



**HAL**  
open science

# A modeling framework for the development of eco-efficient control strategies for airborne pathogens of fruit plants: the case of brown rot and peaches

Andrea Radici

## ► To cite this version:

Andrea Radici. A modeling framework for the development of eco-efficient control strategies for airborne pathogens of fruit plants: the case of brown rot and peaches. Agronomy. Université d'Avignon Pays du Vaucluse, 2023. English. NNT : 2023AVIG0617 . tel-04443924

**HAL Id: tel-04443924**

**<https://theses.hal.science/tel-04443924>**

Submitted on 7 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**THÈSE DE DOCTORAT D'AVIGNON UNIVERSITÉ**

**École Doctorale 536**  
**Agrosciences et sciences**

**INRAE PACA – Unité BioSP**

**Spécialité de doctorat :**  
Sciences Agronomiques

Présentée par  
**Andrea Radici**

---

**Un cadre de modélisation pour le  
développement des stratégies eco-  
efficientes de lutte contre les agents  
pathogènes aériens des plantes**

*Les cas de la rouille noire du blé  
et de la moniliose des pêchers*

---

Soutenue publiquement le 11/12/2023 devant le jury composé de :

Virginie RAVIGNÉ (HDR)	CIRAD (PHIM)	<b>Rapporteuse</b>
Thibaud PORPHYRE (DR)	Université de Lyon 1 (LBBE)	<b>Rapporteur</b>
Cindy MORRIS (DR)	INRAE (PV)	<b>Examinatrice</b>
Suzanne TOUZEAU (CR)	INRAE (ISA)	<b>Examinatrice</b>
Daniele BEVACQUA (HDR)	INRAE (PSH)	<b>Directeur de thèse</b>
Davide MARTINETTI (CR)	INRAE (BioSP)	<b>Co-directeur de thèse</b>



# Résumé

L'émergence de pathogènes végétaux s'accélère dans le monde entier, menaçant la sécurité alimentaire. Il est urgent de concevoir des outils innovants de protection des plantes, en associant des stratégies de surveillance à des stratégies de lutte précoce contre les maladies, afin de garantir la sécurité alimentaire tout en assurant la viabilité environnementale des pratiques agricoles. Dans ce contexte, les pathogènes aériens représentent un enjeu majeur, car ils peuvent se propager sur de longues distances.

Dans cette thèse, je propose un cadre de modélisation pour concevoir des stratégies basées sur les réseaux complexes pour la surveillance et le contrôle des pathogènes des plantes se déplaçant dans les masses d'air. Je prends comme sujets d'étude le pathogène fongique *Puccinia graminis*, agent causal de la rouille noire du blé, et *Monilinia fructicola*, agent causal de la pourriture brune des pêches.

Tout d'abord, je présente la modélisation des interactions entre hôte et pathogène sous forme de réseaux complexes et je retrace la manière dont les scientifiques ont utilisé les propriétés des réseaux pour élaborer des stratégies de protection contre les maladies. Ensuite, je reconstruis le réseau épidémique mondial de la rouille noire du blé, où les régions productrices de blé sont reliées par des connexions potentielles de transport de pathogènes. J'estime ces connexions à l'aide d'un modèle aérobiologique basé sur des simulations de trajectoires Lagrangiennes. Je montre comment un algorithme basé sur les réseaux complexes peut aider à identifier les meilleures sentinelles, c'est-à-dire les endroits où une épidémie peut être détectée rapidement. Troisièmement, j'intègre le modèle aérobiologique dans un cadre épidémiologique métapopulationnel afin de simuler la propagation spatiale d'une épidémie de maladie des plantes. En particulier, je couple un modèle dépendant du climat décrivant la susceptibilité de l'hôte et son épidémiologie à l'intérieur des vergers avec des simulations de trajectoires Lagrangiennes déterminant les transports des pathogènes entre les vergers. J'utilise la pourriture brune des pêches en France comme cas d'étude, pour lequel je produis des cartes de risque épidémiologique afin de faciliter le développement des stratégies de protection. Enfin, j'évalue la perte globale d'efficacité de la surveillance due au manque de communication et de coordination entre les pays dans le cas des maladies transfrontalières. J'utilise le réseau épidémique mondial de la rouille du blé comme cas d'étude. J'évalue les efforts de surveillance (c'est-à-dire le nombre de sentinelles) déployés par chaque pays pour

atteindre un objectif de surveillance donné dans un scénario de coopération (c'est-à-dire sans tenir compte des frontières) puis dans un scénario où chaque pays est indépendant.

Compte tenu de la forte densité du réseau épidémique mondial de *Puccinia*, il est possible de trouver un ensemble restreint de sentinelles (1% du réseau) qui surveillent indirectement la moitié des régions productrices de blé (50% du réseau). Je démontre que la connectivité basée sur les masses d'aire aide la reconstruction des observations de l'incidence de la pourriture brune en France, et identifie les endroits les plus à risque dans la vallée du Rhône. Les avantages d'une stratégie coopérative, qui interprète correctement l'échelle de dispersion de la maladie, sont évidents pour les maladies transfrontalières, mais sont répartis de manière hétérogène entre les pays : des mécanismes de compensation devraient être mis en œuvre pour obtenir un soutien unanime en faveur d'un système de surveillance coopératif international.

# Abstract

Rates of emergence of pathogens accelerates worldwide, threatening food security. Designing innovative crop protection tools, coupling surveillance strategies with timing control of diseases, is a main issue to ensure food security while restoring environmental viability of farming practices. In this context, airborne pathogens represent a challenging case study, as they may be spread over long distances.

In this thesis I propose a modelling framework to support network-based surveillance and control strategies for airborne plant pathogens. I consider the fungal pathogen *Puccinia graminis*, causal agent of stem rust of wheat, and *Monilinia fructicola*, causal agent of brown rot of peaches, as case studies.

First, I introduce the representation of host-pathogens interactions in networks, and I investigate how scientists have used network properties to elaborate disease protection strategies. Secondly, I reconstruct the global epidemic network for stem rust of wheat, where worldwide wheat-producing regions are connected via potential pathogen transport events. I compute such connections using an aerobiological model based on Lagrangian trajectory simulations. I show how a network-based algorithm may help identifying the best sentinels, i.e. those locations, to be frequently monitored, where an outbreak may be early detected. Thirdly, I integrate the aerobiological model within an epidemiological-metapopulation framework to simulate the spatial spread of a plant disease outbreak. In particular, I couple a climate-dependant model describing host susceptibility and in-orchard epidemiology with Lagrangian trajectory simulations determining transport of pathogens between orchards. I use brown rot of peach crops in France as a case study, for which I produce maps of epidemiological risk to inform protection strategies. Finally, I assess the global loss of surveillance efficiency due to lack of communication and cooperation among countries in designing surveillance strategies for transboundary diseases. I use the global epidemic network for stem rust of wheat as a case study. I assess the surveillance efforts (i.e. the number of sentinels) deployed by each country to reach a given monitoring target in a cooperative scenario (i.e. regardless of country border) or each country alone.

Given the high density of the worldwide *Puccinia* epidemic network, it is possible to find a narrow set of sentinels (1% of the network) which indirectly monitor half of the wheat-producing regions (50% of the network). I show that wind-driven connectivity helps recon-

structuring observations of brown rot incidence in France, and help identifying the most risky locations in the Rhône Valley. Benefits of a cooperative strategy, which correctly interpret the dispersal scale of a disease, are evident for transboundary diseases, but are heterogeneously distributed among countries: compensation mechanisms should be implemented to gain unanimous support for an international cooperative surveillance system.

# Acknowledgements

I do acknowledge the support of funding from the French National Research Agency (ANR) for the BEYOND project (contract # 20-PCPA-0002) and SuMCrop Sustainable Management of Crop Health Program of INRAE that supported the work of all authors.



# Remerciements

Bien des années plus tard, face au peloton d'exécution, le colonel Aureliano Buendia devait se rappeler ce lointain après-midi au cours duquel son père l'avait emmené découvrir la glace.

– Gabriel García Márquez, *Cent ans de solitude*

*Cette thèse a marqué un changement soudain dans ma vie. J'ai quitté mon emploi précédent, la ville et le pays où je vivais, pour m'engager dans la recherche scientifique dans un autre pays, répondant ainsi à mon désir de me mettre à l'épreuve. J'attendais cette thèse depuis le 14 septembre 2018, quand Daniele m'a proposé le sujet lors du Congrès de la Société Italienne d'Écologie à Cagliari, jusqu'au 15 février 2021, date à laquelle j'ai signé le contrat doctoral. Cependant, cette attente n'a pas été immédiatement récompensée. Le confinement, le manque de repères et la frustration liée aux difficultés linguistiques étaient plus forts que tout le reste. Trois mois après le début de ma thèse, j'ai appelé Dario, qui se trouvait à l'époque aux Pays-Bas après avoir terminé sa thèse à Avignon. Je lui ai dit, grosso modo:*

*– J'en ai marre de tout ça. Je suis entouré de gens ravis d'être là, mais qui n'ont aucune empathie envers ma situation. J'ai besoin de me défouler.*

*– Je comprends – il m'a répondu. Je crois qu'il préparait le dîner en même temps – Avignon est une ville petite, l'hiver c'est la mort, le Mistral tue, littéralement, les transports avec Milan sont mauvaises, les Français sont un peu têtus et ne font pas beaucoup d'efforts pour parler anglais\*.*

*Il a réfléchi silencieux deux seconds - je crois qu'il a goûté un petit peu de soupe de sa cuillère - et il a continué:*

*– Mais je serais injuste si je disais que ce n'a pas été la meilleure expérience de ma vie. Et je suppose que cela sera pareil pour toi.*

*Avait-il tort? Je laisse la réponse au lecteur, après avoir lu ce chapitre.*

Merci à mes encadrants, qui n'ont jamais abandonné l'espoir de trouver une maudite bourse pour financer ma thèse. En plus, merci à Davide pour sa présence constante et son esprit éclectique, et à Daniele pour sa rigueur scientifique et son regard pointu.

Merci à ceux qui m'ont accueilli et qui m'ont fait me sentir chez moi : mes colocataires Carla et Elena, le groupe des *ritals* composé, parmi les autres, d'Enrico, Marta, Matteo,

---

\*Pour les amis français de Dario : il disait ça pour montrer de l'empathie envers moi, non pas parce qu'il le pensait. Pour tous les autres: ah ah lol.

Dario, Patrizia.

Merci aux membres du laboratoire qui m'ont accompagné, avec une mention particulière pour Sylvie et Amélie – si elles étaient chercheuses, leur *h-index* serait d'environ 60 et 40 – et aux directeurs d'unité : Samuel, Lionel, Edith.

Merci à ceux qui ont traversé les mêmes péripéties avec moi, les *BioSP kids*, avec une mention particulière pour Bartholomé, Dorian, Manon et Jean-Baptiste.

Merci à mon comité de suivi, avec une mention particulière pour Cindy pour son énergie déclenchante.

Merci à Nik Cunniffe (et au groupe d'épidémiologie numérique de l'Université de Cambridge) pour avoir été une source d'inspiration lorsque je tournais en rond.

Merci à ceux qui ont semé les graines de la curiosité dans mes premières 30 années de vie, parmi eux : Patrizio, Riccardo, Francesco, *la prof* Brivio, *la prof* Tranquillo, Paco, Giulia.

Et je ne pourrais pas ne pas remercier ma *mamma* qui a cultivé avec moi le rêve de mon avenir à Avignon.

# Contents

Introduction	1
1 Network-thinking to optimize surveillance and control of crop parasites. A review	11
2 Early-detection surveillance for stem rust of wheat: insights from a global epidemic network based on airborne connectivity and host phenology	36
3 A metapopulation framework integrating landscape heterogeneity to model an airborne plant pathogen: the case of brown rot of peach in France	60
4 Global benefits and domestic costs of a cooperative surveillance strategy to control transboundary crop pathogens	80
Conclusion and perspectives	99
References	108
A Supplementary Information - Chapter 2	128
B Supplementary Information - Chapter 3	159
C Supplementary Information - Chapter 4	174

# Introduction

The story so far:

In the beginning the Universe was created.

This has made a lot of people very angry and been widely regarded as a bad move.

– Douglas Adams, *The Restaurant at the End of the Universe*

## Eating

Food, together with breathing, water, sex, sleep, homeostasis and excretion are physiological fundamental human needs (McLeod, 2007). In the context of motivational theory, these represent the primary needs that must be fulfilled initially to further satisfy a hierarchy of more complex needs which contribute to the idea of personal development, i.e. safety, love and belonging, esteem, and self-actualization.

Broadening the perspective from humans to the whole living world, heterotrophs, and humans among them, take nutrition - mainly energy and carbon - by eating or consuming food (Stuart, 2013). Nutrition is such a fundamental activity among living beings that ecologists classify organisms based on their primary source of nourishment. One first classification divides organisms into primary producers (such as plants, which produce biomass from inorganic compounds), primary consumers (which feed on primary producers), secondary consumers (which feed on primary producers and consumers). This classification, which is called food chain (Elton, 1927), has been revisited several times to include decomposers and scavengers, such as fungi, which feed on complex organic substrate and return carbon to the surrounding environment (Sheldrake, 2021), eventually proposing a circular shape (a food web) rather than a linear one. Although imperfect, this linear representation already allows us a crucial observation: our nutrition, fulfilling a primary physiological fundamental human need, depends ultimately on primary producers.

This is why this thesis concerns plants and their health.

## Agriculture, parasites and society

Historians and archaeologists agree in stating that a fundamental step in the evolution of human societies was the domestication of plants (Harari, 2014). Measures of genetic similarity support the hypothesis that this domestication, also called Agricultural Revolution, happened between 11,000 and 3,000 years ago in distinct parts of the Earth: the Fertile Crescent in Middle East (9,000 BC), the Yangtze and Yellow River basins in Southeast Asia (7,000 BC) and the New Guinea Highlands (7,000 – 5,000 BC), Central Mexico in Central America (3,000 – 2,000 BC), Northern South America (3,000 – 2,000 BC), sub-Saharan Africa (3,000 – 2,000 BC), eastern North America (2,000 – 1,000 BC; Diamond and Bellwood, 2003). Prior to this event, it is widely acknowledged that humans lived in nomad groups engaged in hunting and gathering (Diamond and Ordunio, 1999). Agriculture marked a shift toward a sedentary way of life, at the point that anthropologists prefer talking about “human and plants co-evolution” rather than domestication of plants (Schaal, 2019). The settling of human groups involved major socioeconomic changes, the development of complex social structures (such as cities) or new skills (such as writing), which have dramatic influence on the way we live nowadays (Harari, 2014; Diamond, 2002).

The areas of the world in which agriculture appeared first are called “centers of origin” (Vavilov and Dorofeev, 1992). Different plants (and animals) were domesticated in each “center of origin”. In the Fertile Crescent, which is the cradle of Mediterranean and European societies, these included progenitors of today’s most important crop species, with wheat arguably holding the utmost role. Wheat cultivation and processing have been integral components of European, African and Asian food for several millennia, becoming in the last centuries a central agriculture endeavor worldwide due to European colonization (Palmer, 1996). Nowadays wheat fields cover  $2.15 \times 10^8$  hectares, representing the most abundant agricultural type of land cover (FAO, 2021a), mainly located in Asia, America, and Europe.

This success of wheat, which has resulted in large and continuous areas devoted to its cultivation, has not gone unscathed. Natural ecosystems favour generalist parasites due to the greater biodiversity and environmental heterogeneity. Contrariwise, the increase of homogeneous host density in cultivated areas in new environments, far from the centers of origin, facilitate the transmission of diseases and the emergence of virulent and specialized plant parasites (Stukenbrock and McDonald, 2008). This homogenisation has progressively become a more alarming threat in the last 200 years (Santini et al., 2013) in mono-culture-based agriculture, with the selection of few genetic varieties (Corredor-Moreno and Saunders, 2020) and globalization of trade (Santini et al., 2018). However, although at a slower rate, the roots of this phenomenon can be traced back to the Agriculture Revolution. As evidence of this, it is interesting to note that ancient civilizations clashed with certain plant parasites that continue to challenge us even today. The fungal species *Puccinia graminis*,

causal agent of stem rust of wheat, is an extremely virulent pathogen, capable of provoking great production losses (Olivera et al., 2015), threatening food production worldwide (Ristaino et al., 2021). Plant paleoepidemiology is rarely assisted by direct evidence, but in this very exceptional case *P. graminis* spores have been found in an Israeli granary active in 1,300 B.C. (Kislev, 1982).

With occasional exceptions, in ancient societies, explanations for complex natural phenomena like famines and epidemics often had foundations in myth and religion rather than rigorous scientific structuring (Russo, 2001). One illustrative case is given again by stem rust of wheat. As many other natural phenomena, the emergence of diseases was associated with divine will, which would punish or award humans for their behaviour (van der Eijk et al., 1995; Lucarelli, 2017). We know, for instance, that the Greek Pantheon included Erythibius, also known as the “Apollo of the Rust” (Peterson et al., 2001). More information is available about Romans’ habits: they used to celebrate spring parties, called Robigalia, in which they sacrificed innocent red animals to mitigate the anger of the god Robigus, in the hope this will prevent reddish lesions - revealing stem rust infections - to appear in wheat fields (Zadoks et al., 1985).

Such attempts to explain the presence of stem rust and the consequent preventive actions could trigger the reader’s amusement. If it is the case, note that the first efforts of the scientific community to reconstruct the ethiology of this disease were far from being more convincing. During the XVII century, in England, the most accepted explanations for stem rust of wheat included *i*) high hedges around fields or hollows in the ground, which resulted in stagnant air; *ii*) an atmospheric deposition; or *iii*) the straw of poorly rotted manure (Hartlib, 1655). These theories were elaborated in opposition to a widespread belief - a “foolish superstition” - among farmers, who believed that the lesions were “attracted” by barberry hedges, which people used to plant to separate fields (Ellis, 1741).

Ancient and early modern societies were separated by the so called Copernican Revolution, in the XIV-XV century, a paradigm shift (Kuhn, 1962) in astronomy first, and later in the whole science, which lead to the affirmation of the scientific method as the essential tool to investigate reality. However, it was not before the beginning of the XIX century, thanks also to development of microbiology (Pasteur, 1861), that agricultural scientists started accepting the causal relation between the presence of barberry, around which stem rust lesions were occurring, and the rust disease (Barnes et al., 2020), exactly as farmers argued.

Retracing the whole etiology of stem rust was impossible given the poor conceptual tools of the science of the time. *P. graminis* is an obligate two-host pathogen whose life cycle includes five different stages. During winter, barberry leaves (*Berberis vulgaris*) can be infected by basidiospores, produced by overwintering teliospores. Here, genetic reshuffling can occur within infected barberry by means of pycniospores. When the temperatures start to rise, infection can be transmitted from barberry to nearby susceptible wheat by means of aeciospores. Later in the season, asexual urediniospores are produced from infected wheat.

When temperatures start to decrease, urediniospores are replaced by teliospores for overwintering, and the life cycle starts again (Roelfs, 1992).

Once the fundamental role of barberry in determining the inter-annual persistence of the disease has been widely accepted, removal campaigns have taken place all over Europe and North America along the XIX and XX century (Zadoks, 1967). Barberry is known to persist in remote locations, such as in the Alps mountains (Och and Nowak, 2021) or in the Caucasus (Olivera et al., 2019), or where the warm winter temperatures would allow the survival of urediniospores anyway, such as in Central America (Aylor, 2003) and central Africa (Olivera et al., 2015).

## A noisy spring

This approximately 3,000-year leap in the history of wheat protection against stem rust represents a specific story that tells us how societies have dealt with the complexity of some pathosystems throughout the centuries and it will be taken up in the next section. However, it is not necessarily representative of the story of parasites invasions in general.

Copernican Revolution is far from being the only historical fact that affected the discipline of plant pathology between the XV and XVI century. In the same period, the European exploration and colonization of America lead to the Colombian Exchange, i.e. the relative diffusion of crops between the “Old” and the “New world” (Diamond and Ordunio, 1999), which had been isolated in the precedent 10.000 years. Potatoes, corn and tomatoes were imported to Eurasia-Africa, while wheat, soybean and rice were introduced into America, just to give some examples; and each one with its parasites (Santini et al., 2018).

All these crops have found great diffusion in the new contexts where they were introduced. Potato is arguably one of the crops which met the steepest diffusion in Europe, because of its endurance to cold climatic conditions - which were even enforced by the Little Ice Age (Abel, 2013). The innovations introduced during the Second Agricultural Revolution accelerated its diffusion, which in some contexts abruptly replaced precedent cultures, with alternate outcomes. It is the case, for example, of Ireland, where, around the half of the XIX century, a large outbreak caused by *Phytophthora infestans* devastated the only variety of potato that was cultivated (“lumper”) resulting in a devastating famine and in a dramatic demographic decline, from which the island population has not yet recovered (Carefoot et al., 1969).

Beside sporadic contractions in productivity, an improved agricultural yield and a more reliable food supply accompanied the worldwide demographic boom in the XIX and XX centuries. Another significant transformation in agriculture was represented by the Green Revolution, a set of pioneering farming changes which commenced in the 1940s in Mexico and rapidly spread all over the globe (Pingali, 2012). The cardinal elements of this transformation included the intense input of chemical fertilisers and phytosanitary products, the

introductions of new technologies, irrigation facilities as well as the selection of a narrow set of high-yield varieties. This led to an explosion in the yield and an increase of its reliability, reducing those fluctuations in productivity that were a real constraint on food security (Webb and Eiselen, 2009). For instance, it has been estimated that rice production in South-East Asia doubled within a couple of decades (Zeigler and Mohanty, 2010), helping feeding its unprecedented growing population.

The Green Revolution represented a further step toward the industrialization of agriculture and completely reshaped rural landscapes. With the increase of productivity, fewer farmers were needed to produce the same amount of food, bringing to a progressive decline of the rural population which migrated to cities (Johnson and Lichter, 2019). Increased cultivation surface with less manpower resulted in an expansion of mono-cultured surfaces (Pingali, 2017). In the decades around the half of the last century it was believed that plant parasites could be completely controlled by fungicides and insecticides. This approach started soon to show some limitations (Pingali, 2012).

In the view of the promoters, higher yields deriving from technological innovations would have helped the poorest fraction of population access to food. At the same time, these would have saved natural habitats from agricultural conversion because of a more intense use of existing agricultural surfaces (Pingali, 2012). The first objective was partially achieved, while the second was missed, mainly because of the simultaneous population increase and of the per-capita consumption (Pingali, 2012). Moreover, unexpected difficulties emerged, such as emergence of pesticide resistance of pathogens and pests, favoured by the above-mentioned characteristics of modern agriculture (Stukenbrock and McDonald, 2008), such as the loss of genetic crop diversity within larger host surfaces. Furthermore, increased commodity shipping increased the risk of parasite invasions in new territories. This meant that, in spite of an increased input of pesticides, the disease-related production losses did not decrease (Carson, 1962). This is the case of *Monilinia fructicola*, a fungal pathogen affecting mainly stone fruits, and in particular causing brown rot of peaches (Bryde and Willets, 1977). This species, absent in Europe up to the year 2000, was inadvertently introduced into France, supposedly from the United States (Lichou et al., 2002). *M. laxa* and *M. fructigena* already caused similar symptoms in stone fruits, provoking high production losses (Bryde and Willets, 1977). The spread of this further pathogen throughout the continent has not been controlled, despite the application of high levels of fungicides, which in recent years reached rates of 15 to 25 doses per parcel per year (Cretin et al., 2018). Today *M. fructicola* is endemic in Hungary, Switzerland, Germany, Czech Republic, Slovenia, Italy, Austria, Poland, Slovakia, Serbia and Spain (Oliveira Lino et al., 2016).

Certain unexpected harmful consequences of the Green Revolution arose because there was little awareness about them at the time. Contrariwise to a common belief, the rise of the environmentalist movement was not due to the acknowledgement of climate change or over-exploitation of non-renewable resources, but to the deleterious effects of bio-accumulating



phytosanitary pollutants, witnessed by the American biologist Rachel Carson in her book “Silent spring” (Carson, 1962). Biodiversity loss, soil erosion, eutrophication of water bodies, greenhouse gases emissions and, in general, loss of ecosystem functions are other negative impacts of farming in the second half of the XX century (Power, 2010).

The Green Revolution has shaped agricultural practices in such a way that still nowadays we refer to them as “conventional” farming (Sumberg and Giller, 2022). However, in recent decades, major shifts are taking places, addressing sometimes complementary aspects of farming practices and underlying almost contrasted visions of the role of agriculture. These include, but are not limited to:

1. Engineering techniques to produce more resistant and productive genetically modified organisms (GMOs; Zilberman et al., 2018);
2. Organic farming, a set of practices with the direct aim of reducing chemical inputs in food and in the environment (Seufert et al., 2017);
3. Precision farming, which includes the use of digital sensors and machines to collect data on the state for the crop and support precise and automated management (Finger et al., 2019);

Some of these innovations, such as GMOs, have established in some parts of the world, such as USA, Brazil, Argentina and India (Clive, 2015) while others, such as precision farming, have found limited application and are part of a political road-map to address tomorrow’s agriculture, such as the “Farm to Fork” initiative of the European Commission (Schebesta and Candel, 2020).

Considered in their most commonly held conception, these practices retain some elements of continuity with the precedent paradigms with respect to the issue of crop protection. For example, these innovations seek to transversely address crop protection at the local scale, under the implicit assumption that farming practices have consequences only in the managed plot and in the immediate surroundings.

Does it really work this way? The answer brings us back to *P. graminis*. For the first time after decades, stem rust of wheat was observed in 2013 in Germany (Olivera Firpo et al., 2017), in 2016 in Italy (Bhattacharya, 2017) and in 2020 in Ireland (Tsushima et al., 2022). This happened several decades after the emergence of a virulent and resistant strain, Ug99, in central Africa (Pretorius et al., 2000). Despite a gradual relaxation of its eradication, in Europe barberry is supposed to be conserved only in remote locations; on the other hand, large cultivated surfaces in Europe display poor genetic diversity (Corredor-Moreno and Saunders, 2020) and resistance (Saunders et al., 2019).

Scientists suggest that aerial transport of *Puccinia* spores from endemic zones may explain re-emergence of stem rust in Europe after several decades (Corredor-Moreno and Saunders, 2020) - remarkably, exactly as some English scientists speculated in the XVII century (Hartlib, 1655). Airborne transport of spores over continental scales among a widely

distributed host is known to produce seasonal routes (Brown and Hovmøller, 2002), such as the North-American “*Puccinia* pathway” (Aylor, 2003), a series of long-distance seasonal leaps moving from Mexico to Canada following a thermal gradient. These transport events between fields transform an heterogeneous agricultural landscape into a connected network of cultivated sites where a given action in one site may have consequences in the very distant sites depending on its topology (Strona et al., 2020). It becomes evident that, in case of long-distance dispersed parasites, management practices should be planned considering epidemic interactions occurring at a larger scale than one’s farm.

Regardless of the way crop protection is put in practice (via pesticide application or use of resistant varieties), it becomes evident that long-distance transports - airborne, or human-mediated - over an interactive landscape should be embedded in current crop protection strategies. This shift in the scale of agricultural management asks for a new framework to develop innovative conceptual tools (Morris et al., 2022).

## **Crop protection within the One Health framework**

Covid-19 pandemic has put two important considerations about human pandemics into the public domain. The first concerns the concept of spillover, i.e. the transmission of a pathogen between an infected reservoir population and a novel pristine one. Zoonotic spillover are those passages from an animal population to humans, such as Covid-19, but also Ebola or AIDS; most of human infectious diseases are zoonotic spillovers (Ellwanger and Chies, 2021). The second concerns the spread pattern of emerging pandemics. Epidemics are spread by human contacts. Before globalization, in a less densely inhabited world of limited mobility, plagues spread according to reaction-diffusion patterns, like wild-fire (Noble, 1974); geographic distance ultimately determined the timing of arrival of the disease. Contrariwise, it has been shown that the spread of contemporary pandemics, such as Covid-19 or the 2009 swine flu, can be fully modelled at the large scale by considering the frequency of air-traffic connection among countries (Brockmann and Helbing, 2013). Fuelled by globalized shipping of people and goods, parasites no longer spread solely based on geographic proximity, but are capable of long-distance transports (Banks et al., 2015).

The debate around more frequent zoonotic spillovers lead scientists to elaborate the concept of One Health. This framework acknowledged the interdependence between human and animal health as part of the same system (Hinchliffe, 2017). The One Health framework was first conceived as a bridge between human and veterinary medicine with the aim of anticipating the emergence of new zoonoses. It promoted the idea of an international partnership for the establishment of global bio-surveillance networks to inform political action (Morris et al., 2022).

Given the story that I have developed in the previous sections, readers will not be surprised that it was just a matter of time before plant scientists advocated plant science to

be included in the One Health framework. For instance, this year, the XII meeting of the International Society of Plant Pathology had “One Health for all plants, crops and trees” as motto.

The dependence of human health in plant health in terms of agriculture production and ecosystem services has been summarized in the previous sections. Nevertheless, additional reasons lead scientists to include plant health in the One Health framework. One, for example, is the recent formalization of the concept of “phytonose”, i.e. the passage of pathogens from plants to animals and humans (van der Riet, 1997; van Overbeek et al., 2014), which provides a parallelism with the concept of zoonoses which was at the base of the concept of One Health.

The second reason lies in the observation that increasing ecological imbalances, which ultimately impact human well-being, are represented by biological invasions of alien species and/or of plant parasites, whose rate and related costs are increasing worldwide (Diagne et al., 2021). Current introductions of these species into new territories are mostly mediated by humans (Banks et al., 2015), so more and more research is done under the umbrella of “invasive” species, regardless of whether these are affecting plants or animals (Diagne et al., 2021).

The focus of One Health on surveillance and the inclusion of both animals and plants has induced scientists to look for innovative and versatile tools and methods for preventing the emergence of diseases (Morris et al., 2022). An example of innovative tool for early detection of infectious plant parasites is the use of canine olfactory detection to rapidly survey large plantations without sample collection or laboratory processing (Gottwald et al., 2020), or the use of Lagrangian trajectory simulations of air masses to predict where airborne spores will be deposited (Meyer et al., 2017b) and inform stakeholders and policymakers to design control actions (Allen-Sader et al., 2019). Another example from outside plant pathology is the idea that the emergence of Covid-19 could have been anticipated via the analysis of the occupation of parking lots in Wuhan from satellite images (Nsoesie et al., 2020).

Other innovative tools include text mining and network science. Their versatility consists in the possibility of being used in very different contexts with a reduced cost of adaptation. Text mining techniques, for example, are used to extract information from heterogeneous sources and mapping them into semantic structures. Their application to crop protection is devoted, for example, in building knowledge about pathogens reservoirs (Morris et al., 2022). In France, the animal, food processing and plant epidemic surveillance platforms (Amar and Dupuy, 2020) integrate text mining into the analysis of the impressive amount of documents produced in different languages around the world to anticipate the spread of diseases (the so called “*Veille sanitaire internationale*”). Networks are mathematical objects made of fixed elements, called nodes, and relations among nodes, called links or edges (Newman, 2003). These structures have been used to represent disease spread in humans (Brockmann and Helbing, 2013), animals (Dubé et al., 2011) and plants (Moslonka-Lefebvre

et al., 2011). In such a representation, nodes play as hosts, and epidemic contacts play as edges. Network scientists have explored topological properties that have proved to identify most vulnerable nodes or those nodes which play as disease spreaders (Jeger et al., 2007). Network-based disease surveillance and control strategies have been applied in the case of humans and animals, and a growing interest has risen around application to plants (Garrett et al., 2018).

These methods are all objects of investigation in the ANR BEYOND project, which co-finance my thesis. We are sure, they would have triggered the amusement scientists in the XVII century; how amazed they would have been upon realizing that dogs truly detected infections weeks to years prior to visual and molecular surveys (Gottwald et al., 2020)!

## Objectives of the thesis and outline of the manuscript

In this thesis I aim to explore innovative crop protection strategies against long-distance dispersed parasites. I focus on aerial dispersal as a main transport mechanisms, and I test complex network capabilities to provide management indications in two case studies: stem rust of wheat and brown rot of peaches.

Concerning stem rust of wheat, I am far from being the first to model quantitatively its aerial transport. Recent works have used Lagrangian trajectory simulations of air masses, provided by the NAME (Jones et al., 2007) or HYSPLIT (Draxler and Hess, 1998) models, to reconstruct the aerial introductions of new rust strains in Ethiopia (Meyer et al., 2017a) and Australia (Visser et al., 2019), upgrading past pioneering work about the introduction of soybean rust in the USA (Isard et al., 2005). With respect to these studies, I concentrate my work on reconstructing the worldwide “*Puccinia* pathways” and I use recurrent transport events of spores to identify a set of sentinels, locations to be monitored systematically detect the emergence of an outbreak, to surveil worldwide wheat producing areas, similarly to what Suttrave et al. did at the country scale.

As already mentioned, brown rot of peaches is a fungal disease caused by a set of pathogens of the *Monilinia* genus. These can provoke damages on fruits, blossoms, twigs stone fruit plants in general; however, they have the potential to provoke almost complete losses on peaches and nectarines (Oliveira Lino et al., 2016). Among the main pathogens causing this disease, *M. fructicola* is particularly interesting since, in just one decade, it colonised most of the European continent starting from an outbreak in France (where it was first detected in 2001).

Modelling the spread of brown rot of peaches presents more challenges, although I decided to restrict the geographic domain of the investigation to France only. A first challenge concern its aerobiology; despite airborne transport of *Monilinia* spores has been suggested by literature, there are no specific studies on this topic. Moreover, in order to compare the model with field observation of severity in the last twenty years, I have to build a

climate-dependant dynamical framework reflecting coupled cycles of dispersal, infection, and onward spread of the disease. For this reason I integrate wind-driven parasite transport, represented by connectivity networks, with climate-driven local epidemiology of host, in a unified modelling framework. I then use this framework to identify the most dangerous locations (the ones where a parasites, if introduced, would provoke the highest damage at the landscape level) and the most vulnerable (the ones which would be more often damaged by an outbreak taking place anywhere else) with respect to brown rot spread.

Both stem rust of wheat and brown rot of peaches, even if at different scales, are diseases whose rapid spatial spread concerns more countries. Uncoordinated protection measures at the international scale may lead to an increase of the risk of invasions and its control-related costs. Therefore, the last part of this manuscript is devoted into assessing the benefits of a cooperative international surveillance strategy over a non-cooperative one, where each country optimise its own surveillance. I use the worldwide epidemic network for stem rust to perform this assessment. Moreover, I quantify the degree of heterogeneity of the geographical distribution of benefits among countries due to cooperation.

This manuscript is then structured in four chapters:

1. I define the concepts of crop surveillance and control; I present how network science has been used to represent human, animal and plant epidemiological dynamics; eventually, I investigate how researchers have used network properties to design disease protection strategies.
2. I reconstruct the worldwide epidemic network of stem rust of wheat, based on a framework integrating Lagrangian trajectory simulations of air masses with suitability filters, and I propose a network-based algorithm to identify the best sentinels, i.e. those locations to be monitored systematically to early detect the emergence of an outbreak.
3. I integrate the theoretical work on Lagrangian simulations to inform and calibrate a climate-driven metapopulation model to describe the epidemiological dynamics of brown rot of peaches in France. I test weather directional long-distance dispersal plays a significant role in describing the incidence of brown rot in France in the last twenty years; I produce map of epidemic risk to design crop protection strategies.
4. I use the worldwide epidemic network of stem rust of wheat to test what the overall benefits of a cooperative surveillance strategy over a non-cooperative, country-based one, would be. I demonstrate that a global benefit of a cooperative strategy is accompanied by an heterogeneous distribution of costs.

The second and the fourth chapters correspond to research articles already accepted and published (Radici et al., 2022, 2023b), while the first and the third ones have been submitted to peer-reviewed journals. Each chapter is accompanied by a section that will help the reader contextualise it with respect to the manuscript road-map.

# Chapter 1

## Network-thinking to optimize surveillance and control of crop parasites. A review

The Guide is accurate.

Reality, instead, is often inaccurate.

– Me misquoting *The Hitchhiker's Guide to Galaxy*

### Résumé

L'augmentation des surfaces cultivées, l'homogénéisation des cultures et le commerce mondial des denrées alimentaires ont favorisé la propagation des ravageurs et des maladies des plantes. Il est urgent d'optimiser la protection des cultures pour garantir la sécurité alimentaire. Deux des aspects de la protection des cultures sont la surveillance, qui se concentre sur la détection précoce d'un parasite, et le contrôle, qui vise à le combattre et éventuellement à l'éradiquer. La théorie des réseaux complexes a été largement utilisée pour modéliser la propagation des maladies infectieuses humaines et animales dans des systèmes décrits par des nœuds et des liens. Elle a été utilisée avec succès pour optimiser les campagnes de surveillance et de vaccination. Dans le domaine de la protection des plantes, on trouve de plus en plus d'ouvrages utilisant cette théorie pour décrire la propagation de parasites et pour concevoir des stratégies de protection.

Nous passons ici en revue l'utilisation de la théorie des réseaux complexes en épidémiologie végétale, depuis les approches visant à décrire uniquement la propagation des maladies jusqu'à celles appliquées spécifiquement à l'optimisation de la protection des cul-

---

This chapter has been submitted as a review article (pre-print available at: Radici et al.2023).

tures. Nous retraçons le processus logique qui a conduit les modèles épidémiologiques à s'appuyer sur la théorie des réseaux complexes, et nous donnons des exemples de la façon dont la propagation des parasites des cultures a été représentée par une description de réseau. Nous définissons les objectifs de la surveillance et du contrôle, et nous montrons comment ces concepts ont été déclinés dans la sphère épidémiologique des réseaux, puis adaptés au contexte agricole. Enfin, nous discutons de l'écart que nous observons entre l'application de la théorie des réseaux dans la surveillance et le contrôle, afin d'identifier les failles et les solutions.

Nous constatons que : *i*) les scientifiques ont réussi à interpréter des modes de transmission parasitaire très différents sous la perspective de la théorie des réseaux ; *ii*) alors que l'épidémiosurveillance basée sur les réseaux a progressivement clarifié ses objectifs et que des outils solides ont été proposés, le contrôle basé sur les réseaux a été moins étudié et appliqué ; *iii*) la réflexion autour des réseaux doit porter sur la manière de définir correctement les liens et les nœuds à différentes échelles géographiques afin d'élargir leur application à la protection des cultures.

**Mots-clés** : protection des plantes, parasites, réseaux complexes, épidémiosurveillance, contrôle des épidémies.

## Abstract

Increasing cultivated lands, crop homogenization and global food trade have fostered the spread of crop pests and diseases. Optimizing crop protection is urgently needed to ensure food safety. Two aspects of crop protection are surveillance, which focuses on the early detection of a parasite, and control, aiming to fight and possibly eradicate it.

Network theory has been widely used to model the spread of human and animal infectious diseases in systems described through nodes and edges. It has been successfully used to optimize monitoring and immunization campaigns. In crop protection, there is a growing literature using this theory to describe parasites spread and to conceive protection strategies.

Here we review the use of network theory in plant epidemiology, from the more descriptive to the more applied approaches aimed to optimize crop protection. We retrace the logical process that has led epidemiological models to rely on network theory, and we provide examples of how the spread of crop parasites has been represented via a network description. We define the objectives of surveillance and control, and we show how these have been declined in the network-based epidemiological sphere and then adapted in the agricultural context. We eventually discuss the discrepancy between the application of network theory in surveillance and control to identify culprits and solutions.

We find that: *i*) scientists have successfully interpreted very different modes of parasitic transmission under the lens of network theory; *ii*) while network-based surveillance has

progressively clarified its objectives and sound tools have been proposed, network-based control has been less studied and applied; *iii*) network-thinking must address how to properly define edges and nodes at different geographic scale to broad its application in crop protection.

**Keywords:** crop protection, parasites, networks, epidemic surveillance, epidemic control.

## 1.1 Introduction

Pests and pathogens are responsible for a reduction between 17% and 30% of crop production at the global scale (Savary et al., 2019). In the next decades, crop losses are expected to increase due to the narrowing of diversity in crop species and the increase of food demand, pesticide resistance, and of global trade which determine favorable conditions for spread of pests and diseases as they increase the host abundance and contact probability between hosts (Carvajal-Yepes et al., 2019; Ristaino et al., 2021; Khoury et al., 2014). Climate change could also facilitate the spread of certain pathogens and pests (Corredor-Moreno and Saunders, 2020) and most of the losses are expected in countries with expanding populations, where food supply is already an issue (Tilman et al., 2011; Adam et al., 2021). Moreover, agriculture is asked to mitigate its environmental and health impacts related to the use of chemicals which have been the major tool to protect crop from pest and disease in the last decades (Vitousek, 1994; Tudi et al., 2021). The challenge of reconciling sufficient food production with social and environmental viability is acknowledged as part of the Sustainability Development Goals of the UN 2030 Agenda (Foley et al., 2011; Tilman et al., 2002; Lee et al., 2016).

Therefore, surveillance and control, which are complementary measures to protect crops from parasites, hereinafter intended for both pests and pathogens, ought to be optimized (Carvajal-Yepes et al., 2019; Morris et al., 2022; Ristaino et al., 2021). Surveillance, which is mostly based on data collection and analysis, aims to detect the presence of parasites early and to know the phytosanitary status of a crop. Control aims to reduce parasitic dispersal, either by immunizing hosts, *e.g.* by cultivating resistant varieties, or by limiting parasite dispersal by reducing contacts between individuals. An effective crop protection strategy requires both surveillance and control to work together: in fact, the efficacy of control measures critically depends on the timing of detection of the parasite, which is likely to spread if favorable conditions are present.

Network theory, originally intended to study properties of systems described by a set of nodes connected via edges (Newman, 2003), has been successfully applied to epidemiology. In this context, the nodes of the network represent host individuals or distinct host populations and the edges represent connections between nodes. In recent decades, network theory has been extensively used to model the spread of infectious diseases in humans



and animals and to inform protection policies (Keeling and Rohani, 2011; Brockmann and Helbing, 2013; Kao et al., 2007; Dubé et al., 2011), the recent Covid-19 pandemic being a paradigmatic case (Block et al., 2020; Saunders and Schwartz, 2021; Della Rossa et al., 2020).

Network theory has proven to be a valid tool in crop protection to model disease spread and to design surveillance and control strategies (Moslonka-Lefebvre et al., 2011; Cunniffe et al., 2015; Parnell et al., 2017; Garrett et al., 2018). Unlike animals and humans, plants do not move, so that the transmission of a parasite depends on its displacement. This can be active, such as flying from an orchard to another, or passive, driven by living agents (called vectors) or abiotic elements such as air masses, water currents, or even via transport of goods. In the case of crop parasites, nodes of the network usually identify locations where the host is present, which generally correspond to cultivated areas (Fig. 1.1).

An example of network-thinking in crop protection is the identification of those locations that contribute the most to spreading parasites (Zhang et al., 2016). This identification, hereinafter called prioritization, is needed because crop protection is expensive, so it is important to rank and select a reduced set of locations (nodes) to be inspected for surveillance and/or to be treated for control (as in Fig. 1.1d). Recently, more advanced research has been conducted to investigate useful node properties which allow this prioritization (Martinetti and Soubeyrand, 2019; Sutrave et al., 2012; Hernandez Nopsa et al., 2015) in the case of re-emerging parasites, such as *Puccinia graminis* (Meyer et al., 2017) or *Xylella fastidiosa* (Strona et al., 2017) in Europe. The implications of networks for crop protection are rising (Jeger et al., 2007; Moslonka-Lefebvre et al., 2011; Shaw and Pautasso, 2014; Parnell et al., 2017; Garrett et al., 2018), but little research has focused on rigorously clarifying their properties for surveillance and control (Holme, 2017). Also, the level of application of network-thinking is not the same in surveillance and control: this review aims to present the state of the art in terms of network-thinking in crop protection, investigate the main obstacles to applications and delineate research perspectives.

Cultivated landscapes, crop protection and their representation on networks

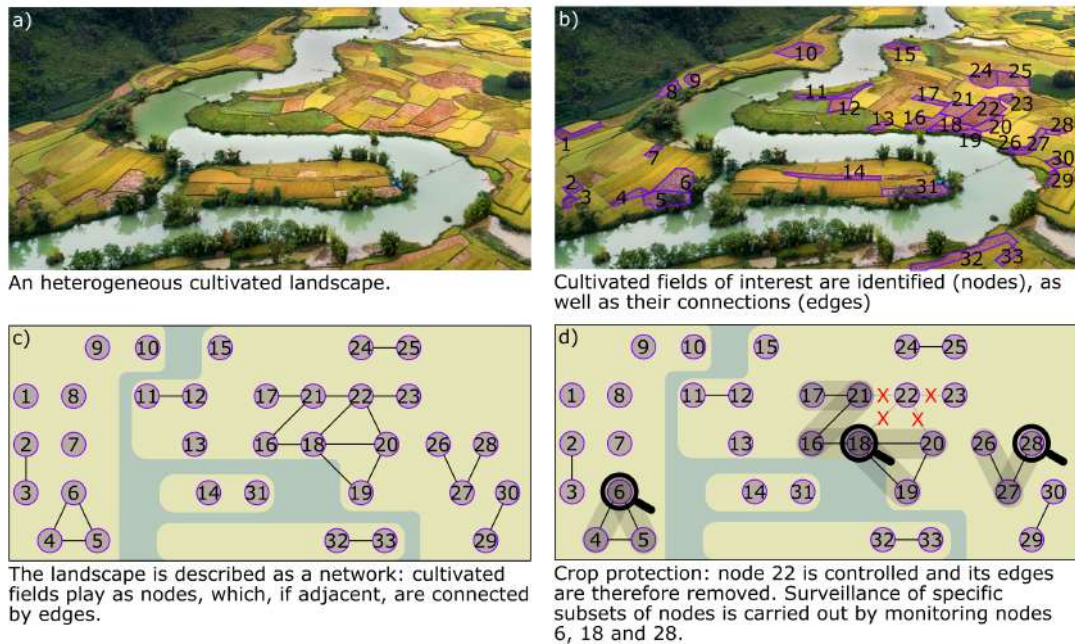


Figure 1.1 – A network of scattered fields over which crop protection is conducted. a) Fields of interest are part of a complex landscape; b) in this example, we imagine that the reddish fields play as nodes, and that parasites can move from a field to an adjacent one (via an edge); c) the network is formalized into a graph made of a set of nodes connected via edges, discarding all the trivial elements (only the river is represented in the background to help readability); d) control of node 22 (via cultivation of resistant cultivars and/or phytosanitary treatment) prevents parasitic flow in and throw this node: its edges are therefore removed. The presence of parasites in the network is acknowledged via surveillance of nodes 6, 18 and 28 (this latter, for example, allows to indirectly monitor connected nodes 26 and 27).

## 1.2 Networks: definitions and properties

Networks are sets of nodes connected via edges (Börner et al., 2007). In epidemiology, nodes represent hosts (individuals or populations), while edges represent infectious contacts. A classic way to represent a network is a square matrix  $\mathbf{W}$ , called adjacency matrix, whose generic element  $w_{i,j}$  represents the edge connecting node  $i$  to node  $j$ . Edges can be either binary ( $w_{i,j} = 1 \vee 0$ ) or associated to weights, generally positive ( $w_{i,j} \in \mathbb{R}^+$ ); either undirected, if a relation from a node  $i$  to a node  $j$  implies the same from  $j$  to  $i$  (it follows that the adjacency matrix is symmetric:  $w_{i,j} = w_{j,i}$ ) or directed; lastly, they can be dynamic (edges arrangement and weight vary with time) or static. Nodes can be characterized in terms of their features, which are quantified via metrics (often called centralities). The simplest metric that one can build is the degree, usually indicated by  $k$ , *i.e.* the number of edges connected to a node. In the case of weighted networks, the notion of degree is usually complemented by that of strength, *i.e.* the sum of the weights of the edges connected to a node. In the case of directed networks, it can be further distinguished between in- and out-degree/strength, considering only in- or out-going edges. The probability distribution of  $k$  in a given network is called degree distribution, and its average is indicated as  $\bar{k}$ . The degree distribution can be used to classify network properties as a whole. For instance, homogeneous networks are built in such a way that all nodes have the same probability of having an edge with another node. On the other hand, the Barabási-Albert algorithm allows to build scale-free networks (Barabási and Albert, 1999), in which the degree distribution follows a power law.

Several node metrics can be defined (Mastin et al., 2020), such as the betweenness of a node  $w$ , *i.e.*, the number of shortest path connecting whatever couple of nodes of the networks  $i$  and  $j$  and passing through  $w$ , where the shortest path is the shortest sequence of consecutive links connecting two nodes, or, in case of weighted networks, the sequence of consecutive links associated with the lowest strength sum (Barabási, 2016). Interested readers can find an exhaustive summary of node metrics in Lü et al. (2016).

## 1.3 Modelling epidemics

Mathematical epidemiological modelling has its origin in the S-I compartmental model (Kermack and McKendrick, 1927). In one of its simplest formulation, it assumes that a population is divided into susceptible (S) and infected (I) individuals and describes its temporal variation via a system of differential equations:

$$\begin{cases} \dot{S} = -\beta SI + \mu I \\ \dot{I} = \beta SI - \mu I \end{cases} \quad (1.1)$$

where  $\beta$  represents the disease transmission rate, that drives S to become I, and  $\mu$  represents the recovery rate, that drives I to return to S. This model can be extended to include other processes (*e.g.* immunity, natural mortality and fertility) and it allows to define the basic reproductive number  $R_0$ , which indicates "the expected number of secondary cases which one case would produce in a completely susceptible population" (Dietz, 1993). In the model described by Eq. 1.1,  $R_0$  is:

$$R_0 = \frac{\beta}{\mu} \quad (1.2)$$

If  $R_0$  of a disease is greater than 1, the introduction of an infected individual into a fully susceptible population will cause the spread of the disease, otherwise it will progressively die out. This model assumes homogeneous mixing, which means that individuals have equal probabilities of transmitting the disease to everybody else. This is not the case, for example, when the spatial location of individuals defines their contact structure, which can be formally represented by a network (Pastor-Satorras et al., 2015). In that case, classic results of epidemic theory may not apply (Pellis et al., 2015).

Moreover, networks can embed SI dynamics. In a time-discrete framework, a node can be either susceptible or infected; at each time step  $\delta$ , each infected node  $i$  can infect its  $k$  neighbors at a probability proportional to  $\beta\delta$ , and recover with a probability  $\mu\delta$ . For instance, in SI dynamics on homogeneous networks it can be shown (Boccaletti et al., 2006) that Eq. 1.2 is modified into:

$$R_0 = \frac{\bar{k}\beta}{\mu} \quad (1.3)$$

So  $R_0$  depends explicitly on the network via the average degree  $\bar{k}$ , which now is spelled out and separated from  $\beta$ . For non-homogeneous networks,  $R_0$  can be recomputed (Boccaletti et al., 2006) instead as:

$$R_0 = \frac{\overline{k^2}\beta}{\bar{k}\mu} \quad (1.4)$$

Where  $\overline{k^2}$  is the 2<sup>nd</sup> moment of the degree distribution of the network. It has been shown that, for epidemic processes in scale-free networks,  $\overline{k^2}$  tends to infinite (Pastor-Satorras and Vespignani, 2001): the disease will spread due to the network topology provided a positive transmission rate (Shaw and Pautasso, 2014; Jeger et al., 2007). Eq.s 1.3 and 1.4 suggest that network-related control strategies can help decreasing  $R_0$  below 1 by acting on the network structure.

## 1.4 Crop-parasite interactions in a network framework

In crop epidemiology, plants are connected via parasitic dispersal (Shaw and Pautasso, 2014). Parasites can move over extremely long distances through different means (air, water, soil, insects; Aylor, 2017, Jordano, 2017) or can be assisted by human activities (Santini et al., 2018; Hulme, 2009; Harwood et al., 2009; Garrett et al., 2018). The definition of the network depends on the scale and mechanism of the contacts (Gilligan, 2008; Shaw and Pautasso, 2014; Garrett et al., 2018). Below is a list of examples on how networks have been used to represent and study crop-parasite interactions.

### 1.4.1 Vector-transmitted parasites

Strona et al. (2017) studied the spread of the pathogenic bacteria *X. fastidiosa* in an olive-orchard-dominated landscape, via its main european vector, the meadow spittlebug *Philaenus spumarius* (Martelli et al., 2016). They considered olive orchards as nodes, which are connected by an edge if they are within 1 km "between the two closest sides", since observations suggested that this is the order of magnitude of adult spittlebug daily flights. They obtained a binary network, since edges are not associated to a weight, and undirected, since connections among two orchards are always reciprocal. Instead, de la Fuente et al. (2018) built a more complex network by modelling the spread of the alien nematode *Bursaphelenchus xylophilus*, causal agent of pine wilt disease, via the longhorn beetle *Monochamus galloprovincialis*. They considered coniferous forests as nodes connected via weighted, directed and time-varying edges. Each edge represents the yearly probability of transmission of the disease from node  $i$  to node  $j$ , whose weight decreases with the distance between  $i$  and  $j$  and increases with the assumed number of nematodes in  $i$  in the previous year.

### 1.4.2 Wind-dispersed

Sutrave et al. (2012) studied a network of soybean fields in the US, susceptible to the airborne pathogen *Phakopsora pachyrhizi*, obtaining a similar network to that of de la Fuente et al. (2018). In this model, each node represents an administrative county, while each edge's intensity is determined by the daily average wind intensity and direction, increases with the density of soybean in the connected nodes, and decreases with the distance. In this case, the network is weighted, directed (because of the prevailing wind direction) and dynamic. Enlarging the scale to the continental one, Meyer et al. (2017) used Lagrangian simulation of air masses to model the airborne dispersal of the fungal pathogen *P. graminis*, causal agent of stem rust of wheat, among African and Asian countries. Countries play as nodes, while the daily mean proportions of airborne spores successfully transmitted from a country to another define the weighted and directed edges.

### **1.4.3 Human-mediated**

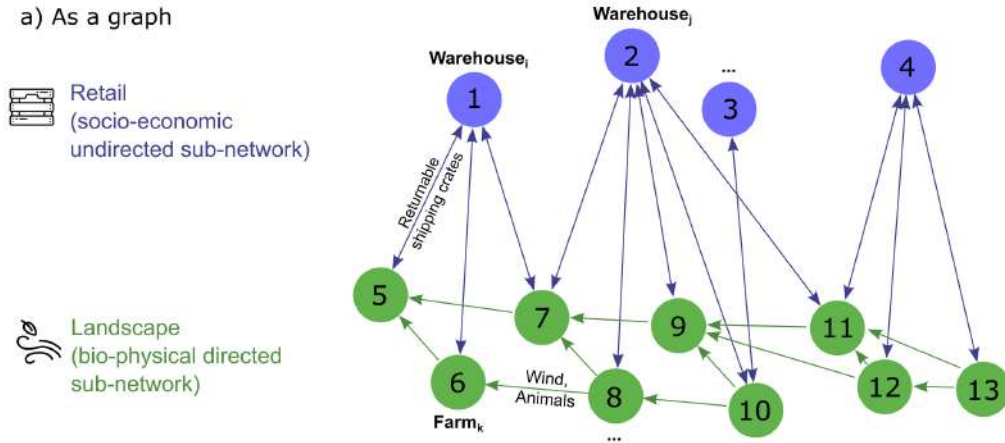
A recent study by Andersen et al. (2019) models the spread dynamics of a hypothetical pathogen affecting sweet potato on a network accounting for local seed trade in Uganda. In such networks, nodes are provided by economic actors or their grouping (sellers, villages), whose interaction, in this case "at least one transaction during the growing season", represent directed edges. Hernandez Nopsa et al. (2015) used the railway networks in the US (and Australia) to model the potential spread of arthropods between grain elevators used to store wheat. In their framework, sets of elevators located in the same state play as nodes, while directed and weighted edges are proportional to the volume of wheat moved from a state to another via railway, which traditionally account for more than 70% of the shipped wheat in the US.

### **1.4.4 Multi-mechanism**

Eventually, networks may also summarize multiple mechanisms involved in parasitic spread (Garrett et al., 2018). For instance, nodes may either be farms, where crops are produced, or warehouses, where crops are stored after harvesting (Fig. 1.2). Edges, representing possible pathways of parasitic spread, can be subdivided into retail sub-network, representing purchase, and landscape sub-network, representing physical movement by wind (or insects or animals). Retail sub-network may be undirected, working both ways between producers and retailers via returnable shipping crates, while landscape sub-network may have a fixed direction, because of prevalent wind.

## A multi-mechanism epidemic network

a) As a graph



b) As an adjacency matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13
Warehouses	1	0	0	0	0	1	1	1	0	0	0	0	0
2	0	0	0	0	0	0	1	1	1	1	0	0	0
3	0	0	0	0	0	0	0	0	0	1	0	0	0
4	0	0	0	0	0	0	0	0	0	0	1	1	1
Farms	5	1	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	1	0	0	0	0	0	0	0	0
7	1	1	0	0	1	0	0	0	0	0	0	0	0
8	0	1	0	0	0	1	1	0	0	0	0	0	0
9	0	1	0	0	0	0	1	0	0	0	0	0	0
10	0	0	1	0	0	0	0	1	1	0	0	0	0
11	0	0	0	1	0	0	0	0	1	0	0	0	0
12	0	0	0	1	0	0	0	0	1	0	1	0	0
13	0	0	0	1	0	0	0	0	0	1	1	0	0

Figure 1.2 – A multi-mechanism epidemic network. Nodes are either farms, where crop is produced, or warehouses, where harvested crop is stoked. Edges, representing possible parasite spread pathways, can be subdivided into two sub-networks: retail, representing purchase, and landscape, representing physical movement by wind insects, or animals. Retail sub-network is undirected, since it may work both way between producers and retailers (for example, via returnable shipping crates), while landscape sub-network has a fixed direction, (for example, because of prevalent wind). Network is represented a) as a graph, b) as an adjacency matrix, in which the retail sub-network is symmetric (in blue) while the landscape one is asymmetric (in green).

## 1.5 Surveillance

Surveillance can be defined as the "collection and analysis of information for detection and successful control of emerging pathogens" (Parnell et al., 2017). Since this collection is expensive, one needs to prioritize a reduced set of locations for monitoring to find clues of crop disease emergence.

### 1.5.1 Prioritize nodes to be monitored

In a network-based surveillance, prioritization of nodes to be monitored (also referred as sentinels) depends on the topology of the network structure. Surveillance may be divided into different objectives (Herrera et al., 2016; Holme, 2018). For surveillance of emerging diseases, Parnell et al. (2017) distinguished between *i) detection*: understanding whether a pathogen has arrived in an area, with the minimum possible delay; *ii) estimation*: estimating the proportion of diseased individuals (incidence) at a given time; *iii) targeting*: maximizing the detection of new cases. These objectives are not pursued without trade-offs. For example, in the case of an already detected parasite, increased "targeting" would imply monitoring those nodes which are connected to the infected one, to the detriment of "estimation". In contrast, good "estimation" can be obtained via homogeneous monitoring, *e.g.* via random sampling of the entire network, but would result in a worse "targeting".

Different prioritization algorithms allow for the pursuit of different objectives. Herrera et al. (2016) compared performances of different algorithms to determine a sentinel set in epidemic dynamics on social networks. The criteria for prioritization are: "most connected" (nodes characterized by the highest degree); "random"; "random acquaintance" of random individuals, *i.e.* a random node connected to a randomly extracted node. As expected, sentinels identified via the "random" strategy yielded indications representative of the populations as a whole, with early warning close to zero (there was no time lag between the surveillance subset and the entire population to reach 1% prevalence) and peak ratios close to one (*i.e.*, peak in incidence was almost identical in the surveillance subset and the entire population). Degree-based strategies, such as "most connected", consistently provided the earliest warnings. Herrera et al. explored different network topologies and observed that surveillance performances tend to be highly differentiated the less homogeneous the network is.

### 1.5.2 Techniques for crop surveillance

Several studies have shown that traditional site prioritization techniques for crop surveillance, based on historical recordings of past epidemics or more complex approaches (Parnell et al., 2017), may be integrated (Sutrave et al., 2012; Sanatkar et al., 2015) or even outperformed (Martinetti and Soubeyrand, 2019) by network metrics.



Sutrave et al. (2012) explicitly explored network metrics for optimizing surveillance in a continental dynamic network among soybean producing counties. The authors explored several sentinel prioritization strategies. They observed that surveillance prediction errors were minimized by combining strength and "infection-frequency" (a metric considering past presence of the disease). In Chapter 2 it is illustrated how we simulated worldwide transport of airborne spores of another cereal pathogen, *P. graminis*, among squared wheat-cultivated cells ( $\approx 2,000 \text{ km}^2$ ) connected by air-mass trajectories computed via a Lagrangian model, obtaining a dynamic directed and weighted connectivity network. We showed that a surveillance strategy based either on betweenness or Set cover algorithm (a modified version of in-degree) minimizes the detection delay of simulated outbreaks. Furthermore, we showed that, given the high density of the epidemic network, surveillance strategies based on separate subset of nodes (representing countries) lead to a sub-optimal allocation of sentinels with respect to considering the complete set of nodes.

Network-based prioritization has been also applied in the case human-mediated crop parasites. Andersen et al. (2019), who studied the spread dynamics of a hypothetical pathogen affecting sweet potato, suggested that degree and betweenness may behave as well as other more complex metrics in node prioritization. Their importance was already noted in previous research on arthropods detection in grain networks (Hernandez Nopsa et al., 2015). Similarly, Buddenhagen et al. (2017) considered a local multi-mechanism potato trade network in Ecuador, and stated that high in-and out-degree nodes were to be surveilled first.

Recent research has been conducted to build more complex network metrics and assess their suitability for surveillance compared to traditional ones. For instance, Martinetti and Soubeyrand (2019) followed an approach similar to that of Herrera et al. (2016) to compare several surveillance options in the case of *X. fastidiosa* in southern France. In this case, Martinetti and Soubeyrand divided the region into  $1 \text{ km}^2$  cells, which represent nodes of the network, while edges' weights are built on notion of risk: each weight  $w_{i,j}$  is the product of a previously computed risk indicator  $r_i, r_j$  in the connected nodes  $i, j$ . They tested prioritisation techniques based on network metrics (such as k-shell, Kitsak et al., 2010; VoteRank Zhang et al., 2016; generalized random walk accessibility, or GRWA, Herrera et al., 2016) and others metrics, such as risk-based and random, to maximize the "detection" (*sensu* Herrera et al., 2016). Strategies based on VoteRank, GRWA and risk-based furnished the earliest detections.

## 1.6 Control

Control can be defined as the set of measures aimed to minimize the disease size or the occurrence of losses within a diseased population (Keeling and Rohani, 2011). Network based epidemic control consists in the preventive removal of a set of nodes (or of their edges) from the epidemic network, in the hope that this will minimize the spread (Shaw

and Pautasso, 2014).

### 1.6.1 Prioritize nodes to be removed

Traditional disease control methods at the population scale, including vaccines (Keeling and Rohani, 2011), may rely on network knowledge. Vaccination levels for reducing transmission usually depend on a threshold determining herd immunity (Fine, 1993), and are in turn calculated via  $R_0$ . This approach allows a first estimation of the immunization level needed to stem the epidemics in homogeneous networks. As effect of the vaccination of a proportion  $g$  of the nodes,  $R_0$  can be tuned below 1, thus extinguishing the spread of the disease (Eq. 1.5):

$$R_0 = \frac{\bar{k}\beta(1-g)}{\mu} \quad (1.5)$$

Network-based vaccination strategies have been studied to stem epidemics on both human and animal populations (Rushmore et al., 2014). It has been shown that neglecting highly heterogeneous networks structures may lead to ineffective or inefficient immunization strategies (Jeger et al., 2007). Identification of the nodes to be immunized via simple metrics, such as node strength, proved to help reducing the vaccine coverage threshold, as in the case of wild chimpanzees (Rushmore et al., 2014). Alternatively to vaccination, isolation (Keeling and Rohani, 2011) may as well target specific nodes (*i.e.* isolation of an infected node by removing all its edges) or random (*i.e.* removal of a fraction of edges independently on the nodes health status, such as social distancing in the case of Covid-19). The first case corresponds to increasing  $g$ , equivalently to an immunization or vaccination. In the second case, the strategy affects  $\bar{k}$ , whose reduction can bring the value of  $R_0$  below 1 (Block et al., 2020).

Several authors highlighted that control strategies should concentrate on immunizing highly connected nodes (Jeger et al., 2007; Shaw and Pautasso, 2014; Pastor-Satorras and Vespignani, 2002; Lloyd-Smith et al., 2005). In the case of human diseases, targeted immunization proved to achieve good results with rubella and mumps, whose heterogeneous transmission structure possesses scale-free properties (Pastor-Satorras and Vespignani, 2001), rather than with pertussis, whose transmission structure is more homogeneous (Trottier and Philippe, 2005). A comprehensive literature exists to support targeted action against sexually transmitted diseases, based on the scale-free properties of sexual partnership networks (Jeger et al., 2007).

Prioritization methods for targeted removal have been examined according to various network metrics, from the simplest (degree, betweenness or closeness, *i.e.* the reciprocal of the average distance from that node to all the other nodes; Sabidussi, 1966, Freeman, 1978) to more complex (random-walker based methods, Zhang et al., 2016). Methods based on k-shell decomposition have proved to produce optimal results compared to classic metrics

in identifying influential spreaders (De Arruda et al., 2014; Kitsak et al., 2010), and some generalizations to dynamic networks have been explored (Galimberti et al., 2018; Ciaperoni et al., 2020). In the specific case of spatial networks, where weights have the meaning of spatial distance, generalized accessibility metrics (Travençolo and Costa, 2008) have proved to be particularly suitable for targeted removal (De Arruda et al., 2014).

Despite the growing research on network-based disease protection strategies, there is a lack of clarification on how to distinguish metrics for determining suitable nodes for surveillance from those suitable for control (Holme, 2017), which can occasionally overlap. Many of the before-mentioned metrics to prioritize nodes to immunize for control are common to those to be detected for surveillance. However, in principle, nothing guarantees that an optimal set of sentinels coincides with an optimal set of candidate nodes to be immunized, and vice versa, since surveillance and control are linked to different properties. Figure 1.3 shows an example of a SI model running on a network, in which in-degree is used to optimize sentinels identification and betweenness is used to identify the best node to immunize.

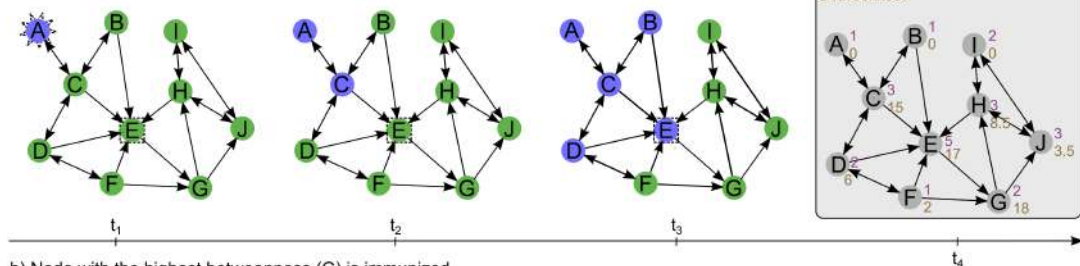
### 1.6.2 Application to crop protections

Optimal crop parasites control methods, *e.g.* allocation of phytosanitary treatments, biological control, resistant varieties or even eradication of suspected hosts, only occasionally is explicitly investigated via network properties, as instead it is the case of targeted vaccination or isolation in humans and animals (Jeger et al., 2007). In one of these example, Strona et al. (2017) found that removing nodes with the highest PageRank (Page, 1998) allows to reduce the size of the largest connected component (the largest set of nodes for which at least one path to any other node in the set exists) faster than using the degree or a random sampling, thus reducing the maximum spread of *X. fastidiosa*.

By contrast, rules of thumb (*sensu* Chadès et al., 2011) loosely related to network-thinking have found greater application. For instance, it has been suggested that prevention planning may discourage growth of certain crop in marginal areas, which afford little profit but may provide chance for disease flow (Margosian et al., 2009). In this case, one could visualize main production areas as nodes, connected by edges representing marginal areas which allow the spread of the disease. Under this perspective, this disease control strategy translates into an edge-removal problem. Shaw and Pautasso (2014) reported successful experiences following this recommendation in reducing yellow rust of wheat, provoked by *P. striiformis*, in the northwest of China (Lu et al., 2011). It has been commonly observed that group of crops can remain uninfected if surrounded by immune individuals, as in the case of the soil-borne pathogen *Rhizoctonia solani* (Bailey et al., 2000). In this case, cultivated areas can play as nodes, immune areas as immunized nodes, while potential disease pathways as edges connecting them. In general, breaking the mono-cultural continuum into disconnected components reduces the opportunity of short-distance pathogen spread (Elton, 1958). Given the definition of betweenness, we argue that this metric may support the identification of those

### Network-based strategies to tackle parasites spread

a) Node with the highest in-degree (E) is the sentinel



b) Node with the highest betweenness (G) is immunized

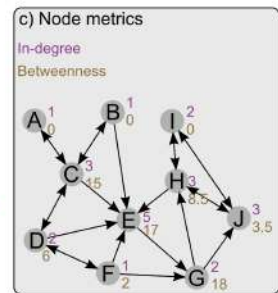
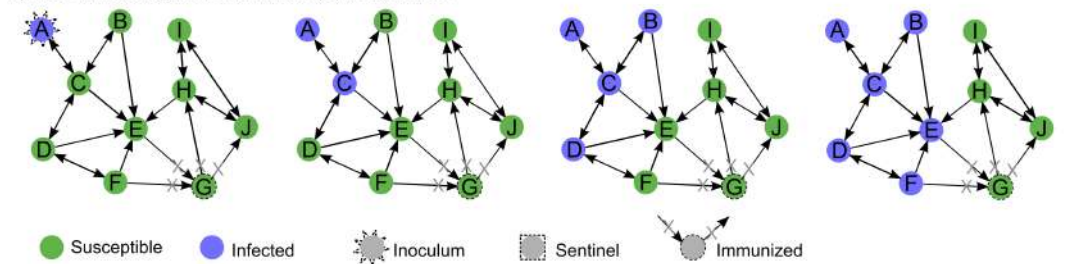


Figure 1.3 – A network representing potential epidemic pathways and strategies based on node metrics to face parasites spread. In panels a) and b) the same SI epidemic process starting from node A (i.e., the inoculated node) is shown. At each time step, each susceptible node becomes infected if it has at least one infected neighbor pointing to it, without recovery. In panel a) node E is a sentinel: it easily detects the disease (i.e., it becomes infected) because of its high in-degree (i.e., the number of edges pointing to it). In panel b) node G is immunized through targeted removal of its edges: this way, any outbreak that starts from the left sub-network (nodes A, B, C, D, E, F) cannot reach the right sub-network (I, H, J), because of its high betweenness. Note that removal of node E result in poor control performances, since its removal is irrelevant with respect to disease spread. Similarly, G is an inefficient sentinel, since it has a low degree and it is reached late in an epidemic process. Panel c) shows the values of in-degree and betweenness computed for each node.

nodes whose removal would more efficiently break the crop network (Fig. 1.3b).

These rules of thumb could gain more consistency if accompanied by network-based prioritization methods, such as a combination of different network metrics. Xing et al. (2020) evaluated the epidemic network given by a hypothetical pathogen affecting major crops across the world and suggested that an index weighting several network metrics (strength, nearest neighbors' degree, betweenness, eigenvector centrality) may give a comprehensive view of the nodes' contribution to connectivity. Additionally, Andersen et al. (2019) suggested that, as well as for surveillance, node characterized by high (out-)degree may be identified as influential spreaders and so prioritized to be immunized.

## 1.7 Conclusion

Although network-thinking is finding wider application in crop protection, its level of diffusion is unbalanced among surveillance and control. Network metrics, such as degree, strength, betweenness, VoteRank, and GRWA are more diffused to support surveillance, while control design in agriculture seems to require a multidisciplinary approach.

This imbalance may be due to several factors. A typical problem about the application of theoretical network tools in real cases is the adequate definition of nodes and edges. For example, network algorithms are very often conceived for unweighted, undirected and static networks, while real epidemic networks are frequently asymmetric, weighted and/or temporal. Generalizations may miss or distort the original meaning. To meet this need, a growing amount of research has been conducted to broaden the application of these metrics to real matrices (Fagiolo, 2007; Wang et al., 2008; Holme, 2005; Galimberti et al., 2018), declining them depending on the physical meaning of the edges, but generally with a loss of straightforwardness.

The challenge of spatial scale is an inveterate obstacle that network-thinking must address in order to expand its applications in agriculture, as pointed out by Shaw and Pautasso (2014). Edges may be defined to describe extremely long-distance transport events along continents, in the order of  $10^7$  m (8,000 km being distance from South Africa to Australia, perhaps the farthest documented periodical incursion of *P. graminis*; Visser et al., 2019), as well as cellular interaction between the hosts and its pathogens, on the order of  $10^{-5}$  m (26.4  $\mu$ m being the spore diameter of *P. graminis*; Eversmeyer and Kramer, 2000), passing through the scale of a host ( $10^0$  m being size of a wheat stem) or a population ( $10^3$  m being the order of the size of an average US farm; Ritchie and Roser, 2021). Moreover, the geographic scale of a node has crucial consequences for the meaning of applying network-based protection practices: to surveil/control a node means to completely monitor/immunize all the hosts within. Eventually, the choice of a node size is a compromise between the resolution of pathogen dispersal mechanisms (setting the biophysical constraints of the network) and the protection measures applicability (setting the management constraints of the network), in

addition to ensure computational viability.

Broadening the scope of investigation, social acceptance may be another reason for the gap in application of network metrics among surveillance and control strategies. This gap may generally reflect a greater social impact of preventive crop control policies in a broad sense, not necessarily related to networks (Eriksson et al., 2019). Whereas optimal surveillance simply requires to geographically arrange the monitoring activities across a network, control thought node treatment implies an invasive intervention, such as the application of phytosanitary products or even the eradication of healthy hosts. Social acceptability of policies aimed at controlling alien pathogen invasions, as *X. fastidiosa* (Strona et al., 2017), relies on direct and indirect risk experiences and problem awareness of stakeholders, which may lead to a social underestimation of the probability of an invasion versus the concrete loss of a cultivated unit. Understanding the factors driving social acceptability of crop protection practices may be the key to the implementation of effective control methods (Marzano et al., 2017), in particular network-based strategies, with a positive feedback on the theoretical research, and *vice versa*.

## Bibliography

- Adam, D. et al. (2021). How far will global population rise? researchers can't agree. *Nature*, 597(7877):462–465.
- Andersen, K., Buddenhagen, C., Rachkara, P., Gibson, R., Kalule, S., Phillips, D., and Garrett, K. (2019). Modeling epidemics in seed systems and landscapes to guide management strategies: the case of sweet potato in northern uganda. *Phytopathology*, 109(9):1519–1532.
- Aylor, D. (2017). *Aerial Dispersal of Pollen and Spores*. The American Phytopathological Society.
- Bailey, D. J., Otten, W., and Gilligan, C. A. (2000). Saprotrophic invasion by the soil-borne fungal plant pathogen rhizoctonia solani and percolation thresholds. *The New Phytologist*, 146(3):535–544.
- Barabási, A.-L. (2016). *Network science*. Cambridge university press.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., and Mills, M. C. (2020). Social network-based distancing strategies to flatten the covid-19 curve in a post-lockdown world. *Nature Human Behaviour*, 4(6):588–596.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4-5):175–308.

- Börner, K., Sanyal, S., and Vespignani, A. (2007). Network science. *Annual review of information science and technology*, 41(1):537–607.
- Brockmann, D. and Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342(6164):1337–1342.
- Buddenhagen, C., Hernandez Nopsa, J.-F., Andersen, K. F., Andrade-Piedra, J., Forbes, G.-A., Kromann, P., Thomas-Sharma, S., Useche, P., and Garrett, K. (2017). Epidemic network analysis for mitigation of invasive pathogens in seed systems: Potato in Ecuador. *Phytopathology*, 107(10):1209–1218.
- Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K. A., Giovani, B., Saunders, D., Kamoun, S., Legg, J., Verdier, V., Lessel, J., et al. (2019). A global surveillance system for crop diseases. *Science*, 364(6447):1237–1239.
- Chadès, I., Martin, T. G., Nicol, S., Burgman, M. A., Possingham, H. P., and Buckley, Y. M. (2011). General rules for managing and surveying networks of pests, diseases, and endangered species. *Proceedings of the National Academy of Sciences*, 108(20):8323–8328.
- Ciaperoni, M., Galimberti, E., Bonchi, F., Cattuto, C., Gullo, F., and Barrat, A. (2020). Relevance of temporal cores for epidemic spread in temporal networks. *Scientific reports*, 10(1):1–15.
- Corredor-Moreno, P. and Saunders, D. G. (2020). Expecting the unexpected: factors influencing the emergence of fungal and oomycete plant pathogens. *New Phytologist*, 225(1):118–125.
- Cunniffe, N. J., Stutt, R. O., DeSimone, R. E., Gottwald, T. R., and Gilligan, C. A. (2015). Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. *PLoS computational biology*, 11(4):e1004211.
- De Arruda, G. F., Barbieri, A. L., Rodríguez, P. M., Rodrigues, F. A., Moreno, Y., and Costa, L. D. F. (2014). Role of centrality for the identification of influential spreaders in complex networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 90(3):1–17.
- de la Fuente, B., Saura, S., and Beck, P. S. (2018). Predicting the spread of an invasive tree pest: The pine wood nematode in Southern Europe. *Journal of Applied Ecology*, 55(5):2374–2385.
- Della Rossa, F., Salzano, D., Di Meglio, A., De Lellis, F., Coraggio, M., Calabrese, C., Guarino, A., Cardona-Rivera, R., De Lellis, P., Liuzza, D., et al. (2020). A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nature communications*, 11(1):1–9.

- Dietz, K. (1993). The estimation of the basic reproduction number for infectious diseases. *Statistical methods in medical research*, 2(1):23–41.
- Dubé, C., Ribble, C., Kelton, D., McNab, B., et al. (2011). Introduction to network analysis and its implications for animal disease modelling. *Revue Scientifique et Technique-OIE*, 30(2):425.
- Elton, C. S. (1958). *The ecology of invasions by animals and plants*. John Wiley: New York, NY, USA.
- Eriksson, L., Boberg, J., Cech, T. L., Corcobado, T., Desprez-Loustau, M.-L., Hietala, A. M., Jung, M. H., Jung, T., Lehtijarvi, H. T. D., Oskay, F., et al. (2019). Invasive forest pathogens in europe: Cross-country variation in public awareness but consistency in policy acceptability. *Ambio*, 48(1):1–12.
- Eversmeyer, M. and Kramer, C. (2000). Epidemiology of wheat leaf and stem rust in the central great plains of the usa. *Annual review of Phytopathology*, 38(1):491–513.
- Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E*, 76(2):026107.
- Fine, P. E. (1993). Herd immunity: history, theory, practice. *Epidemiologic reviews*, 15(2):265–302.
- Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O’Connell, C., Ray, D. K., West, P. C., et al. (2011). Solutions for a cultivated planet. *Nature*, 478(7369):337–342.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.
- Galimberti, E., Barrat, A., Bonchi, F., Cattuto, C., and Gullo, F. (2018). Mining (maximal) span-cores from temporal networks. In *Proceedings of the 27th ACM international Conference on Information and Knowledge Management*, pages 107–116.
- Garrett, K., Alcalá-Briseño, R., Andersen, K., Buddenhagen, C., Choudhury, R., Fulton, J., Hernandez Nopsa, J., Poudel, R., and Xing, Y. (2018). Network analysis: A systems framework to address grand challenges in plant pathology. *Annual review of phytopathology*, 56:559–580.
- Gilligan, C. A. (2008). Sustainable agriculture and plant diseases: an epidemiological perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1492):741–759.
- Harwood, T. D., Xu, X., Pautasso, M., Jeger, M. J., and Shaw, M. W. (2009). Epidemiological risk assessment using linked network and grid based modelling: *Phytophthora ramorum* and *phytophthora kernoviae* in the uk. *Ecological Modelling*, 220(23):3353–3361.



- Hernandez Nopsa, J. F., Daglish, G. J., Hagstrum, D. W., Leslie, J. F., Phillips, T. W., Scoglio, C., Thomas-Sharma, S., Walter, G. H., and Garrett, K. A. (2015). Ecological networks in stored grain: Key postharvest nodes for emerging pests, pathogens, and mycotoxins. *BioScience*, 65(10):985–1002.
- Herrera, J. L., Srinivasan, R., Brownstein, J. S., Galvani, A. P., and Meyers, L. A. (2016). Disease surveillance on complex social networks. *PLoS computational biology*, 12(7):e1004928.
- Holme, P. (2005). Network reachability of real-world contact sequences. *Physical Review E*, 71(4):046119.
- Holme, P. (2017). Three faces of node importance in network epidemiology: Exact results for small graphs. *Physical Review E*, 96(6):062305.
- Holme, P. (2018). Objective measures for sentinel surveillance in network epidemiology. *Physical Review E*, 98(2):022313.
- Hulme, P. E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of applied ecology*, 46(1):10–18.
- Jeger, M. J., Pautasso, M., Holdenrieder, O., and Shaw, M. W. (2007). Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist*, 174(2):279–297.
- Jordano, P. (2017). What is long-distance dispersal? and a taxonomy of dispersal events. *Journal of Ecology*, 105(1):75–84.
- Kao, R. R., Green, D. M., Johnson, J., and Kiss, I. Z. (2007). Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the uk. *Journal of the Royal Society Interface*, 4(16):907–916.
- Keeling, M. J. and Rohani, P. (2011). *Modeling infectious diseases in humans and animals*. Princeton university press.
- Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721.
- Khoury, C. K., Bjorkman, A. D., Dempewolf, H., Ramirez-Villegas, J., Guarino, L., Jarvis, A., Rieseberg, L. H., and Struik, P. C. (2014). Increasing homogeneity in global food supplies and the implications for food security. *Proceedings of the national Academy of Sciences*, 111(11):4001–4006.
- Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., and Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature physics*, 6(11):888–893.

- Lee, B. X., Kjaerulf, F., Turner, S., Cohen, L., Donnelly, P. D., Muggah, R., Davis, R., Realini, A., Kieselbach, B., MacGregor, L. S., et al. (2016). Transforming our world: implementing the 2030 agenda through sustainable development goal indicators. *Journal of public health policy*, 37(1):13–31.
- Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., and Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359.
- Lü, L., Chen, D., Ren, X.-L., Zhang, Q.-M., Zhang, Y.-C., and Zhou, T. (2016). Vital nodes identification in complex networks. *Physics reports*, 650:1–63.
- Lu, N., Wang, J., Chen, X., Zhan, G., Chen, C., Huang, L., and Kang, Z. (2011). Spatial genetic diversity and interregional spread of *puccinia striiformis* f. sp. *tritici* in northwest china. *European journal of plant pathology*, 131(4):685–693.
- Margosian, M. L., Garrett, K. A., Hutchinson, J. S., and With, K. A. (2009). Connectivity of the american agricultural landscape: assessing the national risk of crop pest and disease spread. *BioScience*, 59(2):141–151.
- Martelli, G., Boscia, D., Porcelli, F., and Saponari, M. (2016). The olive quick decline syndrome in south-east italy: a threatening phytosanitary emergency. *European Journal of Plant Pathology*, 144(2):235–243.
- Martinetti, D. and Soubeyrand, S. (2019). Identifying lookouts for epidemio-surveillance: application to the emergence of *xylella fastidiosa* in france. *Phytopathology*, 109(2):265–276.
- Marzano, M., Allen, W., Haight, R., Holmes, T., Keskitalo, E. C. H., Langer, E., Shadbolt, M., Urquhart, J., and Dandy, N. (2017). The role of the social sciences and economics in understanding and informing tree biosecurity policy and planning: a global summary and synthesis. *Biological invasions*, 19(11):3317–3332.
- Mastin, A. J., Gottwald, T. R., van den Bosch, F., Cunniffe, N. J., and Parnell, S. (2020). Optimising risk-based surveillance for early detection of invasive plant pathogens. *PLoS biology*, 18(10):e3000863.
- Meyer, M., Cox, J. A., Hitchings, M. D., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nature Plants*, 3(10):780–786.
- Morris, C. E., Geniaux, G., Nédellec, C., Sauvion, N., and Soubeyrand, S. (2022). One health concepts and challenges for surveillance, forecasting and mitigation of plant disease beyond the traditional scope of crop production. *Plant pathology*.

- Moslonka-Lefebvre, M., Finley, A., Dorigatti, I., Dehnen-Schmutz, K., Harwood, T., Jeger, M. J., Xu, X., Holdenrieder, O., and Pautasso, M. (2011). Networks in plant epidemiology: from genes to landscapes, countries, and continents. *Phytopathology*, 101(4):392–403.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.
- Page, L. (1998). The pagerank citation ranking: Bringing order to the web. technical report. *Stanford Digital Library Technologies Project, 1998*.
- Parnell, S., van den Bosch, F., Gottwald, T., and Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: an epidemiological perspective. *Annual review of phytopathology*, 55:591–610.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of modern physics*, 87(3):925.
- Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200.
- Pastor-Satorras, R. and Vespignani, A. (2002). Immunization of complex networks. *Physical review E*, 65(3):036104.
- Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., and Trapman, P. (2015). Eight challenges for network epidemic models. *Epidemics*, 10:58–62.
- Radici, A., Bevacqua, D., Miele, L., and Martinetti, D. (2023). Network-thinking to optimize surveillance and control of crop parasites. a review. *arXiv preprint arXiv:2310.07442*.
- Ristaino, J. B., Anderson, P. K., Bebber, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., et al. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences*, 118(23).
- Ritchie, H. and Roser, M. (2021). Farm size. *Our World in Data*.
- Rushmore, J., Caillaud, D., Hall, R. J., Stumpf, R. M., Meyers, L. A., and Altizer, S. (2014). Network-based vaccination improves prospects for disease control in wild chimpanzees. *Journal of the Royal Society Interface*, 11(97):20140349.
- Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4):581–603.
- Sanatkar, M., Scoglio, C., Natarajan, B., Isard, S., and Garrett, K. (2015). History, epidemic evolution, and model burn-in for a network of annual invasion: Soybean rust. *Phytopathology*, 105(7):947–955.

- Santini, A., Liebhold, A., Migliorini, D., and Woodward, S. (2018). Tracing the role of human civilization in the globalization of plant pathogens. *The ISME journal*, 12(3):647–652.
- Saunders, H. A. and Schwartz, J.-M. (2021). Covid-19 vaccination strategies depend on the underlying network of social interactions. *Scientific reports*, 11(1):1–10.
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology & evolution*, 3(3):430–439.
- Shaw, M. and Pautasso, M. (2014). Networks and plant disease management: Concepts and applications. *Annual Review of Phytopathology*, 52:477–493.
- Strona, G., Carstens, C. J., and Beck, P. S. (2017). Network analysis reveals why xylella fastidiosa will persist in europe. *Scientific Reports*, 7(1):1–8.
- Sutrave, S., Scoglio, C., Isard, S. A., Hutchinson, J. S., and Garrett, K. A. (2012). Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS One*, 7(6):e37793.
- Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the national academy of sciences*, 108(50):20260–20264.
- Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418(6898):671–677.
- Travençolo, B. A. N. and Costa, L. d. F. (2008). Accessibility in complex networks. *Physics Letters A*, 373(1):89–95.
- Trottier, H. and Philippe, P. (2005). Scaling properties of childhood infectious diseases epidemics before and after mass vaccination in canada. *Journal of theoretical biology*, 235(3):326–337.
- Tudi, M., Daniel Ruan, H., Wang, L., Lyu, J., Sadler, R., Connell, D., Chu, C., and Phung, D. T. (2021). Agriculture development, pesticide application and its impact on the environment. *International journal of environmental research and public health*, 18(3):1112.
- Visser, B., Meyer, M., Park, R. F., Gilligan, C. A., Burgin, L. E., Hort, M. C., Hodson, D. P., and Pretorius, Z. A. (2019). Microsatellite analysis and urediniospore dispersal simulations support the movement of puccinia graminis f. Sp. Tritici from southern Africa to Australia. *Phytopathology*, 109(1):133–144.
- Vitousek, P. M. (1994). Beyond global warming: ecology and global change. *Ecology*, 75(7):1861–1876.

- Wang, H., Hernandez, J. M., and Van Mieghem, P. (2008). Betweenness centrality in a weighted network. *Physical Review E*, 77(4):046105.
- Xing, Y., Hernandez Nopsa, J. F., Andersen, K. F., Andrade-Piedra, J. L., Beed, F. D., Blomme, G., Carvajal-Yepes, M., Coyne, D. L., Cuellar, W. J., Forbes, G. A., et al. (2020). Global cropland connectivity: A risk factor for invasion and saturation by emerging pathogens and pests. *BioScience*, 70(9):744–758.
- Zhang, J. X., Chen, D. B., Dong, Q., and Zhao, Z. D. (2016). Identifying a set of influential spreaders in complex networks. *Scientific Reports*, 6(February):1–10.

**Key points:**

1. Networks are mathematical objects composed of nodes connected by edges. They have been used in epidemiology to model the spread of parasites and pests, under the assumption that nodes represent hosts and edges represent potential epidemic contacts.
2. Network scientists have analysed network properties to identify those nodes which may play a relevant role in epidemic management. For example, those nodes which, if removed, reduce the overall epidemic size in the network (epidemic control); or those which are usually infected very soon, thus helping detecting the emergence of new diseases (epidemic surveillance).
3. A growing number of plant-parasite interactions have been successfully represented by networks and this accomplishment has translated into the design of network-based plant epidemic surveillance strategies.
4. By contrast, compared to the extensive applications of network-based epidemic control in human and animal epidemiology, fewer examples have been proposed in crop and plant protection literature.

**Perspectives:**

1. Adequately define what nodes and edges represent in the case of plant parasites is a modelling compromise between many factors (for example, geographic resolution and computational capability). Network science is currently grappling with these complexities to broaden the applicability of these methods.
2. Insufficient social acceptability might be one of the factors contributing to the limited adoption of network-based control strategies, and this issue warrants further investigation.

## Chapter 2

# Early-detection surveillance for stem rust of wheat: insights from a global epidemic network based on airborne connectivity and host phenology

Man's mind cannot grasp the causes of events in their completeness, but the desire to find those causes is implanted in man's soul. And without considering the multiplicity and complexity of the conditions any one of which taken separately may seem to be the cause, he snatches at the first approximation to a cause that seems to him intelligible and says: 'This is the cause!'

– Leo Tolstoy, *War and Peace*

### Résumé

La rouille noire du blé, causée par le pathogène aérien *Puccinia graminis*, est une maladie ré-émergente des cultures qui représente une préoccupation majeure pour la sécurité alimentaire mondiale. Le transport potentiel sur de longues distances par le vent au-dessus

---

This chapter has been published as a research article (Radici et al., 2022).

d'un hôte réparti dans le monde entier représente un défi pour une surveillance et un contrôle efficaces de cette maladie. Pour surveiller cette maladie, nous avons créé un réseau épidémique mondial en combinant *i*) des simulations Lagrangiennes de trajectoires de masses d'air calculées avec le modèle HYSPLIT de la NOAA ; *ii*) l'utilisation du sol à partir du modèle Map Spatial Production Allocation Model et *iii*) les conditions météorologiques et environnementales qui affectent les processus biophysiques impliqués dans la biologie des spores de *P. graminis*. Nos résultats permettent de reconstruire la bien connue "*Puccinia* pathway" nord-américaine et suggèrent l'existence d'autres voies sous-continentales à l'échelle mondiale. Nous utilisons la théorie des réseaux pour concevoir des stratégies de surveillance visant à la détection précoce des épidémies tout en minimisant le nombre de nœuds à surveiller (également appelés "sentinelles"). Nous avons constaté que l'algorithme "Set cover", en raison de la connectivité élevée du réseau (densité = 0,4 %), est plus performant qu'un certain nombre d'autres métriques de réseau et nous permet d'identifier un ensemble optimal de sentinelles (1 % des nœuds du réseau) pour surveiller 50 % du réseau. Nos résultats montrent également qu'il est possible de concevoir des plans de surveillance efficaces pour la rouille de la tige du blé, mais qu'ils doivent tenir compte de l'échelle géographique réelle du processus épidémiologique sous-jacent et qu'ils nécessitent une approche internationale et transfrontalière.

**Mots-clés:** Dispersion à longue distance, rouille noire du blé, *Puccinia graminis*, réseaux complexes, épidémiosurveillance, détection précoce

## 2.1 Abstract

Stem rust of wheat, caused by the airborne pathogen *Puccinia graminis*, is a re-emerging crop disease representing a major concern to global food security. Potential long-distance transport by wind over a worldwide distributed host represents a challenge to effective surveillance and control of this disease. To monitor this disease, we have created a global epidemic network for stem rust of wheat combining *i*) Lagrangian simulations of air-mass trajectories computed with the HYSPLIT model; *ii*) land use from the Map Spatial Production Allocation Model and *iii*) meteorological and environmental conditions that are known to affect bio-physical processes involved in the biology of *P. graminis* spores. Our findings are in agreement with the well known north-American "*Puccinia* pathway" and suggest the existence of other sub-continental pathways at the global scale. We use network theory to conceive surveillance strategies aimed at early detection of outbreaks while minimizing the number of nodes to be surveilled (also referred to as sentinels). We found that the set cover algorithm, due the high average connectivity of the network (density = 0.4%), performs better than a number of other network metrics and permits us to identify an optimal sentinel set (1% of the network nodes) to surveil 50% of the network. Our results also show that effective surveillance plans for stem rust of wheat can be designed, but that they need to



account for the actual geographical scale of the underlying epidemiological process and call for an international and trans-boundary approach.

**Keywords:** Long-distance dispersal, stem rust of wheat, *Puccinia graminis*, network, surveillance, early detection

## 2.2 Introduction

Epidemics caused by airborne pathogens represent an inveterate challenge to agricultural management (Mahaffee and Stoll, 2016). Moreover, the transition to monocultures characterized by low genetic diversity has made the global farming system less resilient to pathogens (Corredor-Moreno and Saunders, 2020). Airborne pathogens, capable of long-distance transport, create a network of connections among globally diffused crops. For example, *Puccinia graminis*, causing stem rust of wheat, is seasonally dispersed from northern Mexico up to Canada, along the “*Puccinia* pathway” (Aylor, 2003). On the other hand, a single-leap event, the hurricane Ivan, transported *Phakopsora pachyrhizi* spores from Colombia to Alabama, introducing soybean rust in North America (Schneider et al., 2005; Isard et al., 2005). The risk of losses in food production requires to take action against the diffusion of alarming pathogens, such as *P. graminis* (Park et al., 2011; Saunders et al., 2019; Corredor-Moreno and Saunders, 2020) and species of the same genus (Hovmøller et al., 2008). Its aerobiology has been largely studied in the last century (Zadoks, 1967; Burrage, 1970; Maddison and Manners, 1972; Aylor, 1986) and experimental procedures have been recently developed for studying three-dimensional transport and detection of spores (Damialis et al., 2017; Schmale III and Ross, 2015). However, the use of mathematical models to simulate spore transport by wind at the continental scale is very recent (Meyer et al., 2017b,a; Visser et al., 2019; Allen-Sader et al., 2019; Prank et al., 2019; Wang et al., 2021). In fact, the use of such models would allow researchers, policy makers and farmers to design surveillance strategies (Cunniffe et al., 2015; Parnell et al., 2017; Ristaino et al., 2021) capable of detecting outbreaks and take effective countermeasures within an acceptable time-limit (e.g., phytosanitary intervention; Allen-Sader et al., 2019).

Meyer et al. (2017b) assessed the risk of long distance dispersal of *P. graminis* in East Africa and the Middle East via a Lagrangian particle dispersion model. The same modelling framework was then used to investigate the possible origin of virulent strains found in Ethiopia (Meyer et al., 2017a). Allen-Sader et al. (2019) developed a decision support system integrating spore dispersal to help optimize fungicide allocation. Prank et al. (2019) investigated the impacts of climate change on worldwide spores transport patterns. Suttrave et al. (2012) designed a network-based surveillance-system for *P. pachyrhizi*, at the subcontinental scale, by including wind direction and intensity. Despite these advances, to our knowledge, models of spore dispersal have never been coupled to network analysis to design optimal surveillance strategies for *P. graminis* at the global scale.

In this current study we constructed a time-variant connectivity network for *P. graminis* spores at the worldwide scale explicitly considering wind patterns, host phenology and meteorological conditions affecting spores aerobiology. We used biophysical models and graph theory to identify the most susceptible worldwide regions and to reconstruct the epidemic movement through different subcontinents. We validated our findings using available knowledge regarding the North-American “*Puccinia* pathway” (Aylor, 2003). Lastly, we identified those nodes of the network (i.e., regions of the world) that, when monitored, enable early detection of an outbreak. We are confident that our approach is sufficiently generic to be applied to other airborne plant pathogens, provided that basic knowledge on the pathogen aerobiology, host physiology and distribution are available.

## 2.3 Materials and methods

### 2.3.1 Case study

*Puccinia graminis* f. sp. *tritici* is a heteroecious airborne fungal pathogen responsible for stem rust of wheat. Wheat, the main host of *P. graminis*, constitutes a staple food for a great proportion of human population. It covers  $2.15 \times 10^8$  hectares worldwide (1.4% of earth surface), representing the most abundant agricultural type of land cover (FAO, 2021).

In recent decades, yield losses caused by stem rust have been limited by planting resistant cultivars and fungicides application. Nonetheless, the emergence of new strains overcoming plant resistance may provoke severe yield losses, accounting for up to 50-90% of the wheat production at the regional scale (Huerta-Espino et al., 2014; Prank et al., 2019). An extraordinary outbreak in Ethiopia in 2013-2014 caused a complete yield loss in some cultivars (Olivera et al., 2015), while concerns were raised in Europe after the detection of new virulent strains capable of infecting previously resistant cultivars (Bhattacharya, 2017).

### 2.3.2 Geographic domain and air mass trajectories

We extracted the worldwide wheat distribution from the MapSPAM database (International Food Policy Research Institute, 2019) and computed the percentage of wheat cover on a regular grid with a resolution of  $0.5^\circ$ . We accounted only for cells containing at least 2% of wheat land cover. The resulting database contains 7,814 cells (see Section SI A.1.1). We calculated backward air-mass Lagrangian three-dimensional trajectories of 120 hours, considered as a reasonable maximal lifetime of *Puccinia* spores by previous authors (Meyer et al., 2017a), with the HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) model (Draxler and Hess, 1998) from January 1<sup>st</sup>, 2013 to December 31<sup>st</sup>, 2018, using atmospheric GDAS data at a spatial resolution of  $0.5^\circ$ . We ran simulations from the centroids of each cell at 00:00, 06:00, 12:00 and 18:00 UTC +0, at an above-ground altitude set to the minimum between the base of the cloud and the mixed layer depth. Along each simulated

trajectory, we recorded atmospheric variables at an hourly frequency. Overall, out of more than  $6.8 \times 10^7$  potential trajectories, we computed only those  $1.6 \times 10^6$  satisfying the criteria of host availability and environmental suitability specified in Sections 2.3.3 and 2.3.3 (see also Section SI A.1.2).

### 2.3.3 The connectivity network

We built a time-variant connectivity network for stem rust explicitly considering host (i.e. wheat) phenological phases, environmental conditions affecting spore aerobiology (Aylor, 1999) and infection. In this network, nodes represent cells composing the worldwide wheat producing regions, while edges are computed from air-mass trajectories in order to model the dispersal of airborne spores. The nodes (or cells) of the networks remain fixed, while edges are re-computed for each of the 8,764 (4 times a day for 6 years) simulations.

More specifically, edges result from the application of a set of biophysical filters to identify those air-mass movements that correspond to transport events (figure 2.1). We assume that, for a given time  $t$ , an edge exists between an arrival cell  $j$  and any cell  $i$  if the following conditions are satisfied:

1. In cell  $i$ , which at time  $t - x$  ( $x \in \{1, \dots, 120\}$  h) is crossed by the air-mass trajectory arriving at  $j$  at time  $t$ :
  - The host is “available”, i.e. it is present and in a favourable phenological state for infection and sporulation, see paragraph *Host availability*.
  - Environmental conditions are compatible for spore release, see paragraph *Spore release*.
  - The altitude of the trajectory is lower than the planetary boundary layer.
2. Environmental conditions along the trajectory between  $i$  and  $j$  allow spore survival, see paragraph *Aerial spore transport and survival*.
3. The host in cell  $j$  is “susceptible”, which means:
  - The host is “available”.
  - Environmental conditions are favourable for spore deposition and host infection, see paragraph *Spore deposition and host infection*.

#### Host availability

Depending on local climate, wheat is cultivated and harvested at different times across the world. We used a growing degree-day model (McMaster and Smika, 1988; McMaster and Wilhelm, 1997) to compute the period of host availability for each year in each cell of the domain as a function of the year temperature conditions. We used a model conceived for

the US (McMaster and Smika, 1988) and we adjusted it for the Southern Hemisphere (i.e., initializing it at July 1<sup>st</sup>). For tropical countries, we consulted case by case the calendar of the prevailing season provided in the FAO country briefs (FAO, 2021) (see Section SI A.1.3).

### **Spore release**

Spore release in the air column is promoted under unstable atmospheric conditions (Levetin, 2015; Oneto et al., 2020). Following Meyer et al. (2017b), we assumed that release occurs between 9:00 and 15:00, provided precipitations are lower than  $2.54 \text{ mm/h}$  (Allen-Sader et al., 2019). Furthermore, the rate of spore release is affected by temperature (Prank et al., 2019). Eventually, a probability  $P^1(T)$  of release is defined as a function of temperature (see Section SI A.1.5).

### **Aerial spore transport and survival**

Only a fraction of the spores released at canopy level actually enters in the atmospheric layers leading to long-distance transportation. The rest may be dispersed in the air column within few kilometers under turbulent atmospheric conditions (Levetin, 2015; Aylor, 2017). Hence, we considered as suitable release sites those that are crossed by an air-mass trajectory within the mixed layer depth (therefore, able to drag spores distributed in the air column). Furthermore, during the aerial phase, spores must endure critical environmental conditions (Levetin, 2015). The limiting factors affecting the survival of *P. graminis* spores are the exposure to UV radiation (Meyer et al., 2017b) and the washout by rain. We then defined survival probabilities to UV radiation  $P^2(UV)$  and to washout  $P^3(R)$  (see Section SI A.1.5).

### **Spore deposition and host infection**

Since precipitation is the main responsible for spore deposition and infection (Levetin, 2015; Morris et al., 2013; Nagarajan and Singh, 1990; Roelfs, 1992; Rowell et al., 1966; Li et al., 2009), we assumed that an edge can point an arrival cell  $j$  at time  $t$  only if precipitation occurs. In fact, even if spore deposition might occur also in dry conditions (Emerson et al., 2020; Slinn, 1977), wet deposition provides a better environment for the development of infection and it is usually treated as the most important element. Following deposition, infection requires specific environmental conditions (Allen-Sader et al., 2019; Meyer et al., 2017b; Baiocco et al., 2021): in the three following days, spores should enter the phase of germination, requiring dark conditions and mild temperatures, and appressorium formation, requiring sunlight and warm temperature (Roelfs, 1992). We then used these conditions to identify times and sites where infection can occur after deposition (see Section SI A.1.4).

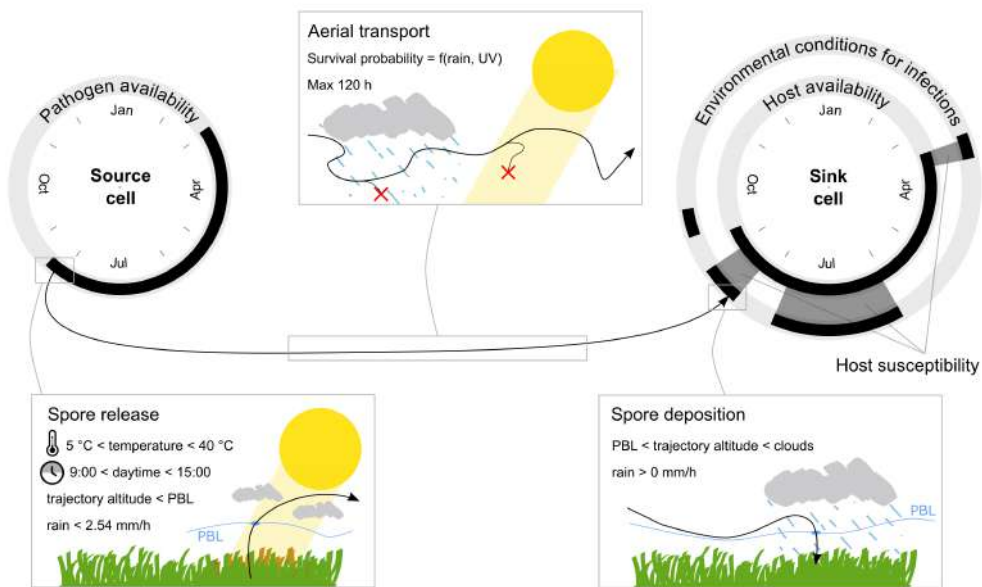


Figure 2.1 – Simplified overview of the transport process between a source (i.e., release) and a sink (i.e., arrival) cell. Pathogen is assumed to be available in the release cell when the host is “available”. “Susceptibility” in the arrival cell is determined by the concurrence of host “availability” and environmental conditions favourable to infections. Conditions for spore release, transport and deposition are summarized in the three boxes along the trajectory.

### Time-variant connectivity networks

In order to represent the connectivity between wheat producing regions, we took advantage of the mathematical formalism of networks (Choufany et al., 2021a), by describing the global epidemic network of stem rust via 6-hourly connectivity networks  $A^t$  whose elements  $A_{ij}^t$  indicate the probability  $P_{ijt}$  of a successful transport event occurring at time  $t$  from node  $i$  to node  $j$ . The probability  $P_{ijt}$  is computed as the product of the mutually independent probabilities  $P_i^1(T)$  of sporulation at release site,  $P_{ij}^2(UV)$  of survival to harmful UV radiation and  $P_{ij}^3(R)$  of washout, provided the concurrence of host availability at site  $i$  and precipitation at site  $j$ . Finally, we aggregate the 8,764 networks to obtain a weighted, directed air-mass connectivity network  $C$  (see Section SI A.1.5) and used a cluster detection algorithm (Clauset et al., 2004) to identify groups of nodes with similar connectivity patterns.

### 2.3.4 Model validation: recovering the “*Puccinia* pathway” in North America

Since the direct observation of long-distance spores transport is beyond current technological capacities, we tested the validity of our model checking if we could use it to identify the well known North-American “*Puccinia* pathway”.

We considered the yearly northward movement of the stem rust infection front in North America by tracking the average onset date ( $\pm 1$  standard deviation) during the 1922–1992 period along the 97<sup>th</sup> meridian west (Aylor, 2003). Then, we computed the weekly cumulative in-strength values of the nodes located along the same meridian in the US from the weekly connectivity networks  $C_w$ . The in-strength is a network metric considering the sum of the weights of the edges pointing a node  $i$ . In this case, the in-strength of a node  $i$  corresponds to the sum of the average weekly frequencies of connection from the other nodes toward  $i$ . We used the cumulative version of this metric, by assuming that local emergence of a disease is observable after sufficiently abundant inoculum has been deposited in that node. We graphically compared observed onset dates and average weekly cumulative in-strength in figure 2.3 (see Section SI A.1.6). Here, we built the frequency distribution of the cumulative in-strength values intersected by 1922–1992 onset date observations and computed the values corresponding to the interquartile range. Lastly, cumulative in-strength values falling in the interquartile range are shaded on the heat map.

### 2.3.5 Network surveillance

#### Exploring “*Puccinia* pathways” worldwide

In order to reconstruct the trajectories of propagation of stem rust outbreaks at world scale, we analyzed how the center of mass of the monthly connectivity networks ( $C_m$ ) in-strength

moves across different subcontinents. Hence, for a given set of nodes representing a subcontinent (and the corresponding geographical coordinates), we defined the center of mass of the monthly connectivity networks in-strength as the geographical point whose coordinates are given by the average latitude and longitude weighted by the nodes monthly in-strengths.

### **Definition of efficient surveillance strategies**

We considered the problem of establishing a reduced set of sentinels, i.e., nodes where the presence of the pathogen is monitored systematically, that should guarantee the largest coverage of the domain and provide an early-warning system for the appearance of new pathogen strains (Parnell et al., 2017). First of all, we defined the coverage of a node  $i$  as the set of nodes that points towards  $i$ , under the assumption that monitoring the presence of the pathogen in  $i$  implies observing all those nodes that are pointing to it. In this case, node  $i$  is referred as the *sentinel* of its coverage. We considered hence the network  $\tilde{C}$  generated by considering those edges of the yearly connectivity networks  $C_y$  recurring at least three times over a 4-year interval 2013-2016 (i.e.,  $\geq 75\%$  of the times), in order to account only for the most frequent connections. The problem of finding the smaller set of sentinels that guarantees the complete coverage is formally equivalent to the set cover problem in graph theory, that happens to have NP-complete computational complexity (Garey and Johnson, 1979; Suttrave et al., 2012). Nonetheless, we used an iterative greedy algorithm, providing a sub-optimal minimum sentinel set. To validate the efficiency of the sentinel set, we assessed its performance in terms of ratio of surveilled domain using the network obtained by the intersection of the yearly connectivity networks of 2017 and 2018. We compared such performances with the ones given by sets of nodes chosen via other network metrics, namely in-strength, betweenness (Freeman, 1978), PageRank (Page, 1998) and random walk generalized accessibility (De Arruda et al., 2014), calculated on the aggregated 2013-2016 networks, and a set of 20 random samplings of nodes (see definitions and procedures in Section SI A.1.7).

### **Measuring early-detection performance of the sentinel set**

Finally, we assessed the performance of different sentinel sets in terms of early detection by means of a compartmental Susceptible-Infected (SI) model based on the intersections of the weekly networks  $C_w$  in the years 2017 and 2018. Within this framework, a node can be either susceptible (S) or infected (I), its state depending on the state of the neighbours in the previous time steps. Namely, in this simplified approach, we assume that, in a given time step  $t$ , a node pass from the state S to I if it has at least one neighbour in state I at time  $t - 1$ , with no recovery. We run 7,814 model simulations, each time letting the outbreak start from a different node (inoculum). Then, we defined the *Disease detection ratio (DDR)* of a given sentinel set as the fraction of the total number of simulations for which the sentinel

set intercepted the epidemics at least once before the end of the 4<sup>th</sup> iteration (i.e., before one month after the first node has been inoculated). This new metric allows to compare the performance of different sentinel sets even when they fail to achieve detection within the first month, something that may happen when the first inoculated node is located in a rather isolated part of the network (see Section SI A.1.7).

## 2.4 Results

### 2.4.1 Worldwide susceptibility and connectivity patterns

The duration of host susceptibility periods, i.e. the concurrence of the host availability and environmental conditions favourable to infection, varies considerably across the world (figure 2.2a; figure A.6). The majority of cells is susceptible for a period between a week and a month. The nodes where susceptibility occurs for more than a month per year are located in northeastern America, southern Brazil and Paraguay, central Europe and central China. The only nodes with susceptibility lasting for more than three months per year are located in Ethiopia (see Section SI A.2.1).

Regions located in the Northern Hemisphere are generally well connected, in such a way that Europe, Asia and North Africa create a unique connected component (figure 2.2b). In spite of the obstacle represented by the Atlantic Ocean, extremely long distance connections may occur from North America to the Mediterranean basin. Conversely, clusters in the Southern Hemisphere are isolated between them, although internally connected. A first representation of the role played by each node within the epidemic network is given by the sum of the probabilities associated to trajectories going into (in-strength: figure 2.2c) and out from (out-strength: figure 2.2d) that node. In biological terms, in-strength (out-strength) is a proxy of the extent to which a region acts as a sink (source) of spore. In-strength appears to be higher in specific regions located in Europe and in the northeastern US, southern Brazil, the Himalaya and central China. Some regions exhibit great variability within relatively short distances, namely Ethiopia, Middle East and Central Asia. Other regions show a smoother and regular gradient, like the US. In this sense, a continuum of intermediate in-strength values extends from Europe to Western Siberia. Out-strength separates more sharply those regions characterised by high and low values and, in particular, Europe is characterized by large regions associated with high out-strength values.

### 2.4.2 Reconstructing *Puccinia* pathways in North America and elsewhere

Our global epidemic network permitted to recover the well known North-American *Puccinia* pathway. In figure 2.3 we compare the average onset date of outbreak along the 97<sup>th</sup> meridian west observed between 1922 and 1992 in North America (Aylor, 2003), with the cumulative in-strength of network nodes along the same meridian. Most of the observed onset



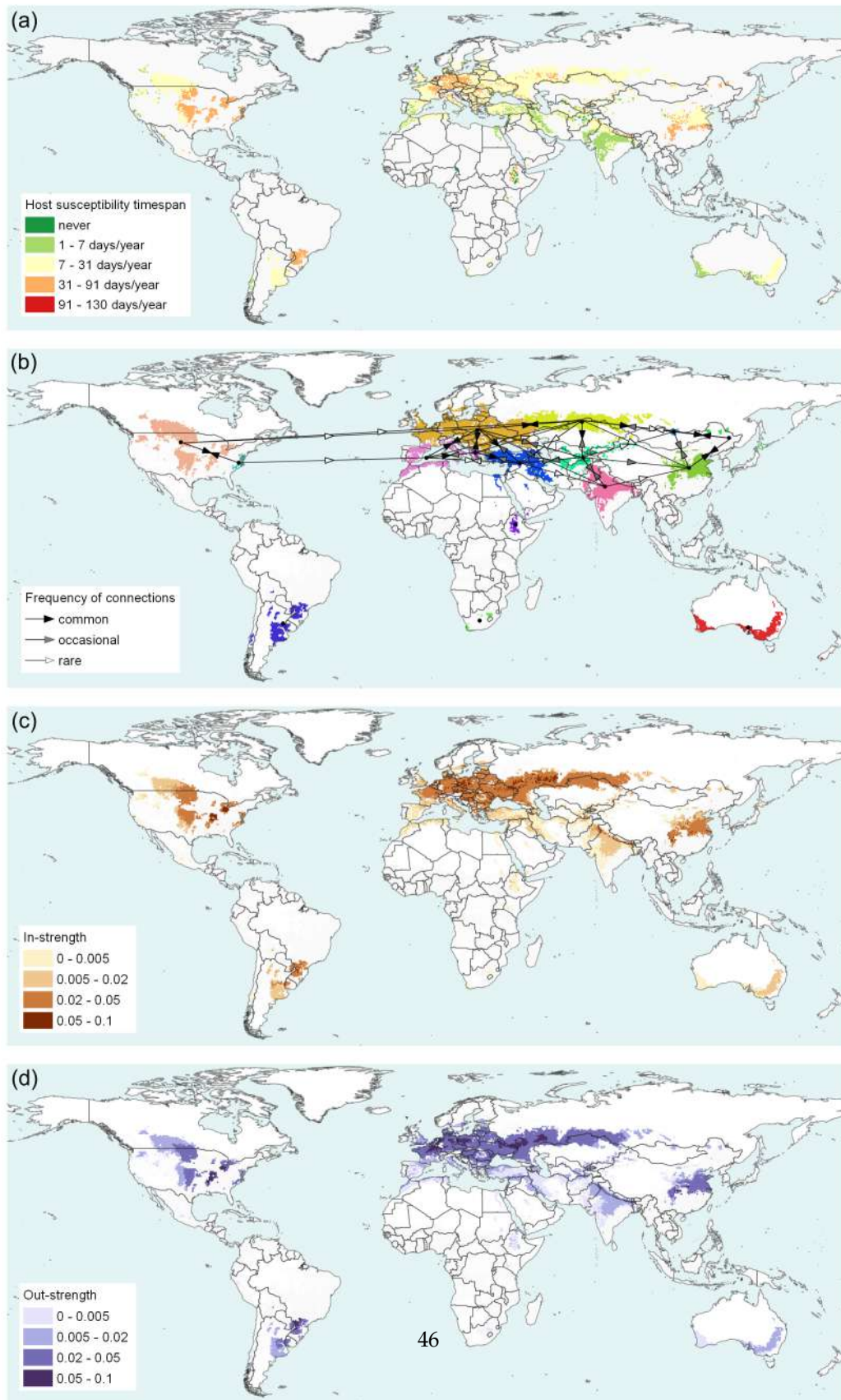


Figure 2.2 – (a) Duration of the host susceptibility period, (b) continental connections among clusters, (c) in-strength and (d) out-strength across the world.

dates occur within a certain interval of values of cumulative in-strength (the interquartile range [0.6,0.9]), suggesting that the cumulative in-strength network metric can proxy the spatio-temporal progression of the “*Puccinia* pathway” in North America. In other words, in this continent one would expect to observe the first sings of an epidemics when the cumulative in-strength has values between 0.6-0.9.

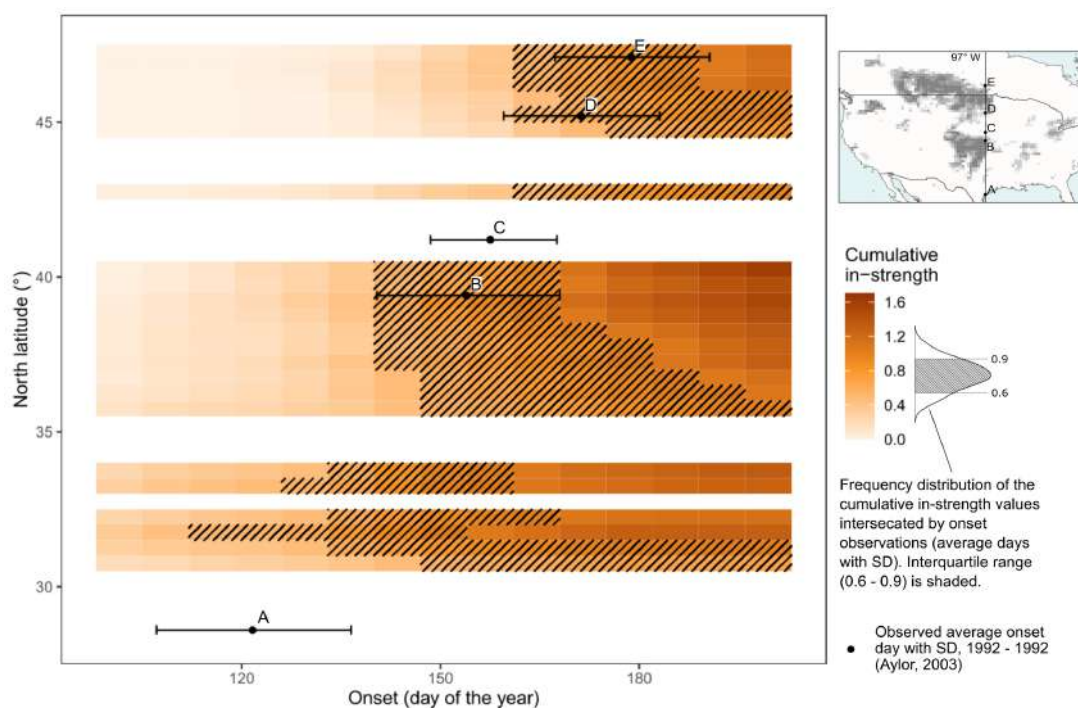


Figure 2.3 – Horizontal error bars, denoted with capital letters, represent the observed average onset dates ( $\pm 1$  standard deviation) of stem rust epidemics along the 97<sup>th</sup> meridian west in the 1922-1992 period (source (Aylor, 2003)). The heat map in the background represents the weekly cumulative in-strength of nodes along the same meridian. In the right panel by the side of the legend, we computed the interquartile range of the cumulative in-strength values intersected by the observed average onset dates and their error bars. The shaded area on the heat map highlights the intervals where the cumulative in-strength value falls within the interquartile range [0.6,0.9]. The top-right panel represents the cells crossed by the 97<sup>th</sup> meridian and the actual wheat distribution in North America.

Globally, we found that the airborne transport estimated via the center of mass of the monthly in-strength always moves poleward from tropical and temperate regions (figure 2.4 and figure A.10), except for the Ethiopian pathway, that follows a southward movement even if it is located in the Northern Hemisphere, likely due to its cropping calendar.

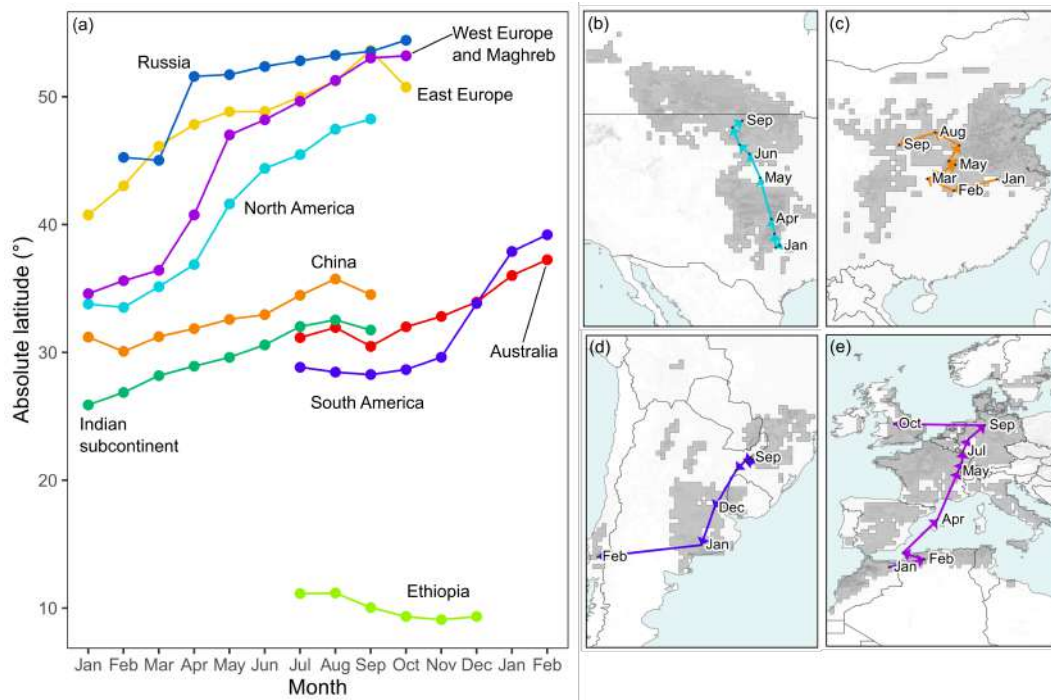


Figure 2.4 – Panel (a) shows the poleward speed of the center of mass of the in-strength of different sub-continents. Panel (b)-(e) depict the monthly pathways for North America, China, South America and Western Europe-Maghreb, respectively.

### 2.4.3 Optimal sentinel set

The sentinel set obtained on the 2013-2016 network using the cover set algorithm allows for a complete coverage of the worldwide wheat producing regions by monitoring 1,007 nodes, i.e., less than 13% of the total (figure A.12). Moreover, a reasonable coverage of 50% of the domain can be obtained monitoring 64 nodes (figure 2.5a), i.e. less than 1% of the total. The first selected nodes are those assuring the greatest coverage, and they do not distribute uniformly across continents (figure A.12 - see Section SI A.2.4).

The sentinel set selected to optimally cover the 2013-2016 epidemic network, via the cover set algorithm, provides satisfactory results also when applied to surveil the epidemic network obtained for the 2017-2018 period (figure 2.5b). In fact, it provides a coverage of 50% of the domain by monitoring 114 nodes (1.5% of the total). For comparison, the same coverage of 50% would require 234 (3%) random sentinels (on average: interquartile range [227; 242]), 475 (6.1%) sentinels if ordered for increasing values of betweenness, or 611 (7.8%) of PageRank.

### 2.4.4 Early detection capabilities of the sentinel set

In terms of *DDR*, the set cover algorithm outperformed all the other methods for sentinel sets containing between 20 and 650 nodes (figure 2.5c). In-strength provides better results for very small sentinel sets, while random sentinel sets are more suitable when larger sentinel sets have to be designed. The *DDR* associated to 275 sentinels is 19.2%, which means that 275 sentinels are able to detect an epidemic process started from any of 19.2% of the world producing regions within a month. A similar *DDR* of 19.1% is obtained with 350 nodes chosen according to their betweenness, or more than 500 nodes according to their PageRank. Between 350 and 500 random sentinels are needed to achieve a *DDR* around 18 – 22%.

We estimated the *DDR* associated to different values of detection delay (2 weeks, 3 month, 6 month, 1 year), showing that the Set Cover strategy improves its performances against the random sentinels sets (figure A.14).

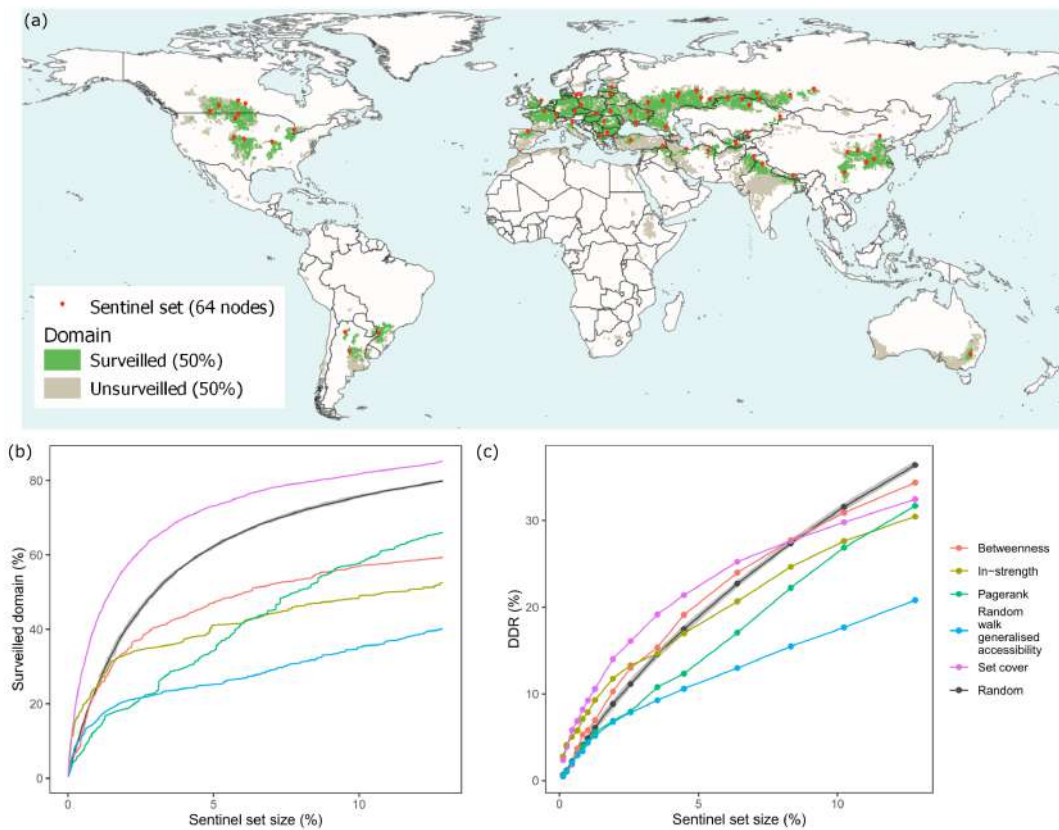


Figure 2.5 – Panel (a) shows the locations of the 64 nodes belonging to the sentinel set associated to a coverage of 50% of the domain. Panel (b) shows the relative extent of the worldwide coverage corresponding to increasing sentinel set sizes for different strategies (network metrics). Panel (c) shows the *DDR* associated to a detection delay of one month assessed with a spatially explicit Susceptible-Infected model corresponding to increasing sentinel set sizes for different strategies. In both panels, the grey area corresponds to the interquartile range of the random strategy.

## 2.5 Discussion

In this study we proposed an original modelling framework based on air mass movements, network analysis, meteorology, ecology and plant physiology to describe the global epidemic network of stem rust of wheat. Eventually, we used the model to identify previously unknown pathways of disease spread at the globe scale and to define a set of sentinels that should be primarily surveyed to achieve early detection of future outbreaks.

### 2.5.1 The most susceptible regions

Our results indicate that the duration of the susceptibility period greatly varies among regions (figure 2.2a; figure A.6). Eastern US, southern Brazil, central Europe and China, which are important areas of wheat production, experience the longest susceptibility periods. On the other hand, other important wheat-producing areas from the Middle East up to the Indian subcontinent have shorter periods of susceptibility. The reason could be the temporal mismatch between the occurrence of environmental conditions for infection and of host availability. This is particularly true for India and in Brazil, where, despite a long-lasting occurrence of environmental conditions for infection, the susceptibility period is constrained by relatively short host availability periods (figure A.6). Ethiopia, which is characterized by exceptional climatic diversity for tropical latitudes (Peel et al., 2007), exhibits a very heterogeneous behaviour, with nodes that are never susceptible in close proximity to others that are highly susceptible (91-130 days per year). Since both host availability and favourable environmental conditions for spore deposition and host infection are climate dependent processes, it is likely that global crop epidemic dynamics will be affected by predicted climate changes (Masson-Delmotte et al., 2021). Prank et al. (2019) predicted a worldwide increase in sporulation period under a RCP8.5 climate change scenario in 2100 - potentially compensated by a general decrease of probability of germination. On the other hand, laboratory experiments conducted on other fungal species suggested that future climate conditions would rather inhibit sporulation, but increase mycelium growth rate (Damialis et al., 2015).

### 2.5.2 On the construction of a global epidemic network

To construct an epidemic network we had to explicitly account for biophysical conditions influencing spore aerobiology. This filtering enabled to discard uninformative air-mass connections and to reduce the complexity of the network, making it simultaneously more significant and manageable.

In principle, by tracking the movement of an air mass for an indefinite long time, one would obtain a completely connected network, i.e. a network with edge density equal to 100%. On the other hand, our epidemic network has an edge density equal to 3.7%, whereas

the network of only the most recurring connections (i.e., those connections found at least 3 times out of 4 in the yearly networks 2013-2016) has a density of 0.4%. For comparison, previous studies using a similar technique but filtering only over a 48-hours time (Choufany et al., 2021b) obtained a density of 28% over the Mediterranean basin for the period 2011-2017. In essence, our filtering allows to make epidemic networks more manageable and informative, presenting itself as alternative to other network-based methods for the extraction of the truly relevant connections (Serrano et al., 2009).

### 2.5.3 From network science to global crop protection

Our results on the frequency of connections (figure 2.2) indicate that regions in the Northern Hemisphere are more densely connected than in the Southern one. This can be explained by the fact that 90% of wheat fields are located in the Northern Hemisphere while the clusters below the Tropic of Cancer, i.e., South America, South Africa, Ethiopia and Australia are thousands of kilometers away from each other. Furthermore, our prediction of possible long-distance connections from North America to Europe, and from Europe towards Central Asia and Russia, are in accordance with previous theoretical and empirical findings (Prank et al., 2019; Wang et al., 2021; Mayol et al., 2017; Visser et al., 2019). On the other hand, due to the time lag in the respective susceptibility seasons, no connection between the Northern and Southern Hemispheres is predicted by our epidemic network. For such a connection to exist, it is necessary the presence of a “green bridge” (Meyer et al., 2017b) in a tropical or sub-tropical region located between the two hemispheres. This could be the case of Ethiopia, where, according to the FAO country brief (FAO, 2021), there exist two wheat growing seasons: *Belg*, from February to July and *Meher*, from May to December. Furthermore, according to our estimates on the host susceptibility duration (figure 2.2a), Ethiopia features some of the nodes with the longest susceptibility period (more than 90 days/year), whereas previous studies already identified this country (and more broadly East Africa and Yemen) as an important stepping stone (Mayol et al., 2017) for the long distance transport of rust spores along the Rift Valley (Meyer et al., 2017b).

The in- and out-strength maps of the epidemic network (figure 2.2c) provide an overview of the epidemic role played by each region, sink rather than source of spores or both. Our results indicate that out-strength tends to be stronger towards the equator compared to in-strength, in line with previous findings (Aylor, 2003) about the “green” and “golden” wave: regions closer to tropics are the first where infection is possible and therefore they are more likely to act as sources of spores rather than sinks. Among the recent seasonal epidemic events moving poleward, the 2013 regional outbreak of stem rust in Germany seems to have moved to Denmark and Sweden later in the season (Saunders et al., 2019).

If one wanted to use our approach to derive effective measures for global crop protection it is worthy to note that in our exercise we privileged generality over local accuracy. Our assumptions to determine host availability and susceptibility are necessarily oversimplifica-

tions of a complex reality, as required by models. For instance, we neglected the presence of alternative hosts, such as spring wheat, whose cropping calendars would increase the host susceptibility period and the frequency of long distance connections. Also, we neglected the existence of the secondary hosts of *P. graminis*, the barberry (*Berberis vulgaris*), which is necessary for sexual reproduction, facilitating the emergence of new strains. It has been successfully eradicated in Western Europe and North America in the last century (Saunders et al., 2019). Yet, it is present in other regions which turn to be sources of new strains (Olivera et al., 2015, 2019) and it has been recently reintroduced in Europe (Barnes et al., 2020).

Our model permitted to design an optimal surveillance strategy capable of covering 50% of wheat cultivated lands while monitoring only 64 sentinel nodes, i.e., less than 1% of the wheat cultivated cells. Although this proportion may seem optimistic for real applications, it is worth recalling the work of Suttrave et al. (2012) on *P. pachyrhizi* indicating that a reliable epidemiological status of soybean rust in the US could be achieved by reducing the sentinel set size from 500 to 12 network optimized nodes.

Of course, the success of the deployment of a sentinel system will also depend on the diagnostic ability of individual sentinels and on the communication between them. In this sense, efforts have recently been devoted to establish international protocols to improve the probability of detection and the timely communication of new outbreaks to a global network of scientists, public authorities and stakeholders (see the Global Cereal Rust Monitoring System (Park et al., 2011; Morris et al., 2022)). Our findings stress the importance of increasing efforts towards a trans-boundary perspective to efficiently contain the emergence of new virulent air-borne crop pathogens (Ristaino et al., 2021).

## Supplementary Information

- A.1 Supplementary materials
- A.2 Supplementary results
- A.3 The published article

## Bibliography

- Allen-Sader, C., Thurston, W., Meyer, M., Nure, E., Bacha, N., Alemayehu, Y., Stutt, R. O., Safka, D., Craig, A. P., Derso, E., Burgin, L. E., Millington, S. C., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2019). An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11).
- Aylor, D. (2017). *Aerial Dispersal of Pollen and Spores*. The American Phytopathological Society.



- Aylor, D. E. (1986). A framework for examining inter-regional aerial transport of fungal spores. *Agricultural and Forest Meteorology*, 38(4):263–288.
- Aylor, D. E. (1999). Biophysical scaling and the passive dispersal of fungus spores: relationship to integrated pest management strategies. *Agricultural and Forest Meteorology*, 97(4):275–292.
- Aylor, D. E. (2003). Spread of plant disease on a continental scale: Role of aerial dispersal of pathogens. *Ecology*, 84(8):1989–1997.
- Baiocco, S., Cavina, F., and Pradolesi, G. (2021). A weather-based simulation model for the development of wheat stem rust epidemics. In *2021 29th Conference of Open Innovations Association (FRUCT)*, pages 22–29. IEEE.
- Barnes, G., Saunders, D. G., and Williamson, T. (2020). Banishing barberry: The history of berberis vulgaris prevalence and wheat stem rust incidence across britain. *Plant Pathology*, 69(7):1193–1202.
- Bhattacharya, S. (2017). Deadly new wheat disease threatens europe’s crops. *Nature*, 542(7640).
- Burrage, S. W. (1970). Environmental factors influencing the infection of wheat by Puccinia graminis. *Annals of Applied Biology*, 66(3):429–440.
- Choufany, M., Martinetti, D., Senoussi, R., Morris, C. E., and Soubeyrand, S. (2021a). Spatiotemporal large-scale networks shaped by air mass movements. *Frontiers in Applied Mathematics and Statistics*, 6:67.
- Choufany, M., Martinetti, D., Soubeyrand, S., and Morris, C. E. (2021b). Inferring long-distance connectivity shaped by air-mass movement for improved experimental design in aerobiology. *Scientific Reports*, 11(1):1–10.
- Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Physical review E*, 70(6):066111.
- Corredor-Moreno, P. and Saunders, D. G. (2020). Expecting the unexpected: factors influencing the emergence of fungal and oomycete plant pathogens. *New Phytologist*, 225(1):118–125.
- Cunniffe, N. J., Stutt, R. O., DeSimone, R. E., Gottwald, T. R., and Gilligan, C. A. (2015). Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. *PLoS computational biology*, 11(4):e1004211.
- Damialis, A., Kaimakamis, E., Konoglou, M., Akritidis, I., Traidl-Hoffmann, C., and Gioulekas, D. (2017). Estimating the abundance of airborne pollen and fungal spores at variable elevations using an aircraft: how high can they fly? *Scientific reports*, 7(1):1–11.

- Damialis, A., Mohammad, A. B., Halley, J. M., and Gange, A. C. (2015). Fungi in a changing world: growth rates will be elevated, but spore production may decrease in future climates. *International Journal of Biometeorology*, 59(9):1157–1167.
- De Arruda, G. F., Barbieri, A. L., Rodríguez, P. M., Rodrigues, F. A., Moreno, Y., and Costa, L. D. F. (2014). Role of centrality for the identification of influential spreaders in complex networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 90(3):1–17.
- Draxler, R. R. and Hess, G. D. (1998). An overview of the HYSPLIT 4 modelling system for trajectories, dispersion and deposition. *Australian Meteorological Magazine*, 47(4):295–308.
- Emerson, E. W., Hodshire, A. L., DeBolt, H. M., Bilsback, K. R., Pierce, J. R., McMeeking, G. R., and Farmer, D. K. (2020). Revisiting particle dry deposition and its role in radiative effect estimates. *Proceedings of the National Academy of Sciences*, 117(42):26076–26082.
- FAO (2021). Fao - country brief. <http://www.fao.org/giews/countrybrief/>. Accessed: 2021-07-20.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.
- Garey, M. R. and Johnson, D. S. (1979). *Computers and intractability*, volume 174. freeman San Francisco.
- Hovmøller, M. S., Yahyaoui, A. H., Milus, E. A., and Justesen, A. F. (2008). Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular ecology*, 17(17):3818–3826.
- Huerta-Espino, J., Singh, R., and Roelfs, A. P. (2014). Rusts fungi of wheat. *Fungi from different substrates*, pages 217–259.
- International Food Policy Research Institute (2019). Global spatially-disaggregated crop production statistics data for 2010, version 2.0.
- Isard, S. A., Gage, S. H., Comtois, P., and Russo, J. M. (2005). Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience*, 55(10):851–861.
- Levetin, E. (2015). Aerobiology of Agricultural Pathogens. *Manual of Environmental Microbiology*, pages 3.2.8–1–3.2.8–20.
- Li, X., Yang, X., Mo, J., and Guo, T. (2009). Estimation of soybean rust uredospore terminal velocity, dry deposition, and the wet deposition associated with rainfall. *European journal of plant pathology*, 123(4):377–386.
- Maddison, A. and Manners, J. (1972). Sunlight and viability of cereal rust uredospores. *Transactions of the British Mycological Society*, 59(3):429–443.

- Mahaffee, W. F. and Stoll, R. (2016). The ebb and flow of airborne pathogens: Monitoring and use in disease management decisions. *Phytopathology*, 106(5):420–431.
- Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. (2021). *IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, volume 1. Cambridge University Press.
- Mayol, E., Arrieta, J. M., Jiménez, M. A., Martínez-Asensio, A., Garcias-Bonet, N., Dachs, J., González-Gaya, B., Royer, S.-J., Benítez-Barrios, V. M., Fraile-Nuez, E., et al. (2017). Long-range transport of airborne microbes over the global tropical and subtropical ocean. *Nature Communications*, 8(1):1–9.
- Mcmaster, G. S. and Smika, D. E. (1988). Estimation and Evaluation of Winter Wheat Phenology in the central great plains. *Agricultural and Forest Meteorology*, 43:1–18.
- Mcmaster, G. S. and Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, 87(1).
- Meyer, M., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017a). Large-scale atmospheric dispersal simulations identify likely airborne incursion routes of wheat stem rust into Ethiopia. *Phytopathology*, 107(10):1175–1186.
- Meyer, M., Cox, J. A., Hitchings, M. D., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017b). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nature Plants*, 3(10):780–786.
- Morris, C. E., Geniaux, G., Nédellec, C., Sauvion, N., and Soubeyrand, S. (2022). One health concepts and challenges for surveillance, forecasting and mitigation of plant disease beyond the traditional scope of crop production. *Plant pathology*.
- Morris, C. E., Sands, D. C., Glaux, C., Samsatly, J., Asaad, S., Moukahel, A. R., Gonçalves, F. L., and Bigg, E. K. (2013). Urediospores of rust fungi are ice nucleation active at >-10 °C and harbor ice nucleation active bacteria. *Atmospheric Chemistry and Physics*, 13(8):4223–4233.
- Nagarajan, S. and Singh, D. (1990). Long-distance dispersion of rust pathogens. *Annual review of phytopathology*, 28(1):139–153.
- Olivera, P., Newcomb, M., Szabo, L. J., Rouse, M., Johnson, J., Gale, S., Luster, D. G., Hodson, D., Cox, J. A., Burgin, L., et al. (2015). Phenotypic and genotypic characterization of race TKTF of *Puccinia graminis* f. sp. *tritici* that caused a wheat stem rust epidemic in southern Ethiopia in 2013-14. *Phytopathology*, 105(7):917–928.

- Olivera, P. D., Sikharulidze, Z., Dumbadze, R., Szabo, L. J., Newcomb, M., Natsarishvili, K., Rouse, M. N., Luster, D. G., and Jin, Y. (2019). Presence of a sexual population of *puccinia graminis* f. sp. *tritici* in georgia provides a hotspot for genotypic and phenotypic diversity. *Phytopathology*, 109(12):2152–2160.
- Oneto, D. L., Golan, J., Mazzino, A., Pringle, A., and Seminara, A. (2020). Timing of fungal spore release dictates survival during atmospheric transport. *Proceedings of the National Academy of Sciences of the United States of America*, 117(10):5134–5143.
- Page, L. (1998). The pagerank citation ranking: Bringing order to the web. technical report. *Stanford Digital Library Technologies Project*, 1998.
- Park, R., Fetch, T., Hodson, D., Jin, Y., Nazari, K., Prashar, M., and Pretorius, Z. (2011). International surveillance of wheat rust pathogens: Progress and challenges. *Euphytica*, 179(1):109–117.
- Parnell, S., van den Bosch, F., Gottwald, T., and Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: an epidemiological perspective. *Annual review of phytopathology*, 55:591–610.
- Peel, M. C., Finlayson, B. L., and McMahon, T. A. (2007). Updated world map of the köppen-geiger climate classification. *Hydrology and earth system sciences*, 11(5):1633–1644.
- Prank, M., Kenaley, S. C., Bergstrom, G. C., Acevedo, M., and Mahowald, N. M. (2019). Climate change impacts the spread potential of wheat stem rust, a significant crop disease. *Environmental Research Letters*, 14(12).
- Radici, A., Martinetti, D., and Bevacqua, D. (2022). Early-detection surveillance for stem rust of wheat: insights from a global epidemic network based on airborne connectivity and host phenology. *Environmental Research Letters*.
- Ristaino, J. B., Anderson, P. K., Bebbler, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., et al. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences*, 118(23).
- Roelfs, A. P. (1992). *Rust diseases of wheat: concepts and methods of disease management*. Cimmyt.
- Rowell, J., Romig, R., et al. (1966). Detection of urediospores of wheat rusts in spring rains. *Phytopathology*, 56(7):807–811.
- Saunders, D. G., Pretorius, Z. A., and Hovmøller, M. S. (2019). Tackling the re-emergence of wheat stem rust in western europe. *Communications biology*, 2(1):51.
- Schmale III, D. G. and Ross, S. D. (2015). Highways in the sky: Scales of atmospheric transport of plant pathogens. *Annual review of phytopathology*, 53:591–611.

- Schneider, R., Hollier, C., Whitam, H., Palm, M., McKemy, J., Hernandez, J., Levy, L., and DeVries-Paterson, R. (2005). First report of soybean rust caused by *phakopsora pachyrhizi* in the continental united states. *Plant disease*, 89(7):774–774.
- Serrano, M. Á., Boguná, M., and Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. *Proceedings of the national academy of sciences*, 106(16):6483–6488.
- Slinn, W. (1977). Some approximations for the wet and dry removal of particles and gases from the atmosphere. *Water, Air, and Soil Pollution*, 7(4):513–543.
- Sutrave, S., Scoglio, C., Isard, S. A., Hutchinson, J. S., and Garrett, K. A. (2012). Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS One*, 7(6):e37793.
- Visser, B., Meyer, M., Park, R. F., Gilligan, C. A., Burgin, L. E., Hort, M. C., Hodson, D. P., and Pretorius, Z. A. (2019). Microsatellite analysis and urediniospore dispersal simulations support the movement of *puccinia graminis* f. Sp. *Tritici* from southern Africa to Australia. *Phytopathology*, 109(1):133–144.
- Wang, M., Kriticos, D. J., Ota, N., Brooks, A., and Paini, D. (2021). A general trait-based modelling framework for revealing patterns of airborne fungal dispersal threats to agriculture and native flora. *New Phytologist*.
- Zadoks, T. (1967). Epidemiology of wheat rust in europe. *International Journal of Pest Management B*, 13(1):29–46.

**Key points:**

- *Puccinia graminis* is a fungal pathogen causing stem rust of wheat, threatening wheat production after decades of apparent calm.
- Using land use data, environmental data affecting host susceptibility and Lagrangian simulations of air-mass trajectories, we reconstruct the worldwide *Puccinia* epidemic network, in which wheat producing regions are connected by airborne spore transport.
- A system of suitability filters applied over Lagrangian simulations allows to identify those trajectories which most likely represented effective transport events, thus reducing the computation complexity and the redundancy of the network. These filters are inspired by experimental knowledge of spore biology, including factors like resistance to UV radiations and rain scavenging.
- We elaborate a new network indicator, called “Set cover”, based on an already existing algorithm. A few nodes, with the highest “Set cover” compose a “sentinel set” able to early detect the emergence of a new disease.

**Perspectives:**

- Validation of the *Puccinia* network may be based on genetic similarity between strains and a measure of proximity of those nodes representing locations where these strains have been found.
- We assume the risk of emergence of new strains to be homogeneously distributed over the nodes. In fact, it depends on the genetic reshuffling occurring on its alternate host *Berberis vulgaris*; incorporating this heterogeneity would enhance the meaningfulness of the network.
- The surveillance assumes perfect cooperation among countries. What is the performance loss due to each country optimising its individual surveillance efforts?
- An upgrade of this model should embed how heterogeneous environmental conditions affect the temporal dynamic of infection and sporulation which ultimately drive the epidemic development.
- Increased knowledge about spore survival and dispersal capability during the airborne phase (e.g. resistance to high or low temperature) would improve the reliability of the model.

## Chapter 3

# A metapopulation framework integrating landscape heterogeneity to model an airborne plant pathogen: the case of brown rot of peach in France

Marco Polo describes a bridge, stone by stone.  
'But which is the stone that supports the bridge?' Kublai Khan asks.  
'The bridge is not supported by one stone or another'  
Marco answers, 'but by the line of the arch that they form'.  
Kublai Khan remains silent, reflecting. Then he adds:  
'Why do you speak to me of the stones? It is only the arch that matters to me.'  
Polo answers: 'Without stones there is no arch.'

– Italo Calvino, *The invisibles cities*

---

This chapter has been submitted as a research article (pre-print available at: Radici et al.2023).

## Résumé

La dynamique des maladies des plantes est déterminée par l'interaction simultanée de la susceptibilité de l'hôte, de la présence de l'agent pathogène et des conditions environnementales. Alors que la susceptibilité de l'hôte et les conditions environnementales locales peuvent être facilement caractérisées, la transmission d'un agent pathogène aérien dépend des conditions biotiques et abiotiques du milieu environnant. Nous proposons ici un cadre original de métapopulation intégrant l'hétérogénéité du paysage, en termes de climat et de densité d'hôtes, où les populations locales d'hôtes végétaux sont connectées par des masses d'air qui permettent la dispersion des agents pathogènes. Nous prenons explicitement en compte les facteurs climatiques affectant la libération et la survie des agents pathogènes tout en modélisant la dispersion aérienne à l'aide de simulations lagrangiennes, ainsi que la phénologie et l'infection des hôtes. Nous calibrons les paramètres du modèle en fonction de la littérature et en utilisant le calcul bayésien par rapport aux observations de l'incidence de la pourriture brune dans les vergers de pêcheurs en France de 2001 à 2020 sur une zone de 50 000  $km^2$ . Nous avons utilisé le modèle pour produire des cartes de risque de la zone d'étude, en distinguant la dangerosité des sites (risque de causer une infection secondaire dans d'autres sites) et leur vulnérabilité (risque d'être infecté). Nous avons constaté que les sites les plus dangereux et les plus vulnérables sont situés le long de la vallée du Rhône, en raison de la concomitance d'une forte densité de vergers de pêches, d'un climat approprié et de connexions aériennes persistantes. Notre travail représente une première étape dans l'intégration de la théorie des métapopulations, de l'épidémiologie et des mouvements de masse d'air pour informer les stratégies de protection des plantes, et pourrait être adapté pour optimiser la protection dans le cadre des projections climatiques futures.

**Mots-clés:** agents pathogènes aériens, pourriture brune de la pêche, protection des plantes, *Monilinia*, métapopulation

## Abstract

Plant disease dynamics are driven by the concurrent interplay of host susceptibility, pathogen presence, and environmental conditions. While host susceptibility and local environmental conditions can readily be characterised, the transmission of an airborne pathogen depends on the biotic and abiotic conditions of the surrounding environment. Here, we propose an original metapopulation framework integrating landscape heterogeneity, in terms of climate and host density, where local populations of plant hosts are connected via air-masses which allow pathogen dispersal. We explicitly account for climatic drivers affecting pathogen release and survival while modelling aerial dispersal using Lagrangian simulations, as well as host phenology and infection. We calibrate the model parameters according to the literature and using Approximate Bayesian Computation against observations of brown rot incidence



in French peach orchards from 2001-2020 across an area of 50,000  $km^2$ . We used the model to produce maps of risk, distinguishing site dangerousness (risk of causing secondary infection in other sites) and vulnerability (risk of becoming infected) across the our study area. We find that most dangerous and vulnerable sites are located along the Rhône Valley, due to the concurrence of high density of peach cultivation, a suitable climate and persistent airborne connections. Our work represents a first step to integrate metapopulation theory, epidemiology and air-mass movements to inform plant protection strategies, and could be adapted to optimize crop protection under future climate projections.

**Keywords:** airborne pathogens, brown rot of peaches, crop protection, *Monilinia*, metapopulation

### 3.1 Introduction

Plant pathogens are a critical issue endangering global food security (Ristaino et al., 2021). Our limited comprehension of long distance dispersed pathogens, *i.e.* transported by wind or other vectors over regional to continental scales (Aylor, 2003; Brown and Hovmøller, 2002), has direct consequences on the implementation of plant protection strategies (Cunniffe et al., 2015; Hyatt-Twynam et al., 2017; Parnell et al., 2017). Given the difficulty to eradicate such pathogens, management strategies should focus on preventing emergence via surveillance (Mastin et al., 2020). The appearance of *Monilinia fructicola* in Europe represents an example of an airborne pathogen whose introduction has evaded conventional crop defense measures (EPPO, 2023).

*Monilinia spp.* are fungal species threatening stone fruit production (Bryde and Willets, 1977; Hrustić et al., 2012). *M. fructicola* is an alien species to Europe, initially observed in France in 2001 (Lichou et al., 2002). Despite being classified as quarantine pathogen, in the following years this new strain was progressively observed in central and southern Europe (Oliveira Lino et al., 2016). Such uncontained invasion may be explained considering the efficacy of its dispersal mechanism. The capability of spores to resist UV radiation (Vilanova et al., 2021) and the compatibility of its aerodynamic diameter (Yamamoto et al., 2014) with airborne transport (Wang et al., 2021) combine to suggest that *Monilinia* spores may spread via air-masses (Bryde and Willets, 1977).

Although epidemiological models already exist to study the local dynamics of brown rot (Bevacqua et al., 2018, 2023), its airborne spread remains unexplored. A possible framework to describe brown rot spread at landscape scale consists in coupling in-site epidemiological dynamics and between-sites pathogen dispersal to create a network of spatially distributed hosts connected via a moving pathogen - a metapopulation. This approach has been widely explored in epidemiology (Keeling and Gilligan, 2000; Thrall and Burdon, 2003) with the general assumption of an isotropic spread (Cunniffe et al., 2016; Rimbaud et al., 2018; Mastin et al., 2020; Fabre et al., 2021). However, in the case of airborne plant diseases, scientific re-

search has recently been extended to include realistic patterns of dispersal thanks to the development of Lagrangian models (Draxler and Hess, 1998; Jones et al., 2007). Studies based on such models have explored the airborne dispersal of plant diseases (Sutrave et al., 2012; Meyer et al., 2017), with management implications (Allen-Sader et al., 2019). Nevertheless, these applications have focused largely on transport between known source and sink locations, with no consideration of the coupling of repeated cycles of dispersal, infection and onward spread that characterises epidemics, which could be embedded in a metapopulation framework. Such a framework would help to conceptually disentangle the local factors (climate) from the connectivity (transport of spores) to understand their relative importance to the emergence of the disease.

In this study, we present a model which integrates current knowledge about climate-dependent fruit phenology (Vanalli et al., 2021) and epidemiology (Bevacqua et al., 2018, 2023) where pathogen dispersal among units is described by Lagrangian simulations of air-mass movements to reproduce the disease dynamics over multiple growing seasons at regional scale. We use brown rot of peach in continental France as a case study. We calibrate the model against observations of disease incidence in the last two decades and we verify the importance of including directional airborne transport by comparing the performances against a null model, in which connectivity is modelled with an isotropic dispersal kernel. We use the metapopulation model to produce maps of epidemiological risk, to find where the most endangered locations for infection are located, and where, if a *Monilinia*-like pathogen were introduced, regional disease size would be maximised. These maps represent one possible application of how this model, based on simulated air-mass movements, can feed into management decisions for plant protection at the national scale.

## 3.2 Materials and methods

### 3.2.1 Model overview

The geographic domain corresponds to the Safran grid (Bertuzzi and Clastre, 2022), made of square cells  $0.11^\circ \times 0.11^\circ$  ( $\sim 8 \times 8 \text{ km}^2$ , hereinafter referred to as “unit”) overlaid on metropolitan France. We considered those 755 units covered by an important peach orchard area ( $> 0.01 \text{ ha/km}^2$ ; hereafter “cultivated area”) and not isolated (Fig. 3.1a; see Subsection B.1.1).

For every unit, for every year in 1996-2020 we computed the ripening period, from pit hardening  $t_0$  to harvest  $t_f$  (Fig. 3.1b), via a phenological temperature dependant model (see Vanalli et al., 2021 for details), which corresponds to the period of susceptibility. Note that  $t_f$  differs among different peach cultivars (i.e. early, mid-early, mid-late, late).

For each unit  $i$ , we run a type SEI climate driven epidemiological model (see Bevacqua et al., 2023 for details) from  $t_{0,i}$  to  $t_{f,i}$  where  $I(t_0) = 0$  and  $S(t_0) + E(t_0) = 15 \text{ fruits/m}^2$

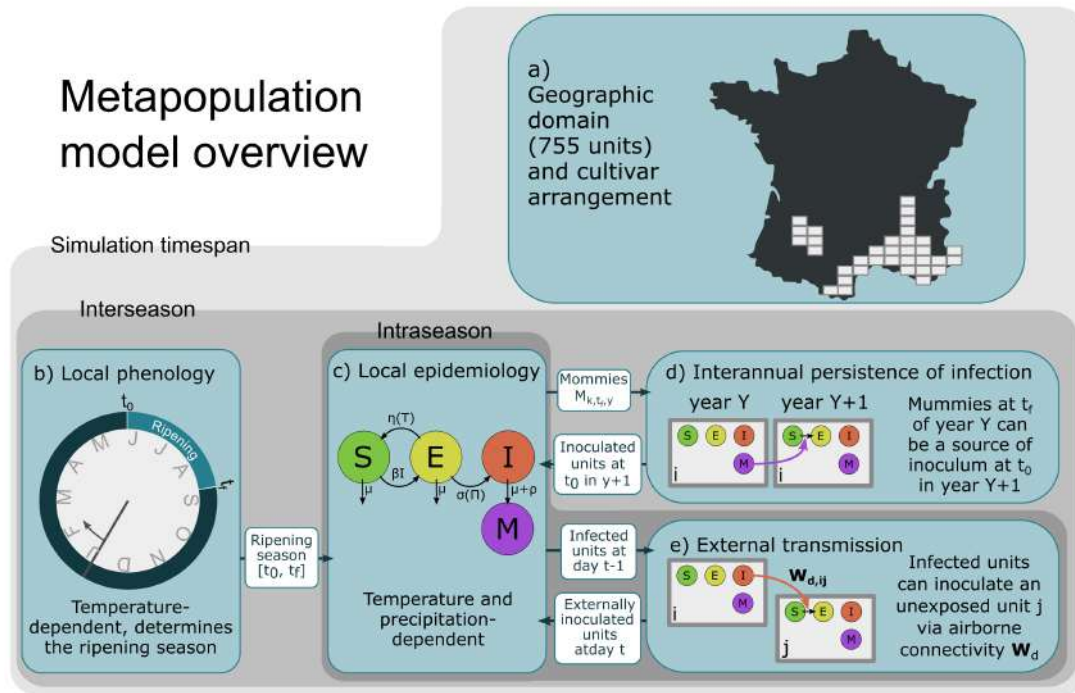


Figure 3.1 – Model overview. a) Geographic domain, composed of 755  $\sim 8 \times 8 \text{ km}^2$  units in France, assigned with a single cultivar for the whole duration of a simulation; in each unit, b) a climate-dependent phenological model is used to estimate the occurrence of the beginning of ripening  $t_0$  (“pit hardening”) and “end of ripening”  $t_f$  (harvest); c) a climate-dependent compartmental model describes the disease dynamics; d) at  $t_0$  of a new season, previous year unharvested mummies determines local inoculum; e) unexposed units can be inoculated at a daily frequency by airborne spores released by infected units.

(Fig. 3.1c). Every year, for each unit, we stochastically assess the value of  $E(t_0)$  as a function of disease incidence in the precedent year. If  $E_i(t_0) > 0$ , the unit is considered as “exposed” and the epidemic dynamics is independent from the epidemic state of the other units. Contrariwise, if  $E_i(t_0) = 0$ , the epidemic cannot spread unless the inoculum comes from connected infected units. In this case, we calculate the probability that it assumes a positive value  $E_i(t) > 0$  with daily frequency (Fig. 3.1d). Such probability depends on the incidence of units that act as a source of airborne inoculum (Fig. 3.1c).

Every year, for each unit, we use the simulated epidemic trajectory to compute the local latent infection, represented by fruits that, following infection, mummify ( $M$ ). We use such a value to assess the probability that, in that unit, a fraction of fruit will be exposed to brown rot at the beginning of the following ripening season (Fig. 3.1d).

### 3.2.2 Model Equations

**Climate dependent phenology and in-unit epidemiological dynamics** We derived initial  $t_0$  and final  $t_f$  fruit ripening time according to Vanalli et al. (2021) (Fig. 3.1b).

We adapted the SEI model for brown rot of peaches developed by Bevacqua et al. (2023) (Eq. 3.1) to deterministically describe the in-unit disease dynamics, with the additional class  $M$  (Fig. 3.1c).

$$\begin{cases} dS(t)/dt = -\beta S(t)I(t) + \eta(T(t))E(t) - \mu S(t) \\ dE(t)/dt = \beta S(t)I(t) - (\eta(T(t)) + \sigma(\Pi(t)))E(t) - \mu E(t) \\ dI(t)/dt = \sigma(\Pi(t))E(t) - (\mu + \alpha)I(t) \end{cases} \quad (3.1)$$

Where  $\beta$  is the transmission rate,  $\eta(T(t))$  is the temperature-dependent spore mortality rate,  $\sigma(\Pi(t))$  is the rain-dependent infection rate,  $\mu$  and  $\alpha$  are the natural and infection-related abscission rates. The value of  $M$  is computed at  $t_f$ :

$$M(t_f) = I(t_f) + \int_{t_0}^{t_f} (\mu + \alpha)I dt \quad (3.2)$$

where we include remaining infected fruits at harvest  $I(t_f)$  among the mummies. We used the climate reanalysis provided by Siclima (Delannoy et al., 2022; Caubel et al., 2015).

**Cultivar distribution across the domain** We inferred the geographical distribution of the cultivars by computing the yield  $Y_{i,z}$  by unit  $i$  in 1991-2005 in absence of disease, for every cultivar  $z$ :

$$Y_{i,z} = \sum_{y=1991}^{2005} e^{-\mu t_{f,y,i,z}} w(t_{f,y,i,z}) \quad (3.3)$$

where  $t_{f,y,i,z}$  is the harvest date of variety  $z$  in unit  $i$  in year  $y$ , and  $w(t_{f,y,i,z})$  is the relative fruit weight (Eq. B.2). We assumed that the probability  $P_{i,z}$  of presence of cultivar  $z$  in  $i$  is proportional to  $Y_{i,z}$  (Fig. B.1):

$$P_{i,z} = \frac{Y_{i,z}}{\sum_z Y_{i,z}} \quad (3.4)$$

Finally, we created 100 randomized geographical distributions of cultivars, where a unit is associated only to one cultivar, extracted with probability  $P_{i,z}$ , for the whole duration of a simulation.

**Interannual persistence of inoculum** At  $t_0$  some units may have some inoculum which originated by overwintering mummies or other sources (Oliveira Lino et al., 2016; Fig. 3.1e). We assumed the presence of inoculum as a fraction of the fruit load being already “ex-

posed":  $E(t_0) = 0.27 \text{ fruits}/m^2$ . We determined the probability of having a primary inoculum  $P_o$  in year  $y$  in unit  $i$  as the complementary of the product of the independent probabilities of not having any inoculum:

$$P_o = 1 - (1 - \tilde{P}_o)[1 - (1 - e^{-\theta_O M_{i,y-1}(t_f)})] \quad (3.5)$$

Where  $\tilde{P}_o$ , fixed to 0.2 to avoid non-identification issues (see Subection B.1.7), represents the probability inoculum due to sources other than peach infections, while  $1 - e^{-\theta_O M_{i,y-1}(t_f)}$  expresses the probability of overwintering. The density of mummies at  $t_f$  of the precedent year  $M_{i,y-1}(t_f)$  is weighted by parameter  $\theta_O$ , to be estimated.

**External transmission of the inoculum** If, instead, the unit is unexposed at  $t_0$ , an epidemic may be triggered by stochastic introduction of inoculum (Fig. 3.1e) from infected units. We modelled such inoculation as a Bernoulli variable, with probability  $P_e$ , depending on the rate of external infection  $R_{i,t+1}$ :

$$P_e = 1 - e^{-R_{i,t+1}\Delta t} \quad (3.6)$$

We extracted new daily inoculated units via Eq. 3.6 ( $\Delta t = 1d$ ). Rate  $R_{i,t+1}$  is defined as:

$$R_{i,t+1} = \begin{cases} \beta\theta_E S_i(t) \sum_{j \neq i} w_{d,ji} I_j(t) A c_j / A & , \text{ if } E_i(t) = 0 \wedge \Pi_i(t) > 0 \\ 0 & , \text{ otherwise} \end{cases} \quad (3.7)$$

where  $\beta$  is the transmission term weighted by  $\theta_E$ , to be estimated,  $S_i(t)$  and  $I_j(t)$  are susceptible and infected fruit loads in units  $i$  and  $j$ , term  $A c_i / A$  is needed to compare units of different cultivated areas,  $w_{d,ji}$  is the airborne transport probability,  $\Pi_i(t)$  is the precipitation in the arrival unit  $i$ , allowing wet deposition.

An inoculation is imagined as a fraction  $0.27 \text{ fruits}/m^2$  of susceptible load that become exposed, as in the case of overwintering. The term  $E_i(t) = 0$  prevents re-introduction of inoculum.

**Airborne epidemic network** To estimate airborne transport probability  $w_{d,ji}$  we set up air-mass simulations with the HYSPLIT model (Draxler and Hess, 1998). We ran forward Lagrangian trajectory simulations from each unit centroid, 1 m above terrain height, 4 times a day, uniformly extracted from daylight hours (since sunlight facilitates atmospheric turbulence and spore escape; Levetin, 2015) in 2008-2019. We set the travel duration to 6 hours, a compromise between the viability of thin-walled spores (of the order of magnitude of one hour; Oneto et al., 2020) and of thick-walled spores (few days; Visser et al., 2019).

Along a trajectory  $v$ , we computed the probability of spore survival to temperature and sunlight-induced mortality. We intersected  $v$  with the domain grid, obtaining a weighted and directed connectivity matrix  $W_t$  in which an element  $w_{t,ij}$  represents the probability

$P_t$  that  $v$  started at time  $t$  in unit  $i$  crosses unit  $j$  along its trajectory. We eventually post-processed matrices  $W_t$  into  $W_d$  by averaging over each day of the year  $d$  and including an additional connectivity for neighbouring units (see Subsection B.1.6).

**Observations of the disease** We collected observations of the disease incidence (as “weak” or “strong”) in different locations from *i*) scientific articles, *ii*) plant health bulletins, *iii*) master thesis reports and *iv*) expert judgement of experimental fields (Radici, 2023).

To compare model outputs with the categorical observations, we mimicked the process by which experts, after examining losses, state whether the incidence has been “weak” or “strong”. Losses  $L$  are expressed as:

$$L = 1 - \frac{S_{t_f}}{S_{t_f}^*} \quad (3.8)$$

Where  $S_{t_f}$  and  $S_{t_f}^*$  are the actual and disease-free fruit loads at harvest time. We assumed there exists a threshold  $\theta_L$ , to be estimated, which distinguishes high and low losses:

$$\text{incidence} = \begin{cases} \text{strong} & , \text{ if } L > \theta_L \\ \text{weak} & , \text{ otherwise} \end{cases} \quad (3.9)$$

### 3.2.3 Training and stratified cross validation of the model’s parameters

To estimate the parameter set  $\theta = (\theta_E, \theta_O, \theta_L)$  we followed an Approximate Bayesian Computation procedure (ABC; Csilléry et al., 2010; Minter and Retkute, 2019). This allows to estimate the posterior distribution of the parameters by running the model several times and selecting those sets whose performances satisfy a proximity threshold  $\epsilon$  to the observations (see Subsection B.1.7). We chose Cohen’s  $\kappa$  index (Fielding and Bell, 1997) as a measure of proximity (ranging from -1, lowest, to 1); we fixed a value (0.475) and set  $\epsilon$  as  $1 - \kappa = 0.525$ .

To further restrict the parameter sets, we used a Stratified K-Fold Cross Validation algorithm (Arlot and Celisse, 2010). We assessed  $\kappa$  both on training ( $\kappa_\phi$ ) and on testing ( $\kappa_\psi$ ) sets, and took the first 100 more frequent accepted sets as our final ensemble  $\theta^*$  (see Subsection B.1.8).

### 3.2.4 Testing the null model

We tested the hypothesis of directional airborne epidemic spread against a null model where we replaced wind-driven matrices  $\mathbf{W}_d$  with an isotropic kernel (matrix  $\mathbf{U}$ ). Each element  $u_{ij}$  depends exclusively on the distance between  $i$  and  $j$ . We run the two models and we compared their performances through a Monte Carlo analysis (Gotelli et al., 2004; see Subsection B.1.9).

### 3.2.5 Estimating risk of brown rot of peaches

We used the model to find the most dangerous units, intended as the ones which, if infected, would maximize the disease size and the most vulnerable units, intended as the ones which become easily infected because of an outbreak elsewhere in the domain. First, we set a unit to be completely infected; then, we associated a random year in 1991-2010, a randomized geographical rearrangement of cultivars, a random parameter set from the posterior distribution, and we ran a 10-years simulation, computing losses in the last simulated year according to Eq. 3.8.

In this experiment we set the parameter  $\tilde{P}_0$  (sources other than peach infections) to 0 until that unit has been infected (Eq. 3.5). From 75,500 simulated epidemics (755 units  $\times$  100 stochastic repetitions for each unit) we computed two indices: *i*) the “vulnerability”, i.e. the average local losses by secondary infection, started anywhere in the region; *ii*) the “dangerousness”, i.e. the overall average losses (weighted by cultivated area) caused by an infection in that unit (see Subsection B.1.10).

## 3.3 Results

### 3.3.1 Wind connectivity matrices

Air-mass connectivity, summarized by the annual matrix  $\mathbf{W}$  (the weighted average of all the daily matrices  $\mathbf{W}_d$ ; see Subsection B.1.6), varies heterogeneously through the study area and can be represented as a spatial network (Fig. 3.2a, rearranged on regions defined in panel b). Beside self loops, an important spread pathway occurs in the Rhône Valley (regions D-E-F; Fig. 3.2b) followed by a Mediterranean coastal pathway (regions B-C-F-G-H). On the opposite, the northeastern part of the domain (A, Guyenne) has very few connections with the rest of the domain.

### 3.3.2 Parameters estimation

Depending on the cross validation splitting, the size of the accepted parameter sets varies from 7 to 67 (average = 29) over 200,000. Cohen’s ( $\kappa_\phi$ ) assessed on the training set has an average value of 0.5 (min = 0.496; max = 0.502), while on the testing set ( $\kappa_\psi$ ) has an average value of 0.2 (0.046, 0.409).

The losses threshold over which an epidemic is considered as “strong”, represented by parameter  $\theta_L$ , has average value is 30.3% (interquartile range: 27.1% to 32.9%; Fig. B.3b and c). It is more complicated to directly visualise the range of accepted values of  $\theta_E$  (the weighting parameter driving external inoculum; average value =  $1.8 \times 10^{-6}$ ;  $6.1 \times 10^{-7}$  to  $2.4 \times 10^{-6}$ ; Fig. B.3a and b) and  $\theta_O$  (the weighting parameter driving overwintering; average value =  $1.0 \times 10^{-2}$ ;  $7.1 \times 10^{-3}$  to  $1.4 \times 10^{-2}$ , Fig. B.3a and c). We therefore computed the

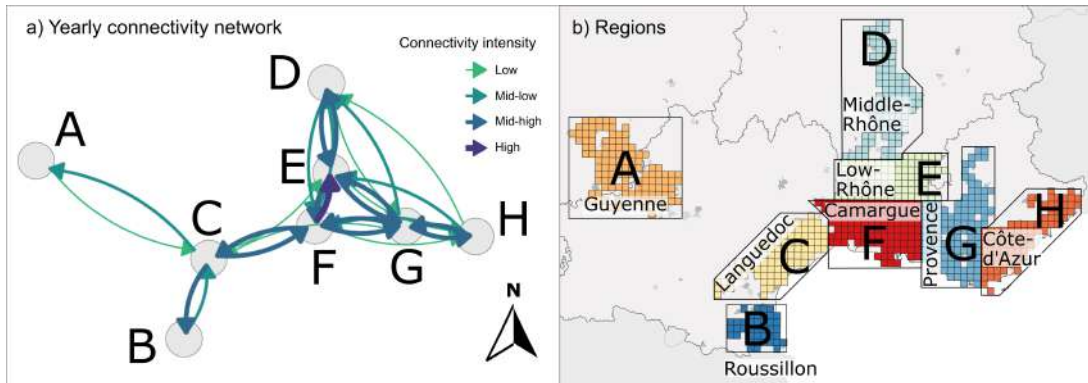


Figure 3.2 – Average connectivity matrix  $W$  a) as a directed network, rearranged grouping units into b) eight regions equal in size (where possible). The colour scale represents connectivity on a logarithmic scale (High: probability of a daily deposition of spore from any unit in a region to any of a second region in the order of  $10^{-4}$ ; Mid-high,  $10^{-5}$ ; Mid-low,  $10^{-5}$ ; Low,  $10^{-7}$ ) consistently with Fig. B.2a.

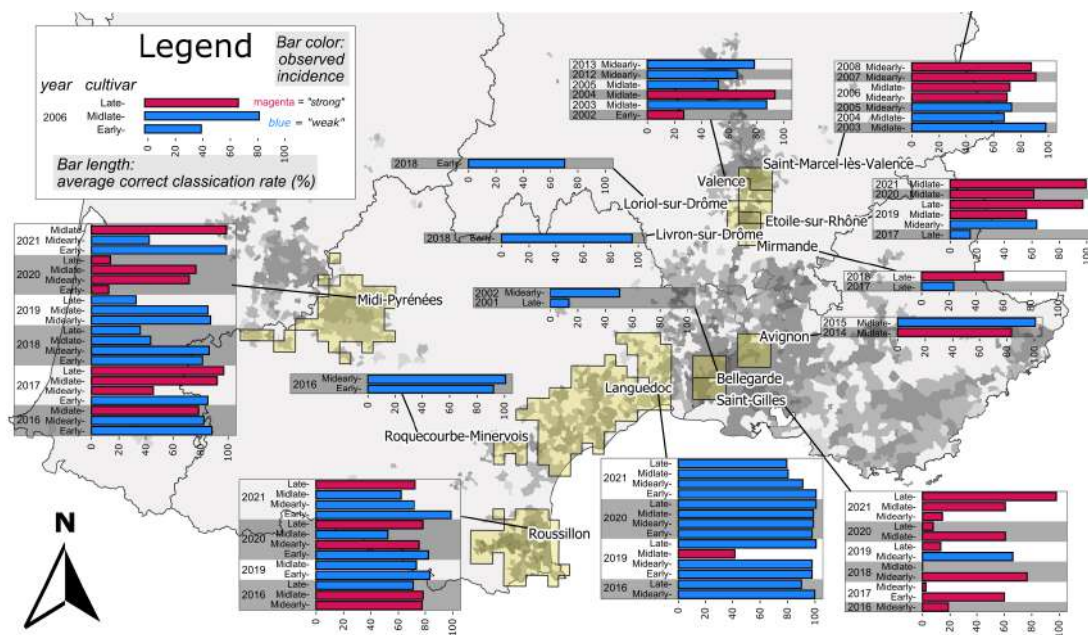


Figure 3.3 – Observations of incidence of brown rot in peach orchards and estimated model accuracy. The observed incidence is represented by the color. Locations are mapped as yellow shades. The bar length represents the correct classification rate of the model, i.e. the % of true positives and true negatives. Gray shades represent cultivated areas per municipality.



“average annual number of external infections”, which is function of  $\theta_E$  (average value =  $1.062 \times 10^2$  out of  $7.55 \times 10^2$ ;  $4.8 \times 10^1$  to  $1.48 \times 10^2$ ; Fig. B.4b) and the “average number of units infected at  $t_0$ ” which is function of  $\theta_O$  (average value =  $1.597 \times 10^2$  out of  $7.55 \times 10^2$ ;  $1.49 \times 10^2$  to  $1.68 \times 10^2$ ; Fig. B.4b).

The average value of  $\kappa$  computed on the whole observations is 0.41 and the correct classification rate (the ratio of true positive and true negative, Fielding and Bell, 1997) is 71.8%. This latter varies throughout each cultivar (Fig. 3.3; Fig. B.5a). It is generally higher for “weak” incidences (79.7%) and lower for “strong” incidences (61.1%). Also, it decreases passing from early (78.4%) to late cultivars (62.7%; Fig. 3.3, B.5c).

The Monte Carlo analysis revealed that the wind-driven model performs significantly better than the null one (p-value  $< 5 \times 10^{-6}$ ).

### 3.3.3 Dangerousness and vulnerability to brown rot

The most dangerous units (average dangerousness, i.e. the average regional losses caused by an infection in that unit) values are found in the Camargue (region F in in Fig. 3.4a, 25.7%, with a maximum of 38.9%), followed by other regions along the Rhône Valley (24%): the Middle-Rhône (D, 23.8%) and Low-Rhône (E, 21.9%). Provence (G, 9.5%), Languedoc (C, 8.7%) and Côte d’Azur (H, 5.9%) display intermediate values, while Roussillon (B, 4.8%) and Guyenne (A, 2.6%) have to the lowest average values.

The most vulnerable units (average vulnerability, i.e. the average local losses by secondary infection, of 19%) are again found in in the Middle-Rhône (Fig. 3.4b), with peaks of 23.9%. From the most to the least vulnerable units, we found in Low-Rhône (average = 16.9%) and Camargue (average = 14.1%) and then eastwards in Provence (average = 11.6%) and Côte d’Azur (H, 11.1%). Guyenne follows with values of vulnerability around 5.2%. Languedoc and Roussillon have again the lowest average values of 3.7% and 2%.

In view of this result we decided to illustrate typical disease dynamics. We therefore recomputed 200 simulations for a specific units. We chose three geographically close units (in Middle-Rhône) but different in term of dangerousness. For each one, we mapped the progress of the advance front of a “rapid spread” simulation, i.e. the earliest year of high losses ( $> 30\%$ ) in the simulation generating the 67<sup>th</sup> (i.e., 2/3) quantile of cumulative losses over the 10 years time span (Fig. 3.4c).

In 11.5% of the simulations where the disease starts in a high-dangerousness the outbreak dies out before the 10<sup>th</sup> year. These percentage increases to 60.5% for a mid-dangerousness and 85.5% in a low-dangerousness unit. A high-dangerousness unit can provoke high losses in 352 units out of 755, with a main north-south spread direction, the double of a mid-dangerousness (180 out of 755), while a low-dangerousness unit would provoke high losses in only one other unit.

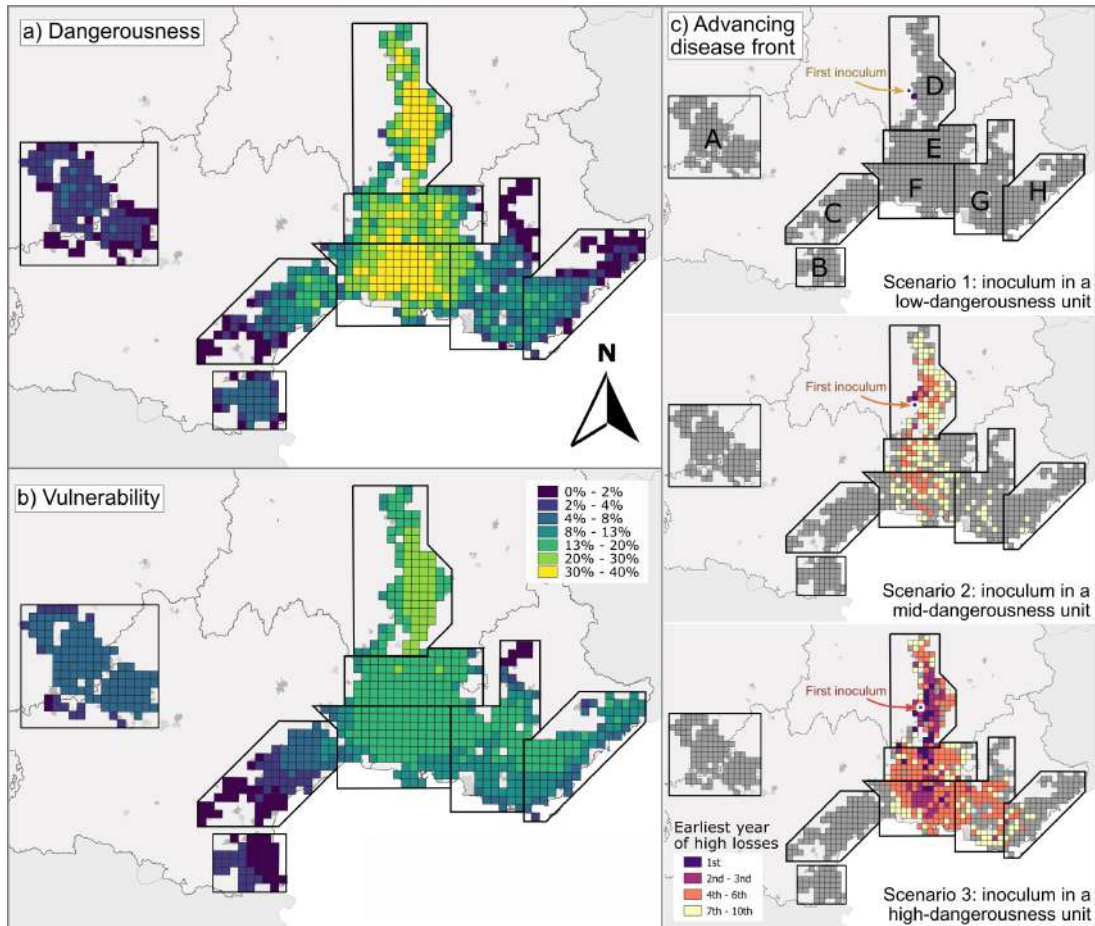


Figure 3.4 – Geographical distribution of epidemiological risk. a) Dangerousness (%), i.e. the average regional losses 10 year after inoculation; b) vulnerability (%), i.e. the average local losses 10 year after inoculation of a random unit; c) earliest year of high losses (> 30%) in a rapid-spread simulation (i.e., the simulation corresponding to the 67<sup>th</sup> quantile of cumulative losses) with a single inoculated unit in Middle Rhône (first inoculated unit has a low, middle and high dangerousness in scenario 1, 2 and 3 respectively).

### 3.4 Discussion

We presented a metapopulation model explicitly including wind-driven dispersal for studying the spatio-temporal dynamics of airborne plant pathogens and we used it to produce maps of disease risk. Our framework allows to match observations of disease incidence while considering the complementary role of local overwintering and the wind-driven transport of inoculum between orchards (Fig. 3.3): the best fit is obtained when extinctions due to missed overwintering are compensated by re-introductions due to air-masses movements (Fig.s B.3 and B.4).

The role of directional wind-driven transport is corroborated by the Monte Carlo analysis with the null model, fed with an isotropic kernel. Such kernel is typically assumed in scientific literature in absence of knowledge on pathogen transports (Cunniffe et al., 2016; Rimbaud et al., 2018; Mastin et al., 2020; Fabre et al., 2021). Lagrangian trajectory simulations of air masses have been already used to investigate aerial connectivity in Chapter 2; our experiment suggests that, for those pathosystems where airborne transportation is ascertained, it be efficiently integrated in a dynamical framework, helping describing the epidemic spread more accurately than a isotropic kernel.

Defining an acceptance threshold is usually a critical point in the ABC (Minter and Retkute, 2019), but a rooted scientific literature helps interpreting the meaning of  $\kappa$ . Negative values characterise models which are worse than random, values over 0.21 are considered as “fair”, while over 0.41 are “good”, “moderate” (McHugh, 2012; Fielding and Bell, 1997). We set  $\kappa > 0.475$  over each cross-validation repetition. The same sets of parameters, if assessed against the testing set, have values of  $\kappa$  close to “fair” (Tab. B.2).

The model predicts “weak” incidences better (average correct classification rate = 80%) than “strong” ones (61%). This is consistent, since a “strong” incidence needs both favourable environmental conditions and inoculum presence, while “weak” incidences simply require one condition to be absent, since the infection will not develop. Geographical distribution of model performances can be discussed under this consideration (Fig.s 3.3, B.5b); “weak” infection levels in Roquecourbe-Minervois, Livron-sur-Rhône, Livron-sur-Rhône and Languedoc are better identified than “strong” levels in Saint Gilles and Valence. Disease dynamics in early cultivars are generally better described than in late cultivars (Fig.s 3.3, B.5c), which have by definition longer ripening duration in which they may be infected.

**The risk of brown rot in France** The model identifies both the most dangerous disease spreaders and most vulnerable pathogen receivers along the Rhône Valley (regions D-E-F).

From a management point of view, vulnerability - which can be compared to the secondary infection risk by Meentemeyer et al. (2011) - calls for direct protection (e.g., via fungicide application) of an epidemic unit. Dangerousness, instead, measures the impact of an outbreak at a broader scale; it underlies a vision in which local management has reper-

cussion all over an interconnected landscape. Independently of its vulnerability, preventing disease in a dangerous unit may be globally advantageous over a larger share of the domain.

Despite being two aspects of the main phenomenon, vulnerability and dangerousness are statistically distributed differently through the domain. The coefficient of variation for the former (0.56) is lower compared to that for the latter (0.84). This means that units are more easily characterised as having high or low dangerousness, rather than for their vulnerability, as in this case they are clustered around to the average value. This suggests that good disease spreaders are more easily distinguishable from bad ones than good or bad pathogen receivers.

Proposed epidemic risk indices depend indirectly from a number of factors: length of ripening, host surface, favourable environmental conditions and air-mass connections. Host density per unit increases proportionally the capacity of infecting connected units, and consequently dangerousness (Eq. 3.7). Rain occurrence can be a measure of suitability of the environmental conditions triggering infections, either due to local inoculum or wind-mediated (due to wet deposition). In our model, rain triggers the rate  $\sigma$  at which exposed fruit load become infected (Fig. 3.1c, Eq. 3.1), proportional also to peach weight. In fact, also plant phytosanitary bulletins warn local producers to pay attention to rain when fruits are close to ripening, as this can facilitate cuticle cracking and subsequent infection (Oliveira Lino et al., 2016).

Climatic factors affecting epidemic risk may be more easily visualized in Guyenne and Roussillon due to their low connectivity to the domain (Fig.s 3.2a, B.2a). In fact, despite scoring low on both indices, their relative isolation from the rest of the domain helps disentangle the importance of local environmental variables from those influenced by connectivity. Guyenne is characterised by little cultivated area (about 5 ha per unit, Fig. B.1e). On the other hand, in this region we observe the largest frequency of rainy days during ripening (24.5 days on average; Fig. B.1e), which is the main factor driving infection in the SEIM model, which may explain why, despite its isolation, its vulnerability is higher compared to other regions such as Languedoc. Roussillon, instead, hosts the largest cultivated areas (33.5 ha per unit) and is slightly more dangerous than Guyenne.

Regions along Rhône Valley, which are then both dangerous and vulnerable, are interested by the typical North-South wind circulation (the Mistral), whose units host large cultivated areas (from 15 to 29.2 ha per unit) and intermediate values of rainy days (12 to 21 days per season). Among mid-risk, but well-connected regions, late varieties are common in Languedoc and Côte d'Azur, and less in Provence (Fig. B.1e). By contrast, they all share little of cultivated areas (5.1 to 7.9 ha per unit) and infrequent rainy events (11.8 to 14 days per season).

Intuitively, the latest varieties should be also the most vulnerable, provides more time for infection. In contrast, these varieties, which typically grow in warmer climates, are relatively common in mid- and low-risk regions (Fig. B.1d). This suggests that frequency of

rainy days (which seems to trigger vulnerability) or host surface and connectivity (which are likely to increase dangerousness) may be more relevant epidemiological factors.

The first European detection of *M. fructicola* occurred in 2001 in the Gard department (Lichou et al., 2002; Fig. B.2b, between regions F and E), which hosts dangerous units (Fig. 3.4a). While epidemics starting from these units persist for several years, their spread is unlikely to cross national borders (Fig. 3.4c). This slow progression may be attributed to the hypothesis underlying the fitting of the model: assuming that the disease is already widespread everywhere may have led to an underestimation of its aggressiveness.

Moreover, it may be imagined that wind is not the only dispersal medium. Brown rot of peaches is also known to cause high post-harvest losses (Oliveira Lino et al., 2016) and it can be hypothesized that the handling of returnable crates used to ship and store harvested fruits contributes to the spread of the disease (Bryde and Willets, 1977). If one wanted to integrate this transportation mode to our modeling framework, this would be represented by an additional connectivity layer reflecting storage and trading (Hernandez Nopsa et al., 2015). Furthermore, this additional spread would not necessarily occur simultaneously with the ripening season but later, thereby extending the period during which susceptible fruits can come into contact with the pathogen. Beside additional dispersal medium, the discrepancy between observed and modeled spread leaves open the debate about possible multiple introductions of *M. fructicola* in Europe.

**Towards metapopulation-based plant protection strategies** Fungal plant disease dynamics have been traditionally interpreted only under the light of local environmental conditions (Juroszek et al., 2020), neglecting transport of spores. The presence itself of *Monilinia* viable spores correlates with local environmental variables (Holb, 2008), but this information has not been used so far to retrace epidemiological dynamics going beyond local descriptors.

Airborne transport has gained increasing scientific evidence to be among the fastest way of pathogen transport (Schmale III and Ross, 2015), provided that given biological protections (melanin, leading to spores resisting UV light) and aerodynamic features are ascertained (Levetin, 2015). The growing availability of computationally efficient Lagrangian models to reconstruct air-mass trajectories, such as HYSPLIT (Draxler and Hess, 1998), allows more research on airborne plant pathogen spread.

Our study represents a first step to use air-mass movements to both study airborne disease spread among fruit trees and inform plant protection strategies at the national range. This research supports a shift in the current paradigm regarding the optimal spatial scale of disease management from the field to the landscape (Thompson et al., 2016). Moreover, it could be adapted to perform optimization of epidemic surveillance and disease control under future climate projections.

## Supplementary Information

- B.1 Supplementary materials
- B.2 Supplementary results

## Bibliography

- Allen-Sader, C., Thurston, W., Meyer, M., Nure, E., Bacha, N., Alemayehu, Y., Stutt, R. O., Safka, D., Craig, A. P., Derso, E., Burgin, L. E., Millington, S. C., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2019). An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11).
- Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4:40 – 79.
- Aylor, D. E. (2003). Spread of plant disease on a continental scale: Role of aerial dispersal of pathogens. *Ecology*, 84(8):1989–1997.
- Bertuzzi, P. and Clastre, P. (2022). Information sur les mailles SAFRAN.
- Bevacqua, D., Quilot-Turion, B., and Bolzoni, L. (2018). A model for temporal dynamics of brown rot spreading in fruit orchards. *Phytopathology*, 108(5):595–601.
- Bevacqua, D., Vanalli, C., Casagrandi, R., and Gatto, M. (2023). A climate-driven compartmental model for fungal diseases in fruit orchards: The impacts of climate change on a brown rot-peach system. *Agricultural and Forest Meteorology*, 332:109293.
- Brown, J. K. and Hovmöller, M. S. (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297(5581):537–541.
- Bryde, R. L. W. and Willets, H. J. (1977). Brown rot fungi of fruit trees. *Nature*, 146(3698):370.
- Caubel, J., García de Cortázar-Atauri, I., Launay, M., de Noblet-Ducoudré, N., Huard, F. F., Bertuzzi, P. P., and Graux, A.-I. (2015). Broadening the scope for ecoclimatic indicators to assess crop climate suitability according to ecophysiological, technical and quality criteria. *Agricultural and Forest Meteorology*, 207:94 – 106.
- Csilléry, K., Blum, M. G., Gaggiotti, O. E., and François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution*, 25(7):410–418.
- Cunniffe, N. J., Cobb, R. C., Meentemeyer, R. K., Rizzo, D. M., and Gilligan, C. A. (2016). Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California. *Proceedings of the National Academy of Sciences*, 113(20):5640–5645.

- Cunniffe, N. J., Stutt, R. O., DeSimone, R. E., Gottwald, T. R., and Gilligan, C. A. (2015). Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. *PLoS computational biology*, 11(4):e1004211.
- Delannoy, D., Maury, O., and Décome, J. (2022). CLIMATIK : système d'information pour les données du réseau agroclimatique INRAE.
- Draxler, R. R. and Hess, G. D. (1998). An overview of the HYSPLIT 4 modelling system for trajectories, dispersion and deposition. *Australian Meteorological Magazine*, 47(4):295–308.
- EPPO (2023). *Monilinia fructicola*. eppto datasheets on pests recommended for regulation. <https://gd.eppo.int>. Accessed: 2023-05-09.
- Fabre, F., Coville, J., and Cunniffe, N. J. (2021). Optimising reactive disease management using spatially explicit models at the landscape scale. In *Plant Diseases and Food Security in the 21st Century*, pages 47–72. Springer.
- Fielding, A. H. and Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1):38–49.
- Gotelli, N. J., Ellison, A. M., et al. (2004). *A primer of ecological statistics*, volume 1. Sinauer Associates Sunderland.
- Hernandez Nopsa, J. F., Daghli, G. J., Hagstrum, D. W., Leslie, J. F., Phillips, T. W., Scoglio, C., Thomas-Sharma, S., Walter, G. H., and Garrett, K. A. (2015). Ecological networks in stored grain: Key postharvest nodes for emerging pests, pathogens, and mycotoxins. *BioScience*, 65(10):985–1002.
- Holb, I. J. (2008). Monitoring conidial density of *Monilinia fructigena* in the air in relation to brown rot development in integrated and organic apple orchards. *European Journal of Plant Pathology*, 120(4):397–408.
- Hrustić, J., Mihajlović, M., Grahovac, M., Delibašić, G., Bulajić, A., Krstić, B., and Tanović, B. (2012). Genus *Monilinia* on pome and stone fruit species. *Pesticidi i fitomedicina*, 27(4):283–297.
- Hyatt-Twynam, S. R., Parnell, S., Stutt, R. O., Gottwald, T. R., Gilligan, C. A., and Cunniffe, N. J. (2017). Risk-based management of invading plant disease. *New Phytologist*, 214(3):1317–1329.
- Jones, A., Thomson, D., Hort, M., and Devenish, B. (2007). *Air pollution modeling and its application XVII*. Borrego, C. & Norman, A.-L.
- Juroszek, P., Racca, P., Link, S., Farhumand, J., and Kleinhenz, B. (2020). Overview on the review articles published during the past 30 years relating to the potential climate change effects on plant pathogens and crop disease risks. *Plant pathology*, 69(2):179–193.

- Keeling, M. J. and Gilligan, C. A. (2000). Metapopulation dynamics of bubonic plague. *Nature*, 407(6806):903–906.
- Levetin, E. (2015). Aerobiology of Agricultural Pathogens. *Manual of Environmental Microbiology*, pages 3.2.8–1–3.2.8–20.
- Lichou, J., Mandrin, J.-F., Breniaux, D., Mercier, V., Giauque, P., Desbrus, D., Blanc, P., and Belluau, E. (2002). Une nouvelle moniliose. *monilia fructicola* s' attaque aux arbres fruitiers à noyaux. *PHYTOMA La Défense des Végétaux*, (547):22–25.
- Mastin, A. J., Gottwald, T. R., van den Bosch, F., Cunniffe, N. J., and Parnell, S. (2020). Optimising risk-based surveillance for early detection of invasive plant pathogens. *PLoS biology*, 18(10):e3000863.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.
- Meentemeyer, R. K., Cunniffe, N. J., Cook, A. R., Filipe, J. A., Hunter, R. D., Rizzo, D. M., and Gilligan, C. A. (2011). Epidemiological modeling of invasion in heterogeneous landscapes: spread of sudden oak death in california (1990–2030). *Ecosphere*, 2(2):1–24.
- Meyer, M., Cox, J. A., Hitchings, M. D., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nature Plants*, 3(10):780–786.
- Minter, A. and Retkute, R. (2019). Approximate Bayesian Computation for infectious disease modelling. *Epidemics*, 29(February):100368.
- Oliveira Lino, L., Pacheco, I., Mercier, V., Faoro, F., Bassi, D., Bornard, I., and Quilot-Turion, B. (2016). Brown Rot Strikes Prunus Fruit: An Ancient Fight Almost Always Lost. *Journal of Agricultural and Food Chemistry*, 64(20):4029–4047.
- Oneto, D. L., Golan, J., Mazzino, A., Pringle, A., and Seminara, A. (2020). Timing of fungal spore release dictates survival during atmospheric transport. *Proceedings of the National Academy of Sciences of the United States of America*, 117(10):5134–5143.
- Parnell, S., van den Bosch, F., Gottwald, T., and Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: an epidemiological perspective. *Annual review of phytopathology*, 55:591–610.
- Radici, A. (2023). Brown rot of peach severity data. <https://www.data.gouv.fr/fr/datasets/brown-rot-of-peach-severity-data/>. Accessed: 2023-02-01.
- Radici, A., Martinetti, D., Vanalli, C., Cunniffe, N., and Bevacqua, D. (2023). A metapopulation framework integrating landscape heterogeneity to model an airborne plant pathogen: the case of brown rot of peach in france. *bioRxiv*, pages 2023–10.



- Rimbaud, L., Papaïx, J., Rey, J.-F., Barrett, L. G., and Thrall, P. H. (2018). Assessing the durability and efficiency of landscape-based strategies to deploy plant resistance to pathogens. *PLoS computational biology*, 14(4):e1006067.
- Ristaino, J. B., Anderson, P. K., Bebbler, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., et al. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences*, 118(23).
- Schmale III, D. G. and Ross, S. D. (2015). Highways in the sky: Scales of atmospheric transport of plant pathogens. *Annual review of phytopathology*, 53:591–611.
- Sutrave, S., Scoglio, C., Isard, S. A., Hutchinson, J. S., and Garrett, K. A. (2012). Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS One*, 7(6):e37793.
- Thompson, R. N., Cobb, R. C., Gilligan, C. A., and Cunniffe, N. J. (2016). Management of invading pathogens should be informed by epidemiology rather than administrative boundaries. *Ecological Modelling*, 324:28–32.
- Thrall, P. H. and Burdon, J. J. (2003). Evolution of virulence in a plant host-pathogen metapopulation. *Science*, 299(5613):1735–1737.
- Vanalli, C., Casagrandi, R., Gatto, M., and Bevacqua, D. (2021). Shifts in the thermal niche of fruit trees under climate change: The case of peach cultivation in France. *Agricultural and Forest Meteorology*, 300(September 2020).
- Vilanova, L., Valero-Jiménez, C. A., and van Kan, J. A. (2021). Deciphering the monilinia fructicola genome to discover effector genes possibly involved in virulence. *Genes*, 12(4).
- Visser, B., Meyer, M., Park, R. F., Gilligan, C. A., Burgin, L. E., Hort, M. C., Hodson, D. P., and Pretorius, Z. A. (2019). Microsatellite analysis and urediniospore dispersal simulations support the movement of puccinia graminis f. Sp. Tritici from southern Africa to Australia. *Phytopathology*, 109(1):133–144.
- Wang, M., Kriticos, D. J., Ota, N., Brooks, A., and Paini, D. (2021). A general trait-based modelling framework for revealing patterns of airborne fungal dispersal threats to agriculture and native flora. *New Phytologist*.
- Yamamoto, N., Nazaroff, W. W., and Peccia, J. (2014). Assessing the aerodynamic diameters of taxon-specific fungal bioaerosols by quantitative PCR and next-generation DNA sequencing. *Journal of Aerosol Science*, 78:1–10.

**Key points:**

- *Monilinia fructicola*, a fungal pathogen causing brown rot of peaches - arguably the main cause of production losses - has appeared in France in 2001 and in a decade has colonized orchards in half of the European continent almost undisturbed.
- Via the presented metapopulation model, which includes weather series affecting in-orchard epidemiological dynamics and an explicit description of the mobility of the pathogen between orchards, we demonstrate that directional wind movement is essential to represent the spatio-temporal dynamics of brown rot of peaches.
- In France, most “vulnerable” locations are located along the Rhône Valley, especially in the north. Vulnerability is associated to frequent rains during ripening.
- Locations where an introduction of a *Monilinia-like* pathogen would cause the greatest losses at the regional level correspond to the most densely cultivated area in the whole Rhône Valley.

**Perspectives:**

- Our level of knowledge on *Monilinia's* survival and dispersal capability during the airborne phase is limited and anecdotal, way poorer compared to what is currently known about *Puccinia*. Experiments are needed to address this gap.
- The calibrated spatial spread rates appear too low to explain alone such a large-scale invasion in a decade. It seems plausible that other means helped in the displacement, such as returnable shipping crates that are used to stock harvested fruits - given the high losses that usually occur in the post-harvest phase.
- Brown rot of peaches is already everywhere and is treated with chemicals at a high economic and environmental costs; the presented model can be used to design the spatial optimisation of fungicide applications.

## Chapter 4

# Global benefits and domestic costs of a cooperative surveillance strategy to control transboundary crop pathogens

The end of this story can only be related in metaphors since it takes place in the kingdom of heaven, where there is no time. Perhaps it would be correct to say that Aurelian spoke with God and that He was so little interested in religious differences that He took him for John of Pannonia. This, however, would imply a confusion in the divine mind. It is more correct to say that in Paradise, Aurelian learned that, for the unfathomable divinity, he and John of Pannonia (the orthodox believer and the heretic, the abhorrer and the abhorred, the accuser and the accused) formed one single person.

– Jorge Luis Borges, *The Theologians*

### Résumé

Les maladies transfrontalières sont extrêmement complexes à contrôler et peuvent causer des dommages socio-économiques à l'échelle mondiale. Dans le contexte de la protection des plantes, les stratégies de surveillance sont généralement conçues en fonction des fron-

---

This chapter has been published as a research article (Radici et al., 2023).

tières nationales, sans tenir compte de l'échelle spatiale de la propagation de la maladie. Dans cette étude, nous nous intéressons à la pertinence de cette échelle pour la surveillance des agents pathogènes dispersés sur de longues distances. Nous utilisons un réseau épidémique décrivant le transport potentiel mondial de *Puccinia graminis*, l'agent causal de la rouille noire du blé, modélisé dans un travail antérieur. Sur la base des propriétés du réseau, nous concevons deux stratégies de priorisation des zones à surveiller pour la présence de la maladie, soit en coopération, soit chaque pays seul, et nous comparons leurs performances en termes de minimisation de l'effort déployé pour atteindre des objectifs de surveillance donnés, au niveau mondial et national. Nous constatons qu'une stratégie de coopération est plus efficace à l'échelle mondiale. Cependant, son adoption implique une distribution géographique hétérogène des coûts et des bénéfices liés à l'effort de surveillance. Les pays de taille moyenne d'Europe centrale et d'Asie en tireraient le plus grand bénéfice ; en revanche, les pays situés sur des voies de propagation importantes devraient déployer davantage d'efforts de surveillance qu'ils ne le feraient en l'absence de coopération. Parmi les principaux producteurs de blé, la Chine est le seul pays pour lequel une stratégie de coopération pourrait avoir un coût, tandis que l'Inde, la Russie, les États-Unis, la France et l'Ukraine en tireraient le plus grand avantage. La reconnaissance de la manière dont les coûts et les bénéfices d'une gouvernance mondiale seraient partagés entre les pays est nécessaire pour obtenir un soutien unanime en faveur d'un système de surveillance coopératif international.

**Mots-clés:** protection des cultures, dispersion à longue distance, réseaux complexes, *Puccinia graminis*, surveillance transfrontalière

## 4.1 Abstract

Transboundary diseases are extremely complex to control and can cause global socioeconomic damage. In the context of crop protection, surveillance strategies are usually designed according to country boundaries, regardless of the spatial scale of the spread of the disease. In this study, we investigate the suitability of this scale for surveilling long-distance dispersed pathogens. We use an epidemic network describing worldwide potential transport of *Puccinia graminis*, the causal agent of stem rust of wheat, modelled in Chapter 2. Based on network properties, we conceive two strategies for prioritizing areas to be monitored for the presence of the disease, either cooperative or each country alone, and we compare their performances in terms of minimizing the effort deployed in achieving given surveillance targets at global and domestic level. We find that a cooperative strategy is more efficient at the global scale. However, its adoption implies a heterogeneous geographic distribution of surveillance effort-related costs and benefits. Medium-sized countries in central Europe and Asia would benefit the most; on the other hand, countries placed in important spreading pathways should deploy more surveillance effort than they would place without

cooperation. Among the major wheat producers, China is the only country that may have a cost from a cooperative strategy, while India, Russia, United States, France and Ukraine would have the most benefits. The acknowledgement of how costs and benefits of a global governance would be shared among countries is needed to gain unanimous support for an international cooperative surveillance system.

**Keywords:** crop protection, long distance dispersal, networks, *Puccinia graminis*, trans-boundary surveillance

## 4.2 Introduction

The issue of surveillance of transboundary diseases, hereinafter intended as infectious diseases whose rapid spatial spread is likely to concern more than a country, has recently come in the spotlight due to the Covid-19 pandemic (Dhama et al., 2020; Chinazzi et al., 2020; Soubeyrand et al., 2020; Mohamed et al., 2020). New outbreaks of such diseases (Brockmann and Helbing, 2013; Saunders et al., 2019), as well as biological invasions of alien species (Diagne et al., 2021), are hardly predictable events. They can be shaped by different dissemination pathways (human transportation, commodity shipping, animal vectors or atmospheric agents) and cause socioeconomic and health issues. Furthermore, lack, mismatch or delay in the communication of first detection among countries, together with uncoordinated control measures, may lead to inefficient management (Carvajal-Yepes et al., 2019; Thompson et al., 2020). Notably, the threat posed by airborne crop pathogens represents a paradigmatic case of transboundary spread (Isard et al., 2005; Xing et al., 2020; Corredor-Moreno and Saunders, 2020). The risk of large losses in food production due to unexpected outbreaks has prompted researchers and institutions to explore international surveillance systems to timely tackle the diffusion of the most alarming crop pathogens (Park et al., 2011; Carvajal-Yepes et al., 2019). The spatio-temporal persistence of large-scale seasonal movements, such as the well known *Puccinia* pathway from Mexico to Canada (Brown and Hovmøller, 2002; Aylor, 2003), has recently emerged as a major source of inspiration for devising such innovative surveillance systems (Sutrave et al., 2012; Meyer et al., 2017; Allen-Sader et al., 2019). In spite of such efforts, standard surveillance of transboundary crop diseases has frequently been performed according to country boundaries, without a cooperative perspective, regardless of the actual scale of spread of the disease, lacking international, and timely, communication of first detections (Park et al., 2011; Carvajal-Yepes et al., 2019; Ristaino et al., 2021). Yet, benefits from a possible general reduction of surveillance effort of a global, cooperative and communicative strategy (Thompson et al., 2016) over a non cooperative one, i.e. each country alone, have never been quantified in the case of long-distance dispersed pathogens.

In this study we investigate to what extent, and under which conditions, country boundaries represent a suitable scale for surveillance of long-distance dispersed crop pathogens,

and whether international cooperation would make crop protection more effective. We use stem rust of wheat, caused by *Puccinia graminis*, an airborne fungal pathogen whose spores can be transported over long distances by wind (Levetin, 2015), as a case study. In the majority of wheat producing countries the presence of this pathogen has been controlled by the use of resistant cultivars and the eradication of its secondary host, *Berberis vulgaris*, which enables overwintering in temperate regions. This pathogen reappeared in western Europe after several decades of absence (Barnes et al., 2020; Corredor-Moreno and Saunders, 2020; Saunders et al., 2019) and is considered a threat to global food security due to the rapid spread of virulent races through a worldwide distributed host (Singh et al., 2015; Ristaino et al., 2021). In Chapter 2, we retraced its global epidemic network across worldwide wheat-producing countries. In the present study, we use this epidemic network to conceive two surveillance strategies, a “non-cooperative” one, representing a within-boundary scenario with no collaboration and communication between countries, and a “cooperative” one, where countries collaborate surveilling each other and timely communicate the detection of the disease. We compare their performances in terms of surveillance effort needed to achieve given targets both at the global and domestic scale.

## 4.3 Materials and methods

### 4.3.1 The worldwide *Puccinia* epidemic network

In order to evaluate the performances of different surveillance strategies, we used the epidemic networks obtained in Chapter 2. Here we present a summary of the methodology proposed there. We simulated worldwide transport of *P. graminis* spores among wheat producing countries, obtaining a time-varying directed and weighted connectivity network  $\mathbf{W}$ . In  $\mathbf{W}$ , the 7,814 nodes represent  $0.5^\circ \times 0.5^\circ$  cells ( $\approx 2,000 \text{ km}^2$ ) in wheat-producing countries, while edges represent likely air-mass connections among cells, computed at a time resolution of 6 hours for the time span 2013-2016. More specifically, each weighted edge  $w_{ijt}$  of  $\mathbf{W}$  is computed in such a way to account for the likelihood of air-mass trajectories (computed via NOAA’s HYSPLIT model; Draxler and Hess, 1998) which potentially disseminate spores from a release node  $i$  to an arrival node  $j$  at time  $t$ . In both  $i$  and  $j$ , host availability and favourable environmental conditions (for sporulation and/or infection) are determined via a climate-dependent suitability model and validated via a comparison with cropping calendar from the FAO country briefs (FAO, 2021a). Seventy-two-hour (72 h) trajectories (Meyer et al., 2017) are filtered according to different criteria (rain washout, cumulative UV radiation, flight duration and altitude) to exclude those air-mass movements that are less likely to lead to an effective spore transport event.

We then projected this time-varying epidemic networks in a static, directed and binary design network  $\mathbf{W}_D$ , generated by considering only recurring connections, *i.e.* occurring  $i$ )

at least once a year and *ii*) at least three times over the 4-year interval 2013-2016 (*i.e.*,  $\geq 75\%$  of the years). Network  $\mathbf{W}_D$  identifies only highly likely direct spore dissemination events on a seasonal timescale.

### 4.3.2 Surveillance strategy design

We further considered the problem of establishing a reduced set of *sentinels*, nodes where the presence of the pathogen is systematically monitored (*i.e.*, the surveillance effort), that should guarantee the largest aggregated coverage of the domain (*i.e.*, the surveillance target) and provide an early-warning system for the detection of the pathogens. First of all, we defined the *coverage* of a sentinel as the set of nodes that points directly towards it, under the assumption that, by monitoring the presence of the pathogen in a sentinel, we can indirectly observe the possible presence in all those nodes that are pointing to it in one step. We leveraged on an iterative heuristic algorithm to determine sub-optimal solutions to the problem of finding the smallest set of sentinels  $\mathbf{s}_\sigma$  that guarantees the maximum aggregated coverage (associated to a surveillance target  $\sigma$ ).

The iterative heuristic algorithm (or “Set cover”) to determine sub-optimal solutions to the problem of finding the smallest set of sentinels consists in: *i*) finding the node associated to the largest coverage, *ii*) add this node to the sentinel set  $\mathbf{s}_\sigma$ , initially empty, *iii*) label its coverage as surveilled and remove all the edges pointing to it, *iv*) repeat steps *i-iii* until the proportion of nodes in the aggregated coverage reaches the desired target  $\sigma$ . The optimal set of sentinels  $\mathbf{s}_\sigma$  is ranked by growing aggregated coverage. The size of  $\mathbf{s}_\sigma$  defines the surveillance effort  $x_\sigma$ .

We designed two surveillance strategies, a “cooperative” and a “non-cooperative” one. In the “cooperative” strategy, the Set cover algorithm was run on all nodes of the network. By contrast, in the “non-cooperative” strategy, we *i*) labelled each node with the country where it is placed and *ii*) ran the Set cover algorithm separately for each country by considering only the corresponding sub-block of the network. We thus obtained the optimal sentinel sets  $\mathbf{s}_{\sigma,c}^{-T}$  for each country  $c$ , where  $^{-T}$  stands for “without Transboundary edges”, ranked by growing aggregated domestic coverage. To compare the performances of the “cooperative” and “non-cooperative” strategies, we computed the number of sentinels needed to achieve different global targets (Fig. 4.1).

### 4.3.3 Measuring benefits and costs of cooperation at domestic scale.

To investigate how the burden of surveillance is shared among countries, for each country  $c$ , we calculated the number of sentinels  $x_{c,\sigma,s}$  needed to achieve a domestic surveillance target of  $\sigma$  under a given strategy  $s$  ( $s = \text{“cooperative” or “non-cooperative”}$ ). Then, we defined the cost-benefit index  $\alpha_{c,\sigma}$  as the ratio between the number of domestic sentinels needed to

achieve  $\sigma$  in the “cooperative” and in the “non-cooperative” strategy, for a given country  $c$ :

$$\alpha_{c,\sigma} = \frac{x_{c,\sigma,s=cooperative}}{x_{c,\sigma,s=non-cooperative}} \quad (4.1)$$

We evaluated it for  $\sigma = 1\%, 2\%, \dots, 100\%$  and then we computed the average ( $\bar{\alpha}_c$ ) by country. We ascribe to a country  $c$  the label of “CoopBeneficial” if  $\bar{\alpha}_c < 1$ , “CoopAdverse” if  $\bar{\alpha}_c > 1$  and “CoopNeutral” if  $\bar{\alpha}_c = 1$ . After having computed  $\bar{\alpha}_c$  by country, we aggregated it by continent weighting each country’s contribution by its wheat production (FAO, 2021b) to investigate geographical heterogeneity of benefits and costs of cooperative surveillance.

#### 4.3.4 Robustness of the sentinel sets

To assess the temporal robustness of the results to slight changes in the epidemic network, we set up a validation procedure of the performances of the sentinel sets. We recomputed the connectivity network  $\mathbf{W}$  on years 2017-2018 and projected it into a validation (directed, binary, static) network  $\mathbf{W}_V$ , obtained by considering only those connections occurring at least once a year both in 2017 and 2018.

We then recomputed the aggregated coverage and  $\alpha_{c,\sigma}$  of the sentinels sets  $\mathbf{s}_\sigma$  and  $\mathbf{s}_{\sigma,c}^{-T}$  using network  $\mathbf{W}_V$ .

## 4.4 Results

Our global epidemic network, together with the applications of the Set cover algorithm, allowed us to identify those sentinels that would best perform to detect disease presence within a certain portion of the network. Note that sentinels might not be included in the network portion that one wants to surveil. For example, if the objective is to monitor the portion of the network corresponding to all wheat-producing regions in Germany, regardless of where the sentinels are placed (the “cooperative” strategy), the optimal sentinel set would comprise only three domestic sentinels (see Fig. 4.1a). On the other hand, it would be necessary to place six sentinels if surveillance could be provided only by domestic sentinels (the “non-cooperative” strategy, see Fig. 4.1b), not contributing to transboundary surveillance. Our results indicates that, for a  $\sigma$  of 100%, Germany would benefit from a cooperative strategy as the number of domestic sentinels needed to monitor its territory would pass from 6 to 3, thus meaning a cost-benefit index of  $= 3/6 = 0.5$ . Indeed, the interpretation of the cost-benefit index is rather straightforward: if  $\alpha_{c,\sigma} < 1$ , country  $c$  requires less sentinels within its borders in the “cooperative” scenario than in the “non-cooperative” one for achieving the same surveillance target  $\sigma$ . If  $\alpha_{c,\sigma} > 1$ , the opposite is true, while if  $\alpha_{c,\sigma} = 1$ , country  $c$  needs the same number of sentinels in both the strategies for achieving surveillance target  $\sigma$ .



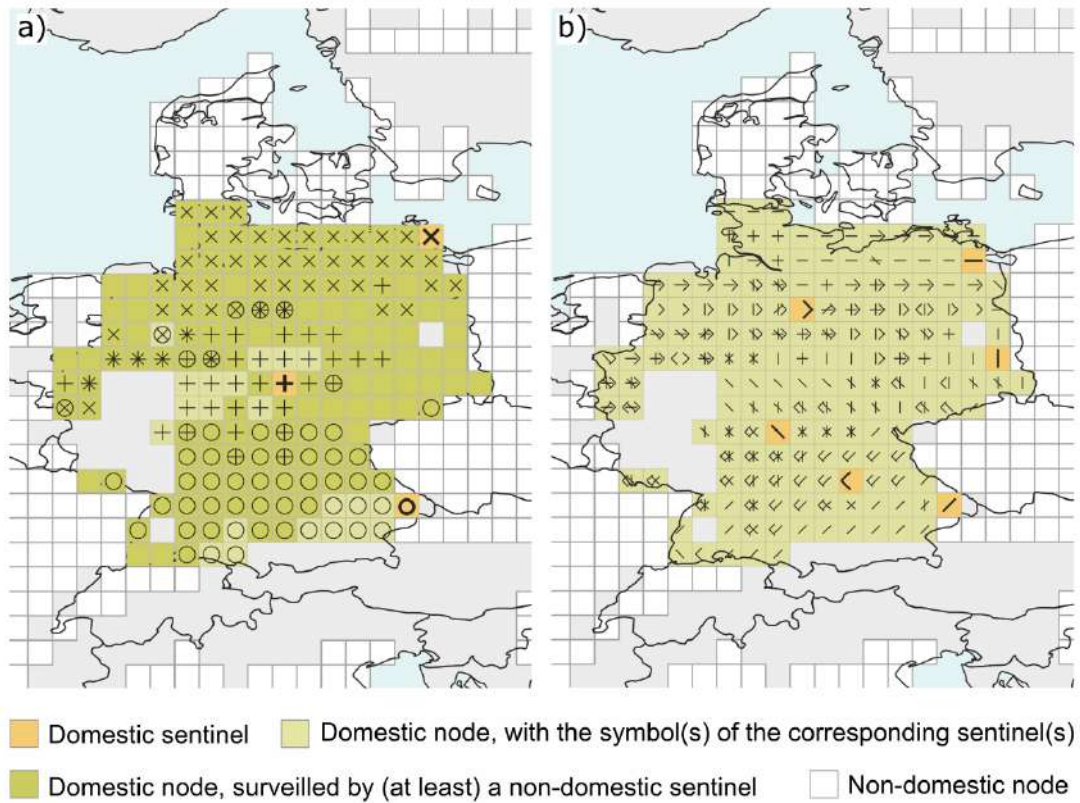


Figure 4.1 – A graphic example to compare “cooperative” and “non-cooperative” strategies when the surveillance target is set to  $\sigma = 100\%$ . Square cells represent nodes of the network, corresponding to wheat producing regions. (a) In the “cooperative” strategy, three domestic sentinels (orange nodes:  $x$ ,  $o$ ,  $+$ , surveilling light green cells) in addition to others placed abroad (which surveil dark green nodes), are needed to cover all nodes in Germany. Each node is associated to one or more symbols, each for the sentinel(s) monitoring it. (Note that the sentinel  $x$  has a domestic cover set which is also surveilled by international sentinels. Yet, in a cooperative framework its role is essential to efficiently surveil nodes out of Germany). (b) In the “non-cooperative” strategy, six domestic sentinels ( $l$ ,  $-$ ,  $l$ ,  $\backslash$ ,  $>$ ,  $<$ ) are needed to surveil German nodes (light green cells). They do not contribute to transboundary surveillance.

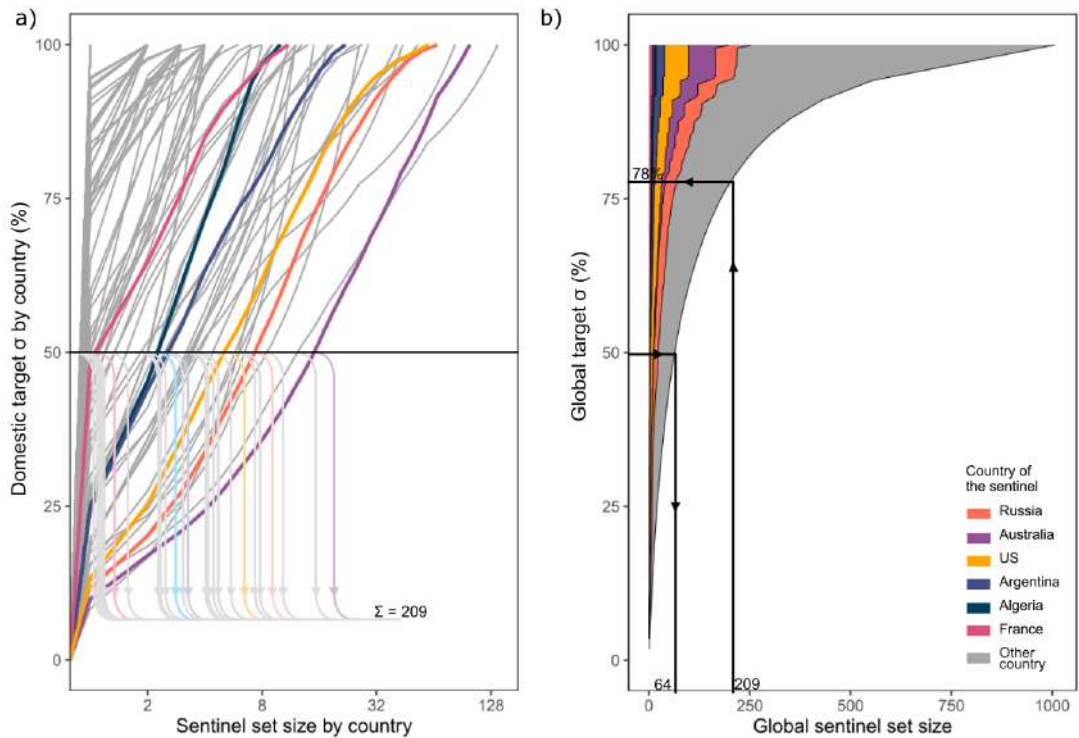


Figure 4.2 – Each line in panel (a) represents the surveillance effort (x-axis, in  $\log_2$  scale) needed by a country to achieve increasing domestic surveillance targets  $\sigma$  (y-axis) in the “non-cooperative” strategy. We highlighted, via colouring, one representative country for each continent. The intersection of each line with a given surveillance target (e.g. horizontal line at  $\sigma = 50\%$ ) gives the minimum size of the sentinel set for that country. The global effort can be obtained by summing all intersections (209 for  $\sigma = 50\%$ ). Panel (b) shows the number of sentinels needed in the “cooperative” strategy to achieve increasing global surveillance targets  $\sigma$ . In this case, the target  $\sigma = 50\%$  is achieved with just 64 sentinels, while 209 sentinels ensure a global coverage of 78%.

#### 4.4.1 Global surveillance effort reduction due to cooperation

In a context of non-cooperation between countries, a coverage of half of the worldwide wheat producing regions (i.e.  $\sigma = 50\%$ ) would be achieved by placing 209 sentinels (Fig. 4.2a), corresponding to 2.7% of the nodes of the global epidemic network. Due to the discrete nature of each coverage, this would correspond to a worldwide target of about  $\sigma = 58\%$  (Fig. 4.2a). Note that with the same amount of sentinels, within a “cooperative” strategy, one would achieve a worldwide coverage of  $\sigma = 78\%$ . On the other hand, the coverage target of  $\sigma = 50\%$  would require only 64 sentinels (Fig. 4.2b). An aggregated coverage of 58% would be obtained with 87 sentinels. If the coverage target were a complete coverage of the worldwide wheat producing regions (i.e.  $\sigma = 100\%$ ), in a “cooperative” framework it would need 1,007 sentinels (Fig. 4.2b) and 1,148 otherwise.

#### 4.4.2 Heterogeneity in the distribution of surveillance effort reduction due to cooperation

Overall, out of 87 countries, 55 (63%) are classified as CoopBeneficial, 23 (27%) as CoopNeutral, and 9 (10%) as CoopAdverse. In terms of wheat production, around 71% is located in CoopBeneficial countries, 6% in CoopNeutral countries while 23% in CoopAdverse ones (Fig. 4.3). A large variety exists in the cost-benefit index by differentiating countries with large (at least 45 nodes), medium (between 44 and 13 nodes) and small producing regions (12 or less nodes; Fig. 4.3 and C.2). For 47 countries, mainly medium (*e.g.*, Czechia or Uruguay) or large (*e.g.*, India or Russia), the cost-benefit index is always  $\leq 1$ , thus implying an advantage in adopting a “cooperative” strategy independently of  $\sigma$ . Only 4 countries (Morocco, Greece, Finland and Nepal) are always discouraged from adopting a “cooperative” strategy. Great part of the small countries (such as Yemen or New Zealand) display  $\alpha_{i,\sigma} = 1$  for any value of  $\sigma$ , for which the two strategies are equivalent. For a few number of large (*e.g.* US, China or Iran) or medium countries (*e.g.* Moldova or Tunisia), the cost-benefit index is lower or larger than one depending on the value of  $\sigma$ . Their qualification as beneficial or adverse to cooperation depends on the surveillance target.

At the world scale, each continent (except Australia) has at least one CoopBeneficial, one CoopNeutral and one CoopAdverse country (Fig. 4.4a). In North America, countries are typically CoopBeneficial, while South America is more balanced. Continental Europe is mainly CoopBeneficial, with some countries (Belgium, Luxembourg, Austria, Slovenia, Croatia, Bosnia and Herzegovina, Albania, North Macedonia) having  $\bar{\alpha}_c = 0$ . Finland has the highest  $\bar{\alpha}_{\text{Finland}}$  of 1.3, followed by Greece ( $\bar{\alpha}_{\text{Greece}} = 1.2$ ). Asia has a composition similar to Europe, with few CoopAdverse countries (China, Mongolia, Nepal), some isolated CoopNeutral (*e.g.* Japan) and a majority of CoopBeneficial ones, mainly in inner parts of the continent. Africa is almost entirely CoopNeutral, with the exception of the Maghreb and Tanzania that are CoopBeneficial. Due to geographic isolation, island states such as Australia and New Zealand are CoopNeutral.

#### 4.4.3 Robustness of the surveillance strategies

Overall, there is good agreement between the values of  $\bar{\alpha}_c$  obtained via the design and the validation network for all countries  $c$  (correlation coefficient of 0.89; p-value  $\ll 0.001$ ). A visual comparison is also provided in Figs C.1, 3-5.

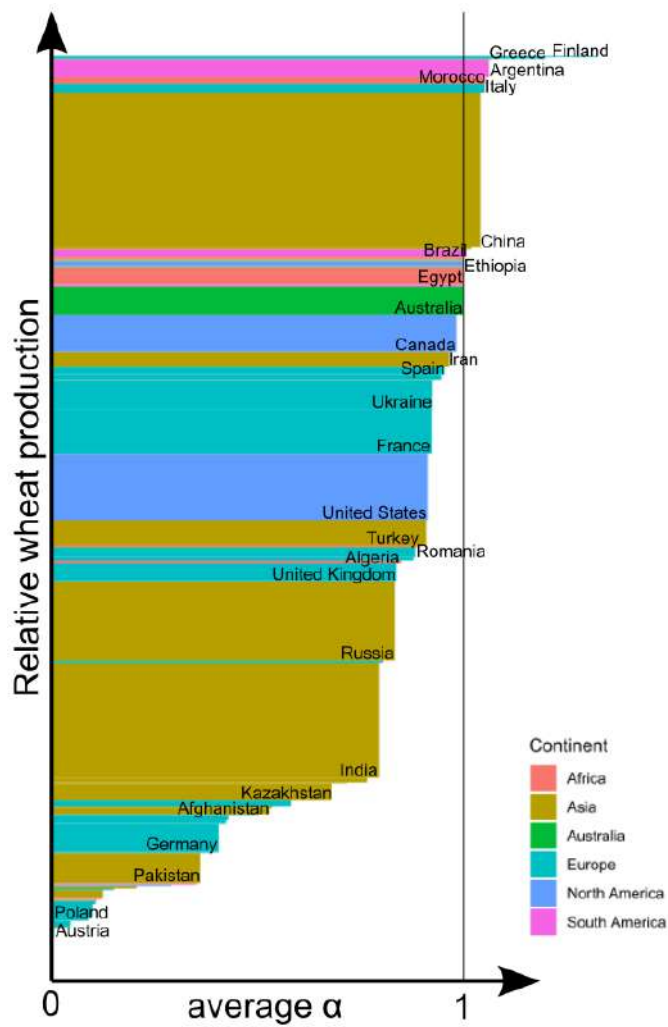


Figure 4.3 – Bar chart of the average cost-benefit index for all wheat-producing countries considered in the study. Each country is represented by a rectangle where the base is proportional to  $\bar{\alpha}_c$  and the height to wheat production in 2010 - 2020 according to FAO (2021b).

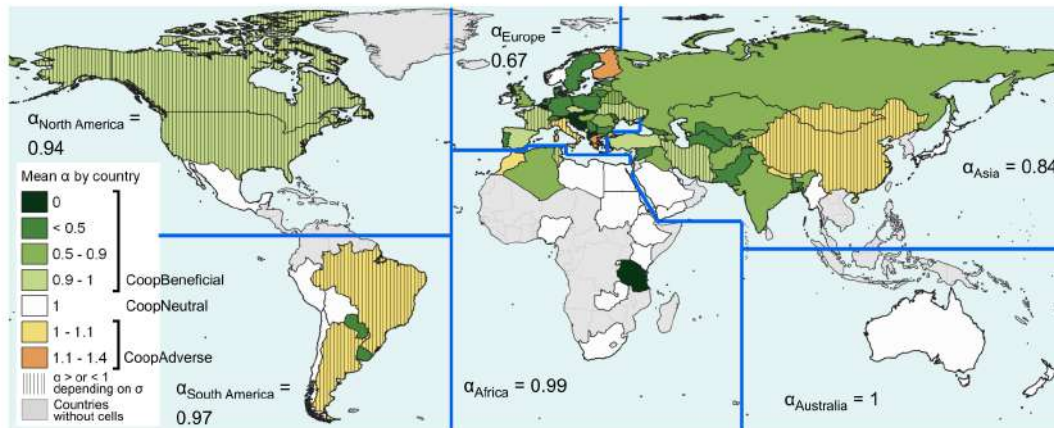


Figure 4.4 – Global map of the average cost-benefit index by country. Average values by continents, weighted by country wheat production 2010 - 2020 (FAO, 2021b) are also displayed. Europe and Asia, and in particular their innermost countries, display the lowest values of  $\bar{\alpha}_c$ . Insular countries (Australia, New Zealand, Japan) or those with limited wheat producing surface (mostly African countries) tend to be CoopNeutral.

## 4.5 Discussion

### 4.5.1 From domestic to global cooperative crop protection

As previous research have stressed, the scale of disease management should correspond to that of the spread of the disease of interest, regardless of country boundaries (Thompson et al., 2016). We have collected evidence that, in the case of long-distance dispersed diseases, a “cooperative” approach allows significant reduction in the surveillance effort needed to achieve a global coverage (-69% and -12% for a global coverage of  $\sigma = 50\%$  and  $100\%$ , respectively). This outcome agrees with previous studies, which underlined that neglecting long-distance connectivity leads to an underestimation of the disease spread capacity (Jeger et al., 2007).

Despite increasing evidence of a global advantage in cooperative international surveillance, crop surveillance design is still mostly dictated by country boundaries, rather than the actual scale of the pathogen spread (Thompson et al., 2016; Carvajal-Yepes et al., 2019). The mismatch between optimal and actual scale of action affects also other kinds of transboundary natural threats, such as biological invasions by alien species. In this regard, Diagne *et al.* (Diagne et al., 2021) recently outlined that invasion-related economic damages are projected to increase in the next decades; one reason behind the inertia in the implementation of international and coordinated protection strategies may lie in the underestimation of the costs by the general public, stakeholders and decision-makers. This may be particularly true in the case of airborne diseases, where the direct observation of their dispersal is ac-

tually unfeasible (Jordano, 2017; Barnes et al., 2020), and may discourage consideration by decision-makers.

#### 4.5.2 Network-thinking in crop surveillance

The use of networks to support crop protection strategies have been largely advocated in recent studies (Jeger et al., 2007; Sutrave et al., 2012; Shaw and Pautasso, 2014; Parnell et al., 2017; Garrett et al., 2018). One advantage of networks is that they are “asemantic”, *i.e.* they can represent whatever relationship, contact or flow mediated by different means (air-masses as well as human transportation (e.g. Brockmann and Helbing, 2013) or animal trade (e.g. Bernini et al., 2019) in a topological space which can correspond to the physical one. In the most simplistic way, crop protection strategies rely on the identification of the nodes of the network that most contribute to spread the disease, or those that, if successfully treated, would reduce the disease size. Other methods rely on the identification of certain recurrent network patterns, where the disease spread is the fastest (Chadès et al., 2011). Concerning surveillance, relevant nodes correspond to those that may allow early disease detection if systematically monitored (Sutrave et al., 2012; Holme, 2017; Neufeld et al., 2018).

Despite the risk of incurring local minima, we used the Set cover algorithm to prioritize nodes to be monitored, *i.e.*, sentinels. Set cover iteratively selects the node associated with the highest coverage, solving the otherwise unsolvable Set cover problem in finite time. This algorithm only ensures that a node is surveilled by at least one sentinel. A less error-prone procedure may request that nodes are surveilled by at least  $n > 1$  sentinels. This would increase the reliability of the sentinel set by reducing the risk of imperfect surveillance (Chadès et al., 2011), but consequently increasing the surveillance effort. Furthermore, in our exercise we assume that the risk of emergence of new strains (made possible by the alternate host *B. vulgaris*, which allows sexual recombination of *P. graminis*), the costs of surveillance, distribution of resistant varieties and crop management practices are the same in all the nodes. Relaxing these assumptions would ask for a different modelling framework, referable to a multi-constrained and multi-objective problem (such as a multi-dimensional knapsack problem Kulik and Shachnai, 2010), with increasing complexity of the solution with respect to that of the Set Cover algorithm.

This algorithm assumes that sentinel locations are chosen regardless of country borders, while it may not be the case. For these reasons, we named the solution of the above-mentioned algorithm as the “cooperative” strategy and we built a second strategy, where surveillance is designed mimicking a more realistic scenario. This strategy, named “non-cooperative”, differs from the previous as the algorithm is carried out each country independently of the others, which means that the Set cover algorithm is solved at the country level. In turn, coverage can be thought as a step-by-step updated version of the in-degree, *i.e.* the number of the edges pointing to a node, penalising those nodes whose coverage overlaps with that of nodes already labelled as sentinels. Other studies already noted that

in-degree (or simply degree for undirected networks) is, as a general rule of thumb, a good proxy of both a good sentinel and a potential disease spreader (Herrera et al., 2016; Holme, 2018).

Moreover, in our work we proposed an hybrid network and geographical approach, in which metadata are associated to network components: each node is associated to the label of the corresponding country, and each edge is consequently labelled as “transboundary” or not. To our knowledge, this is one of the first attempts to compare non-topological surveillance strategies, *i.e.* “cooperative” and “non-cooperative”, and to quantify the heterogeneity in the allocation of the burden of “cooperative” surveillance.

Our results thus indicate that the cooperative strategy becomes more valuable when the surveillance target is intermediate. This is mainly due to the fact that this strategy reduces overlapping among coverages. Overlapping is negligible also for the “non-cooperative” strategy for moderate target of surveillance, and becomes important for both strategies approaching  $\sigma = 100\%$ .

### 4.5.3 Sharing benefits and costs of cooperation

From a global perspective, a “cooperative” strategy is necessarily more efficient compared to a “non-cooperative” one, since it corresponds to an optimization subjected to fewer constraints. However, it is interesting to quantify how such strategy performs against a “non-cooperative” strategy at country level, since benefits and burden may not be equally shared; similarly, wheat production is valuable differently according to each country’s food system.

We found that medium-sized countries located in an inner continental position, such as in central Europe or central Asia, are associated to the lowest  $\bar{\alpha}_c$  values, since they benefit of transboundary potential transport events among a landscape dominated by wheat producing areas. Insular countries, such as Australia, New Zealand or Japan, having no recurrent edges with other countries, are CoopNeutral. Due to the low presence of wheat, many African and South American countries are CoopNeutral. By contrast, it is more difficult to determine general characteristics for CoopAdverse countries, even keeping in mind that connections are mostly north-eastward in the Northern Hemisphere and south-westward in the Southern Hemisphere. Finland and Nepal are small-medium sized wheat producing countries, located at the point of arrival of western-eastern European (Zadoks, 1967) and Indian (Brown and Hovmøller, 2002) “*Puccinia* pathways”, respectively. Given the relatively small size of their wheat producing regions, they are forced to assume more sentinels in the benefits of upwind countries, whose food systems are probably much more wheat-based, than they would need if left alone. By contrast, Canada, the final destination of the north-American pathway, is a large wheat producing country, hence it would need several sentinels no matter the strategy. We may suppose that Italy and Greece, due to their location in the middle of the Mediterranean basin, may play as stepping stones for epidemics spreading northward from Africa towards central Europe (Mehta et al., 2007); furthermore, both

have relatively low wheat productions, hence they would need less sentinels if not cooperating. Brazilian and Argentinian large wheat producing surfaces are located just poleward compared to those of their smaller neighbours (Paraguay and Uruguay, respectively). In the same way, due to the general eastward circulation in the Northern Hemisphere, Chinese wheat-producing regions might act as sink for trajectories from their western neighbours (that are, indeed, CoopBeneficial).

By averaging the cost-benefit index by continent it is possible to highlight those continents which would benefit the most of a cooperative surveillance. Europe and Asia display the lowest cost-benefit index values (0.6-0.8), while for other continents it is generally around 1. To sum up, the connectivity network of this airborne disease creates a heterogeneous distribution of costs and benefits, but Asia and Europe would certainly take advantage of an international and cooperative surveillance system (Fig.s 4.4 and C.5).

The heterogeneous geographical distribution of benefits and costs of cooperation in surveillance has already been highlighted by other studies (Bacon et al., 2012) and suggests that a compensating mechanism should be set up to make it acceptable. This compensation mechanism should take into account different costs of surveillance among countries (Augustin et al., 2012). This idea can be borrowed from the socioeconomic concept of “burden sharing” (Suhrke, 1998; Sandler and Forbes, 1980), which is finding application in the management of environmental goods. Differentiate greenhouse gas emissions reduction in the framework of the Conference of the Parties to achieve climate targets (Ringius et al., 2002), as well as in the multi-stakeholders management of marine resources (Bennett et al., 2021), may be two notably example. Furthermore, other fields of crop protection may benefit of a network-based transboundary perspective. For example, the deployment of resistant varieties to both contain pathogens spread and delay resistance overcoming (Rimbaud et al., 2018) is another spatial optimization problem; whether it should be approached at the national or international scale is an interesting issue that can benefit from the approach proposed here.

While our study tries to push towards a change in the perspective of governance of crop disease surveillance, we believe that proper identification of spatial distribution of costs and benefits can help facilitate international agreement for a global crop epidemic surveillance and gain support of all stakeholders.

## Supplementary Information

- C.1 Supplementary materials
- C.2 The published article



## Bibliography

- Allen-Sader, C., Thurston, W., Meyer, M., Nure, E., Bacha, N., Alemayehu, Y., Stutt, R. O., Safka, D., Craig, A. P., Derso, E., Burgin, L. E., Millington, S. C., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2019). An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11).
- Augustin, S., Boonham, N., De Kogel, W. J., Donner, P., Faccoli, M., Lees, D. C., Marini, L., Mori, N., Petrucco Toffolo, E., Quilici, S., Roques, A., Yart, A., and Battisti, A. (2012). A review of pest surveillance techniques for detecting quarantine pests in Europe. *EPPO Bulletin*, 42(3):515–551.
- Aylor, D. E. (2003). Spread of plant disease on a continental scale: Role of aerial dispersal of pathogens. *Ecology*, 84(8):1989–1997.
- Bacon, S. J., Bacher, S., and Aebi, A. (2012). Gaps in border controls are related to quarantine alien insect invasions in Europe. *PLoS One*, 7(10):e47689.
- Barnes, G., Saunders, D. G., and Williamson, T. (2020). Banishing barberry: The history of *Berberis vulgaris* prevalence and wheat stem rust incidence across Britain. *Plant Pathology*, 69(7):1193–1202.
- Bennett, N. J., Blythe, J., White, C. S., and Campero, C. (2021). Blue growth and blue justice: Ten risks and solutions for the ocean economy. *Marine Policy*, 125:104387.
- Bernini, A., Bolzoni, L., and Casagrandi, R. (2019). When resolution does matter: Modelling indirect contacts in dairy farms at different levels of detail. *PLoS one*, 14(10):e0223652.
- Brockmann, D. and Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342(6164):1337–1342.
- Brown, J. K. and Hovmøller, M. S. (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297(5581):537–541.
- Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K. A., Giovani, B., Saunders, D., Kamoun, S., Legg, J., Verdier, V., Lessel, J., et al. (2019). A global surveillance system for crop diseases. *Science*, 364(6447):1237–1239.
- Chadès, I., Martin, T. G., Nicol, S., Burgman, M. A., Possingham, H. P., and Buckley, Y. M. (2011). General rules for managing and surveying networks of pests, diseases, and endangered species. *Proceedings of the National Academy of Sciences*, 108(20):8323–8328.
- Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore Piontti, A., Mu, K., Rossi, L., Sun, K., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489):395–400.

- Corredor-Moreno, P. and Saunders, D. G. (2020). Expecting the unexpected: factors influencing the emergence of fungal and oomycete plant pathogens. *New Phytologist*, 225(1):118–125.
- Dhama, K., Khan, S., Tiwari, R., Sircar, S., Bhat, S., Malik, Y. S., Singh, K. P., Chaicumpa, W., Bonilla-Aldana, D. K., and Rodriguez-Morales, A. J. (2020). Coronavirus disease 2019–covid-19. *Clinical microbiology reviews*, 33(4):e00028–20.
- Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I., Salles, J.-M., Bradshaw, C. J., and Courchamp, F. (2021). High and rising economic costs of biological invasions worldwide. *Nature*, 592(7855):571–576.
- Draxler, R. R. and Hess, G. D. (1998). An overview of the HYSPLIT 4 modelling system for trajectories, dispersion and deposition. *Australian Meteorological Magazine*, 47(4):295–308.
- FAO (2021a). Fao - country brief. <http://www.fao.org/giews/countrybrief/>. Accessed: 2021-07-20.
- FAO (2021b). World food and agriculture—statistical yearbook 2021. *World Food and Agriculture-Statistical Yearbook*.
- Garrett, K., Alcalá-Briseño, R., Andersen, K., Buddenhagen, C., Choudhury, R., Fulton, J., Hernandez Nopsa, J., Poudel, R., and Xing, Y. (2018). Network analysis: A systems framework to address grand challenges in plant pathology. *Annual review of phytopathology*, 56:559–580.
- Herrera, J. L., Srinivasan, R., Brownstein, J. S., Galvani, A. P., and Meyers, L. A. (2016). Disease surveillance on complex social networks. *PLoS computational biology*, 12(7):e1004928.
- Holme, P. (2017). Three faces of node importance in network epidemiology: Exact results for small graphs. *Physical Review E*, 96(6):062305.
- Holme, P. (2018). Objective measures for sentinel surveillance in network epidemiology. *Physical Review E*, 98(2):022313.
- Isard, S. A., Gage, S. H., Comtois, P., and Russo, J. M. (2005). Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience*, 55(10):851–861.
- Jeger, M. J., Pautasso, M., Holdenrieder, O., and Shaw, M. W. (2007). Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist*, 174(2):279–297.
- Jordano, P. (2017). What is long-distance dispersal? and a taxonomy of dispersal events. *Journal of Ecology*, 105(1):75–84.

- Kulik, A. and Shachnai, H. (2010). There is no eptas for two-dimensional knapsack. *Information Processing Letters*, 110(16):707–710.
- Levetin, E. (2015). Aerobiology of Agricultural Pathogens. *Manual of Environmental Microbiology*, pages 3.2.8–1–3.2.8–20.
- Mehta, S. V., Haight, R. G., Homans, F. R., Polasky, S., and Venette, R. C. (2007). Optimal detection and control strategies for invasive species management. *Ecological Economics*, 61(2-3):237–245.
- Meyer, M., Cox, J. A., Hitchings, M. D., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nature Plants*, 3(10):780–786.
- Mohamed, K., Rodríguez-Román, E., Rahmani, F., Zhang, H., Ivanovska, M., Makka, S. A., Joya, M., Makuku, R., Islam, M. S., Radwan, N., et al. (2020). Borderless collaboration is needed for covid-19—a disease that knows no borders. *Infection Control & Hospital Epidemiology*, 41(10):1245–1246.
- Neufeld, K., AP, K., Gugino, B., McGrath, M., Sikora, E., Miller, S., Ivey, M., Langston, D., Dutta, B., Keever, T., Sims, A., and Ojiambo, P. (2018). Predicting the risk of cucurbit downy mildew in the eastern United States using an integrated aerobiological model. *International Journal of Biometeorology*, 62(4):655–668.
- Park, R., Fetch, T., Hodson, D., Jin, Y., Nazari, K., Prashar, M., and Pretorius, Z. (2011). International surveillance of wheat rust pathogens: Progress and challenges. *Euphytica*, 179(1):109–117.
- Parnell, S., van den Bosch, F., Gottwald, T., and Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: an epidemiological perspective. *Annual review of phytopathology*, 55:591–610.
- Radici, A., Martinetti, D., and Bevacqua, D. (2023). Global benefits and domestic costs of a cooperative surveillance strategy to control transboundary crop pathogens. *Plants, People, Planet*, (April):1–10.
- Rimbaud, L., Papaix, J., Rey, J.-F., Barrett, L. G., and Thrall, P. H. (2018). Assessing the durability and efficiency of landscape-based strategies to deploy plant resistance to pathogens. *PLOS Computational Biology*, 14(4):1–33.
- Ringius, L., Torvanger, A., and Underdal, A. (2002). Burden sharing and fairness principles in international climate policy. *International Environmental Agreements*, 2(1):1–22.
- Ristaino, J. B., Anderson, P. K., Bebber, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., et al. (2021). The persistent threat

- of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences*, 118(23).
- Sandler, T. and Forbes, J. F. (1980). Burden sharing, strategy, and the design of nato. *Economic inquiry*, 18(3):425–444.
- Saunders, D. G., Pretorius, Z. A., and Hovmøller, M. S. (2019). Tackling the re-emergence of wheat stem rust in western europe. *Communications biology*, 2(1):51.
- Shaw, M. and Pautasso, M. (2014). Networks and plant disease management: Concepts and applications. *Annual Review of Phytopathology*, 52:477–493.
- Singh, R. P., Hodson, D. P., Jin, Y., Lagudah, E. S., Ayliffe, M. A., Bhavani, S., Rouse, M. N., Pretorius, Z. A., Szabo, L. J., Huerta-Espino, J., et al. (2015). Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. *Phytopathology*, 105(7):872–884.
- Soubeyrand, S., Demongeot, J., and Roques, L. (2020). Towards unified and real-time analyses of outbreaks at country-level during pandemics. *One Health*, 11:100187.
- Suhrke, A. (1998). Burden-sharing during refugee emergencies: The logic of collective versus national action. *Journal of refugee studies*, 11(4):396–415.
- Sutrave, S., Scoglio, C., Isard, S. A., Hutchinson, J. S., and Garrett, K. A. (2012). Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS One*, 7(6):e37793.
- Thompson, R. N., Cobb, R. C., Gilligan, C. A., and Cunniffe, N. J. (2016). Management of invading pathogens should be informed by epidemiology rather than administrative boundaries. *Ecological Modelling*, 324:28–32.
- Thompson, R. N., Hollingsworth, T. D., Isham, V., Arribas-Bel, D., Ashby, B., Britton, T., Challenor, P., Chappell, L. H., Clapham, H., Cunniffe, N. J., et al. (2020). Key questions for modelling covid-19 exit strategies. *Proceedings of the Royal Society B*, 287(1932):20201405.
- Xing, Y., Hernandez Nopsa, J. F., Andersen, K. F., Andrade-Piedra, J. L., Beed, F. D., Blomme, G., Carvajal-Yepes, M., Coyne, D. L., Cuellar, W. J., Forbes, G. A., et al. (2020). Global cropland connectivity: A risk factor for invasion and saturation by emerging pathogens and pests. *BioScience*, 70(9):744–758.
- Zadoks, T. (1967). Epidemiology of wheat rust in europe. *International Journal of Pest Management B*, 13(1):29–46.

**Key points:**

- The high rate of spread of infectious diseases or invasive species questions the conventional paradigm whereby each country alone can efficiently handle such threats.
- In the case of stem rust of wheat, we use the global epidemic network modelled in Chapter 2 to show that a cooperative surveillance strategy would help reducing the global monitoring effort, notably at intermediate target of surveillance (e.g. when 50% of the global wheat surfaces are monitored), compared to a scenario where each country optimises its own surveillance.
- Such benefits are distributed heterogeneously across the globe. Depending on the surveillance target, some countries would be asked to increase their domestic surveillance effort in the global interest: compensation mechanisms should be put in place to reach an equitable share of benefits and costs.

**Perspectives:**

- Further research may consider modifying the “Set cover” algorithm to ensure a node is surveilled by more than one sentinel. This would probably shrink the coverage of each sentinel, and would reasonably lead to reduced transboundary interactions.
- As in Chapter 2, presence of the secondary host, which provide genetic reshuffling and favours the emergence of new strains, is neglected.

# Conclusion and perspectives

The answer, my friend, is blowin' in the wind  
The answer is blowin' in the wind

– Bob Dylan, *Blowin' in the wind*

## Reducing epidemiological complexity at its essential

The first article I read at the beginning of my theses was proposed me by Davide. It was the paper by Brockmann and Helbing (2013), *The hidden geometry of complex, network-driven contagion phenomena*, which I already mentioned in the Introduction. This paper stated that spatio-temporal spread of contemporary pandemics is guided by the emergent transport mechanism, which is air-traffic, with epidemiological parameters playing a secondary or even negligible role. Moving up one level of abstraction, this paper stated that, at least in spatial epidemiology, very complex phenomena may become simple if reduced at its essential.

This reduction to essential is one Holy Grail for modelers aiming to capture the “big picture” of phenomena. The other approach is at the precise antipode: including every detail. I was fascinated by this visionary article: from one side, February 2021 was still “Covid time”, and so this topic was very present to our lives; from the other, it seemed the starting point of a broad applicability of its tools - complex networks, to which Chapter 1 is entirely dedicated - in whatever framework in epidemiology.

With anecdotal exceptions, contrary to humans, plants do not fly (Keeling and Gilligan, 2000), though they sometimes unintentionally take the plane for being traded (Banks et al., 2015). Nevertheless, their parasites can move: they can be transported by various means, including their own motion, animals, water bodies, rain-splash, and wind (Damschen et al., 2008; Nagarajan and Singh, 1990). This latter means of transportation is particularly interesting from various perspectives. It is certainly one means of transportation which allows long-distance transport of pathogens (Aylor, 2003). Although many harmful plant pathogens are wind-dispersed, traditionally the lack of methods to study such a complex mode of transportation has been accompanied by the belief that its study would not yield

practical results due to the stochastic nature of air masses (Kling and Ackerly, 2021).

On the contrary, new methods based on Lagrangian trajectory simulations have emerged, initially designed to study the possible impact of air pollutants (Draxler and Hess, 1998). Models such as HYSPLIT or NAME (Jones et al., 2007) have provided scientists with tools that integrate a wide range of environmental data at increasing spatial resolutions. For instance, the main environmental database used by HYSPLIT, called Global Data Assimilation System (GDAS), increased its resolution from 1° to 0.25° from 2006 to 2019. These models have progressively garnered the scientific interest for the application in studying airborne pathogens, in particular stem rust of wheat, caused by *P. graminis* (Meyer et al., 2017b,a; Visser et al., 2019; Allen-Sader et al., 2019; Prank et al., 2019).

I was originally supposed to work on brown rot of peaches. However, the lack of a modelling framework for approaching the airborne dispersal of *Monilinia* spores, primarily due to the anecdotal knowledge of its aerobiology, lead me to concentrate initially on *P. graminis*. I had the precise objective of building a replicable modelling framework, flexible enough to be adapted to whatever wind-transported pathogen by tuning some parameters, keeping Brockmann and Helbing's philosophy: think at the big picture (so, moving to a global scale) by taking the essential.

I soon found myself in a situation where I had to make a compromise. In fact, Lagrangian trajectories alone may provide superfluous information. Characterizing them, especially in terms of spatio-temporal coordinates of release or simulation duration, helps assign them the role of representing the dispersal of a specific pathogen. Moreover, the intersection of such trajectories with the host grids just produced a over-connected network - in jargon: a *hairball* graph (Dianati, 2016) - in which every node was connected to almost any other.

The idea of using suitability filters emerged from the dual necessity of characterizing air-mass trajectories and simplifying networks. These filters reduce the amount of unwarranted information by mapping when and where wheat is present and susceptible, or launching backward trajectory simulations only from those place and time when precipitation is falling, to capture only wet deposition events (the one where there is the highest probability of infection; Emerson et al., 2020). Other "filters" include the fact that airborne spores should resist UV radiation and rain washout (Isard et al., 2005); trajectories should be within the planetary boundary layer to have the possibility to effectively drag released spores; and so on. Great part of the Methods section of Chapter 2 is dedicated to their description.

The suitability filters represent a smart solution to deal with the above-mentioned issues. Other studies simulating spore dispersal leveraged on the so called "dispersion model" of NAME (or HYSPLIT), in which turbulent motion of a large number of Lagrangian particles is taken into account, producing then the so called "plume", and are also provided with a "dry deposition module" simulating effect of gravity on spores (Meyer et al., 2017b). The "trajectory model", which I used, simply integrates one curvilinear trajectory in the atmo-

sphere, along which environmental variables are computed. While the first is enormously demanding in computational terms, allowing a limited set of release sites, the second is more handy, and with the help of the suitability filters I've been able to run more and more meaningful simulations covering the world surface at greater density and frequency with respect to studies which used the "dispersion model" (Meyer et al., 2017b).

I was starting to getting far from the article by Brockmann and Helbing. On the one hand, these suitability filters helped in reducing the great amount of information HYSPLIT provided to the essential; on the other hand, these filters were selecting the emerging mechanisms at the cost of adding multiple layers of biophysical knowledge.

## **Modelling plant epidemics deals with multiple layers of complexities**

I stored the information about intensity and frequency of airborne epidemic connections in a network object. This allowed to take advantage of the knowledge summarised in Chapter 1, which inspired me to conceive a surveillance prioritization algorithm with which I identified a set of "sentinels", i.e. easily infected nodes to monitor frequently\*. It is, indeed, a greedy algorithm, which means that its solution is locally optimal at each step: but all the attempts of Davide of finding a more effective algorithm failed. In fact, this algorithm addresses the problem of redundancy in spatial networks, which tends to have rich-club structures (Colizza et al., 2006) - highly connected nodes tends to be clustered, and are consequently less useful under the perspective of surveillance, where one would like to differentiate the sources of information by checking distant subsets of the network. The idea behind this algorithm is close to that of the VoteRank (Zhang et al., 2016), an iterative algorithm which defines, at each step, the best "spreader" among the nodes, decreasing the weight of the spreader's connections in the following iterations.

The sentinel set identified via the "Set cover" performs well also with respect to slight changes of the connectivity network (I defined the set starting from the network summarising airborne connections in the period 2013-2016, while I tested it with the network summarising airborne connections in 2017-2018). However, I acknowledged that an "on field" validation of the sentinel set would present serious challenges, which deal with the necessarily broad definition of "node" and its monitoring, since this includes all wheat fields within a square of around 2,000  $km^2$ . Even before coming to the sentinel set, the validation of the original network itself would be complicate. I believe that a possible validation procedure could be based on comparing the genetic similarity between strains and between those nodes of the networks in which the strains have been found. However, this method would

---

\*Fun fact: I had the idea of the "Set cover" algorithm, found on Wikipedia - [https://en.wikipedia.org/wiki/Set\\_cover\\_problem#Greedy\\_algorithm](https://en.wikipedia.org/wiki/Set_cover_problem#Greedy_algorithm) - which I later interpreted as a step-wise updated version of the in-degree network centrality, during my holidays in Côte d'Azur, some days before the 2021 Footbal Euro Final.



require to have a widespread measures of *Puccinia* genotypes, while the emergence of new strains is (luckily) relegated to those locations where the secondary host, the barberry, is present (Barnes et al., 2020; Olivera et al., 2015).

A reviewer proposed a further crash test to determine whether the proposed sentinel set was also capable of detecting the emergence of the disease earlier than sentinel sets defined according to other ranking criteria. This test implied a shift of the perspective of the network, from static to dynamic, through the creation of an elementary metapopulation model, which I built under trivial assumptions. Eventually, it resulted that the “Set cover” generally performed better in providing the earliest detection than other standard metrics.

The rudimentary metapopulation model developed for *P. graminis* called for a more sophisticated approach, which, beside simulating pathogen jump from one node to another, embedded environmental heterogeneity in favour of more adherence to reality. Daniele was already finalising a climate-based local phenological temperature-driven model for determining the period of susceptibility of peaches (Vanalli et al., 2021) and on a climate-dependent epidemiological model for describing brown rot dynamics in an orchard. This idea was far from that of Brockmann and Helbing (2013): integrating all these dynamics, would have really meant “getting off the plane” and include many epidemiological, environmental and biological layers in the model.

The model presented in Chapter 3 is far from a “reduction to its essential”. I eventually included more dynamics compared to what I expected, such as the stochastic presence of primary inoculum due to overwintering rotten fruits and expert judgement about the incidence of the disease-related losses. Moreover, the Monte Carlo approach implies unhandy and time-consuming simulations. On the other hand, I kept almost unchanged all the conceptual framework about Lagrangian trajectory simulations and suitability filters, conceived for stem rust of wheat. Eventually, these modifications were necessary to calibrate the parameters, given the extremely heterogeneous available database of observations. These efforts produced a more ambitious and complete model, able to describe the climate-driven multi-seasonal spread of brown rot of peach balancing in-orchard disease transmission with inter-units spread. With respect to the global epidemic *Puccinia* network, the proposed metapopulation network is calibrated and cross-validated with direct observations.

The epidemic risk indices proposed in Chapter 3 present complementary roles related to crop protection. A vulnerable node calls for direct protection, while a dangerous node have an explicit role in determining outbreak at the landscape scale. Despite this distinction, it should be acknowledged that these two indices are spatially correlated (one can’t expect vulnerable cells being far from the dangerous ones!)

Chapter 3 ends presenting the potential use of the metapopulation model in spatial epidemic management design. The spatial optimization of control methods (such as application of fungicides) and testing the effectiveness through a Monte Carlo approach represents

the most immediate perspective of the presented work. Even if I did not have the opportunity of exploring it properly during my Ph.D., in the next section I am illustrating a possible draft of procedure.

## Network-based optimisation of fungicide application (some more equations)

The starting point is the observation that brown rot of peaches is already everywhere in France and *M. fructicola* is no more in the EU quarantine list (Oliveira Lino et al., 2016). Well established within the Green Revolution paradigm, farmers use to apply between 15 and 25 fungicides treatments on average to conventional orchards every year (Cretin et al., 2018). One typical fungicide which is applied during ripening is Azoxystrobin Martini and Mari (2014), a respirator inhibitor which reduces the infection (in the presented modelling framework, it reduces the rate  $\sigma$  of the transition from “exposed” to “infected”).

From a mathematical point of view (Hobbelen et al., 2011; Elderfield et al., 2018), the parameter  $\alpha$ , which tunes the infection rate, might depend on the concentration of fungicide  $F$  according to Eq. 2:

$$\alpha = \alpha_{max}(1 - e^{\theta F}) \quad (2)$$

Where  $\alpha_{max}$  is the maximum rate (usually assumed as unitary) and  $\theta$  determines the curvature of the dose-response curve. Fungicide concentration dynamics, instead, follows an exponential decay depending on a decay rate  $\nu$  - except when a new dose is applied, when its value is usually reset to its maximum:

$$dF(t)/dt = -\nu F(t) \quad (3)$$

The system of equations 3.1 introduced in Chapter 3, can be re-defined as Eq. 4, where the transition from exposed to infected is then tuned by parameter  $\alpha$ , whose value goes from 0 to 1 and decreases with increasing fungicide concentration.

$$\begin{cases} dS(t)/dt = -\beta S(t)I(t) + \eta(T(t))E(t) - \mu S(t) \\ dE(t)/dt = \beta S(t)I(t) - (\eta(T(t)) + (1 - \alpha)\sigma(\Pi(t)))E(t) - \mu E(t) \\ dI(t)/dt = (1 - \alpha)\sigma(\Pi(t))E(t) - (\mu + \alpha)I(t) \end{cases} \quad (4)$$

By using a parameterization assumed in previous studies for simulating the effects of Azoxystrobin in another fungal disease, powdery mildew (*Erysiphe necator*) on grapevines (Hobbelen et al., 2011; Elderfield et al., 2018), I am able to assess the effects of the application of fungicides in any unit of the domain.

One thing this modelling framework allows to do is to choose a set of nodes where fungicides are applied and then assess the overall effect on peach production - in other words, to prioritize units of the network to be treated first. This prioritization is proposed to explore possible compromises between increasing farmer's income (here simplified in "reducing disease-related production losses") and reducing environmental impact of chemicals.

These objectives are conflicting since they require either to maintain high fungicide application or to reduce it, respectively, at least at the short/medium run. In the long run, due to a number of factors, including the development of fungicide resistance (Lucas et al., 2015), it is plausible that these two objectives are less conflicting (Olitaa et al., 2023).

A possible option to prioritise units to be treated involves ranking units according to an index or a metric - exactly as proposed in Chapter 1 - and simulate the expected fruit production with Monte Carlo repetitions with growing treated areas. This idea led me to run an exploratory study in which I compare two different strategies, "rain-based", where the units where rain is most frequent during ripening are treated first, and "in-degree-based", where the units with highest in-degree are treated first (Fig. 5). As anticipated in Chapter 3, rain is related to vulnerability, because it causes cuticle cracking. In-degree, instead, is a network metric presented in Chapter 1 which measures the number of edges entering a node. In this case, the in-degree of a unit  $i$  is the number of units which are expected to disseminate their spores to that unit  $i$  according to the wind-driven matrix  $W$  presented in Chapter 3.

If there were no "network effect" - that is, if an action on a node only affects that node, without interactions with the others - I would expect the production to increase linearly with the treated surface (the solid line in Fig. 5). What I observe instead is that, at least for the "rain-based" strategy, the curve describing the expected production in function of treated surfaces has a concavity downwards. This non-linear behaviour has interesting consequences: it implies that, in a scenario where all orchards are already treated, a slight reduction of fungicides would keep the overall production almost unchanged; or that, at the opposite, in a fungicide-free scenario (but where losses are constantly between 35% and 65%), a reasoned spatial allocation of fungicide application would significantly increase the production. Instead, compared to the "rain-based" strategy, the "in-degree-based" do not seem to represent an interesting ranking criteria. Indeed, the trajectory of the expected production is almost convex.

Further research is needed, but this first preliminary experiment suggests a couple of observations. The first is that disease control strategies display a "network effect", and it is possible to test them on a network basis. The second is that, for brown rot of peaches, the most promising strategies seem based on local climate characteristics (frequency of rainy days) rather than on a network-mediated wind index (the in-degree of the airborne connectivity network). At least in this pathosystem, the dependence of disease control actions with weather conditions is crucial (Bevacqua et al., 2023).

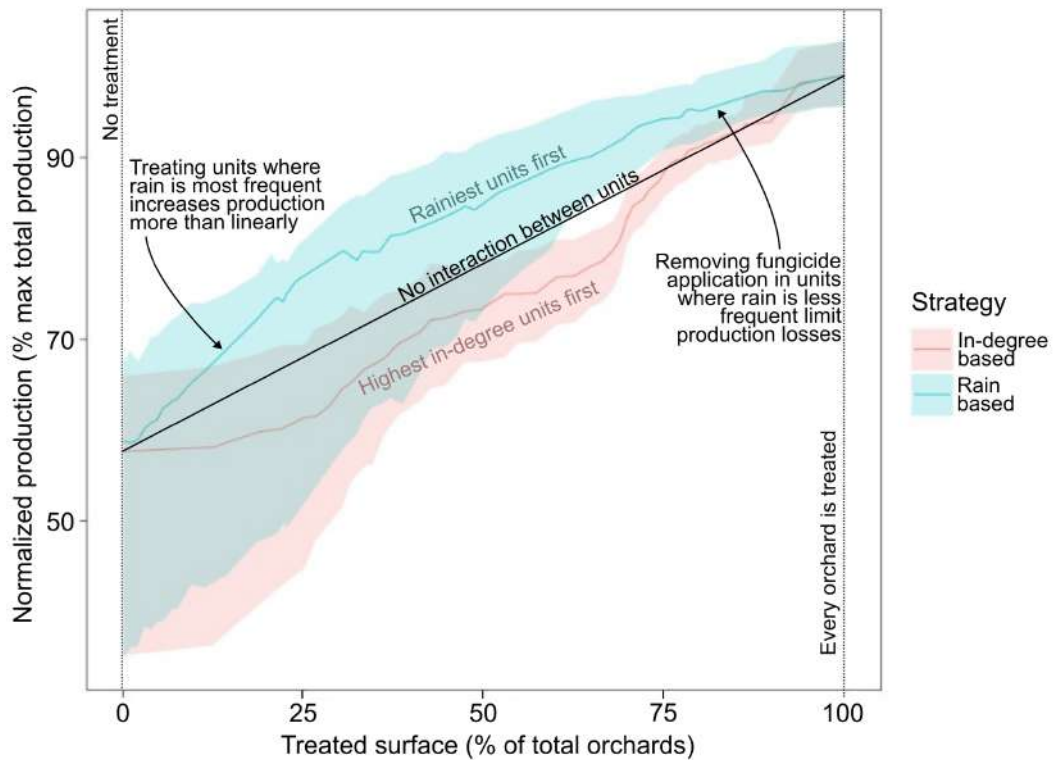


Figure 5 – Normalized peach production (in terms of total weight) with different amount of treated surfaces (in terms of hectares) according to two strategies. The shaded area represent the interquartile interval of the Monte Carlo simulations, while the solid line represent the average.

## Let us zoom out, it is a matter of scale

In the case of the dissemination of an airborne pathogen whose capability to endure harsh atmospheric conditions among a worldwide distributed host, the above-mentioned “network effect” between units may translate into a transboundary effect between countries. Chapter 4 is entirely devoted to that, focused again on “*P. graminis*” and stem rust of wheat. I used the methodological framework developed in Chapter 2 to design a country-based surveillance strategy, based on the solution of the “Set cover algorithm” for each country separately. For doing that, I removed all edges connecting nodes placed in different countries, so that to each country corresponds (at most) an isolated sub-network component (John and Allan, 1995). This procedure allowed me to state that the performance of a “non-cooperative” strategy, i.e. each country alone, are globally worse than the one developed in Chapter 2, labelled as “cooperative”, in terms of effort (i.e., the number of deployed sentinels) needed to reach the same surveillance target.

Is this a “happily ever after” statement? Not exactly, since also a “cooperative” strategy has some drawbacks. Beside the overlooked costs of coordinating several countries under a unified framework, the global reduction of surveillance effort is not equally distributed. I estimated that, in this scenario, a few countries should increase their surveillance efforts for the global benefit. This heterogeneous distribution of efforts should be acknowledged to facilitate international agreement for a global crop protection and gain support of all stakeholders.

The studies conducted in this thesis have the aim of exploring the advantages of broadening the spatial scale in current paradigm of crop protection in agriculture. I have investigated this possibility for long-distance dispersed parasites, in particular on airborne pathogens, for both surveillance and control. In doing this, I have been inspired by a visionary study from which I gradually moved away in favour of a more accurate description of the aerobiological, phenological, epidemiological and agronomic dynamics that made me more aware of the complexity of the problems involved in contemporary agriculture, of which I initially had limited knowledge.

The Green Revolution has profoundly reshaped farming practices over the last 80 years. While it has brought significant benefits, it has also given rise to unexpected issues, some of which I partially discussed in the introduction.

One problematic aspect concerns the way crop protection is approached, which sometimes conflicts with the long-term sustainability of agriculture and human activities in general. The indiscriminate use of chemicals and phytosanitary products has fostered the hope of finding a panacea for an age-old problem that has plagued humanity since prehistoric times: the constant battle against pathogens and pests. In the case of airborne pathogens, this approach was also driven by the unpredictable emergence of their symptoms.

The emergence of pathogen resistance to pesticides and growing environmental concerns

related to the vast resources conventional agriculture consumes have prompted scientists and institutions to reconsider the paradigm established by the Green Revolution. This shift has led to technological and methodological innovations that are being applied in various ways across the world.

Through this thesis, I aimed to contribute to one of the most urgent methodological changes in crop protection, which involves zooming out from individual plots and recognizing in the network of interactions between farms, fields, and orchards a fundamental ecological function. Identifying these connections, which summarize potential epidemic connections, is essential for collectively designing innovative, viable and effective surveillance and control strategies.

# References

## Bibliography

- Abel, W. (2013). *Agricultural Fluctuations in Europe: From the Thirteenth to twentieth centuries*. Routledge.
- Adam, D. et al. (2021). How far will global population rise? researchers can't agree. *Nature*, 597(7877):462–465.
- Allen-Sader, C., Thurston, W., Meyer, M., Nure, E., Bacha, N., Alemayehu, Y., Stutt, R. O., Safka, D., Craig, A. P., Derso, E., Burgin, L. E., Millington, S. C., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2019). An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11).
- Amar, H. and Dupuy, C. (2020). Les plateformes d'épidémiosurveillance: un concept novateur au service de l'efficacité des dispositifs de surveillance. *Bulletin de l'Académie Vétérinaire de France*, 173(1):200–205.
- Andersen, K., Buddenhagen, C., Rachkara, P., Gibson, R., Kalule, S., Phillips, D., and Garrett, K. (2019). Modeling epidemics in seed systems and landscapes to guide management strategies: the case of sweet potato in northern Uganda. *Phytopathology*, 109(9):1519–1532.
- Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4:40 – 79.
- Augustin, S., Boonham, N., De Kogel, W. J., Donner, P., Faccoli, M., Lees, D. C., Marini, L., Mori, N., Petrucco Toffolo, E., Quilici, S., Roques, A., Yart, A., and Battisti, A. (2012). A review of pest surveillance techniques for detecting quarantine pests in Europe. *EPPO Bulletin*, 42(3):515–551.
- Aylor, D. (2017). *Aerial Dispersal of Pollen and Spores*. The American Phytopathological Society.
- Aylor, D. E. (1986). A framework for examining inter-regional aerial transport of fungal spores. *Agricultural and Forest Meteorology*, 38(4):263–288.

- Aylor, D. E. (1999). Biophysical scaling and the passive dispersal of fungus spores: relationship to integrated pest management strategies. *Agricultural and Forest Meteorology*, 97(4):275–292.
- Aylor, D. E. (2003). Spread of plant disease on a continental scale: Role of aerial dispersal of pathogens. *Ecology*, 84(8):1989–1997.
- Bacon, S. J., Bacher, S., and Aebi, A. (2012). Gaps in border controls are related to quarantine alien insect invasions in europe. *PLoS One*, 7(10):e47689.
- Bailey, D. J., Otten, W., and Gilligan, C. A. (2000). Saprotrophic invasion by the soil-borne fungal plant pathogen *rhizoctonia solani* and percolation thresholds. *The New Phytologist*, 146(3):535–544.
- Baiocco, S., Cavina, F., and Pradolesi, G. (2021). A weather-based simulation model for the development of wheat stem rust epidemics. In *2021 29th Conference of Open Innovations Association (FRUCT)*, pages 22–29. IEEE.
- Banks, N. C., Paini, D. R., Bayliss, K. L., and Hodda, M. (2015). The role of global trade and transport network topology in the human-mediated dispersal of alien species. *Ecology letters*, 18(2):188–199.
- Barabási, A.-L. (2016). *Network science*. Cambridge university press.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Barnes, G., Saunders, D. G., and Williamson, T. (2020). Banishing barberry: The history of *berberis vulgaris* prevalence and wheat stem rust incidence across britain. *Plant Pathology*, 69(7):1193–1202.
- Bennett, N. J., Blythe, J., White, C. S., and Campero, C. (2021). Blue growth and blue justice: Ten risks and solutions for the ocean economy. *Marine Policy*, 125:104387.
- Bernini, A., Bolzoni, L., and Casagrandi, R. (2019). When resolution does matter: Modelling indirect contacts in dairy farms at different levels of detail. *PloS one*, 14(10):e0223652.
- Bertuzzi, P. and Clastre, P. (2022). Information sur les mailles SAFRAN.
- Bevacqua, D., Quilot-Turion, B., and Bolzoni, L. (2018). A model for temporal dynamics of brown rot spreading in fruit orchards. *Phytopathology*, 108(5):595–601.
- Bevacqua, D., Vanalli, C., Casagrandi, R., and Gatto, M. (2023). A climate-driven compartmental model for fungal diseases in fruit orchards: The impacts of climate change on a brown rot-peach system. *Agricultural and Forest Meteorology*, 332:109293.



- Bhattacharya, S. (2017). Deadly new wheat disease threatens europe's crops. *Nature*, 542(7640).
- Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., and Mills, M. C. (2020). Social network-based distancing strategies to flatten the covid-19 curve in a post-lockdown world. *Nature Human Behaviour*, 4(6):588–596.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4-5):175–308.
- Börner, K., Sanyal, S., and Vespignani, A. (2007). Network science. *Annual review of information science and technology*, 41(1):537–607.
- Brockmann, D. and Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342(6164):1337–1342.
- Brown, J. K. and Hovmöller, M. S. (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297(5581):537–541.
- Bryde, R. L. W. and Willets, H. J. (1977). Brown rot fungi of fruit trees. *Nature*, 146(3698):370.
- Buddenhagen, C., Hernandez Nopsa, J.-F., Andersen, K. F., Andrade-Piedra, J., Forbes, G.-A., Kromann, P., Thomas-Sharma, S., Useche, P., and Garrett, K. (2017). Epidemic network analysis for mitigation of invasive pathogens in seed systems: Potato in ecuador. *Phytopathology*, 107(10):1209–1218.
- Burrage, S. W. (1970). Environmental factors influencing the infection of wheat by *Puccinia graminis*. *Annals of Applied Biology*, 66(3):429–440.
- Calbó, J., Pages, D., and González, J.-A. (2005). Empirical studies of cloud effects on uv radiation: A review. *Reviews of Geophysics*, 43(2).
- Carefoot, G. L., Sprott, E. R., et al. (1969). Famine on the wind. plant disease and human history. *Famine on the wind. Plant disease and human history*.
- Carson, R. (1962). *Silent spring*. Penguin Books.
- Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K. A., Giovani, B., Saunders, D., Kamoun, S., Legg, J., Verdier, V., Lessel, J., et al. (2019). A global surveillance system for crop diseases. *Science*, 364(6447):1237–1239.
- Caubel, J., García de Cortázar-Atauri, I., Launay, M., de Noblet-Ducoudré, N., Huard, F. F., Bertuzzi, P. P., and Graux, A.-I. (2015). Broadening the scope for ecoclimatic indicators to assess crop climate suitability according to ecophysiological, technical and quality criteria. *Agricultural and Forest Meteorology*, 207:94 – 106.

- Chadès, I., Martin, T. G., Nicol, S., Burgman, M. A., Possingham, H. P., and Buckley, Y. M. (2011). General rules for managing and surveying networks of pests, diseases, and endangered species. *Proceedings of the National Academy of Sciences*, 108(20):8323–8328.
- Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore Piontti, A., Mu, K., Rossi, L., Sun, K., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science*, 368(6489):395–400.
- Choufany, M., Martinetti, D., Senoussi, R., Morris, C. E., and Soubeyrand, S. (2021a). Spatiotemporal large-scale networks shaped by air mass movements. *Frontiers in Applied Mathematics and Statistics*, 6:67.
- Choufany, M., Martinetti, D., Soubeyrand, S., and Morris, C. E. (2021b). Inferring long-distance connectivity shaped by air-mass movement for improved experimental design in aerobiology. *Scientific Reports*, 11(1):1–10.
- Ciaperoni, M., Galimberti, E., Bonchi, F., Cattuto, C., Gullo, F., and Barrat, A. (2020). Relevance of temporal cores for epidemic spread in temporal networks. *Scientific reports*, 10(1):1–15.
- Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Physical review E*, 70(6):066111.
- Clive, J. (2015). *20th anniversary (1996 to 2015) of the global commercialization of biotech crops and biotech crop highlights in 2015*. International Service for the Acquisition of Agri-Biotech Applications.
- Colizza, V., Flammini, A., Serrano, M. A., and Vespignani, A. (2006). Detecting rich-club ordering in complex networks. *Nature physics*, 2(2):110–115.
- Corredor-Moreno, P. and Saunders, D. G. (2020). Expecting the unexpected: factors influencing the emergence of fungal and oomycete plant pathogens. *New Phytologist*, 225(1):118–125.
- Cretin, L., Triquenot, A., and Espinosa, M. (2018). Apports de produits phytopharmaceutiques en arboriculture : nombre de traitements et indicateur de fréquence de traitements. Technical report, Agreste.
- Csilléry, K., Blum, M. G., Gaggiotti, O. E., and François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution*, 25(7):410–418.
- Cunniffe, N. J., Cobb, R. C., Meentemeyer, R. K., Rizzo, D. M., and Gilligan, C. A. (2016). Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in california. *Proceedings of the National Academy of Sciences*, 113(20):5640–5645.

- Cunniffe, N. J., Stutt, R. O., DeSimone, R. E., Gottwald, T. R., and Gilligan, C. A. (2015). Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. *PLoS computational biology*, 11(4):e1004211.
- Damialis, A., Kaimakamis, E., Konoglou, M., Akritidis, I., Traidl-Hoffmann, C., and Gioulekas, D. (2017). Estimating the abundance of airborne pollen and fungal spores at variable elevations using an aircraft: how high can they fly? *Scientific reports*, 7(1):1–11.
- Damialis, A., Mohammad, A. B., Halley, J. M., and Gange, A. C. (2015). Fungi in a changing world: growth rates will be elevated, but spore production may decrease in future climates. *International Journal of Biometeorology*, 59(9):1157–1167.
- Damschen, E. I., Brudvig, L. A., Haddad, N. M., Levey, D. J., Orrock, J. L., and Tewksbury, J. J. (2008). The movement ecology and dynamics of plant communities in fragmented landscapes. *Proceedings of the National Academy of Sciences*, 105(49):19078–19083.
- De Arruda, G. F., Barbieri, A. L., Rodríguez, P. M., Rodrigues, F. A., Moreno, Y., and Costa, L. D. F. (2014). Role of centrality for the identification of influential spreaders in complex networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 90(3):1–17.
- de la Fuente, B., Saura, S., and Beck, P. S. (2018). Predicting the spread of an invasive tree pest: The pine wood nematode in Southern Europe. *Journal of Applied Ecology*, 55(5):2374–2385.
- Delannoy, D., Maury, O., and Décome, J. (2022). CLIMATIK : système d’information pour les données du réseau agroclimatique INRAE.
- Della Rossa, F., Salzano, D., Di Meglio, A., De Lellis, F., Coraggio, M., Calabrese, C., Guarino, A., Cardona-Rivera, R., De Lellis, P., Liuzza, D., et al. (2020). A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nature communications*, 11(1):1–9.
- Dhama, K., Khan, S., Tiwari, R., Sircar, S., Bhat, S., Malik, Y. S., Singh, K. P., Chaicumpa, W., Bonilla-Aldana, D. K., and Rodriguez-Morales, A. J. (2020). Coronavirus disease 2019–COVID-19. *Clinical microbiology reviews*, 33(4):e00028–20.
- Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I., Salles, J.-M., Bradshaw, C. J., and Courchamp, F. (2021). High and rising economic costs of biological invasions worldwide. *Nature*, 592(7855):571–576.
- Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature*, 418(6898):700–707.
- Diamond, J. and Bellwood, P. (2003). Farmers and their languages: the first expansions. *science*, 300(5619):597–603.

- Diamond, J. M. and Ordunio, D. (1999). *Guns, germs, and steel*, volume 521. Books on Tape New York.
- Dianati, N. (2016). Unwinding the hairball graph: Pruning algorithms for weighted complex networks. *Physical Review E*, 93(1):012304.
- Dietz, K. (1993). The estimation of the basic reproduction number for infectious diseases. *Statistical methods in medical research*, 2(1):23–41.
- Draxler, R. R. and Hess, G. D. (1998). An overview of the HYSPLIT 4 modelling system for trajectories, dispersion and deposition. *Australian Meteorological Magazine*, 47(4):295–308.
- Dubé, C., Ribble, C., Kelton, D., McNab, B., et al. (2011). Introduction to network analysis and its implications for animal disease modelling. *Revue Scientifique et Technique-OIE*, 30(2):425.
- Elderfield, J. A., Lopez-Ruiz, F. J., van den Bosch, F., and Cunniffe, N. J. (2018). Using epidemiological principles to explain fungicide resistance management tactics: Why do mixtures outperform alternations? *Phytopathology*, 108(7):803–817.
- Ellis, w. (1741). *The Timber Tree Improved*. London: D. Evans.
- Ellwanger, J. H. and Chies, J. A. B. (2021). Zoonotic spillover: Understanding basic aspects for better prevention. *Genetics and Molecular Biology*, 44.
- Elton, C. S. (1927). *Animal ecology*. University of Chicago Press.
- Elton, C. S. (1958). *The ecology of invasions by animals and plants*. John Wiley: New York, NY, USA.
- Emerson, E. W., Hodshire, A. L., DeBolt, H. M., Bilsback, K. R., Pierce, J. R., McMeeking, G. R., and Farmer, D. K. (2020). Revisiting particle dry deposition and its role in radiative effect estimates. *Proceedings of the National Academy of Sciences*, 117(42):26076–26082.
- EPPO (2023). *Monilinia fructicola*. eppo datasheets on pests recommended for regulation. <https://gd.eppo.int>. Accessed: 2023-05-09.
- Eriksson, L., Boberg, J., Cech, T. L., Corcobado, T., Desprez-Loustau, M.-L., Hietala, A. M., Jung, M. H., Jung, T., Lehtijarvi, H. T. D., Oskay, F., et al. (2019). Invasive forest pathogens in europe: Cross-country variation in public awareness but consistency in policy acceptability. *Ambio*, 48(1):1–12.
- Eurostat (2021). Peach and apricot trees - area by density classes and group of varieties (area in ha). [shorturl.at/afNVY](https://shorturl.at/afNVY). Accessed: 2022-07-07.
- Eversmeyer, M. and Kramer, C. (2000). Epidemiology of wheat leaf and stem rust in the central great plains of the usa. *Annual review of Phytopathology*, 38(1):491–513.

- Fabre, F., Coville, J., and Cunniffe, N. J. (2021). Optimising reactive disease management using spatially explicit models at the landscape scale. In *Plant Diseases and Food Security in the 21st Century*, pages 47–72. Springer.
- Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E*, 76(2):026107.
- FAO (2021a). Fao - country brief. <http://www.fao.org/giews/countrybrief/>. Accessed: 2021-07-20.
- FAO (2021b). World food and agriculture—statistical yearbook 2021. *World Food and Agriculture-Statistical Yearbook*.
- FAO (2026). "crops". [http://www.fao.org/faostat/en/#rankings/countries\\_by\\_commodity](http://www.fao.org/faostat/en/#rankings/countries_by_commodity). Retrieved 2018-08-12. Countries - Select All; Regions - World + (Total); Elements - Production Quantity; Items - Wheat; Years - 2016.
- Fielding, A. H. and Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1):38–49.
- Fine, P. E. (1993). Herd immunity: history, theory, practice. *Epidemiologic reviews*, 15(2):265–302.
- Finger, R., Swinton, S. M., El Benni, N., and Walter, A. (2019). Precision farming at the nexus of agricultural production and the environment. *Annual Review of Resource Economics*, 11:313–335.
- Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O’Connell, C., Ray, D. K., West, P. C., et al. (2011). Solutions for a cultivated planet. *Nature*, 478(7369):337–342.
- Foyo-Moreno, I., Vida, J., and Alados-Arboledas, L. (1999). A simple all weather model to estimate ultraviolet solar radiation (290–385 nm). *Journal of Applied Meteorology*, 38(7):1020–1026.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.
- Galimberti, E., Barrat, A., Bonchi, F., Cattuto, C., and Gullo, F. (2018). Mining (maximal) span-cores from temporal networks. In *Proceedings of the 27th ACM international Conference on Information and Knowledge Management*, pages 107–116.
- Garey, M. R. and Johnson, D. S. (1979). *Computers and intractability*, volume 174. freeman San Francisco.

- Garrett, K., Alcalá-Briseño, R., Andersen, K., Buddenhagen, C., Choudhury, R., Fulton, J., Hernandez Nopsa, J., Poudel, R., and Xing, Y. (2018). Network analysis: A systems framework to address grand challenges in plant pathology. *Annual review of phytopathology*, 56:559–580.
- Gilligan, C. A. (2008). Sustainable agriculture and plant diseases: an epidemiological perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1492):741–759.
- Godging, D., Stutt, R. O. J. H., Alicai, T., Okao-okuja, G., and Gilligan, C. A. (2022). Developing a predictive model for an emerging epidemic on cassava in sub-Saharan Africa. *bioRxiv*, pages 1–24.
- Gotelli, N. J., Ellison, A. M., et al. (2004). *A primer of ecological statistics*, volume 1. Sinauer Associates Sunderland.
- Gottwald, T., Poole, G., McCollum, T., Hall, D., Hartung, J., Bai, J., Luo, W., Posny, D., Duan, Y.-P., Taylor, E., et al. (2020). Canine olfactory detection of a vectored phyto-bacterial pathogen, *liberibacter asiaticus*, and integration with disease control. *Proceedings of the National Academy of Sciences*, 117(7):3492–3501.
- Harari, Y. N. (2014). *Sapiens: a brief history of humankind* by Yuval Noah Harari. *The Guardian*.
- Hartlib, S. (1655). *Samuel Hartlib, his Legacy of Husbandry*. London: Richard Wodnothe.
- Harwood, T. D., Xu, X., Pautasso, M., Jeger, M. J., and Shaw, M. W. (2009). Epidemiological risk assessment using linked network and grid based modelling: *Phytophthora ramorum* and *Phytophthora kernoviae* in the UK. *Ecological Modelling*, 220(23):3353–3361.
- Hernandez Nopsa, J. F., Daglish, G. J., Hagstrum, D. W., Leslie, J. F., Phillips, T. W., Scoglio, C., Thomas-Sharma, S., Walter, G. H., and Garrett, K. A. (2015). Ecological networks in stored grain: Key postharvest nodes for emerging pests, pathogens, and mycotoxins. *BioScience*, 65(10):985–1002.
- Herrera, J. L., Srinivasan, R., Brownstein, J. S., Galvani, A. P., and Meyers, L. A. (2016). Disease surveillance on complex social networks. *PLoS computational biology*, 12(7):e1004928.
- Hinchliffe, S. (2017). More than one world, more than one health: Re-configuring inter-species health. In *Global health and geographical imaginaries*, pages 159–175. Routledge.
- Hobbelen, P., Paveley, N., Fraaije, B., Lucas, J., and Van den Bosch, F. (2011). Derivation and testing of a model to predict selection for fungicide resistance. *Plant Pathology*, 60(2):304–313.

- Holb, I. J. (2008). Monitoring conidial density of *Monilinia fructigena* in the air in relation to brown rot development in integrated and organic apple orchards. *European Journal of Plant Pathology*, 120(4):397–408.
- Holme, P. (2005). Network reachability of real-world contact sequences. *Physical Review E*, 71(4):046119.
- Holme, P. (2017). Three faces of node importance in network epidemiology: Exact results for small graphs. *Physical Review E*, 96(6):062305.
- Holme, P. (2018). Objective measures for sentinel surveillance in network epidemiology. *Physical Review E*, 98(2):022313.
- Hovmøller, M. S., Yahyaoui, A. H., Milus, E. A., and Justesen, A. F. (2008). Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular ecology*, 17(17):3818–3826.
- Hrustić, J., Mihajlović, M., Grahovac, M., Delibašić, G., Bulajić, A., Krstić, B., and Tanović, B. (2012). Genus *Monilinia* on pome and stone fruit species. *Pesticidi i fitomedicina*, 27(4):283–297.
- Huerta-Espino, J., Singh, R., and Roelfs, A. P. (2014). Rusts fungi of wheat. *Fungi from different substrates*, pages 217–259.
- Hulme, P. E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of applied ecology*, 46(1):10–18.
- Hyatt-Twynam, S. R., Parnell, S., Stutt, R. O., Gottwald, T. R., Gilligan, C. A., and Cunniffe, N. J. (2017). Risk-based management of invading plant disease. *New Phytologist*, 214(3):1317–1329.
- International Food Policy Research Institute (2019). Global spatially-disaggregated crop production statistics data for 2010, version 2.0.
- Isard, S. A., Gage, S. H., Comtois, P., and Russo, J. M. (2005). Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience*, 55(10):851–861.
- Jeger, M. J., Pautasso, M., Holdenrieder, O., and Shaw, M. W. (2007). Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist*, 174(2):279–297.
- John, C. and Allan, H. D. (1995). *A first look at graph theory*. Allied Publishers.
- Johnson, K. M. and Lichter, D. T. (2019). Rural depopulation: Growth and decline processes over the past century. *Rural Sociology*, 84(1):3–27.

- Jones, A., Thomson, D., Hort, M., and Devenish, B. (2007). *Air pollution modeling and its application XVII*. Borrego, C. & Norman, A.-L.
- Jordano, P. (2017). What is long-distance dispersal? and a taxonomy of dispersal events. *Journal of Ecology*, 105(1):75–84.
- Juroszek, P., Racca, P., Link, S., Farhumand, J., and Kleinhenz, B. (2020). Overview on the review articles published during the past 30 years relating to the potential climate change effects on plant pathogens and crop disease risks. *Plant pathology*, 69(2):179–193.
- Kao, R. R., Green, D. M., Johnson, J., and Kiss, I. Z. (2007). Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the uk. *Journal of the Royal Society Interface*, 4(16):907–916.
- Keeling, M. J. and Gilligan, C. A. (2000). Metapopulation dynamics of bubonic plague. *Nature*, 407(6806):903–906.
- Keeling, M. J. and Rohani, P. (2011). *Modeling infectious diseases in humans and animals*. Princeton university press.
- Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721.
- Khoury, C. K., Bjorkman, A. D., Dempewolf, H., Ramirez-Villegas, J., Guarino, L., Jarvis, A., Rieseberg, L. H., and Struik, P. C. (2014). Increasing homogeneity in global food supplies and the implications for food security. *Proceedings of the national Academy of Sciences*, 111(11):4001–4006.
- Kislev, M. E. (1982). Stem rust of wheat 3300 years old found in israel. *Science*, 216(4549):993–994.
- Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., and Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature physics*, 6(11):888–893.
- Kling, M. M. and Ackerly, D. D. (2021). Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees. *Proceedings of the National Academy of Sciences*, 118(17):e2017317118.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago press.
- Kulik, A. and Shachnai, H. (2010). There is no optimal for two-dimensional knapsack. *Information Processing Letters*, 110(16):707–710.



- Large, E. C. (1954). Growth Stages in Cereals Illustration of the Feekes Scale. *Plant pathology*, 62(1):207.
- Lee, B. X., Kjaerulf, F., Turner, S., Cohen, L., Donnelly, P. D., Muggah, R., Davis, R., Realini, A., Kieselbach, B., MacGregor, L. S., et al. (2016). Transforming our world: implementing the 2030 agenda through sustainable development goal indicators. *Journal of public health policy*, 37(1):13–31.
- Levetin, E. (2015). Aerobiology of Agricultural Pathogens. *Manual of Environmental Microbiology*, pages 3.2.8–1–3.2.8–20.
- Li, X., Yang, X., Mo, J., and Guo, T. (2009). Estimation of soybean rust uredospore terminal velocity, dry deposition, and the wet deposition associated with rainfall. *European journal of plant pathology*, 123(4):377–386.
- Lichou, J., Mandrin, J.-F., Breniaux, D., Mercier, V., Giauque, P., Desbrus, D., Blanc, P., and Belluau, E. (2002). Une nouvelle moniliose. *monilia fructicola* s' attaque aux arbres fruitiers à noyaux. *PHYTOMA La Défense des Végétaux*, (547):22–25.
- Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., and Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359.
- Lü, L., Chen, D., Ren, X.-L., Zhang, Q.-M., Zhang, Y.-C., and Zhou, T. (2016). Vital nodes identification in complex networks. *Physics reports*, 650:1–63.
- Lu, N., Wang, J., Chen, X., Zhan, G., Chen, C., Huang, L., and Kang, Z. (2011). Spatial genetic diversity and interregional spread of *puccinia striiformis* f. sp. *tritici* in northwest china. *European journal of plant pathology*, 131(4):685–693.
- Lucarelli, R. (2017). Illness as divine punishment: the nature and function of the disease-carrier demons in the ancient egyptian magical texts. In *Demons and Illness from Antiquity to the Early-Modern Period*, pages 53–60. Brill.
- Lucas, J. A., Hawkins, N. J., and Fraaije, B. A. (2015). The evolution of fungicide resistance. *Advances in applied microbiology*, 90:29–92.
- Maddison, A. and Manners, J. (1972). Sunlight and viability of cereal rust uredospores. *Transactions of the British Mycological Society*, 59(3):429–443.
- Mahaffee, W. F. and Stoll, R. (2016). The ebb and flow of airborne pathogens: Monitoring and use in disease management decisions. *Phytopathology*, 106(5):420–431.
- Margosian, M. L., Garrett, K. A., Hutchinson, J. S., and With, K. A. (2009). Connectivity of the american agricultural landscape: assessing the national risk of crop pest and disease spread. *BioScience*, 59(2):141–151.

- Martelli, G., Boscia, D., Porcelli, F., and Saponari, M. (2016). The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. *European Journal of Plant Pathology*, 144(2):235–243.
- Martinetti, D. and Soubeyrand, S. (2019). Identifying lookouts for epidemio-surveillance: application to the emergence of *Xylella fastidiosa* in France. *Phytopathology*, 109(2):265–276.
- Martini, C. and Mari, M. (2014). *Monilinia fructicola*, *Monilinia laxa* (*Monilinia Rot*, *Brown Rot*). Elsevier.
- Marzano, M., Allen, W., Haight, R., Holmes, T., Keskitalo, E. C. H., Langer, E., Shadbolt, M., Urquhart, J., and Dandy, N. (2017). The role of the social sciences and economics in understanding and informing tree biosecurity policy and planning: a global summary and synthesis. *Biological Invasions*, 19(11):3317–3332.
- Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. (2021). *IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, volume 1. Cambridge University Press.
- Mastin, A. J., Gottwald, T. R., van den Bosch, F., Cunniffe, N. J., and Parnell, S. (2020). Optimising risk-based surveillance for early detection of invasive plant pathogens. *PLoS biology*, 18(10):e3000863.
- Mayol, E., Arrieta, J. M., Jiménez, M. A., Martínez-Asensio, A., Garcías-Bonet, N., Dachs, J., González-Gaya, B., Royer, S.-J., Benítez-Barrios, V. M., Fraile-Nuez, E., et al. (2017). Long-range transport of airborne microbes over the global tropical and subtropical ocean. *Nature Communications*, 8(1):1–9.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.
- McLeod, S. (2007). Maslow's hierarchy of needs. *Simply psychology*, 1(1-18).
- McMaster, G. S. and Smika, D. E. (1988). Estimation and Evaluation of Winter Wheat Phenology in the central great plains. *Agricultural and Forest Meteorology*, 43:1–18.
- McMaster, G. S. and Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, 87(1).
- Meentemeyer, R. K., Cunniffe, N. J., Cook, A. R., Filipe, J. A., Hunter, R. D., Rizzo, D. M., and Gilligan, C. A. (2011). Epidemiological modeling of invasion in heterogeneous landscapes: spread of sudden oak death in California (1990–2030). *Ecosphere*, 2(2):1–24.

- Mehta, S. V., Haight, R. G., Homans, F. R., Polasky, S., and Venette, R. C. (2007). Optimal detection and control strategies for invasive species management. *Ecological Economics*, 61(2-3):237–245.
- Meyer, M., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017a). Large-scale atmospheric dispersal simulations identify likely airborne incursion routes of wheat stem rust into Ethiopia. *Phytopathology*, 107(10):1175–1186.
- Meyer, M., Cox, J. A., Hitchings, M. D., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017b). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nature Plants*, 3(10):780–786.
- Michael, P. R., Johnston, D. E., and Moreno, W. (2020). A conversion guide: solar irradiance and lux illuminance. *Journal of Measurements in Engineering*, 8(4):153–166.
- Ministère de l’Agriculture (2010). Recensement général de l’agriculture. <https://agreste.agriculture.gouv.fr/agreste-web/disaron/Pri2105/detail/>. Accessed: 2021-03-01.
- Minter, A. and Retkute, R. (2019). Approximate Bayesian Computation for infectious disease modelling. *Epidemics*, 29(February):100368.
- Mohamed, K., Rodríguez-Román, E., Rahmani, F., Zhang, H., Ivanovska, M., Makka, S. A., Joya, M., Makuku, R., Islam, M. S., Radwan, N., et al. (2020). Borderless collaboration is needed for covid-19—a disease that knows no borders. *Infection Control & Hospital Epidemiology*, 41(10):1245–1246.
- Morris, C. E., Geniaux, G., Nédellec, C., Sauvion, N., and Soubeyrand, S. (2022). One health concepts and challenges for surveillance, forecasting and mitigation of plant disease beyond the traditional scope of crop production. *Plant pathology*.
- Morris, C. E., Sands, D. C., Glaux, C., Samsatly, J., Asaad, S., Moukahel, A. R., Gonçalves, F. L., and Bigg, E. K. (2013). Urediospores of rust fungi are ice nucleation active at >-10 °c and harbor ice nucleation active bacteria. *Atmospheric Chemistry and Physics*, 13(8):4223–4233.
- Moslonka-Lefebvre, M., Finley, A., Dorigatti, I., Dehnen-Schmutz, K., Harwood, T., Jeger, M. J., Xu, X., Holdenrieder, O., and Pautasso, M. (2011). Networks in plant epidemiology: from genes to landscapes, countries, and continents. *Phytopathology*, 101(4):392–403.
- Nagarajan, S. and Singh, D. (1990). Long-distance dispersion of rust pathogens. *Annual review of phytopathology*, 28(1):139–153.
- Neufeld, K., AP, K., Gugino, B., McGrath, M., Sikora, E., Miller, S., Ivey, M., Langston, D., Dutta, B., Keever, T., Sims, A., and Ojiambo, P. (2018). Predicting the risk of cucurbit downy mildew in the eastern United States using an integrated aerobiological model. *International Journal of Biometeorology*, 62(4):655–668.

- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.
- Noble, J. V. (1974). Geographic and temporal development of plagues. *Nature*, 250(5469):726–729.
- Nsoesie, E. O., Rader, B., Barnoon, Y. L., Goodwin, L., and Brownstein, J. (2020). Analysis of hospital traffic and search engine data in wuhan china indicates early disease activity in the fall of 2019. Havard Library Office for Scholarly Communication.
- Och, A. and Nowak, R. (2021). Barberry (*berberis vulgaris*)—traditional and contemporary use. *Medicinal Plants: Domestication, Biotechnology and Regional Importance*, pages 797–825.
- Olitaa, H. T., Sunga, B., Hoopera, B., Cacao, Z., Lopez-Ruiza, F., and Gibberda, M. (2023). The socio-economic impact of fungicide resistance in west australia’s wheatbelt. *Advances in Agronomy*, 180:1.
- Oliveira Lino, L., Pacheco, I., Mercier, V., Faoro, F., Bassi, D., Bornard, I., and Quilot-Turion, B. (2016). Brown Rot Strikes Prunus Fruit: An Ancient Fight Almost Always Lost. *Journal of Agricultural and Food Chemistry*, 64(20):4029–4047.
- Olivera, P., Newcomb, M., Szabo, L. J., Rouse, M., Johnson, J., Gale, S., Luster, D. G., Hodson, D., Cox, J. A., Burgin, L., et al. (2015). Phenotypic and genotypic characterization of race TKTTF of *Puccinia graminis* f. sp. *tritici* that caused a wheat stem rust epidemic in southern Ethiopia in 2013-14. *Phytopathology*, 105(7):917–928.
- Olivera, P. D., Sikharulidze, Z., Dumbadze, R., Szabo, L. J., Newcomb, M., Natsarishvili, K., Rouse, M. N., Luster, D. G., and Jin, Y. (2019). Presence of a sexual population of *puccinia graminis* f. sp. *tritici* in georgia provides a hotspot for genotypic and phenotypic diversity. *Phytopathology*, 109(12):2152–2160.
- Olivera Firpo, P., Newcomb, M., Flath, K., Sommerfeldt-Impe, N., Szabo, L., Carter, M., Luster, D., and Jin, Y. (2017). Characterization of *puccinia graminis* f. sp. *tritici* isolates derived from an unusual wheat stem rust outbreak in germany in 2013. *Plant Pathology*, 66(8):1258–1266.
- Oneto, D. L., Golan, J., Mazzino, A., Pringle, A., and Seminara, A. (2020). Timing of fungal spore release dictates survival during atmospheric transport. *Proceedings of the National Academy of Sciences of the United States of America*, 117(10):5134–5143.
- Page, L. (1998). The pagerank citation ranking: Bringing order to the web. technical report. *Stanford Digital Library Technologies Project, 1998*.
- Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.

- Palmer, A. W. (1996). *Dictionary of the British Empire and Commonwealth*. John Murray.
- Park, R., Fetch, T., Hodson, D., Jin, Y., Nazari, K., Prashar, M., and Pretorius, Z. (2011). International surveillance of wheat rust pathogens: Progress and challenges. *Euphytica*, 179(1):109–117.
- Parnell, S., van den Bosch, F., Gottwald, T., and Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: an epidemiological perspective. *Annual review of phytopathology*, 55:591–610.
- Pasteur, L. (1861). *Recherches sur la dissymétrie moléculaire des produits organiques naturels*. Livingstone.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of modern physics*, 87(3):925.
- Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200.
- Pastor-Satorras, R. and Vespignani, A. (2002). Immunization of complex networks. *Physical review E*, 65(3):036104.
- Peel, M. C., Finlayson, B. L., and McMahon, T. A. (2007). Updated world map of the köppen-geiger climate classification. *Hydrology and earth system sciences*, 11(5):1633–1644.
- Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., and Trapman, P. (2015). Eight challenges for network epidemic models. *Epidemics*, 10:58–62.
- Peterson, P. D. et al. (2001). *Stem rust of wheat: from ancient enemy to modern foe*. American Phytopathological Society (APS Press).
- Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. *Proceedings of the national academy of sciences*, 109(31):12302–12308.
- Pingali, P. L. (2017). The green revolution and crop biodiversity. In *Routledge Handbook of Agricultural Biodiversity*, pages 213–223. Routledge.
- Power, A. G. (2010). Ecosystem services and agriculture: tradeoffs and synergies. *Philosophical transactions of the royal society B: biological sciences*, 365(1554):2959–2971.
- Prank, M., Kenaley, S. C., Bergstrom, G. C., Acevedo, M., and Mahowald, N. M. (2019). Climate change impacts the spread potential of wheat stem rust, a significant crop disease. *Environmental Research Letters*, 14(12).
- Pretorius, Z., Singh, R., Wagoire, W., and Payne, T. (2000). Detection of virulence to wheat stem rust resistance gene sr31 in puccinia graminis. f. sp. tritici in uganda. *Plant disease*, 84(2):203–203.

- QGIS Development Team (2022). *QGIS Geographic Information System*. QGIS Association.
- Radici, A. (2023). Brown rot of peach severity data. <https://www.data.gouv.fr/fr/datasets/brown-rot-of-peach-severity-data/>. Accessed: 2023-02-01.
- Radici, A., Bevacqua, D., Miele, L., and Martinetti, D. (2023a). Network-thinking to optimize surveillance and control of crop parasites. a review. *arXiv preprint arXiv:2310.07442*.
- Radici, A., Martinetti, D., and Bevacqua, D. (2022). Early-detection surveillance for stem rust of wheat: insights from a global epidemic network based on airborne connectivity and host phenology. *Environmental Research Letters*.
- Radici, A., Martinetti, D., and Bevacqua, D. (2023b). Global benefits and domestic costs of a cooperative surveillance strategy to control transboundary crop pathogens. *Plants, People, Planet*, (April):1–10.
- Radici, A., Martinetti, D., Vanalli, C., Cunniffe, N., and Bevacqua, D. (2023c). A metapopulation framework integrating landscape heterogeneity to model an airborne plant pathogen: the case of brown rot of peach in france. *bioRxiv*, pages 2023–10.
- Rimbaud, L., Papaïx, J., Rey, J.-F., Barrett, L. G., and Thrall, P. H. (2018a). Assessing the durability and efficiency of landscape-based strategies to deploy plant resistance to pathogens. *PLoS computational biology*, 14(4):e1006067.
- Rimbaud, L., Papaïx, J., Rey, J.-F., Barrett, L. G., and Thrall, P. H. (2018b). Assessing the durability and efficiency of landscape-based strategies to deploy plant resistance to pathogens. *PLOS Computational Biology*, 14(4):1–33.
- Ringius, L., Torvanger, A., and Underdal, A. (2002). Burden sharing and fairness principles in international climate policy. *International Environmental Agreements*, 2(1):1–22.
- Ristaino, J. B., Anderson, P. K., Bebbler, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., et al. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences*, 118(23).
- Ritchie, H. and Roser, M. (2021). Farm size. *Our World in Data*.
- Roelfs, A. P. (1992). *Rust diseases of wheat: concepts and methods of disease management*. Cimmyt.
- Rowell, J., Romig, R., et al. (1966). Detection of urediospores of wheat rusts in spring rains. *Phytopathology*, 56(7):807–811.
- Rushmore, J., Caillaud, D., Hall, R. J., Stumpf, R. M., Meyers, L. A., and Altizer, S. (2014). Network-based vaccination improves prospects for disease control in wild chimpanzees. *Journal of the Royal Society Interface*, 11(97):20140349.

- Russo, L. (2001). *La rivoluzione dimenticata: il pensiero scientifico greco e la scienza moderna*. Feltrinelli Editore.
- Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4):581–603.
- Sanatkar, M., Scoglio, C., Natarajan, B., Isard, S., and Garrett, K. (2015). History, epidemic evolution, and model burn-in for a network of annual invasion: Soybean rust. *Phytopathology*, 105(7):947–955.
- Sandler, T. and Forbes, J. F. (1980). Burden sharing, strategy, and the design of nato. *Economic inquiry*, 18(3):425–444.
- Santini, A., Ghelardini, L., De Pace, C., Desprez-Loustau, M.-L., Capretti, P., Chandelier, A., Cech, T., Chira, D., Diamandis, S., Gaitniekis, T., et al. (2013). Biogeographical patterns and determinants of invasion by forest pathogens in europe. *New Phytologist*, 197(1):238–250.
- Santini, A., Liebhold, A., Migliorini, D., and Woodward, S. (2018). Tracing the role of human civilization in the globalization of plant pathogens. *The ISME journal*, 12(3):647–652.
- Saunders, D. G., Pretorius, Z. A., and Hovmøller, M. S. (2019). Tackling the re-emergence of wheat stem rust in western europe. *Communications biology*, 2(1):51.
- Saunders, H. A. and Schwartz, J.-M. (2021). Covid-19 vaccination strategies depend on the underlying network of social interactions. *Scientific reports*, 11(1):1–10.
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology & evolution*, 3(3):430–439.
- Schaal, B. (2019). Plants and people: Our shared history and future. *Plants, People, Planet*, 1(1):14–19.
- Schebesta, H. and Candel, J. J. (2020). Game-changing potential of the eu’s farm to fork strategy. *Nature Food*, 1(10):586–588.
- Schmale III, D. G. and Ross, S. D. (2015). Highways in the sky: Scales of atmospheric transport of plant pathogens. *Annual review of phytopathology*, 53:591–611.
- Schneider, R., Hollier, C., Whitam, H., Palm, M., McKemy, J., Hernandez, J., Levy, L., and DeVries-Paterson, R. (2005). First report of soybean rust caused by *phakopsora pachyrhizi* in the continental united states. *Plant disease*, 89(7):774–774.
- Serrano, M. Á., Boguná, M., and Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. *Proceedings of the national academy of sciences*, 106(16):6483–6488.

- Seufert, V., Ramankutty, N., and Mayerhofer, T. (2017). What is this thing called organic?—how organic farming is codified in regulations. *Food Policy*, 68:10–20.
- Shaw, M. and Pautasso, M. (2014). Networks and plant disease management: Concepts and applications. *Annual Review of Phytopathology*, 52:477–493.
- Sheldrake, M. (2021). *Entangled life: How fungi make our worlds, change our minds & shape our futures*. Random House Trade Paperbacks.
- Singh, R. P., Hodson, D. P., Jin, Y., Lagudah, E. S., Ayliffe, M. A., Bhavani, S., Rouse, M. N., Pretorius, Z. A., Szabo, L. J., Huerta-Espino, J., et al. (2015). Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. *Phytopathology*, 105(7):872–884.
- Slinn, W. (1977). Some approximations for the wet and dry removal of particles and gases from the atmosphere. *Water, Air, and Soil Pollution*, 7(4):513–543.
- Soubeyrand, S., Demongeot, J., and Roques, L. (2020). Towards unified and real-time analyses of outbreaks at country-level during pandemics. *One Health*, 11:100187.
- Strona, G., Carstens, C. J., and Beck, P. S. (2017). Network analysis reveals why xylella fastidiosa will persist in europe. *Scientific Reports*, 7(1):1–8.
- Strona, G., Castellano, C., Fattorini, S., Ponti, L., Gutierrez, A. P., and Beck, P. S. (2020). Small world in the real world: Long distance dispersal governs epidemic dynamics in agricultural landscapes. *Epidemics*, 30:100384.
- Stuart, H. (2013). *Essential microbiology*. John Wiley & Sons.
- Stukenbrock, E. H. and McDonald, B. A. (2008). The origins of plant pathogens in agroecosystems. *Annu. Rev. Phytopathol.*, 46:75–100.
- Suhrke, A. (1998). Burden-sharing during refugee emergencies: The logic of collective versus national action. *Journal of refugee studies*, 11(4):396–415.
- Sumberg, J. and Giller, K. E. (2022). What is ‘conventional’agriculture? *Global Food Security*, 32:100617.
- Sutrave, S., Scoglio, C., Isard, S. A., Hutchinson, J. S., and Garrett, K. A. (2012). Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS One*, 7(6):e37793.
- Thompson, R. N., Cobb, R. C., Gilligan, C. A., and Cunniffe, N. J. (2016). Management of invading pathogens should be informed by epidemiology rather than administrative boundaries. *Ecological Modelling*, 324:28–32.



- Thompson, R. N., Hollingsworth, T. D., Isham, V., Arribas-Bel, D., Ashby, B., Britton, T., Challenor, P., Chappell, L. H., Clapham, H., Cunniffe, N. J., et al. (2020). Key questions for modelling covid-19 exit strategies. *Proceedings of the Royal Society B*, 287(1932):20201405.
- Thrall, P. H. and Burdon, J. J. (2003). Evolution of virulence in a plant host-pathogen metapopulation. *Science*, 299(5613):1735–1737.
- Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the national academy of sciences*, 108(50):20260–20264.
- Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418(6898):671–677.
- Travençolo, B. A. N. and Costa, L. d. F. (2008). Accessibility in complex networks. *Physics Letters A*, 373(1):89–95.
- Trottier, H. and Philippe, P. (2005). Scaling properties of childhood infectious diseases epidemics before and after mass vaccination in canada. *Journal of theoretical biology*, 235(3):326–337.
- Tsushima, A., Lewis, C. M., Flath, K., Kildea, S., and Saunders, D. G. (2022). Wheat stem rust recorded for the first time in decades in ireland. *Plant Pathology*, 71(4):890–900.
- Tudi, M., Daniel Ruan, H., Wang, L., Lyu, J., Sadler, R., Connell, D., Chu, C., and Phung, D. T. (2021). Agriculture development, pesticide application and its impact on the environment. *International journal of environmental research and public health*, 18(3):1112.
- van der Eijk, P., Horstmanshoff, H. F., and Schrijvers, P. (1995). *Ancient medicine in its socio-cultural context: papers read at the congress held at Leiden University, 13-15 April 1992*, volume 2. Rodopi.
- van der Riet, F. D. S. J. (1997). Diseases of plants transmissible between plants and man (phytonoses) exist. *Medical hypotheses*, 49(4):359–361.
- van Overbeek, L. S., van Doorn, J., Wichers, J. H., van Amerongen, A., van Roermund, H. J., and Willemsen, P. T. (2014). The arable ecosystem as battleground for emergence of new human pathogens. *Frontiers in Microbiology*, 5:104.
- Vanalli, C., Casagrandi, R., Gatto, M., and Bevacqua, D. (2021). Shifts in the thermal niche of fruit trees under climate change: The case of peach cultivation in France. *Agricultural and Forest Meteorology*, 300(September 2020).
- Vavilov, N. I. and Dorofeev, V. F. (1992). *Origin and geography of cultivated plants*. Cambridge University Press.

- Vilanova, L., Valero-Jiménez, C. A., and van Kan, J. A. (2021). Deciphering the monilinia fructicola genome to discover effector genes possibly involved in virulence. *Genes*, 12(4).
- Visser, B., Meyer, M., Park, R. F., Gilligan, C. A., Burgin, L. E., Hort, M. C., Hodson, D. P., and Pretorius, Z. A. (2019). Microsatellite analysis and urediniospore dispersal simulations support the movement of puccinia graminis f. Sp. Tritici from southern Africa to Australia. *Phytopathology*, 109(1):133–144.
- Vitousek, P. M. (1994). Beyond global warming: ecology and global change. *Ecology*, 75(7):1861–1876.
- Wang, H., Hernandez, J. M., and Van Mieghem, P. (2008). Betweenness centrality in a weighted network. *Physical Review E*, 77(4):046105.
- Wang, M., Kriticos, D. J., Ota, N., Brooks, A., and Paini, D. (2021). A general trait-based modelling framework for revealing patterns of airborne fungal dispersal threats to agriculture and native flora. *New Phytologist*.
- Webb, P. and Eiselen, H. (2009). Fiat panis: For a world without hunger. *Hampp Media/Balance Publications, Stuttgart, Germany*, pages 410–434.
- Xing, Y., Hernandez Nopsa, J. F., Andersen, K. F., Andrade-Piedra, J. L., Beed, F. D., Blomme, G., Carvajal-Yepes, M., Coyne, D. L., Cuellar, W. J., Forbes, G. A., et al. (2020). Global cropland connectivity: A risk factor for invasion and saturation by emerging pathogens and pests. *BioScience*, 70(9):744–758.
- Yamamoto, N., Nazaroff, W. W., and Peccia, J. (2014). Assessing the aerodynamic diameters of taxon-specific fungal bioaerosols by quantitative PCR and next-generation DNA sequencing. *Journal of Aerosol Science*, 78:1–10.
- Zadoks, J. et al. (1985). Cereal rusts, dogs and stars in antiquity. *Cereal Rusts Bulletin*, 13(1):1–10.
- Zadoks, T. (1967). Epidemiology of wheat rust in europe. *International Journal of Pest Management B*, 13(1):29–46.
- Zeigler, R. S. and Mohanty, S. (2010). Support for international agricultural research: current status and future challenges. *New biotechnology*, 27(5):565–572.
- Zhang, J. X., Chen, D. B., Dong, Q., and Zhao, Z. D. (2016). Identifying a set of influential spreaders in complex networks. *Scientific Reports*, 6(February):1–10.
- Zilberman, D., Holland, T. G., and Trilnick, I. (2018). Agricultural gmos—what we know and where scientists disagree. *Sustainability*, 10(5):1514.

# Appendix A

## Supplementary Information - Chapter 2

### A.1 Supplementary materials

#### A.1.1 Geographical domain

The geographical domain used in this work consists of 7,814 cells covered with a surface of at least 2% of wheat (average 10%, standard deviation 8%), according to the 2010 MapSPAM database release (International Food Policy Research Institute, 2019). Physical area (cod. *A*) was used to perform this calculation.

Due to the grid extraction procedure (resolution of the grid:  $0.5^\circ \times 0.5^\circ$ ), cells size varies across latitudes, increasing in surface from the poles to the Equator. Here the cell size varies from a minimum of  $1,495 \text{ km}^2$  (Finland) to a maximum of  $3,075 \text{ km}^2$  (Kenya).

The final land surface considered in this work is  $1.77 \times 10^7 \text{ km}^2$ , approximately 12.1% of the total Earth's land surface.

#### A.1.2 Lagrangian trajectory simulation settings

HYSPLIT configuration setup is reported in table A.1.

**Simulation outputs.** The following meteorological and physical outputs have been stored for each trajectory at hourly frequency: day, hour, age, latitude, longitude, altitude, air temperature (TAMB), rainfall (RAIN), mixed layer depth (MIXD), relative humidity (RELH), terrain height (TERR), downward solar radiation flux (DSWF; Hysplit calculates meteorological variables at given positions from GDAS, <https://www.ncei.noaa.gov/products/weather-climate-models/global-data-assimilation>)

**Clouds altitude estimation.** Lagrangian simulation were run every 6 h starting at mid-

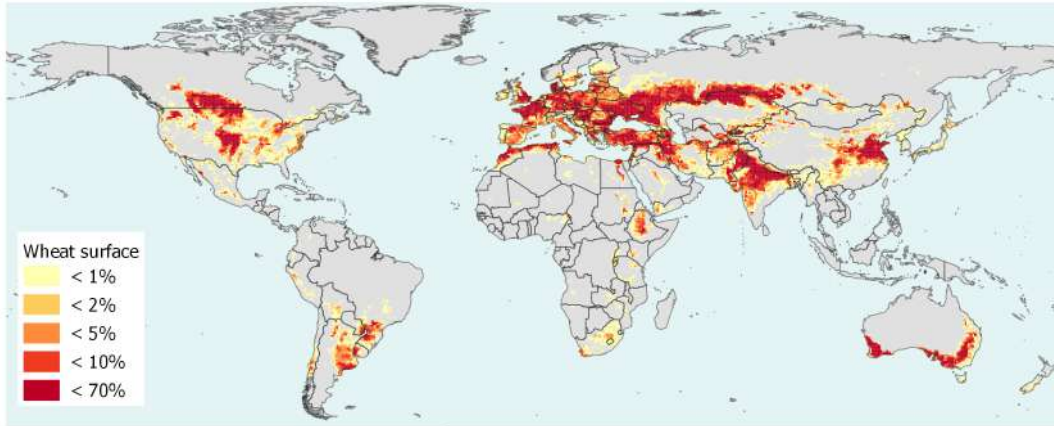


Figure A.1 – Percentage of wheat as land cover worldwide, elaborated from the SPAM database.

Parameter	Value	Meaning
tratio	0.75	advection stability ratio
delt	0 (autoset)	integration time step
mgmin	10	minimum meteorological subgrid size
khmax	9999	maximum duration (h) for a particle or trajectory
kmixd	0 (default)	Source of mixed layer calculation
kmsl	1 (AGL)	Starting heights default
kagl	1 (AGL)	trajectory output heights are written as above ground level
k10m	1	use surface 10m winds / 2m temperature
nstr	0	trajectory restart time interval in hours
mhrs	9,999	trajectory restart duration limit (hours)
nver	0	trajectory vertical split number
tout	60	trajectory output interval in minutes
dx	1	horizontal X-grid adjustment factor for ensemble
dy	1	horizontal Y-grid adjustment factor for ensemble
dz	0.009999	vertical (0.01 ~ 250m) factor for ensemble

Table A.1 – HYSPLIT configuration setup.

night from the centroids of the cells at an above-ground altitude set to the minimum between the base of the cloud and the mixed layer depth, above the mostly mixed layers of the troposphere. Clouds altitude is not provided as output of HYSPLIT simulations. However, HYSPLIT user manual explains how to estimate the base of the cloud (RHB) and the top of the clouds (RHT) from relative humidity reconstruction, typically used to calculate deposition. The base of the clouds is estimated to be at the height where the relative humidity profile exceeds 80%, while top of the clouds is the height where the relative humidity profile falls below 60%.

A specific simulation was run from each cell centroid for each release date to collect RH values from 0 up to 5,000 m above ground levels each 100 m. The value of 5,000 m was chosen since it represents the maximum distance above ground level that could be calculated in the highest point in the domain. In other words, since HYSPLIT domain altitude limit is set to 10,000 m above sea level, and the highest point in the domain is 4,899 m above sea level (77° 45' E, 31° 45' N, Himalayan mountains in the north of India), 5,000 is the maximum value that enable meteorological variables computation in each point of the domain.

**Temperature estimation at ground level.** Air temperature values are estimated along the trajectories, at the trajectory's altitude. However, air temperature at ground level is necessary since sporulation is limited between 5 and 40 °C. Temperature at ground level has been estimated assuming a typical gradient of 6.5 °C/km along the vertical trajectory, so that:

$$T(z_{ground}) = -0.0065 \times T(z_{trajectory}) \times (z_{trajectory} - z_{ground}) \quad (A.1)$$

Where  $T$  is temperature at different elevations,  $z_{trajectory}$  is the trajectory elevation and  $z_{ground}$  is the ground elevation.

### A.1.3 Farming Calendar

**Temporal host availability detection** Host availability is defined as the period of the year in which wheat has reached the phenological phase suitable for infection, in presence of suitable environmental conditions and of a pathogen. We used the growing degree-days formula (GDD) suggested by McMaster & Smika (McMaster and Smika, 1988) to reproduce the main wheat season calendar. We have chosen the formulation proposed to these authors because it allows us to identify certain physiological phases of wheat from weather variable from the first day of the year rather than from the sowing date. This guarantees the possibility of calculating the growth of wheat even in the absence of specific information relative to the sowing calendar, in particular the sowing date, which can be spatially heterogeneous and susceptible to mutation due to climate change. Here, the following phases were considered:

1. Dormancy end (DE), or phase 3.0 of the Feekes (Large, 1954) growth scale, is the last phase before leaves and pseudo stem erection. Here it is considered as the beginning of host availability.
2. Maturity (M), or phase 11.0 of the Feekes growth scale, is the last phase before harvest. It defines the last period of host availability (2 month are added at the end to account for harvest procedures, as explained below).

Using the dormancy end as the date of onset of host availability may lead to an overestimation of susceptibility because at this phase the plant has not yet fully developed the stem, which can be infected by *P. graminis*. However, if other stages had been chosen to set the beginning of host availability, they would have led to an underestimation of the susceptibility period. We chose the dormancy end because it represents a choice in favour of safety.

Phenological phase	Lower limit ( $T_b$ )	Upper limit ( $T_B$ )	GDD threshold ( $T_h$ )
Dormancy end (DE)	2	25	38
Maturity (M)	0	25	1622

Table A.2 – Parameters of the growing degree-days formula.

The Growing degree-days formula is written as follows (Mcmaster and Wilhelm, 1997):

$$GDD(D) = \sum_{d=1}^D \min(T_B, \max(\frac{T_M(d) + T_m(d)}{2} - T_b, 0)) \quad (\text{A.2})$$

Where, beside coefficients defined in table A.2,  $T_M(d)$  and  $T_m(d)$  are the minimum and the maximum daily temperature. Phenological states are identified when the GDD exceed the corresponding threshold:

$$DE = \arg \min_D \mathbf{1}_{GDD(D) > Th_{DE}}(GDD(D)) \quad (\text{A.3})$$

$$M = \arg \min_D \mathbf{1}_{GDD(D) > Th_M}(GDD(D)) \quad (\text{A.4})$$

Where  $\mathbf{1}_{[arg]}(f(\cdot))$  is the indicator function, which is equal 1 if the condition in  $[arg]$  is verified, 0 elsewhere.

Air temperature data at ground level were obtained from the Global Data Assimilation System – GDAS (<https://www.ncei.noaa.gov/products/weather-climate-models/global-data-assimilation>), whose temporal resolution is of 6 hours (00 – 06 – 12 – 18 UTC).

In countries of the southern hemisphere the computation of the cumulative degree days starts on the 1<sup>st</sup> of July, to solve the seasonal time lag between hemispheres.

**Growing season in tropical regions.** Growing season are assigned to tropical regions according to FAO (<http://www.fao.org/giews/countrybrief/>) or USDA (<https://ipad>).

fas.usda.gov/rssiws/al/crop\_calendar/us.aspx) calendars, summarized in table A.3. For instance, the longest farming season (*meher*) in Ethiopia goes from July-August to December-January (Meyer et al., 2017). As a consequence, despite being located in the Northern Hemisphere, the same regime of southern countries was adopted, and we neglected the shortest season (*belg*, June-July).

Tropical country	Main growing season	Assigned calendar season
Ethiopia	May to January	South (1 <sup>st</sup> of July)
Eritrea	June to November	South (1 <sup>st</sup> of July)
Sudan	November to March	North (1 <sup>st</sup> of January)
Kenya	July to November	South (1 <sup>st</sup> of July)
Rwanda	September to Jan	North (1 <sup>st</sup> of January)
Nigeria*	From December	North (1 <sup>st</sup> of January)
India	October to May	North (1 <sup>st</sup> of January)
Peru	December to September	North (1 <sup>st</sup> of January)
Yemen	June to September	South (1 <sup>st</sup> of July)
Saudi Arabia	November to May	North (1 <sup>st</sup> of January)
Myanmar	September to May	North (1 <sup>st</sup> of January)

Table A.3 – Growing season calendar in tropical regions.

**Length of the harvest season.** Length of the harvest season in the first 10 wheat producing countries worldwide according to FAO calendars (table A.4) (FAO, 2026). This length was used to determine the period of host availability after maturation, here chosen as 2 month.

Country	Length of harvest season
China	1.5 months
India	2.5 months
Russia	1 month
USA	2 months
France	2.5 months
Canada	2 months
Pakistan	2.5 months
Ukraine	2 months
Germany	2 months
Australia	2 months

Table A.4 – Length of the harvest season in the top 10 wheat producing countries.

#### A.1.4 Environmental conditions for infection

In the three days following deposition, spores should enter two consecutive phases (Roelfs, 1992) in order to complete infection: i) germination, which requires high relative humidity,

dark or dull conditions (solar radiation  $< 26.7 \text{ W}/m^2$ ) and temperatures between  $15^\circ\text{C}$  and  $24^\circ\text{C}$ ; ii) appressorium formation, which requires high relative humidity, sunlight (solar radiation  $> 133.3 \text{ W}/m^2$ ) and temperatures between  $15^\circ\text{C}$  and  $35^\circ\text{C}$ .

As specified in Chapter 2, rain is considered a necessary condition for deposition and infection. Solar radiation limits for the detection of infection are expressed in lux in the article by Allen Sader *et al.* (Allen-Sader et al., 2019). More specifically, germination solar radiation upper threshold was set to  $3200 \text{ lux}$ , while appressorium formation solar radiation lower threshold was set to  $16.000 \text{ lux}$ . Here a conversion factor of  $120 \text{ W}/(m^2 \times \text{lux})$  was adopted, since solar irradiance of 1 Sun ( $1,000 \text{ W}/m^2$ ) equals approximately  $120,000 \text{ lux}$  (Michael et al., 2020).

All environmental variables used to detect environmental conditions for infection come from the GDAS data base, whose temporal resolution is of 6 hours (00 – 06 – 12 – 18).

### A.1.5 Connectivity

The results of every backward Lagrangian simulation from site  $i$  is a set of 120 points ( $j = 1, 2, \dots, 120$ ), corresponding to each hour of a 5-days trajectory, each one characterized by latitude, longitude, altitude, local time, UV solar radiation, rain. From here, cumulative UV solar radiation and rain are computed. These variables contributed to the definition of connection probability  $P_{ij}$  of successful dissemination between a location  $i$  where the host is available and a location  $j$  where host is susceptible, as defined in the Chapter 2 (Fig. 2.1).

Provided that all the other conditions are satisfied (cell  $i$  being crossed by a trajectory at a altitude lower than the mixed layer depth, between 9:00 and 15:00 and with precipitation  $< 2.54 \text{ mm}/h$ ),  $P_{ij}$  is the product of the mutually independent probability of release  $P^1(T)$ , survival to UV radiation  $P^2(UV)$  and rain scavenging  $P^3(R)$ . Probability  $P^1(T)$  of spore release is defined as follows

$$P^1(T) = \begin{cases} 0 & \text{if } T < T_l \text{ or } T > T_u \\ \frac{T-T_l}{T_{opt}-T_l} & \text{if } T_l \leq T < T_{opt} \\ \frac{T_u-T}{T_u-T_{opt}} & \text{if } T_{opt} \leq T < T_u \end{cases} \quad (\text{A.5})$$

$T_l = 5^\circ\text{C}$  being the lower limit,  $T_u = 40^\circ\text{C}$  the upper and with maximal sporulation at  $T_{opt} = 30^\circ\text{C}$  (Prank et al., 2019).

The probability  $P^2(UV)$  of spore survival to UV radiation in the air column is defined as:

$$P^2(UV) = e^{-\lambda_{UV}[f_{UV}SR(1+z_fz)]} \quad (\text{A.6})$$

Where  $\lambda_{UV} = 1.1$  represents the spore sensibility to UV radiation (Meyer et al., 2017; Maddison and Manners, 1972),  $f_{UV}$  is the fraction of UV radiation in the cumulative solar



radiation  $SR$  (MJ) reaching the atmosphere (Foyo-Moreno et al., 1999), and  $f_z$  is a correction term due to the altitude  $z$  (Calbó et al., 2005).

Similarly, rain scavenging is accounted as (Isard et al., 2005):

$$P^3(R) = e^{-R/K_R} \quad (\text{A.7})$$

Where  $R$  is the cumulative rainfall along the trajectory and  $K_R = 25.4$  mm (Isard et al., 2005) represents the constant describing the loss of spores to cumulative rainfall along the trajectory. For sake of numerical tractability, we arbitrarily discarded trajectories associated with  $P_{ij} < 0.01$  to limit the number of retained connections between cells. For what concerns the remaining edges, probabilities  $P_{ij}$  are rounded up into 5 classes (0.01 – 0.21; 0.21 – 0.41; 0.41 – 0.61; 0.61 – 0.81; 0.81 – 1). This operation allowed consecutive release points  $i$  belonging to the same probability class to be linearly interpolated not to omit suitable release locations between two calculated points  $j$ . For simplicity and for continuity reasons, it was assumed that the segment linking two consecutive suitable release locations  $i$  is, in turn, made of suitable locations.

An unweighted version  $\hat{C}$  was generated by imposing the condition  $C > 0$ . For specific applications, the matrices  $A_t$  were also averaged on a weekly basis, obtaining 52  $C_w$  weekly networks, on a monthly basis, obtaining 12  $C_m$  monthly networks, and on a yearly basis, obtaining 6  $C_y$  networks representing each simulated year. Finally, in order to retain only the most recurrent connections in years 2013-2016, we also aggregated the adjacency matrices  $C_y$  considering only those connections happening at least 3 times out of 4. Finally, for testing, we also aggregated the adjacency matrices  $C_y$  using the minimum function in the years 2017 and 2018. In these two last cases, the principle was to consider only those connections happening with frequency  $\geq 75\%$ .

**Cluster detection algorithm.** The cluster detection algorithm used to find highly connected clusters based on the average connectivity network  $C$  (Clauset et al., 2004). A cluster detection algorithm is a procedure enabling to identify groups of nodes with similar connectivity patterns. Nodes belonging to the same cluster share more connections among them rather than with nodes of other clusters. Clusters with less than 25 nodes were discarded to display only the most prevalent ones.

### A.1.6 Network validation

**Puccinia pathway.** In order to reconstruct the Puccinia pathway in North America, we modelled the recorded onset date of Puccinia along the 97<sup>th</sup> meridian in 1922 – 1992 with an index based on cumulative in-strength. Given the connectivity matrix  $C^w$  of week  $w$ , the

cumulative in-strength  $I_i(W)$  of the cell  $i$  is given by:

$$I_i(W) = \sum_{w=1}^W \sum_{j \neq i} C_{ji}^w \quad (\text{A.8})$$

In Fig. 2.3 in Chapter 2 we compare the cumulative in-strength  $I_i(W)$  with the observed onset date.

### A.1.7 Surveillance strategy

**Greedy algorithm to solve the set cover problem.** A greedy algorithm is an iterative procedure enabling to find a locally optimal solution at each step. The greedy algorithm for determining the minimum sentinel set consist in: 1) finding the node  $i$  with the highest in-degree, i.e., the number of cells sending spores towards  $i$ , 2) add this node to the sentinel set, 3) label the coverage of  $i$  as covered and remove all its ingoing edges. Repeat the procedure until the proportion of nodes in the surveilled set reaches the desired threshold or all nodes are labelled as either sentinels or covered. The sentinels compose the optimal set.

Are there other nodes indices which can behave better than the set cover in identifying an efficient sentinel set able to surveil the widest possible domain? In Chapter 2 we addressed these questions by modifying the criteria underlying the definition of the sentinel set. Beside the set cover algorithm, we chose 5 networks metrics, namely:

1. In-strength, i.e., the sum of the weights of the edges incoming a node.
2. Betweenness, which quantifies the number shortest paths between two other nodes of the network passing for a specific node (Freeman, 1978).
3. PageRank, influence ranking method based on the random walk concept, mostly used to provide quality ranking for each web page (Page et al., 1999).
4. Random walk generalized accessibility, a centrality measure based on the concept of random walk on networks (De Arruda et al., 2014).
5. Random samplings of node, repeated 20 times.

**SI model.** The delay in detection was estimated by assuming an abstract SI model integrating a binary version of the weekly connectivity network  $C^w$  and no recovering is allowed. The networks  $C^w$  are obtained as the intersection of the weekly networks of validation years 2017-2018. For instance, network  $C^{w=1}$  is the binary network made of all nodes and edges that exist in both  $C_{year=2017}^{w=1}$  and  $C_{year=2018}^{w=1}$ . For each node, the state of the cell  $x_i$  at time  $w + 1$  is defined as:

$$x_{i,w+1} = \begin{cases} 1, & \text{if } \sum_j^N C_{ji}^w x_{j,w} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (\text{A.9})$$

Where  $x_{i,w} = 0$  means that, at time  $t$ , node  $i$  is susceptible (S), while  $x_{i,w} = 1$  means infected and infectious (I). The diagonal of matrix  $C$  is set to 1 so that an infected node cannot recover. In each simulation,  $x_{i,w=0}$  is set to 0 except for a specific “inoculated” node, which is set to 1 at the beginning of its susceptibility season ( $w = 0$ ). The detection delay is computed as the number of steps needed for a node belonging to the sentinel node to become infected, and correspond to the “time to detection” index by Holme (Holme, 2018). The SI models runs for one year (from  $w = 0$  to  $w = 52$ ).

The following sizes of the sentinels’ sets were chosen to estimate detection delay: 10, 20, 35, 50, 65, 80, 100, 150, 200, 275, 350, 500, 650, 800, 1,000.

The computation of the average detection delay takes into account only those nodes of the domain for which the detection delay is feasible. To do that, completely unconnected nodes have been excluded from the computation of the average detection delay. This means that only 7,251 cells (out of 7,814) able to infect the smallest sentinel set (size = 10) are considered for delay calculation. This is due to the fact that the 10-sentinel set has no sentinel in unconnected components (such as Australia), whose observation delay would provide infinite values.

To overcome calculation problems due to infinite delays, we also elaborated a new performance index, called *Disease detection ratio*. It is defined as the fraction of the total number of simulations for which the sentinel set intercepted the epidemics at least once before the end of the 52<sup>nd</sup> iteration (i.e., before one year after the first node has been inoculated).

We compared the values of such index with the ones that we would have observed choosing a set of nodes by means of other network metrics, namely in- and out-degree, in- and out-strength, betweenness, PageRank, random walk generalized accessibility obtaining that the Set Cover algorithm provides the best performances but it is outperformed by the betweenness strategy for larger sentinel set sizes the networks metrics defined before (in-strength, betweenness, PageRank, Random walk generalized accessibility, random sampling).

## A.2 Supplementary results

### A.2.1 Host availability and susceptibility calendar definition

Estimated date of dormancy end (begin of host availability). Mean over 6 years (figure A.2)

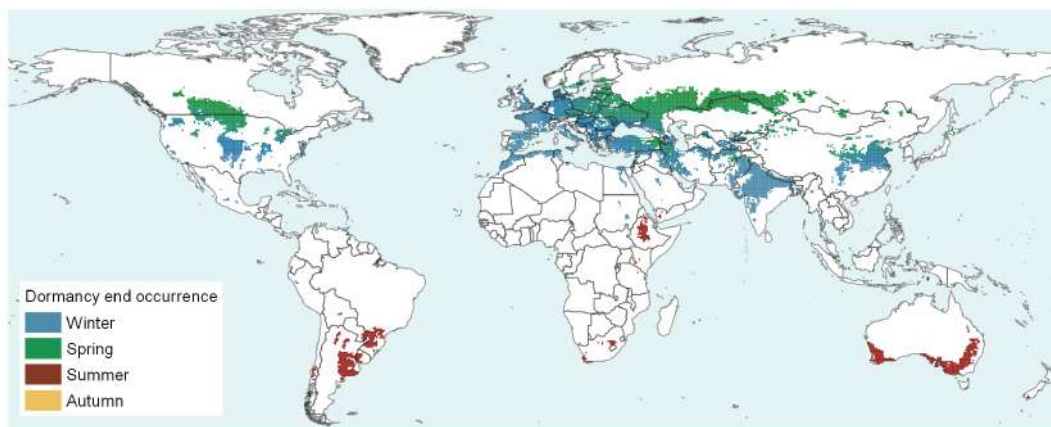


Figure A.2 – Estimated dormancy end occurrence.

Estimated date of wheat completely harvested (end of host availability). Mean over 6 years (figure A.3).

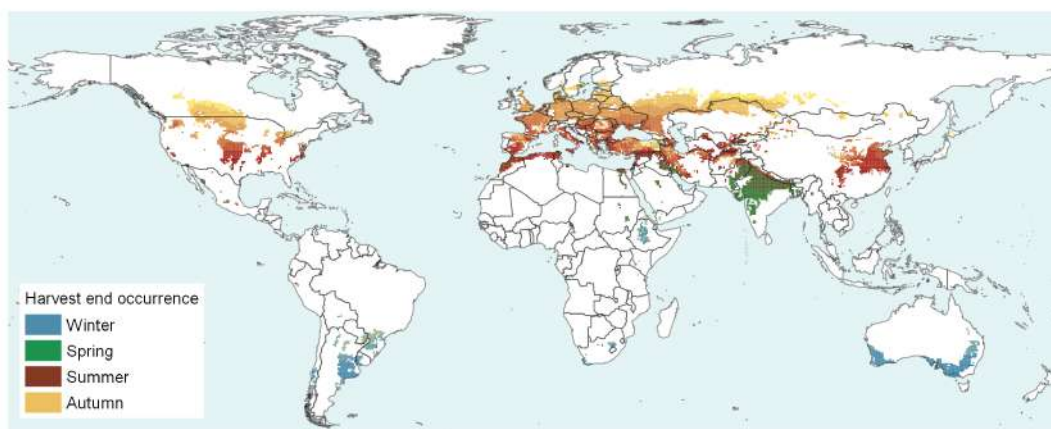


Figure A.3 – Estimated harvest end occurrence (seasons expressed according to the Northern Hemisphere).

Estimated frequency of environmental conditions for infection, in days per year. Mean over 6 years (figure A.4).

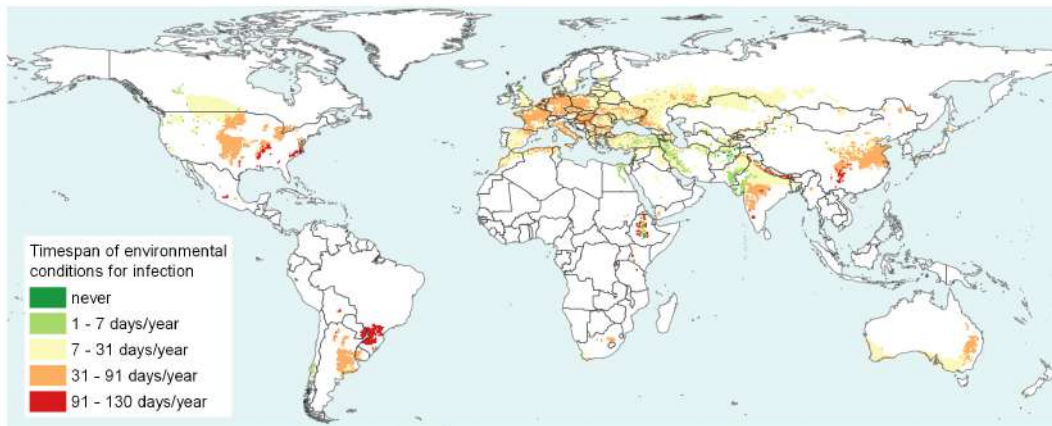


Figure A.4 – Estimated timespan of environmental conditions for infection.

Estimated length of growth season, in month per year (host availability). Mean over 6 years (figure A.5).

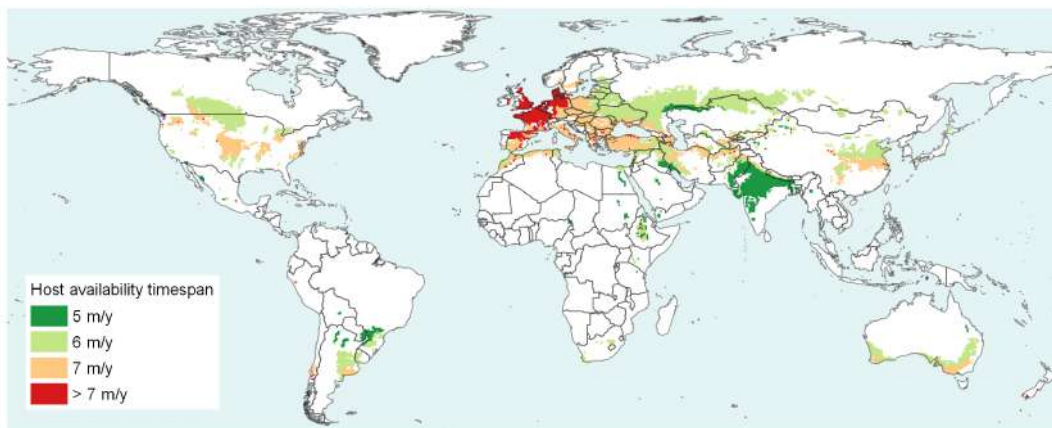


Figure A.5 – Estimated length of the growing season (or host availability) as difference between the end of harvest and dormancy end.

Representation of host availability timespan (green), environmental suitability for infection (yellow) and host susceptibility timespan (red) for 25 locations across the world for each one of the 6 years take into consideration (figure A.6).

## A.2.2 Network clusters

16 clusters detected with the cluster detection algorithm are reported in figure A.8. Topological representation of the network is reported in figure A.9.

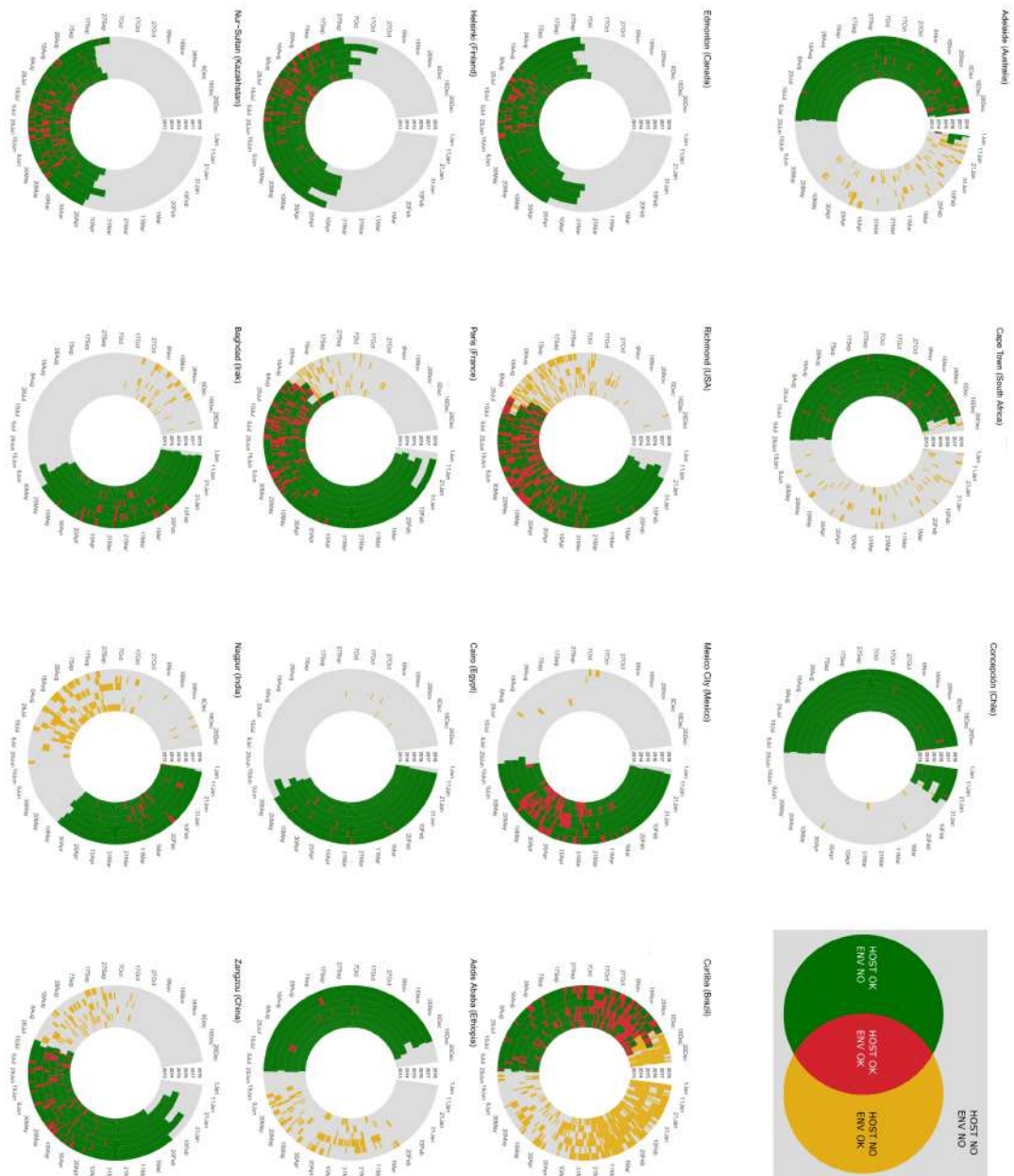


Figure A.6 – Rings of wheat availability, environmental conditions for infection, wheat susceptibility of 16 representative cities across the world.



Figure A.7 – Where are representative cities of figure A.6 located?

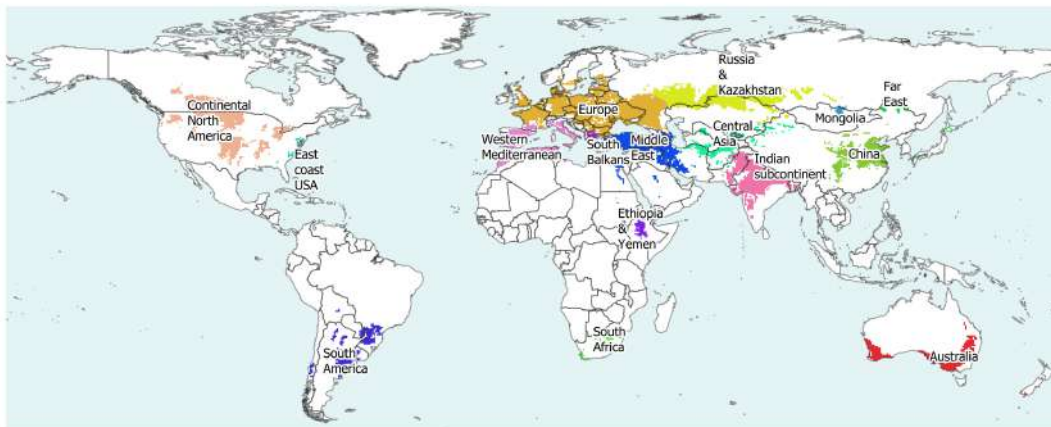


Figure A.8 – 16 clusters detected with the detection algorithm.

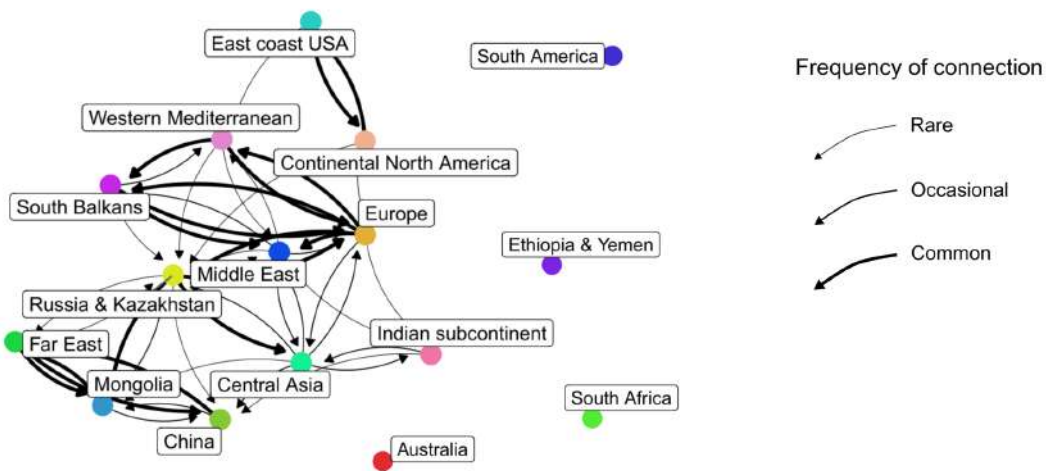


Figure A.9 – Topologic representation of wind connections worldwide after applying the cluster detection algorithm.

### A.2.3 Pathways

Additional pathways to those of Chapter 2 are reported in figure A.10.

Interestingly, we found that the average poleward speed tends to be lower in regions between  $25^\circ$  and  $35^\circ$  absolute latitude (South America, Australia, Indian subcontinent, China at 6.0 km/d, 3.5 km/d, 3.2 km/d and 3.1 km/d respectively), and higher in regions closer to the poles (absolute latitude higher than  $35^\circ$ , Western Europe and Maghreb, North America, Eastern Europe and Russia at 9.1 km/d, 8.4 km/d, 5.4 km/d, 3.9 km/d respectively).

### A.2.4 Network surveillance

A complete coverage of the worldwide wheat producing regions by means of the recurring connections can be achieved by monitoring 1,007 sentinels, i.e., less than 13% of the total number of nodes of the network A.12a. A coverage of 20% of the entire domain can be obtained with 15 sentinels (figure A.11).

Fig. A.12b shows the continental coverage for increasing sentinel set sizes, relative to the size of the continent. It may be interpreted as the relative efficiency of detection in different continents, which in turns can be interpreted as different wheat distribution patterns. Europe and North America, whose wheat fields are densely distributed, are more easily observed rather than Africa or Australia. Around 6 – 9% of nodes are necessary to observe entire Europe or North America, while around 27 – 35% are needed for Australia or Africa. Europe, North and South America are the first continent being observed as the number of sentinels increase, while Asia, Australia or Africa are only covered by the last chosen sentinels (figure A.12b).



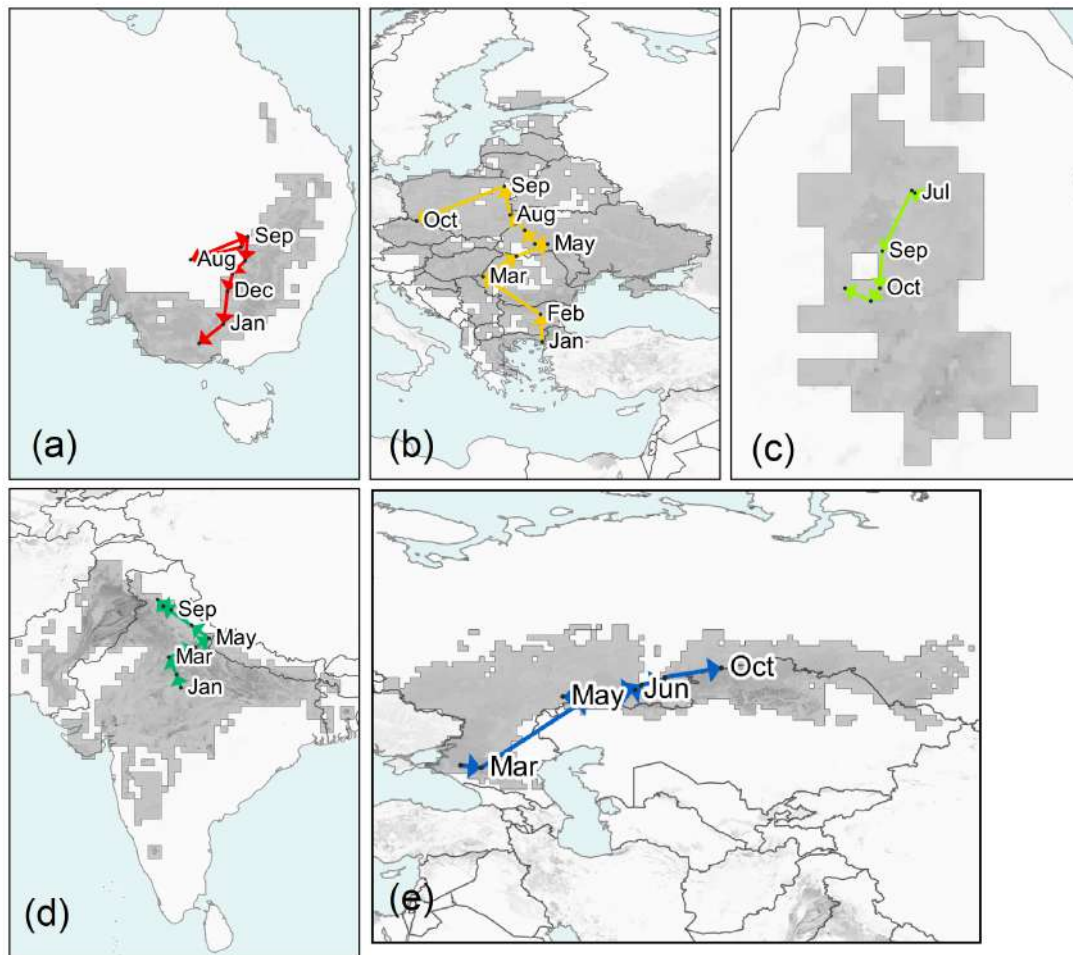


Figure A.10 – Other continental pathways: First row, from the left: (a) Australia, (b) east Europe, (c) Ethiopia. Second row: (d) Indian subcontinent, (e) Russia and Kazakhstan.

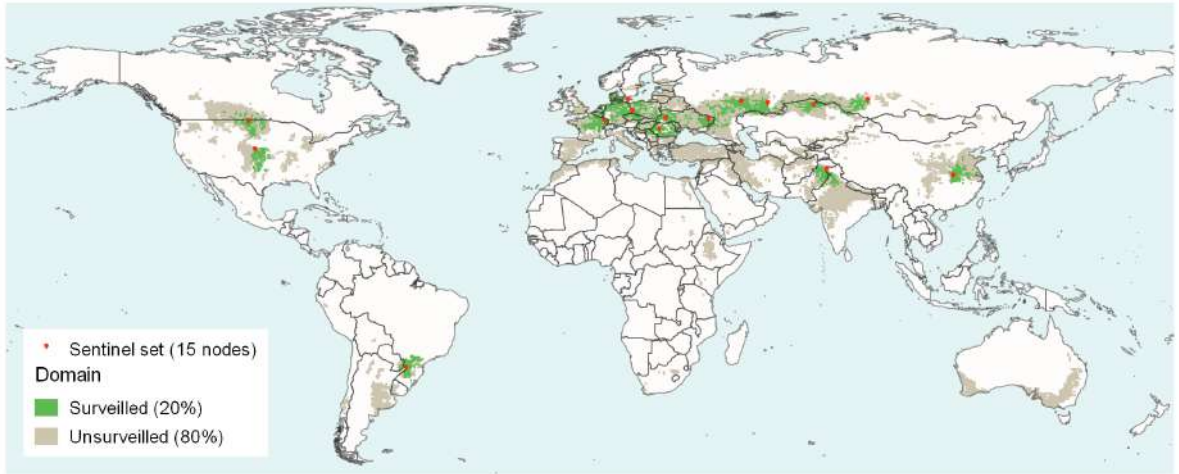


Figure A.11 – Sentinel set observing 20% of the domain.

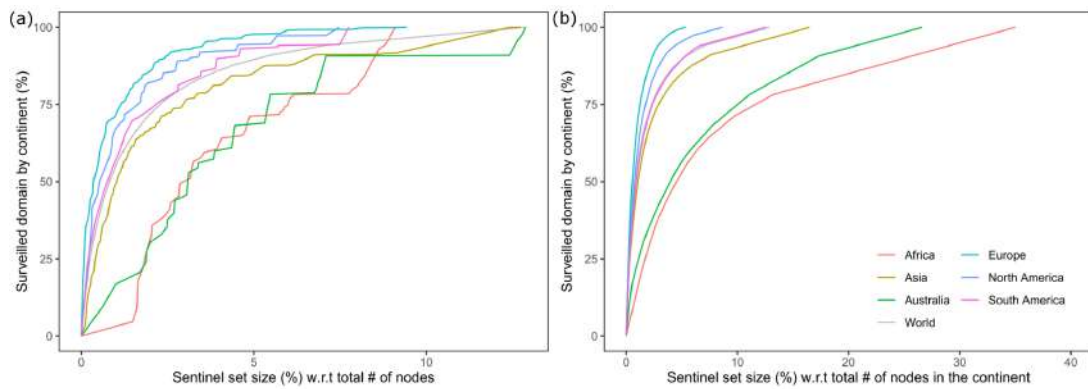


Figure A.12 – Panel (a) shows the relative extent of the coverage in each continent corresponding to different sentinel set sizes (considered as % of worldwide nodes). Panel (b) shows the relative extent of the coverage in each continent corresponding to different sentinel set sizes (considered as % of nodes in that continent)

The average detection delay of the sentinel set obtained via the set cover greedy algorithm, computed for the 7,251 nodes (out of 7,814) for which the delay is always finite, decreases significantly as the number of sentinels increases, going from 46.9 iterations for 10 sentinels to just 28.9 for the 1,000 sentinels that survey the entire domain. Interestingly, for sentinel set larger than 50 nodes (0.6% of the total nodes), the average detection delay overlaps with the third quartile (figure A.13).

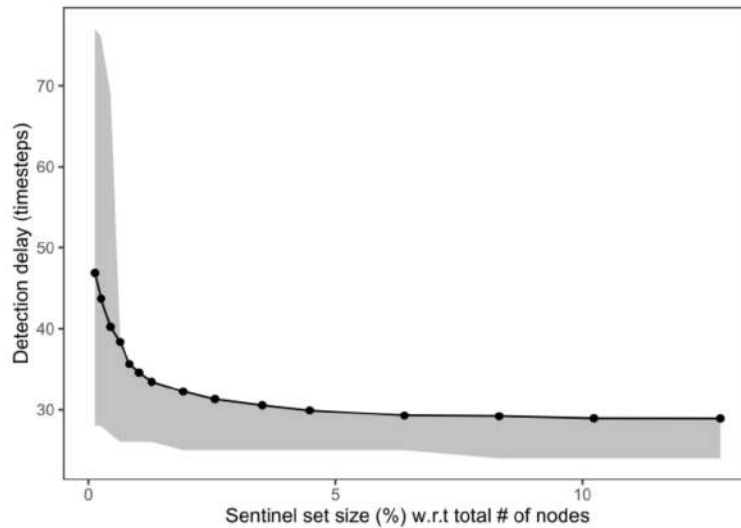


Figure A.13 – Detection delay assessed with a spatially explicit SI model corresponding to different sentinel set sizes. Grey area correspond to the 1<sup>st</sup> and 3<sup>rd</sup> quartile.

Furthermore, we repeated the computations the *DDR* with different values of detection delay in figure A.14 (complementary to Fig. 1.5c in Chapter 2). Interestingly, for the 2-weeks *DDR* (figure A.14a), the set cover shows the best performances for relatively small sentinel sets (< 8%), while it is outperformed by betweenness. 3-months, 6-months and 1-years delays *DDRs* (figure A.14b, c, d) are qualitatively similar, Set cover being always the best choice (except for extremely small sentinel sets, in which in-strength is preferred), followed by random sentinel sets. PageRank improves its *DDR* performances for sentinels sets larger than 10% while increasing the value of detection delay, performing better than random selection (figure A.13c, d).

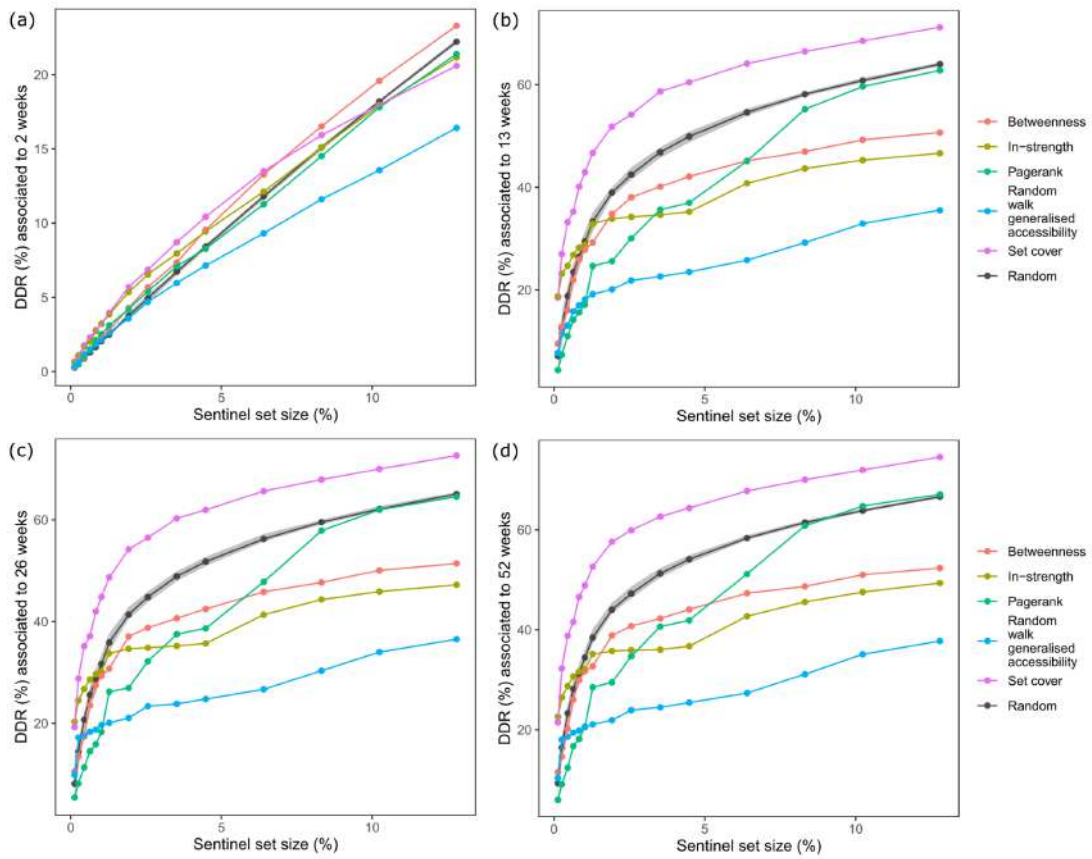


Figure A.14 – *DDR* assessed with a spatially explicit Susceptible-Infected model corresponding to increasing sentinel set sizes for different strategies, associated to different values of detection delay: (a) two weeks (b) three months (c) six months (d) one year. Grey area correspond to the 1<sup>st</sup> and 3<sup>rd</sup> quartile.

## Bibliography

- Allen-Sader, C., Thurston, W., Meyer, M., Nure, E., Bacha, N., Alemayehu, Y., Stutt, R. O., Safka, D., Craig, A. P., Derso, E., Burgin, L. E., Millington, S. C., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2019). An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11).
- Calbó, J., Pages, D., and González, J.-A. (2005). Empirical studies of cloud effects on uv radiation: A review. *Reviews of Geophysics*, 43(2).
- Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Physical review E*, 70(6):066111.
- De Arruda, G. F., Barbieri, A. L., Rodríguez, P. M., Rodrigues, F. A., Moreno, Y., and Costa, L. D. F. (2014). Role of centrality for the identification of influential spreaders in complex networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 90(3):1–17.
- FAO (2026). "crops". [http://www.fao.org/faostat/en/#rankings/countries\\_by\\_commodity](http://www.fao.org/faostat/en/#rankings/countries_by_commodity). Retrieved 2018-08-12. Countries - Select All; Regions - World + (Total); Elements - Production Quantity; Items - Wheat; Years – 2016.
- Foyo-Moreno, I., Vida, J., and Alados-Arboledas, L. (1999). A simple all weather model to estimate ultraviolet solar radiation (290–385 nm). *Journal of Applied Meteorology*, 38(7):1020–1026.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.
- Holme, P. (2018). Objective measures for sentinel surveillance in network epidemiology. *Physical Review E*, 98(2):022313.
- International Food Policy Research Institute (2019). Global spatially-disaggregated crop production statistics data for 2010, version 2.0.
- Isard, S. A., Gage, S. H., Comtois, P., and Russo, J. M. (2005). Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience*, 55(10):851–861.
- Large, E. C. (1954). Growth Stages in Cereals Illustration of the Feekes Scale. *Plant pathology*, 62(1):207.
- Maddison, A. and Manners, J. (1972). Sunlight and viability of cereal rust uredospores. *Transactions of the British Mycological Society*, 59(3):429–443.
- Mcmaster, G. S. and Smika, D. E. (1988). Estimation and Evaluation of Winter Wheat Phenology in the central great plains. *Agricultural and Forest Meteorology*, 43:1–18.

- McMaster, G. S. and Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, 87(1).
- Meyer, M., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017). Large-scale atmospheric dispersal simulations identify likely airborne incursion routes of wheat stem rust into Ethiopia. *Phytopathology*, 107(10):1175–1186.
- Michael, P. R., Johnston, D. E., and Moreno, W. (2020). A conversion guide: solar irradiance and lux illuminance. *Journal of Measurements in Engineering*, 8(4):153–166.
- Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.
- Prank, M., Kenaley, S. C., Bergstrom, G. C., Acevedo, M., and Mahowald, N. M. (2019). Climate change impacts the spread potential of wheat stem rust, a significant crop disease. *Environmental Research Letters*, 14(12).
- Roelfs, A. P. (1992). *Rust diseases of wheat: concepts and methods of disease management*. Cimmyt.

### **A.3 The published article**

ENVIRONMENTAL RESEARCH  
LETTERS

## LETTER

## OPEN ACCESS

RECEIVED  
4 March 2022REVISED  
23 May 2022ACCEPTED FOR PUBLICATION  
26 May 2022PUBLISHED  
7 June 2022

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.



## Early-detection surveillance for stem rust of wheat: insights from a global epidemic network based on airborne connectivity and host phenology

Andrea Radici<sup>1</sup> , Davide Martinetti<sup>1,\*</sup> and Daniele Bevacqua<sup>2</sup> <sup>1</sup> BioSP UR 546, INRAE, 228 route de l'Aérodrome, CS 40509 Domaine Saint-Paul—Site Agroparc, 84914 Avignon, France<sup>2</sup> PSH UR 1115, INRAE, 228 route de l'Aérodrome, CS 40509 Domaine Saint-Paul—Site Agroparc, 84914 Avignon, France

\* Author to whom any correspondence should be addressed.

E-mail: [davide.martinetti@inrae.fr](mailto:davide.martinetti@inrae.fr)**Keywords:** long-distance dispersal, stem rust of wheat, *Puccinia graminis*, network, surveillance, early detectionSupplementary material for this article is available [online](#)**Abstract**

Stem rust of wheat, caused by the airborne pathogen *Puccinia graminis*, is a re-emerging crop disease representing a major concern to global food security. Potential long-distance transport by wind over a worldwide distributed host represents a challenge to effective surveillance and control of this disease. To monitor this disease, we have created a global epidemic network for stem rust of wheat combining (a) Lagrangian simulations of air-mass trajectories computed with the NOAA's HYSPLIT model; (b) land use from the Map Spatial Production Allocation Model and (c) meteorological and environmental conditions that are known to affect bio-physical processes involved in the biology of *P. graminis* spores. Our findings are in agreement with the well known north-American 'Puccinia pathway' and suggest the existence of other sub-continental pathways at the global scale. We used network theory to conceive surveillance strategies aimed at early detection of outbreaks while minimizing the number of nodes to be surveilled (also referred to as sentinels). We found that the set cover algorithm, due the high average connectivity of the network (density = 0.4%), performs better than a number of other network metrics and permits us to identify an optimal sentinel set (1% of the network nodes) to surveil 50% of the network. Our results also show that effective surveillance plans for stem rust of wheat can be designed, but that they need to account for the actual geographical scale of the underlying epidemiological process and call for an international and trans-boundary approach.

**1. Introduction**

Epidemics caused by airborne pathogens represent an inveterate challenge to agricultural management [1]. Moreover, the transition to monocultures characterized by low genetic diversity has made the global farming system less resilient to pathogens [2]. Airborne pathogens, capable of long-distance transport, create a network of connections among globally diffused crops. For example, *Puccinia graminis*, causing stem rust of wheat, is seasonally dispersed from northern Mexico up to Canada, along the 'Puccinia pathway' [3]. On the other hand, a single-leap event, the hurricane Ivan, transported

*Phakopsora pachyrhizi* spores from Colombia to Alabama, introducing soybean rust in North America [4, 5]. The risk of losses in food production requires to take action against the diffusion of alarming pathogens, such as *P. graminis* [2, 6, 7] and species of the same genus [8]. Its aerobiology has been largely studied in the last century [9–12] and experimental procedures have been recently developed for studying three-dimensional transport and detection of spores [13, 14]. However, the use of mathematical models to simulate spore transport by wind at the continental scale is very recent [15–20]. In fact, the use of such models would allow researchers, policy makers and farmers to design surveillance strategies [21–23]

capable of detecting outbreaks and take timely countermeasures (e.g. phytosanitary intervention [18]).

Meyer *et al* [15] assessed the risk of long distance dispersal of *P. graminis* in East Africa and the Middle East via a Lagrangian particle dispersion model. The same modelling framework was then used to investigate the possible origin of virulent strains found in Ethiopia [16]. Allen-Sadder *et al* [18] developed a decision support system integrating spore dispersal to help optimize fungicide allocation. Prank *et al* [19] investigated the impacts of climate change on worldwide spores transport patterns. Suttrave *et al* [24] designed a network-based surveillance-system for *P. pachyrhizi*, at the subcontinental scale, by including wind direction and intensity. Despite these advances, to our knowledge, models of spore dispersal have never been coupled to network analysis to design optimal surveillance strategies for *P. graminis* at the global scale.

In this current study we constructed a time-variant connectivity network for *P. graminis* spores at the worldwide scale explicitly considering wind patterns, host phenology and meteorological conditions affecting spores aerobiology. We used biophysical models and graph theory to identify the most susceptible worldwide regions and to reconstruct the epidemic movement through different subcontinents. We validated our findings using available knowledge regarding the North-American ‘*Puccinia* pathway’ [3]. Lastly, we identified those nodes of the network (i.e. regions of the world) that, when monitored, enable early detection of an outbreak. We are confident that our approach is sufficiently generic to be applied to other airborne plant pathogens, provided that basic knowledge on the pathogen aerobiology, host physiology and distribution are available.

## 2. Materials and methods

### 2.1. Case study

*Puccinia graminis* f. sp. *tritici* is a heteroecious airborne fungal pathogen responsible for stem rust of wheat. Wheat, the main host of *P. graminis*, constitutes a staple food for a great proportion of human population. It covers  $2.15 \times 10^8$  hectares worldwide (1.4% of earth surface), representing the most abundant agricultural type of land cover [25].

In recent decades, yield losses caused by stem rust have been limited by planting resistant cultivars and fungicides application. Nonetheless, the emergence of new strains overcoming plant resistance may provoke severe yield losses, accounting for up to 50%–90% of the wheat production at the regional scale [19, 26]. An extraordinary outbreak in Ethiopia in 2013–2014 caused a complete yield loss in some cultivars [27], while concerns were raised in Europe after the detection of new virulent strains capable of infecting previously resistant cultivars [28].

### 2.2. Geographic domain and air mass trajectories

We extracted the worldwide wheat distribution from the MapSPAM database [29] and computed the percentage of wheat cover on a regular grid with a resolution of  $0.5^\circ$ . We accounted only for cells containing at least 2% of wheat land cover. The resulting database contains 7814 cells (see section SI 1.1). We calculated backward air-mass Lagrangian three-dimensional trajectories of 120 hours [16] with the HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) model [30] from 1st January 2013 to 31st December 2018, using atmospheric GDAS data at a spatial resolution of  $0.5^\circ$ . We ran simulations from the centroids of each cell at 00:00, 06:00, 12:00 and 18:00 UTC +0, at an above-ground altitude set to the minimum between the base of the cloud and the mixed layer depth. Along each simulated trajectory, we recorded atmospheric variables at an hourly frequency. Overall, out of more than  $6.8 \times 10^7$  potential trajectories, we computed only those  $1.6 \times 10^6$  satisfying the criteria of host availability and environmental suitability specified in sections 2.3.1 and 2.3.4 (see also section SI 1.2).

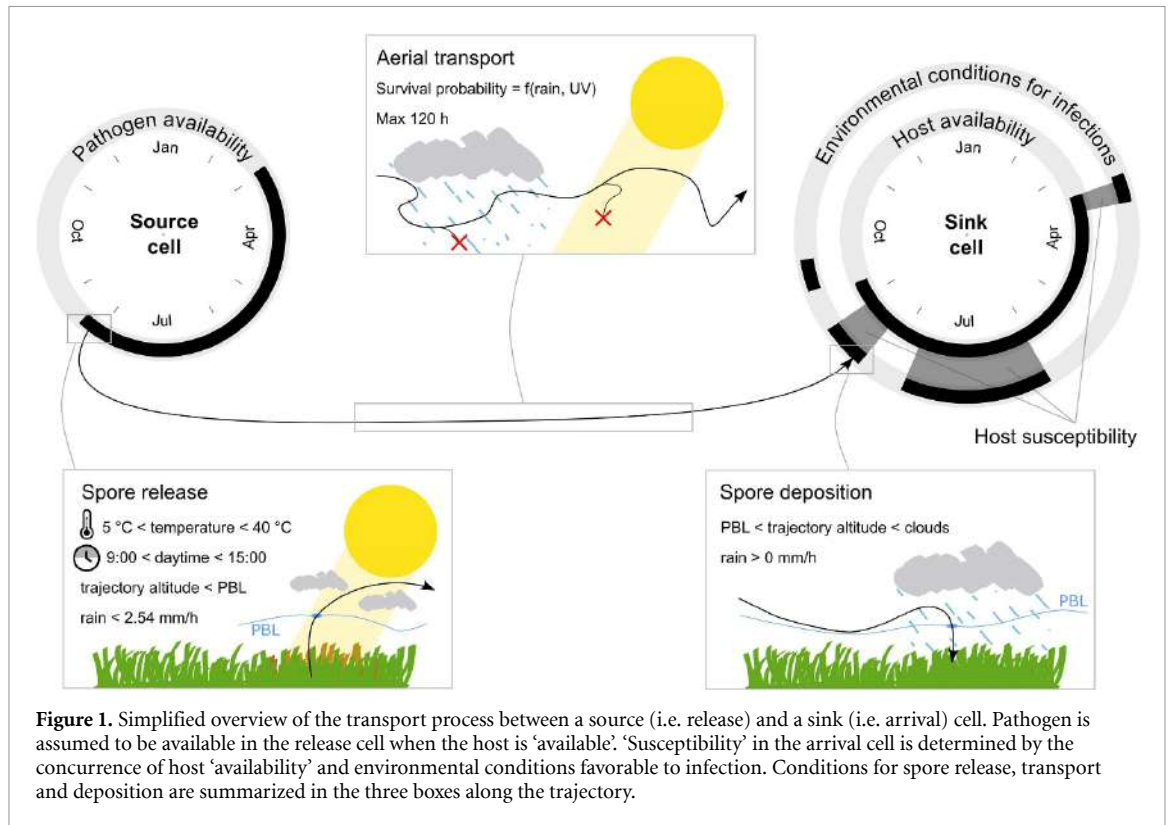
### 2.3. The connectivity network

We built a time-variant connectivity network for stem rust explicitly considering host (i.e. wheat) phenological phases, environmental conditions affecting spore aerobiology [31] and infection. In this network, nodes represent cells composing the worldwide wheat producing regions, while edges are computed from air-mass trajectories in order to model the dispersal of airborne spores. The nodes (or cells) of the networks remain fixed, while edges are re-computed for each of the 8764 (4 times a day for 6 years) simulations.

More specifically, edges result from the application of a set of biophysical filters to identify those air-mass movements that correspond to transport events (figure 1). We assume that, for a given time  $t$ , an edge exists between an arrival cell  $j$  and any cell  $i$  if the following conditions are satisfied:

- (a) In cell  $i$ , which at time  $t - x$  ( $x \in \{1, \dots, 120\}$  h) is crossed by the air-mass trajectory arriving at  $j$  at time  $t$ :
  - The host is ‘available’, i.e. it is present and in a favourable phenological state for infection and sporulation, see section 2.3.1.
  - Environmental conditions are compatible for spore release, see section 2.3.2.
  - The altitude of the trajectory is lower than the planetary boundary layer.
- (b) Environmental conditions along the trajectory between  $i$  and  $j$  allow spore survival, see section 2.3.3.
- (c) The host in cell  $j$  is ‘susceptible’, which means:
  - The host is ‘available’.





- Environmental conditions are favourable for spore deposition and host infection, see section 2.3.4.

### 2.3.1. Host availability

Depending on local climate, wheat is cultivated and harvested at different times across the world. We used a growing degree-day model [32, 33] to compute the period of host availability for each year in each cell of the domain as a function of the year temperature conditions. We used a model conceived for the US [32] and we adjusted it for the Southern Hemisphere (i.e. initializing it at July 1st). For tropical countries, we consulted case by case the calendar of the prevailing season provided in the FAO country briefs [25] (see section SI 1.3 (available online at [stacks.iop.org/ERL/17/064045/mmedia](https://stacks.iop.org/ERL/17/064045/mmedia))).

### 2.3.2. Spore release

Spore release in the air column is promoted under unstable atmospheric conditions [34, 35]. Following Meyer *et al* [15], we assumed that release occurs between 9:00 and 15:00, provided precipitations are lower than  $2.54 \text{ mm h}^{-1}$  [18]. Furthermore, the rate of spore release is affected by temperature [19]. Eventually, a probability  $P^1(T)$  of release is defined as a function of temperature (see section SI 1.5).

### 2.3.3. Aerial spore transport and survival

Spores are released at canopy level, but only a fraction actually enters in the atmospheric layers leading

to long-distance transportation. The rest may be dispersed in the air column within few kilometers under turbulent atmospheric conditions [34, 36]. Hence, we considered as suitable release sites those that are crossed by an air-mass trajectory within the mixed layer depth (therefore, able to drag spores distributed in the air column). Furthermore, during the aerial phase, spores must endure critical environmental conditions [34]. The limiting factors affecting the survival of *P. graminis* spores are the exposure to UV radiation [15] and the washout by rain. We then defined survival probabilities to UV radiation  $P^2(UV)$  and to washout  $P^3(R)$  (see section SI 1.5).

### 2.3.4. Spore deposition and host infection

Since precipitation is the main responsible for spore deposition and infection [34, 37–41], we assumed that an edge can point an arrival cell  $j$  at time  $t$  only if precipitation occurs. In fact, even if spore deposition might occur also in dry conditions [42, 43], wet deposition provides a better environment for the development of infection and it is usually treated as the most important element. Following deposition, infection requires specific environmental conditions [15, 18, 44]: in the three following days, spores should enter the phase of germination, requiring dark conditions and mild temperatures, and appressorium formation, requiring sunlight and warm temperature [39]. We then used these conditions to identify times and sites where infection can occur after deposition (see section SI 1.4).

### 2.3.5. Time-variant connectivity networks

In order to represent the connectivity between wheat producing regions, we took advantage of the mathematical formalism of networks [45], by describing the global epidemic network of stem rust via 6-hourly connectivity networks  $A^t$  whose elements  $A_{ij}^t$  indicate the probability  $P_{ijt}$  of a successful transport event occurring at time  $t$  from node  $i$  to node  $j$ . The probability  $P_{ijt}$  is computed as the product of the mutually independent probabilities  $P_i^1(T)$  of sporulation at release site,  $P_{ij}^2(UV)$  of survival to harmful UV radiation and  $P_{ij}^3(R)$  of washout, provided the concurrence of host availability at site  $i$  and precipitation at site  $j$ . Finally, we aggregate the 8764 networks to obtain a weighted, directed air-mass connectivity network  $C$  (see section SI 1.5) and used a cluster detection algorithm [46] to identify groups of nodes with similar connectivity patterns.

## 2.4. Model validation: recovering the ‘Puccinia pathway’ in North America

Since the direct observation of long-distance spores transport is beyond current technological capacities, we tested the validity of our model checking if we could use it to identify the well known North-American ‘Puccinia pathway’.

We considered the yearly northward movement of the stem rust infection front in North America by tracking the average onset date ( $\pm 1$  standard deviation) during the 1922–1992 period along the 97th meridian west [3]. Then, we computed the weekly cumulative in-strength values of the nodes located along the same meridian in the US from the weekly connectivity networks  $C_w$ . The in-strength is a network metric considering the sum of the weights of the edges pointing a node  $i$ . In this case, the in-strength of a node  $i$  corresponds to the sum of the average weekly frequencies of connection from the other nodes toward  $i$ . We used the cumulative version of this metric, by assuming that local emergence of a disease is observable after sufficiently abundant inoculum has been deposited in that node. We graphically compared observed onset dates and average weekly cumulative in-strength in figure 3 (see section SI 1.6). Here, we built the frequency distribution of the cumulative in-strength values intersected by 1922–1992 onset date observations and computed the values corresponding to the interquartile range. Lastly, cumulative in-strength values falling in the interquartile range are shaded on the heat map.

## 2.5. Network surveillance

### 2.5.1. Exploring ‘Puccinia pathways’ worldwide

In order to reconstruct the trajectories of propagation of stem rust outbreaks at world scale, we analyzed how the center of mass of the monthly connectivity networks ( $C_m$ ) in-strength moves across different subcontinents. Hence, for a given set of nodes representing a subcontinent (and the corresponding

geographical coordinates), we defined the center of mass of the monthly connectivity networks in-strength as the geographical point whose coordinates are given by the average latitude and longitude weighted by the nodes monthly in-strengths.

### 2.5.2. Definition of efficient surveillance strategies

We considered the problem of establishing a reduced set of sentinels, i.e. nodes where the presence of the pathogen is monitored systematically, that should guarantee the largest coverage of the domain and provide an early-warning system for the appearance of new pathogen strains [22]. First of all, we defined the coverage of a node  $i$  as the set of nodes that points towards  $i$ , under the assumption that monitoring the presence of the pathogen in  $i$  implies observing all those nodes that are pointing to it. In this case, node  $i$  is referred as the *sentinel* of its coverage. We considered hence the network  $\tilde{C}$  generated by considering those edges of the yearly connectivity networks  $C_y$  recurring at least three times over a 4 year interval 2013–2016 (i.e.  $\geq 75\%$  of the times), in order to account only for the most frequent connections. The problem of finding the smaller set of sentinels that guarantees the complete coverage is formally equivalent to the set cover problem in graph theory, that happens to have NP-complete computational complexity [24, 47]. Nonetheless, we used an iterative greedy algorithm, providing a sub-optimal minimum sentinel set. To validate the efficiency of the sentinel set, we assessed its performance in terms of ratio of surveilled domain using the network obtained by the intersection of the yearly connectivity networks of 2017 and 2018. We compared such performances with the ones given by sets of nodes chosen via other network metrics, namely in-strength, betweenness [48], PageRank [49] and random walk generalized accessibility [50], calculated on the aggregated 2013–2016 networks, and a set of 20 random samplings of nodes (see definitions and procedures in section SI 1.7).

### 2.5.3. Measuring early-detection performance of the sentinel set

Finally, we assessed the performance of different sentinel sets in terms of early detection by means of a compartmental Susceptible-Infected (SI) model based on the intersections of the weekly networks  $C_w$  in the years 2017 and 2018. Within this framework, a node can be either susceptible (S) or infected (I), its state depending on the state of the neighbors in the previous time steps. Namely, in this simplified approach, we assume that, in a given time step  $t$ , a node pass from the state S to I if it has at least one neighbor in state I at time  $t - 1$ , with no recovery. We run 7814 model simulations, each time letting the outbreak start from a different node (inoculum). Then, we defined the *Disease detection ratio (DDR)* of a given sentinel set as the fraction of the total number

of simulations for which the sentinel set intercepted the epidemics at least once before the end of the fourth iteration (i.e. before one month after the first node has been inoculated). This new metric allows to compare the performance of different sentinel sets even when they fail to achieve detection within the first month, something that may happen when the first inoculated node is located in a rather isolated part of the network (see section SI 1.7).

### 3. Results

#### 3.1. Worldwide susceptibility and connectivity patterns

The duration of host susceptibility periods, i.e. the concurrence of the host availability and environmental conditions favorable to infection, varies considerably across the world (figures 2(a) and SI6). The majority of cells is susceptible for a period between a week and a month. The nodes where susceptibility occurs for more than a month per year are located in northeastern America, southern Brazil and Paraguay, central Europe and central China. The only nodes with susceptibility lasting for more than three months per year are located in Ethiopia (see section SI 2.1).

Regions located in the Northern Hemisphere are generally well connected, in such a way that Europe, Asia and North Africa create a unique connected component (figure 2(b)). In spite of the obstacle represented by the Atlantic Ocean, extremely long distance connections may occur from North America to the Mediterranean basin. Conversely, clusters in the Southern Hemisphere are isolated between them, although internally connected. A first representation of the role played by each node within the epidemic network is given by the sum of the probabilities associated to trajectories going into (in-strength: figure 2(c)) and out from (out-strength: figure 2(d)) that node. In biological terms, in-strength (out-strength) is a proxy of the extent to which a region acts as a sink (source) of spore. In-strength appears to be higher in specific regions located in Europe and in the northeastern US, southern Brazil, the Himalaya and central China. Some regions exhibit great variability within relatively short distances, namely Ethiopia, Middle East and Central Asia. Other regions show a smoother and regular gradient, like the US. In this sense, a continuum of intermediate in-strength values extends from Europe to Western Siberia. Out-strength separates more sharply those regions characterized by high and low values and, in particular, Europe is characterized by large regions associated with high out-strength values.

#### 3.2. Reconstructing *Puccinia* pathways in North America and elsewhere

Our global epidemic network permitted to recover the well known North-American *Puccinia* pathway. In figure 3 we compare the average onset date of

outbreak along the 97th meridian west observed between 1922 and 1992 in North America [3], with the cumulative in-strength of network nodes along the same meridian. Most of the observed onset dates occur within a certain interval of values of cumulative in-strength (the interquartile range [0.6, 0.9]), suggesting that the cumulative in-strength network metric can proxy the spatio-temporal progression of the '*Puccinia* pathway' in North America. In other words, in this continent one would expect to observe the first signs of an epidemics when the cumulative in-strength has values between 0.6 and 0.9.

Globally, we found that the airborne transport estimated via the center of mass of the monthly in-strength always moves poleward from tropical and temperate regions (figures 4 and SI10), except for the Ethiopian pathway, that follows a southward movement even if it is located in the Northern Hemisphere, likely due to its cropping calendar.

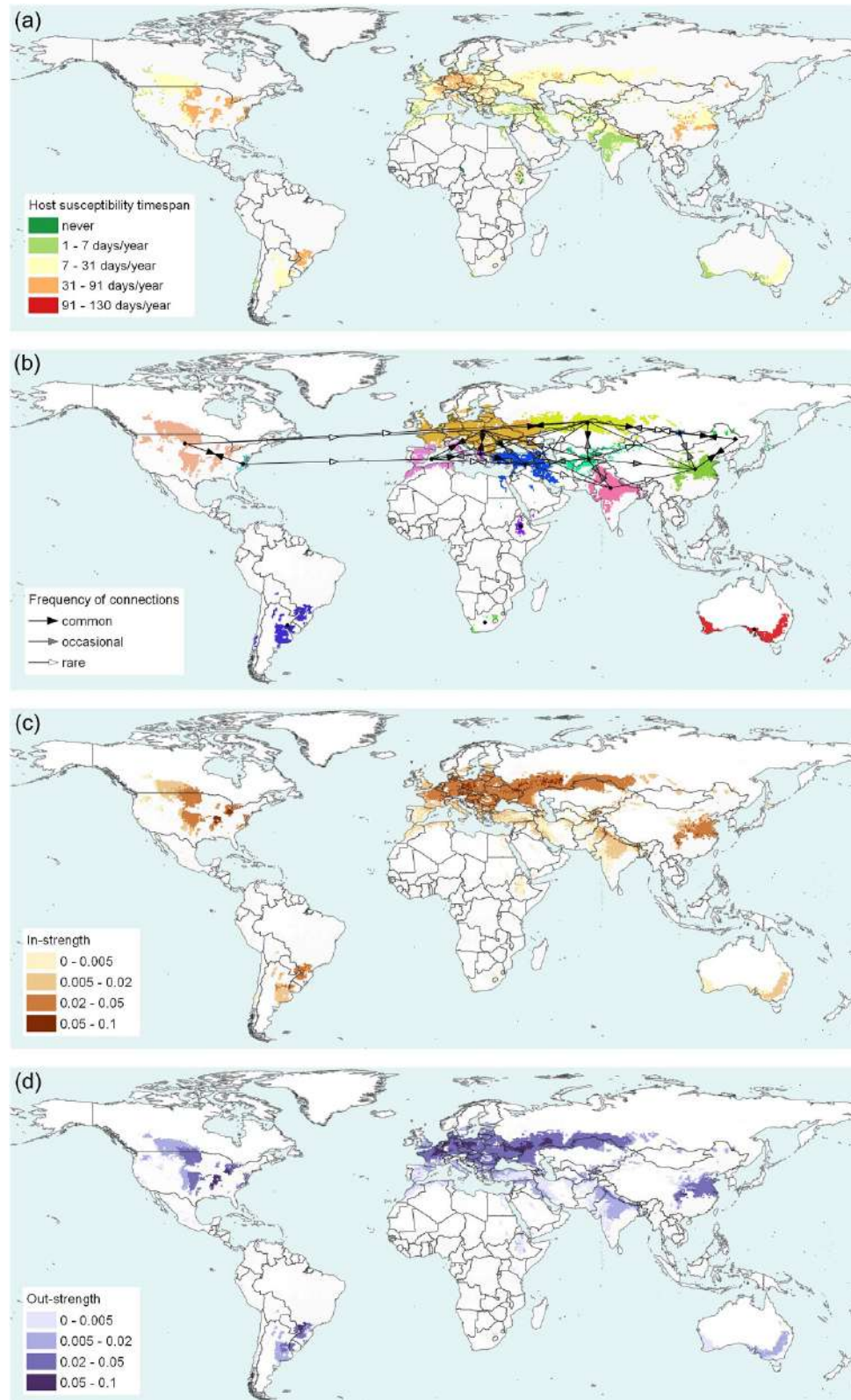
#### 3.3. Optimal sentinel set

The sentinel set obtained on the 2013–2016 network using the set cover algorithm allows for a complete coverage of the worldwide wheat producing regions by monitoring 1007 nodes, i.e. less than 13% of the total (figure SI12). Moreover, a reasonable coverage of 50% of the domain can be obtained monitoring 64 nodes (figure 5(a)), i.e. less than 1% of the total. The first selected nodes are those assuring the greatest coverage, and they do not distribute uniformly across continents (figure SI12—see section SI 2.4).

The sentinel set selected to optimally cover the 2013–2016 epidemic network, via the set cover algorithm, provides satisfactory results also when applied to surveil the epidemic network obtained for the 2017–2018 period (figure 5(b)). In fact, it provides a coverage of 50% of the domain by monitoring 114 nodes (1.5% of the total). For comparison, the same coverage of 50% would require 234 (3%) random sentinels (on average: interquartile range [227; 242]), 475 (6.1%) sentinels if ordered for increasing values of betweenness, or 611 (7.8%) of PageRank.

#### 3.4. Early detection capabilities of the sentinel set

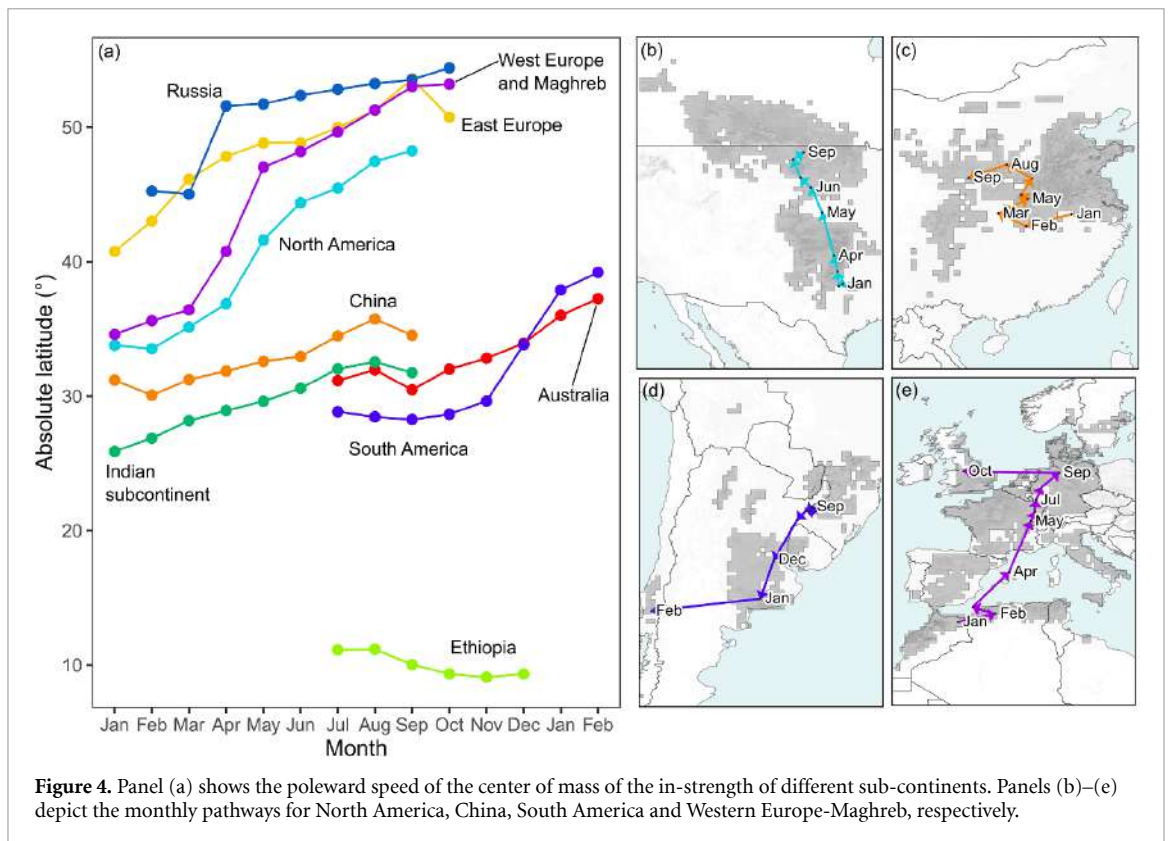
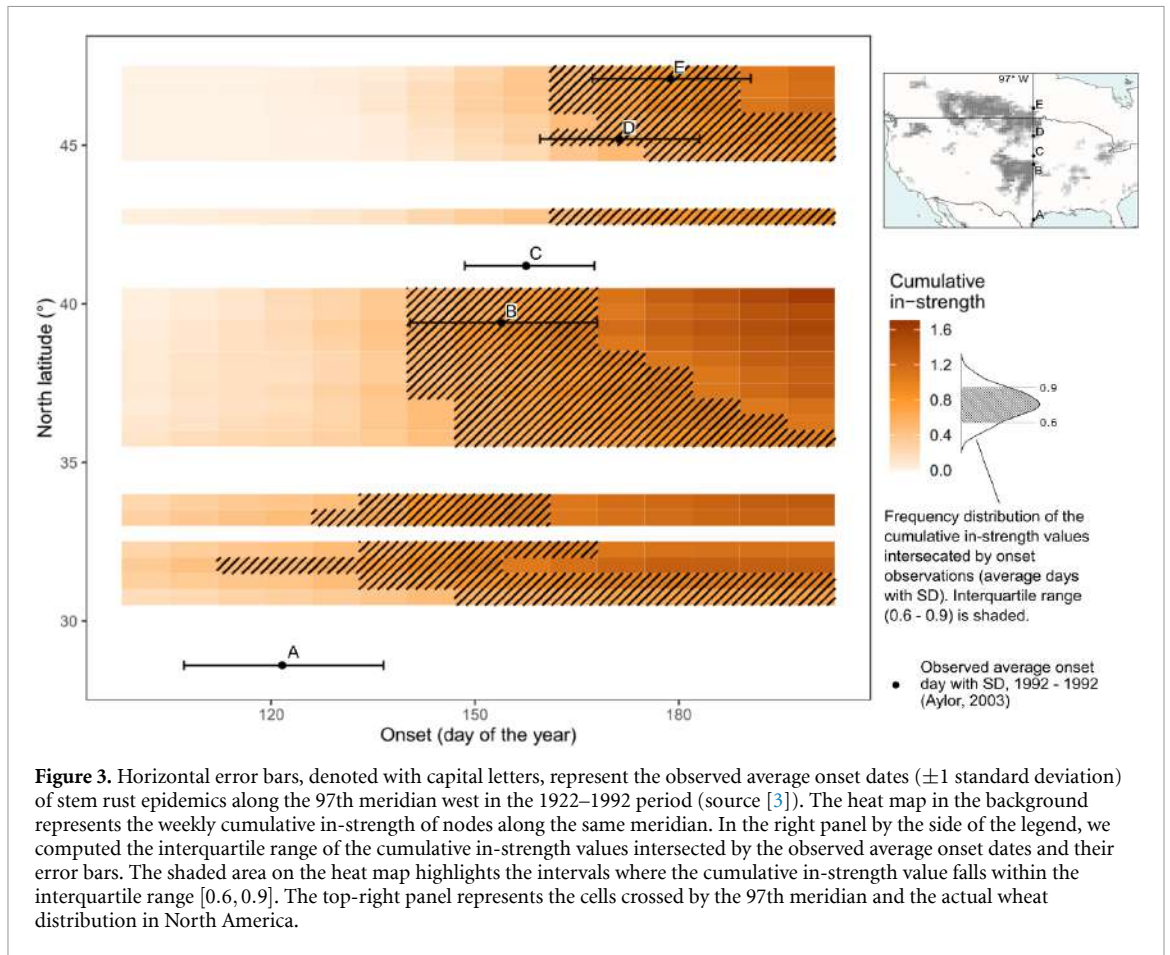
In terms of *DDR*, the set cover algorithm outperformed all the other methods for sentinel sets containing between 20 and 650 nodes (figure 5(c)). In-strength provides better results for very small sentinel sets, while random sentinel sets are more suitable when larger sentinel sets have to be designed. The *DDR* associated to 275 sentinels is 19.2%, which means that 275 sentinels are able to detect an epidemic process started from any of 19.2% of the world producing regions within a month. A similar *DDR* of

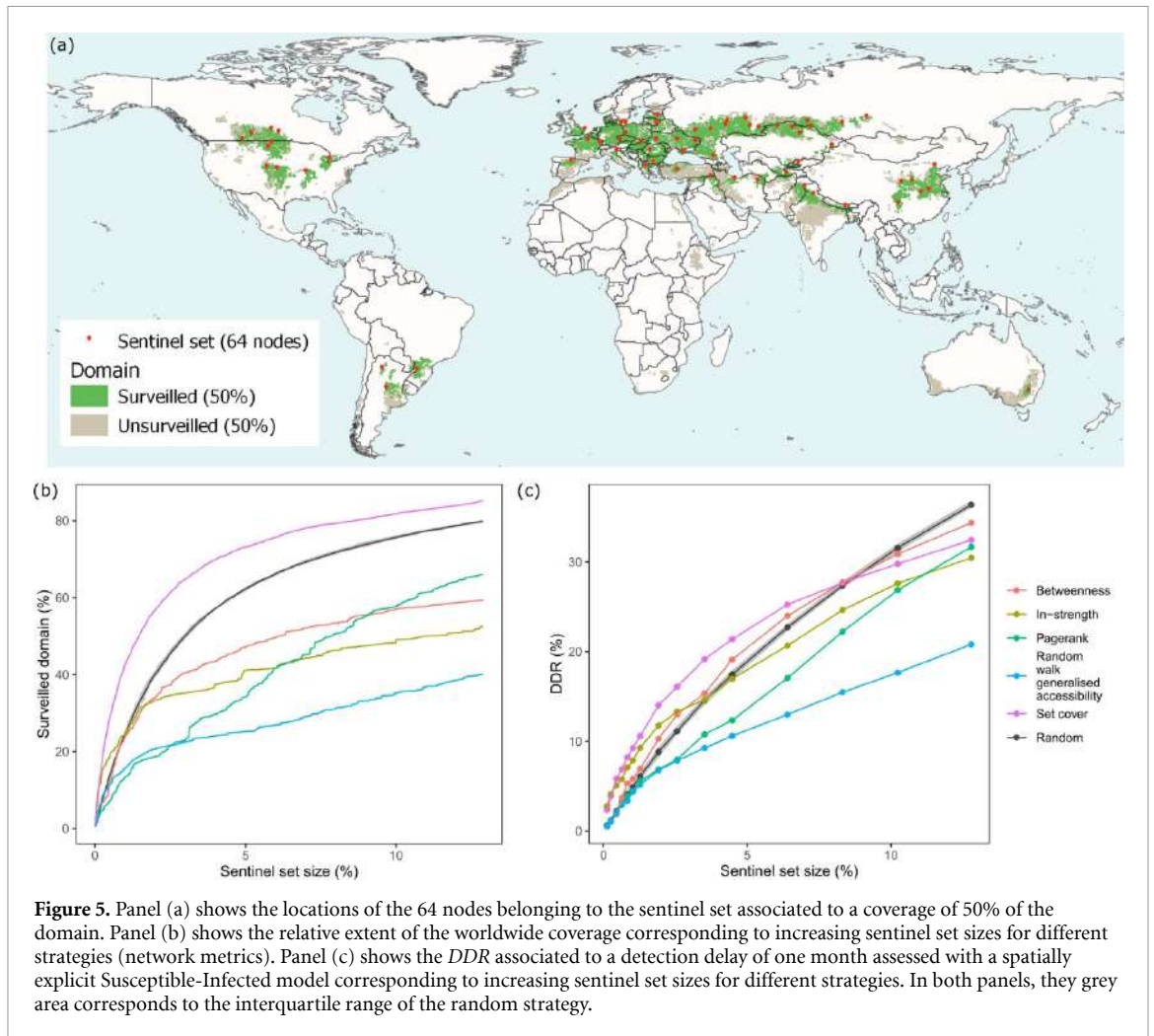


**Figure 2.** (a) Duration of the host susceptibility period, (b) continental connections among clusters, (c) in-strength and (d) out-strength across the world.

19.1% is obtained with 350 nodes chosen according to their betweenness, or more than 500 nodes according to their PageRank. Between 350 and 500 random sentinels are needed to achieve a *DDR* around 18%–22%.

We estimated the *DDR* associated to larger values of detection delay (3 month, 6 month, 1 year), showing that the set cover strategy improves its performances against the random sentinels sets (figure SI14).





**Figure 5.** Panel (a) shows the locations of the 64 nodes belonging to the sentinel set associated to a coverage of 50% of the domain. Panel (b) shows the relative extent of the worldwide coverage corresponding to increasing sentinel set sizes for different strategies (network metrics). Panel (c) shows the *DDR* associated to a detection delay of one month assessed with a spatially explicit Susceptible-Infected model corresponding to increasing sentinel set sizes for different strategies. In both panels, the grey area corresponds to the interquartile range of the random strategy.

## 4. Discussion

In this study we proposed an original modelling framework based on air mass movements, network analysis, meteorology, ecology and plant physiology to describe the global epidemic network of stem rust of wheat. Eventually, we used the model to identify previously unknown pathways of disease spread at the globe scale and to define a set of sentinels that should be primarily surveyed to achieve early detection of future outbreaks.

### 4.1. The most susceptible regions

Our results indicate that the duration of the susceptibility period greatly varies among regions (figures 2(a) and SI6). Eastern US, southern Brazil, central Europe and China, which are important areas of wheat production, experience the longest susceptibility periods. On the other hand, other important wheat-producing areas from the Middle East up to the Indian subcontinent have shorter periods of susceptibility. The reason could be the temporal mismatch between the occurrence of environmental

conditions for infection and of host availability. This is particularly true for India and in Brazil, where, despite a long-lasting occurrence of environmental conditions for infection, the susceptibility period is constrained by relatively short host availability periods (figure SI6). Ethiopia, which is characterized by exceptional climatic diversity for tropical latitudes [51], exhibits a very heterogeneous behavior, with nodes that are never susceptible in close proximity to others that are highly susceptible (91–130 days per year). Since both host availability and favorable environmental conditions for spore deposition and host infection are climate dependent processes, it is likely that global crop epidemic dynamics will be affected by predicted climate changes [52]. Prank *et al* [19] predicted a worldwide increase in sporulation period under a RCP8.5 climate change scenario in 2100—potentially compensated by a general decrease of probability of germination. On the other hand, laboratory experiments conducted on other fungal species suggested that future climate conditions would rather inhibit sporulation, but increase mycelium growth rate [53].

#### 4.2. On the construction of a global epidemic network

To construct an epidemic network we had to explicitly account for biophysical conditions influencing spore aerobiology. This filtering enabled to discard uninformative air-mass connections and to reduce the complexity of the network, making it simultaneously more significant and manageable.

In principle, by tracking the movement of an air mass for an indefinite long time, one would obtain a completely connected network, i.e. a network with edge density equal to 100%. On the other hand, our epidemic network has an edge density equal to 3.7%, whereas the network of only the most recurring connections (i.e. those connections found at least 3 times out of 4 in the yearly networks 2013–2016) has a density of 0.4%. For comparison, previous studies using a similar technique but filtering only over a 48 hours time [54] obtained a density of 28% over the Mediterranean basin for the period 2011–2017. In essence, our filtering allows to make epidemic networks more manageable and informative, presenting itself as alternative to other network-based methods for the extraction of the truly relevant connections [55].

#### 4.3. From network science to global crop protection

Our results on the frequency of connections (figure 2) indicate that regions in the Northern Hemisphere are more densely connected than in the Southern one. This can be explained by the fact that 90% of wheat fields are located in the Northern Hemisphere while the clusters below the Tropic of Cancer, i.e. South America, South Africa, Ethiopia and Australia are thousands of kilometers away from each other. Furthermore, our prediction of possible long-distance connections from North America to Europe, and from Europe towards Central Asia and Russia, are in accordance with previous theoretical and empirical findings [17, 19, 20, 56]. On the other hand, due to the time lag in the respective susceptibility seasons, no connection between the Northern and Southern Hemispheres is predicted by our epidemic network. For such a connection to exist, it is necessary the presence of a ‘green bridge’ [15] in a tropical or subtropical region located between the two hemispheres. This could be the case of Ethiopia, where, according to the FAO country brief [25], there exist two wheat growing seasons: *Belg*, from February to July and *Meher*, from May to December. Furthermore, according to our estimates on the host susceptibility duration (figure 2(a)), Ethiopia features some of the nodes with the longest susceptibility period (more than 90 days/year), whereas previous studies already identified this country (and more broadly East Africa and Yemen) as an important stepping stone [56] for the long distance transport of rust spores along the Rift Valley [15].

The in- and out-strength maps of the epidemic network (figure 2(c)) provide an overview of the epidemic role played by each region, sink rather than source of spores or both. Our results indicate that out-strength tends to be stronger towards the equator compared to in-strength, in line with previous findings [3] about the ‘green’ and ‘golden’ wave: regions closer to tropics are the first where infection is possible and therefore they are more likely to act as sources of spores rather than sinks. Among the recent seasonal epidemic events moving poleward, the 2013 regional outbreak of stem rust in Germany seems to have moved to Denmark and Sweden later in the season [7].

If one wanted to use our approach to derive effective measures for global crop protection it is worthy to note that in our exercise we privileged generality over local accuracy. Our assumptions to determine host availability and susceptibility are necessarily oversimplifications of a complex reality, as required by models. For instance, we neglected the presence of alternative hosts, such as spring wheat, whose cropping calendars would increase the host susceptibility period and the frequency of long distance connections. Also, we neglected the existence of the secondary hosts of *P. graminis*, the barberry (*Berberis vulgaris*), which is necessary for sexual reproduction, facilitating the emergence of new strains. It has been successfully eradicated in Western Europe and North America in the last century [7]. Yet, it is present in other regions which turns to be sources of new strains [27, 57] and it has been recently reintroduced in Europe [58].

Our model permitted to design an optimal surveillance strategy capable of covering 50% of wheat cultivated lands while monitoring only 64 sentinel nodes, i.e. less than 1% of the wheat cultivated cells. Although this proportion may seem optimistic for real applications, it is worth recalling the work of Suttrave *et al* [24] on *P. pachyrhizi* indicating that a reliable epidemiological status of soybean rust in the US could be achieved by reducing the sentinel set size from 500 to 12 network optimized nodes.

Of course, the success of the deployment of a sentinel system will also depend on the diagnostic ability of individual sentinels and on the communication between them. In this sense, efforts have recently been devoted to establish international protocols to improve the probability of detection and the timely communication of new outbreaks to a global network of scientists, public authorities and stakeholders (see the Global Cereal Rust Monitoring System [6, 59]). Our findings stress the importance of increasing efforts towards a trans-boundary perspective to efficiently contain the emergence of new virulent air-borne crop pathogens [23].

## Data availability statement

The code that support the findings of this study is available from the corresponding author upon reasonable request. The data that support the findings of this study are available upon reasonable request from the authors.

## Acknowledgments

The authors acknowledge the support of funding from the French National Research Agency (ANR) for the BEYOND project (Contract No. 20-PCPA-0002) and SuMCrop Sustainable Management of Crop Health Program of INRAE that supported the work of all authors. We thank Cindy Morris, Renato Casagrandi, Lorenzo Mari and Samuel Soubeyrand for fruitful discussions, two anonymous referees for extremely valuable suggestions, Luke Riley for valuable language tips as well as the technical support of Loïc Houde for the computation of HYSPLIT trajectories.

## Conflict of interest

The authors have no competing interests.

## ORCID iDs

Andrea Radici  <https://orcid.org/0000-0001-5852-0015>

Davide Martinetti  <https://orcid.org/0000-0003-2047-1793>

Daniele Bevacqua  <https://orcid.org/0000-0002-3341-1696>

## References

- [1] Mahaffee W F and Stoll R 2016 *Phytopathology* **106** 420–31
- [2] Corredor-Moreno P and Saunders D G O 2020 *New Phytol.* **225** 118–25
- [3] Aylor D E 2003 *Ecology* **84** 1989–97
- [4] Schneider R, Hollier C, Whitam H, Palm M, McKemy J, Hernandez J, Levy L and DeVries-Paterson R 2005 *Plant Dis.* **89** 774–774
- [5] Isard S A, Gage S H, Comtois P and Russo J M 2005 *BioScience* **55** 851–61
- [6] Park R, Fetch T, Hodson D, Jin Y, Nazari K, Prashar M and Pretorius Z 2011 *Euphytica* **179** 109–17
- [7] Saunders D G, Pretorius Z A and Hovmøller M S 2019 *Commun. Biol.* **2** 9–11
- [8] Hovmøller M S, Yahyaoui A H, Milus E A and Justesen A F 2008 *Mol. Ecol.* **17** 3818–26
- [9] Zadoks J C 1967 *Neth. J. Plant Pathol.* **73** 61–80
- [10] Burrage S W 1970 *Ann. Appl. Biol.* **66** 429–40
- [11] Maddison A and Manners J 1972 *Trans. Br. Mycol. Soc.* **59** 429–43
- [12] Aylor D E 1986 *Agric. Forest Meteorol.* **38** 263–88
- [13] Damialis A, Kaimakamis E, Konoglou M, Akritidis I, Traidl-Hoffmann C and Gioulekas D 2017 *Sci. Rep.* **7** 44535
- [14] Schmale III D G and Ross S D 2015 *Annu. Rev. Phytopathol.* **53** 591–611
- [15] Meyer M, Cox J A, Hitchings M D, Burgin L, Hort M C, Hodson D P and Gilligan C A 2017 *Nat. Plants* **3** 780–6
- [16] Meyer M, Burgin L, Hort M C, Hodson D P and Gilligan C A 2017 *Phytopathology* **107** 1175–86
- [17] Visser B, Meyer M, Park R F, Gilligan C A, Burgin L E, Hort M C, Hodson D P and Pretorius Z A 2019 *Phytopathology* **109** 133–44
- [18] Allen-Sader C et al 2019 *Environ. Res. Lett.* **14** 115004
- [19] Prank M, Kenaley S C, Bergstrom G C, Acevedo M and Mahowald N M 2019 *Environ. Res. Lett.* **14** 124053
- [20] Wang M, Kriticos D J, Ota N, Brooks A and Paini D 2021 *New Phytol.* **232** 1506–18
- [21] Cunniffe N J, Koskella B, Metcalf E C J, Parnell S, Gottwald T R and Gilligan C A 2015 *Epidemics* **10** 6–10
- [22] Parnell S, van den Bosch F, Gottwald T and Gilligan C A 2017 *Annu. Rev. Phytopathol.* **55** 591–610
- [23] Ristaino J B et al 2021 *Proc. Natl Acad. Sci.* **118** e2022239118
- [24] Suttrave S, Scoglio C, Isard S A, Hutchinson J S and Garrett K A 2012 *PLoS One* **7** e37793
- [25] FAO 2021 FAO—country brief (available at: <http://www.fao.org/giews/countrybrief/>) (Accessed 20 July 2021)
- [26] Huerta-Espino J, Singh R and Roelfs A P 2014 *Fungi From Different Substrates* (Boca Raton, FL: CRC Press) pp 217–59
- [27] Olivera P et al 2015 *Phytopathology* **105** 917–28
- [28] Bhattacharya S 2017 *Nat. News* **542** 145
- [29] Institute I F P R 2019 Global spatially-disaggregated crop production statistics data for 2010, version 2.0 (<https://doi.org/10.7910/DVN/PRFF8V>)
- [30] Draxler R R and Hess G D 1998 *Aust. Meteorol. Mag.* **47** 295–308
- [31] Aylor D E 1999 *Agric. Forest Meteorol.* **97** 275–92
- [32] McMaster G S and Smika D E 1988 *Agric. Forest Meteorol.* **43** 1–18
- [33] McMaster G S and Wilhelm W W 1997 *Agric. Forest Meteorol.* **87** 291–300
- [34] Levetin E 2015 *Manual of Environmental Microbiology* (Washington, DC: ASM Press) pp 3.2.8-1–20
- [35] Oneto D L, Golan J, Mazzino A, Pringle A and Seminara A 2020 *Proc. Natl Acad. Sci. USA* **117** 5134–43
- [36] Aylor D 2017 *Aerial Dispersal of Pollen and Spores* (St. Paul, MN: The American Phytopathological Society)
- [37] Morris C E, Sands D C, Glaux C, Samsatly J, Asaad S, Moukahel A R, Gonçalves F L and Bigg E K 2013 *Atmos. Chem. Phys.* **13** 4223–33
- [38] Nagarajan S and Singh D 1990 *Annu. Rev. Phytopathol.* **28** 139–53
- [39] Roelfs A P 1992 *Rust Diseases of Wheat: Concepts and Methods of Disease Management* (Mexico City: CIMMYT)
- [40] Rowell J and Romig R et al 1966 *Phytopathology* **56** 807–11
- [41] Li X, Yang X, Mo J and Guo T 2009 *Eur. J. Plant Pathol.* **123** 377–86
- [42] Emerson E W, Hodshire A L, DeBolt H M, Bilsback K R, Pierce J R, McMeeking G R and Farmer D K 2020 *Proc. Natl Acad. Sci.* **117** 26076–82
- [43] Slinn W 1977 *Water Air Soil Pollut.* **7** 513–43
- [44] Baiocco S, Cavina F and Pradoles G 2021 A weather-based simulation model for the development of wheat stem rust epidemics 2021 29th Conf. Open Innovations Association (FRUCT) (IEEE) pp 22–29
- [45] Choufany M, Martinetti D, Senoussi R, Morris C E and Soubeyrand S 2021 *Front. Appl. Math. Stat.* **6** 67
- [46] Clauset A, Newman M E and Moore C 2004 *Phys. Rev. E* **70** 066111
- [47] Garey M R and Johnson D S 1979 *Computers and Intractability* vol 174 (San Francisco, CA: Freeman)
- [48] Freeman L C 1977 *Sociometry* **40** 35–41
- [49] Page L, Brin S, Motwani R and Winograd T 1999 The pagerank citation ranking: bringing order to the web *Technical Report* Stanford InfoLab
- [50] De Arruda G F, Barbieri A L, Rodriguez P M, Rodrigues F A, Moreno Y and Costa L D F 2014 *Phys. Rev. E* **90** 032812
- [51] Peel M C, Finlayson B L and McMahon T A 2007 *Hydrol. Earth Syst. Sci.* **11** 1633–44
- [52] IPCC 2021 *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment*



- Report of the Intergovernmental Panel on Climate Change* vol 1, ed Masson Delmotte *et al* (Cambridge: Cambridge University Press)
- [53] Damialis A, Mohammad A B, Halley J M and Gange A C 2015 *Int. J. Biometeorol.* **59** 1157–67
- [54] Choufany M, Martinetti D, Soubeyrand S and Morris C E 2021 *Sci. Rep.* **11** 11093
- [55] Serrano M A, Boguná M and Vespignani A 2009 *Proc. Natl Acad. Sci.* **106** 6483–8
- [56] Mayol E *et al* 2017 *Nat. Commun.* **8** 201
- [57] Olivera P D, Sikharulidze Z, Dumbadze R, Szabo L J, Newcomb M, Natsarishvili K, Rouse M N, Luster D G and Jin Y 2019 *Phytopathology* **109** 2152–60
- [58] Barnes G, Saunders D G and Williamson T 2020 *Plant Pathol.* **69** 1193–202
- [59] Morris C E, Geniaux G, Nédellec C, Sauvion N and Soubeyrand S 2021 *Plant Pathol.* **71** 86–97

## Appendix B

# Supplementary Information - Chapter 3

### B.1 Supplementary Methods

#### B.1.1 Domain extraction

We built the domain upon units extracted from the so-called Safran grid (Bertuzzi and Clastre, 2022), which is composed of regular square cells with spatial resolution of  $0.11^\circ \times 0.11^\circ$  ( $\sim 8 \times 8 \text{ km}^2$ ). For each unit, we computed stone fruit orchard area and peach orchard area (this latter is referred simply as “cultivated area” in Chapter 3: here, for sake of clarity, we specify that we mean the peach cultivated area). We extracted stone fruit cultivated areas from data collected by the 2010 national French agricultural survey (RGA) conducted by the French Ministry of Agriculture (*Ministère de l’Agriculture et de la Souveraineté alimentaire*; Ministère de l’Agriculture (2010)). From this database, we summed “stone fruits” (column `G_1013_LI_DIM2` == “fruits à noyau”) area data (column “Superficie correspondante (hectares)”) of all farms (column “G\_1013\_LIB\_DIM1” == “Ensemble des exploitations (hors pacages collectifs)”) by municipality (expressed by the INSEE code).

This operation allowed us to obtain the area  $A_{s_m}$  covered with stone fruit orchards by municipality  $m$ .

We computed the stone fruits density by municipality and geographically intersected the corresponding shapefile with Safran grid. This operation allowed to associate each unit of the Safran grid with a stone fruit density. Open source software QGIS 3 (QGIS Development Team, 2022) was used for this operation.

We obtained the spatial data about peach orchards from Eurostat at NUTS2 (Eurostat, 2021), which correspond to groups of French regions (as defined before the 2016 reform: FR1 *Île de France*, FR2 *Bassin Parisien*, FR3 *Nord-Pas-de-Calais*, FR4 *Est*, FR5 *Ouest*, FR6 *Sud-Ouest*,

FR7 *Centre-Est*, FR8 *Méditerranée* and FR9 the *Département d'Outre Mer*). We considered data about “Peach and apricot trees - Area by density classes and group of cultivars (area in ha)”.

We grouped together all areas from following categories: “Dessert peach and nectarine trees [PCD]”; “Peaches for fresh consumption [PCD\_PEA]”; “Yellow flesh peaches [PCD\_PEAY]”; “Very early yellow flesh peaches [PCD\_PEAY\_VE]”; “Early yellow flesh peaches [PCD\_PEAY\_E]”; “Medium-early yellow flesh peaches [PCD\_PEAY\_M]”; “Late yellow flesh peaches [PCD\_PEAY\_L]”; “White flesh peaches [PCD\_PEA\_W]”; “Very early white flesh peaches [PCD\_PEA\_W\_VE]”; “Early white flesh peaches [PCD\_PEA\_W\_E]”; “Medium-early white flesh peaches [PCD\_PEA\_W\_M]”; “Late white flesh peaches [PCD\_PEA\_W\_L]”; “Doughnut peaches [PCD\_PEA\_D]”; “Medium-early yellow flesh nectarines [PCD\_NECY\_M]”; “Late yellow flesh nectarines [PCD\_NECY\_L]”; “White flesh nectarines [PCD\_NECW]”; “Very early white flesh nectarines [PCD\_NECW\_VE]”; “Early white flesh nectarines [PCD\_NECW\_E]”; “Medium-early white flesh nectarines [PCD\_NECW\_M]”; “Late white flesh nectarines [PCD\_NECW\_L]”; “Peach and nectarine trees for industrial processing (including group of Pavie) [PCI]”.

This operation allowed us to obtain the area  $Ac_n$  covered with peach orchards within NUTS2 region  $n$ . We estimated area  $Ac_m$  covered with peach orchards by municipality  $m$  located in NUTS2 region  $n$  via Eq. B.1. This equation assumes that peach cultivated area in municipality  $m$  is proportional both *i*) to the stone fruit cultivated area in that municipality  $As_m$  and *ii*) to the peach cultivated area in the region where that municipality is located  $Ac_n$  and that *iii*) the sum of the peach cultivated area by municipality  $m$  is equal to the peach cultivated area in that region  $n$ .

$$Ac_m = Ac_n \frac{As_m}{\sum_{m \in n} As_m} \quad (B.1)$$

Each unit of the grid is thus associated with *i*) its area, *ii*) the stone fruits cultivated area and *iii*) the peach cultivated area. Since keeping all the elements of the grid in the domain would imply considerably slower computations describing the dynamic of a great amount of poorly informative units, we decided to keep only those units respecting a threshold of density (i.e., at least  $0.01 \text{ ha/km}^2$ ) and continuity (i.e., only units which form groups of at least 4 contiguous elements; Fig. B.1e). By excluding cells with a low density and with no neighbours we discarded a great amount of cells where peach cultivation is known to be absent, for climatic reasons (for example in central and northern France) but Eq. B.1 may still assign a residual  $Ac$  due to a misclassification error. For instance, we found few, isolated cells in Alsace-Lorraine (north-east of France) where we computed a relevant peach cultivated area. The same cells correspond to a well-known production basin of mirabelle plums, a stone fruit typical of the region and well adapted to its climate. We found more reasonable to conclude that these cells were hosting important mirabelle plum orchards, rather than peach ones; therefore, we proceeded to remove them.

### B.1.2 Pit hardening date

We computed peach growth increase according to equation (Bevacqua et al., 2023):

$$w(t) = \frac{w_B w_M}{w_B + (w_M - w_B) e^{-h(t-t_B)}} \quad (\text{B.2})$$

Where  $w(t)$  is the weight at time  $t$ ,  $w_B = 0.49$  g is the fresh fruit weight at bloom time,  $w_M = 214$  g is the fresh maximum fruit weight,  $h = 0.056$  ( $d^{-1}$ ) is, at first approximation, the conversion rate of resources into fruit mass,  $t_B$  is the bloom time.

We used this equation to compute  $w_H = 164$ g, that is weight at harvest time  $t_f$ , from 2014 and 2015 data ( $t_{f,2014} = 196$ ,  $t_{B,2014} = 59$ ,  $t_{f,2015} = 198$ ,  $t_{B,2014} = 74$ , (Vanalli et al., 2021); times are here expressed in terms of day of the year).

As suggested in Bevacqua et al. (2023), once the weight at harvest is known, the conversion rate  $h$  should be iteratively recomputed to be applied in other ripening seasons. Therefore, we recomputed  $h'$  from the known ripening date  $t_R$  for the computation of the pit hardening date ( $t_0$ ):

$$h' = -\log(w_B / (w_M - w_B) * (w_M / w_H - 1)) / (t_R - t_B) \quad (\text{B.3})$$

And  $t_0$  as follows:

$$t_0 = t_B - \log(w_B / (w_M - w_B) * (w_M / w_C - 1)) / h' \quad (\text{B.4})$$

Where  $w_C = 61$  g is the fresh fruit weight threshold for cuticle cracking.

### B.1.3 Initialization of the SEIM model

In inoculated units, we initialised fruit load as  $S_0 = 14.73$  *fruits/m<sup>2</sup>*,  $E_0 = 0.27$  *fruits/m<sup>2</sup>*,  $I = 0$  *fruits/m<sup>2</sup>* (to account for the observation according to which, in infected orchards, 1.8% of the fruits are considered as exposed; Bevacqua et al. (2023)) while unexposed units are initialized as  $S_0 = 15$  *fruits/m<sup>2</sup>*,  $E = I = 0$  *fruits/m<sup>2</sup>*. We chose an initial total fruit load of 15 *fruits/m<sup>2</sup>* since this is the average of 2014 and 2015 data in Avignon (Bevacqua et al., 2023).

### B.1.4 Weather variables

We used Safran weather reanalysis data (via the Siclima portal: Delannoy et al. (2022); Caubel et al. (2015)), which are arranged according to the Safran grid (Bertuzzi and Clastre, 2022) the same used for the metacommunity model.

Daily precipitations (“preliq\_q, meaning daily liquid precipitations 06-06 UTC”), daily mean temperature (“preliq\_q, 01-00 UTC”) and max temperature (“tinf\_h\_q”) from the 1<sup>st</sup> of January 1980 to the the 31<sup>st</sup> of December 2021 have been stored.

The Safran grid has been recently rearranged: with respect to the original grid, comprising 8602 units, Safran data managers added 379 new units, located on the coasts. This rearrangement created some mismatches with the previous one, since stone fruits have been mapped with the most recent version, while available weather reanalysis are consistent with the previous one. To solve this mismatch, we attributed to the newly introduced units, all located next to coastal environments, the weather series of the neighbouring unit that shares the longest geographical border. The IDs (consistent with the old Safran grid) of the stored variables are indicated in “ID\_safran\_grid.csv” in supplementary material.

### B.1.5 Determination of the probability of spore deposition $P_v$

A trajectory  $v$  as emitted from our HYSPLIT simulations is made up of points in the atmosphere  $x_{v,t}, y_{v,t}, z_{v,t}$ , with  $t = 0, 1, \dots, 6$ , which we linearly interpolate to obtain a continuous trajectory  $x_v(t), y_v(t), z_v(t)$ . Along a trajectory we assessed spore viability to temperature  $T$  (Bevacqua et al., 2023) and other losses, such as solar radiation, deposition and dilution, combined with an exponential kernel, as in eq. B.5:

$$P_{v0}(x_v(t), y_v(t), z_v(t), T) = e^{-\alpha t} \int_0^t e^{\frac{-r}{24T(x_v(\tau), y_v(\tau), z_v(\tau))}} d\tau \quad (\text{B.5})$$

With  $r = 1.4e + 4$   $K$  is the reciprocal of the Boltzmann’s constant,  $T(x_v(\tau), y_v(\tau), z_v(\tau))$  is the temperature in a point  $(x_v(t), y_v(t), z_v(t))$  along the trajectory  $v$ , and  $\alpha$  being chosen so that  $P_{v0}(t = 6h, T = 0) = 0.01$ . The exponent is divided by 24 because  $t$  is measured in hours. We made this choice because we wanted to mimic a progressive loss of spores due to dilution and dry deposition, assuming that viable spores become negligible after 6 hours.

To obtain the probabilities of deposition  $P_v$ , we corrected  $P_{v0}$  to remove those parts of the trajectories whose altitude is higher than the planetary boundary layer.

Eventually, we performed a last correction to account for the fact that spore deposition is directly proportional to the time spent by the air mass over the crops. Let us consider  $l_h$  the Euclidean distance between two consequent points  $(x_h, y_h)$  and  $(x_{h+1}, y_{h+1})$  of a trajectory, where  $h$  indicates the time in hours. This proportionality means that, the longer the segment  $l_h$ , the lower the probability of spore deposition in a point of that segment. Therefore, we computed  $P_{v,t}$  as:

$$P_{v,t} = P_{v0,t} \left( 1 - \frac{l_h - \min_{\forall h \in 0,1\dots 6} l_h}{\max_{\forall h \in 0,1\dots 6} l_h} \right) \quad (\text{B.6})$$

In this, way, in the shortest segment ( $\min_{\forall h \in 0,1\dots 6} l_h$ ) the value of  $P_{v,t}$  is equal to that of  $P_{v0,t}$ , since this latter is multiplied by 1, while the longest ( $\max_{\forall h \in 0,1\dots 6} l_h$ ) is multiplied for the ratio between the shortest and the longest segment ( $\min_{\forall h \in 0,1\dots 6} l_h / \max_{\forall h \in 0,1\dots 6} l_h$ ), which is less than (or equal) to 1.

### B.1.6 Connectivity matrix

Once we have obtained matrices  $W_t$ , which consider all trajectories  $v$  that started at time  $t$ , we first performed a temporal aggregation on a daily basis  $d$  (where  $t$  is intended as a discrete measure of time, while  $d$  refers to a specific day of the year) and secondly corrected to account for neighbouring units:

1. The temporal aggregation consists of averaging all the connectivity matrices  $W_t$  obtained on the same day of the year. For instance  $W'_d$  on the 15<sup>th</sup> of May (i.e.,  $W_{15-05}$ ) contains the element wise average of all the  $11 \times 4 = 44$  corresponding  $W_t$  computed on the 15<sup>th</sup> of May 2008, 2009,... 2018, computed from Lagrangian simulations launched four times a day;
2. The correction for neighbouring units accounts for geographic proximity. We consider that two bordering units may reciprocally infect even when we have not found any connection among them. We first considered all couples of 8-neighbouring units  $i$  and  $j$ , then computed the average connectivity  $\bar{w}$  among all  $W'_{d,ij}$ . We then corrected matrices  $W'$  as follow:

$$W_{d,ij} = \max(W'_{d,ij}, \bar{w}) \quad (\text{B.7})$$

$W_d$  is the matrix used in the metacommunity model. We calculated also the annual matrix  $W$ , which is the average of all  $W_t$  during all ripening season in 2008-2019, where a filter has been added to consider only rainy events in the arrival unit, which can lead to wet deposition (which means that each element of  $W$  is determined by  $W_{ij} = (1/T) \sum_t^T (r_{t,j} > 0) W_{t,ij}$ )

### B.1.7 Determination of the prior probability function

We set a tentative prior distribution of  $\theta$ , which we iteratively refined via simulations (see Godding et al., 2022).

We estimated the shape of the prior distribution of parameter  $\theta_O$ , which weights the probability of overwintering (i.e., the probability of having an inoculum at the beginning of the next season because of the presence of mummies from the previous ripening reason), by considering the observation dataset. Such a probability may be approximated by the ratio between *i*) the occurrences of a “strong” disease incidence in year  $y + 1$  in early or mid-early cultivar after a “strong” incidence in year  $y$  in the same location over *ii*) the occurrences of a high incidence in year  $y$ . This ratio is equal to  $1/3$  (2 occurrences in the observation dataset over 6). So, we started with the hypothesis that the expected value  $E[.]$  of the random variable corresponding to presence of local inoculum at the beginning of the

season ( $E[E_y(t_0) = 0.27 \text{ fruits}/m^2]$ , hereafter  $P_o$ ) is defined as follows:

$$P_o = E[1 - (1 - \tilde{P}_o)(e^{-\theta_O M_{t_f, y-1}})] = 1/3 \quad (\text{B.8})$$

Where  $\tilde{P}_o$  represents the presence of spores due to external causes, while  $1 - e^{-\theta_O M_{t_f, y-1}}$  expresses the probability of overwintering, with the density of mummies at harvest time of the precedent year  $M_{t_f, y-1}$  weighted by parameter  $\theta_O$ . We estimated  $\tilde{P}_o = 0.2$  directly from the observation dataset as the frequency of “weak” disease severities followed by a “strong” disease severity in early and mid-early cultivars of the following year in the same place (2 occurrences in the observation dataset over 10), under the simplification assumption that, due to shorter time, inoculum from overwintered mummies could have been the main cause of infection in early and mid-early cultivars.

Considering  $M_{t_f, y-1} = M_0 = 10 \text{ fruits}/m^2$  as a strong infection, we obtained:

$$E[\theta_O] = \frac{1}{M_0} \log\left(\frac{1 - \tilde{P}_o}{1 - P_o}\right) = 0.018 \quad (\text{B.9})$$

We then considered  $\theta_O$  as extracted from an exponential variable with average  $1/E[\theta_O]$ , which is  $\approx 54.8$ . The exponential variable allows us to obtain *i)* positive values, *ii)* scattered around  $E[\theta_O]$ , *iii)* allowing us to explore also few very large values, with a parameter only ( $1/E[\theta_O]$ ). While launching preliminary ABC algorithms with a reduced number of simulations, we progressively brought the value of  $1/E[\theta_O]$  up to 81, since this increased the number of accepted parameter sets.

We have several simulations spanning several order of magnitudes to find the interval in which  $\theta_E$  describes both *i)* almost no new introductions and *ii)* almost the maximum of introduction each year. Eventually, the prior distribution of  $\theta_E$  was set as an exponential of a uniform distribution between  $-8\log(10)$  and  $-5\log(10)$ .

Parameter  $\theta_L$  is to be considered as the losses threshold over which a disease incidence is considered as “strong”. From consistency with how the observation dataset was built, we assumed that this parameter should be placed between 0.2 and 0.4. We therefore considered a lognormal distribution, with parameters  $\mu_L = -1.26$  and  $\sigma_L = 0.21$ , for which 0.2 and 0.4 represent the 5<sup>th</sup> and the 95<sup>th</sup> percentiles, respectively. We selected a lognormal distribution for similar reasons that brought us to choose an exponential variable for overwintering, but with enough indications to guess two parameters instead of one. While launching preliminary ABC algorithms with a reduced number of simulations, we updated the values of  $\mu_L$  and  $\sigma_L$  respectively to 1.2 and 0.15.

### B.1.8 Training and stratified cross validation

In our training, we ran 200,000 simulations, spanning over 1996-2021, under the assumption that initial inoculum is randomly present in each unit with probability  $\tilde{P}_o$ .

We split the observations  $\Omega = 5$  times into a training set  $\phi$  (containing  $1 - 1/\Omega = 80\%$  of the observations, over which ABC was applied) and a testing set  $\psi$ , representative of the observation structure.

We formed  $\Omega = 5$  testing sets so that each one satisfies all the following criteria: *i*) it includes  $1/\Omega = 20\%$  of observation dataset; *ii*) it includes all the locations of the observation dataset (Fig. 3 of Chapter 3) and cultivars; *iii*) where possible (given the reduced degrees of freedom due to criteria *i* and *ii*), locations and cultivars are represented proportionally to the complete dataset .

We relaxed a fourth criteria, i.e. *iv*) disease incidences represented proportionally to the complete dataset, since it was not compatible with *i*-*iii*); we considered only testing sets in which proportions the ratio of “weak” and “strong” are not too distant from the proportions in the complete dataset: they should not exceed 60% and 45% of the testing set, respectively.

Criteria *i* to *iv* allow us to “stratify” the cross validation, i.e. they assure that each testing set is a good representation of the whole dataset. However, this “stratification” has the drawback that some observations repeat systematically in the different testing sets. If this repetition becomes important, to the limit where all testing sets contain the same observation, this can represent a problem, since it means that the model validation is performed against only one the same testing set, which reduces the model’s confidence. Therefore, we computed the grade of similarity among the testing sets (i.e., the ratio of observations which are identical for two sets), which is between 18.5% and 37.5%. We considered it acceptable since it never represent the largest part (its grade of similarity being always strictly lower than 50%).

In order to choose the 100 more frequent accepted sets, we weighted frequencies by the corresponding  $\kappa_{\psi,\omega}$  in split  $\omega$ .

### B.1.9 Null model: isotropic kernel

We computed matrix  $\mathbf{U}$  by computing each element  $u_{ij}$  connecting  $i$  and  $j$  (which are found  $d$  units away), as the average of all  $w_{mn}$ , where  $m$  and  $n$  are all the couples of nodes which are  $d$  units away. A representation of the kernel is given in Fig. B.7a.

The Monte Carlo analysis is performed this way:

1. the posterior distribution  $\theta^U = (\theta_E^U, \theta_O^U, \theta_L^U)$  is computed via ABC with 200.simulations;
2. Named  $\kappa_w$  and  $\kappa_u$  the vectors of  $\kappa$  computed with the prior distribution of the parameters, we computed  $D_o = \bar{\kappa}_w - \bar{\kappa}_u$ ;
3. To check if  $D_o$  is larger than it would be by chance, we created 20,000 vectors  $\hat{\kappa}_w$  and  $\hat{\kappa}_u$  by reshuffling elements of  $\kappa_w$  and  $\kappa_u$ , and computed  $\hat{D}_i$  as  $\bar{\hat{\kappa}}_{w,i} - \bar{\hat{\kappa}}_{u,i}$ ;
4. We compared the distribution of  $\hat{D}_i$  with respect to  $D_o$ .



### B.1.10 Definition of “vulnerability” and “dangerousness”

Given any unit  $j$  ( $j = 1, 2, \dots, N = 755$ ) characterised by its cultivated area  $Ac_j$ , a time horizon defined by  $y = 10$  years, a stochastic iteration  $\lambda$  of the model ( $\lambda = 1, 2, \dots, \Lambda = 100$ ), local losses  $L_{ij,y}$  (defined by Eq. 5 in Chapter 3) due to an inoculated unit  $i$  ( $i = 1, 2, \dots, N = 755$ ), vulnerability  $v_j$  and dangerousness  $d_i$  are defined by Eq.s B.10 and B.11:

$$v_j = \frac{1}{\Lambda(N-1)} \sum_{\lambda, i \neq j} L_{ij,y,\lambda} \quad (\text{B.10})$$

$$d_i = \frac{1}{\Lambda \sum_{j \neq i} Ac_j} \sum_{\lambda, j \neq i} L_{ij,y,\lambda} Ac_j \quad (\text{B.11})$$

Note that, since vulnerability computes the local losses, it is normalized by the number of cells reduced by one ( $N - 1$ ), while in dangerousness one wants to weight losses according to the cultivated surface of the other cells ( $\sum_{j \neq i} Ac_j$ ).

In these equations we imposed  $j \neq i$  so that to exclude the contribution of the first inoculated units. Otherwise, these two indices would be artificially biased toward units with large  $Ac$ .

## B.2 Supplementary Results

### B.2.1 Domain extraction

Peach surface by NUTS2 region is summarized in Tab. B.1. The final values of peach orchards surface per unit is reported in Fig. B.1e.

### B.2.2 Spatial distribution of cultivars

The probability for each cultivar (early, mid-early, mid-late, late) of being chosen in the metacommunity model is represented in Fig. B.1a-d. The sum of the probabilities for each

NUTS2	Peach surface (ha)
FR1 (Île-de-France)	0
FR2 (Champagne-Ardenne, Picardy, Upper Normandy, Centre, Lower Normandy, Burgundy)	24.8
FR3 (Nord-Pas-de-Calais)	0
FR4 (Lorraine, Alsace, Franche-Comté)	12.5
FR5 (Pays de la Loire, Brittany, Poitou-Charentes)	27.6
FR6 (Aquitaine, Midi-Pyrénées, Limousin)	629.2
FR7 (Rhône-Alpes, Auvergne)	2,007.8
FR8 (Languedoc-Roussillon, Provence-Alpes-Côte d’Azur, Corsica)	7,879.0

Table B.1 – Amount of peach surface by French NUTS2 region in 2017.

cultivar is 100%.

### B.2.3 Average connectivity matrix

The value of  $\tilde{w}$  is 0.053. A detailed version of the average air-masses driven connectivity matrix  $W$  (depicted as network in Chapter 3, Fig. 2b) is represented in Fig. B.2.

### B.2.4 ABC, stratified cross validation

The results of the ABC applied over the  $\Omega = 5$  stratified splitting of the observation dataset is reported in Tab. B.2. The overall 100 parameters sets are represented as marginalized posterior function in Fig. B.3. To facilitate interpretability of the accepted parameters  $\theta_E$  and  $\theta_W$ , the “average annual occurrences of local primary inocula” and the “average annual introductions of external inocula” are represented in Fig. B.4.

Repetition ( $\omega$ )	1	2	3	4	5
Cohen’s $\kappa_\phi$ (training set)	0.496	0.497	0.502	0.492	0.502
Cohen’s $\kappa_\psi$ (testing set)	0.046	0.086	0.289	0.409	0.161
Accepted parameter set size	67	34	19	7	18

Table B.2 – Parameter’s performances in training and testing.

### B.2.5 Model accuracy

The average correct classification rate (Fielding and Bell, 1997), i.e. the frequency of correspondences (true positives and true negatives) between the observation dataset and the model’s output, grouped by observation, is represented in Fig. B.5.

### B.2.6 The null model

Depending on the cross validation splitting  $\omega$ , the size of the accepted parameter sets varies from 10 to 38 (average = 18). Values of  $\kappa_\phi$  have an average of 0.5 (min = 0.495; max = 0.510), while  $\kappa_\psi$  has an average value of 0.11 (min = 0.07, max = 0.16). The size of the accepted parameter set is 68.

Accuracy of the parameterization of the null model per location, year and variety is reported in Fig. B.6, while accepted parameters are reported in Tab. B.3. Matrix  $U$  used in the null model is represented in Fig. B.7a.

Results of Monte Carlo analysis are reported in Fig. B.7b. The statistical distribution of the difference  $\hat{D}_i = \hat{\kappa}_{w,i} - \hat{\kappa}_{u,i}$ , where  $\hat{\kappa}_{w,i}$  and  $\hat{\kappa}_{u,i}$  ( $i = 1, \dots, 20,000$ ) contain randomly reshuffled elements of  $\kappa_w$  and  $\kappa_u$ , is always lower than the observed difference  $D_o = \kappa_w - \kappa_u$ . There is no reshuffled distances  $\hat{D}_i$  (average value =  $-3.01 \times 10^{-6}$ ; interquartile range =

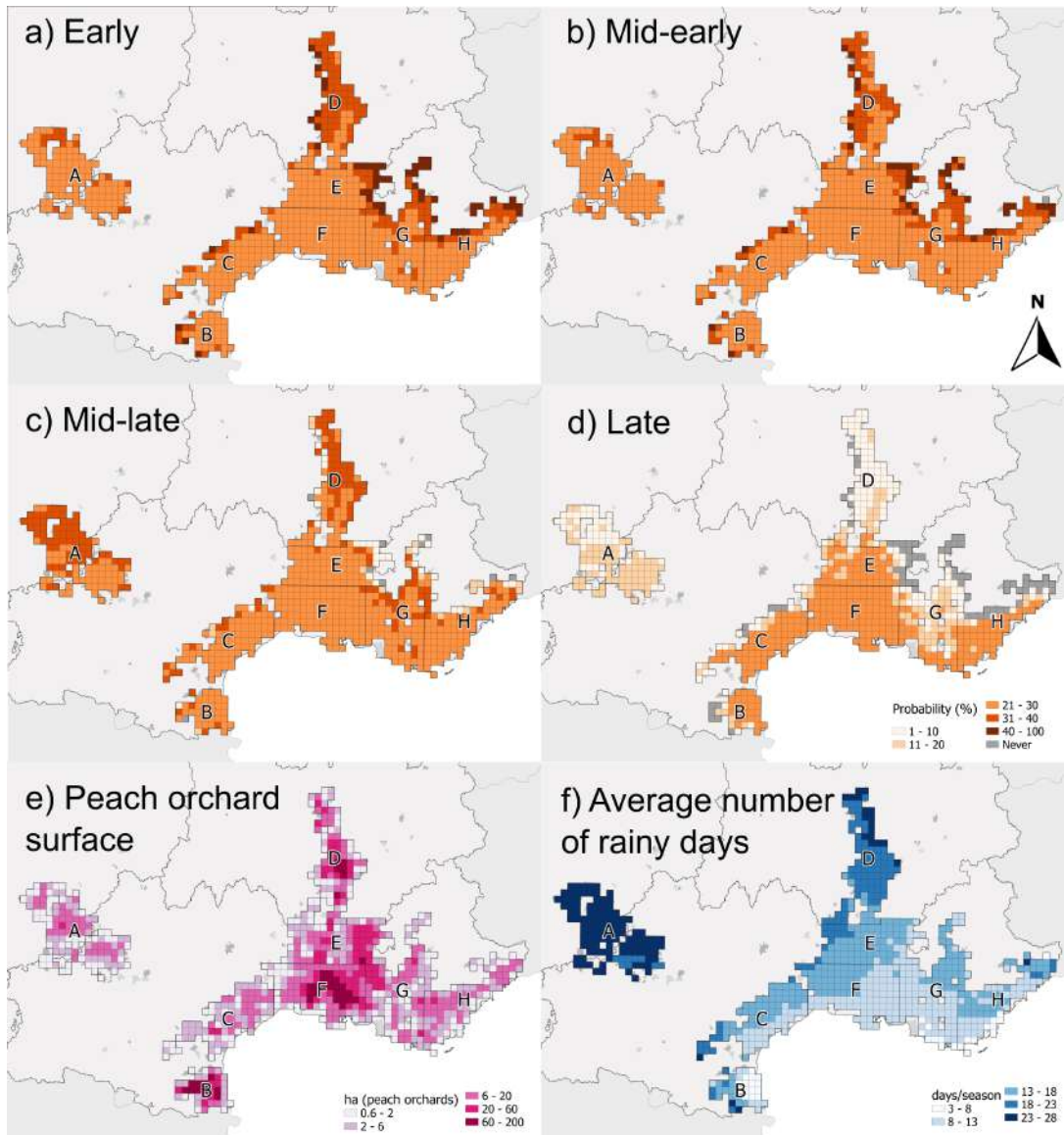


Figure B.1 – a - d) Probability of being chosen for each cultivar e) peach orchard surface by unit in *ha*; f) average number of rainy days during each ripening season. Regions of Fig. 2b of Chapter 3 are reported.

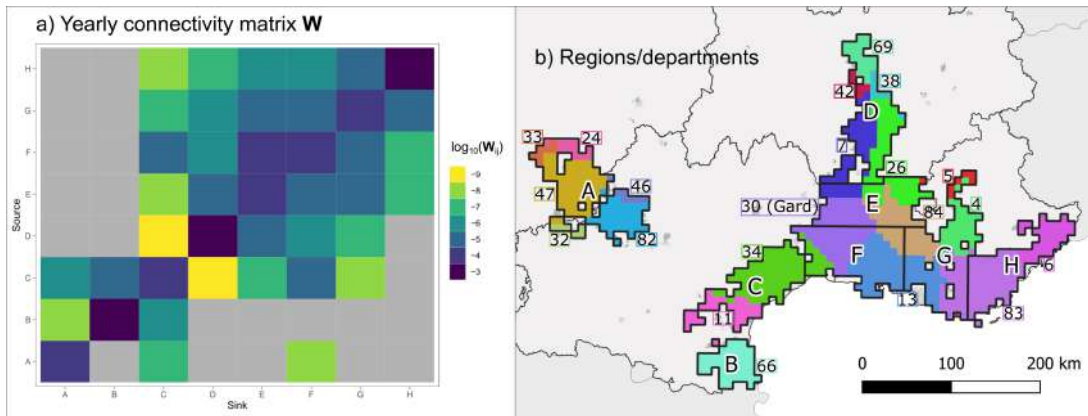


Figure B.2 – a) Average yearly connectivity matrix  $W$ , represented grouping units according to regions in panel b), where, in addition to regions, also French administrative departments are mapped with different colours. The colour scale of panel a) is built to represent the  $\log_{10}$  of each element  $W_{ij}$ . In panel b), department 30 (Gard, in violet in the center) is the one where *M. fructicola* has been detected first in 2001 (EPPO, 2023).

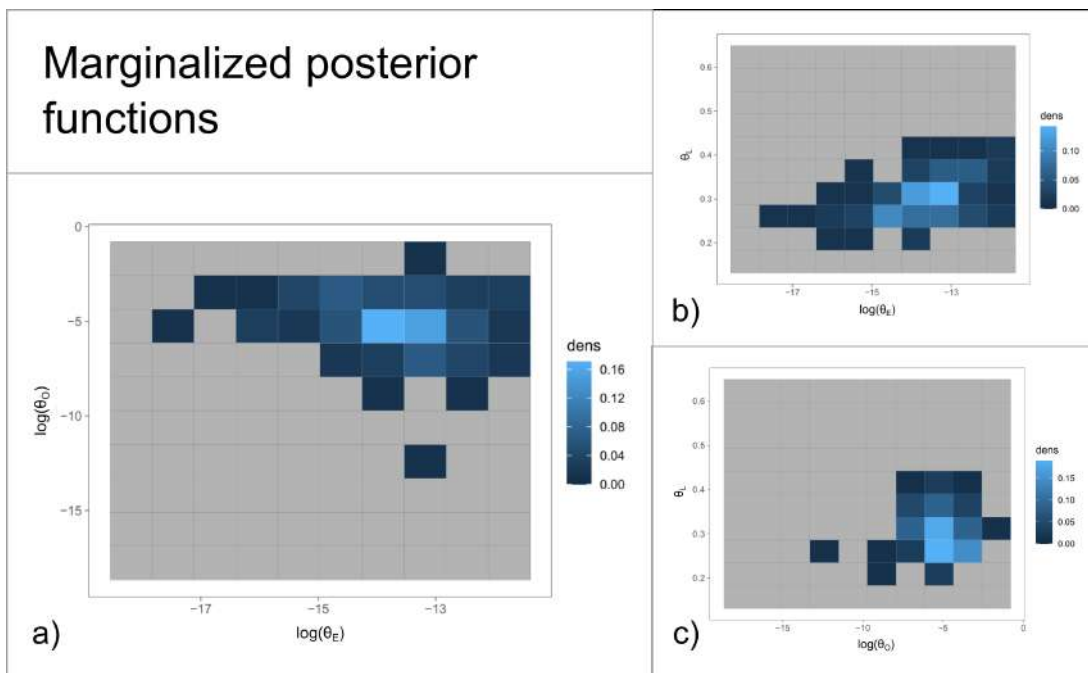


Figure B.3 – Panels a) to c) show the marginal distribution of the posterior function for all possible a couple of the three parameters  $\theta$ . Lighter the colour, higher the probability. Note that  $\theta_L$ 's axis is linear, while the others' are logarithmic. In a) the posterior distribution is normalized over  $\theta_E$  and  $\theta_O$ ; this panel is larger since it represents the parameters of the main modules of the metacommunity model (i.e., external inoculum, Fig. 3.1e, and interannual persistence of inoculum, Fig. 3.1d). In b) the posterior distribution is marginalized over  $\theta_E$  and  $\theta_L$ , while in c) the posterior distribution is marginalized over  $\theta_O$  and  $\theta_L$ .

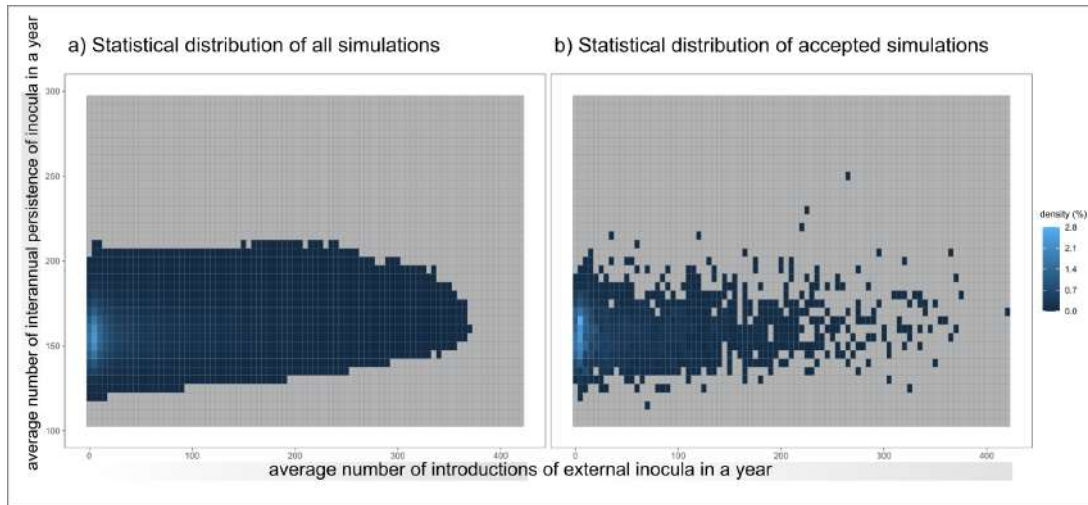


Figure B.4 – Statistical distribution of “average number of introductions of external inocula in a year” and “average number of interannual persistence of inocula in a year” in a) all the simulations and b) the simulations of the ensemble of the accepted parameter sets. Density values  $< 0.1\%$  have been omitted.

$-2.11 \times 10^{-4}$  to  $2.04 \times 10^{-4}$ ) which is larger than the observed distance between the mean performance of the full model  $D_o 9.50 \times 10^{-3}$ .

Repetition ( $\omega$ )	1	2	3	4	5
Cohen's $\kappa_\phi$ (training set)	0.497	0.510	0.496	0.498	0.495
Cohen's $\kappa_\psi$ (testing set)	0.076	0.068	0.155	0.113	0.152
Accepted parameter set size	38	12	15	10	13

Table B.3 – Parameter's performances in training and testing of the null model.

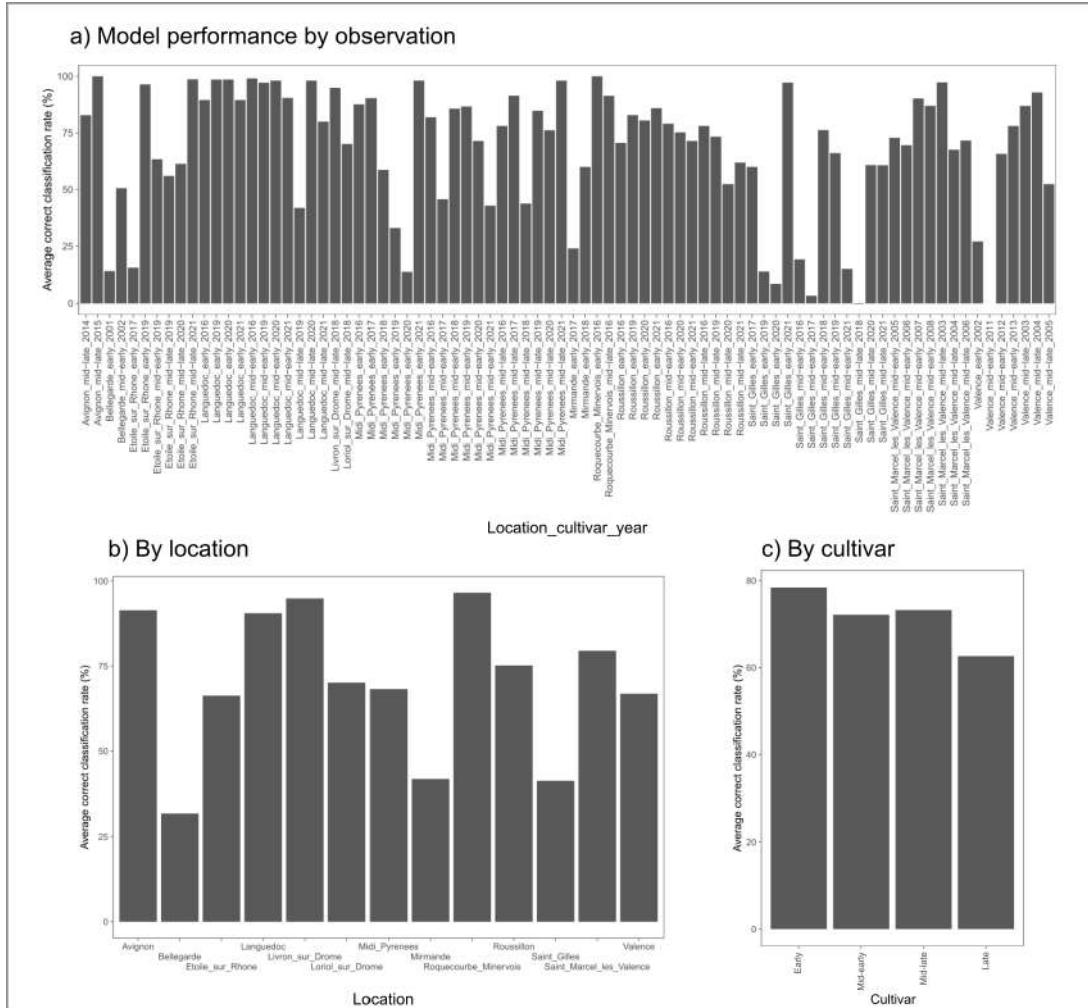


Figure B.5 – Average correct classification rate (Fielding and Bell, 1997), expressed as %, a) by observation b) by location and c) by cultivar. Note that panel a) represents the same rates as in Fig. 3 in Chapter 3.

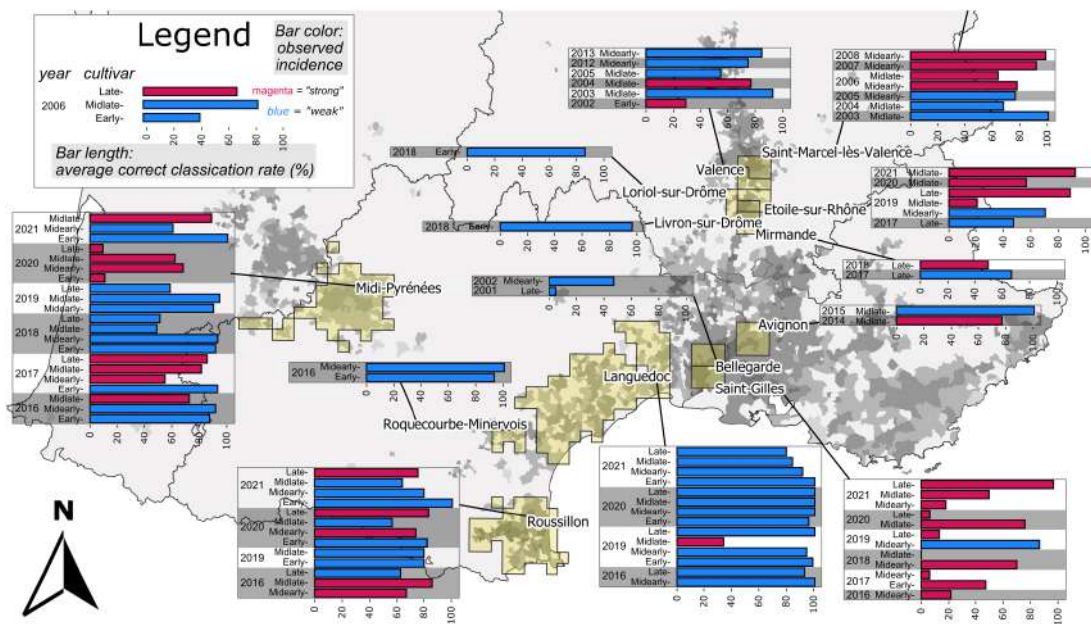


Figure B.6 – Summary of the performances of the null model throughout the domain. The length of the bars represented the average correct classification rate of the null model with the accepted parameter set. The legend is equivalent of that of Fig. 3 of Chapter 3.

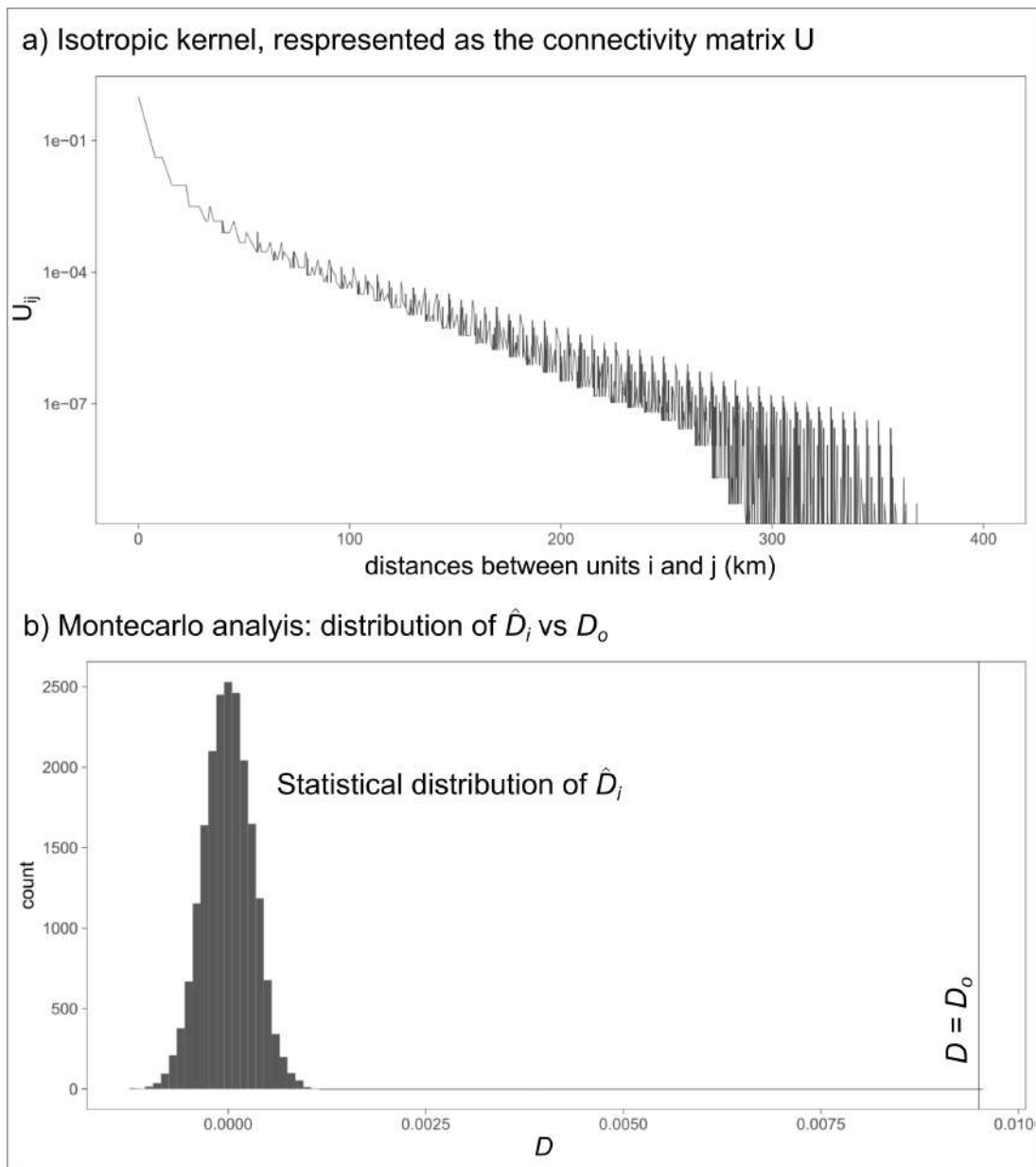


Figure B.7 – Null model: a) representation of each element of the matrix  $U_{ij}$  as a function of the distance (in km) between units  $i$  and  $j$ ; b) representation of the Monte Carlo analysis. The statistical distribution of the difference  $\hat{D}_i = \hat{\kappa}_{w,i} - \hat{\kappa}_{u,i}$ , where  $\hat{\kappa}_{u,i}$  and  $\hat{\kappa}_{w,i}$  ( $i = 1, \dots, 20,000$ ) contain randomly reshuffled elements of  $\kappa_w$  and  $\kappa_u$ , is compared with the observed difference  $D_o = \kappa_w - \kappa_u$ . No element  $\hat{D}_i$  is greater than  $D_o$ .



# Appendix C

## Supplementary Information - Chapter 4

### C.1 Supplementary methods

This section provides additional details about the relation among country wheat regions size and value of the cost-benefit index (par. “Local impacts of cooperation vs country size”) and about robustness of the performances of the sentinels sets (par. “Robustness of the surveillance strategies”) and of the cost-benefit index (par. “Robustness of the costs and benefit distribution among countries”)

#### C.1.1 Local impacts of cooperation vs country size

In order to explore possible relations among surveillance effort distribution (i.e., CoopBeneficial, CoopNeutral and CoopAdverse) and country size, after having calculated  $\alpha_{c,\sigma}$ , we characterised each country by its wheat-producing surface: large (at least 45 cells), medium (between 44 and 13) and small (12 or less). Fig. C.2 reports a selection of the curves describing the relationship among country sentinel set size and aggregated coverage for some representative examples. The function  $\alpha_{c,\sigma} = f(\sigma)$  is shown as well.

#### C.1.2 Robustness of the surveillance strategies

We assessed the performances of same sentinels vector  $\mathbf{s}_c^{-T}$  and  $\mathbf{s}$  against the validation network  $\mathbf{C}_V$ . In both cases, the maximum achievable target  $\sigma$  is lower than 100% (Fig.s C.3c,d) due to the reduction of the edges of the networks. However, it is still possible to estimate the sentinel set size needed to achieve a target of  $\sigma_i = 50\%$  by country in the “Non-cooperative” strategy, which is 459 (5.9% of the nodes; Fig. C.3c). This computation excludes

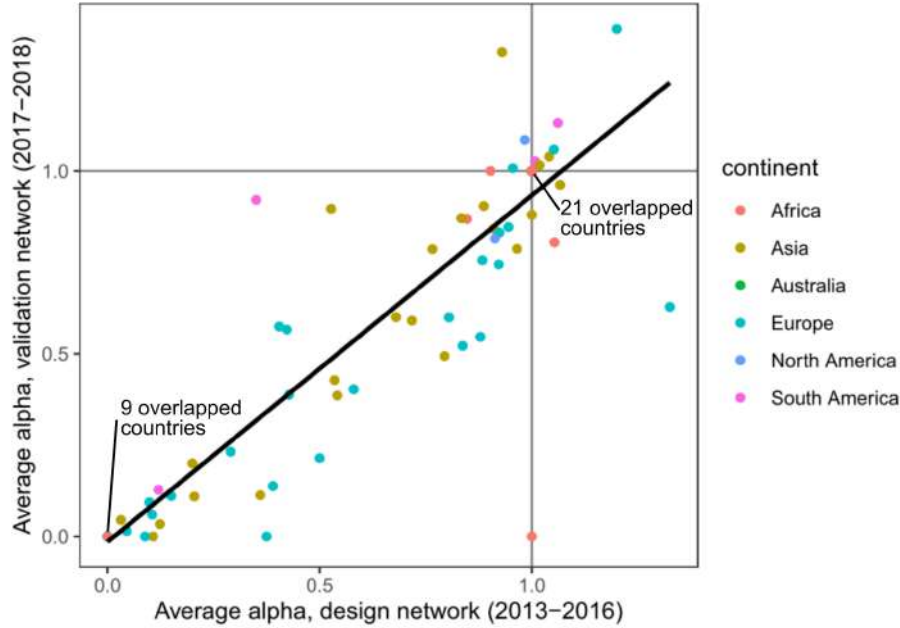


Figure C.1 – Scatter plot representing the change in  $\bar{\alpha}_c$  by country from design  $C_D$  to validating networks  $C_V$ . The linear correlation coefficient among  $\bar{\alpha}_c$  is 0.89, and the resulting p-value via the Pearson test is  $< 2.2e - 16$ . 77 out of 87 countries (i.e., 89% of the countries) keep qualitatively the same performances (i.e.,  $\bar{\alpha}_c$  is either  $<$ ,  $=$ , or  $>$  1 in both design  $C_D$  and validation network  $C_V$ ).

Finland and Tunisia, for which the maximum aggregated coverage in  $C_V$  are  $\sigma_{Finland} = 25\%$  and  $\sigma_{Tunisia} = 47\%$ . Again, due to the discrete nature of the coverage, 459 sentinels are associated to a global coverage of 54%. In the “Cooperative” strategy, 459 sentinels would enable to increase the observed domain up to 75%, while the  $x_{54}$  and  $x_{50}$  are 138 (1.8% of the nodes) and 114 (1.5%) respectively, i.e. around a quarter of the sentinels needed in the “Non-cooperative” strategy (Fig. C.3d).

### C.1.3 Robustness of the costs and benefit distribution among countries

In a similar way compared to the previous paragraph, we assessed the robustness of the outcomes of the distribution of costs and benefits of a cooperative surveillance against the network  $C_V$ . A visual comparison is given by observing Fig.s C.4 and C.5.

In terms of wheat production (FAO, 2021), passing from the design to the validating network implies that the fraction of CoopAdverse wheat producing countries slightly increases from 23% to 26%, CoopNeutral countries increase from 6% to 7% while the majority of the countries remain CoopBeneficial (from 71% to 66%, Fig. C.4a). We observe that 77 out of 87 countries (i.e., 89% of the countries) keep the same label as CoopBeneficial, CoopAdverse and CoopNeutral. The linear correlation coefficient in the scatterplot which compares  $\bar{\alpha}_i$

computed with  $C_D$  and  $C_V$  is 0.89 (Fig. C.1), and the resulting p-value via the Pearson test is  $< 2.2e - 16$ .

## **Bibliography**

FAO (2021). World food and agriculture—statistical yearbook 2021. *World Food and Agriculture-Statistical Yearbook*.

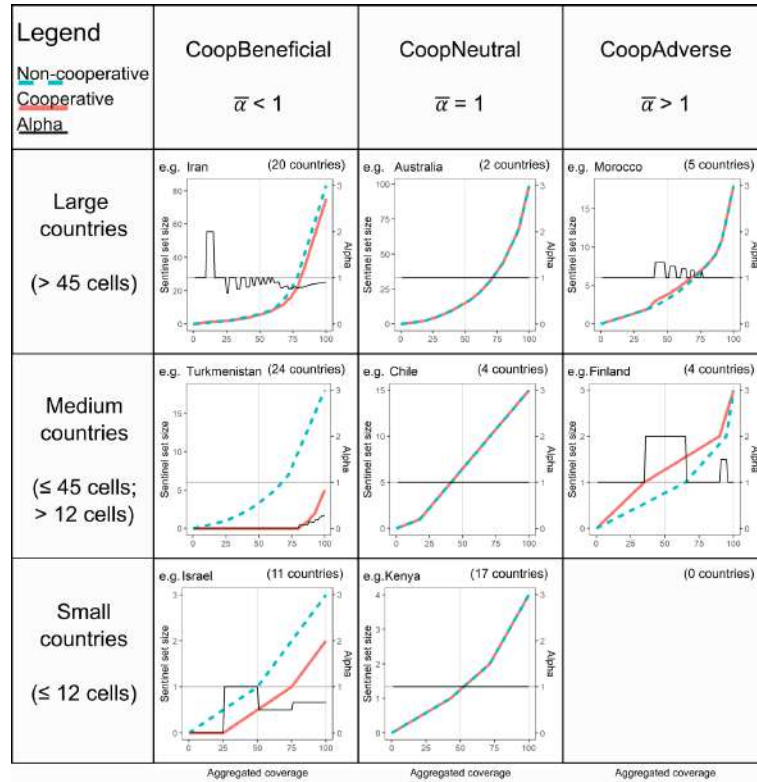


Figure C.2 – Patterns of surveillance effort vs sentinel set size area investigated for different country size (rows) and values of  $\bar{\alpha}$  (columns) for the two surveillance strategies. 63% of the countries is CoopBeneficial, and among them in 54% the “Cooperative” curve is always below the “Non-cooperative” ( $\alpha \leq 1 \forall \sigma$ ). For 27% of the countries, the two lines correspond, indicating no difference in terms of surveillance effort between the two strategies (central column). Eventually 10% are CoopAdverse (for 5% of the countries  $\alpha \geq 1 \forall \sigma$ ). Few countries (15%) vary their  $\alpha$  depending on  $\sigma$  (Such as Iran, in the first row and first column). For each combination of country size and pattern of  $\bar{\alpha}$ , curves representing sentinel set size (first y axis) and  $\alpha$  (second y axis) versus  $\sigma$  (x axis) for a typical country are displayed, as well as the number of countries with a similar qualitative pattern.

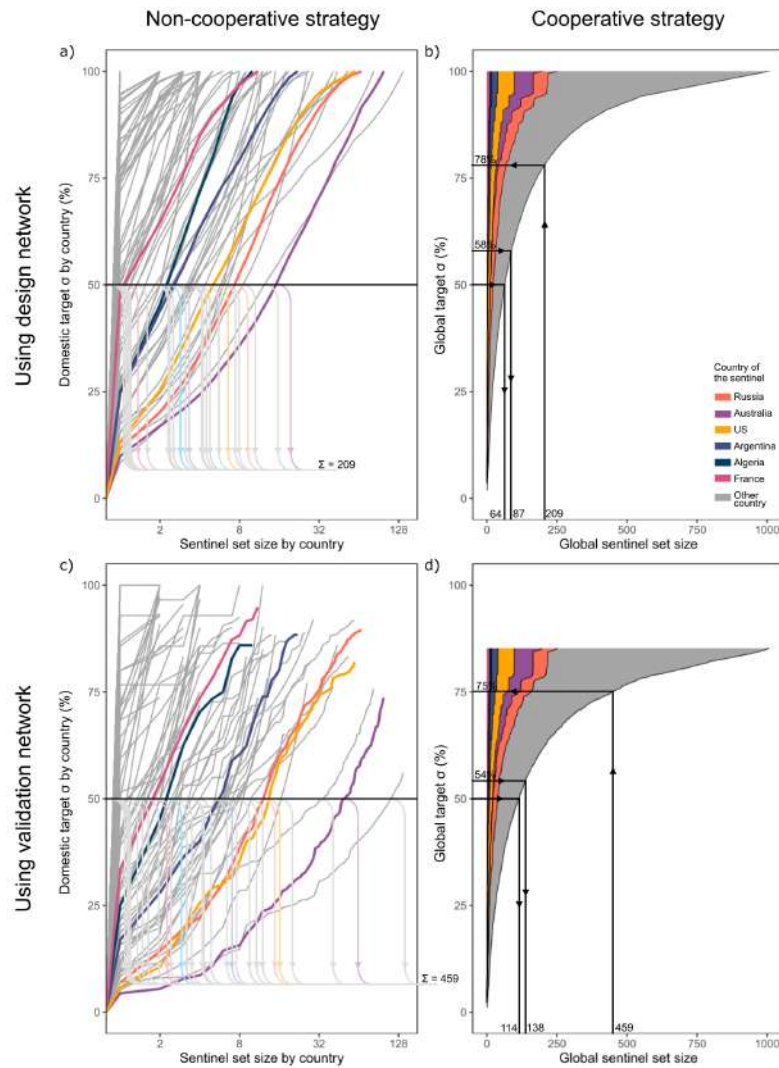


Figure C.3 – Curves describing the relationship between increasing sentinel set size and relative aggregate coverage under the “Non-cooperative” or “Cooperative” strategies (by columns) and using the design or validation networks (by rows). The first row is identical to Fig. 2 in Chapter 4, except for panel b), where we highlight the fact that 87 sentinels are needed to achieve a target of 58%, corresponding to the effective aggregated coverage obtained with 209 sentinels in the “Non-cooperative” scenario. In the second row, the performances of the same sentinels set are assessed using the validation network  $C_V$ . In c) the “Non-cooperative” strategy is reported; the number of sentinels needed to achieve a target of at least  $\sigma = 50\%$  by country increases to 459 (Finland and Tunisia excluded). In d), the “Cooperative” strategy, where worldwide coverage of  $\sigma = 50\%$  needs 114 sentinels, and 459 sentinels would allow a global coverage of  $\sigma = 75\%$ . A target of 54% (corresponding to the aggregated coverage obtained with 459 sentinels in the “non-cooperative” scenario) needs 138 sentinels.

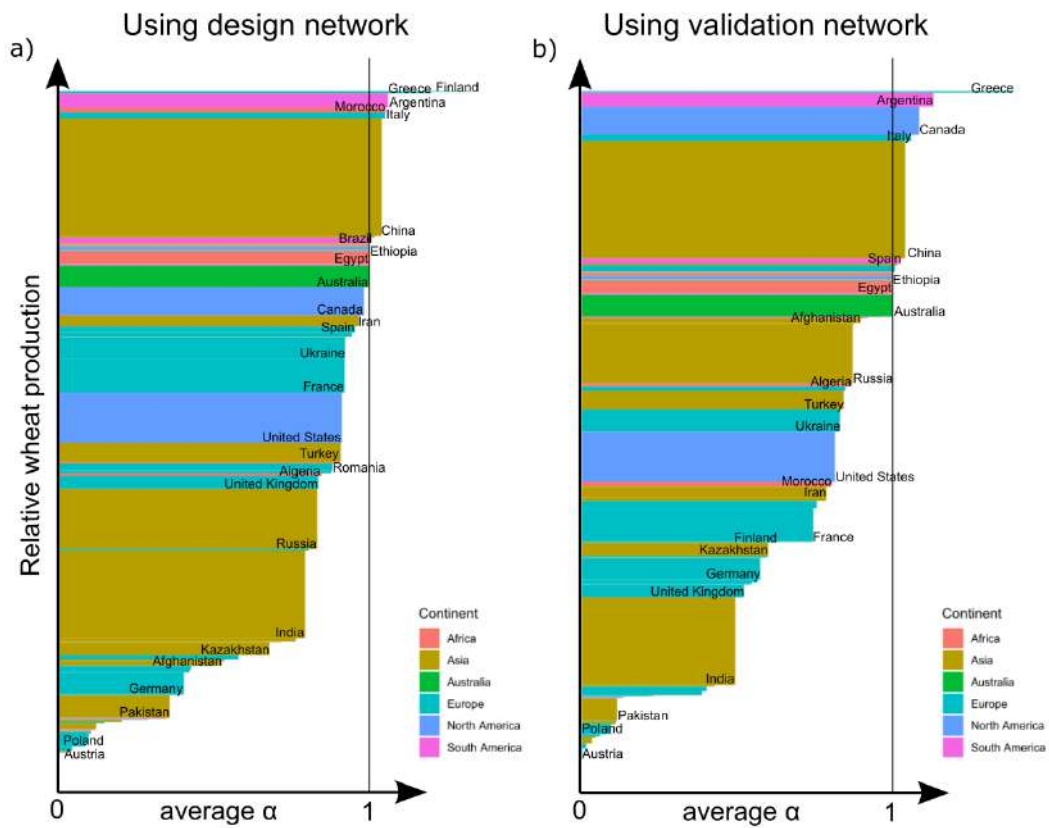


Figure C.4 – Barplot of  $\bar{\alpha}_c$  all over the world. Each country is represented by a rectangle where the base is proportional to  $\bar{\alpha}$  and the height to wheat production in 2010 - 2020 (FAO, 2021). In a)  $\bar{\alpha}_c$  is calculated against the design network, in b) against the validation network.

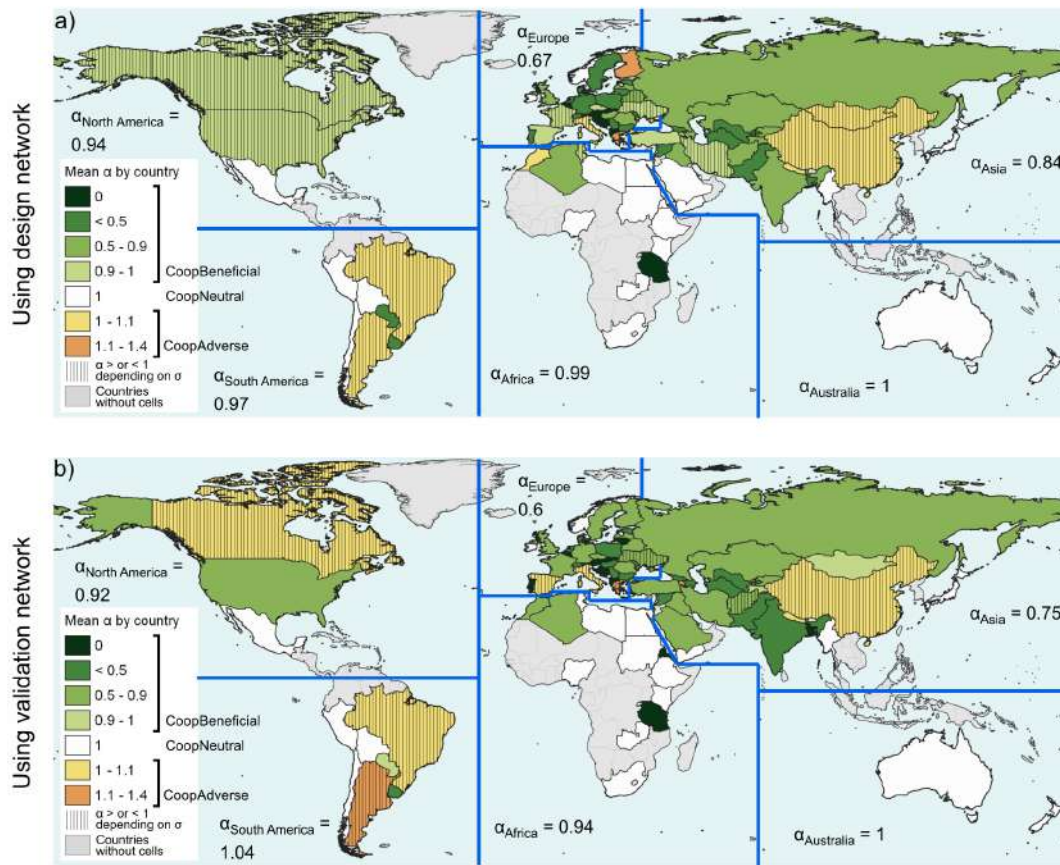


Figure C.5 – Map of  $\bar{\alpha}_c$  all over the world. On the top a) surveillance performances are assessed against the the design network, while on the bottom b) surveillance performances are assessed against the the validation network. Average values by continents, weighted by country wheat production 2010 - 2020 (FAO, 2021) are displayed.

## **C.2 The published article**



## RESEARCH ARTICLE

# Global benefits and domestic costs of a cooperative surveillance strategy to control transboundary crop pathogens

Andrea Radici<sup>1</sup>  | Davide Martinetti<sup>1</sup> | Daniele Bevacqua<sup>2</sup><sup>1</sup>BioSP UR 546, INRAE, Avignon, France<sup>2</sup>PSH UR 1115, INRAE, Avignon, France**Correspondence**Davide Martinetti, BioSP UR 546, INRAE,  
228 route de l'Aérodrome, CS 40509 Domaine  
Saint Paul - Site Agroparc, 84914, Avignon,  
France.Email: [davide.martinetti@inrae.fr](mailto:davide.martinetti@inrae.fr)**Funding information**French National Research Agency,  
Grant/Award Number: 20-PCPA-0002; INRAE**Societal Impact Statement**

Surveillance of plant pathogens is usually designed according to country boundaries. Benefits of a global surveillance system to tackle long-distance dispersed crop pathogens are unquantified. Here, a 'non-cooperative' and a 'cooperative' strategy are compared in terms of minimizing the surveillance effort to achieve given domestic and global targets. Although a 'cooperative' strategy is always more suitable, impacts of its adoption are not equally distributed among countries. Medium-sized countries in central Europe and Asia would benefit the most from reducing the domestic effort, whereas others would need to deploy more sentinels than they would place in their own interests.

**Summary**

- Transboundary diseases are extremely complex to control and can cause global socio-economic damage. In the context of crop protection, surveillance strategies are usually designed according to country boundaries, regardless of the spatial scale of the spread of the disease.
- In this study, we investigate the suitability of this scale for surveilling long-distance dispersed pathogens. We use an epidemic network describing worldwide potential transport of *Puccinia graminis*, the causal agent of stem rust of wheat, modelled in a previous work. Based on network properties, we conceive two strategies for prioritizing areas to be monitored for the presence of the disease, either cooperative or each country alone, and we compare their performances in terms of minimizing the effort deployed in achieving given surveillance targets at global and domestic level.
- We find that a cooperative strategy is more efficient at the global scale. However, its adoption implies a heterogeneous geographic distribution of surveillance effort-related costs and benefits. Medium-sized countries in central Europe and Asia would benefit the most; on the other hand, countries placed in important spreading pathways should deploy more surveillance effort than they would place without cooperation. Among the major wheat producers, China is the only country that may have a cost from a cooperative strategy, whereas India, Russia, the United States, France and Ukraine would have the most benefits.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Plants, People, Planet* published by John Wiley & Sons Ltd on behalf of New Phytologist Foundation.

- The acknowledgement of how costs and benefits of a global governance would be shared among countries is needed to gain unanimous support for an international cooperative surveillance system.

**KEYWORDS**crop protection, long distance dispersal, network, *Puccinia graminis*, transboundary surveillance**1 | INTRODUCTION**

The issue of surveillance of transboundary diseases, hereinafter intended as infectious diseases whose rapid spatial spread is likely to concern more than a country, has recently come in the spotlight due to the Covid-19 pandemic (Chinazzi et al., 2020; Dhama et al., 2020; Mohamed et al., 2020; Soubeyrand et al., 2020). New outbreaks of such diseases (Brockmann & Helbing, 2013; Saunders et al., 2019), as well as biological invasions of alien species (Diagne et al., 2021), are hardly predictable events. They can be shaped by different dissemination pathways (human transportation, commodity shipping, animal vectors or atmospheric agents) and cause socio-economic and health issues. Furthermore, lack, mismatch or delay in the communication of first detection among countries, together with uncoordinated control measures, may lead to inefficient management (Carvajal-Yepes et al., 2019; Thompson et al., 2016). Notably, the threat posed by airborne crop pathogens represents a paradigmatic case of transboundary spread (Corredor-Moreno & Saunders, 2020; Isard et al., 2005; Xing et al., 2020). The risk of large losses in food production due to unexpected outbreaks has prompted researchers and institutions to explore international surveillance systems to timely tackle the diffusion of the most alarming crop pathogens (Carvajal-Yepes et al., 2019; Park et al., 2011). The spatio-temporal persistence of large-scale seasonal movements, such as the well-known *Puccinia* pathway from Mexico to Canada (Aylor, 2003; Brown & Hovmøll, 2002), has recently emerged as a major source of inspiration for devising such innovative surveillance systems (Allen-Sader et al., 2019; Meyer et al., 2017; Radici et al., 2022; Suttrave et al., 2012). In spite of such efforts, standard surveillance of transboundary crop diseases has frequently been performed according to country boundaries, without a cooperative perspective, regardless of the actual scale of spread of the disease, lacking international, and timely, communication of first detections (Carvajal-Yepes et al., 2019; Park et al., 2011; Ristaino et al., 2021). Yet, benefits from a possible general reduction of surveillance effort of a global, cooperative and communicative strategy (Thompson et al., 2016) over a non-cooperative one, that is, each country alone, have never been quantified in the case of long-distance dispersed pathogens.

In this study, we investigate to what extent, and under which conditions, country boundaries represent a suitable scale for surveillance of long-distance dispersed crop pathogens and whether international cooperation would make crop protection more effective. We use stem rust of wheat, caused by *Puccinia graminis*, an airborne fungal pathogen whose spores can be transported over long distances by wind

(Levetin, 2015), as a case study. In the majority of wheat-producing countries, the presence of this pathogen has been controlled by the use of resistant cultivars and the eradication of its secondary host, *Berberis vulgaris*, which enables overwintering in temperate regions. This pathogen reappeared in western Europe after several decades of absence (Barnes et al., 2020; Corredor-Moreno & Saunders, 2020; Saunders et al., 2019) and is considered a threat to global food security due to the rapid spread of virulent races through a worldwide distributed host. In a recent article, we retraced its global epidemic network across worldwide wheat-producing countries (Radici et al., 2022). In the present study, we use this epidemic network to conceive two surveillance strategies, a 'non-cooperative' one, representing a within-boundary scenario with no collaboration and communication between countries, and a 'cooperative' one, where countries collaborate surveilling each other and timely communicate the detection of the disease. We compare their performances in terms of surveillance effort needed to achieve given targets both at the global and domestic scales.

**2 | MATERIALS AND METHODS****2.1 | The worldwide *Puccinia* epidemic network**

In order to evaluate the performances of different surveillance strategies, we used the epidemic networks obtained in a previous study. Here, we present a summary of the methodology proposed there. In Radici et al. (2022), we simulated worldwide transport of *P. graminis* spores among wheat-producing countries, obtaining a time-varying directed and weighted connectivity network **W**. In **W**, the 7814 nodes represent  $0.5^\circ \times 0.5^\circ$  cells ( $\approx 2000 \text{ km}^2$ ) in wheat-producing countries, whereas edges represent likely air-mass connections among cells, computed at a time resolution of 6 h for the time span 2013–2016. More specifically, each weighted edge  $w_{ijt}$  of **W** is computed in such a way to account for the likelihood of air-mass trajectories (computed via NOAA's HYSPLIT model; Draxler & Hess, 1998), which potentially disseminate spores from a release node  $i$  to an arrival node  $j$  at time  $t$ . In both  $i$  and  $j$ , host availability and favourable environmental conditions (for sporulation and/or infection) are determined via a climate-dependent suitability model and validated via a comparison with cropping calendar from the FAO country briefs (FAO, 2021a). Seventy-two-hour trajectories (Meyer et al., 2017) are filtered according to different criteria (rain washout, cumulative UV radiation, flight duration and altitude) to exclude those air-mass movements that are less likely to lead to an effective spore transport event.

We then projected this time-varying epidemic networks in a static, directed and binary design network  $\mathbf{W}_D$ , generated by considering only recurring connections, that is, occurring (i) at least once a year and (ii) at least three times over the 4-year interval 2013–2016 (i.e.  $\geq 75\%$  of the years). Network  $\mathbf{W}_D$  identifies only highly likely direct spore dissemination events on a seasonal timescale.

## 2.2 | Surveillance strategy design

We further considered the problem of establishing a reduced set of *sentinels*, nodes where the presence of the pathogen is systematically monitored (i.e. the surveillance effort), that should guarantee the largest aggregated coverage of the domain (i.e. the surveillance target) and provide an early-warning system for the detection of the pathogens (Radici et al., 2022). First of all, we defined the *coverage* of a sentinel as the set of nodes that points directly towards it, under the assumption that, by monitoring the presence of the pathogen in a sentinel, we can indirectly observe the possible presence in all those nodes that are pointing to it in one step. We leveraged on an iterative heuristic algorithm to determine sub-optimal solutions to the problem of finding the smallest set of sentinels  $\mathbf{s}_\sigma$  that guarantees the maximum aggregated coverage (associated with a surveillance target  $\sigma$ ).

The iterative heuristic algorithm (or ‘Set cover’) to determine sub-optimal solutions to the problem of finding the smallest set of sentinels consists in (i) finding the node associated with the largest coverage; (ii) adding this node to the sentinel set  $\mathbf{s}_\sigma$ , initially empty; (iii) labelling its coverage as surveilled and remove all the edges pointing to it; and (iv) repeating steps i–iii until the proportion of nodes in the aggregated coverage reaches the desired target  $\sigma$ . The optimal set of sentinels  $\mathbf{s}_\sigma$  is ranked by growing aggregated coverage. The size of  $\mathbf{s}_\sigma$  defines the surveillance effort  $x_\sigma$ .

We designed two surveillance strategies, a ‘cooperative’ and a ‘non-cooperative’ one. In the ‘cooperative’ strategy, the Set cover algorithm was run on all nodes of the network. By contrast, in the ‘non-cooperative’ strategy, we (i) labelled each node with the country where it is placed and (ii) ran the Set cover algorithm separately for each country by considering only the corresponding sub-block of the network. We thus obtained the optimal sentinel sets  $\mathbf{s}_{\sigma,c}^{-T}$  for each country  $c$ , where  $^{-T}$  stands for ‘without Transboundary edges’, ranked by growing aggregated domestic coverage. To compare the performances of the ‘cooperative’ and ‘non-cooperative’ strategies, we computed the number of sentinels needed to achieve different global targets (Figure 1).

## 2.3 | Measuring benefits and costs of cooperation at domestic scale

To investigate how the burden of cooperative surveillance is shared among countries, for each country  $c$ , we calculated the number of sentinels  $x_{c,\sigma,s}$  needed to achieve a domestic surveillance target of  $\sigma$

under a given strategy  $s$  ( $s = \text{‘cooperative’}$  or ‘non-cooperative’). Then, we defined the cost–benefit index  $\alpha_{c,\sigma}$  as the ratio between the number of domestic sentinels needed to achieve  $\sigma$  in the ‘cooperative’ and in the ‘non-cooperative’ strategy, for a given country  $c$ :

$$\alpha_{c,\sigma} = \frac{x_{c,\sigma,s = \text{cooperative}}}{x_{c,\sigma,s = \text{non-cooperative}}}$$

We evaluated it for  $\sigma = 1\%, 2\%, \dots, 100\%$  and then we computed the average ( $\bar{\alpha}_c$ ) by country. We ascribe to a country  $c$  the label of ‘CoopBeneficial’ if  $\bar{\alpha}_c < 1$ , ‘CoopAdverse’ if  $\bar{\alpha}_c > 1$  and ‘CoopNeutral’ if  $\bar{\alpha}_c = 1$ . After having computed  $\bar{\alpha}_c$  by country, we aggregated it by continent weighting each country’s contribution by its wheat production (FAO, 2021b) to investigate geographical heterogeneity of benefits and costs of cooperative surveillance.

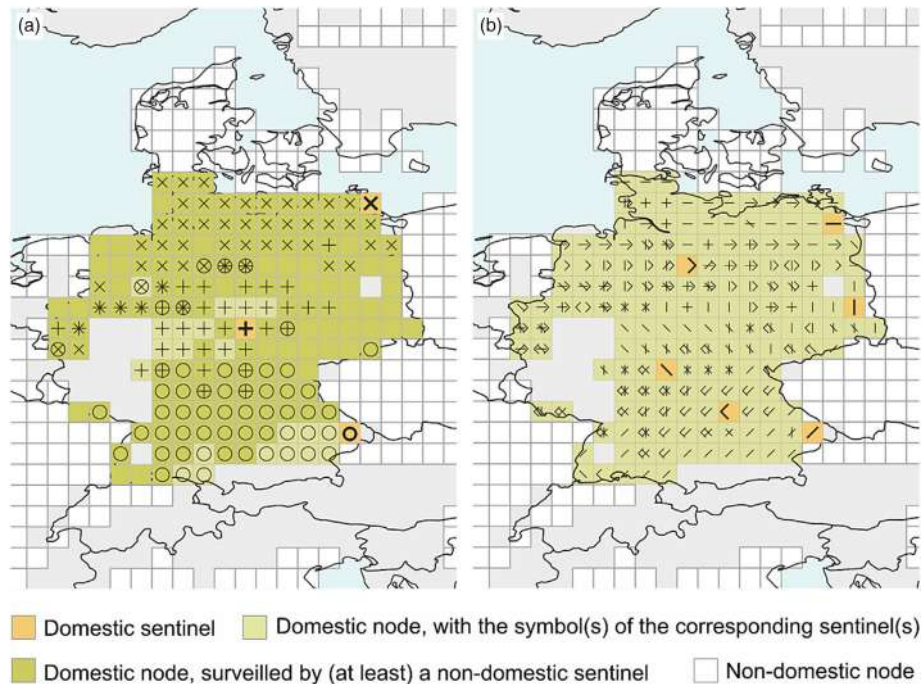
## 2.4 | Robustness of the sentinel sets

To assess the temporal robustness of the results to slight changes in the epidemic network, we set up a validation procedure of the performances of the sentinel sets. We recomputed the connectivity network  $\mathbf{W}$  on years 2017–2018 and projected it into a validation (directed, binary, static) network  $\mathbf{W}_V$ , obtained by considering only those connections occurring at least once a year both in 2017 and 2018.

We then recomputed the aggregated coverage and  $\alpha_{c,\sigma}$  of the sentinels sets  $\mathbf{s}_\sigma$  and  $\mathbf{s}_{\sigma,c}^{-T}$  using network  $\mathbf{W}_V$ .

## 3 | RESULTS

Our global epidemic network, together with the applications of the Set cover algorithm, allowed us to identify those sentinels that would best perform to detect disease presence within a certain portion of the network. Note that sentinels might not be included in the network portion that one wants to surveil. For example, if the objective is to monitor the portion of the network corresponding to all wheat-producing regions in Germany, regardless of where the sentinels are placed (the ‘cooperative’ strategy), the optimal sentinel set would comprise only three domestic sentinels (see Figure 1a). On the other hand, it would be necessary to place six sentinels if surveillance could be provided only by domestic sentinels (the ‘non-cooperative’ strategy; see Figure 1b), not contributing to transboundary surveillance. Our results indicates that, for a  $\sigma$  of 100%, Germany would benefit from a cooperative strategy as the number of domestic sentinels needed to monitor its territory would pass from 6 to 3, thus meaning a cost–benefit index of  $= 3/6 = 0.5$ . Indeed, the interpretation of the cost–benefit index is rather straightforward: if  $\alpha_{c,\sigma} < 1$ , country  $c$  requires less sentinels within its borders in the ‘cooperative’ scenario than in the ‘non-cooperative’ one for achieving the same surveillance target  $\sigma$ . If  $\alpha_{c,\sigma} > 1$ , the opposite is true, whereas if  $\alpha_{c,\sigma} = 1$ , country  $c$  needs the same number of sentinels in both the strategies for achieving surveillance target  $\sigma$ .



**FIGURE 1** A graphic example to compare surveillance strategies of transboundary crop pathogens when the surveillance target is set to  $\sigma = 100\%$  of the nodes, that is, all nodes of the network points to at least a sentinel. Square cells represent nodes, corresponding to wheat-producing regions, which can be infected by the airborne pathogen *Puccinia graminis*. (a) In the ‘cooperative’ strategy (i.e. surveillance is optimized as if there were no country borders), three domestic sentinels (orange nodes: x, o, +, surveilling light green cells), in addition to others placed abroad (which surveil dark green nodes), are needed to cover all nodes in Germany. Each node is associated with one or more symbols, each for the sentinel(s) monitoring it. (Note that the sentinel x has a domestic cover set which is also surveilled by international sentinels. Yet, in a cooperative framework its role is essential to efficiently surveil nodes out of Germany). (b) In the ‘non-cooperative’ strategy (i.e. each country optimizes its own surveillance and does not communicate the others the detection of the disease), six domestic sentinels (light green cells) are needed to surveil German nodes (light green cells). They do not contribute to transboundary surveillance.

### 3.1 | Global surveillance effort reduction due to cooperation

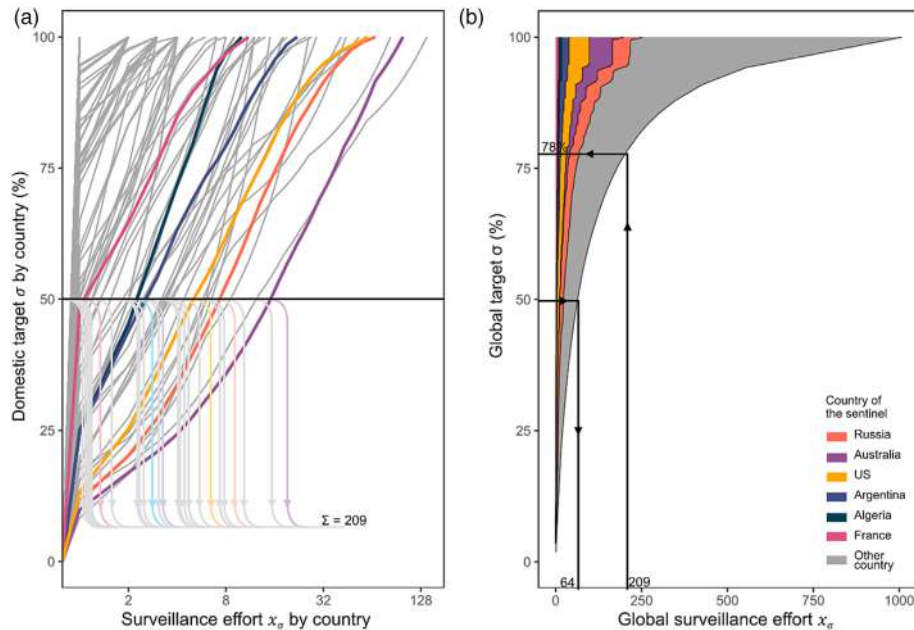
In a context of non-cooperation between countries, a coverage of half of the worldwide wheat-producing regions (i.e.  $\sigma = 50\%$ ) would be achieved by placing 209 sentinels (Figure 2a), corresponding to 2.7% of the nodes of the global epidemic network. Due to the discrete nature of each coverage, this would correspond to a worldwide target of about  $\sigma = 58\%$  (Figure 2a). Note that with the same amount of sentinels, within a ‘cooperative’ strategy, one would achieve a worldwide coverage of  $\sigma = 78\%$ . On the other hand, the coverage target of  $\sigma = 50\%$  would require only 64 sentinels (Figure 2b). An aggregated coverage of 58% would be obtained with 87 sentinels. If the coverage target were a complete coverage of the worldwide wheat-producing regions (i.e.  $\sigma = 100\%$ ), in a ‘cooperative’ framework, it would need 1007 sentinels (Figure 2b) and 1148 otherwise.

### 3.2 | Heterogeneity in the distribution of surveillance effort reduction due to cooperation

Overall, out of 87 countries, 55 (63%) are classified as CoopBeneficial, 23 (27%) as CoopNeutral and nine (10%) as CoopAdverse. In terms of

wheat production, around 71% is located in CoopBeneficial countries, 6% in CoopNeutral countries and 23% in CoopAdverse ones (Figure 3). A large variety exists in the cost–benefit index by differentiating countries with large (at least 45 nodes), medium (between 44 and 13 nodes) and small producing regions (12 or less nodes; Figures 3 and S1; see Methods S1). For 47 countries, mainly medium (e.g. Czechia or Uruguay) or large (e.g. India or Russia), the cost–benefit index is always  $\leq 1$ , thus implying an advantage in adopting a ‘cooperative’ strategy independently of  $\sigma$ . Only four countries (Morocco, Greece, Finland and Nepal) are always discouraged from adopting a ‘cooperative’ strategy. Great part of the small countries (such as Yemen or New Zealand) display  $\alpha_{i,\sigma} = 1$  for any value of  $\sigma$ , for which the two strategies are equivalent. For a few number of large (e.g. the United States, China or Iran) or medium countries (e.g. Moldova or Tunisia), the cost–benefit index is lower or larger than one depending on the value of  $\sigma$ . Their qualification as beneficial or adverse to cooperation depends on the surveillance target.

At the world scale, each continent (except Australia) has at least one CoopBeneficial, one CoopNeutral and one CoopAdverse country (Figure 4a). In North America, countries are typically CoopBeneficial, whereas South America is more balanced. Continental Europe is mainly CoopBeneficial, with some countries (Belgium, Luxembourg, Austria, Slovenia, Croatia, Bosnia and Herzegovina, Albania, North



**FIGURE 2** Increasing the surveillance target  $\sigma$  (i.e. the proportion of surveilled nodes) requires a surveillance effort  $x_\sigma$ , which varies by country and strategy (i.e. the size of the sentinel set). Each line in panel (a) represents the surveillance effort  $x_\sigma$  ( $x$  axis, in  $\log_2$  scale) needed by each country to achieve increasing domestic surveillance targets  $\sigma$  ( $y$  axis) in the ‘non-cooperative’ strategy, where each country optimizes its own surveillance strategy for monitoring airborne crop pathogens. We highlighted, via colouring, one representative country for each continent. The intersection of each line with a given surveillance target (e.g. horizontal line at  $\sigma = 50\%$ ) gives the minimum size of the sentinel set for that country (arrows) to reach that given surveillance target. The global effort can be obtained by summing all intersections (209 for  $\sigma = 50\%$ ). Panel (b) shows the global surveillance effort  $x_\sigma$  needed in the ‘cooperative’ strategy (where we run optimization as if there were no borders) to achieve increasing global surveillance targets  $\sigma$ . In this case, the target  $\sigma = 50\%$  is achieved with just 64 sentinels, whereas 209 sentinels ensure a global coverage of 78%.

Macedonia) having  $\bar{\alpha}_c = 0$ . Finland has the highest  $\bar{\alpha}_{\text{Finland}}$  of 1.3, followed by Greece ( $\bar{\alpha}_{\text{Greece}} = 1.2$ ). Asia has a composition similar to Europe, with few CoopAdverse countries (China, Mongolia, Nepal), some isolated CoopNeutral (e.g. Japan) and a majority of CoopBeneficial ones, mainly in inner parts of the continent. Africa is almost entirely CoopNeutral, with the exception of the Maghreb and Tanzania that are CoopBeneficial. Due to geographic isolation, island states such as Australia and New Zealand are CoopNeutral.

### 3.3 | Robustness of the surveillance strategies

Overall, there is good agreement between the values of  $\bar{\alpha}_c$  obtained via the design and the validation network for all countries  $c$  (correlation coefficient of 0.89;  $p$ -value  $\ll 0.001$ ; see Methods S1). A visual comparison is also provided in Figures S2–S5.

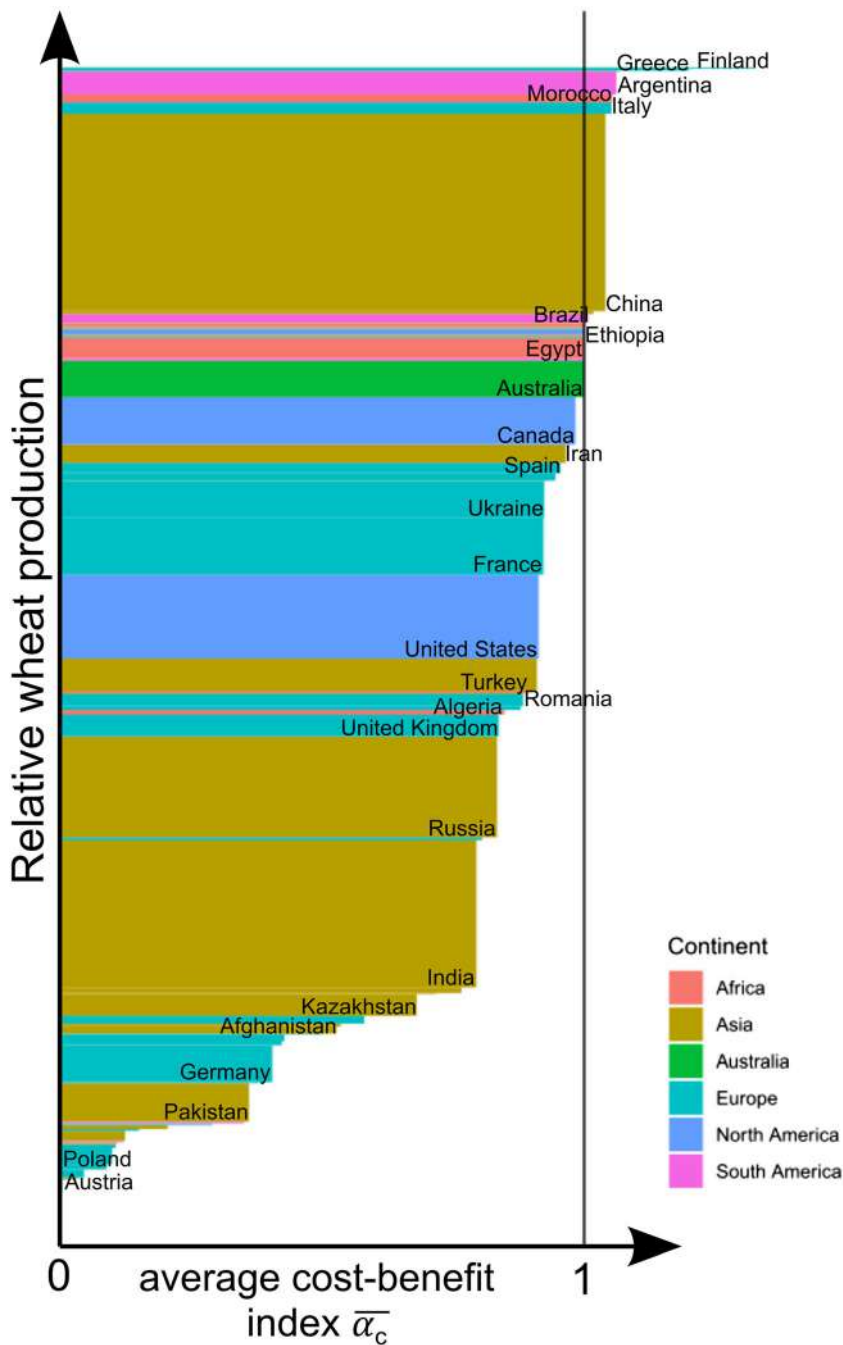
## 4 | DISCUSSION

### 4.1 | From domestic to global cooperative crop protection

As previous research has stressed, the scale of disease management should correspond to that of the spread of the disease of

interest, regardless of country boundaries (Thompson et al., 2016). We have collected evidence that, in the case of long-distance dispersed diseases, a ‘cooperative’ approach allows significant reduction in the surveillance effort needed to achieve a global coverage ( $-69\%$  and  $-12\%$  for a global coverage of  $\sigma = 50\%$  and  $100\%$ , respectively). This outcome agrees with previous studies, which underlined that neglecting long-distance connectivity leads to an underestimation of the disease spread capacity (Jeger et al., 2007).

Despite increasing evidence of a global advantage in cooperative international surveillance, crop surveillance design is still mostly dictated by country boundaries, rather than the actual scale of the pathogen spread (Carvajal-Yepes et al., 2019; Thompson et al., 2016). The mismatch between optimal and actual scale of action affects also other kinds of transboundary natural threats, such as biological invasions by alien species. In this regard, Diagne et al. (2021) recently outlined that invasion-related economic damages are projected to increase in the next decades; one reason behind the inertia in the implementation of international and coordinated protection strategies may lie in the underestimation of the costs by the general public, stakeholders and decision-makers. This may be particularly true in the case of airborne diseases, where the direct observation of their dispersal is actually unfeasible (Barnes et al., 2020; Jordano, 2017), and may discourage consideration by decision-makers.



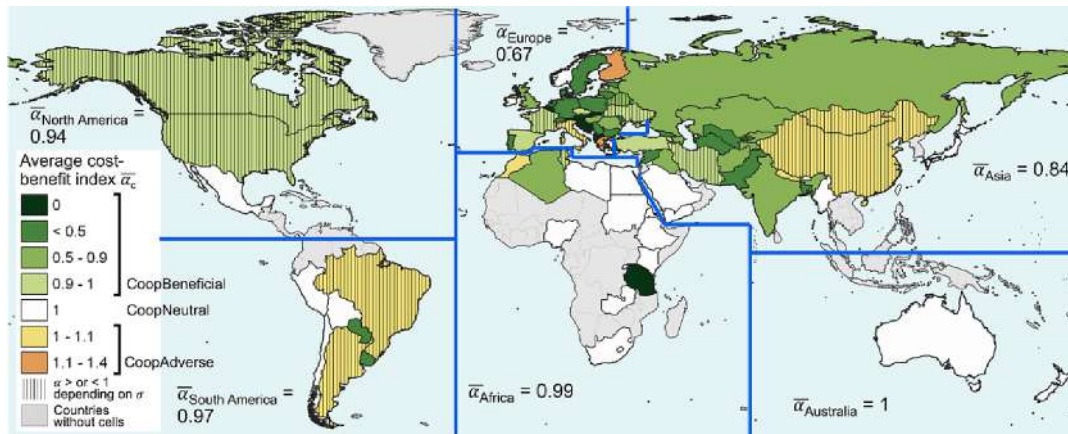
**FIGURE 3** Bar chart of the average cost-benefit index  $\bar{\alpha}_c$  (x axis) for all wheat-producing countries considered in the study. The cost-benefit index investigates how the burden of cooperative surveillance is shared among countries. Each country is represented by a rectangle where the base is proportional to  $\bar{\alpha}_c$  and the height is proportional to wheat production in 2010–2020 (y axis) according to (FAO, 2021b). Countries with  $\bar{\alpha}_c < 1$ , such as the United States, Russia or India, benefit from cooperative surveillance and are labelled as CoopBeneficial, that is, in a ‘cooperative’ scenario would need, on average, less sentinels than in a ‘non-cooperative’ to surveil their wheat production regions against airborne crop pathogen *Puccinia graminis*. On the opposite,  $\bar{\alpha}_c > 1$  identifies CoopAdverse countries, such as China. CoopNeutral countries, such as Australia, are indifferent towards cooperation ( $\bar{\alpha}_c = 1$ ).

## 4.2 | Network thinking in crop surveillance

The use of networks to support crop protection strategies has been largely advocated in recent studies (Garrett et al., 2018; Jeger et al., 2007; Parnell et al., 2017; Shaw & Pautasso, 2014; Suttrave et al., 2012). One advantage of networks is that they are ‘asemantic’, that is, they can represent whatever relationship, contact or flow mediated by different means (air masses as well as human transportation (e.g. Brockmann & Helbing, 2013) or animal trade (e.g. Bernini et al., 2019) in a topological space, which can correspond to the physical one. In the most simplistic way, crop protection strategies rely on

the identification of the nodes of the network that most contribute to spread the disease, or those that, if successfully treated, would reduce the disease size. Other methods rely on the identification of certain recurrent network patterns, where the disease spread is the fastest (Chadès et al., 2011). Concerning surveillance, relevant nodes correspond to those that may allow early disease detection if systematically monitored (Holme, 2017; Neufeld et al., 2018; Suttrave et al., 2012).

Despite the risk of incurring local minima, we used the Set cover algorithm to prioritize nodes to be monitored, that is, sentinels. Set cover iteratively selects the node associated with the highest coverage, solving the otherwise unsolvable Set cover problem in finite time.



**FIGURE 4** Global map of the average cost–benefit index  $\bar{\alpha}_c$  by country. Average values by continents (identified by blue lines) weighted by country wheat production 2010–2020 are also displayed. Europe and Asia, and in particular their innermost countries, display the lowest values of  $\bar{\alpha}_c$  (they are CoopBeneficial, i.e. they benefit from cooperative surveillance). Insular countries (Australia, New Zealand, Japan) or those with limited wheat-producing surface (mostly African countries) tend to be CoopNeutral. Few countries, often located along or at the end of dissemination pathways (Finland, Argentina, China), are CoopAdverse (i.e. in a cooperative scenario, they would need to deploy more sentinels than they would place in their own interests).

This algorithm only ensures that a node is surveilled by at least one sentinel. A less error-prone procedure may request that nodes are surveilled by at least  $n > 1$  sentinels. This would increase the reliability of the sentinel set by reducing the risk of imperfect surveillance (Chadès et al., 2011), but consequently increasing the surveillance effort. Furthermore, in our exercise, we assume that the risk of emergence of new strains (made possible by the alternate host *B. vulgaris*, which allows sexual recombination of *P. graminis*), the costs of surveillance, distribution of resistant varieties and crop management practices are the same in all the nodes. Relaxing these assumptions would ask for a different modelling framework, referable to a multi-constrained and multi-objective problem (such as a multi-dimensional knapsack problem; Kulik & Shachnai, 2010), with increasing complexity of the solution with respect to that of the Set cover algorithm.

This algorithm assumes that sentinel locations are chosen regardless of country borders, although it may not be the case. For these reasons, we named the solution of the above-mentioned algorithm as the ‘cooperative’ strategy, and we built a second strategy, where surveillance is designed mimicking a more realistic scenario. This strategy, named ‘non-cooperative’, differs from the previous as the algorithm is carried out each country independently of the others, which means that the Set cover algorithm is solved at the country level. In turn, coverage can be thought as a step-by-step updated version of the in-degree, that is, the number of the edges pointing to a node, penalizing those nodes whose coverage overlaps with that of nodes already labelled as sentinels. Other studies already noted that in-degree (or simply degree for undirected networks) is, as a general rule of thumb, a good proxy of both a good sentinel and a potential disease spreader (Herrera et al., 2016; Holme, 2018).

Moreover, in our work, we proposed a hybrid network and geographical approach, in which metadata are associated with network components: Each node is associated with the label of the

corresponding country, and each edge is consequently labelled as ‘transboundary’ or not. To our knowledge, this is one of the first attempts to compare non-topological surveillance strategies, that is, ‘cooperative’ and ‘non-cooperative’, and to quantify the heterogeneity in the allocation of the burden of ‘cooperative’ surveillance.

Our results thus indicate that the cooperative strategy becomes more valuable when the surveillance target is intermediate. This is mainly due to the fact that this strategy reduces overlapping among coverages. Overlapping is negligible also for the ‘non-cooperative’ strategy for moderate target of surveillance and becomes relevant for both strategies approaching  $\sigma = 100\%$ .

### 4.3 | Sharing benefits and costs of cooperation

From a global perspective, a ‘cooperative’ strategy is necessarily more efficient compared with a ‘non-cooperative’ one, because it corresponds to an optimization subjected to fewer constraints. However, it is interesting to quantify how such strategy performs against a ‘non-cooperative’ strategy at country level, because benefits and burden may not be equally shared; similarly, wheat production is valuable differently according to each country’s food system.

We found that medium-sized countries located in an inner continental position, such as in central Europe or central Asia, are associated with the lowest  $\bar{\alpha}_c$  values, because they benefit of transboundary potential transport events among a landscape dominated by wheat-producing areas. Insular countries, such as Australia, New Zealand or Japan, having no recurrent edges with other countries, are CoopNeutral. Due to the low presence of wheat, many African and South American countries are CoopNeutral. By contrast, it is more difficult to determine general characteristics for CoopAdverse countries, even keeping in mind that connections are mostly north-eastward in the Northern

Hemisphere and south-westward in the Southern Hemisphere (Radici et al., 2022). Finland and Nepal are small-medium-sized wheat-producing countries, located at the point of arrival of western-eastern European (Zadoks, 1967) and Indian (Brown & Hovmøller, 2002) ‘Puccinia pathways’, respectively. Given the relatively small size of their wheat-producing regions, they are forced to assume more sentinels in the benefits of upwind countries, whose food systems are probably much more wheat based, than they would need if left alone. By contrast, Canada, the final destination of the North American pathway, is a large wheat-producing country; hence, it would need several sentinels no matter the strategy. We may suppose that Italy and Greece, due to their location in the middle of the Mediterranean basin, may play as stepping stones for epidemics spreading northward from Africa towards central Europe (Mehta et al., 2007); furthermore, both have relatively low wheat productions; hence, they would need less sentinels if not cooperating. Brazilian and Argentinian large wheat-producing surfaces are located just poleward compared with those of their smaller neighbours (Paraguay and Uruguay, respectively). In the same way, due to the general eastward circulation in the Northern Hemisphere, Chinese wheat-producing regions might act as sink for trajectories from their western neighbours (that are, indeed, CoopBeneficial).

By averaging the cost–benefit index by continent, it is possible to highlight those continents which would benefit the most of a cooperative surveillance. Europe and Asia display the lowest cost–benefit index values (0.6–0.8), whereas for other continents, it is generally around 1. To sum up, the connectivity network of this airborne disease creates a heterogeneous distribution of costs and benefits, but Asia and Europe would certainly take advantage of an international and cooperative surveillance system (Figures 4 and S4).

The heterogeneous geographical distribution of benefits and costs of cooperation in surveillance has already been highlighted by other studies (Bacon et al., 2012) and suggests that a compensating mechanism should be set up to make it acceptable. This compensation mechanism should take into account different costs of surveillance among countries (Augustin et al., 2012). This idea can be borrowed from the socio-economic concept of ‘burden sharing’ (Sandler & Forbes, 1980; Suhrke, 1998), which is finding application in the management of environmental goods. Differentiate greenhouse gas emissions reduction in the framework of the Conference of the Parties to achieve climate targets (Ringius et al., 2002), as well as in the multi-stakeholder management of marine resources (Bennett et al., 2021), may be two notably example. Furthermore, other fields of crop protection may benefit of a network-based transboundary perspective. For example, the deployment of resistant varieties to both contain pathogens spread and delay resistance overcoming (Rimbaud et al., 2018) is another spatial optimization problem; whether it should be approached at the national or international scale is an interesting issue that can benefit from the approach proposed here.

Although our study tries to push towards a change in the perspective of governance of crop disease surveillance, we believe that proper identification of spatial distribution of costs and benefits can help facilitate international agreement for a global crop epidemic surveillance and gain support of all stakeholders.

## AUTHOR CONTRIBUTIONS

Andrea Radici, Davide Martinetti and Daniele Bevacqua designed the research; Andrea Radici and Davide Martinetti performed research; Andrea Radici, Davide Martinetti and Daniele Bevacqua analysed the data; Andrea Radici, Davide Martinetti and Daniele Bevacqua wrote the paper.

## ACKNOWLEDGEMENTS

The authors acknowledge the support of funding from the French National Research Agency (ANR) for the BEYOND project (contract no. 20-PCPA-0002) and SuMCrop Sustainable Management of Crop Health Program of INRAE that supported the work of all authors. We thank prof. Nik J. Cunniffe for valuable suggestions.

## CONFLICT OF INTEREST STATEMENT

The authors have no competing interests.

## DATA AVAILABILITY STATEMENT

The code that supports the findings of this study is available at the repository in <https://github.com/radiciandrea/PgraminisTransboundary.git>.

## ORCID

Andrea Radici  <https://orcid.org/0000-0001-5852-0015>

## REFERENCES

- Allen-Sader, C., Thurston, W., Meyer, M., Nure, E., Bacha, N., Alemayehu, Y., Stutt, R. O. J. H., Safka, D., Craig, A. P., Derso, E., Burgin, L. E., Millington, S. C., Hort, M. C., Hodson, D. P., & Gilligan, C. A. (2019). An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11), 115004. <https://doi.org/10.1088/1748-9326/ab4034>
- Augustin, S., Boonham, N., De Kogel, W. J., Donner, P., Faccoli, M., Lees, D. C., Marini, L., Mori, N., Toffolo, E. P., Quilici, S., Roques, A., Yart, A., & Battisti, A. (2012). A review of pest surveillance techniques for detecting quarantine pests in Europe. *EPPO Bulletin*, 42(3), 515–551. <https://doi.org/10.1111/epp.2600>
- Aylor, D. E. (2003). Spread of plant disease on a continental scale: Role of aerial dispersal of pathogens. *Ecology*, 84(8), 1989–1997. <https://doi.org/10.1890/01-0619>
- Bacon, S. J., Bacher, S., & Aebi, A. (2012). Gaps in border controls are related to quarantine alien insect invasions in Europe. *PLoS ONE*, 7(10), e47689. <https://doi.org/10.1371/journal.pone.0047689>
- Barnes, G., Saunders, D. G. O., & Williamson, T. (2020). Banishing barberry: The history of *Berberis vulgaris* prevalence and wheat stem rust incidence across Britain. *Plant Pathology*, 69(7), 1193–1202. <https://doi.org/10.1111/ppa.13231>
- Bennett, N. J., Blythe, J., White, C. S., & Campero, C. (2021). Blue growth and blue justice: Ten risks and solutions for the ocean economy. *Marine Policy*, 125, 104387. <https://doi.org/10.1016/j.marpol.2020.104387>
- Bernini, A., Bolzoni, L., & Casagrandi, R. (2019). When resolution does matter: Modelling indirect contacts in dairy farms at different levels of detail. *PLoS ONE*, 14(10), e0223652. <https://doi.org/10.1371/journal.pone.0223652>
- Brockmann, D., & Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342(6164), 1337–1342. <https://doi.org/10.1126/science.1245200>



- Brown, J. K. M., & Hovmøll, M. S. (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297(5581), 537–541. <https://doi.org/10.1126/science.1072678>
- Brown, J. K. M., & Hovmøller, M. S. (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297(5581), 537–541. <https://doi.org/10.1126/science.1072678>
- Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K. A., Giovani, B., Saunders, D. G. O., Kamoun, S., Legg, J. P., Verdier, V., Lessel, J., Neher, R. A., Day, R., Pardey, P., Gullino, M. L., Records, A. R., Bextine, B., Leach, J. E., Staiger, S., & Tohme, J. (2019). A global surveillance system for crop diseases. *Science*, 364(6447), 1237–1239. <https://doi.org/10.1126/science.aaw1572>
- Chadès, I., Martin, T. G., Nicol, S., Burgman, M. A., Possingham, H. P., & Buckley, Y. M. (2011). General rules for managing and surveying networks of pests, diseases, and endangered species. *Proceedings of the National Academy of Sciences*, 108(20), 8323–8328. <https://doi.org/10.1073/pnas.1016846108>
- Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Piontti, A. P. Y., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M. E., Longini, I. M. Jr., & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489), 395–400. <https://doi.org/10.1126/science.aba9757>
- Corredor-Moreno, P., & Saunders, D. G. O. (2020). Expecting the unexpected: Factors influencing the emergence of fungal and oomycete plant pathogens. *New Phytologist*, 225(1), 118–125. <https://doi.org/10.1111/nph.16007>
- Dhama, K., Khan, S., Tiwari, R., Sircar, S., Bhat, S., Malik, Y. S., Singh, K. P., Chaicumpa, W., Bonilla-Aldana, D. K., & Rodríguez-Morales, A. J. (2020). Coronavirus disease 2019–COVID-19. *Clinical Microbiology Reviews*, 33(4), e00028–20. <https://doi.org/10.1128/CMR.00028-20>
- Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I., Salles, J.-M., Bradshaw, C. J. A., & Courchamp, F. (2021). High and rising economic costs of biological invasions worldwide. *Nature*, 592(7855), 571–576. <https://doi.org/10.1038/s41586-021-03405-6>
- Draxler, R. R., & Hess, G. D. (1998). An overview of the HYSPLIT\_4 modeling system for trajectories, dispersion and deposition. *Australian Meteorological Magazine*, 47(4), 295–308.
- FAO. (2021a). FAO - Country brief.
- FAO. (2021b). *World food and agriculture—Statistical yearbook 2021*. World Food and Agriculture-Statistical Yearbook. FAO. <https://doi.org/10.4060/cb4477en>
- Garrett, K. A., Alcalá-Briseño, R. I., Andersen, K. F., Buddenhagen, C. E., Choudhury, R. A., Fulton, J. C., Nopsa, J. F. H., Poudel, R., & Xing, Y. (2018). Network analysis: A systems framework to address grand challenges in plant pathology. *Annual Review of Phytopathology*, 56, 559–580. <https://doi.org/10.1146/annurev-phyto-080516-035326>
- Herrera, J. L., Srinivasan, R., Brownstein, J. S., Galvani, A. P., & Meyers, L. A. (2016). Disease surveillance on complex social networks. *PLoS Computational Biology*, 12(7), e1004928. <https://doi.org/10.1371/journal.pcbi.1004928>
- Holme, P. (2017). Three faces of node importance in network epidemiology: Exact results for small graphs. *Physical Review E*, 96(6), 062305. <https://doi.org/10.1103/PhysRevE.96.062305>
- Holme, P. (2018). Objective measures for sentinel surveillance in network epidemiology. *Physical Review E*, 98(2), 22313. <https://doi.org/10.1103/PhysRevE.98.022313>
- Isard, S. A., Gage, S. H., Comtois, P., & Russo, J. M. (2005). Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience*, 55(10), 851–861. [https://doi.org/10.1641/0006-3568\(2005\)055\[0851:POTAPF\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2005)055[0851:POTAPF]2.0.CO;2)
- Jeger, M. J., Pautasso, M., Holdenrieder, O., & Shaw, M. W. (2007). Modeling disease spread and control in networks: Implications for plant sciences. *New Phytologist*, 174, 279–297. <https://doi.org/10.1111/j.1469-8137.2007.02028.x>
- Jordano, P. (2017). What is long-distance dispersal? And a taxonomy of dispersal events. *Journal of Ecology*, 105(1), 75–84. <https://doi.org/10.1111/1365-2745.12690>
- Kulik, A., & Shachnai, H. (2010). There is no EPTAS for two-dimensional knapsack. *Information Processing Letters*, 110(16), 707–710. <https://doi.org/10.1016/j.ipl.2010.05.031>
- Levetin, E. (2015). Aerobiology of agricultural pathogens. In *Manual of environmental microbiology* (pp. 3.2.8-1–3.2.8-20). ASM Press. <https://doi.org/10.1128/9781555818821.ch3.2.8>
- Mehta, S. V., Haight, R. G., Homans, F. R., Polasky, S., & Venette, R. C. (2007). Optimal detection and control strategies for invasive species management. *Ecological Economics*, 61(2–3), 237–245. <https://doi.org/10.1016/j.ecolecon.2006.10.024>
- Meyer, M., Cox, J. A., Hitchings, M. D. T., Burgin, L., Hort, M. C., Hodson, D. P., & Gilligan, C. A. (2017). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nature Plants*, 3(10), 780–786. <https://doi.org/10.1038/s41477-017-0017-5>
- Mohamed, K., Rodríguez-Román, E., Rahmani, F., Zhang, H., Ivanovska, M., Makka, S. A., Joya, M., Makuku, R., Islam, M. S., Radwan, N., Rahmah, L., Goda, R., Abarikwu, S. O., Shaw, M., Zoghi, S., Irtsyan, S., Ling, I., Csepkekal, O., Faten, A.-B., ... Rezaei, N. (2020). Borderless collaboration is needed for COVID-19—A disease that knows no borders. *Infection Control and Hospital Epidemiology*, 41(10), 1245–1246.
- Neufeld, K. N., Keinath, A. P., Gugino, B. K., McGrath, M. T., Sikora, E. J., Miller, S. A., Ivey, M. L., Langston, D. B., Dutta, B., Keever, T., Sims, A., & Ojiambo, P. S. (2018). Predicting the risk of cucurbit downy mildew in the eastern United States using an integrated aerobiological model. *International Journal of Biometeorology*, 62(4), 655–668. <https://doi.org/10.1007/s00484-017-1474-2>
- Park, R., Fetch, T., Hodson, D., Jin, Y., Nazari, K., Prashar, M., & Pretorius, Z. (2011). International surveillance of wheat rust pathogens: Progress and challenges. *Euphytica*, 179(1), 109–117. <https://doi.org/10.1007/s10681-011-0375-4>
- Parnell, S., van den Bosch, F., Gottwald, T., & Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: An epidemiological perspective. *Annual Review of Phytopathology*, 55, 591–610. <https://doi.org/10.1146/annurev-phyto-080516-035334>
- Radici, A., Martinetti, D., & Bevacqua, D. (2022). Early-detection surveillance for stem rust of wheat: Insights from a global epidemic network based on airborne connectivity and host phenology. *Environmental Research Letters*, 17, 064045. <https://doi.org/10.1088/1748-9326/ac73aa>
- Rimbaud, L., Papaix, J., Rey, J. F., Barrett, L. G., & Thrall, P. H. (2018). Assessing the durability and efficiency of landscape-based strategies to deploy plant resistance to pathogens. *PLoS Computational Biology*, 14, e1006067. <https://doi.org/10.1371/journal.pcbi.1006067>
- Ringius, L., Torvanger, A., & Underdal, A. (2002). Burden sharing and fairness principles in international climate policy. *International Environmental Agreements*, 2(1), 1–22. <https://doi.org/10.1023/A:1015041613785>
- Ristaino, J. B., Anderson, P. K., Bebbler, D. P., Brauman, K. A., & Cunniffe, N. J. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Agricultural Sciences*, 118(23), 1–9. <https://doi.org/10.1073/pnas.2022239118>
- Sandler, T., & Forbes, J. F. (1980). Burden sharing, strategy, and the design of NATO. *Economic Inquiry*, 18(3), 425–444. <https://doi.org/10.1111/j.1465-7295.1980.tb00588.x>
- Saunders, D. G. O., Pretorius, Z. A., & Hovmøller, M. S. (2019). Tackling the re-emergence of wheat stem rust in Western Europe. *Communications Biology*, 2(1), 51. <https://doi.org/10.1038/s42003-019-0294-9>
- Shaw, M. W., & Pautasso, M. (2014). Networks and plant disease management: Concepts and applications. *Annual Review of Phytopathology*, 52, 477–493. <https://doi.org/10.1146/annurev-phyto-102313-050229>

- Soubeyrand, S., Demongeot, J., & Roques, L. (2020). Towards unified and real-time analyses of outbreaks at country-level during pandemics. *One Health*, 11, 100187. <https://doi.org/10.1016/j.onehlt.2020.100187>
- Suhrke, A. (1998). Burden-sharing during refugee emergencies: The logic of collective versus national action. *Journal of Refugee Studies*, 11(4), 396–415. <https://doi.org/10.1093/jrs/11.4.396>
- Sutrave, S., Scoglio, C., Isard, S. A., Hutchinson, J. M. S., & Garrett, K. A. (2012). Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS ONE*, 7(6), e37793. <https://doi.org/10.1371/journal.pone.0037793>
- Thompson, R. N., Cobb, R. C., Gilligan, C. A., & Cunniffe, N. J. (2016). Management of invading pathogens should be informed by epidemiology rather than administrative boundaries. *Ecological Modelling*, 324, 28–32. <https://doi.org/10.1016/j.ecolmodel.2015.12.014>
- Xing, Y., Hernandez Nopsa, J. F., Andersen, K. F., Andrade-Piedra, J. L., Beed, F. D., Blomme, G., Carvajal-Yepes, M., Coyne, D. L., Cuellar, W. J., Forbes, G. A., Kreuze, J. F., Kroschel, J., Kumar, P. L., Legg, J. P., Parker, M., Schulte-Geldermann, E., Sharma, K., & Garrett, K. A. (2020). Global cropland connectivity: A risk factor for invasion and saturation by emerging pathogens and pests. *Bioscience*, 70(9), 744–758. <https://doi.org/10.1093/biosci/biaa067>
- Zadoks, J. C. (1967). Internationale verspreiding van schimmels. *Netherlands Journal of Plant Pathology*, 73(1 Supplement), 61–80. <https://doi.org/10.1007/BF01974423>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Radici, A., Martinetti, D., & Bevacqua, D. (2023). Global benefits and domestic costs of a cooperative surveillance strategy to control transboundary crop pathogens. *Plants, People, Planet*, 1–10. <https://doi.org/10.1002/ppp3.10379>