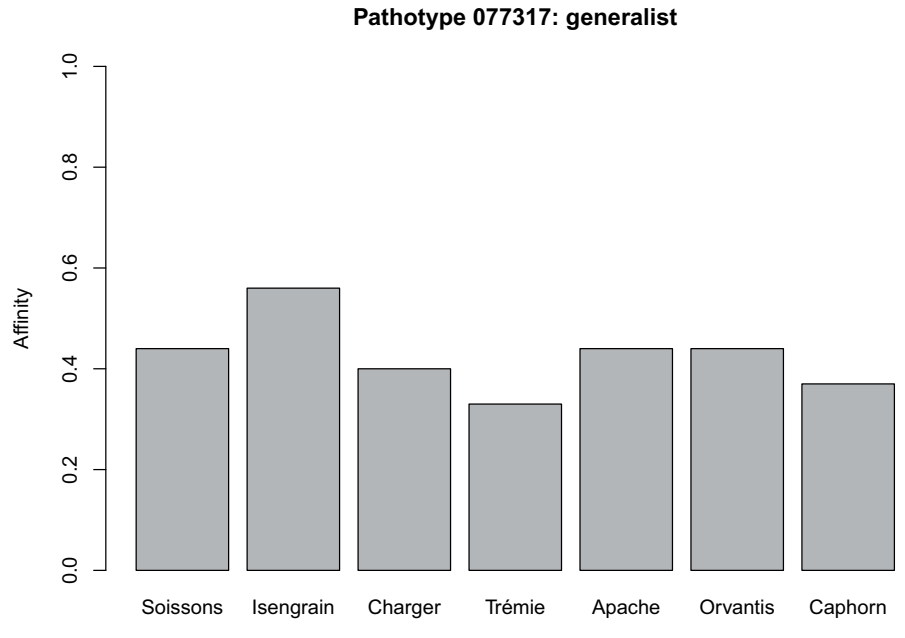
Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

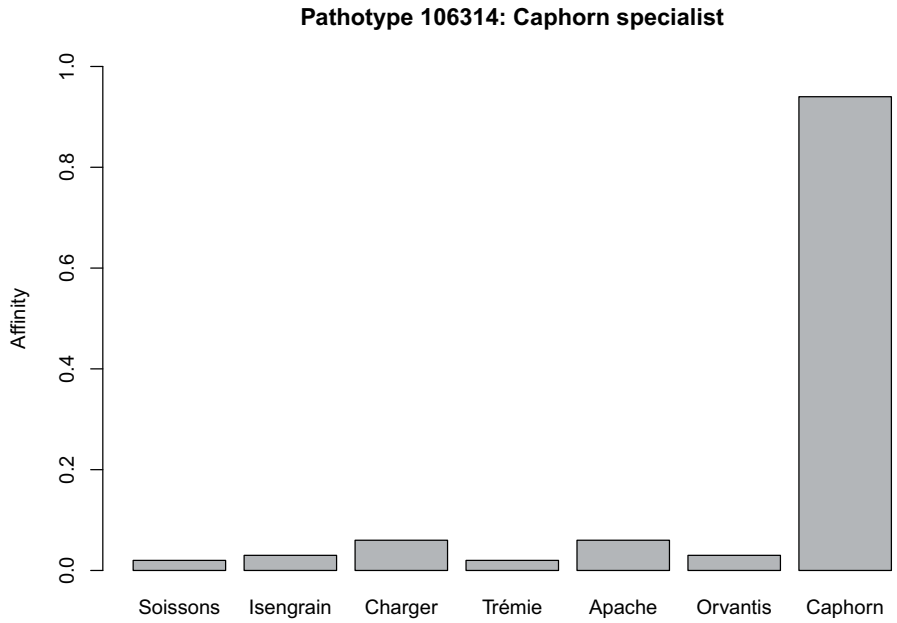
Results

Pattern of ecological specialisation



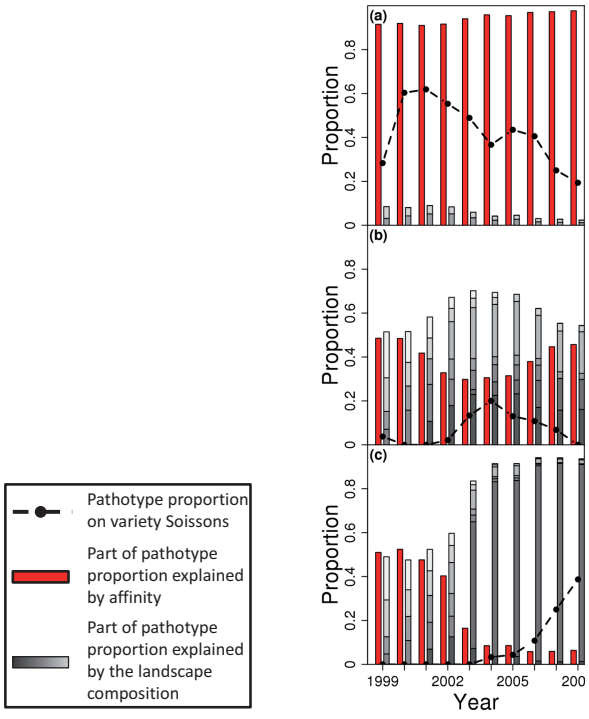
Results

Pattern of ecological specialisation



Results

Affinity *vs* landscape composition



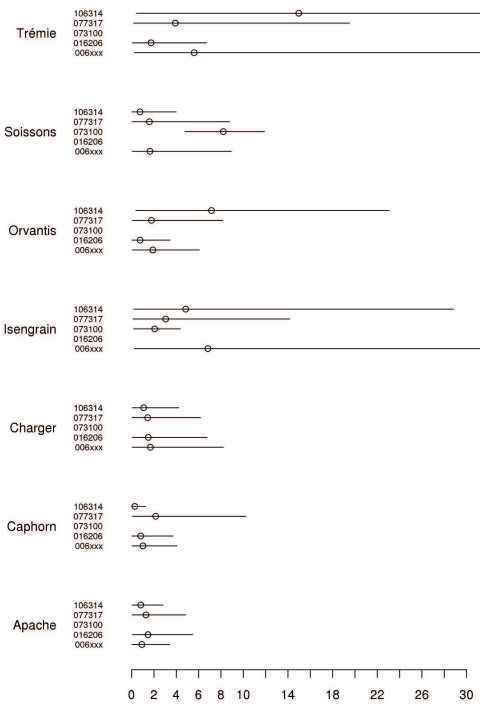
073100 proportion mainly explained by its affinity for Soissons.

077317 proportion explained both by its affinity for Soissons and the other varieties.

106314 proportion mainly explained by the presence of Caphorn in the landscape.

Results

Pathogen population composition and observed disease



The pathogen population composition living on a particular variety influences the observed disease level.

Major results

The landscape varietal composition influences the observed resistance level of the most grown wheat varieties by shaping *P. triticina* populations.

- Quantification of plant-pathogen interactions,
- spatial scale and conditions of production,
- towards predictive models for agricultural advice.

Questions of interest

1. **Dispersal was homogeneous and no explicit disease dynamics was investigated**
   $\Rightarrow$ How does the epidemic spread over the landscape when dispersal is limited?

Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

Questions of interest

1. **Dispersal was homogeneous and no explicit disease dynamics was investigated**
   $\Rightarrow$ How does the epidemic spread over the landscape when dispersal is limited?
2. **Composition of *P. triticina* population depended on affinity and landscape composition**
   $\Rightarrow$ Which are the respective roles of local fitness and migration in the competition between pathogen genotypes?

Questions of interest

1. **Dispersal was homogeneous and no explicit disease dynamics was investigated**
   $\Rightarrow$ How does the epidemic spread over the landscape when dispersal is limited?
2. **Composition of *P. triticina* population depended on affinity and landscape composition**
   $\Rightarrow$ Which are the respective roles of local fitness and migration in the competition between pathogen genotypes?
3. **Host spectrum was found highly variable (from highly specialized genotypes to generalist ones)**
   $\Rightarrow$ Which conditions favour the pathogen specialisation?

Contents

Modelling approaches

- **Simulation of pathogen population spread on an agricultural landscape:**
  1. spread of pathogen population,
  2. competition among pathogen genotypes.
- **Adaptive dynamics in a metapopulation context:**
  3. evolution of pathogen population.

Epidemics simulation

Epidemics simulation

What is an epidemic ?

- The environment (host landscape)

Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

Epidemics simulation

What is an epidemic ?

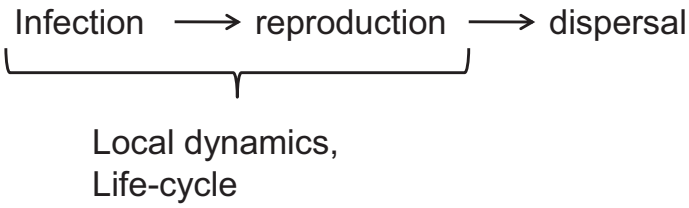
- The environment (host landscape)
- Spatio-temporal spread of a pathogen population:

$$\text{Infection} \longrightarrow \text{reproduction} \longrightarrow \text{dispersal}$$

Epidemics simulation

What is an epidemic ?

- The environment (host landscape)
- Spatio-temporal spread of a pathogen population:



Landscape structure and epidemic risk

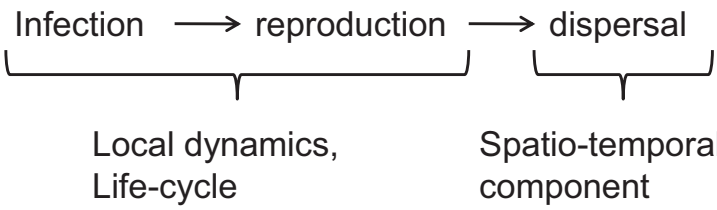
Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

Epidemics simulation

What is an epidemic ?

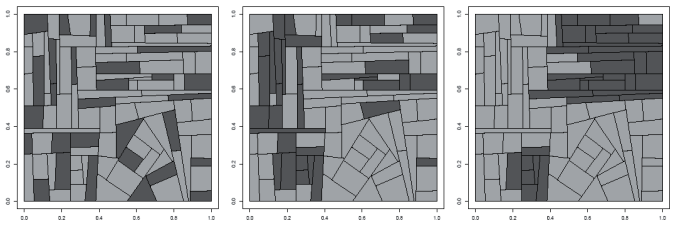
- The environment (host landscape)
- Spatio-temporal spread of a pathogen population:



Epidemics simulation

- The environment as a stochastic variable:

A patron



Same field pattern with an increasing aggregation level

Landscape structure and epidemic risk

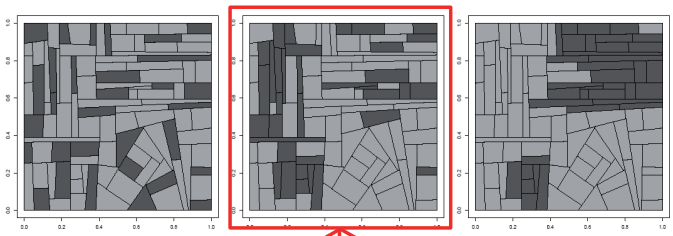
Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

Epidemics simulation

- The environment as a stochastic variable:

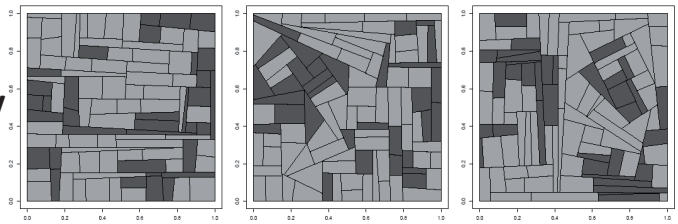
A patron



Same field pattern with an increasing aggregation level

+

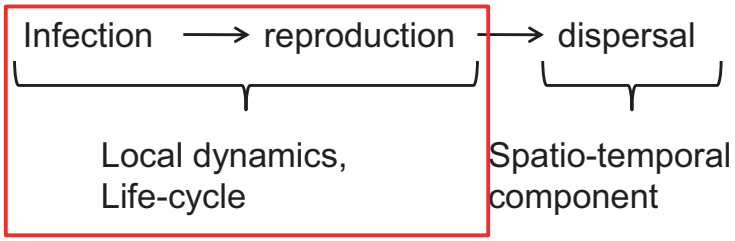
Variability around

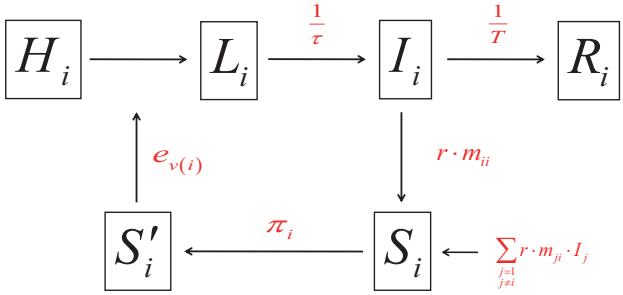


Same aggregation level with different field patterns and variety allocations

Epidemics simulation

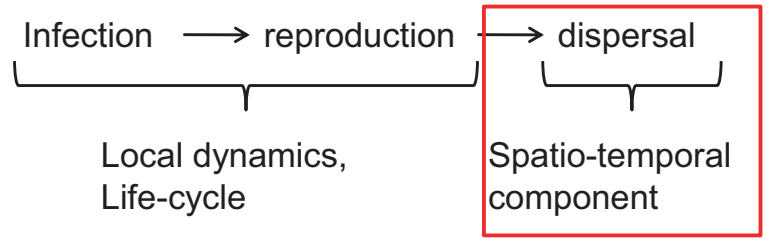
- Spatio-temporal spread of a pathogen population:

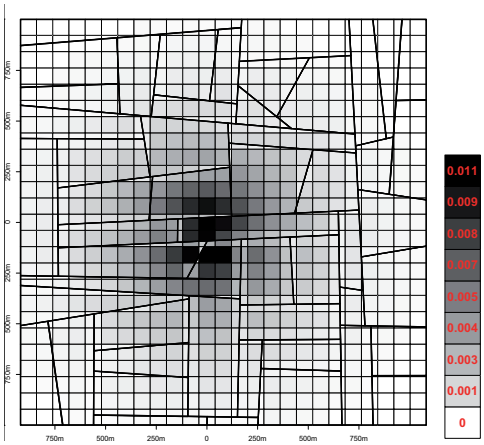




Epidemics simulation

- Spatio-temporal spread of a pathogen population:





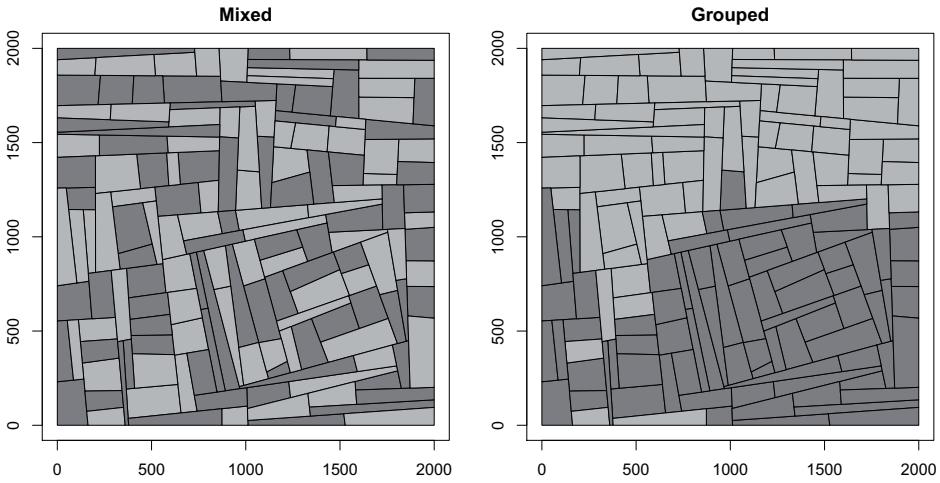
$$\mu(x, x') = \frac{2\pi}{\mu_0^2} \exp\left(\frac{2\pi}{\mu_0} \left\| x - x' \right\|\right)$$

1. Spread of pathogen population over a heterogeneous agricultural landscape

Spread of pathogen population

Effect of variety aggregation level

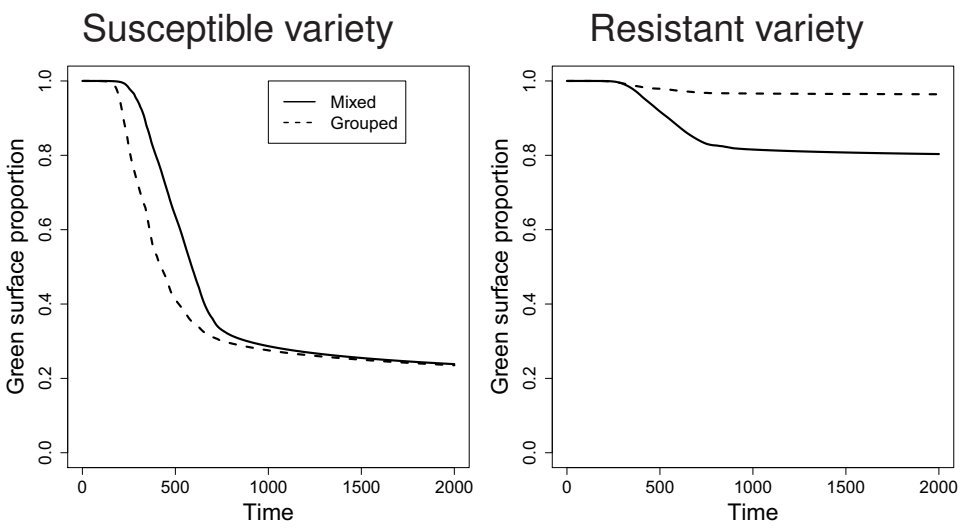
- One pathogen genotype.
- Two varieties: a susceptible and a partially resistant.



Papaïx et al., in prep.

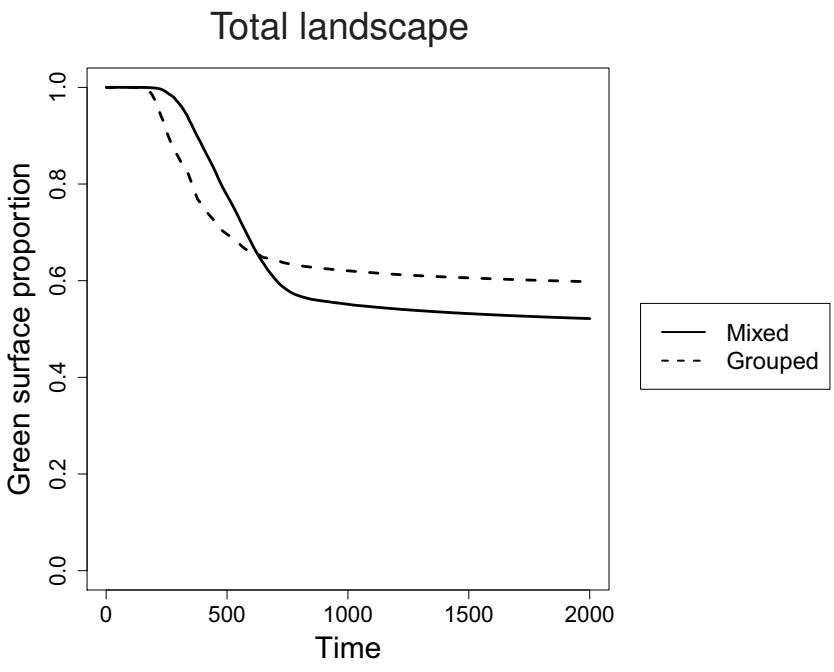
Spread of pathogen population

Effect of variety aggregation level



Spread of pathogen population

Effect of variety aggregation level



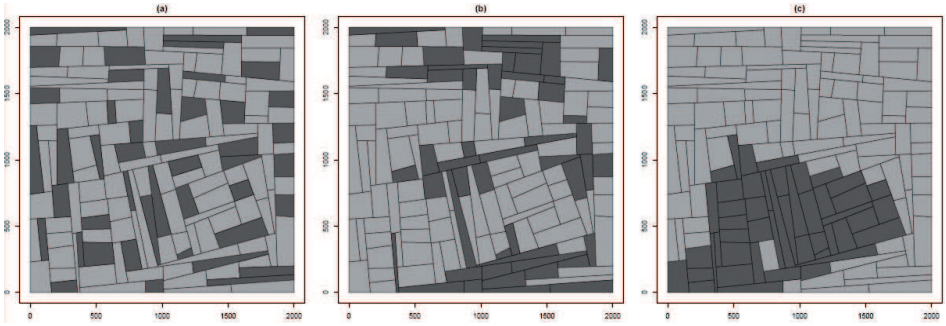
2. Competition among pathogen genotypes in a heterogeneous agricultural landscape

Competition among pathogen genotypes

Landscapes

- Two varieties: V_1 (30%) and V_2 (70%).
- Three aggregation levels:

Mixed strategy | Mosaic strategy | Grouped strategy

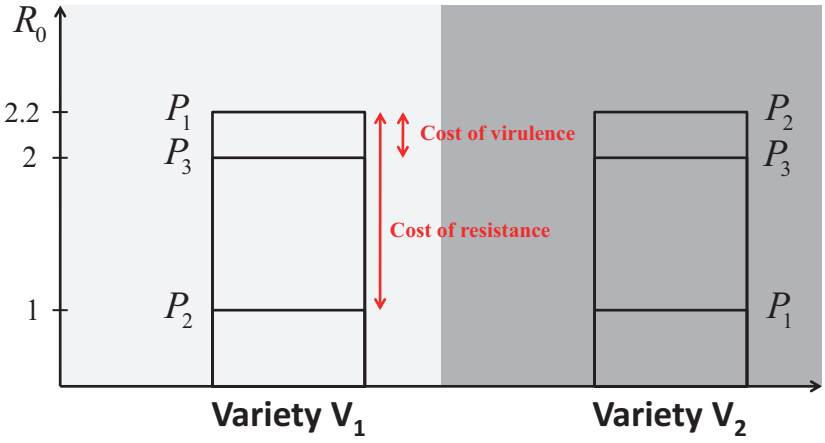


Competition among pathogen genotypes

Pathogen genotypes

Three pathogen genotypes:

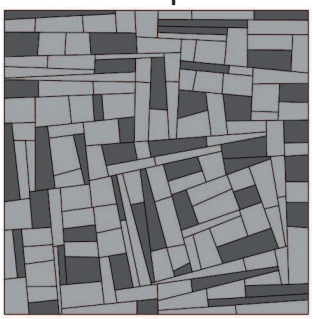
- V_1 specialist (P_1),
- V_2 specialist (P_2),
- generalist (P_3).



Competition among pathogen genotypes

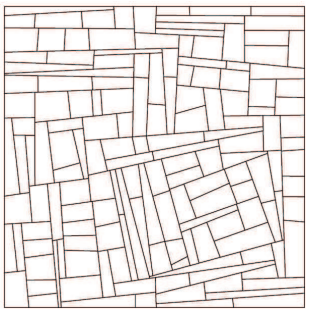
Mixed strategy

Landscape

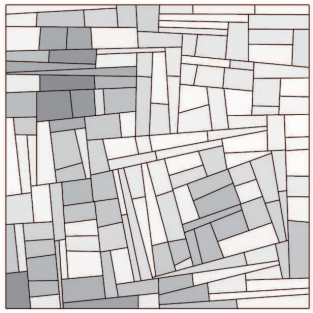


Generalist (P_3)

V_1 specialist (P_1)



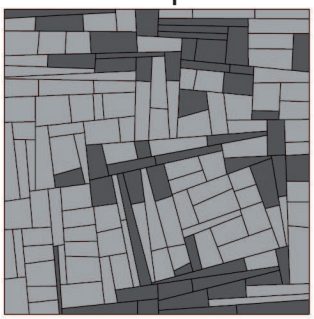
V_2 specialist (P_2)



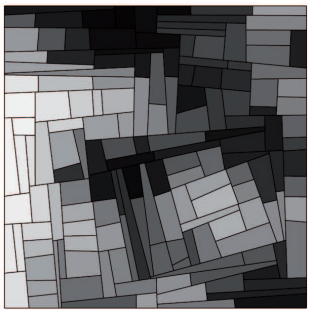
Competition among pathogen genotypes

Mosaic strategy

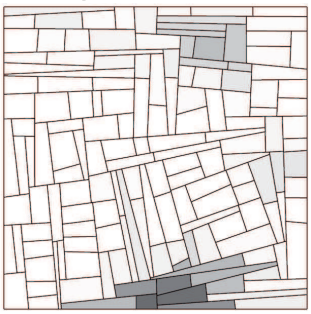
Landscape



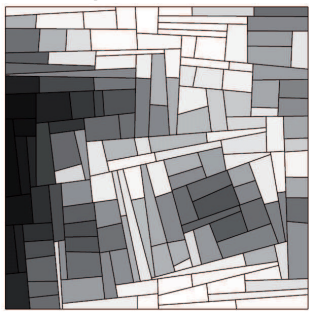
Generalist (P_3)



V_1 specialist (P_1)



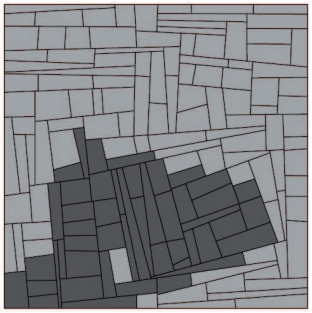
V_2 specialist (P_2)



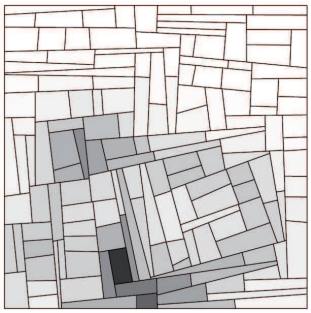
Competition among pathogen genotypes

Grouped strategy

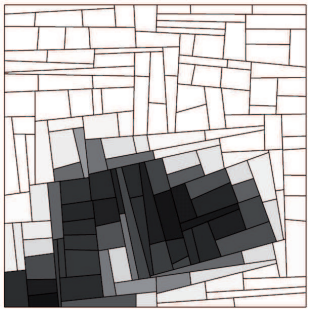
Landscape



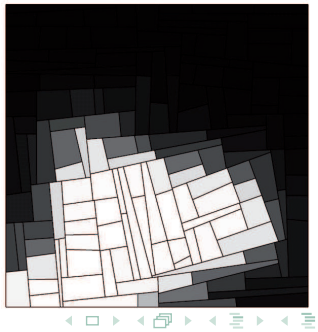
Generalist (P_3)



V_1 specialist (P_1)

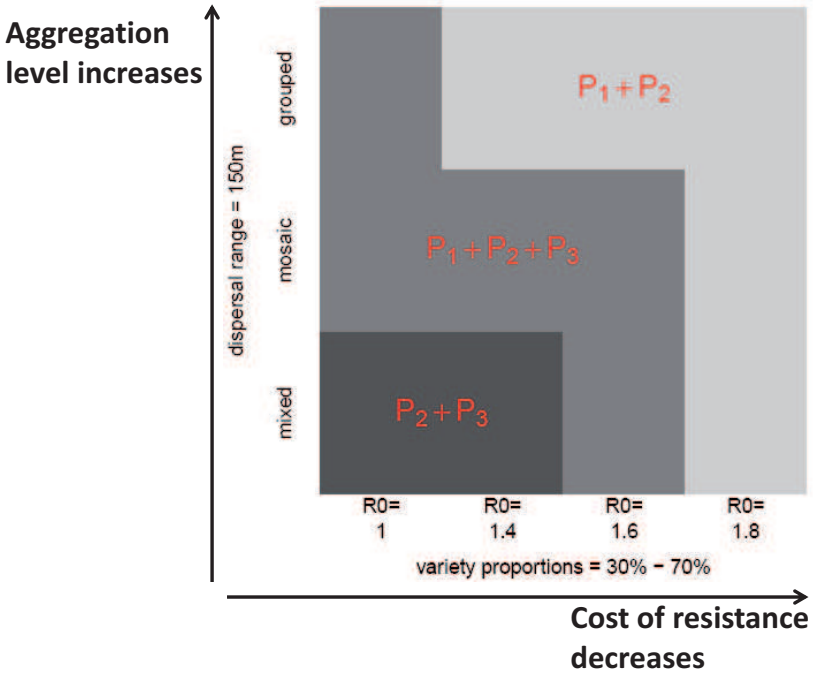


V_2 specialist (P_2)



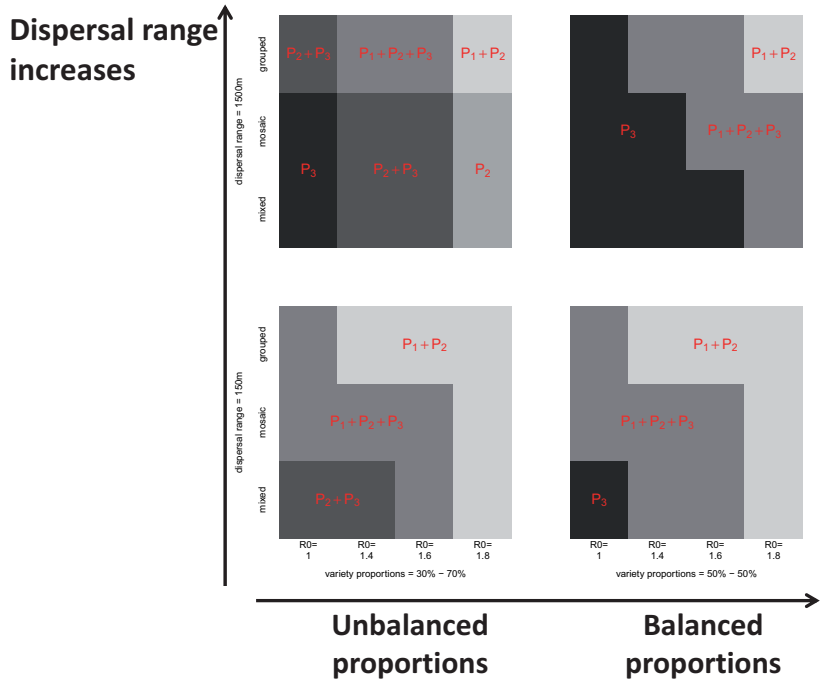
Competition among pathogen genotypes

Summary



Competition among pathogen genotypes

Summary



3. Evolution of specialisation in a spatially heterogeneous environment

Adaptive dynamics

What is it ?

A theoretical approach for studying phenotypic changes of an evolving population.

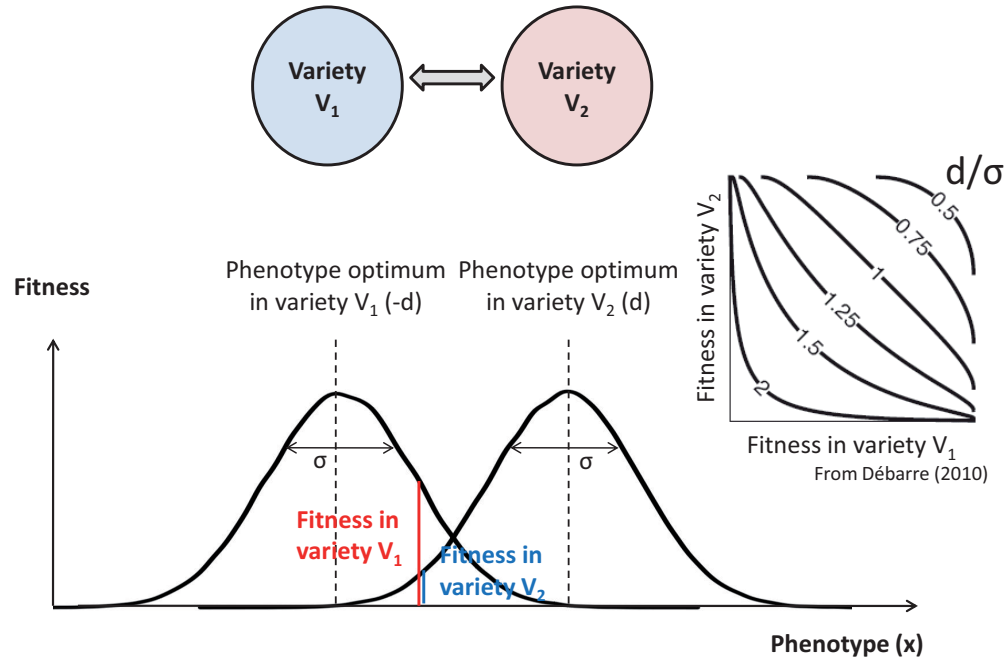
Main hypotheses:

- clonal reproduction,
- non-overlapping generations,
- phenotype characterized by a 1D continuous trait,

- mutants are rare and evolution is gradual,
- demographic equilibrium reached before the introduction of a new mutant.

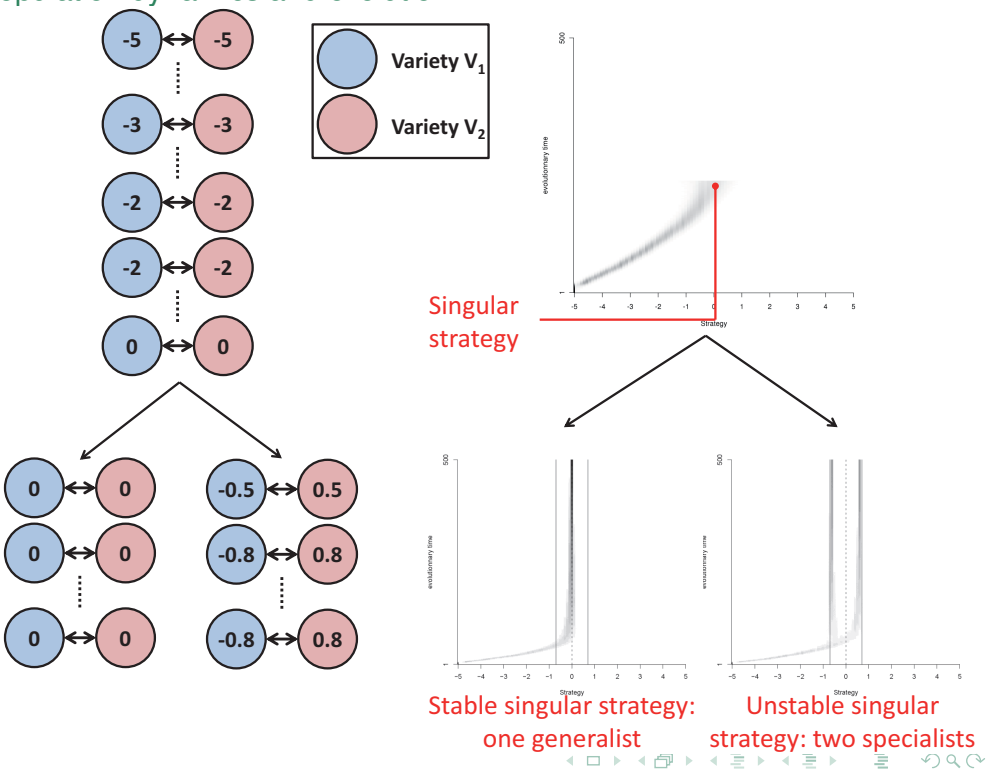
Adaptive dynamics

Fitness and trade-off



Adaptive dynamics

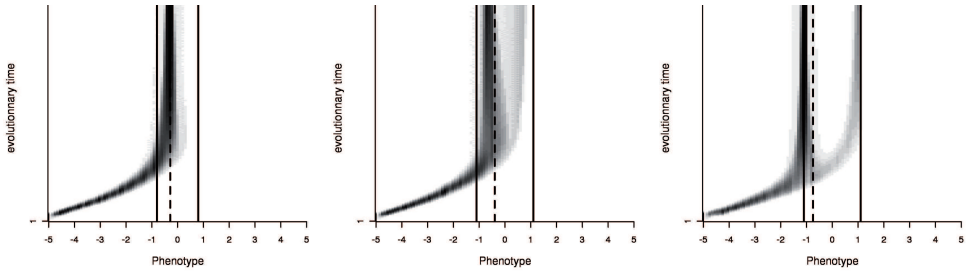
Population dynamics and evolution



Adaptive dynamics

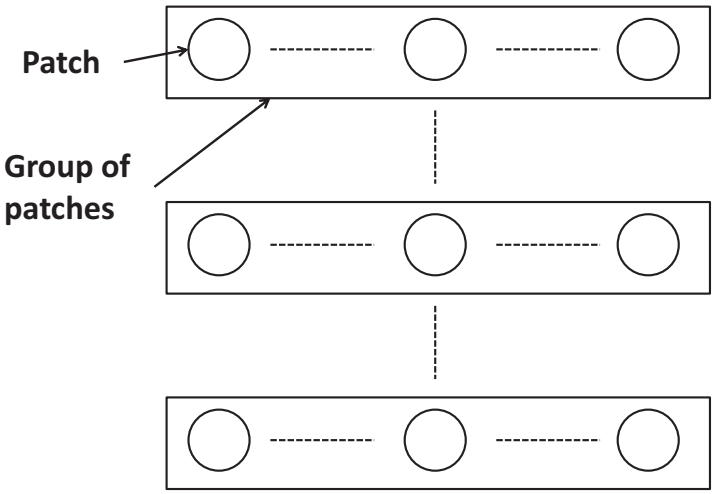
Branching criterion

Branching criterion < 1 | Branching criterion > 1 | Branching criterion $>> 1$



Evolution of pathogen population

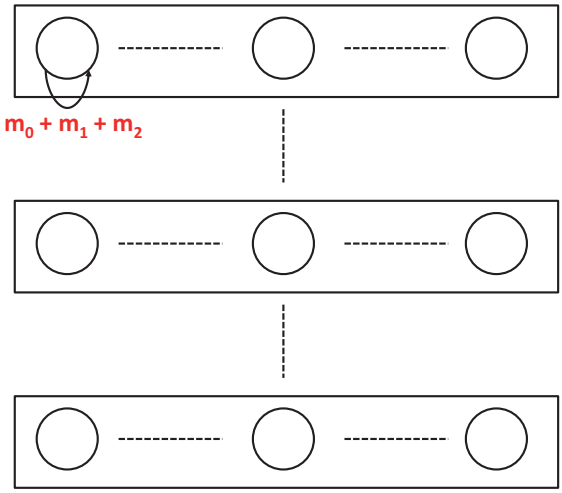
Decomposition of the branching criterion



Papaïx et al., in prep.

Evolution of pathogen population

Decomposition of the branching criterion



Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents

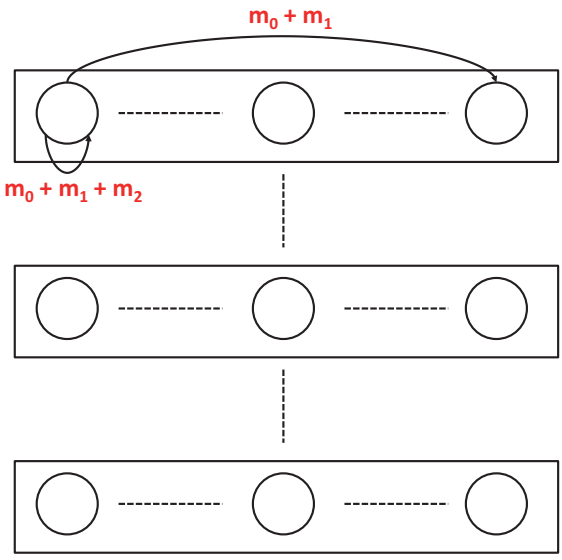
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest

Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population

General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

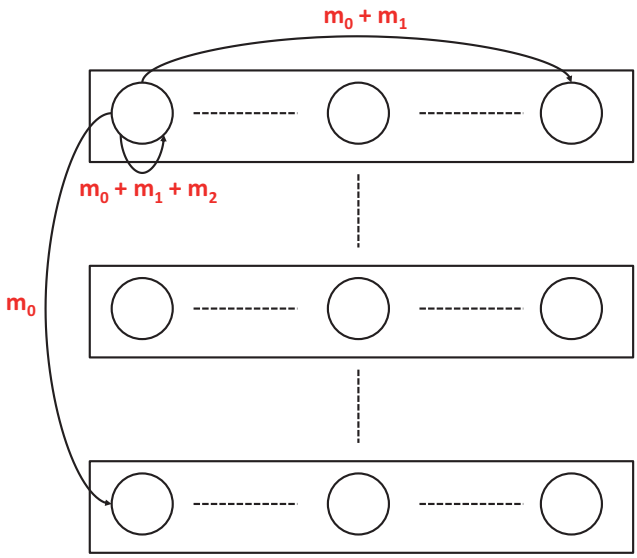
Evolution of pathogen population

Decomposition of the branching criterion



Evolution of pathogen population

Decomposition of the branching criterion



Evolution of pathogen population

Decomposition of the branching criterion

$$\sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\xi \sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\upsilon \sum_{k=1}^{H} \sum_{k'=1}^{H} v_{kk'} \frac{\left(x^* - \beta_k\right)\left(x^* - \beta_{k'}\right)}{\sigma^2} < 1$$

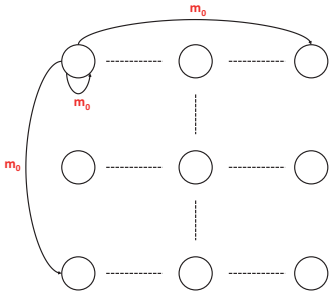
Evolution of pathogen population

Decomposition of the branching criterion

- homogeneous dispersal ($m_1 = m_2 = 0$)

$$\sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\xi \sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\upsilon \sum_{k=1}^{H} \sum_{k'=1}^{H} v_{kk'} \frac{\left(x^* - \beta_k\right)\left(x^* - \beta_{k'}\right)}{\sigma^2} < 1$$

Environment composition



Evolution of pathogen population

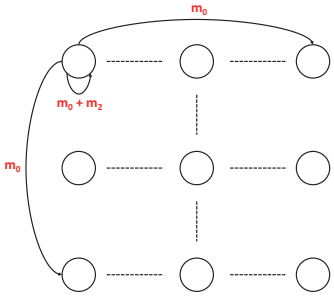
Decomposition of the branching criterion

- hindered dispersal ($m_1 = 0$)

Patch isolation

$$\sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\xi \sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\upsilon \sum_{k=1}^{H} \sum_{k'=1}^{H} v_{kk'} \frac{\left(x^* - \beta_k\right)\left(x^* - \beta_{k'}\right)}{\sigma^2} < 1$$

Environment composition



Evolution of pathogen population

Decomposition of the branching criterion

- group structure

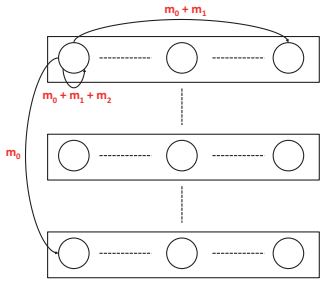
Patch isolation

Group isolation

$$\sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\xi \sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\upsilon \sum_{k=1}^{H} \sum_{k'=1}^{H} v_{kk'} \frac{\left(x^* - \beta_k\right)\left(x^* - \beta_{k'}\right)}{\sigma^2} < 1$$

Environment composition

Group compositions



Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents

Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest

Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population

General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

Evolution of pathogen population

Decomposition of the branching criterion

- group structure

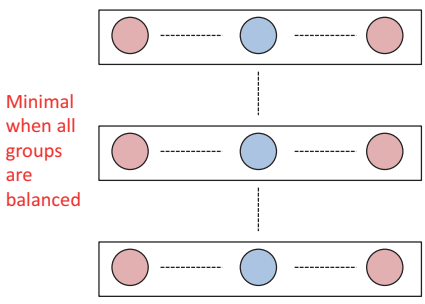
Patch isolation

Group isolation

$$\sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\xi \sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\upsilon \sum_{k=1}^{H} \sum_{k'=1}^{H} v_{kk'} \frac{\left(x^* - \beta_k\right)\left(x^* - \beta_{k'}\right)}{\sigma^2} < 1$$

Environment composition

Group compositions



Minimal when all groups are balanced

Evolution of pathogen population

Decomposition of the branching criterion

- group structure

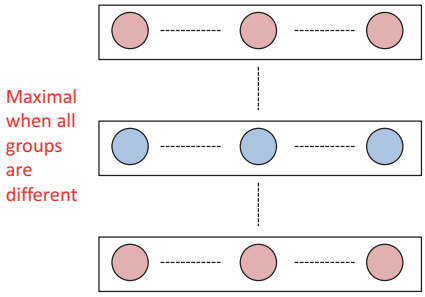
Patch isolation

Group isolation

$$\sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\xi \sum_{k=1}^{H} \pi_k \frac{\left(x^* - \beta_k\right)^2}{\sigma^2} + 2\upsilon \sum_{k=1}^{H} \sum_{k'=1}^{H} v_{kk'} \frac{\left(x^* - \beta_k\right)\left(x^* - \beta_{k'}\right)}{\sigma^2} < 1$$

Environment composition

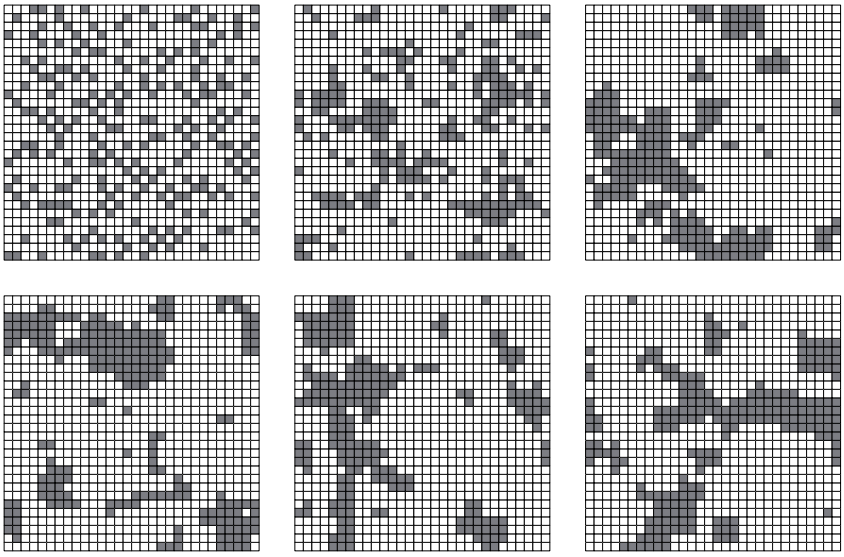
Group compositions



Evolution of pathogen population

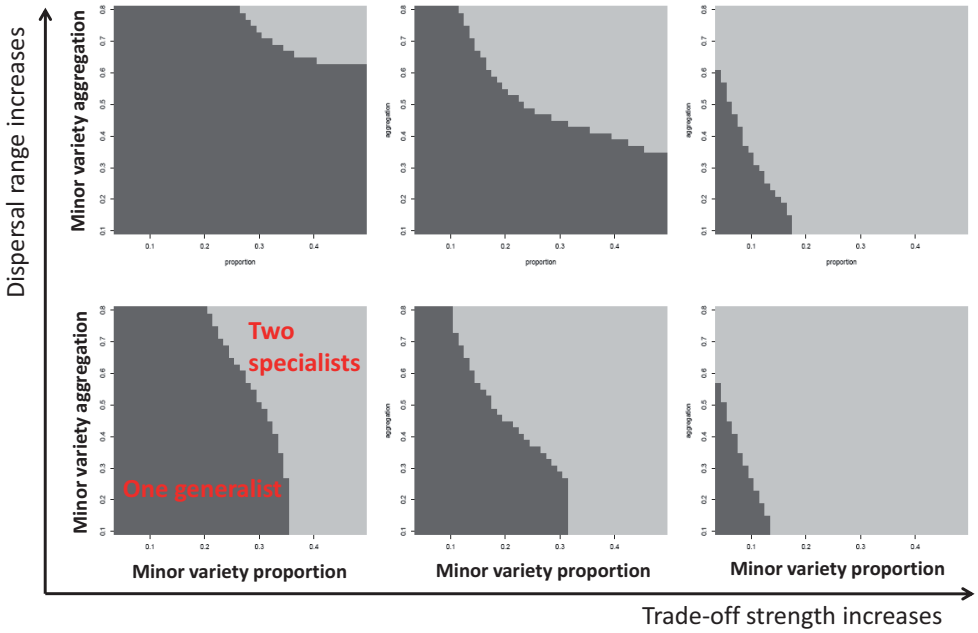
Landscape effect on specialisation

- Two varieties (dark grey: minor variety)
- Different proportion and aggregation level



Evolution of pathogen population

Landscape effect on specialisation



Contents

Agro-ecological conclusions

Host spatial structure influences

- Landscape connectivity
- Exposure of resistant varieties

Question 1

- Mutant establishment
- Competition

Question 2

- Host spectrum
- Level of local adaptation
- Evolutionary speed

Question 3

Landscape structure and epidemic risk

Julien Papaïx

Introduction
Agricultural context
Ecological questions
Spatial heterogeneity
Contents
Observations on wheat leaf rust
The wheat leaf rust pathosystem
The Soissons-073100 case
Generalisation
Questions of interest
Host diversity and epidemic risk
Modelling approaches
Spread of pathogen population
Competition among pathogen genotypes
Long term evolution of pathogen population
General discussion
Agro-ecological conclusions
Methodological conclusions
Perspectives

Agro-ecological conclusions

High global and local diversity

Host spatial structure influences

- Landscape connectivity → Impede the spread of epidemics.
- Exposure of resistant varieties → Expose RV to a higher inoculum but implies a progressive erosion of R.

Question 1

- Mutant establishment → Lead to diffuse mutant populations.
- Competition → Enhance inter-genotypes competition (limit coexistence).

Question 2

- Host spectrum → Favour generalists.
- Level of local adaptation → Decrease the level of local adaptation.
- Evolutionary speed → Slow down the evolution of pathogen populations.

Question 3

Methodological conclusions

Effect of host spatial structure on pathogen ecology and evolution for controlling epidemics in agricultural systems

- Data analysis (Bayesian inference)
- Simulation models (numerical experiment and sensitivity analysis)
- Analytical approaches

$\Rightarrow$ role of data *vs* theoretical approaches,
$\Rightarrow$ complex ecological modelling.

Perspectives

- Quantitative and/or qualitative **validation** of theoretical models via large scale experiments.
- **Towards agricultural advice:**
  - temporal dimension (internship F. Claeys 2011),
  - socio-economical models, collective strategies,
  - interactions between practices and parasites, efficiency of diversification strategies,
  - sexual recombination,...
- **Towards natural systems:** plant-pathogen coevolution.

Thanks to everyone who contributed to this PhD !

Thank you for your attention !

Non, non, anuèit vòli fugir l'ostal !
Vòli lo fial de fum que s'estira suls camps
Quand lo lauraire aluca un fuòc d'erbassas.
O fial de fum, vèni ligar un raive,
Un raive que m'escapa
-coma tu, fial de fum-
Per fugir cap a las estelas.

Loisa Paulin, Fum (Sorgas, 1940)

RESUME

L'intensification de l'agriculture a amélioré de façon considérable la production alimentaire ces dernières cinquante années mais elle s'est accompagnée d'un impact croissant sur l'environnement. En particulier, la modernisation de l'agriculture a impliqué une simplification de la structure des paysages agricoles rendant nos agro-écosystèmes plus sensibles au risque épidémique. L'utilisation de la diversité génétique des cultures est une solution prometteuse pour réduire le risque d'occurrence et de propagation des maladies des cultures. Elle nécessite cependant une gestion collective des espaces agricoles. En conséquence, l'échelle d'étude ne doit plus se focaliser sur la parcelle mais sur le paysage. Dans cette thèse, nous nous interessons aux processus se déroulant à l'échelle du paysage et au rôle de la diversité des plantes cultivées pour le contrôle des épidémies. Nous avons identifié trois questions: comment les populations pathogènes se propagent-elles dans un paysage d'hôtes hétérogène ? Comment les différents génotypes composant la population pathogène entrent-ils en compétition au sein d'une population hôte diversifiée ? et, à plus long terme, comment les populations pathogènes évoluent-elles en réponse à la structure des populations hôtes ? Chacune de ces questions a été approfondie grâce à l'analyse de données obtenues en condition de production mais aussi par des approches théoriques. Nous avons montré que la composition et la structure spatiale des populations hôtes influence fortement la population pathogène. Cependant, les recommandations que peut fournir ce travail pour gérer la diversité génétique dépendent de l'objectif visé.

Evolution | Epidémiologie du paysage | Hétérogénéité spatiale | Spécialisation

ABSTRACT

Agriculture intensification has improved food production impressively in the past 50 years but this came with an increasing impact on the environment. In particular, modern agriculture has led to the simplification of the environmental structure over vast areas. As a consequence, agro-ecosystems are particularly susceptible to epidemics. The increase of crop genetic diversity is a promising way for reducing the risk of occurrence and development of diseases in crops but the technical and organisational conditions required to manage the genetic resources at this scale have not been established yet. This will require shifting the scale of crop protection investigations from the field to the agricultural landscape. In this PhD thesis we focus on landscape-scale processes and on the potential role of functional diversity in cultivated landscapes to better control plant diseases. We identified three questions: how does a pathogen population spread over a heterogeneous host landscape? How do pathogen genotypes compete in a diversified host population? And, in a longer term, how do pathogen populations evolve in response to host landscape structure? Each of these questions is investigated through the analysis of real data and the development of theoretical approaches. We demonstrate that the composition and the spatial structure of the host landscape greatly influence the pathogen population dynamics and evolution. The recommendations that this work may provide in order to practically manage the genetic resources will depend on the desired aim and will request further collaborative work with the professional operators.

Evolution | Landscape epidemiology | Spatial heterogeneity | Specialisation