Simple models for complicated epidemics: Epidemiologically-based surveillance for early detection of invading plant diseases

Thomas Taylor

This thesis is submitted to the University of Salford

School of Science, Engineering & Environment

In fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

i. Acknowledgements

Pursuing this PhD award has been by far the most challenging experience of my life; It has also been the most liberating and self-assuring experience and would not have been possible without the tireless support of my family, friends, colleagues, and supervisors. I am and will remain 100% grateful.

I would like to begin by thanking Stephen Parnell, whose expertise in plant epidemiology guided my understanding of the subject from beginning to finish. The expansion of my knowledge base and my confidence in applying this knowledge is owed to Stephen. His kindness and patience throughout the PhD journey have never faltered, thank you.

Secondly, I would like to thank Katherine Yates. As a supervisor, her tireless commitment to all facets of my student experience, including my mental wellbeing and living circumstances, has given me the space I needed to complete this project. In addition, I thank Katherine for her contributions to our many supervisory meetings and her consistent and helpful feedback with my thesis.

Thirdly, I would like to thank Samantha Grant, my Defra Industry supervisor. Sam never hesitated in throwing me into Defra meetings and trusting in my capacity to lead my *Oak Processionary Moth* research. Without Sam, I would not have learned how an NPPO functions on a day-to-day basis and would not have met any of the fantastic people who make Defra happen. I thank Sam also for her guidance in my exploration of *Oak Processionary Moth* and her statistical expertise.

Fourthly, I would like to thank Robin Beck, who took on the role of supervisor late in the PhD project timeframe. Robin's insightful feedback during the writing process of the thesis was always welcome and added much value.

Fifthly, I thank Yoann Bourhis for contributing the initial epidemiological simulation model and showing me how to use it. My thesis would not be possible without your contribution.

Sixthly, I thank all my colleagues, but especially Alexander Mastin, whose feedback and support for my research during this process has left me more confident as a researcher. Thank you for your support.

Seventhly, I'd like to thank all the staff at the University of Salford. Your tireless work behind the scenes makes these projects happen.

Eighthly, I'd like to express my gratitude to all my friends and acquaintances I've met along the way. We are all going somewhere, and having the opportunity to cross paths during this PhD journey has been priceless. A special thank you to my friend Adam, who has never left my side during this PhD journey.

Ninthly, I'd like to thank my family. Without whom, I would not be the person capable of following through on such a momentous project. A special thanks to my uncle who helped me to keep a clear head. Of course, my deepest gratitude goes to my mum, I'm sure my mum will keep a copy of this PhD somewhere near to the family photographs. I've put you through this PhD journey alongside me; it will be worth it I promise. This thesis is dedicated to you.

Finally, to every person who stumbles across this thesis, I thank you for being a part of the bigger picture.

ii. Thesis Summary

Plant disease epidemiology is the study of plant pests and their related diseases in populations of plants over time and space. In plant disease epidemiology, mathematical and computer simulation modelling is often deployed to provide insight into the factors that drive spread, as well as to identify effective intervention strategies. A key intervention for the successful eradication and control of invasive plant pests is surveillance for early detection of invading populations. Models have been used to address outstanding research questions, such as what prevalence will a pest have reached when first detected and how much surveillance resources need to be allocated for early detection. This thesis addresses how generalisable early detection models are when applied to realistic epidemiological scenarios. Here I show that epidemiological parameters including the dispersal ability of the pest, landscape heterogeneity, detection assay sensitivity and surveillance intensity influence the degree to which a simple epidemic model, termed the 'rule of thumb', can be used to predict detection-prevalence in complex epidemics. I also apply both the rule of thumb and a spatially-explicit stochastic epidemiological model to the case study of *Oak Processionary Moth (Thaumetopoea processionea*). These results indicated in particular that the rule of thumb is less accurate for short wavefront (spatially-compressed epidemic spread), highly virulent (infectious) pests where the frequency between surveillance rounds is long. These findings show that the rule of thumb benefits from increased surveillance frequency because the rule of thumb assumes constant sampling efforts in a non-spatial context with 100% detection rate. The findings also indicate that the effects of landscape heterogeneity are largely mitigated in the context of high distance pest dispersal and that a modification to the rule of thumb can increase generalisability drastically. I anticipate that the findings of this thesis will demonstrate that early detection models can be broadly applicable to a range of diseases

when dispersal distances of pests are high, virulence is low, landscapes are largely homogenous and detection sensitivity is high.

iii. Table of Contents

iv. List of Figures

v. List of tables

1 Chapter one: Introduction

1.1. Background

Plants are essential to life on earth. They are foundational to entire ecosystems, support national economies, feed the global community, provide pharmaceutical solutions, promote mental health and wellbeing, and supply a wealth of scientific insight (Hall & Knuth, 2019; Kumar, 2004; De Luca *et al.*, 2012; Lortie *et al.*, 2022; Balderman *et al.*, 2016). Yet studying the impact of plant systems on both the local and global scale is often overlooked in favour of human-focused activities (Grierson *et al.*, 2011). Overcoming this plant-blindness is perhaps the most important aspect of generating a sustainable future for all. Fortunately, as people are increasingly looking for healthy ways to interact with nature, awareness of plants continues to grow (Stagg & Dillon, 2022).

However, an aspect that people do not often consider when thinking about plants is of the pests that affect them. In reality, plant pests pose a significant threat to our environment, agriculture and even the global economy (He *et al.*, 2016). This is ironic as it is through awareness that a willingness to investigate plant disease arises; without investigation, there would not be surveillance. Surveillance is foundational to the development of effective plant pest management strategies (Parnell *et al.*, 2017). The rate at which novel pests are establishing in regions is constantly increasing (Spence *et al.*, 2020). The acquisition of surveillance data must match this pace because timely detection and intervention are crucial to mitigate their impacts on agriculture, biodiversity, and local economies (He *et al.*, 2016). Therefore, the dependency on scientists taking the lead on data collection has never been greater. Co-ordinating scientists, stakeholders and other non-specialists requires accurate assessments of current knowledge on the status of pests. This is only possible through generating meaningful insights based on available data. The role of National Plant Protection

Organisations (NPPOs) has been well designated by the International Plant Protection Convention (IPPC) mandate; it is imperative that NPPOs adhere to and enforce phytosanitary requirements to ensure the safeguarding of global plants and trade (IPPC, 1997). NPPOs require sufficient resources and expertise to facilitate the gathering of evidence of plant disease, and the capacity to act on data to reduce the overall introduction and burden of plant disease (IPPC, 2016). The focus of plant disease specialists within this framework should be on translating data collected in the field into transparent insights for bilateral interactions between stakeholders and NPPOs. By these ends, the development of a repertoire of scientific models that are transparent and transferable is sorely needed. Therefore, it is the role of plant disease specialists to investigate biological phenomena and integrate their findings into the wider economic, social, and ecological framework through the effective use of models. This framework encompasses the economic considerations of the impact of plant disease on agriculture and trade, the social consequences of the loss of crop species according to cultural and trade value and the ecological ramifications for biodiversity and the delicate balance of ecosystems (He *et al.*, 2016; Yang *et al.*, 2006).

Few attempts have been made to synthesise the overall impact of plant pests in the context of the above considerations, though Yang *et al.* proposed a system based on the risk assessment of disease (Yang *et al.*, 2006). The narrow definition of risk assessment is considered as determining the potential epidemiological and economic impact of emerging or new diseases (Yang *et al.*, 2006). Risk is a suitable index for summarising the causal interactions within the plant disease network because it relates the spatio-temporal aspects of disease with global stability (Yang *et al.*, 2006; Hywatt-Twynam *et al.*, 2017). Risk is also predictive and stochastic, and so is a generalisable parameter that is applicable in many plant disease scenarios (Yang *et al.*, 2006; Hywatt-Twynam *et al.*, 2017). Decomposing risk by specific criteria (such as likelihood of disease establishment, environmental suitability and extent to

seasonal dispersal) allows for targeted assessments according to the relevant stake holders affected (Yang *et al.,* 2006). By using biological principles as a foundation for risk assessment, multiple objectives can be addressed. Therefore, it is logical that plant disease specialists have focused on quantifying risk in an epidemiological context. Ultimately, an ability to quantify risk in the context of specific objectives is vital to developing plant disease management strategies that can effectively reduce disease burden (Hywatt-Twynam *et al.,* 2017; Cunniffe *et al.,* 2016).

Examples of risk models include the parameterisation of sanitation methods, disease contact rates, dispersal, virulence, pathogenicity, economic burden, weather conditions etc. (Magarey & Sutton, 2007; Sundström *et al.*, 2014; Newlands *et al.*, 2018). Besides risk, it is also pertinent to consider impact, i.e., the degree to which a disease will reduce crop yield or damage ecosystems. With this consideration integrated into risk models, it has been predicted that the poorest communities globally will be disproportionally affected by the global increase in plant disease (Sundström *et al.*, 2014). Expected global food demand is predicted to increase by 35% to 56% before 2050 because of the global population explosion and developing status of many countries (Van Dijk *et al.*, 2021; Kc *et al.,* 2018). Producing sufficient nutrition for all is not just a question of how to produce food efficiently, how all this food production will be protected from pests must also be considered. Currently, there is too much reliance on a few select crops for sustaining human needs. Whilst it is enticing to use a single commodity crop that can serve the needs of nutrition and support local economies, many plant disease models have highlighted the risk associated with monoculture practises and agricultural intensification (Dun-chun *et al.*, 2021; Uekoetter, 2011). The development of risk models therefore should look to towards a more holistic interdisciplinary approach, whereby the quality of prediction is improved over time by introducing or removing parameters based on sound scientific investigation.

Evidence indicates that the increase in plant pest establishment globally is a consequence of human-mediated activities (Bradley *et al.,* 2010). For example, elevated CO_2 levels have been demonstrated experimentally to increase the proliferation of invasive plant species, especially in arid regions (which constitutes 20% of available land globally) (Bradley *et al.,* 2010; Smith *et al.,* 2000; Nagel *et al.,* 2004). Higher global temperatures have been demonstrated to increase the likelihood of insect pests surviving throughout the year, including overwintering, especially in colder climates (Dukes *et al.,* 2009). In addition, increases in the global movement of material has presented many novel opportunities for plant pests to establish in new climates (Hellmann *et al.,* 2007). Overall, climate change and globalisation are driving the establishment of plant disease through the transportation of material and changing the environment in favour of the plant pests (Pautasso *et al.*, 2012; Evans & Waller, 2009). Examples of human-mediated changes include the expansion of agricultural land use, the loss of native habitats, altering water ecosystems and increases in urbanisation (Zabel *et al.,* 2019; Hanski, 2011; de Barros Ruas, 2022). Modelling these trends is difficult to do mechanistically due to the extent to which stochasticity impacts investigation efforts. However, the localised co-ordination of NPPOs is the best option for managing the impact of these human-mediated activities, and effective plant disease management strategies hinges on transparency of data exchange between governments with unified objectives to managing plant disease (Yeh *et al.,* 2017). A unified effort to managing global biosecurity will require further incentivisation in order to promote sustainable practises (Muhie, 2022). This overall strategy for global bio-security can only be effectively put in place with the use of models that can quantify the impact of NPPO activities. Simple models are now available to present the risk associated with sanitary practises and surveillance protocol, and the adoption of these models can be a priority for NPPOs co-ordinating their activities together moving forward (Magarey & Sutton, 2007; Parnell *et al.,* 2012). Therefore, the focus of this thesis is on the

validation of simple transferable models that can assist NPPO activities across a broad range of epidemics and the implications of epidemiological modelling in the context of surveillance.

1.2. Aims of thesis

The purpose of this thesis is to assess the quality of simple early detection models that have broad application to many plant diseases. Parnell *et al.* (2012) derived a simple early detection model that predicts the prevalence of disease when it is first detected, which from this point forward will be referred to as the rule of thumb (named so for its simplicity). The rule of thumb is founded on the early growth rate of disease and the corresponding rate of surveillance, directly measuring the association between the surveillance rate and the growth rate of an epidemic (Parnell *et al.*, 2012). The output of this model is a quantitative prediction of the total prevalence of disease when first detected. Therefore, validating this model for generalisability could be of benefit for future researchers and stake holders. Previous validation of this model has focused on non-spatial assumptions, i.e. measuring the accuracy of the model without spatial reference to host dynamics (with some exceptions) (Parnell *et al.*, 2015; Parnell *et al.*, 2012). Validation of spatially-explicit epidemiological models is of vital importance because dispersal mechanics are a fundamental aspect of the disease cycle (Wolf & Isard, 2007). By including dispersal mechanics within an epidemiological model, it is possible to build better representations of disease dynamics for further investigation. In this thesis, I will explore many aspects relevant to disease dynamics models with the use of spatial constraints and dispersal mechanics. There is much literature demonstrating how spatial assumptions introduce different epidemiological phenomena within models as compared to non-spatial assumptions (Riley *et al.*, 2015). However, how spatial behaviours affect the accuracy of simple early detection models, and consequently how this impacts the development of pest surveillance management strategies, is largely unexplored territory.

Methods from the field of geostatistics have commonly been used to explore prevalence of disease, based on a population sample protocol (Parnell *et al.*, 2017). However, these methods are not mechanistic, and hence do not capture underlying epidemiological phenomena. Species distribution models present different obstacles, primarily attributing equal probability to infection across landscapes (Parnell *et al.*, 2017). By using a spatially explicit stochastic epidemiological model, realistic representations of a wide range of epidemics can be explored to infer underlying mechanistic principles, which can be applied to allow better early detection surveillance.

Early detection of plant disease is of vital importance. Given the natural exponential growth rate of many polycyclic diseases (Gilbert & Parker, 2023), it is critical that a disease is detected before a disease can establish within a region (Augustin *et al.*, 2012). Early detection depends on the effective co-ordination of NPPOs. Using the rule of thumb can inform NPPOs with meaningful insight into the appropriate allocation of surveillance resources dependent on the disease in question (Parnell *et al.*, 2012; Parnell *et al.*, 2015). This is a major step towards optimising surveillance allocation, so that the focus can shift towards the quality of surveillance as opposed to quantity. There is still, however, outstanding questions that need to be answered with the rule of thumb. For example, how does detection sensitivity affect the accuracy of the rule of thumb? How does landscape heterogeneity affect the accuracy of the rule of thumb? How do epidemiologically relevant parameters such as the dispersal of pathogens affect the accuracy of the rule of thumb? Finally, can the rule of thumb be used with real application to a novel pest? This thesis seeks to address these outstanding questions.

1.3. Structure of Thesis

Chapter two, the literature review, will first explore various aspects of plant disease and how it is linked to National Plant Protection Organisation activities. Chapter three then explores epidemiological parameters and their impact on homogenous landscapes to reduce variability in landscape effects. The objective here is to determine to what extent dispersal and transmission affect the accuracy of the model. Chapter four then explores landscape heterogeneity to explore how the aggregation of hosts in space affects the accuracy of the model. The objective is to introduce incremental levels of aggregation to quantify how this impacts the accuracy of the simple detection model. In chapter five, detection sensitivity is introduced into the picture to determine to what extent detection sensitivity impacts the accuracy of Parnell et al's early detection model, and a derivative developed by Mastin *et al*. (Parnell *et al.*, 2012; Mastin *et al*., 2020). Chapter six then goes on to explore a case study using derived epidemiological parameters on the novel pest *Oak Processionary Moth* (*Thaumetopoea processionea*) to highlight how current surveillance strategy is impacting the prevalence at detection of this pest. The objective is to demonstrate a methodology to deriving epidemiological relevant parameters to use within simple detection models. Chapter seven bring all these findings together to discuss the findings of the thesis, synthesising the findings together and highlighting their significance.

Hopefully the reader will see the broader picture in validation of models and the methodology of investigation deployed here is left as a stepping-stone to the validation of other models of scientific interest. With the increasing prevalence of plant disease globally, the need for generalist expertise across a broad range of epidemics has never been greater. I intend to inform stake holders and scientists with the most thorough analysis to date of the accuracy of this simple epidemiological detection model by introducing various epidemiological factors such as dispersal distance and landscape aggregation, surveillance parameters such as

detection sensitivity and applying the models explored in this thesis to a real case study. The findings from this thesis will hopefully be of benefit to policy makers, agricultural communities and conservationists in optimising early detection and response strategies, thereby safeguarding our ecosystems, food security and economic interests from the escalating threat of plant disease.

2. Chapter two: Literature review

2.1. Introduction

2.1.1. History of plant disease

Historically, there has been much interest in the occurrence of plant disease, with evidence that plant disease has been a constituent part of human life since agriculture developed (Agrios, 2005). Theophrastus, considered as the "father of botany", wrote extensively about plants, and gave descriptions of their diseases in "The nature of Plants" around 300 B.C. He noted that rusts were more common on cereal crops than on legumes, and that disease was more common and severe in lowlands (Stakman, 1957; Agrios, 2005). His contribution to the knowledge of plant disease was very limited however due to beliefs at the time, and plant disease management strategies were restricted to pleasing the Gods (Barnes *et al.*, 2020). The first recognised control strategy for a plant disease was suggested by Albertus Magnus around 1200 A.D (Agrios, 2005). Albertus Magnus recognised mistletoe as a plant pest that reduced the quality of timber by swelling the trunks and causing breakage, and efforts to cure the tree by pruning marked the first documented control strategy (Siraisi, 2001).

The first case of plant disease regulatory legislation arrived in the 1660s, when French farmers petitioned the French Government to remove barberry bushes near wheat crops to reduce the wheat rust agent *Puccinia graminis* f. sp. *tritici* (Stakman, 1957; Agrios, 2005; Barnes *et al.*, 2020). It was still a century before the *Puccinia* genus was first defined, and more than a century after this before the general link between fungi and disease was established (Barnes *et al.*, 2020). In the late 19th to early 20th century, wheat rust epidemics were very common in the US, in many years leading to a near total loss of wheat yield (Peterson, 2013). Outbreaks were common across the US and most of Europe, including the

UK, prompting drastic culling of barberry bushes through government legislation (Barnes *et al.*, 2020).

One of the most notable plant epidemics in history is the Irish potato famine in the 19th century, in which an estimated 1.5 million people died, and nearly the same emigrated from Ireland due to the devasting effects of late potato blight (Agrios, 2005). Late potato blight in Europe was caused by the HERB-1 strain of *Phytophthora infestans*, which was introduced in the 19th century from the Americas (Goss *et al.*, 2014). Anton deBarry is best known for elucidating the life cycle of *P. infestans* in the mid eighteenth century, and the role of *P. infestans* in the development of symptoms associated with potato blight disease (Matta, 2010). His experiments predated the discovery of germ theory by Louis Pasteur, as he demonstrated that spores from the previous season were the causal agents of the next season's epidemic (Matta, 2010; Agrios, 2005).

The debate regarding spontaneous production and the causality of disease continued into the 20th century, though the gradual acceptance of Koch's postulates eventually superseded previous theory (Zadoks, 2001). The 20th century was a pinnacle era in the development of the understanding of the conditions surrounding disease development (Zadoks, 2001). Early in the 20th century, discussions moved to the nature of overwintering pests and pathogens, the transmission of pathogens between sub-species of crops and other plants and, perhaps most important to the general field of epidemiology, the establishment of epidemic growth curves (Zadoks, 2001, Brauer, 2017). Advances in plant pathology can be attributed to advancements in technologies such as microscopy and better understanding of the relationship between pathogen, host, and environment. Progress in the field of microbiology further facilitated advancement in quantitative epidemiology through the work of scientists such as Kermack and McKendrick (Zadoks, 2001; Brauer, 2017). After 1963, plant

epidemiology itself became a field of study, with the insights of scientists such as J. E. Vanderplank firmly advancing ideas such as the latent and infectious periods in plant disease epidemics, and better understanding of the theory underpinning quantitative plant disease epidemiology (Zadoks, 2001; Madden *et al.*, 2017; Vanderplank, 1963) (see section 2.4 for further details).

2.1.2. Biosecurity & significance of infectious plant disease

Today, the significance of infectious plant diseases across the globe is undeniable (IPPC, 2021; Antonelli *et al.*, 2020; Global Plant Health Assessment & Savary, 2023). Globalisation has facilitated massive increases in trade and travel as well as the onset of climate change, and hence the risk of novel plant pests invading national borders is increasing yearly (Defra, 2014; IPPC, 2021). The Food and Agricultural Organisation of the United Nations estimates that between 20 and 40 percent of global crop production is lost to pests annually, with these losses to plant disease costing the global economy around £170 billion (IPPC, 2021). Furthermore, plants directly contribute to quality of life and provide sustainable ecosystems, which can be disrupted by the destructive nature of plant disease (Spence *et al.*, 2020; Turner-Skoff & Cavender, 2019). Infectious plant diseases can be identified by the characteristic symptoms that lead to damage: wilting, spotting (necrosis), rotting, mould, among others (Nazarov *et al.*, 2020). Infectious plant diseases can also be identified by their causal agents: bacteria, fungi, viruses, protozoa, parasitic plants, oomycetes, nematodes and other vectors (Nazarov *et al.*, 2020). Determining the causal agent of a disease is essential to further control measures. Firstly, treatment measures may vary in their effectiveness dependent on the individual pathogen species' biochemistry. This is a major factor in prescribing the most effective mixture to mitigate losses, and other factors including multiple infections by disparate pathogens on the same host must also be considered (Jetiyanon & Kloepper, 2002). Secondly, identifying the causal agent is crucial to isolating the pathogen's

mode of dispersal and reproductive potential. This information is vital for modelling the spread of a pathogen and must be considered for the purposes of prediction and estimation of yield losses within a plant population.

For example, the fungus species *Magnaporthe oryzae* has been identified as the causal agent of rice blast disease, which is the most destructive disease of rice world-wide (Dean *et al.*, 2012). Rice is the primary calorie intake for half the world's population, and it is estimated than rice production will have to increase by 40% by 2030 to meet demand (Khush, 2005). Total crop losses of between 10%-30% annually due to *M. oryzae* are common, with sudden regional epidemics representing the typical scenario (Dean *et al.*, 2012). Cisgenic breeding processes for rice cultivars resistant to rice blast disease have been proposed recently, and modelling the impact of such interventions is of much relevance to National Plant Protection Organisations (NPPOs) (Nalley *et al.*, 2016). Considering the global significance of rice blast disease, amongst other prevalent diseases, optimising the link between modelling and the implementation of disease management strategies is vital if biosecurity measures are to be utilised effectively.

Plant viruses also pose a major threat; for example, the establishment of *tomato spotted wilt virus* (TSWV; carried by its most important vector, *Frankliniella occidentalis*) has caused significant economic losses globally (Scholthof *et al.*, 2011). TSWV is prevalent globally, where it has consistently reduced the yield of several important plant species such as tomato, peanut, pepper, and potato (Culbreath & Srinivasan, 2011; Roggero & Masenga, 2002; Parrella *et al.*, 2003). The symptoms of this disease vary from species to species, and even from host to host, due to factors such as the age of the infected plant, the level of nutrition and environmental conditions (Best, 1968). Firstly, this degree of prevalence indicates that plant pathogens have sufficient capacity to spread across distances that span national borders.

Secondly, the unique presentation of symptoms on hosts due to the presence of TSWV demonstrates that the interactions between host, plant and environment largely determines the presentation of symptoms on plant species, which lead to our causal assumptions of the chances of damage or risk to entire populations (Kaniyassery *et al.*, 2022). Fundamentally, these interactions between host, pathogen and environment underlie even the most basic of epidemiological models. These interactions are ubiquitous across all taxa responsible for the symptomatic presentation of disease.

A final example raised here for the diversity and significance of plant disease is an example from Nematoda. Pine wood nematode, (*Bursaphelenchus xylophilus*), is an example of a parasitic nematode that is devastating conifer trees globally, leading to significant economic losses (Mota *et al.*, 2009). The destruction caused by *B. xylophilus* is comparable to the virtual elimination of the American chestnut by chestnut blight caused *by Cryphonectria parasitica* and elimination of ash trees in Europe by ash dieback caused by *Hymenoscyphus fraxineus* (McMullan *et al.*, 2018). Prioritising disease management on a local scale should be based on a comprehensive understanding of local pathogenic species and ecological interactions from which targeted preventative efforts should be used (based on efficient modelling approaches) and the financial, ecological, and social value produced by a plant species. Perhaps more generally, global scale prioritisation should focus on sorely neglected plant diseases that have historically been overlooked in favour of the more popular human and animal diseases (Flood, 2010).

Control of plant diseases is classified as high priority today, and is managed at a local, regional, and global scale (Stack, 2008). Challenges to plant biosecurity include human population growth, globalisation, climate change, bioterrorism and changing agribusiness trade networks (Stack, 2008). There is a need for standardised yet robust strategies for the

detection and eradication of plant disease backed by scientific knowledge of plant epidemiology with an open dialog between scientists, stakeholders and government to construct effective plant disease management strategies (Spence *et al.*, 2020). Plant disease management strategies rely on anticipating occurrence of disease and attacking vulnerable points in the disease cycle, which can occur before or after infection has been established. Therefore, knowledge of plant epidemiology can be tremendously beneficial to informing plant disease management strategies (Filho *et al.*, 2016).

To address these challenges, there has been increasing emphasis on Pest Risk Assessments (PRAs). PRAs are analytical tools that evaluate the risks associated with the introduction with specific pests in specific areas. They consider the potential for entry, establishment, and spread of pests, and based on these assessments, offer informed decisions about the best course of action (Defra, 2023). Examples of PRAs include the UK Plant Health Risk Register, which has assessed over 1400 pests and pathogens as potential threats to the UK agricultural and horticultural landscape (Defra, 2023).

One of the most effective risk reduction options that emerge from PRAs is the identification and closing of pathways. For example, if a specific pathogen is found to be prevalent in a trading partner's country, then the trade of specific plants may be prohibited or restricted (Defra, 2023). Furthermore, the practise of horizon scanning is being adopted by many countries. Horizon scanning is a continuous, systematic process to identifying and assess future threats and opportunities for pest expansion. This continuous effort to identify potential pest risks pre-border is consistently advancing in scope and capacity with the help of machine learning (Sutherland *et al.*, 2023). The emphasis lies on both pre-border and post-border measures; while pre-border measures include rigorous inspections, certifications and quarantine protocol, post border measures involve surveillance, early detection, and

management of established pests (Defra, 2023). Machine learning techniques could be applied to many areas of plant disease management such as identification, classification, quantification and prediction of plant disease (Yang & Guo, 2017). Briefly, detection, quantification etc. of diseases can be performed by identifying key classifiers such as discoloration or leaf pattern changes caused by the symptomatic presentation of disease. Furthermore, coupled with advanced tools such as near-infrared (NIR) spectroscopy, changes in leaf content can be detected and quantified using machine learning techniques for the identification of disease before symptoms present themselves (Conrad *et al.*, 2020). Advancing the use of machine learning requires using methods that are transferable from one disease scenario to another; for example, machine learning algorithms "trained" on canopy identifiers in a particular region may not be transferable without additional modifications to another region.

To ensure the success of these measures, it's important that there is constant and transparent dialogue between scientists, stakeholders, and government. Only by pooling the collective knowledge of plant pests and resources for pest management can it be possible to construct effective pest management strategies against the myriad challenges within plant biosecurity (Defra, 2023).

2.2. Epidemiology of plant disease

The purpose of plant epidemiology is to address plant disease epidemics as ecological phenomena, and then derive knowledge from which effective plant disease management strategies can be designed for tactical implementation (Filho *et al.*, 2016). With the advancements in machine learning algorithms and computational power it is possible to derive effective modelling solutions with the use of data alone, though this does not provide as much causative inference as with building mechanistic models (Tourinho & Vale, 2023).

A fundamental understanding of the plant disease triangle is essential to modelling the disease dynamics within plant epidemiology (Figure 2.1) (Agrios, 2005).

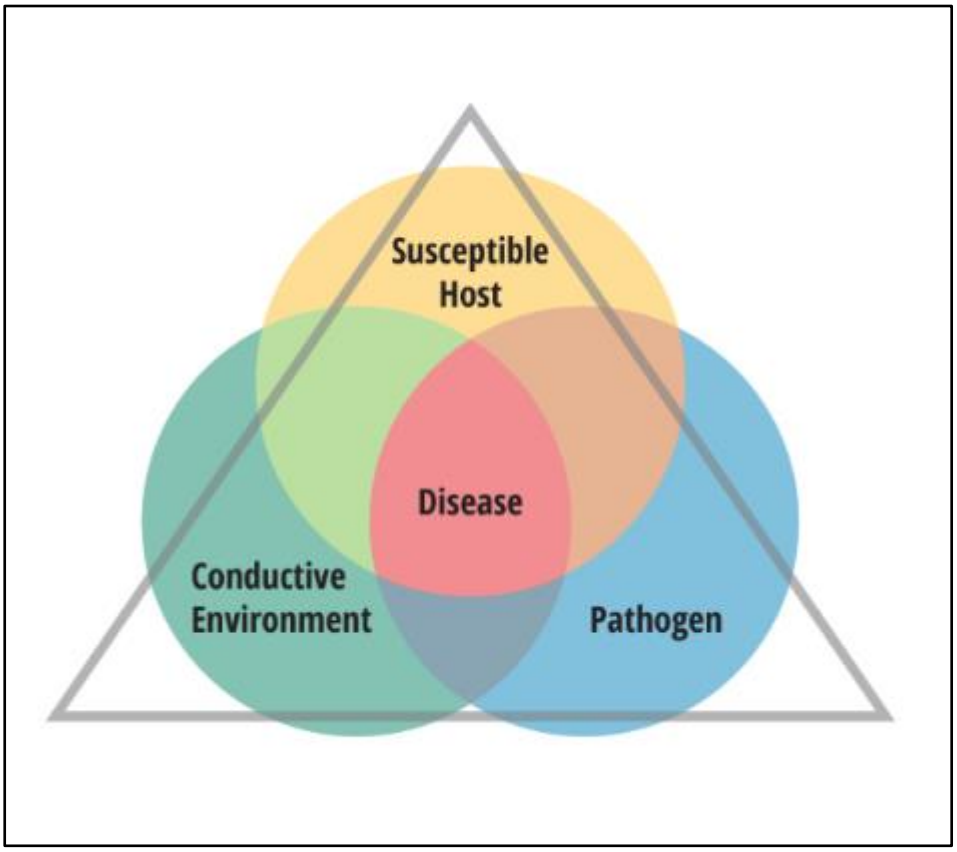


Figure 2.1. The plant disease triangle. All three components must be present for the occurrence of disease. Image taken from *precisionfarmingdealer.com* (1st September 2023).

To make an estimation of how a pathogen will spread we need to address fundamental questions relevant to host susceptibility, pathogen virulence, pathogenicity, and the conducive nature of the environment (Barrett *et al.*, 2009; Sacristán & García-Arenal, 2008; Meentemeyer *et al.*, 2012; Lapin & van den Ackerveken, 2013). Historically, the distinction between pathogenicity and virulence has been obscured in plant disease literature; recently. virulence has been defined as the reduction in host fitness caused by the interaction between host and pathogen. and pathogenicity as the capability of a pathogen to establish itself on or within host material (Barret *et al.*, 2009). Modern development in plant epidemiology does

not consider these factors in isolation. Host susceptibility is affected by a pathogen's evolutionary capabilities, pathogenicity is altered by the landscape distribution of hosts or the variability in climate parameters, and the environmental distribution of plants is changed by the establishment of endemic disease (Turner, 2005; Meentemeyer *et al.*, 2012; Lapin & van den Ackerveken, 2013). Furthermore, the picture becomes increasingly complex as we consider multiple infections on individual hosts, spillover to non-target species, the evolutionary "arms race" between host and pathogen, and the influence of human activity in the context of landscape ecology (Barret *et al.*, 2009, Turner, 2005; Lapin & van den Ackerveken, 2013). Furthermore, with advances in knowledge regarding the introduction of novel plant diseases as a direct consequence of human-mediated climate change and globalisation, plant epidemiology must consider and anticipate factors that remain largely unknown. Therefore, plant epidemiology must rapidly adapt as a field in which prediction is possible across categories of taxa larger than ever before. Plant epidemiology has a strong mathematical and statistical basis on which to build realistic models to be utilised across spatial, temporal, and spatio-temporal populations (Cunniffe *et al.,* 2015); how these models can be translated into meaningful tools for stakeholders to embrace within the schema of plant disease management still remains a fundamental challenge of plant epidemiology in the 21st century (Madden, 2006).

2.2.1. Pests & pathogens

The number of species capable of damaging plants is well recorded. Over 10,000 species of fungi damage plants through the causation of diseases such as the rusts, mildews, and leafspots (Nazarov *et al.*, 2020). In addition, the diversity of oomycetes (as defined by their diploid life cycle), as currently known is vast and continuing to expand (Kamoun *et al.*, 2014). Approximately 4500 plants are considered to live parasitically on other plants (Gogoi *et al.*, 2020). There are nearly 2000 different species of virus currently known to cause

damage in plants (Gaur *et al.*, 2021). Approximately 150 bacterial species are known to cause diseases such as rots, wilts, and blights in plants (Kannan & Bastas, 2015). Parasitic nematodes are also causative agents of plant disease: at least several hundred are known to feed on plants, causing a variety of diseases worldwide (Agrios, 2005).

The proliferation of plant disease is largely dependent on factors integrated within the plant disease triangle; however, here I will focus on the biology of pathogenic species as a means of classification and characterising epidemiological behaviour. Modelling the proliferation of pathogenic species, and to an extent the total virulence within a population, requires detailed inference on the reproductive capacity of the pathogen, the dormancy period in which a pest requires to become virulent, mechanisms of dispersal, and overall pathogenesis (Wolf & Isard, 2007). It logically follows that the duration of time necessary for a complete replication cycle consequently plays a significant role in the growth rates of epidemics (Gergerich & Dolja, 2006; Agrios, 2005). Fungi, bacteria, and viruses typically have short reproduction cycles (replicate rapidly) and produce relatively large quantities of inoculum, thus are likely to proliferate quickly. A short reproduction cycle leads to a polycyclic epidemic whereupon plants that are infected act as secondary sources of infection during an epidemic expansion phase (Sacristán & García-Arenal, 2008). Longer reproduction cycles, such as those for parasitic plants, nematodes, and some species of fungi, result in slower offspring generation (Agrios, 2005). Infected plants often cannot act as secondary sources of inoculum in these cases, and the epidemic is classified as monocyclic. Whether a disease is polycyclic, monocyclic, or some variant of either, should inform stakeholder's disease management decisions. van der Plank initially proposed over sixty years ago that the key to effective disease management for monocyclic diseases lies in targeting the primary source of infection that drives the epidemic (van der Plank, 1963). This approach, whilst a foundational insight to the field of mathematical plant epidemiology, requires a more modern analysis. Recent

advances in modelling and pathogen control techniques have demonstrated that the quantity and relationship between primary and secondary inoculum is context dependent on the particular pathogen strain, the host distribution (through both time and space) and the environment in which the interactions between pathogen and host occurs (González-Domínguez *et al.*, 2020). For example, whilst the polycyclic disease cycle is traditionally defined by the logistic equation (with y_0 representing primary inoculum), study of sexual and asexual stages of oomycetes have led scientists to include additional terms to reflect the variable disease intensity contributions of numerous phases (offset by temporal considerations) (González-Domínguez *et al.*, 2020). This emphasizes the evolving nature of epidemiological strategies and the need for their continual refinement based on the latest scientific insights. The parameter *r* is calculated as the rate in which inoculum will replicate given the current infection status of hosts. Therefore, for disease cases where rapid replication of inoculum is prevalent, focusing on the host-pathogen interactions is prioritised over identifying the initial source of inoculum (Nutter, 2007).

The dispersal mechanisms of plant pathogens are additionally important considerations when characterising diseases within the context of plant epidemiology. In some cases, fungi and parasitic plants produce their inoculum on the surface of the aerial parts of the host (such as spores and seeds), and therefore dispersal can be facilitated over long distances through environmental factors associated with weather conditions (Magarey & Sutton, 2007). In contrast, viruses depend entirely on systematic relationships with vector carriers or other means of co-transportation to be actively transmitted from plant to plant (Gergerich & Dolja, 2006). For example, Tomato spotted wilt virus (TSWV) multiplies in cells of thrip vectors (such as *Frankiniella spp.*); however, dispersal is restricted to adult thrips because larvae and pupae do not feed on plant material (Salvalaggio *et al.*, 2017). Understanding this phenological behaviour is one of many critical factors relevant to mechanistic modelling;

including other components of thrip life cycles and preferred weather conditions for both pathogen and vector are also essential (Salvalaggio *et al.*, 2017; Chappell *et al.*, 2013). Synthesising all of these parameters into meaningful models that can be considered generalisable (i.e., transferable across spatial and temporal boundaries) even for the single TSWV virus-vector relationship presents a challenge to scientists, and consequently difficulties for stakeholders (Magarey & Sutton, 2007). Discussions of these challenging phenomena will be presented later in this section, though it suffices to state a lack of standardisation of parameters, a dependence on modelling specialisation in select crops, and the inherent challenges presented by modelling overall risk in terms of summary variables collectively create barriers to the development of transferable models (Magarey & Sutton, 2007).

Whilst fungal spores will not utilise an active vector seeking host plants, the dispersal (release) mechanisms can be effective with small perturbations in the environment (Mukherjee *et al.*, 2021). The passive removal of spores from their hosts because of climatic forces such as wind force is often sufficient enough for the successful propagation and establishment of diseases (Chaudhary *et al.*, 2022). Hence, the largest proportion of plant diseases that are known are fungal, simply due to the successful mechanisms of reproduction driven by evolutionary forces (Magarey & Sutton, 2007). Field experiments have identified spore movements of several hundred kilometres in what is known as a long-distance dispersal event (Rieux *et al.*, 2014; Prospero *et al.*, 2005). The pattern of spore dispersal is also cyclical, with peaks in spore concentration in the middle of the day (Oneto *et al.*, 2020). Other mechanisms of dispersal have been recorded across the fungal taxa. These mechanisms include responses to the wetness of the host or pathogen, changes in air temperature, changes in humidity, and total exposure to irradiance (both short and long wavelengths) (Chaudhary *et al.*, 2022). Choosing the appropriate model for fungal epidemics depends on the effective

characterising of dispersal with the use of the most appropriate mathematical representation. The mechanistic modelling of take-off, transport and deposition processes depends on the interactions between transmission parameters and dispersal parameters within mathematical frameworks.

Plant diseases caused by bacterial species can be dispersed in a variety of ways, though on short dispersal scales splash dispersal is often a key factor. Here the binding of bacteria to plant cells by mixing with water reduces the likelihood of removal via wind intermittence (Jones & Harrison, 2004). Similarly, plant nematodes are adapted to heavier moisture conditions as they typically move through the film of water that is attached to soil particles, or otherwise thrive in light, sandy soils (Kim *et al.*, 2017). Modelling soil-borne plant nematodes must particularly account for human mediated transportation activities, as soil-born nematodes have very restricted mechanisms of dispersal via natural movement. Unlike other plant pathogens, aquatic plant nematode species have unique dispersal mechanisms constrained by hydrodynamics (Ptatscheck & Traunsprunger, 2020). Plant nematodes typically thrive in warmer climates where longer growing seasons extend feeding periods and increase reproduction rates (Somasekhar & Prasad, 2011). Examples of bacterial and nematode species of significant economic or research importance include *Xanthomonas* spp., *Xylella fastidiosa, Globodera* spp., and *Radopholus similis* (Mansfield *et al.*, 2012; Jones *et al.*, 2013). The diversity of plant pathogens globally demonstrates the repertoire of (largely species dependent) variables and processes that researchers must consider. Hence, researchers (particularly practitioners) within plant pathology are usually specialised in their focus on a select few pathogenic species, leading to difficulties in amalgamating generalised mechanistic data for transferable model synthesis (Magarey & Sutton, 2007).

The vertical transmission of pests (i.e. vegetative transmission) are equally important considerations. Vegetative transmission is a major concern for most field crops, fruit, ornamental trees, and shrubbery where readily accessible plant organs are available for transplantation (He *et al.*, 2019; Cobos *et al.*, 2019; Galperin *et al.*, 2003). Recent research has indicated however that vertical transmission of pathogens acts as a double-edged sword for plant species. Vertical transmission of certain micro-organisms can promote plant health; and incidences where pathogen virulence is decreased and plant resistance is increased through selection processes that promote these micro-organisms have been recorded (Pagán *et al.*, 2014; Truyens *et al.*, 2015). Mechanical transmission of pathogens through sap transfer is less common, though plant viruses such as *potato virus X, tobacco mosaic virus* and *cucumber mosaic virus* can cause significant crop losses through mechanical transportation (Gergerich & Dolja, 2006; Cruz *et al.*, 1998; Murakishi *et al.*, 1971). Mechanical transmission of pathogenic bacteria can also occur through abrasions or wounds on leaves, stems or roots; the pathogen *Xanthomonas campestris* pv. *pruni*, for example, is transmitted from cankers on infected plum trees to nursery trees when pruned under suitable weather conditions (Goodman & Hattingh, 1988). Given the role of human-mediated activities in the increased transmission of pathogens; some scientists have considered the plant disease triangle as a tetrahedron instead (with human-mediated influences such as farming practises as the fourth vertex) (Francl *et al.*, 2001). Alternatively, time is often expressed as the fourth variable to the plant disease relationship instead of human activity because of the temporal considerations over the duration of plant disease epidemics (for example, host senescence).

The transmission of plant disease from the preceding generation of epidemic inoculum, often referred to as overwintering, is also a major concern to biosecurity. Seeds not fully submerged in soil surfaces account for the largest proportion of overwintering events, though volunteer plant species and plant debris also contribute to the indirect transmission of plant

disease over seasons (Jaspers *et al.*, 2015). The presence of indirect overwintering sources is contextual on the geographic landscape in which a plant disease has established. The establishment of disease in an urban region may lead to rapid proliferation in surrounding rural areas, so overwintering management efforts should ideally require consideration of landscape epidemiology, as in the case of *Huanglongbing* and psyllid survival over winter months in Florida (Martini *et al.*, 2020).

2.2.2. Environment & host

Environmental factors significantly contribute to the proliferation of plant disease epidemics and therefore must be considered as integral to the plant-pathogen paradigm (Kalaris *et al.*, 2014). Weather conditions such as wind intermittence, irradiance, humidity, and temperature play a significant role in the proliferation of most plant pathogens, as well as environmental gas concentrations of CO_2 and CH_3 (Kalaris *et al.*, 2014; Chaudhary *et al.*, 2022; Velásquez *et al.*, 2018). Fundamentally, studying the interactions between disease presence and weather conditions is currently the most viable method of developing generalisable models due to the availability of climate data and the relatively simple methods of identifying parameter ranges in which pathogens will proliferate (Magarey & Sutton, 2007). Typically, it is a rule of thumb that conditions suitable for a plant host are suitable for their pathogens.

Leaf wetness for example encourages plant growth but contributes to the overall susceptibility of plants to infection (Velásquez *et al.*, 2018; Kalaris *et al.*, 2014). In contrast we find some diseases, such as the causative agent of rice blast disease (*Magnaporthe oryzae*), are significantly more prevalent during drought conditions (Velásquez *et al.*, 2018). The virulence of many other diseases is known to be affected by precipitation and humidity (Clarkson *et al.*, 2014; Thompson *et al.*, 2013). Effectively developing sufficient systems of modelling these interactions involves a solid understanding of not only host-environment

interactions (i.e. leaf wetness), but also environment-environment interactions (i.e. soil composition-precipitation) under which the host is interacting (Thompson *et al.*, 2013). If a mechanistic model focuses on partial environmental interactions, there is a risk that insufficient complexity has been captured. This risk can lead to confounding factors remaining unknown; thus, a model that captures some behaviour in a particular species may not be transferable to another. If the model parameters are designed for the purposes of prediction, then this can lead to disastrous consequences (Ioannidis *et al.*, 2022).

Temperature is another key environmental driver of disease (Magarey & Sutton, 2007). For every plant pathogen interaction, there is an optimal temperature at which disease develops (Velásquez *et al.*, 2018). For a simplistic comparison of this generalised principle, the potato cyst nematode *Globodera pallida* has an optimal temperature of approximately 15 °C for disease growth, whereas the temperature range of 27 °C to 35 °C is optimal for the virus *Xanthomonas oryzae* (Horino et al, 1982). Variability in optimal temperature is not restricted to the unique classification of plant pest species. Optimal temperatures can significantly vary between species and even sub-species of the same genera (Rossi *et al.*, 2001). Furthermore, the host-pathogen interactions influence the degree to which temperature is a determining causal factor. Optimal temperature growth for a plant pathogen may vary across host species. The temperature influence may itself be influenced by co-infections or exposure to other hazards as in the case of *Iflavirus* whereby multiple pathogen infections alters the minimum requirements for proliferation (Jakubowska *et al.*, 2016; Ponnuvel *et al.*, 2022).

Weather conditions also impact overall vector density and survival rates throughout the year. The survival rate of vectors over winter months is well-explored within the literature due to the relevance it bares on the expansion of disease (Roos *et al.*, 2011; Zeilinger *et al.*, 2017). As the trend in global warming continues to increase positively, insects are most likely to

benefit from increasing temperatures based on favourable changes in vector biology and changes to agricultural practise (Roos *et al.*, 2011). Vector density is also demonstrated to be positively associated with the current trend in changing climate conditions, albeit this is a complex dynamic to model (Zeilinger *et al.*, 2017).

How host plants interact within the plant disease system is another aspect that must be considered when delineating causal relationships for the purposes of disease management. Plants have natural defence systems that can cope with many different pathogens (Heil & Bostock, 2002; Lapin & van den Ackerveken, 2013). However, the interactions between plant host and pathogen are complex and the consequence of the genetic arms race between plant host and pathogen can lead to dire consequences for plants (Lapin & van den Ackerveken, 2013). For example, components of plant surface waxes can promote the germination of fungal spores and flavonoids released by the host plant into the soil can attract soil-borne pathogens (Lapin & van den Ackerveken, 2013). Given the typical genetic uniformity of crop species within a cultivar, the aggregation of genetic predisposition to infection can compound to more severe epidemics (Brown *et al.*, 2004). Conversely, the introduction of genetic resistance is a possible strategy to strengthening crops and reducing the likelihood of pathogen establishment (Lapin & Ban den Ackerveken, 2013, Nutter *et al.*, 2007; Adams *et al.*, 1971). Genetic resistance has been utilised in various crops (such as barley, potatoes and rice) to reduce the impact of their associated fungal pathogens, significantly increasing yield as a result (Parlevliet, 1979; Rimbaud *et al.*, 2021).

The composition of an effective plant resistance system must consider the spatio-temporal scale that includes the genetic deployment strategy (Parlevliet, 1979; Rimbaud *et al.*, 2021). For example, the dispersal of resistant genes may be aggregated within one cultivar, introduced across a time interval through crop rotation or segregated as a mosaic of resistance

across an entire landscape (Parlevliet, 1979; Rimbaud *et al.,* 2021). These considerations are not limited to genetic resistance; the general principles of landscape distribution (i.e. how plants are distributed throughout space) and how this impacts host susceptibility to disease are well studied (Brown & Bolker, 2004; Rottstock *et al.*, 2014; Plantegenest, 2007). Broadening the scale from field to landscape, aggregated clusters of hosts emerge in what are often referred to as "patches" (Malavasi *et al.,* 2016; Plantegenest, 2007; Borer *et al.*, 2016). By using a multiscale approach that incorporates larger landscape behaviours, more accurate models of plant host-pathogen interactions can be derived (Borer *et al.*, 2016). Disease progression is different from pathogen to pathogen, how these interactions between host and pathogen lead to disease and further quantifying the disease progression curve itself is the foundation to developing effective disease strategy. It is therefore vital to develop techniques that can either deduce the degree to which landscape plays a role within this foundation or to inductively explore how landscape shapes disease behaviour through experimentation. Patch effects are well documented in literature (Johnson & Haddad, 2011; Plantegenest, 2007; Mundt *et al.*, 2011). The challenge remains to synthesise models that can account for patch effects in nature and only with a plant epidemiologist or plant disease ecologists perspective is this possible.

Overall, the interactions between host, pathogen and environment are complex, dynamic, and unique. This leads to difficulties in generalising models for application across a broad range of taxa. However, progress in the understanding of the underlying mechanisms of disease establishment and proliferation does mean that there are plenty of models and experimental data to inform concerned scientists and stakeholders when considering effective disease plant disease management strategies (Zaffaroni *et al.,* 2020; Parnell *et al.,* 2015). It is important to consider the spatio-temporal scale of these interactions when understanding the entire system and both the direct and indirect consequences of intervention, from gene to landscape. One

could visualise the plant disease triangle becoming a Sierpinski tetrahedron (Figure 2.2). A fourth element is introduced (time) to become a tetrahedron and the aggregation of hosts as space between systems is removed approaches the Sierpinski tetrahedron. The tetrahedron represents the addition of the fifth variable space. However, given the restrictions on the formation of a Sierpinski tetrahedron, the spatial dimension can only be represented as uniformly distributed elements; the reality of spatial considerations is indeed more complex than this would suggest. Accounting for the spatial components of disease proliferation remains a challenge for 21st century plant disease research.

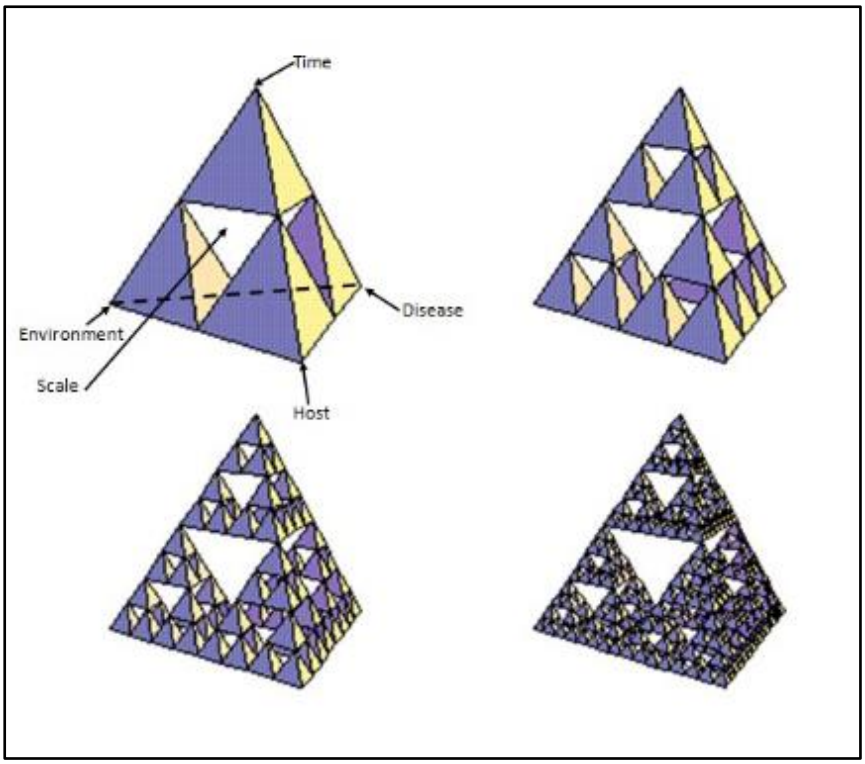


Figure 2.2. Sierpinski tetrahedron to demonstrate the fourth and fifth variables of time and scale respectively within the plant disease system. A Sierpinski tetrahedron is a fractal form of a regular tetrahedron. Inner space is removed iteratively as the dimensions of the shape approaches 2. By summarising the entire disease system within a single host and representing this knowledge as the original tetrahedron, every removal of space within the tetrahedron represents a shifting of scale tending towards the entire landscape. The original system becomes an aggregation of individual sub-units representing a complex mosaic of landscape across time. This representation reflects both an increasing understanding of the interactions involved in plant disease systems and the increasing complexity required to meet the challenges of plant disease management. Image adapted from *www.robertdickau.com/tetrahedron.html* (1st September 2023).

2.3. Plant disease surveillance

2.3.1. Surveillance objectives

Surveillance is a critical component of plant disease management strategies, and has multiple purposes for biosecurity; consequently, it is a mandatory activity of NPPOs (IPPC, 2016). New tools have been developed to address surveillance strategies such as the deployment of molecular tools, geographic information systems and remote sensing (Bronzato Badial *et al.*, 2018; Madden *et al.*, 2017; Zhang *et al.*, 2019). However, designing surveillance strategies sufficient to control plant disease is multi-faceted and complex (Parnell *et al.*, 2017). There are statistical design requirements (i.e. quantifying confidence intervals to produce meaningful insights) that must sufficiently meet the standards of the inspectorate. The minimum level of confidence in surveillance output is often sacrificed to meet financial or logistical constraints; often ideal surveillance objectives are not met due to limited resources or budget allocation (Parnell *et al.*, 2017). Likewise, if a targeted region of surveillance outmatches surveillance resources, then the focus of surveillance efforts often shifts instead to critical pathways of entry (IPPC, 2007). Furthermore, how many surveillance resources to allocate to a particular plant disease shift depending on economic and social variables which may be difficult to quantify (Garret *et al.*, 2011; Cunniffe *et al.*, 2015). Overall, developing surveillance strategies is a challenge even without considerations of the plant disease cycle. Model based approaches have been used to address these outstanding issues in plant disease surveillance (Cunniffe *et al.*, 2015). The inclusion of disease dynamics within surveillance models remains a challenge that sorely requires investigating. Disease dynamics provide insight into how diseases spread over time. Incorporating these dynamics into surveillance models enhances the likelihood that predictions are more aligned with real-world scenarios, thus enhancing the accuracy of forecasts. Furthermore, incorporating disease dynamics into surveillance models ensures that surveillance models account for fluctuations and

stochasticity in natural disease spread. Including disease dynamics also ensures that resources are used efficiently, without missing critical information on disease spread such as high-risk pests or particularly vulnerable regions.

2.3.2. Surveillance operations

Surveillance operations have broad application; typically categorised as early detection, monitoring, delimiting and pest free confirmation (Kalaris *et al.*, 2014). Surveillance protocols, coupled with analysis, are how data are gathered to address the on-going challenges of managing and preventing losses due to plant disease. With the use of effective surveillance, data are available to estimate the impact a disease will have. With the use of analysis, theses data ar translated into meaningful insights to overcome economic, social and health challenges related to plant disease (Kalaris *et al.*, 2014).

The ultimate objective of detection surveys is to discover invasive species before they have become firmly established in a region (Kalaris *et al.*, 2014). However, if an incursion is already established then surveillance protocol is shifted to a monitoring process, where classification of the prevalence of disease per area is determined (James, 1974). If a pathogen has been detected, then it is the role of monitoring surveys to maintain an ongoing assessment of disease (James, 1974). The means by which this is done is relatively straight-forward and common. Firstly, field experiments are conducted to establish the virulence and pathogenicity of disease, secondly this information is used as the foundation for assessment in another population (James, 1974).

Detection surveys are like monitoring surveys in that they can provide the foundation for detection and diagnosis; coupled together detection and monitoring surveys provide a powerful means for risk analysis and quantifying the impact of disease. The purpose of these surveys is to assist by informing plant disease management strategy development through an

effective a warning system to nearby regions in the event of another epidemic and informing stakeholders as to the appropriate level of disease control necessary (Kalaris *et al.*, 2014).

On the other hand, delimiting surveys are used when NPPOs and stakeholders need to determine the extent to which a region is infected by a pest when knowledge of the pest distribution or prevalence is unknown (Kalaris *et al.*, 2014; IPPC, 2016). Delimiting surveys often follow the detection of a novel incursion via a detection survey (Lázaro *et al.*, 2021). Logically, the objective of a surveillance strategy is a result of the known spatio-temporal distribution of a pathogen; Figure 2.3 is a simple analysis of selecting the most appropriate surveillance strategy (Figure 2.3).

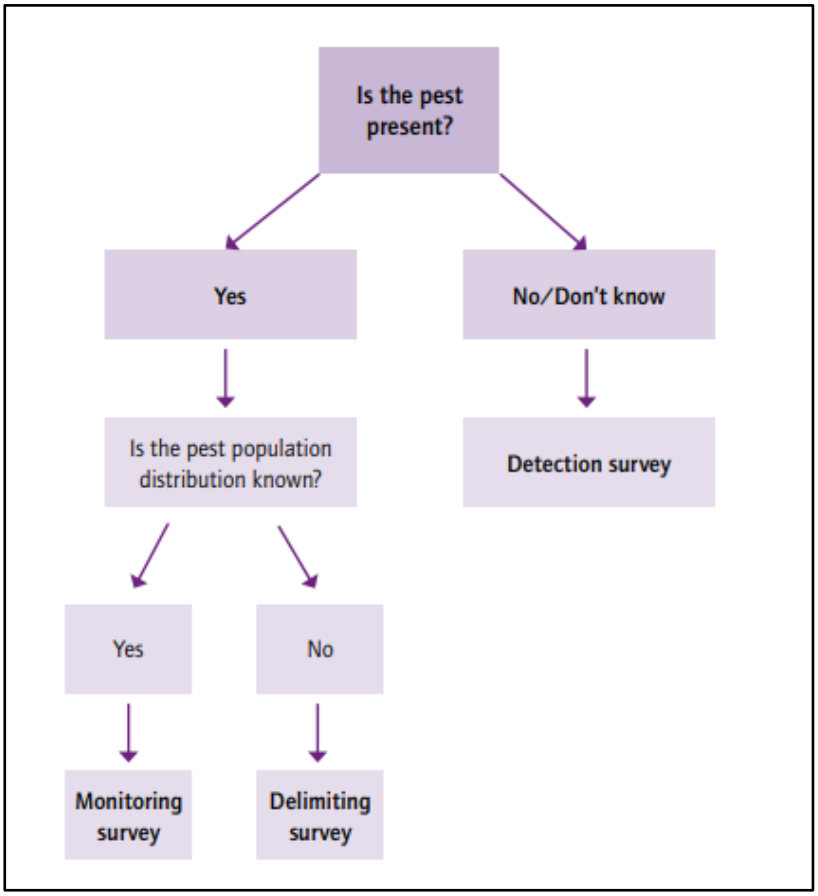


Figure 2.3. The decision support process for selecting the most appropriate survey objective. Taken from the 2016 IPPC surveillance guidance report (IPPC, 2016).

2.3.3. Surveillance methods

The three fundamental methodologies available for surveying plant disease are classified as: sampling surveys, trapping surveys and visual examinations (IPPC, 2016). These three fundamental methodologies vary in their method of collecting data and range in specificity and sensitivity. In sampling surveys host material, pests or soil are collected for identification and analysis (IPPC, 2016; Dean *et al.*, 2005; Lázaro *et al.*, 2021; Le & Vu, 2017). Prior to the first decade of the 21st century, molecular techniques for identification and analysis were expensive and time-consuming to operate. However, thanks to technological advances such as real time PCR and ELISA, the rapid and inexpensive identification of pathogens has become a viable option for surveyors (Defra, 2014; Schaad & Frederick, 2002). In contrast sampling with molecular diagnostic tools, trapping surveys such as those used for vectored diseases are used prior to disease symptom onset (Jackson & Bayliss, 2011). Trapping tools are far easier to operate than molecular diagnostic sampling tools and have the additional benefit of being useful without human-facilitated operation. However, data requires specialised expertise to identify pests and pathogenic spores and therefore will be less accurate overall than molecular tools. Furthermore, as trapping tools are often designed to capture specific pests or pathogenic spores, the transferability of these methods is therefore lacking. Coupling trapping methods with molecular diagnostic methods has demonstrated promising results with real application to managing plant disease (Klosterman *et al.*, 2014).

Early detection of plant disease pivots on the use of appropriate and efficient surveillance measures; utilising trapping tools is useful for vector-borne and fungal diseases in the early phase of epidemic management but these tools have limited application for other pathogens such as viruses and bacteria (Jackson & Bayliss, 2011). For the purposes of biosecurity, confidence in positive results, reliability with repeated use and rapid identification of pest or pathogen are all critical components of any novel survey device (Jackson & Bayliss, 2011).

Visual surveys are a simple form of disease inspection, but results are often subjective and require the onset of symptoms before being applicable (Martinelli *et al.*, 2015). With the aid of technological advancement, visual inspection has become more consistent via the use of visual light photography and digital image analysis (Martinelli *et al.*, 2015). The development of deep-learning software now presents a feasible opportunity to rapidly diagnose plant diseases based on automated image recognition through images captured on smart phones (Mohanty *et al.*, 2016). In addition, recent advances in artificial intelligence (AI) have prompted scientists to investigate the combination of AI and Remote Sensing (RS) technology. This promising avenue of research could be potentially useful in detecting plant disease through the measurements of electromagnetic wave readings; though this has yet to be proven as operationally feasible (Martinelli et al., 2015). However, RS could potentially provide highly accurate visualisation of the spatial distribution of plant disease, and these data could directly feed into a plant disease surveillance strategy (Martinelli *et al.*, 2015).

The structure of the sampling method may also affect the outcome of the survey. For example, sampling may be randomly selecting hosts in a population, systemically selecting hosts according to some ordering scheme, separated according to some "strata" or selection based on clustering (Kalaris *et al.*, 2014). Sample size, confidence in the results and detection threshold are connected; typically, if the sample size is larger, then the confidence in the results below a certain threshold is greater (Kalaris *et al.*, 2014). The sampling scheme used must be specified by the relevant stakeholders and NPPOs, taking into consideration the results required given resources available whilst meeting the obligations of the IPPC (Kalaris *et al.*, 2014; IPPC, 2016). This requires in part a fundamental understanding of the biological impact of the disease in question, which can be estimated using epidemiological modelling. Developments in plant pest modelling have integrated sampling protocol with epidemiological growth to predict the extent of disease prevalence within a population when

first detected (Parnell *et al.*, 2012). Modelling approaches such as this one are very simple to parameterise and thus their predictions have much utility if accurate.

Plant disease "behaviour" is fundamentally unique to each system contained within the plant Sierpinski tetrahedron as described above. Therefore, capturing epidemic "behaviour" within our surveillance approach is critical to developing successful plant disease management strategies. Investigations must be made at each variable within the disease systems and integrated as a whole. This requires collaboration and transparency between stakeholders, government, agencies, and organisations (IPPC, 2016). Adhering to standard practise protocol will facilitate meaningful transference of data between countries. This is vital as it is obvious that plant disease does not recognise borders.

2.4. Epidemiological modelling

2.4.1. Growth curve modelling & parameters

The development of exponential, monomolecular and logistic models has resulted in effective means to describe epidemics within the plant biosecurity global community (van Maanen & Xu, 2003). Historically, polycyclic diseases were described by logistic models whilst monocyclic diseases by monomolecular models; though in more recent times, experimental evidence has demonstrated that these two general classes of epidemic progression curves cannot always be described exclusively by a particular mathematical form (Pfender, 1981). The situation is usually context dependent: the initial amount of inoculum, the life cycle of the host and pest, and environmental conditions may affect the most appropriate mathematical model to use that best reflects disease progression (Pfender, 1981). Other models such as Gompertz can be applied to disease progress curves dependent on field study validation (van Maanen & Xu, 2003). These models are characterised by the initial quantity of inoculum (Y_0), the disease growth rate (r) and, except for unbounded distributions such as

the exponential, the final capacity of disease (Y_{max}) (Nutter *et al.*, 2007). Other parameters that can affect disease proliferation are considered in modified models which include environmental considerations, host susceptibility, biological growth phases; generally if it is within the mechanistic system then it can be parameterised (van Maanen & Xu, 2003). Parameters have been used to represent plant disease epidemic growth since their initial derivation by Vanderplank in later half of the 20th century, this parameterisation has contributed much to the recent development within plant epidemiological modelling (Jeger, 2000). However, the difference between parameterising mechanistic models and parameterising phenomenological models is that phenomenological models must reduce the parameters to an effectively quantifiable space i.e. noise will remain within the system. By doing so, explanatory reasoning is reduced though it is now possible to quantify a complex and possibly stochastic system.

Susceptible-Infected-Recovered (SIR) compartmental models are another set of models that instead of using disease growth data to feed *apriori* into prediction, are parameterised to measure the rate of change in disease state and characterise disease behaviour *aposteriori*. Their form can be either deterministic or stochastic, or some combination of both. The units of the basic SIR model are the total sum of susceptible plants, infected plants and plants that are removed (Jeger, 2000). Further components have been developed as necessary to include latent and cryptic hosts within the modelling framework (Leclerc *et al.*, 2014). These additional units have been included to account for greater complexity in disease dynamics such as when hosts are exposed to disease but not infected or present the disease asymptomatically (van Maanen & Xu, 2003; Cunniffe *et al.*, 2015; Leclerc *et al.*, 2014). The transition from susceptible (S) to infected (I) depends on the number of infected individuals, the underlying population contact structure and the probability of transmission given contact (Keeling & Rohani, 2007). The transition from infected (I) to removed (R) is dependent on

the total time in the infectious class (Keeling & Rohani, 2007). The fundamental principle here is that the entire system is contained within a few parameters that can be predicted with well-designed field experiments. However, generally, these parameters cannot typically be approximated *via* lab experiment and therefore are typically derived from previous epidemic data. Even with accurate parameterisation, without the inclusion of distance as a variable within the compartmental model, it would not be possible to reflect the influence of more realistic contact rates between plants and the distribution of plants as a whole population (Keeling & Rohani, 2007).

The selection of either deterministic or stochastic models is also of critical importance when determining how knowledge of the important growth curve will be used. A deterministic set of equations can be useful to demonstrate principles such as endemic periodicity and equilibrium over time; however, when exploring realistic and often complex landscapes (i.e. with the inclusion of landscape heterogeneity, multiple patho-systems or variable host dynamics) they can fail to provide meaningful insight (Roberts *et al.*, 2015). From this perspective, it is understandable why the literature on stochastic processes to describe plant epidemiological phenomena has rapidly expanded in recent years (Richter-Heitmann *et al.*, 2020; Fabre *et al.*, 2021).

2.4.2. Spatially explicit modelling

Recent advances in technology allow for the spatio-temporal representation of epidemics using compartmental models, facilitating detailed inference about disease dynamics across disparate landscapes (Plantegenest *et al.*, 2007; Jeger, 2000; Bisin & Moro, 2022). Under most circumstances, for directly transmitted diseases, transmission will be a highly localised process and hence the spatial distribution of pathogens is directly related to transmission between infected and susceptibles hosts (Bisin & Moro, 2022; Keeling & Rohani, 2007).

Landscape heterogeneity can therefore affect disease progression in often unanticipated ways, though research demonstrates that landscape heterogeneity typically slows down epidemic growth progression under modelling assumptions (Bisin & Moro, 2022; Keeling & Rohani, 2007).

It is logical that the distances between populations should affect the likelihood of detection and this assumption is described in dispersal kernel constructs (see Figure 2.4) (Brown *et al.*, 2004; Phillips *et al.*, 2008). The dispersal of various diseases has been characterised by dispersal models based on mathematical principles such as the exponential, gaussian or power law assumptions (Keeling & Rohani, 2007). These kernels are selected either by experimentation conducted in the field or by matching the criteria for disease dispersal i.e. a long-distance kernel will require a leptokurtic tail. There are often underlying assumptions to most kernels: likelihood of transmission is reduced with distance from the source, disease is dispersed in all directions, dispersal mechanisms are consistent as they spread across the landscape, inoculum cannot be spontaneous etc. These assumptions are based on experimental evidence, though exceptions to the rules are always points of interest. For example, wind direction may change the dispersal direction of disease, and host resistance mechanisms and interactions with other pathogens change the virulence of pathogens over time and space. Whilst in an ideal world it would be perfect to use a dispersal kernel that allows for perfect prediction and explains the underlying phenomena at play, in reality it is a trade-off between predictability and fitting parameters to match the data. Dispersal kernels grounded in scaling invariance properties, such as the power law kernel, with close approximation to real biological phenomena, have been demonstrated to be better generalised in this regards; allowing scaling to match disease behaviour to provide predictions on larger spatial areas and across longer durations of time (Severns *et al.*, 2019; Farber *et al.*, 2019).

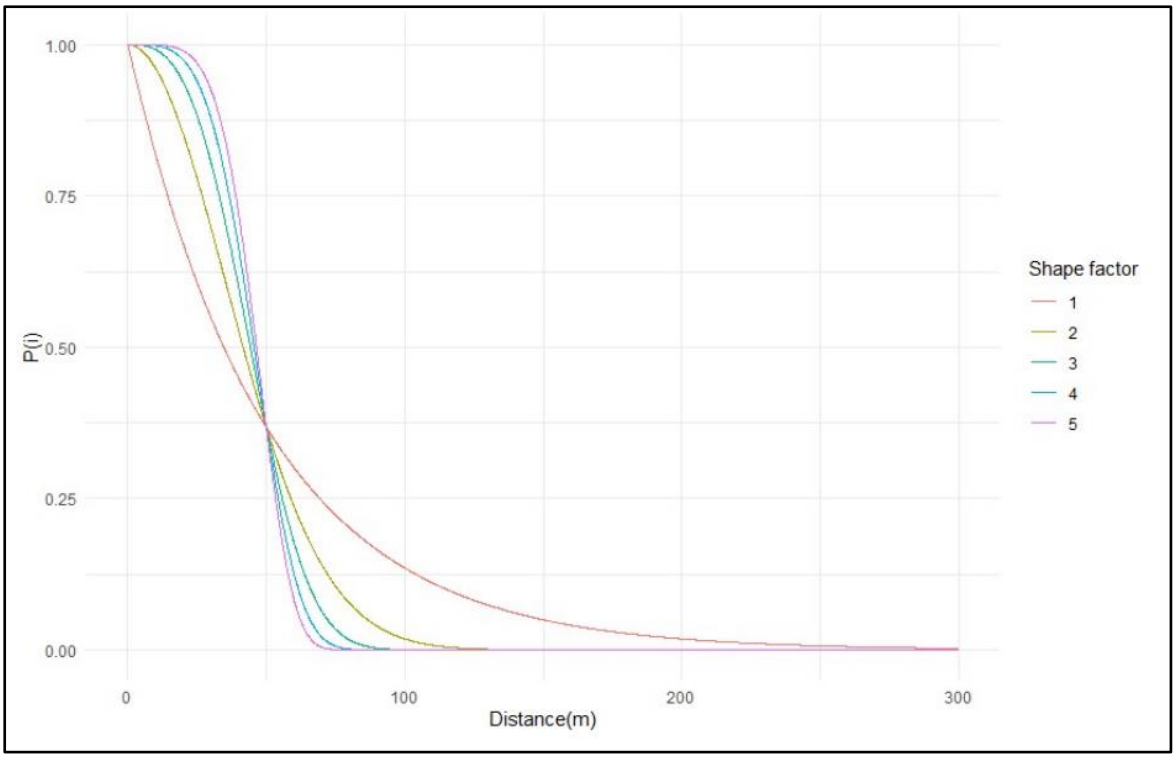


Figure 2.4. five dispersal kernels ranging in kurtic behaviour. Gamma distributed. Kernels become more thin-tailed as the shape parameter increases; the more fat-tailed kernels have a larger probability density at distances beyond the inflection point and vice versa for the higher shape parameter.

Spatial models incorporate dispersal kernels to reflect the natural difference in the probability of disease transmission from a source over distance. For example, localised extinction occurring in subpopulations that are separate to the infection wave can be modelled in stochastic spatially-explicit models (Keeling & Rohani, 2007; Bisin & Moro, 2022). Calculations can be useful in determining R_0 for a subpopulation based on the average prevalence of all populations (Keeling & Rohani, 2007). By calculating R_0, it is possible to infer other useful predictions for disease hence prediction is not constrained by a lack of mathematical inference. If using a stochastic model, there is expected to be much variation in spatially-explicit simulations, and sub-populations of plant hosts contained within strata across the landscape may affect disease dynamics in each simulation depending on the initial

host infection location and the motion of disease based on probability events (Keeling & Rohani, 2007). In effect, landscape heterogeneity adds a layer of complexity that only stochastic spatially-explicit models can capture.

2.4.3. Modelling for surveillance strategy

NPPOs often do not have the necessary resources to conduct surveillance programmes across all vulnerable plant species (Kalaris *et al.*, 2014). Therefore, besides necessary prioritisation and decision-making protocol, surveillance programmes require predictive models to determine spatial and temporal targeting for survey efforts (Kalaris *et al.*, 2014; IPPC, 2016). It is imperative that disease is detected at the earliest possible interval to implement an effective plant disease management strategies (Parnell *et al.*, 2017). In plant health, statistical modelling methods have historically drawn on binomial distribution sampling theory to inform surveillance strategy (Parnell *et al.*, 2017). As stated previously, understanding how the disease cycle alters predictions is critical to developing efficient and effective surveillance strategy.

Parnell *et al.* (2017) demonstrated epidemiological variables could be utilised to inform surveillance via a model that predicted the prevalence of a disease given a given surveillance protocol (Parnell *et al.*, 2017). Likewise, Potts *et al.* (2013) developed a model parametrised for citrus canker that could determine an effective "adaptive radius" in which to search for infected individuals (Potts *et al.*, 2013). Though the "adaptive radius" strategy is not 100% efficient and requires a total exploration of susceptible nodes (Potts *et al.*, 2013). This reflects a limitation of models; it is currently not feasible to model disease dynamics with 100% accuracy, the best that can be done is stating confidence limits around a prediction (Keeling & Rohani, 2007). Application of models to informing surveillance strategy has also been based on ecological principles. Russell *et al.* developed a rapid eradication assessment model

to determine the minimal sufficient monitoring effort required to confirm eradication of invasive species (Russell *et al.*, 2017). It is a general rule of thumb that the earlier detection occurs, the more successful plant disease management strategies will be, therefore models with predictive application to facilitate early detection are preferable to untargeted surveillance approaches (Russell *et al.*, 2017).

Overall, how surveillance is co-ordinated and how surveillance informs NPPOs and other agencies hinges on the development of models that can synthesise relevant epidemiological parameters with meaningful output based on objectives. Hence surveillance models are typically specific to requirements, though effective if utilised properly. Generalising and quantifying the objective of surveillance surveys across NPPOs is critical to increasing the usefulness of these surveillance models. By shifting the focus to the simplicity of objective there will be less urgency in developing a repertoire of models. Therefore, standardisation of simplistic models with thorough investigation into the accuracy of predictive output is the approach that I recommend.

2.4.4. Transferability of models

Many NPPOs are in sore need of models that are robust, transparent and fit for purpose (Defra, 2014). The use of generalisable models in anticipatory prediction are increasingly being used to support policy and decision making with regards to biosecurity (Yates *et al.*, 2018). However, an outstanding challenge of transferable models is the parameterisation of pathogen behaviour which has been repeatedly demonstrated to not be uniform (Yates *et al.*, 2018; Nutter *et al.*, 2007; Magarey & Sutton, 2007). Therefore, parameterisation that reduces these ecological and systematic complexities is a necessity. Furthermore, were an epidemic to be repeatable within a region, it may behave differently according to the nature of complex systems. Therefore, a heavier reliance on stochastic modelling would appear to be appropriate

in moving forwards to more generalisable models. Transferable models will benefit from accounting for sampling bias of individuals. Surveillance data are sometimes pooled from different areas or biased in selection based on spatial and temporal targeting (Kalaris *et al.*, 2014). Therefore, models that can account for these dynamics are more transferable. In addition, models that can account for heterogeneity whilst remaining parsimonious will have greater transferability. When considering the spatial and temporal aspects of a transferable model, it may be useful to define a range from which a model is no longer pragmatic to use. This will help NPPOs define their objectives with a more realistic outlook. Overall, a transferable model will benefit from higher data quality and quantity (Yates *et al.*, 2018). Hence the use of modern surveillance technology such as remote sensing provides an opportunity to fill these criteria. Models that are built with a better appreciation of underlying epidemiological phenomena will also increase the chances in success (Yates *et al.*, 2018). Furthermore, training plant pathologists to understand the entire patho-system globally will increase the likelihood of improving the transferability of modelling approaches (Magarey & Sutton, 2007). This is being observed in reality as the budding field of pathogen ecology becomes more prominent within the 21st century.

In conclusion, the biosecurity challenges of the modern era are based on a lack of knowledge such as data on pest prevalence in nearby regions and an inability to currently synthesise the body of available knowledge into meaningful translatable insight. Traditional scientific methods are addressing the first aspect of this problem, though an interdisciplinary approach and an application of simplicity should be the priority now as we address the transferability of models within plant epidemiology. This thesis seeks to address a broad range of epidemiological values within a set range of parameters relevant to epidemics to synthesis together the applicability of the rule of thumb. It is intended that the utility of the rule of thumb is well defined and explored thoroughly. Thus, exploring epidemiological parameters,

landscape heterogeneity, detection sensitivity, surveillance strategies and applying this research to the real-world case scenario of a current pest in the UK. As the underlying phenomena are better understood from the context of a systems approach, and more people become aware and subsequently interested in plant disease, progress will inevitably be made to improve the lives of millions of people globally who are affected by the proliferation of plant disease.

3. Chapter three: Simple models for complicated epidemics - exploring the use of epidemiologically relevant parameters in parsimonious models to inform early detection surveillance

3.1. Chapter Summary

Plant Epidemiology is the study of disease transmission from plant to plant on a population scale. This discipline is often concerned with the consequences of emerging plant pest threats and developing models to manage invasive species that have not previously established within the focus region. In the UK, the estimation of the cost of invasive species is £1.7 billion per year. To manage these emerging plant diseases, biosecurity measures against plant pathogens employ a combination of surveillance and predictive modelling. How effective predictive models are at informing surveillance strategy across a range of spatially stochastic epidemiological conditions has yet to be extensively explored. By using a spatially-explicit stochastic epidemiological simulation model combined with a spatially-stochastic surveillance model, I explore a fixed surveillance strategy with a range of epidemiological parameters to evaluate the accuracy of one such early detection model, labelled the rule of thumb. I show that decreasing the dispersal distance parameter θ (metres) and increasing transmission coefficient β (rate of infection) leads to a reduction in the accuracy of the rule of thumb. Previously it was thought that increasing the growth rate parameter had a diminishing effect on the accuracy of the rule of thumb. However, my results indicated that the effect is largely complex and dependent on both the dispersal distance and the transmission coefficient parameters. Results suggest the rule of thumb is more applicable to emerging plant diseases that are dispersed across longer distances i.e. >140 meters. My results also demonstrate that the rule of thumb has variable accuracy under a range of realistic epidemic scenarios. These results should be used to inform plant health managers as to the effectiveness of using this specific model in real world scenarios.

3.2. Introduction

Plant diseases are a major problem globally, both economically and in terms of food security (Dun-chun *et al.*, 2021; Burdon *et al.*, 2020). The Food and Agricultural Organisation (FAO) of the United Nations estimates that between 20 and 40 percent of global crop production is lost to pests annually, and that these losses cost the global economy approximately $220 billion annually (FAO, 2021). This estimation of crop loss has been consistent for at least a decade, though these impact costs reflect a complex series of interacting networks within a fabric of ecological, social, and economic constructs (Savary *et al*., 2012; Dun-chun *et al.*, 2021). The introduction of exotic plant diseases to new regions contributes significantly to this crisis, and this trend is increasing globally over time (Spence, 2020). Many nations are investing heavily in efforts to minimise spread and contain these emerging plant diseases (Burdon *et al.*, 2020; Savary *et al.*, 2012). This increase in the threat of emerging plant diseases is largely due to the globalisation of trade, the movement of people, and the unprecedented volume of traded material leading to the proliferation of the causal agents (often referred to as plant pests) (Spence, 2020). Furthermore, with the unpredictability of climate change due to global warming, uncertainty in the risk of plant pest establishment and the associated potential costs are increasing significantly (Coakley *et al.*, 1999). The major concern is that plant pests remain a persistent threat to global security despite continual action to mitigate the impact of the responsible plant pests (Mansfield *et al.*, 2012; Scholthof *et al.*, 2011; Dean *et al.*, 2012; Jones *et al.*, 2013).

How well nations and their National Plant Protection Organisations (NPPOs) can respond to emerging plant diseases is contingent on effective surveillance of plants and control for early detection (Epanchin-Niell *et al.*, 2014; Parnell *et al.*, 2015; Mastin *et al.*, 2020). Tools to facilitate early detection of emerging plant pests are sorely needed to inform optimal strategies for surveillance and control (Gottwald *et al.*, 2002; Gottwald, 2010; Rizzo *et al.*,

2005; Legg *et al.*, 2011; EPPO, 2021). A fundamental challenge to surveillance for emerging plant pests is allocating the appropriate resources to detect and monitor pests, and with specific timing appropriate to the emergence of disease symptoms (Epanchin-Niell *et al.*, 2014). Typically, either the amount of surveillance is too little, leading to inaccurate reporting of disease prevalence and missed opportunities to detect epidemics early, or too much, leading to costly and unnecessary allocation of limited resources (Epanchin-Niell *et al.*, 2014; Ristaino *et al.*, 2021; Anderson *et al.*, 2004). For example, despite the well-known status of ash dieback across Europe, a lack of surveillance meant it was already widespread in the UK when first detected in 2012 (Parnell *et al.*, 2015; Woodward & Boa, 2013). Dieback of ash trees (*Fraxinus excelsior*) is now expected to cost £15 billion in the UK in trade losses, and subsequent damage to vital ecosystems (Hill *et al.*, 2019).

Recently, efforts have been made to use epidemiological modelling to develop optimal strategies to overcome current challenges to surveillance for emerging plant diseases (Mastin *et al.*, 2020; Parnell *et al.*, 2014; Cunniffe *et al.*, 2015; Hauser & McCarthy, 2009). However, many of these recent modelling approaches focused on "where" to allocate surveillance resources, and not "how much" surveillance resources should be allocated to effectively prevent emerging plant disease from establishing in a new region (Mastin *et al.*, 2020; Parnell *et al.*, 2014; Hauser & McCarthy, 2009). Previous efforts to determine how much surveillance is to be allocated to an emerging plant disease have focused on endemic status diseases, drawing on binomial theory to estimate whether a disease is beyond a prevalence threshold for effective management or eradication (Parnell *et al.*, 2017; Madden & Hughes, 1995). However, these statistical approaches are limited in that they only consider a one-off round of sampling, and do not incorporate any available epidemiological information on the rate or pattern of disease spread. Furthermore, other attempts to model the appropriate allocation of surveillance resources based on cost or high priority epidemiologically relevant

parameters have resulted in models that are difficult to understand for the typical stakeholder; whose expertise is required and will be based on their fundamental understanding of the relevant model (Hauser & McCarthy, 2009).

How much surveillance should be allocated before a disease was to be detected was a question that was addressed by Parnell *et al.* (2012). The authors developed a generalised 'rule of thumb' that determined the prevalence of an emerging plant disease when first detected based on the surveillance allocation (Parnell *et al.*, 2012). This rule of thumb utilised the exponential growth rate of specific emerging plant disease epidemics (Methods section 3.3.4.). Utilising the growth rate of epidemics is progression towards making epidemiologically informed decisions of surveillance allocation (Parnell *et al.*, 2015). However, the approach makes several simplifying assumptions, including ignoring spatial effects of disease spread caused by dispersal limitations and heterogeneity in host population distribution. The rule of thumb thus raised important issues around model transferability, namely how far simple epidemic models can be used to make useful predictions to inform surveillance programmes in practice across different epidemic scenarios. Given the non-spatial nature of the rule of thumb, transferring this model to more realistic scenarios has important implications for applicability in real world scenarios. For example, whilst the simple model was demonstrated as useful in predicting Citrus Canker and *Phytophthora ramorum* prevalence at detection, how the model performs generally across a range of different dispersal and transmission coefficient interactions has yet to be explored in further detail (Parnell *et al.*, 2015; Mastin *et al.*, 2020).

This chapter explores different configurations of plant disease transmission coefficients and plant disease dispersal distances within a spatially explicit stochastic simulation model to determine how the rule of thumb translated into different realistic scenarios. To explore these

scenarios, a comparison of both a surveillance model and a spatially-explicit stochastic epidemic simulation model with the rule of thumb was made. In this chapter the intention is to inform researchers and stakeholders how generalisable this early detection model is when used in combination with other epidemiologically relevant models, which are necessary to inform scientists of the complex behaviours inherent in epidemiological predictions.

The primary research question that was addressed in this chapter therefore was:

- How do epidemiologically relevant parameters such as dispersal distance and transmission coefficient impact the accuracy of early detection models?

3.3. Methods

3.3.1. Epidemiological model

The model tracked simulated host plants in a two compartmental model (Keeling & Rohani, 2007): susceptible (*S*) hosts were uninfected and infected (*I*) hosts were infectious. If host *i* was susceptible at time *t*, then the rate of infection of *i* was given by:

$$\varphi_i(t) = \beta \sum_j K(d_{ij}; \theta)$$

(Eqn 3.1)

$\phi_i(t)$ was the rate of infection at time *t*, β was the transmission coefficient that defined the rate for any given distance within the kernel, and the summation ran over all infectious hosts *j* at time *t*. The dispersal $K(d_{ij}; \theta)$ set the rate of disease transmission between a pair of hosts separated by distance d_{ij}, and was parameterised by the dispersal scale parameter θ. Therefore, β is a rate and θ has the units of distance. To allow robustness in dispersal mechanics if needed based on the dispersal kernel tail, I included the power exponential function dispersal kernel explored by previous authors (Bourhis *et al.*, 2019; Rieux *et al.*, 2014). I fixed the scale parameter θ as the thin-tailed exponential kernel in this study, and then normalised so that the probability of infection translated to a total probability of 1 for the entire kernel at any given time:

$$\theta_{norm} = \frac{1}{2\pi\theta^2 \Gamma(2)}$$

(Eqn 3.2)

In order to reduce the time taken to run epidemic simulations, I used the tau-leap Gillespie algorithm (Gillespie, 2001). This algorithm was defined as:

$$\delta I_T \approx \sum_{T-\tau} Poisson(S_t \varphi_i(t) \tau)$$

(Eqn 3.3)

Where τ was the time leap between the initial state and updated state of the simulation process, $(T - t)$.

Table 3.1. List of the epidemiological parameter values used within the epidemiological model. β values were selected as a range between 50-150 and then distributed across seven even intervals to match θ for the purposes of analysis (see results).

Transmission Coefficient (β) (rate of infection)	Dispersal Distance (θ) (metres)
50	10
66.7	20
83.3	30
100	40
116.7	50
133.3	60
150	70

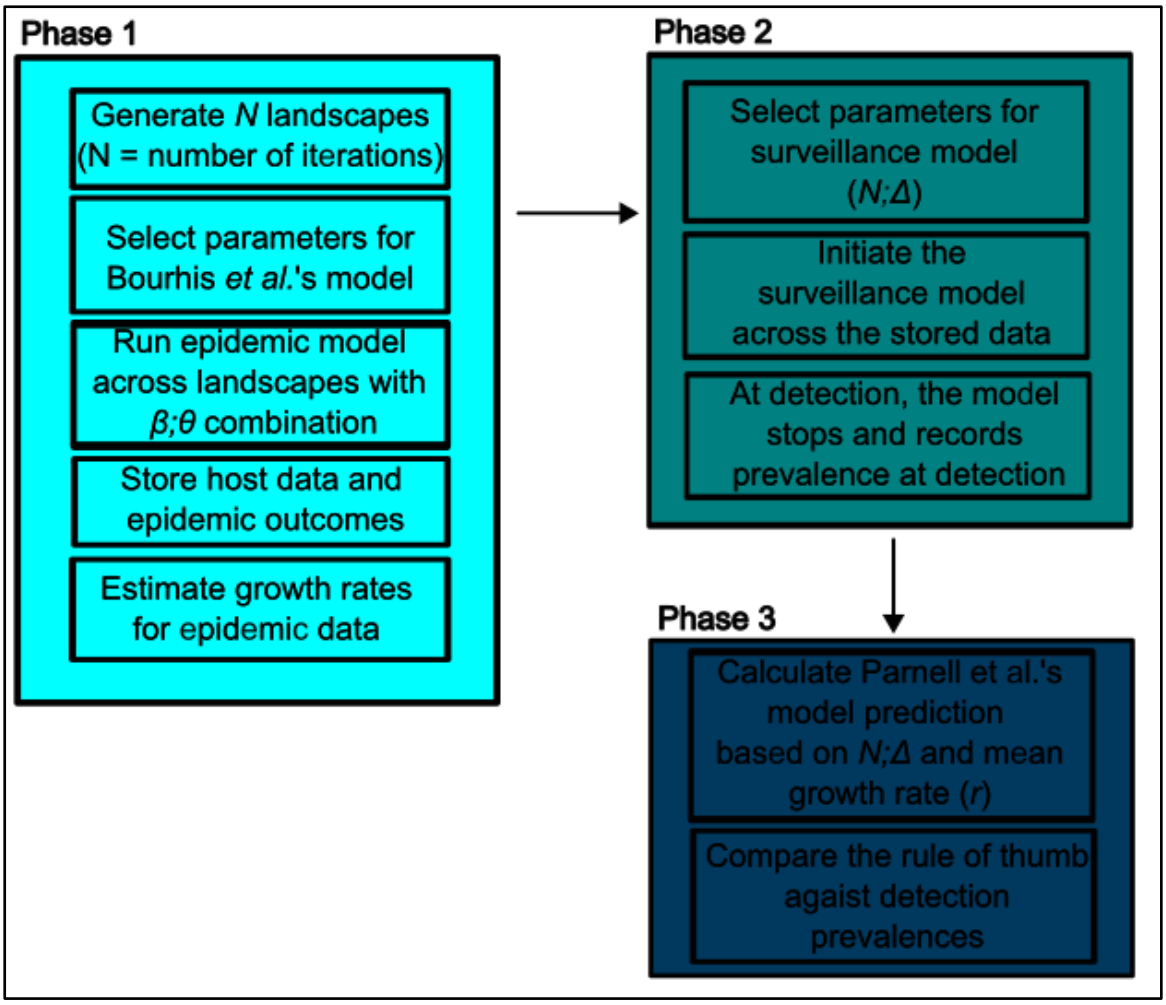


Figure 3.1. Schematic of the experimental procedure to calculate the accuracy of the rule of thumb in chapter 3.

The mean dispersal distance for the kernel is given by the equation stated in Bourhis *et al.* (2019):

$$\phi = \frac{\theta\Gamma(\frac{3}{b})}{\Gamma(\frac{2}{b})}$$

(Eqn 3.4)

Where θ was the dispersal parameter, b was the shape parameter, Γ was the gamma function, and φ was the mean dispersal distance. The shape parameter b is set to 1 (which is defined for

the power exponential kernel), therefore the mean dispersal distance is reduced to $\phi = 2\theta$; the mean dispersal distance of inoculum is twice the dispersal parameter, θ.

3.3.2. Landscape generation and epidemic simulation

The landscape was modelled as a 1km^2 area containing 900 hosts, which was considered a reasonable density for a variety of commercial crops (i.e. citrus or pine species) (Moreira *et al.,* 2019; Gottwald *et al.*, 2002; Kullman, 2006). These landscapes were generated in the R environment using the spatstat package (Baddeley & Turner, 2005). I initiated epidemics with 1 random host infection at time $t = 0$, and the initial host was random for each epidemic run within the simulation, which continued until all hosts were infected. I selected the range of dispersal distances (θ*;* metres) to reflect a mean dispersal of between 20m and 140m and a transmission coefficient value (β; rate of infection) of between 50 and 150 to reflect realistic epidemic behaviour (Alonso-Chavez *et al.,* 2016). These values produced epidemics that reached total prevalence approximately between 1-10 years, which reflected typical epidemic growth curves (see Figure 3.2). The range of epidemiological parameters in this chapter produced growth rates of between 0.011 to 0.022. 2000 landscapes were generated using a random distribution algorithm, with 900 hosts allocated across a 1 km^2 area (Alonso-Chavez *et al.,* 2016).

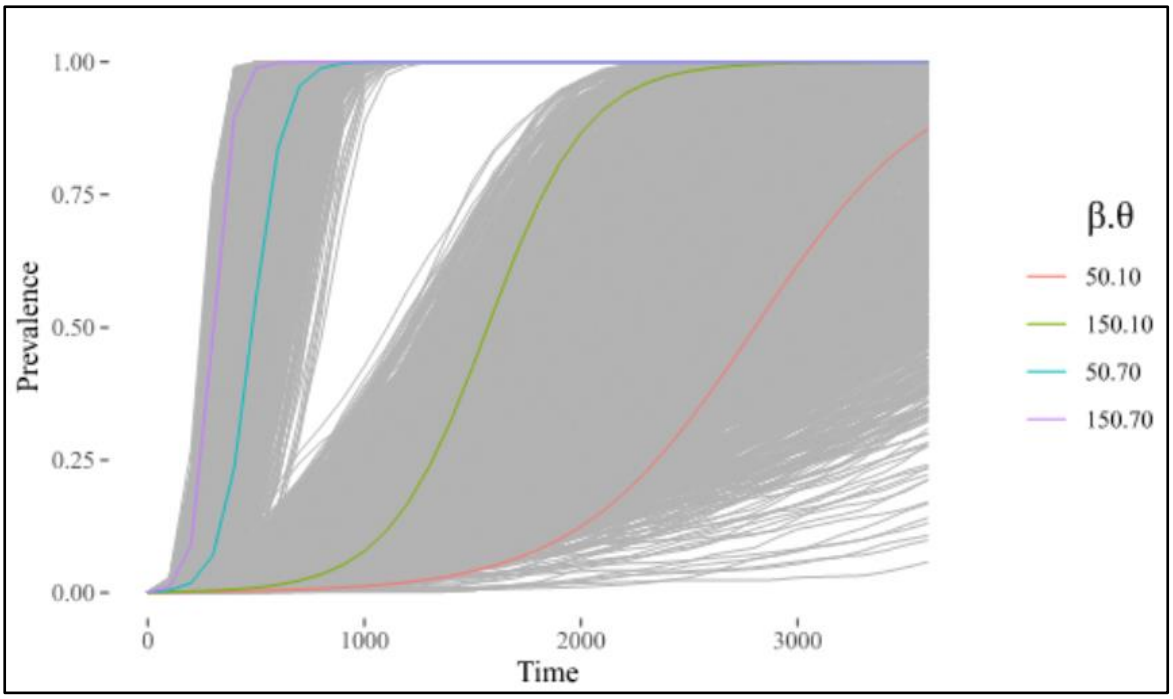


Figure 3.2. Growth curves for the upper and lower bounds of the epidemiological parameter set in this study. β (50;150) and θ (10;70) are the transmission co-efficient and dispersal distance parameters respectively.

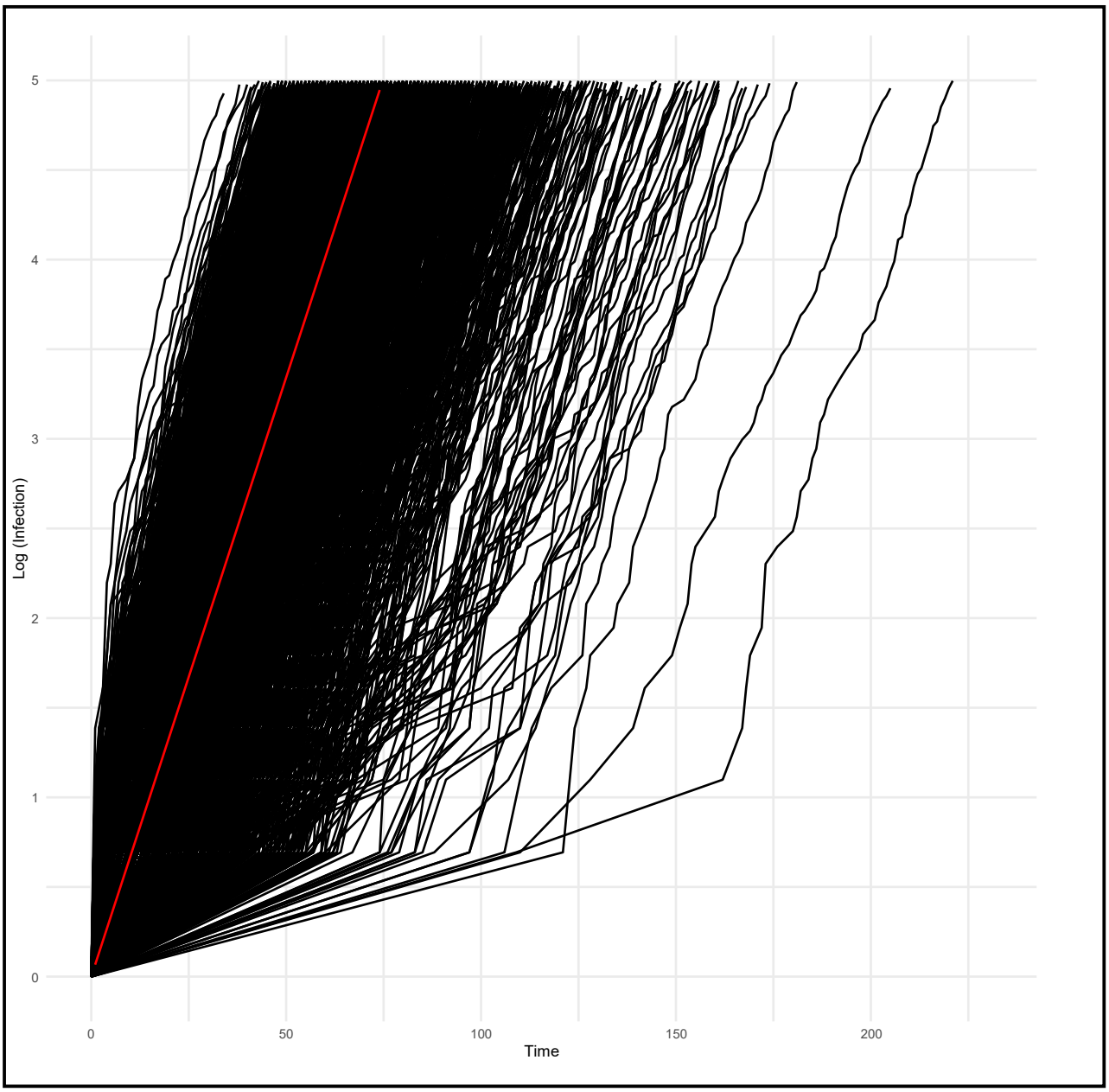


Figure 3.3. Linearised growth curve for an example simulation set across the time period studied for the extrapolation of the epidemiological growth parameter (r). The red line is the plotted exponential growth model with respect to time (t) and (r).

3.3.3. Surveillance model

I developed a spatially stochastic surveillance model by randomly sampling hosts at set time intervals using the stored epidemic data generated previously to estimate the mean growth rate of epidemics. I stopped the surveillance model at the point at which the epidemic was

detected for the first time, and total prevalence in the simulation run was recorded. I then calculated the average prevalence for the complete simulation set. The initial sample size per sampling round was 30 hosts (N), with the interval between surveillance rounds (Δ) set to 30 time units. The compatibility of the surveillance model with the epidemiological model ensures that time units remain coherent, i.e. the time units were set as days and required no conversion between models.

3.3.4. The rule of thumb model

The Parnell *et al.* (2012) rule of thumb is defined as:

$$q^* = \frac{r\Delta}{N}$$

(Eqn 3.5)

Where q^* was the theoretical detection prevalence which was determined by the relationship between the growth rate of the epidemic r and the surveillance protocol, Δ was the interval between surveillance rounds, and N was the sample size taken for inspection each surveillance round. The derivation of this equation was discovered as follows: by the assumption that if an epidemic grows exponentially, then the probability of detecting the disease at any given time was a simple interaction between rate of surveillance and the exponential growth of the epidemic (Parnell *et al.*, 2012). By solving this equation via integration, Parnell *et al.* (2012) arrived at the total probability of detection at the point of detection. Using a random variable transformation of the initial time with an exponential growth assumption, the prevalence at the time of detection was translated from this probability using the Jacobian of the transformation. A useful approximation then followed from this. A full derivation of the approach can be found in Parnell et al (2012; 2015). Thereby, the above rule of thumb stated that given the epidemic growth parameter (r), the

surveillance frequency (Δ) and the sample size (N), then the prevalence of disease at detection (q^*) was estimated as an approximation of the true prevalence of disease at detection. Key assumptions from the rule of thumb are exponential increase of the pathogen population and continuous monitoring of the host population with time.

For every unique epidemiological parameter combination of θ (metres) and β (rate of infection), I performed 2000 simulation runs, providing a unique set of a growth rate values (r). I estimated r by using a linear transformation on the initial growth phase of the epidemic, corresponding to the first quartile of transition events and measuring the mean gradient of the curve. The average of these 2000 r values was then used as input within the rule of thumb (Eqn 4). For N and Δ in the above equation, these values were taken from the surveillance model as defined above.

3.4. Results

Results demonstrated a distinct pattern in the relationship between the epidemiological parameters θ (dispersal distance; metres), β (rate of infection) and the accuracy of the rule of thumb prediction of prevalence at first detection: this is shown by the relative differences between the rule of thumb prediction of prevalence at first detection and the prevalence at detection within the simulation model (Figure 3.3c). As epidemiological parameters θ and β increased, the prevalence at detection within the simulation model increased (Figure 3.3b). This was reflected in the predictive model, which also increased in output value when approximating prevalence at detection within the simulation model (Figure 3.3a). The rule of thumb predictive model consistently underestimated the prevalence at detection within the simulation model, although this effect decreased as dispersal θ increased, as shown in Figure

3.3a & b (Figure 3.3a & b). The pattern for the difference between the rule of thumb predictive model and the prevalence at detection within the simulation model when β was changed demonstrated convergence, where variability in β decreased as θ increased.

My results demonstrated several important findings regarding the impact of epidemiological parameters on the accuracy of the Parnell *et al.*, (2012) method (Figure 3.4). As the epidemiological parameters increased in magnitude, the absolute magnitude in the prevalence at detection within the simulation model increased (Figure 3.3b). The rule of thumb prediction was also consistent in increasing approximation in parallel with increasing epidemiological parameters (Figure 3.3a) However, the rule of thumb prediction generally under-reported the actual prevalence at detection within the simulation model. This result was consistent across the epidemiological parameter space: these underestimations were observed in the relative differences between the predictions and prevalence at detection within the simulation model (Figure 3.3c).

Another key finding was that, as dispersal θ decreased, the relative difference between the rule of thumb prediction and the prevalence at detection within the simulation model increased. For example, the mean relative difference between the rule of thumb approximation and the prevalence at detection within the simulation model was 0.183 for the maximum value of θ (θ = 70). This was contrasted against the mean relative difference of 0.620 for the minimum value of θ (θ = 10).

Another important result was that changing β had a converging effect on the relative difference; at lower values of θ, the increase in β led to larger variance in the relative difference for any given θ value, but as θ moved through the mid-range there was less variability when altering β value (Figure 3.4a & b). Generally, for the range of β values, the mean relative difference across the θ values was approximately 0.28 (Figure 3.4b). The

impact of changing β was not as great as changing θ in the relative difference between the rule of thumb prediction and the prevalence at detection within the simulation model (Figure 3.4a & b).

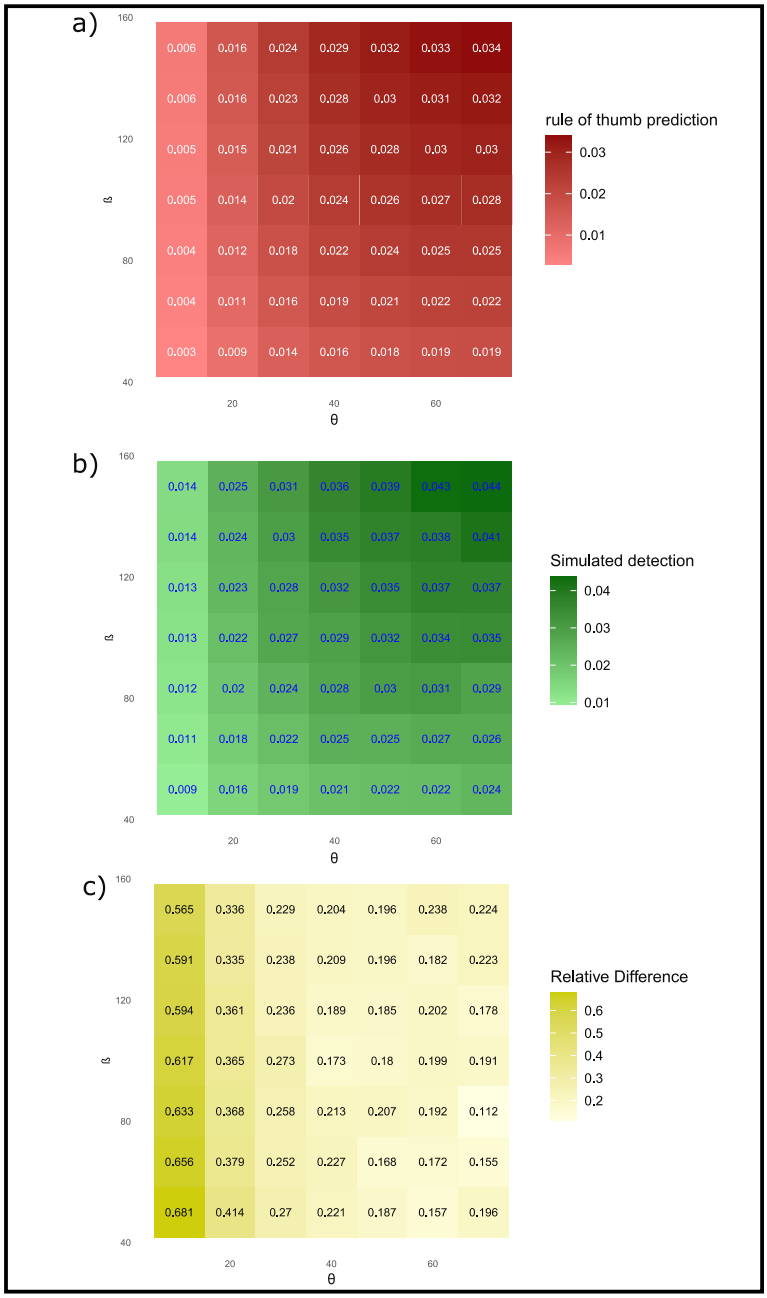


Figure 3.3. Heatmaps demonstrating the output of the rule of thumb, prevalence at detection within the simulation model, and the relative difference between the rule of thumb and prevalence at detection. a) The rule of thumb predictions; b) the prevalence at detection within the simulation model; c) the relative comparison between the rule of thumb and simulated detection prevalences. Surveillance was fixed at 30 N and 30 Δ. Parameters investigated were β and θ with values ranging from 10 to 70 for both. 2000 simulations were used per parameter combination.

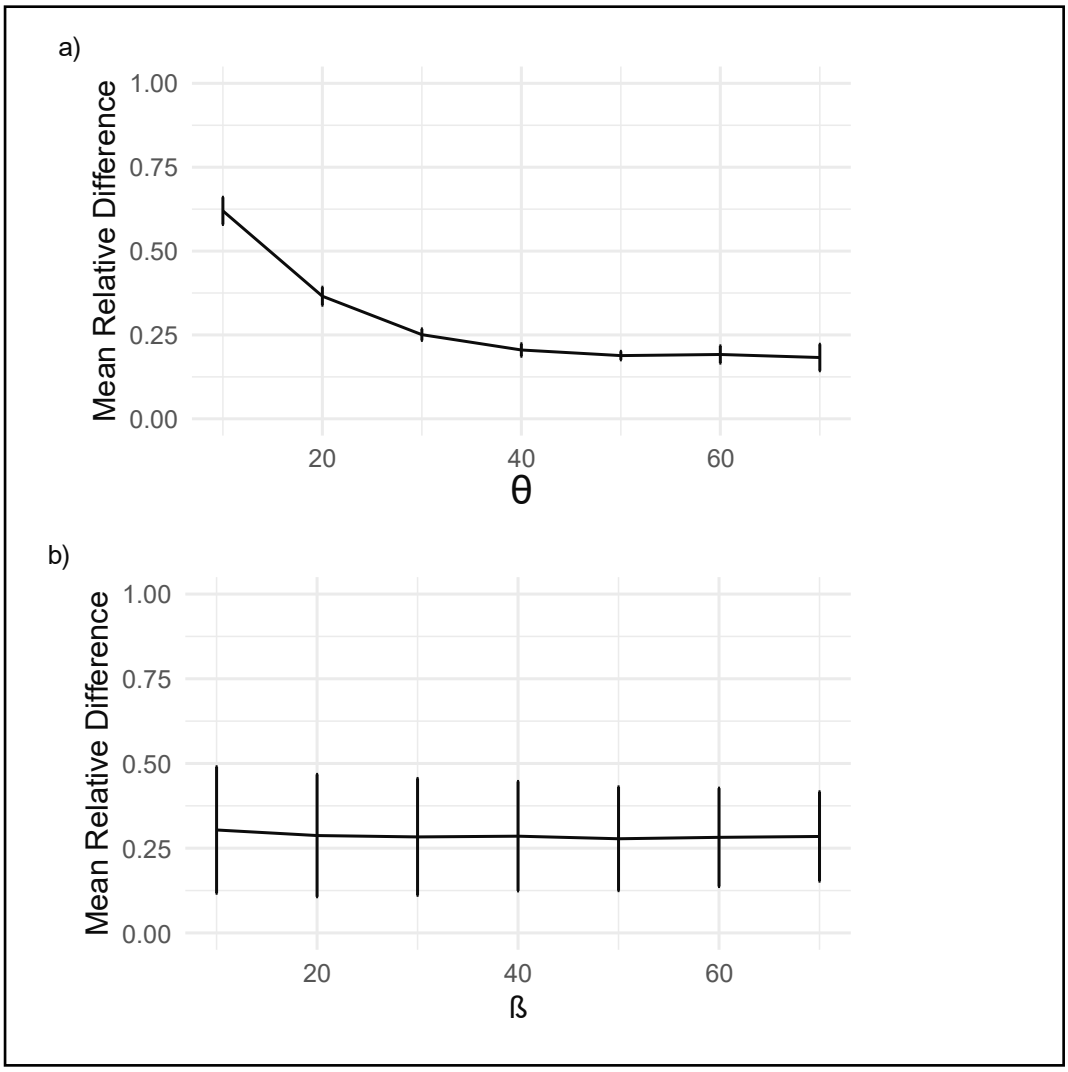


Figure 3.4. Mean relative differences between the prevalence at detection within the simulation model and the rule of thumb approximation for a range of parameter values. Each point reflects the mean relative difference for all values of a fixed parameter, error bars indicating the range of values of the other parameter. a) For each θ value, the β values are reported as the standard deviation with their mean as the point; b) for each β value, the θ values are reported as the standard deviation with their mean as the point. Parameters are β and θ with values ranging from 50 to 150 and 10 to 70 respectively. Surveillance was fixed at 30 N and 30 Δ. 2000 simulations per parameter combination.

3.5. Discussion

Both the transmission coefficient (β; rate of infection) and the dispersal distance (θ; metres) epidemiological parameters affect the accuracy of the rule of thumb when explored within spatially-explicit stochastic models. These results are an extension of previous studies that have explored the rule of thumb in a non-spatial and spatial context, with both simulated and real data (Parnell *et al.*, 2012; 2015; Mastin *et al.*, 2020). The rule of thumb is a promising prediction model, given the simplicity of the model and the applicability of the model to epidemics, as has been demonstrated previously (Parnell *et al.*, 2015, Mastin *et al.*, 2020). This chapter emphasises that simplifying epidemics using the characteristic epidemiological growth rate variable (r) can lead to varying levels of accuracy in early detection predictions. Figure 3.3. demonstrates the linearised fit of the the exponential function to the epidemic data. Given that the exponential fit was applied to the logistic epidemiological data, consideration of a potentially poor fit is merited (Figure 3.3.). The rule of thumb model uses the growth parameter which is applicable to an exponential model and logistic model, though estimates to derive the early phase fit will be different. Judging by the linearised data in Figure 3.3., the stochastic form of the simulation model does somewhat mitigate the risk of misfitting. This chapter has demonstrated that epidemiological parameters such as the transmission coefficient (β) and the dispersal distance (θ) which are directly linked to the estimation of r exhibit their own behaviours both independently and dependently of one another. For the purposes of this chapter, surveillance was fixed at a constant rate for all epidemiological parameter combinations (Δ=30, N=30) to clarify the effects of epidemiological parameters on the accuracy of simple epidemiologically relevant predictive models such as the rule of thumb.

The results showed that starting with a low value of θ in the epidemic simulation model led to the rule of thumb prediction having the lowest relative accuracy. These small dispersal

distances correspond to some real-world disease agents, such as explosive fungal spores of *Pyricularia oryzae* (rice blast disease) or the soil-borne parasite *Rhizoctonia solani* (Meredith, 1973; Gregory, 1968). Previous research has indicated that short distance dispersal of inoculum results in lower final prevalence of disease (Mundt & Sackett, 2012). This result was supported in this chapter, where lower θ values resulted in lower prevalence at detection. While the rule of thumb prediction reflected the reduction in disease prevalence with smaller values of θ, the difference between the predictions and detection prevalences remained relatively consistent across parameter space. This led to the larger relative differences observed between prediction and simulation.

There could be several reasons for this observed consistency. The rule of thumb assumes equal probability of detection across the entire landscape. However, in my simulation model there was a period of time during the early phase of epidemic in which inoculum was not fully distributed across the landscape. Whilst the disease was spreading through the landscape during this period, the sampling strategy may have missed early detection in areas where the disease had not yet spread. The simulation model was also designed so that initial surveillance did not begin at the time infection was introduced (t_0). This would lead to further disease progression not accounted for within the rule of thumb prediction. The consistent underprediction by the rule of thumb raises questions about the assumptions underlying early detection models, and the impact of epidemiological parameters on disease spread and surveillance. By incorporating additional factors into the simulation models and surveillance strategies, such as more realistic assumptions about disease progression related to environmental conditions and pathogenic behaviours, the complexity of disease spread is better described and the accuracy of early detection methods is potentially improved (De wolf & Isard, 2007).

By changing β with respect to θ, there was a converging effect on the accuracy of the rule of thumb prediction. While smaller values of θ had a larger range of impact when altering β, the mean relative difference across the parameter range was approximately 0.28. This indicated several findings. Firstly, that dispersal distance had a larger impact on the accuracy of the rule of thumb prediction. The larger impact of θ on the accuracy of the rule of thumb prediction suggests that epidemiological detection models are influenced more by the spatial distribution of hosts or spatial constraints of the landscape than by the transmission coefficient of disease. This finding supported previous research that indicated that the spatial distribution of host populations and spatial constraints of the landscape can have a significant impact on disease spread, and consequently an impact on the accuracy of early detection models (Holdenrieder *et al.*, 2004; Plantegenest *et al.*, 2007; Parnell *et al.*, 2015; Mastin *et al.*, 2020; Bourhis *et al.*, 2019).

The second finding this result indicated was that, as θ increased, the effects of the spatial distribution or constraints of the landscape became less influential, leading to a flattening of the impact of the transmission coefficient. This flattening was a result of the infection wave front becoming less pronounced, with a larger degree of homogeny resulting from more mixing of infected and uninfected hosts. When the inoculum reached further distances in each time step, this facilitated greater mixing between uninfected and infected hosts, leading to a more uniform disease spread across the landscape. This effect, coupled with the normalisation constant leading to a thinner spread of inoculum across the landscape, resulted in reductions in the influence of β, and hence less variability with increased θ. This result has several important implications. Regarding long-distance dispersal, the disease dynamics changed significantly as the infection wave front became less distinct and the spread of disease became more uniform: diseases proliferated faster, which led to higher prevalence

values. Fortunately, this simple detection model demonstrated higher accuracies for such scenarios, which led to the assumption that this model is suited for such events.

Secondly, this finding also indicated that with long distance dispersal of inoculum, the loci of disease hotspots became more difficult to determine. This could potentially disrupt the implementation of targeted control methods. Thus, with long-distance dispersal of inoculum, alternative management strategies may become more important in management of disease, such as aerial spraying or enhancing host resistance, as opposed to construction of barriers (Filipe *et al.*, 2012; Norelli *et al.*, 2012).

Thirdly, the simulation model in the paper focused on an exponential decay kernel, which has been described previously as one of the most useful dispersal kernels in modelling disease progression behaviour (Bullock *et al.*, 2017). However, in specific long-distance dispersal scenarios, a power law kernel may be more appropriate, and how the rule of thumb prediction aligns with simulations of this nature remains to be explored (Severns *et al.*, 2019). This consideration highlights the need for additional exploration of the impact of other dispersal kernels, such as the power law kernel mentioned, on epidemiologically relevant parameters such as the transmission coefficient and dispersal distance of inoculum. By doing so, researchers would have a deeper understanding of how these kernels impacted the accuracy of the rule of thumb model.

The Parnell *et al.* (2015) study featured growth rates ranging between 0.014 and 0.002, while more recent research focusing on *Phythophthora ramorum* exponential growth rates ranging from 0.001 to 0.005 (Mastin *et al.*, 2020). In this chapter, the range of epidemiological parameters produced growth rates of between 0.011 to 0.022. Growth rates of above 0.02 have been estimated for particularly infectious diseases such as *Xylella fastidiosa* (White *et al.*, 2020). As *Xylella fastidiosa* poses an emerging plant disease threat to the British Isles,

my findings have highlighted the importance of accurately determining the dispersal distance of such infectious diseases (Mabbett, 2018). A previous study of the main vector of *X. fastidiosa* in Apulia in Italy (*Philaenus spumarius*) applied a mean dispersal distance of 100 meters between olive trees to explore the epidemiology of *X. fastidiosa* (White *et al.*, 2017). If *X. fastidiosa* establishes within the British Isles and native *P. spumarius* become infected, then the results of this chapter indicate that the rule of thumb prediction will have a promising chance of being applicable ($\theta = 50$) with an accuracy of approximately 80%. This is also with the assumption that the surveillance protocol of 30 samples every 30 days is used. If the rule of thumb is to be applied for the early detection of such plant disease threats, then incorporating realistic assumptions about the spatial dynamics of disease dispersal and realistic surveillance allocation should be a top priority for disease management and surveillance strategies.

The results of this chapter have shown that the relative accuracy of the rule of thumb was more effective for long-distance dispersal epidemic events, and that there was always some degree of inaccuracy even when spatial constraints and dynamics were not interfering with the predictive accuracy. The results of this chapter also indicated that the rule of thumb will consistently under-estimate the true prevalence of disease due to spatial and temporal factors. These findings directly related to the research question "Do epidemiologically relevant parameters such as dispersal distance and transmission coefficient impact the accuracy of the rule of thumb?" by illustrating the varying degrees of accuracy of the rule of thumb depending on dispersal distance and transmission coefficient. The results contribute valuable insight into the performance of the rule of thumb and hopefully highlight its applicability and limitations in different disease scenarios for future researchers and stakeholders alike. In summary, the findings of this chapter have indicated that the rule of thumb is more applicable for longer dispersed pests and diseases, and for shortly dispersed pests it is wise to consider

how virulent a pest may be; if pest virulence is high, then there is an increased likelihood that the rule of thumb model is less likely to provide accurate estimates of true prevalence at detection.

4. Chapter four: Simple models for complicated epidemics - how introducing spatial dynamics affect parsimonious plant pest models that inform early detection surveillance

4.1. Summary

Early detection models are important for effective disease management and control, but advancing models from non-spatial assumptions to spatially-explicit models remains a challenge. This chapter explored the effects of changing host patterns, landscape scales, and epidemiologically relevant parameters on the accuracy of Parnell *et al.*'s (2012) rule of thumb within a spatially-explicit stochastic epidemiological model. Surveillance simulations were performed over epidemiological simulation data, with the frequency of surveillance Δ and the sample size for each surveillance round N parameterised. These findings indicated that the dispersal distance parameter (θ*;* metres) was the most influential factor in determining the accuracy of the model, with a larger θ resulting in higher accuracy. Furthermore, the impact of landscape clustering on the accuracy of the model was context-dependent, but generally the model was more accurate on homogeneous landscapes than on clustered landscapes, especially in the context of short dispersal θ magnitudes. The low accuracy in estimation of prevalence at detection for diseases that have limited dispersal may lead stakeholders to underestimate the true spread of disease, leading to inappropriate resource allocation to manage the outbreak. This chapter also highlighted the need for further research to refine understanding of dispersal parameters and their complex interactions with other landscape variables. Overall, this chapter contributes to the growing body of research on the role of landscape variables in shaping disease dynamics and control efforts.

4.2. Introduction

Trees and other plant species provide major ecological and economic benefits to many species and microorganisms globally, including humans (Turner-Skoff & Cavender, 2019). Why it is significant to effectively manage plant diseases has been highlighted in Chapter 3. The successful control of large-scale epidemics relies not only on the potential effectiveness of disease management strategies but also on our actions, priorities, resource allocation and the influence of political factors. These overall disease management strategies could be categorised as disease management strategies deployed before the epidemic expansion (for example pre-border prediction and prevention), disease management strategies during the early phase of the epidemic expansion (for example post-border prediction and early detection) and post-hoc disease management strategies, during the later phase of epidemic (for example adaption and analysis, including cost analysis) (Parnell et al., 2015; Boyd et al., 2013; Cunniffe et al., 2016; Ristaino et al., 2021).

Epidemiological modelling is increasingly used to inform disease management strategies (Defra, 2023). Surveillance effort during the early phase of emerging pest epidemics is often not sufficient due to a general lack of understanding of appropriate allocation to meet management objectives of containment or eradication (Parnell et al., 2015). Early phase epidemic management is often required to contain or eradicate disease; this can difficult to implement due to limited information available about an emerging disease, the rapid spread of disease due to lack of detection and diagnosis or the mobilisation and allocation of resources during the early phase of disease establishment (Severns et al., 2019; Parnell et al., 2015). The development of early detection diagnostic tools for many emerging plant diseases have prompted epidemiological modellers to match the technology with models that can facilitate early detection and consequently better management of emerging pest epidemics (Peng et al., 2022; Liu & Wang, 2020). The development of more robust models capable of

directing the activities of disease management efforts during the early phase of epidemics are thus required.

As stated previously, the rule of thumb for predicting the prevalence of disease when it is first detected is a novel tool that plant health managers can use to direct surveillance allocation (Parnell et al., 2012; Parnell et al., 2015; Mastin et al., 2020). Surveillance models have been developed to direct where surveillance should be allocated during the early phase of an epidemic previously (Mastin et al., 2020). Matching spatially targeted allocation of surveillance resources with the appropriate quantity of surveillance will better guide plant health managers when dealing with an emerging plant pest threat. Furthermore, knowing when to allocate the surveillance resources is essential to maximising management efforts (Parnell et al., 2015; Mastin et al., 2020).

The focus of this chapter was to explore the impact of changing landscape dynamics in epidemic simulation scenarios on the accuracy of the rule of thumb. Combining a spatially stochastic epidemic simulation model, a spatially stochastic surveillance simulation model and the rule of thumb allows exploration of the application of early detection models. In particular, landscape heterogeneity is known to significantly influence the dynamics of an epidemic however how landscape heterogeneity influences early-stage epidemic spread and thus our ability to estimate resources required for early detection is largely dependent on available data (Plantegenest et al., 2007; Meentemeyer et al., 2008; Václavík et al., 2010). With advances in tools such as spatial statistics and geographic information systems (GIS) it is possible to quantify the impact of landscape dynamics on disease dynamics (Plantegenest et al., 2007; Meentemeyer et al., 2008; Václavík, 2010). Many recent studies have attempted to understand how spatial dynamics will fit into our current understanding of predictive modelling and overall model transferability (Condeso & Meentemeyer, 2007; Allouche et al.,

2008; Meentemeyer et al., 2008a; Václavík, 2010). For example, by using invasive species distribution models (iSDMs), it is possible to assign "weight" to variables that may be influencing the distribution of invasive species (Meentemeyer et al., 2008a; Václavík, 2010). These models utilise machine learning techniques such as maximum entropy (MAXENT), and are flexible with the input of parameters, allowing different "weights" dependent on a particular invasive species. These models account specifically for the distribution of plants, including symptomatic and asymptomatic hosts to model both the risk of establishment and the possible distribution of disease. Condeso & Meentemeyer used an iSDM to assess the impact of landscape heterogeneity on the proliferation of Sudden Oak Death (caused by the pathogen *Phytophthora ramorum*) (Condeso & Meentemeyer, 2007). By predicting environmental suitability along the Western Coast of the United States, targeted surveillance of disease severity was delivered. The results of this study indicated that trees surrounded by large contiguous forests were most severely affected (Condeso & Meentemeyer, 2007). This is an excellent example of how predictive modelling approaches can be utilised to serve the interests of stake holders in disease management whilst simultaneously delivering novel information to inform future studies and disease management efforts.

How early detection models that incorporate dispersal patterns and host aggregation will perform has yet to be studied. That is, to what extent these factors drive early-stage epidemic and our ability to predict the performance of early detection surveillance strategies is poorly understood. This chapter explores the concepts of landscape heterogeneity with regards to the reliability and uncertainty of predictive modelling (Parnell et al., 2012; Parnell et al., 2015; Newlands, 2018). Given that the rule of thumb used in this chapter was developed with non-spatial assumptions, how well it transfers across a range of possible landscapes is a critical question that is sought to be answered within this chapter. This information is intended to equip plant health managers with a more flexible and robust tool for managing an emerging

or ongoing epidemic, analysing the rule of thumb utility under diverse epidemic conditions where landscape heterogeneity and the distance scale of landscape are factors.

Therefore, the research questions of this chapter are:

- How does changing the host pattern impact the ability to predict detection prevalence with simple epidemic models?
- How does changing the dispersal distance parameter within different landscapes, hence reflecting a range of epidemics, impact the ability to predict detection prevalence with simple epidemic models?

4.3. Methods

The methodology of this section covers the mechanics of the modelling process used in this chapter to explore and evaluate the accuracy of the simple rule of thumb under changing landscape conditions. Initially the models are lain out, including the structure of the dispersal model, landscape generation and surveillance processes. The rule of thumb equation is then described in detail, it's derivation and what it represents are described. Finally, the indices used to quantify the impact of landscape heterogeneity and homogeneity are explained.

4.3.1. Epidemiological model

My model tracked simulated host plants in a two compartmental model (Keeling & Rohani, 2007): susceptible (*S*) hosts were uninfected and infected (*I*) hosts were infectious. If host *i* was susceptible at time *t*, then the probability of infection of *i* was given by:

$$\varphi_i(t) = \beta \sum_j K(d_{ij}; \theta)$$

(Eqn 4.1)

$\phi_i(t)$ was the rate of infection at time t, β was the transmission coefficient that defined the rate for any given distance within the kernel, and the summation ran over all infectious hosts j at time t. The dispersal $K(d_{ij}; \theta)$ set the rate of disease transmission between a pair of hosts separated by distance d_{ij}, and was parameterised by the dispersal scale parameter θ. Therefore, β is a rate and θ has the units of distance. To allow robustness in dispersal mechanics if needed based on the dispersal kernel tail, I included the power exponential function dispersal kernel explored by previous authors (Bourhis *et al.*, 2019; Rieux *et al.*, 2014). I fixed the scale parameter θ as the thin-tailed exponential kernel in this study, and then normalised so that the probability of infection translated to a total probability of 1 for the entire kernel at any given time:

$$\theta_{norm} = \frac{1}{2\pi\theta^2 \Gamma 2}$$

(Eqn 4.2)

In order to reduce the time taken to run epidemic simulations, I used the tau-leap Gillespie algorithm (Gillespie, 2001). This algorithm was defined as:

$$\delta I_T \approx Poisson(S_t \varphi_i(t) \tau)$$

(Eqn 4.3)

Where τ was the time leap between the initial state and updated state of the simulation process, $(T - t)$.

Table 4.1. List of the epidemiological parameter values for chapter four. For each square matrix of parameter combinations, a random landscape modifier (ξ) is assigned. ξ values are selected as a range between 0-1 and then 11 θ values are selected to create a square matrix of parameter combinations for the purposes of analysis (see results).

Random Landscape Modifier (ξ)	Dispersal Distance (θ) (metres)
0	10 m
.1	20 m
.2	30 m
.3	40 m
.4	50 m
.5	60 m
.6	70 m
.7	80 m
.8	90 m
.9	100 m
1	110 m

I modelled the landscapes as $1km^2$ areas containing 1000 hosts, which reflected a realistic plant host density. I initiated epidemics with a single randomly selected host infection at time $t = 0$, continuing until a specified time limit of ten years (t_{max}). This limit was included to account for the occurrence of simulated epidemics where transmission between hosts became effectively zero because of the distance between host clusters; thus, preventing simulations from running indefinitely. The range of dispersal distances (θ) and the transmission

coefficient (β=100) were chosen based on the results of chapter three which reflected realistic parameter values. All parameter values are specified in table one.

4.3.2. Generating landscapes

I developed the landscape patterns by generating clustered landscapes points with a proportionality parameter labelled the randomisation factor ξ that reallocates points using a randomiser function ($\xi = 1$, random landscape; $\xi = 0$, clustered landscape). Each simulation was configured with a unique landscape distribution whilst parameter combinations did not vary. I generated clustered landscape patterns using a Matérn cluster process, a special class of Neyman-Scott processes (Neyman & Scott, 1958). This process defines parent points λ_p and subsequent daughter points λ_d via Poisson estimation. The total set of points on each landscape was given by:

$$\lambda_T = \lambda_{p_1} + \cdots + \lambda_{p_i} = \sum_{i=1}^{\lambda_p} \lambda_{d_i}$$

(Eqn 4.4)

Where both λ_p and λ_d are Poisson distributed. I scaled the radius of the plant host clusters λ_r to the 1 km^2 area and fixed these parameters for subsequent scales. λ_p and λ_d were set to 50 and 25 by default to explore clustering effects.

The reallocation of points within the Matérn cluster process was defined as:

$$(x, y) = (\lambda_{d_i}|X)$$

(Eqn 4.5)

Where λ_{d_i} is replaced by co-ordinates defined as:

$$X \sim \text{Uniform}(0, \sqrt{K})$$

Where *K* is defined as the scale of the landscape in meters squared. The landscape modifier ξ is the proportion of hosts that undergo this transformation.

For simplicity, I removed edge effects from the study, subsequently the total set of points was consistent and remained within the study region.

4.3.3. Surveillance model

I developed a spatially explicit, stochastic surveillance model by randomly sampling hosts at set time intervals using the stored epidemic data generated previously to estimate the mean growth rate of epidemics. At the point of disease detection, I stopped surveillance and then I recorded and stored the total prevalence. I then calculated the average of these prevalence values. The initial sample size per sampling round was 30 hosts (*N*), with the interval between surveillance rounds (Δ) set to 30 (arbitrary time units). The compatibility of the surveillance model with the epidemiological model ensures that time units remain coherent i.e. the time units were arbitrary and required no conversion between models; here time units are assumed as days.

4.3.4. The rule of thumb model

Parnell *'s* rule of thumb was defined as (Parnell *et al.*, 2012):

$$q^* = \frac{r\Delta}{N} \qquad \text{(Eqn 4.6)}$$

Where $q^*$was the theoretical detection prevalence which was determined by the relationship between the growth rate of the epidemic r and the surveillance protocol, Δ was the interval

between surveillance rounds, and N was the sample size taken for inspection each surveillance round. The derivation of this equation was discovered as follows: by the assumption that if an epidemic grows exponentially, then the probability of detecting the disease at any given time was a simple interaction between rate of surveillance and the exponential growth of the epidemic (Parnell *et al.*, 2012). By solving this equation via integration, Parnell *et al.* (2012) arrived at the total probability of detection at the point of detection. Using a random variable transformation of the initial time with an exponential growth assumption, the prevalence at the time of detection was translated from this probability using the Jacobian of the transformation. A full derivation of the approach can be found in Parnell *et al* (2012; 2015). Thereby, the above rule of thumb stated that given the epidemic growth parameter (r), the surveillance frequency (Δ) and the sample size (N), then the prevalence of disease at detection (q^*) was estimated as an approximation of the true prevalence of disease at detection. Key assumptions from the rule of thumb were exponential increase of the pathogen population and continuous monitoring of the host population with time.

For every unique epidemiological parameter combination of θ and β, 2000 simulation runs were performed, providing a unique set of a growth rate values (r). I estimated r by using a linear transformation on the initial growth phase of the epidemic, corresponding to the first quartile of transition events and measuring the mean gradient of the curve. The average of these 2000 r values was then used as input within the rule of thumb (Eqn 5). For N and Δ in the above equation, these values were taken from the surveillance model as defined above.

4.3.5. The comparison metrics

The comparison metrics used were absolute difference and relative difference. The absolute difference was calculated as the difference between q^* and the average epidemic prevalence

at detection. The relative difference was the absolute difference further divided by the average epidemic prevalence.

4.3.6. Nearest Neighbour Index

To approximate the total degree of clustering within the landscapes, The Nearest Neighbour index (NNI) was used. The NNI represents the relative difference between the observed allocation of points and the expected allocation of points under an expected random distribution. This is an expression of the observed mean distance between points to the expected mean distance between points as is defined as:

$$NNI = \frac{D_0}{D_E}$$

(Eqn 4.7)

Where,

$$D_0 = \frac{\sum_{i=1}^{n} d_i}{n}$$

Where d_i represented the distance between each host and its nearest neighbour and,

$$D_E = \frac{.5}{\sqrt{n/A}}$$

Where n corresponded to the number of features in a landscape and A represented the area of the landscape. If the output was less than a value of 1, then the landscape exhibited some degree of clustering.

4.4. Results

Results demonstrated that the accuracy of the rule of thumb model increased with increased dispersal distance and decreased with clustering. This was shown by the rule of thumb prediction of prevalence at first detection and the prevalence of detection within the simulation model (Figure 4.4). The prevalence at detection within the simulation model was generally consistent across the parameter space for the wide range of landscape configurations and dispersal distance combinations explored when the transmission coefficient remained constant (Figure 4.2). The rule of thumb reflected this generally constant trend in prevalence, hence the absolute difference between the rule of thumb prediction and the simulated detection model did not vary significantly across the parameter space (see Figure 4.3). However, the relative difference between the rule of thumb and the simulated detection model decreased as dispersal distance (θ; metres) increased due to the increasing prevalence at detection within the simulation model. Thus the accuracy of the model increased with dispersal distance (Figure 4.4). In addition, the relative difference between the rule of thumb prediction and the simulated detection model increased as clustering increased, thus the accuracy of the rule of thumb model decreased with clustering (Figure 4.4).

Figure 4.2 shows the mean prevalence of detection for each parameter combination across the parameter range (Figure 4.2). These values ranged from 0.01 to 0.05, where ξ represented the host distribution modifier (0=total clustering, 1=total randomness) and θ represented the mean dispersal distance of inoculum (metres). The results indicated that the mean prevalence of detection increased as θ and ξ were increased. One of the parameter combinations ($\xi = 0.1$, $\theta = 90$) was omitted from the analysis because the data was corrupted for this combination. The reason for this data corruption has yet to be determined. Figure 4.3 presents the absolute difference between simulated detection prevalence and the rule of thumb for each parameter

combination across the parameter range (Figure 4.3). The values range between 0.01 to 0.02 (at two decimal points), therefore the impact of ξ on the absolute accuracy of the rule of thumb appeared to be limited, this limited impact appears to be consistent for θ also. However, the heatmaps did indicate a weak relational pattern between decreasing ξ and increasing absolute difference based on heat intensity between 0.01 and 0.02. Parameter combinations were omitted from this analysis as code did not function whilst generating epidemic growth curves for a selection of the lowest dispersal distances (ξ = 0.2, 0.3, 0.4, 0.5, 0.7, θ = 10). Whilst calculating growth rate values, several simulation estimations produced NA values. Why these simulations produced NAs has yet to be determined. Given the omission of the parameter combination in the previous data, it was impossible to calculate the absolute difference for that same point again. Figure 4.4 shows the relative difference in detection estimation between the rule of thumb and the simulated detection model (Figure 4.4). This was estimated as the division of the absolute difference by detection prevalence. The values range from 0.25 to 0.81 and both ξ and θ play a role in determining the relative difference. The heatmap observational analysis indicated that increasing θ and decreasing ξ led to increased relative difference in detection estimation between the rule of thumb and the simulated detection model.

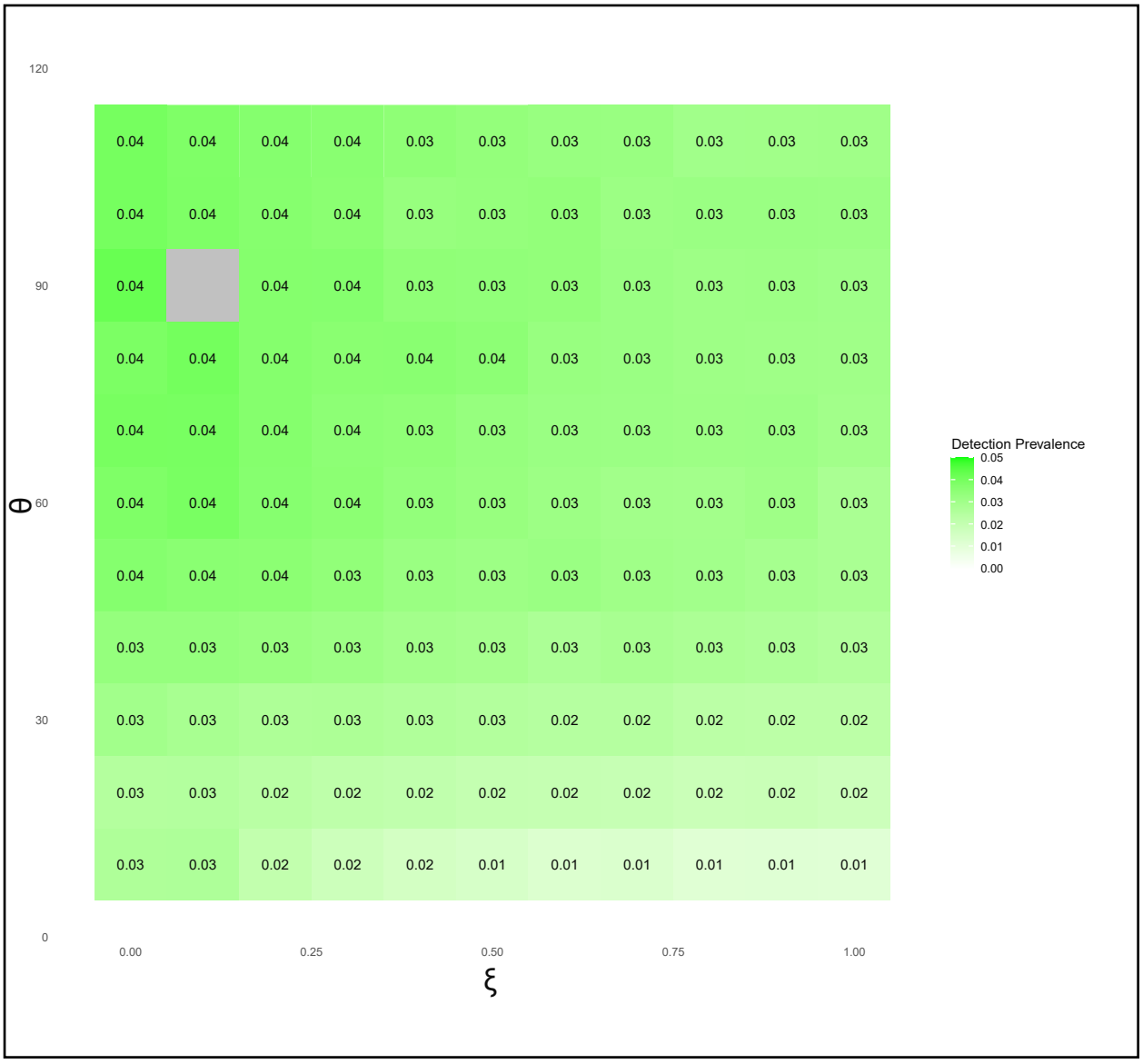


Figure 4.1. The mean prevalence of detection for each parameter combination across the parameter range. Values range from 0.01 to 0.05. ξ represents the landscape modifier (0 = total clustering, 1 = total randomness) and θ represents the mean dispersal distance of inoculum (meters). Surveillance strategy is determined by sample size and sampling frequency interval (N = 30, Δ = 30). Values rounded to two decimal places.

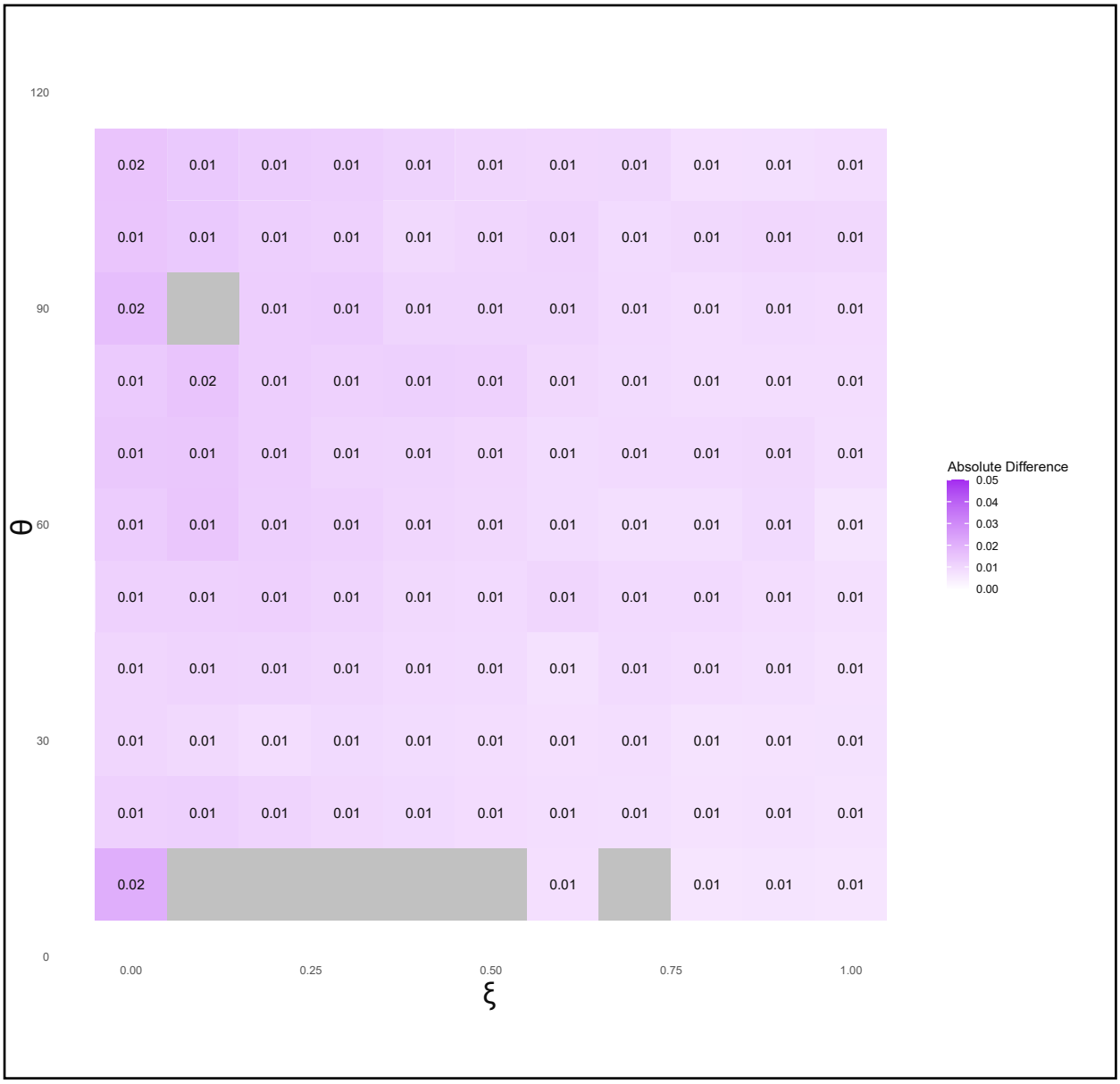


Figure 4.2. The absolute difference in detection estimation between the rule of thumb and the simulated epidemic detection model for each parameter combination across the parameter range. Values range from 0.01 to 0.05. ξ represents the landscape modifier (0 = total clustering, 1 = total randomness) and θ represents the mean dispersal distance of inoculum (meters). Surveillance strategy is determined by sample size and sampling frequency interval (N = 30, Δ = 30). Values rounded to two decimal places.

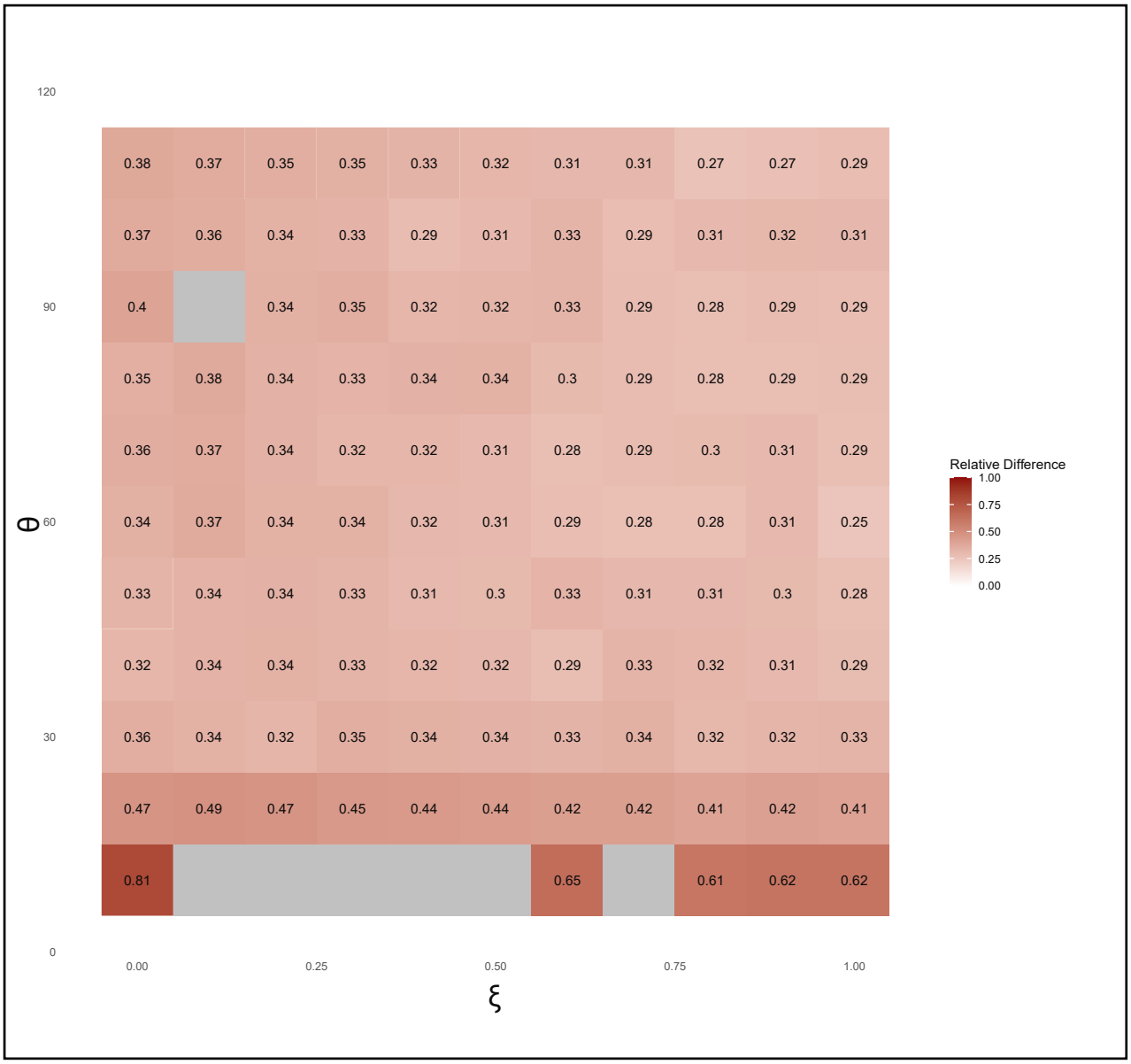


Figure 4.3. The relative difference in detection estimation between the rule of thumb and the simulated epidemic detection model for each parameter combination across the parameter range. Estimated as the division of absolute difference by detection prevalence. Values range from 0.25 to 0.81. ξ represents the landscape modifier (0 = total clustering, 1 = total randomness) and θ represents the mean dispersal distance of inoculum (meters). Surveillance strategy is determined by sample size and sampling frequency interval (N = 30, Δ = 30). Values rounded to two decimal places.

To investigate the degree of clustering in the landscape, the Nearest Neighbour Index (NNI) was used (see Figure 4.5). It is evident from the figure the decrease in clustering observed with increasing ξ.

Figures 4.6-4.9 demonstrated the epidemic behaviour represented spatially, with snapshots at the time of detection for a range of parameters. For epidemics with small values of θ, there was a similarity between the clustered and random landscape configurations. i.e. the influence of the clusters was not as significant due to the limited dispersal of inoculum.

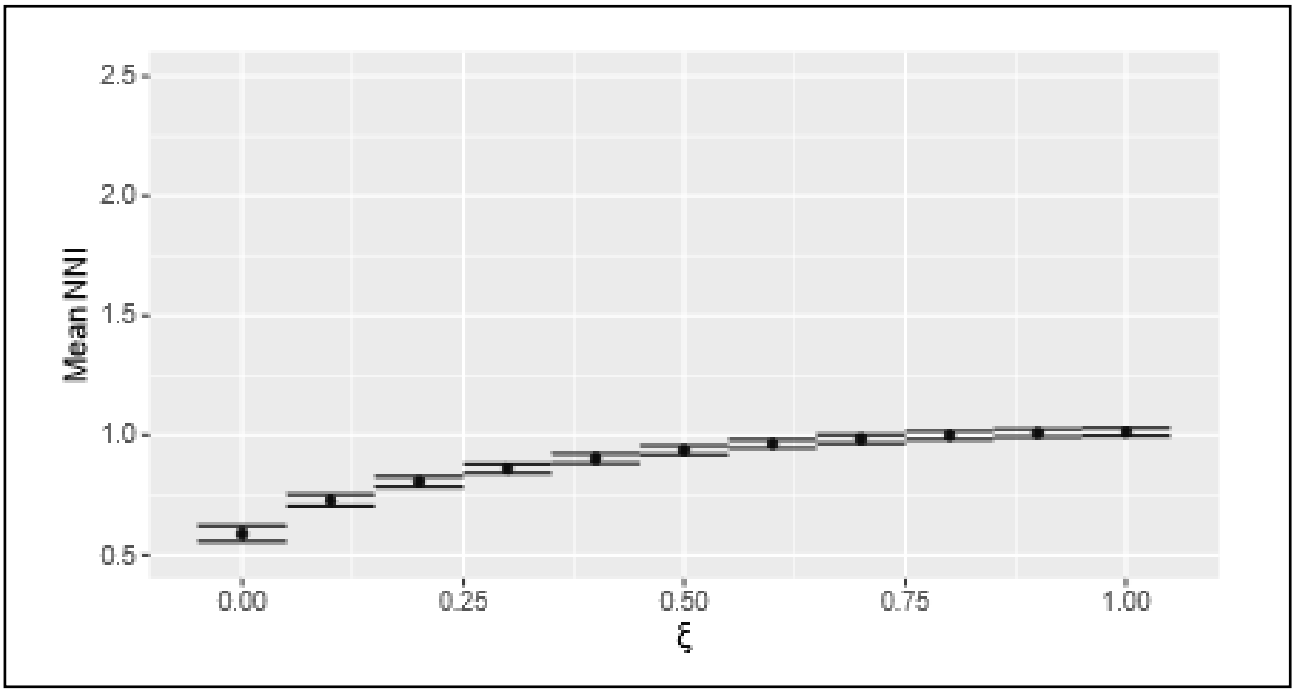


Figure 4.4. The sequence of Nearest Neighbour Index's for the random landscape factor on the 1 km^2 scale. Values less than 1 indicate clustering, however, the accuracy of the NNI in describing clustering depends on several factors, including the scale of the analysis, the spatial distribution of the points, and the assumptions underlying the NNI calculation. Standard Deviation is reported as solid bars above and below the mean NNI.

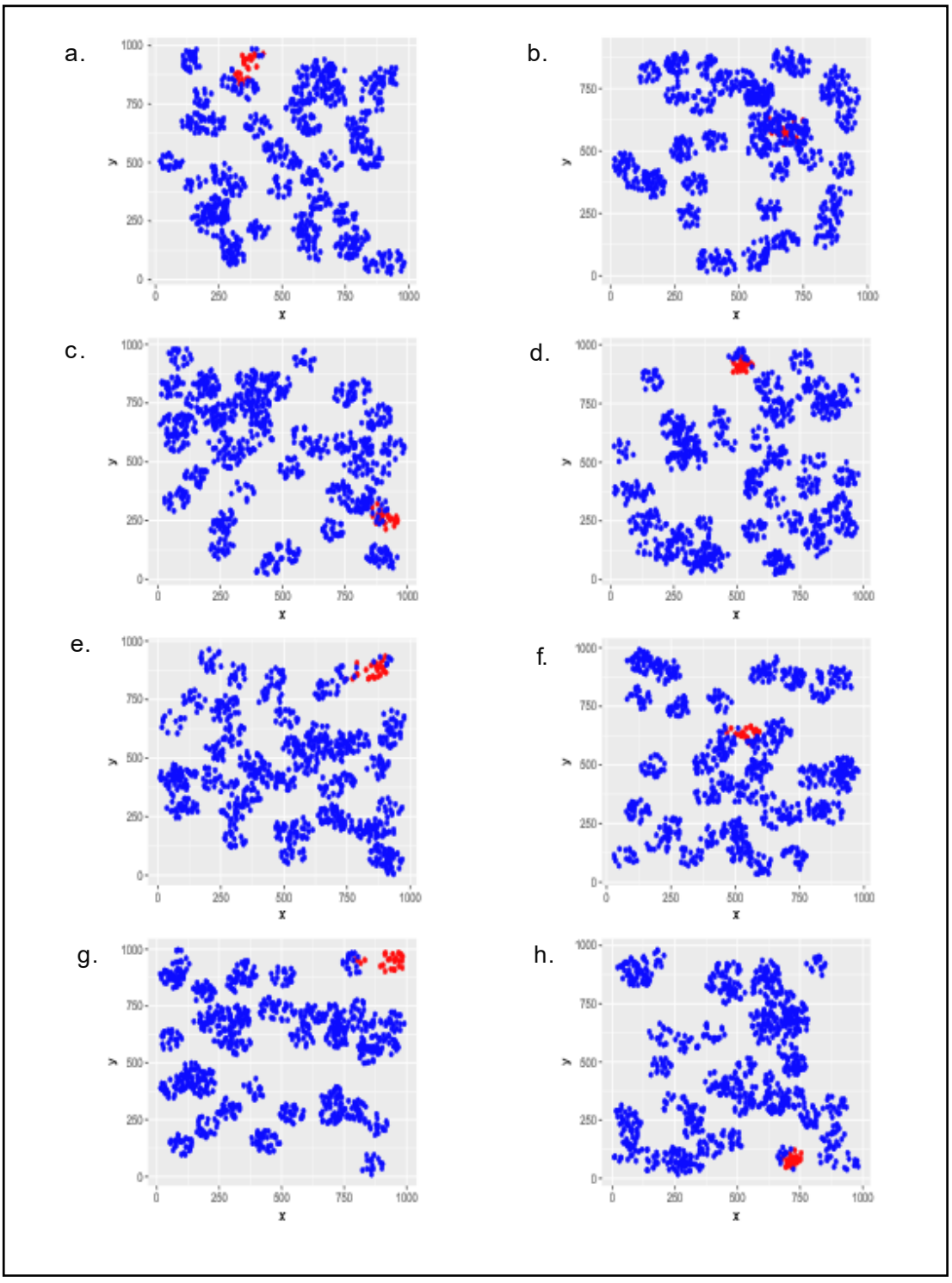


Figure 4.5. A selection of epidemic snapshots at the point of detection for $\xi = 0$, $\theta = 10$. All other parameters including β, Δ and N are fixed ($N = 30$, $\Delta = 30$). Red dots indicate infected hosts at detection, blue dots indicate hosts that remain susceptible at detection. Shown are randomly selected realisations from the epidemic model (a-h) to demonstrate the range of epidemic behaviour.

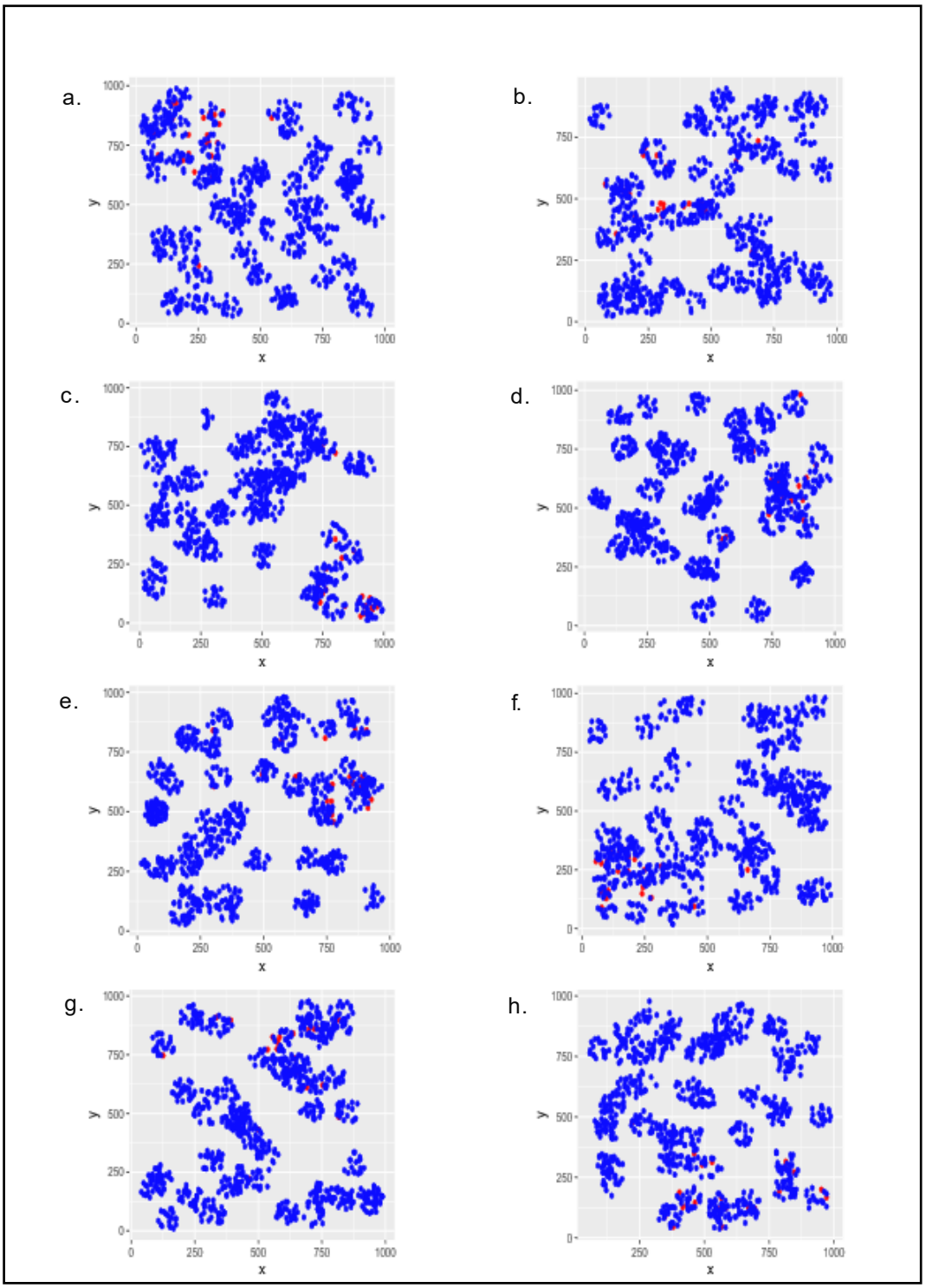


Figure 4.6. A selection of epidemic snapshots at the point of detection for $\xi = 0$, $\theta = 110$. All other parameters including β, Δ and N are fixed ($N = 30$, $\Delta = 30$). Red dots indicate infected hosts at detection, blue dots indicate hosts that remain susceptible at detection. Shown are randomly selected realisations from the epidemic model (a-h) to demonstrate the range of epidemic behaviour.

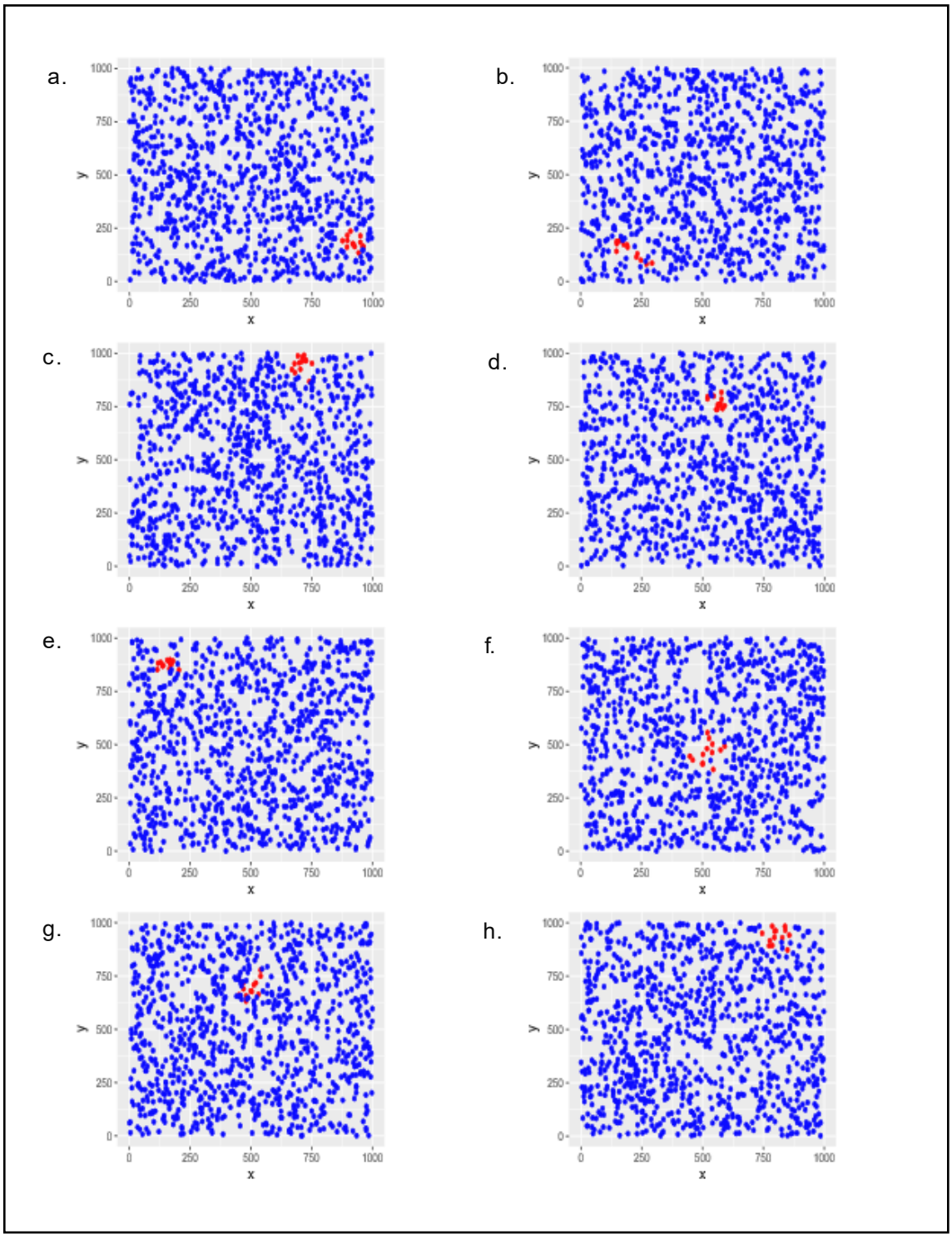


Figure 4.7. A selection of epidemic snapshots at the point of detection for $\xi = 1$, $\theta = 10$. All other parameters including β, Δ and N are fixed ($N = 30$, $\Delta = 30$). Red dots indicate infected hosts at detection, blue dots indicate hosts that remain susceptible at detection. Shown are randomly selected realisations from the epidemic model (a-h) to demonstrate the range of epidemic behaviour.

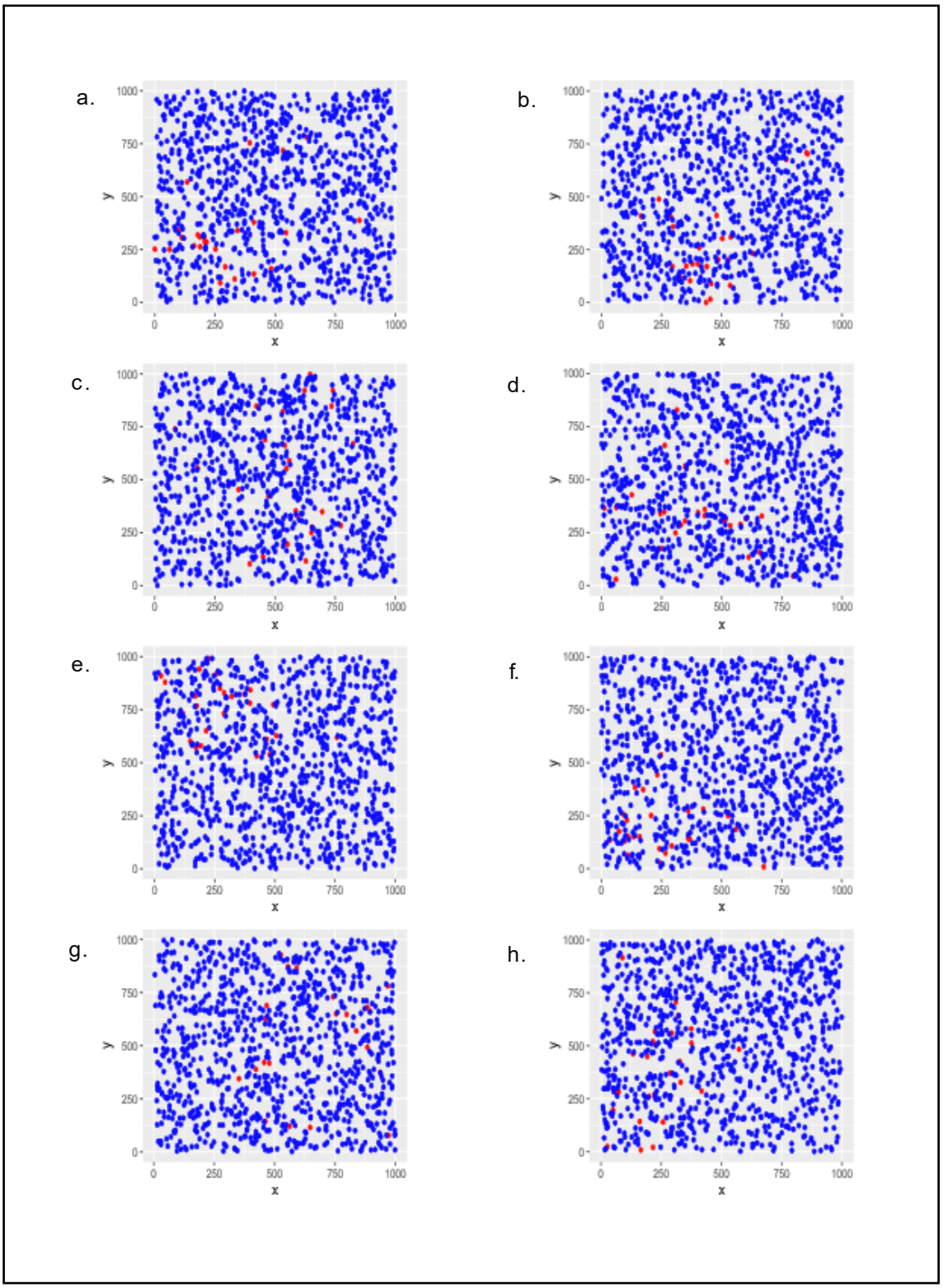


Figure 4.8. A selection of epidemic snapshots at the point of detection for $\xi = 1$, $\theta = 110$. All other parameters including β, Δ and N are fixed ($N = 30$, $\Delta = 30$). Red dots indicate infected hosts at detection, blue dots indicate hosts that remain susceptible at detection. Shown are randomly selected realisations from the epidemic model (a-h) to demonstrate the range of epidemic behaviour.

4.5. Discussion

The objectives of this thesis chapter were to explore the effects of landscape variables and their interactions across a range of epidemics; specifically how these interactions influenced the accuracy of an early detection model (Parnell et al., 2012). These objectives could be addressed by thus answering the following questions:

- How does changing the host pattern impact the ability to predict detection prevalence with simple epidemic models?
- How does changing the dispersal distance parameter within different landscapes, hence reflecting a range of epidemics, impact the ability to predict detection prevalence with simple epidemic models?

The main findings of this chapter indicated that the most influential factor on the accuracy of the "rule of thumb" was the dispersal distance parameter, which specified the distance inoculum would spread. This was described using a special form of the gamma dispersal kernel analogous to the exponential decay model (Bourhis et al., 2019). The mean dispersal distance corresponded to twice the value of θ (metres), where the probability decayed proportional to the distance from source. β (rate of infection) corresponded to the likelihood of a new host being infected upon contact with inoculum, which in this epidemic model translated into the probability of infection for each host given their distance from source. Initially, the extensive experiments across the landscape modifier scale (ξ) demonstrated that as ξ decreased and as θ decreased, there was an observable reduction in the accuracy of the rule of thumb as described by the relative differences (Figure 4.4). This observable reduction was consistent though the absolute differences remained almost entirely constant across the parameter range (Figure 4.3). The reasoning behind the range of relative differences is that the even though the prevalence of epidemic at detection increased as ξ increased and θ

increased, the absolute differences remained constant, hence reducing the overall relative drop in accuracy of the "rule of thumb". The implication of this finding contrasts previous literature that indicates modelling long distance disease dispersal is more stochastic and more unpredictable as compared to short distance disease dispersal (Robinet et al., 2012). In respect to restricting this study to the power exponential form of the gaussian kernel; the kernel chosen for this study (the gaussian kernel) may need to be parameterised with different shape values to improve realism when modelling the spread of specific plant pathogens (as well as to provide additional meaningful insights). Were the rule of thumb to be applied to a wide distribution of kernels, unravelling the nature in the increase in accuracy of the rule of thumb with larger dispersal kernels would be less challenging. For now, it is hypothesised that the change in the dispersal kernels scaling parameter does not significantly alter the behaviour of epidemic expansion (Farber *et al*., 2019). If the scaling of the epidemic expansion does not significantly affect the resulting prevalence pattern beyond the increase in prevalence, the rule of thumb models assumptions will continue to hold and this will be reflected in the increased accuracy with increasing prevalence. The clustering behaviour reduced the accuracy of the rule of thumb because host aggregation may not be accounted for within the rule of thumb model. Previous research has indicated that edge effects have a significant impact on the final prevalence status of a diseased crop or plant landscape (Johnson & Haddad, 2011). Johnson & Haddad demonstrated that crops that were patchier would have lower disease prevalence (albeit specifically a fungal pathogen case study). Following this logic, were there no significant model assumption violations, a lower prevalence would predict a greater decrease in accuracy of the rule of thumb model when observing relative differences.

According to the model this effect was observed, indicating that the clustering of hosts in a landscape will decrease the rule of thumb's accuracy without violating the assumption of exponential growth (Figure 4.4). This is further supported by the constant absolute differences across the landscape aggregation parameter space (Figure 4.3). Again, the epidemiological model deployed in this study was relatively simple if representative of a large group of epidemics. How the rule of thumb model would perform with further compartmentation of individual host status would more thoroughly explore the effects of landscape dynamics on the accuracy of the rule of thumb.

The relative amount of clustering was indicated by the Nearest Neighbour Index (NNI), demonstrating a relationship between clustering and the value of ξ across the landscape range (Figure 4.5). The NNI analysis demonstrated that the effects of clustering decreased with increasing ξ. Values less than 1 indicate some degree of clustering, though the NNI indicates some degree of variability. The compounded effect of less clustering leads to additional accuracy in the rule of thumb model. An additional experiment was performed to explore the effects of clustering on differing scales of landscape for different values of β, however the cluster formations using the previous clustering parameter values were smaller on larger landscapes and hence removed from the experiment and consequently is in the appendix (Appendix, Figure 4.10, 4.11a & b). It is important to note that the impact of β on the accuracy of the rule of thumb was limited; β had some degree of influence on the $1km^2$ scale but was restricted for larger landscape scales (Appendix, Figure 4.11a & b). This finding indicates that for long dispersal distance events, the transmission coefficient will not greatly impact the rule of thumb approximation, hence increasing transferability across epidemics. A further consideration is that the normalisation constant within this epidemiological model represents falling inoculum between hosts, further decreasing the effects of the transmission coefficient between hosts on larger scales. This is considered largely realistic behaviour and

hence reinforces the finding that long distance dispersal of inoculum will reduce the impact of the transmission coefficient on the accuracy of the rule of thumb model. This has been investigated previously with similar findings (Arias et al., 2018). However, it is important to note that until the clustering behaviour in the landscape is standardised, it is difficult to draw any meaningful conclusion of clustering behaviour on larger scales.

Figures 4.6-4.9 demonstrated a series of 8 random iterations for contrasting sets of epidemiological relevant parameters (Figures 4.6-4.9) The parameters were selected at the maximum and minimum range values to demonstrate the contrast in prevalence at detection. When explored as a set of 2000 iterations per parameter combination, Figure 4.6 and 4.8 revealed that some change in the accuracy of the rule of thumb occurred when changing between a homogeneous landscape and a heterogenous landscape on a shorter scale, as interpreted visually (Figure 4.6 & 4.8). How influential this effect of landscape clustering is likely dependent on the epidemiological parameters of the epidemiological model used.

In reflecting on the findings of this chapter, I believe it is pertinent to acknowledge that the effects of landscape clustering might also be impacted by the surveillance intensity. If a range of surveillance frequencies and sample sizes are to be deployed, then the effects of clustering on the accuracy of early detection models might become more prominent because of the inoculum of infection becoming trapped within the clusters. I propose that a future investigation into the effects of changing surveillance intensity in tandem with changing landscape aggregation would be of benefit to future researchers.

The impact of θ on the accuracy of this simple epidemic model was explored thoroughly in the chapter and was determined to be the most influential parameter. The effects of dispersal parameters in spatially stochastic models have been explored previously and demonstrated to have a significant effect on the dynamics of epidemics (Filipe & Maule, 2004; Arias et al.,

2018). For example Arias *et al*., (2018) demonstrated that longer dispersal kernels promote the proliferation of epidemics and Filipe & Maule, (2004) observed similarly that epidemic size is proportional to the size of the dispersal distance parameter (Arias et al., 2018; Filipe & Maule, 2004). Given the importance of θ in determining the accuracy of the early detection model, future research could focus on refining understanding of how this parameter varies across different landscapes and epidemic scenarios by collecting data on dispersal distances in the field or using more sophisticated modeling approaches to simulate different types of epidemics. Moreover, as noted by Filipe and Maule (2004) and Arias et al. (2018), the effects of dispersal parameters on epidemic dynamics can be complex and context-dependent (Filipe & Maule, 2004; Arias et al., 2018). Therefore, future studies could also investigate how other factors, such as host population densities or environmental conditions, interact with dispersal parameters to shape disease spread and control efforts.

Overall, this chapter provided valuable insights into the role of landscape variables and their interactions in shaping the accuracy of early detection models for epidemics. Of particular importance are the influences of both dispersal distance and clustering dynamics. Therefore, the results are overall indicative that the rule of thumb model maintains a high degree of accuracy across a range of epidemics. By building on these findings, researchers and stakeholders can work towards developing more effective strategies for preventing and controlling disease outbreaks in a variety of contexts. Further research efforts should focus on varying the range of compartmental models such as SIR and SIS models and a range of dispersal kernel behaviours to further elucidate the accuracy of early detection models such as the rule of thumb within changing landscape environments.

5. Chapter five: Simple models for complicated epidemics - exploring the interactions of surveillance and epidemiological parameters in parsimonious models to inform early detection surveillance

5.1. Chapter summary

The early detection of pests in a plant population, such as a crop or a forest, is essential to mitigating significant losses at regional or national levels through effective plant pest management. Plant pest modellers concern themselves with reductions in the impact of these pests by informing stakeholders with appropriate action. Currently, there is much investigation into early detection models that integrate surveillance intensity with the epidemiological growth rate of an epidemic to inform stakeholders what the prevalence of pest will be when first detected. This study explores the accuracy of two early detection models by changing surveillance intensity, detection method sensitivities, and resource allocation with the use of a spatially-explicit stochastic epidemiological simulation model. My findings reveal that the accuracy of the early detection models are sensitive to both changing the epidemiological parameters transmission coefficient (β; rate of infection) and dispersal distance(θ; metres) and changing the surveillance intensity (frequency Δ and sample size N). My findings also reveal that one of the early detection models produced higher accuracies when accounting for changing detection assay sensitivity. These findings could inform stakeholders how changing surveillance intensity will affect the accuracy of the prediction of pest prevalence across a broad range of epidemics. The insights from this chapter could also inform stakeholders to a preferred modelling approach when accounting for detection assay sensitivity. As such, my research should contribute to improved biosecurity strategies and predictive modeling approaches for effective plant pest management.

5.2. Introduction

Plant pests are a substantial, global problem, environmentally, economically and in terms of global food security (Dun-chun *et al.*, 2021; Burdon *et al.*, 2020). The importance of effectively managing plant diseases is stated in Chapter 3. Pests can be defined as insects, other invertebrate, bacteria, fungi, viruses or pathogens which affect the health of plants or plant products (Defra, 2023). Current pests that are of increasing global concern include Emerald Ash borer (*Agrilus planipennis*), Banana Fusarium Wilt (*Fusarium oxysporum*) and the bacterium *Xylella fastiodisa* (IPPC Secretariat, 2021; Buja *et al*., 2020; Volkavitsh *et al*., 2021; Pegg *et al*., 2019). The early detection of these plant pests is crucial for the successful management and control of destructive pests (Buja *et al*., 2021). However, predicting the performance of surveillance strategies, and thus ensuring appropriate surveillance design, is not trivial. Recently, modelling approaches have been developed as a potential solution to improve our ability to predict the performance of plant pest detection strategies (Newlands, 2018). Increasing the accuracy of plant pest detection models increases their applicability to managing novel pests that are introduced because of, in increasing importance, ongoing evolutionary processes, climate change and globalisation of trade (IPPC Secretariat, 2021; Spence, 2020; Parker and Gilbert, 2004). Increasing the accuracy of plant pest detection models is an important issue for NPPOs and relevant plant stakeholders impacted by the negative effects of plant pest epidemics (Newlands, 2018; Almieda, 2018). Plant pest detection models need to be flexible enough to be responsive to different scenarios whilst still producing consistent and accurate estimations. If there is a potential for reductions in accuracy, stakeholders should be informed beforehand (Newlands, 2018). The appropriate use of surveillance parameters in a plant pest detection model must also be considered; these are contextual and dependent on the intended use of the model. Examples of plant pest surveillance parameters vary; in the context of plant pest detection models, parameters

include the number of plants sampled on a day, the frequency of days in which surveillance is conducted and the sensitivity of the surveillance method (Parnell *et al*., 2012; Parnell *et al*., 2015; Mastin *et al*., 2020). Previous early detection models have utilised these parameters as foundations to mathematical analysis of pest prevalence when first detected (Parnell *et al.*, 2012; Parnell *et al.*, 2015; Mastin *et al*., 2020). As stated previously, one of the most documented early detection models is the rule of thumb derived by Parnell *et al.* (Parnell et al., 2012). This model accounts for sampling efforts and epidemiological growth rates to predict how much pest will be present when a pest is first detected (Parnell *et al.*, 2012). Mastin *et al.* (2019) expanded on this model to account for sensitivity structure (Mastin *et al.*, 2020).

Chapters three and four investigated the influence of epidemiology and landscape structure on the ability to predict the plant pest detection using the rule of thumb. However, this also depends on the implementation of the surveillance program itself. In practice surveillance programs differ markedly depending on the resources available to an NPPO, the awareness of the pest threat and logistical issues that influence the frequency and intensity of surveillance resources (Moore *et al.,* 2010; Barnes *et al.,* 2019; Cacho *et al.,* 2010). In addition, they also differ in the sensitivity of the detection methods available for the pest (EFSA, 2020). Assays used for detecting plant pests differ in their effectiveness to identify these pests, referred to as the assay's sensitivity. Choosing the most appropriate assay is often dependent on the scale and objectives of assessment; visual assays may be less sensitive than molecular techniques such as Electrochemical Impedance Spectroscopy (EIS) when detecting asymptomatic pests (Mohd Ali *et al.,* 2019). However, when there is a need to monitor the progression of symptoms or minimise costs for surveying pests more frequently, visual assays are often more suitable (Buja *et al*., 2020; Balodi *et al*., 2017; Mastin *et al*., 2020). Therefore, the ability to detect a pest based on the sensitivity of the assay chosen is contingent on the

general objective of the chosen surveillance strategy. Recently, investment has been directed to innovations in novel detection techniques, with an emphasis on standardising assays that are fast, accurate and inexpensive (Buja *et al*., 2020). These assays work effectively only when strategy is designated prior to deployment to account for pest progression, which should be based on appropriate modelling (Parnell *et al*., 2017; Bebber & Gurr, 2015; Cunniffe *et al*., 2015). Ultimately, the overall sensitivity of a detection method will depend on both the sampling effectiveness (the probability to collect infected samples from an infected host unit) and the diagnostic sensitivity (the probability to correctly identify an infected sample as infected) (EFSA, 2020).

Therefore, the research questions this thesis chapter seeks to address are:

- How much does the sensitivity of detection methods impact the accuracy of simple early detection models such as the rule of thumb when using a spatially-explicit stochastic epidemiological model (Parnell et al., 2012)?
- How do epidemiological phenomena interact with these simple early detection model surveillance modifications?
- How accurate is the Mastin *et al.* derivation of the rule of thumb when applied to a spatially-explicit stochastic epidemiological model (Mastin *et al.*, 2020)?
- Is the accuracy of the early detection models affected by a restricted pool of surveillance resource allocation?

To address these research questions, in this chapter I will expand upon the spatially-explicit stochastic epidemic simulation model developed in chapters three and four. This model is sufficiently complex enough to include epidemic behaviours i.e. the dispersal θ and transmission coefficient β, whilst flexible enough to represent a wide range of plant pest epidemics (Bourhis *et al*., 2019). I will also modify the model to include a range of detection

sensitivities. Furthermore, I will explore varying surveillance strategies given a limited set of surveillance resources. I will be performing the spatially-explicit stochastic simulation model in tandem with a spatially-explicit stochastic surveillance model to simulate the prevalence of pests at detection and compare this against Parnell *et al.* early detection model for a range of surveillance strategies (Methods, Figure 5.1). Furthermore, I will be investigating the Mastin *et al.* rule of thumb model with sensitivity modification for a range of detection sensitivities within the surveillance model; by doing so expand the understanding of the applicability of these early detection models across myriad epidemics. To my knowledge, this has not been explored previously.

Furthermore, I will integrate pest dispersal and likelihood of infection based on spatial constraints within this model thereby making it possible to explore pest and surveillance interactions over time and space (Rohani & Keeling, 2013). For the purposes of simplicity, I will use an exponential dispersal kernel to define the range of expected epidemics, which is commonly referenced in the literature to approximate the spread of plant pests (Bourhis *et al*., 2019). Therefore, the model developed for this research will demonstrate realism within a range of epidemics that could potentially prepare appropriate stake holders for novel emerging plant epidemics. Given the precedent of biosecurity standards many NPPOs wish to encompass, this research chapter could contribute useful predictive insights that will assist strategic decision making within the context of plant pest management.

5.3. Methods

5.3.1. Epidemiological model

The epidemiological model used in this chapter for the purposes of simulating epidemics was identical to the model used in previous chapters. The model tracked plants represented as points in space, where host status could either be susceptible (*S*) or infected (*I*) within a

simple SI compartmental model. If host *i* was susceptible at time *t*, then the probability of infection of *i* was given by:

$$\varphi_i(t) = \beta \sum_j K(d_{ij}; \theta)$$

(Eqn 5.1)

$\phi_i(t)$ was the rate of infection at time *t*, β was the transmission coefficient that defined the rate for any given distance within the kernel, and the summation ran over all infectious hosts *j* at time *t*. The dispersal $K(d_{ij}; \theta)$ set the rate of disease transmission between a pair of hosts separated by distance d_{ij}, and was parameterised by the dispersal scale parameter θ. Therefore, β is a rate and θ has the units of distance. To allow robustness in dispersal mechanics if needed based on the dispersal kernel tail, I included the power exponential function dispersal kernel explored by previous authors (Bourhis *et al.*, 2019; Rieux *et al.*, 2014). I fixed the scale parameter θ as the thin-tailed exponential kernel in this study, and then normalised so that the probability of infection translated to a total probability of 1 for the entire kernel at any given time:

$$\theta_{norm} = \frac{1}{2\pi\theta^2 \Gamma 2}$$

(Eqn 5.2)

In order to reduce the time taken to run epidemic simulations, I used the tau-leap Gillespie algorithm (Gillespie, 2001). This algorithm was defined as:

$$\delta I_T \approx Poisson(S_t \varphi_i(t) \tau)$$

(Eqn 5.3)

Where τ was the time leap between the initial state and updated state of the simulation process, ($T - t$).

The mean dispersal distance for the exponential kernel was equal to twice the value of the θ (metres) parameter, as given by the equation stated in Bourhis *et al.*, (2019):

$$\phi = \frac{\theta\Gamma(\frac{3}{b})}{\Gamma(\frac{2}{b})}$$

(Eqn 5.4)

Where θ is the dispersal parameter, b is the shape parameter, Γ is the gamma function and φ is the mean dispersal distance.

5.3.2. Simulating Host Landscapes

1000 hosts were given co-ordinates drawn from a uniform distribution in a 1km^2 area using the R environment using the spatstat package (Baddeley & Turner, 2005). This process was repeated for 2000 iterations per epidemiological parameter combination. The chapter followed the previously established generation of clustered landscapes and reallocation of points according to random landscape factor ξ set to 1 (Chapter 3.3 methods). This was the chosen method given time constraints.

5.3.3. Storing epidemic growth data

I initiated epidemics with one random host infection at time $t = 0$, and the initial host was randomly selected for each epidemic run, which continued until all hosts were infected or the maximum time had been reached (t_{max} = 3650 days). I selected a range of realistic dispersal distances (θ) and transmission coefficients (β) based on the findings of chapter three and four (Table 5.1 & 5.2). These findings were based on realistic epidemic growth curves as observed

in literature (Alonso Chavez, 2016). For example, the range of r selected in this chapter were between 0.011-0.022, which would represent the growth rate of Citrus Canker. These values were produced through iterative exploration of the model initially, to determine growth rates that reflected real proliferation of plant diseases with suitably realistic dispersal. After epidemics were stopped, epidemiological data was stored and the mean growth rate of each set of 2000 simulations was estimated. I estimated the epidemiological growth rate (r) by using a linear transformation on the initial growth phase of the epidemic, corresponding to the first quartile of transition events and measuring the mean gradient of the curve. The average of these 2000 r values was then used as input within the rule of thumb models (Eqn 5 & Eqn 6). For N and Δ in the above equation, these values were taken from the surveillance model as defined above.

5.3.4. Surveillance model

A spatially stochastic surveillance model was developed by randomly sampling hosts at set time intervals using the stored epidemic data generated previously to estimate the mean growth rate of epidemics. At the point of pest detection, I stopped surveillance and the prevalence at the time point of detection was recorded and stored. An average of these prevalence values was then calculated for each set of 2000 simulations. The compatibility of the surveillance model with the epidemiological model ensured that time units remain coherent i.e. the time units were arbitrary and required no conversion between models; here time units are discussed as if days.

5.3.5. Detection Method Sensitivity

To investigate the degree to which sensitivity of assays affect the accuracy of the rule of thumb model, I included a sensitivity modifier within the detection procedure. This modifier, σ, stated that should a host be inspected for its pest status and the plant was already infected

then the probability of an inspection detecting the pest was determined by a probability between 0 and 1.

5.3.6. Varying surveillance frequency and intensity

To investigate the effect of different sampling factors, I varied sampling size and sampling frequency whilst keeping the surveillance intensity fixed at one (i.e. so that the average number of samples per day is equal to one, but how they are distributed over time was allowed to vary). Specifically, I explored several factors of 180 as my maximum sampling size and frequency: including 1, 2, 3, 12, 36, and 180. By doing so, it was possible to examine how changes in the sampling rounds whilst intensity was fixed affected the accuracy of the rule of thumb model.

5.3.7. The rule of thumb model

Parnell *et al*.'s rule of thumb was defined as (Parnel *et al*., 2012):

$$q^* = \frac{r\Delta}{N}$$

(Eqn 5.5)

Where $q^*$was the theoretical detection prevalence which was determined by the relationship between the growth rate of the epidemic r and the surveillance protocol, Δ was the interval between surveillance rounds, and N was the sample size taken for inspection each surveillance round. The derivation of this equation was discovered as follows: by the assumption that if an epidemic grows exponentially, then the probability of detecting the disease at any given time was a simple interaction between rate of surveillance and the exponential growth of the epidemic (Parnell *et al.*, 2012). By solving this equation via integration, Parnell *et al.* (2012) arrived at the total probability of detection at the point of

detection. Using a random variable transformation of the initial time with an exponential growth assumption, the prevalence at the time of detection was translated from this probability using the Jacobian of the transformation. A full derivation of the approach can be found in Parnell et al (2012; 2015). Thereby, the above rule of thumb stated that given the epidemic growth parameter (r), the surveillance frequency (Δ) and the sample size (N), then the prevalence of disease at detection (q^*) is estimated as an approximation of the true prevalence of disease at detection. Key assumptions from the rule of thumb are exponential increase of the pathogen population and continuous monitoring of the host population with time.

5.3.8. The sensitivity rule of thumb model

The sensitivity rule of thumb model was derived based on the formulation provided in Mastin *et al.* (2019), including the sensitivity of the assay within the underlying assumption of the model (Mastin *et al*., 2020). This equation is modelled as:

$$q^{sen} = \frac{r\Delta}{N\sigma}$$

(Eqn 5.6)

5.3.9. The comparison metrics

To compare the differences in accuracy between the rule of thumb models and the simulated detection prevalence, I used the indices of absolute difference and relative difference. I calculated the absolute difference as the difference between $q*$ or q^{sen} and the average epidemic prevalence at detection. The relative difference was the absolute difference further divided by the average epidemic prevalence. Therefore, the absolute difference informs the reader as to the difference between epidemic sizes regardless of the total epidemic size itself.

Whereas, the relative difference also took into account the relative size of the epidemic at detection, which may be more heavily influenced by the epidemiological parameters within the spatially-explicit stochastic epidemiological model.

Table 5.1. List of epidemiological parameter values for chapter five. Two transmission coefficient (β; rate of infection) values and two dispersal distance (θ; metres) values are selected and for each epidemiological parameter combination, a range of surveillance frequencies (Δ) and sample sizes (N) are selected.

β	θ
50	10
150	70

Table 5.2. List of the surveillance parameter values for chapter five. Frequency (Δ) is in days and sample size (N) is individual plant units.

Δ	N
15	15
30	30
45	45
60	60
75	75
90	90
105	105

Table 5.3. List of the sensitivity values for chapter five. σ values are selected as a range between 0.1-1 and then three θ values are selected to create figures for analysis.

Sensitivity of assay (σ)
.1
.2
.3
.4
.5
.6
.7
.8
.9
1

Table 5.4. List of the sampling protocols selected for this chapter. Sample rounds with a fixed intensity ($N/\Delta = 1$) are varied by choosing a maximum N and Δ of 180 (approximately equal to a sample round once every half a year) and then calculating a range of factors of 180, so that sampling intensity remained constant.

Sample rounds (*Fixed intensity*, $N/\Delta = 1$)
1
2
3
5
12
36
180

Table 5.5. List of dispersal distance parameter values (θ; metres) selected for this chapter.

Dispersal Distance (θ)
10 m
60 m
110 m

5.4. Results

My findings indicated that there were interactions between the surveillance parameters and the epidemiological parameters when evaluating the accuracy of Parnell *et al*.'s rule of thumb model (Parnell *et al*., 2012) (Figures 5.2 & 5.3). Furthermore, Figures 5.4 and 5.5 demonstrated a relationship in between detection sensitivity (σ) and Parnell *et al*.'s rule of thumb model and Mastin *et al*.'s model as indicated by the absolute and relative differences (Figures 5.4b & 5.5b). In addition, Figure 5.4 and 5.5 explore varying sampling rounds with fixed sampling intensity (Figures 5.4a & 5.5a). These figures collectively provide a comprehensive overview of the performance and implications of different early detection models with regards to surveillance strategy across a broad range of epidemics.

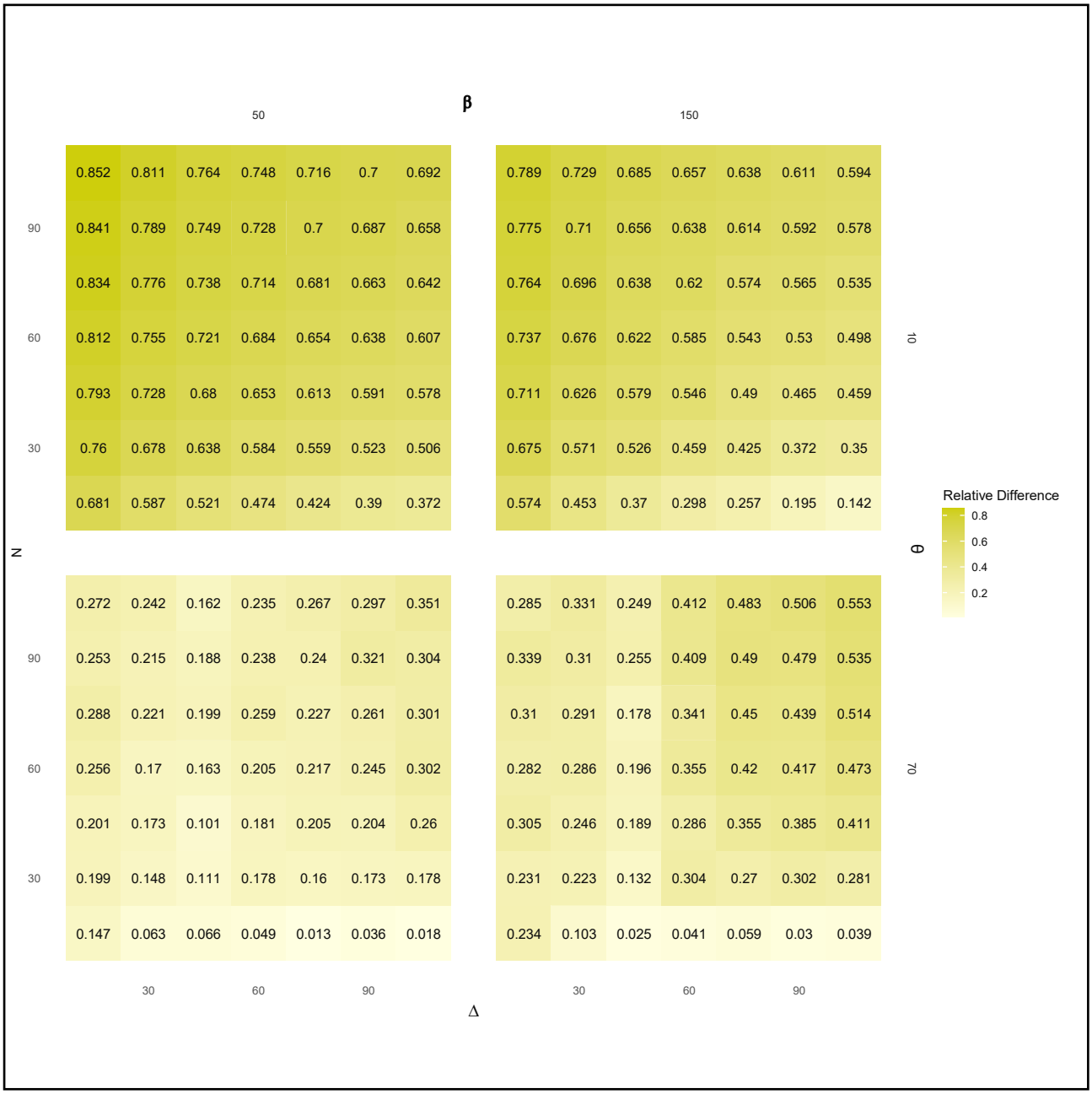


Figure 5.1. The relative differences between the mean detection prevalence for the rule of thumb predictive model and the simulated detection model for different samples sizes (N) sampling intervals (Δ). (β = 50, 150; θ = 10, 70).

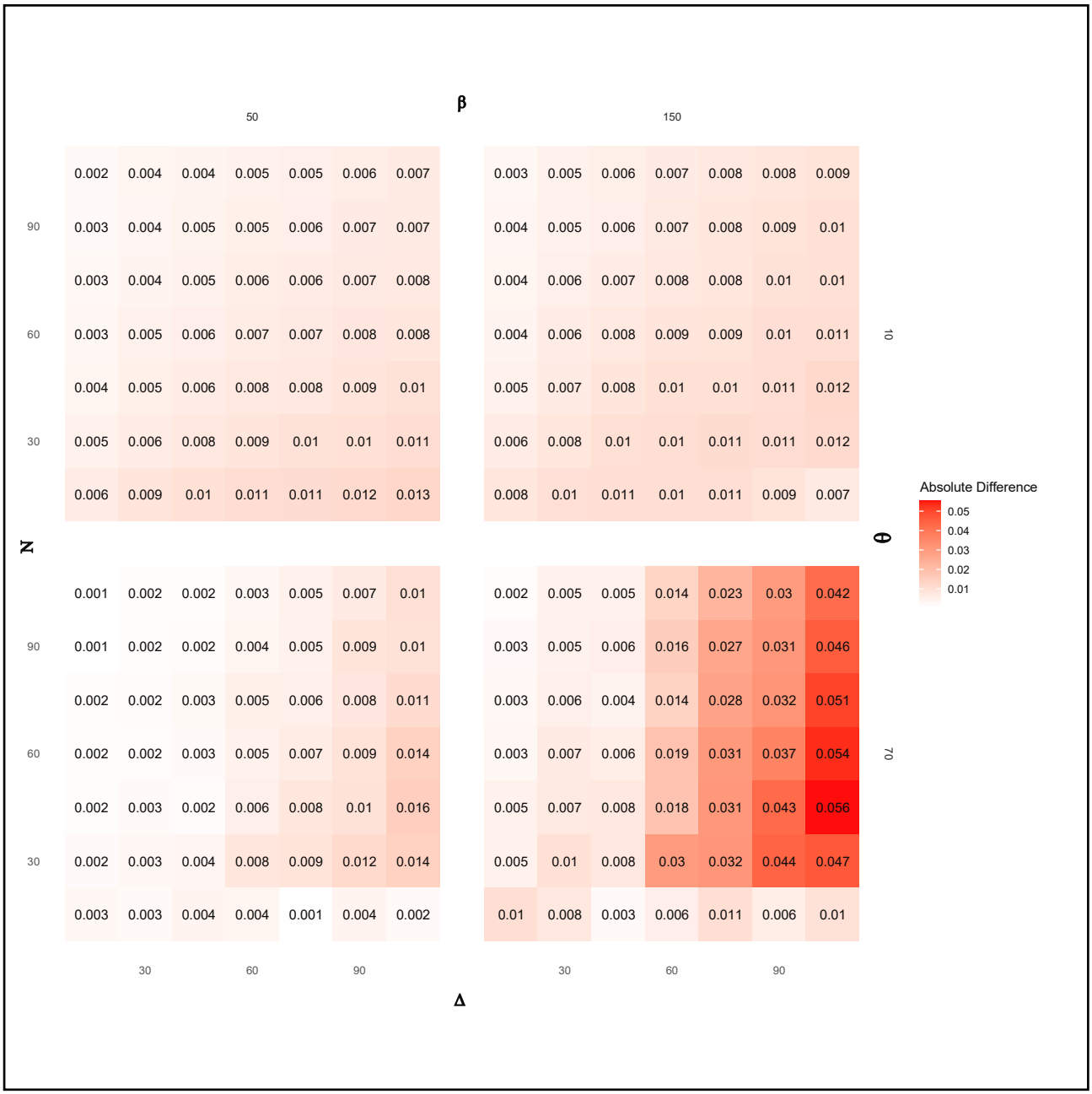


Figure 5.2. The absolute differences between the rule of thumb predictions and the simulated detection prevalences for different samples sizes (N) and sampling intervals (Δ). (β = 50, 150; θ = 10, 70).

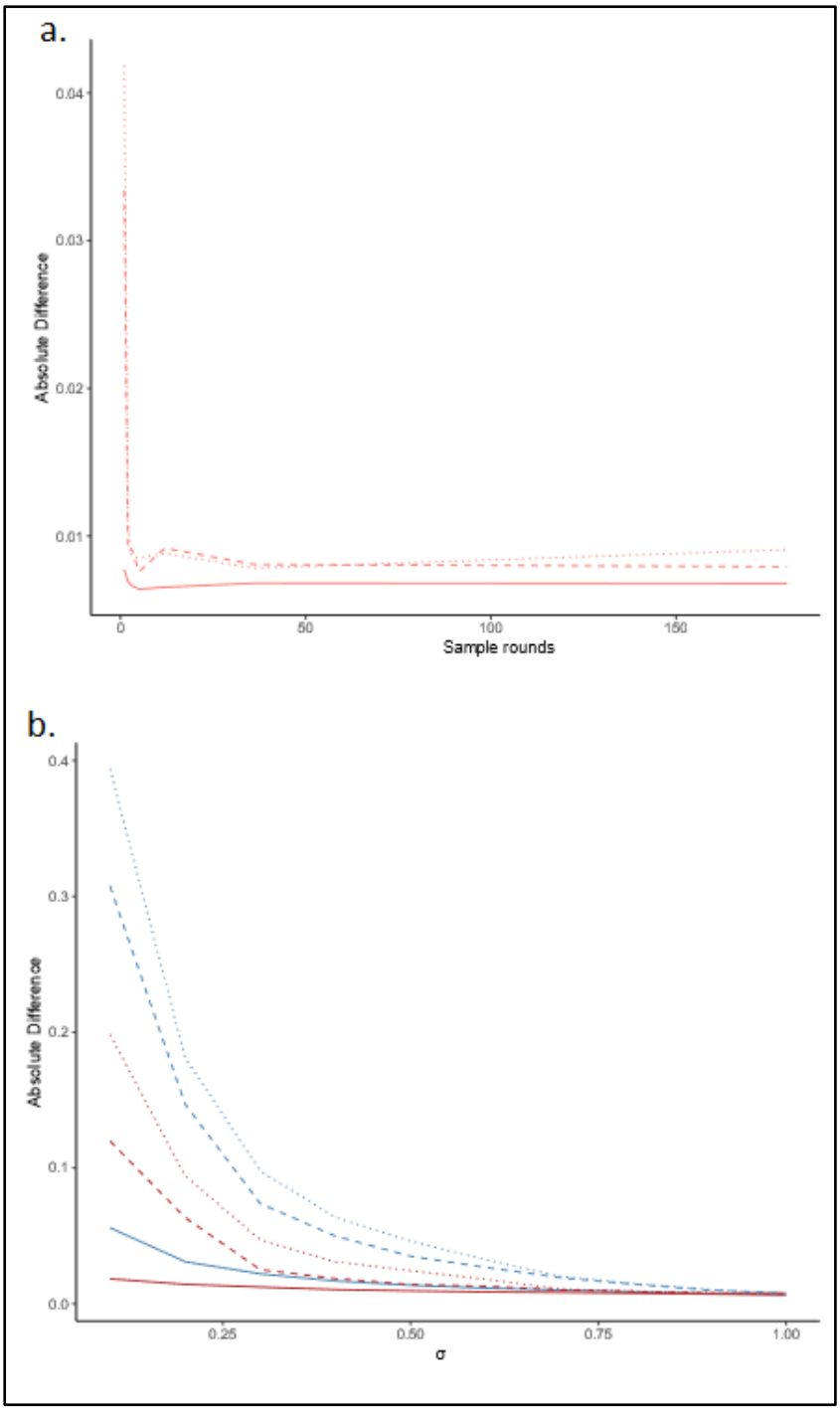


Figure 5.3. Absolute differences in the mean detection prevalence. a. the relationship between sampling allocation with fixed surveillance intensity and the absolute difference between Parnell *et al*.'s rule of thumb and the estimated simulated prevalence of pest at detection. Solid line is θ of 10, dashed line is θ of 60 and dotted line is θ of 110. b. the relationship between the detection sensitivity (σ) and the absolute difference between the estimated

simulated detection prevalence and the two models (Mastin *et al*.'s and Parnell *et al*.'s) for three different θ values. Solid line is θ of 10, dashed line is θ of 60 and dotted line is θ of 110. Blue line is Parnell *et al.*'s rule of thumb and red line is Mastin *et al.*'s rule of thumb. Sensitivity of detection assay within the simulation σ is on the x-axis.

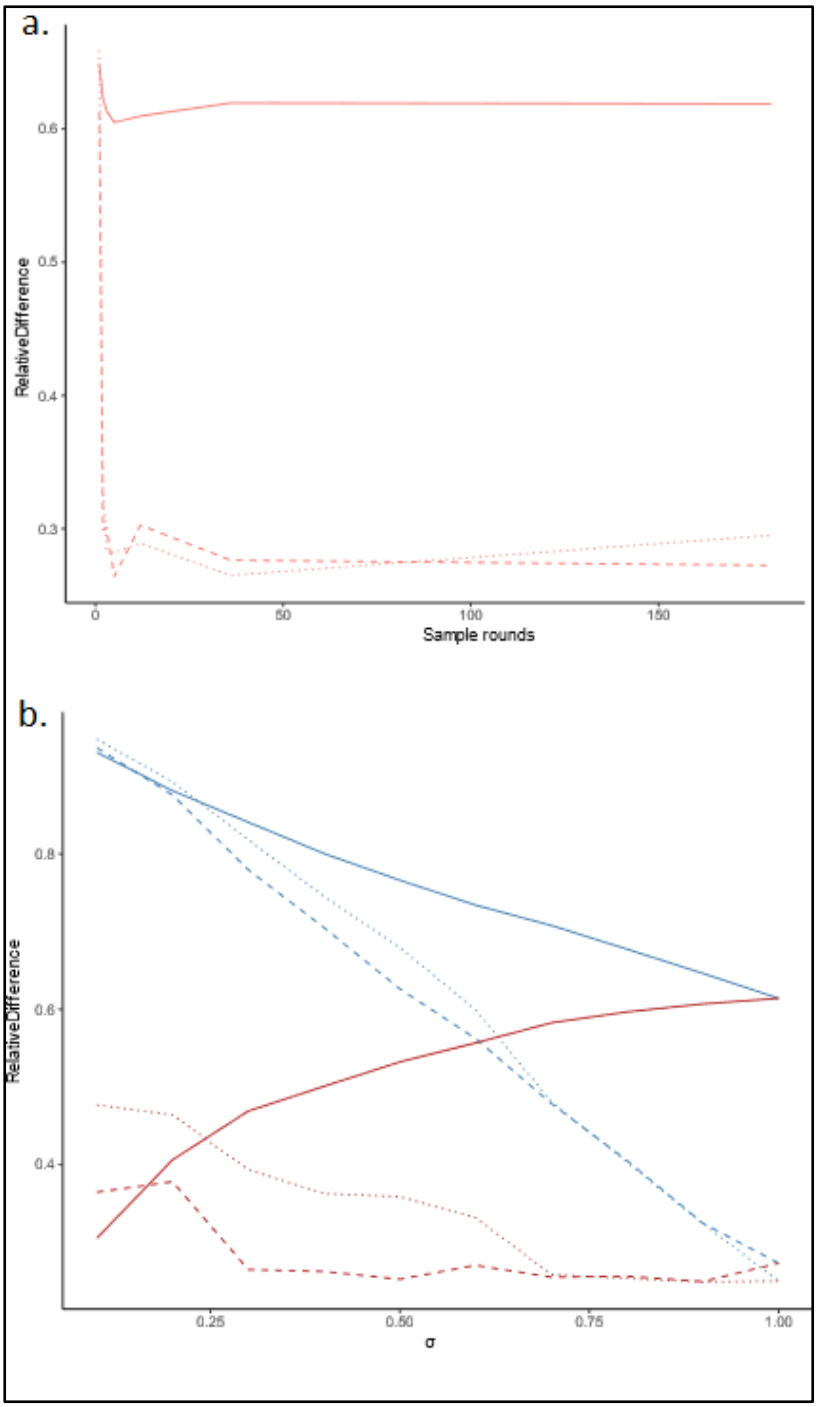


Figure 5.4. Relative differences in the mean detection prevalences. a. the relationship between sampling allocation with fixed surveillance intensity and the relative difference between Parnell *et al.*'s rule of thumb and the estimated simulated prevalence of pest at detection. Solid line is θ of 10, dashed line is θ of 60 and dotted line is θ of 110. b. the relationship between the detection sensitivity (σ) and the relative difference between the

estimated simulated detection prevalence and the two models (Mastin *et al*.'s and Parnell *et al*.'s) for three different θ values. Solid line is θ of 10, dashed line is θ of 60 and dotted line is θ of 110. Blue line is Parnell *et al.*'s rule of thumb and red line is Mastin *et al.*'s rule of thumb. Sensitivity of detection assay within the simulation σ is on the x-axis.

Figure 5.2 illustrated the relative differences between the rule of thumb approximation and the simulated prevalence of pest at detection, whereas Figure 5.3 illustrated the absolute differences between the rule of thumb approximation and the simulated prevalence of pest at detection. Decreasing the pest dispersal parameter (θ) led to a reduction in the relative accuracy of the rule of thumb. However, when considering the absolute differences I observed that this increase in difference is due to smaller final epidemic sizes, which impacts the final output of the relative differences. These findings highlighted that the choice of measurement of accuracy is complex, since it relied on an interplay between the epidemic and surveillance parameters, and must be suited for the research question as necessary. In addition, these findings indicated that the relationship between the epidemiological parameters and the accuracy of the Parnell *et al.* model were complex. Figure 5.2 and 5.3 also highlighted that increasing the interval between sampling rounds (Δ) led to a reduction in the accuracy of the rule of thumb. The general trend indicated that the accuracy of the rule of thumb also decreased as the sample size within the sampling round decreased (N). This was also more impactful as the epidemiological parameters increased. Of course, given the that the final epidemic sizes were larger in these faster epidemics, the relative differences did not adequately capture this effect. It is important to note that for the smallest sample size in the epidemic set with the largest values of epidemiological parameters, the absolute differences were small. As to why this may be the case requires further investigation. Overall these results suggested that increasing Δ and decreasing N would lead to a reduction in the accuracy

of the rule of thumb, and that this effect was intensified by increasing epidemiological parameter values.

Figures 5.4 and 5.5 were the absolute and relative differences between the rule of thumb approximation and the simulated detection prevalence of pest for a range of detection sensitivities and sample rounds with surveillance intensity fixed. Figure 5.4a and 5.5a showed the absolute and relative differences between the rule of thumb approximation and the simulated detection prevalence of pest for a range of sampling rounds whilst keeping the sampling intensity fixed for three values of θ. There was no significant difference between the sampling round choices for the range of θ. The relative difference between the rule of thumb approximation and the simulated detection prevalence of pest was larger for the shortest dispersal distance ($\theta = 10$) than the longest dispersal distance ($\theta = 70$), and this difference remained constant across the sampling rounds. Figure 5.4b and Figure 5.5b showed the relative differences between two different early detection models predictions and the simulated detection prevalence of pest. Here it was observed that the Mastin *et al*. model had the greater accuracy, except for small detection assay values on a short dispersed landscape. The model generally performed better than the Parnell *et al.* model, with the largest increase in accuracy being observed for the smallest sensitivity values in epidemics that were not modelled for short dispersal ($\theta = 10$). For the smallest value of θ, the accuracy of Mastin *et al*. model converges upon a relative difference of approximately 0.6, though when sensitivity (σ) is much smaller than 1, the difference between the early detection models is greater.

5.5. Discussion

In this chapter, I sought to address the challenge of model verification and reliability in the context of early plant pests detection modeling whilst also providing insight into optimising surveillance strategy with limited resource allocation. Assay sensitivity plays a crucial role in determining the likelihood of the early detection of pests, and different assay techniques may be more suitable depending on the scale and objectives of assessment (Martin *et al*., 2000). In this chapter I aimed to explore the challenge of allocating the distribution of resources from a limited sampling effort pool. This question is outstanding in the literature and is of significant concern for designing surveillance strategies within the context of plant pest management. Therefore, my research questions aimed to address the impact of detection method sensitivity on the accuracy of simple detection models, the interactions between epidemiological phenomena, early detection model considerations such as limited resource allocation, and the potential for selecting the most appropriate early detection model. To tackle these questions, I deployed a spatially-explicit stochastic epidemic simulation model which captured dynamic epidemic behaviours whilst accommodating for a range of plant pest scenarios. By integrating pest dispersal and the transmission coefficient of pests, I aimed to unravel the complex interactions between pathogens and surveillance strategy in a spatially-explicit stochastic environment through simulating realistic epidemic behaviour and measuring the performance of the rule of thumb models.

Here, I showed that decreasing the pest dispersal parameter (θ) heightened the impact of reducing surveillance intensity on the reduction in accuracy of the Parnell *et al.* early detection model. This result indicated that for shorter dispersed pests, such as soil-borne nematodes and pests carried by short flight-distance vectors, an increased surveillance intensity protocol is recommended. By increasing the surveillance intensity, the reduction in the accuracy of the Parnell *et al.* model is minimised, thus facilitating more accurate

estimations of pest at first detection. However, it is crucial to consider the absolute differences in this context. The increase in relative differences can be attributed to smaller final epidemic sizes, which have an impact on the final output of relative differences. These findings highlighted the complexity of accurately measuring and interpreting accuracy in the context of surveillance models. The absolute accuracy represented the true prevalence of pests at detection, and this may be more informative when determining tolerable levels of prevalence when first detected (Magarey & Sutton, 2007). It is also important to note that quantifying detection sensitivity in practise is difficult; EFSA define detection sensitivity as a combination of sampling effectiveness and diagnostic sensitivity, however the commonly used visual assay detection method is very difficult to quantify. In practice, although lab sensitivity is often quantified, it is very rare for sampling effectiveness of the sensitivity of visual inspection to be quantified. There are however a small number of exceptions to this generalisation e.g. Futch *et al.,* 2009 (Futch *et al.,* 2009).

However, the results were complicated by the interaction with the epidemic parameters and less clear in some instance when the differences were expressed through the relative differences. The relative differences between the rule of thumb and epidemic simulation model contradicted the findings of the absolute differences between the rule of thumb and epidemic simulation model. This is because whilst there were larger absolute differences between the epidemic simulation model and the rule of thumb, the total size of the epidemic reflected a relatively smaller drop in accuracy for the less intense surveillance strategies with larger epidemic parameter values. This means that if epidemic size when first detected is to be considered for future scientists and stakeholders, then the results indicate that the rule of thumb is more accurate for wider spread epidemics than for thinner epidemics. The coupled use of relative differences and absolute differences in this chapter is perhaps a recommended strategy to determining the accuracy of pest detection prevalence prediction and the limits on

accuracy for tolerable pest threshold respectively. The use of absolute values and relative transformations have been utilised previously to highlight the importance of variables of particular interest within plant science (Jung *et al*., 2000; Henson & French, 1993; Patil & Bodhi, 2011). For example, Jung *et al*., 2000 demonstrated that when the relative health of oak roots in association with *Phytophthora* spp. was measured relative to weather and seasonal conditions, the correlative index indicating the soil type and pH values associated with pest increased in significance than absolute values alone (Jung *et al*., 2000). Similar methods have been used in microbiological techniques to quantify the differences in plant DNA and RNA scarcity and in quantifying the degree of pest severity in plant samples using imaging techniques (Henson & French, 1993; Patil & Bodhi, 2011).

It is important to note that the smallest N in the epidemic set with the largest values of epidemiological parameters exhibited small absolute differences, justifying further investigation. Overall, these results suggest that the choice of surveillance parameters can impact the accuracy of the early detection models, and this effect is intensified by higher epidemiological parameter values when comparing the absolute difference between prevalence at detection and Parnell *et al.*'s model. Whilst the accuracy of the model improves with increasing epidemiological parameter values, it is important to note that prevalence at detection may be beyond acceptable levels and an increase in accuracy is coupled with a sacrifice in early pest management. My findings indicated that less virulent ($\beta = 50$) epidemics with longer pest dispersal ($\theta = 70$) led to greater predictions in the accuracy of pest prevalence when detected for the first time. This difference between virulence estimations was more significant when surveillance resources were scarcer. This indicated that resource allocation should be prioritised for longer-dispersal pests which proliferate faster because it is hypothesised that the increase in the number of hosts infected via increased pest virulence leads to more stochastic events occurring, which would lead to less accurate approximations

of prevalence when first detected. This has implications for management strategies, as often control methods are labour intensive and expensive, or the damage caused by a pest may be significant. Therefore, surveillance allocation should be proportional to the total cost of pest incursion should a pest invade (Kompas *et al*., 2019). If the modelling attempts to quantify the total amount of pest prevalence become less accurate when a pest is deemed a worse threat for biosecurity, then additional surveillance resources should be assigned, going beyond the anticipated required resources based on the damage a pest may cause. Ultimately, it would be prudent to take a more cautionary approach when allocating resources towards the most virulent and highly dispersed plant pest that could potentially invade.

My chapter also provided valuable insight into the effects of changing sampling rounds whilst fixing surveillance intensity. This was equivalent to having a set amount of surveillance resources and deciding how frequently to allocate those surveillance resources during the early monitoring process of a potential pest incursion. My results indicated that apart from applying all resources in one round of monitoring, varying the allocation of resources across time does not significantly impact the accuracy of Parnell *et al.*'s model (Figure 5.4a and 5.5a). This is a promising result with implications for pest management; my results indicated that if the appropriate amount of surveillance is allocated within sampling rounds, the stake holder can afford to use fewer sampling rounds without a drop in accuracy in the prediction of prevalence of pest when first detected. This would be useful in scenarios where sampling may need to be "time-dependent", i.e. timing the surveillance in line with the emergence of plant pest during the crop season. Other research has indicated that these "time-dependent" sampling events are of relevance for pests beyond those affecting plant species (Magnani *et al*., 2005). This is of course contingent on the appropriate surveillance intensity being applied, less frequent sampling efforts will not be effective if the sample size does not match the required intensity for sufficiently accurate prediction. Furthermore, epidemics may

behave differently if the early expansion phase is too fast to afford a relaxed surveillance frequency.

My results also provided key insights into the impact of detection sensitivity and sampling round frequency on the accuracy of plant pest detection models. It is evident that the Mastin *et al*. variant consistently outperforms the Parnell *et al*. variant, particularly as the sensitivity decreases (Figures 5.4b & 5.5b). Notably, for smallest pest dispersal ($\theta = 10$), the Mastin *et al*. model converges upon a relative difference of approximately 0.6. Consistent with the findings in Figure 5.4, the Mastin *et al*. variant exhibits superior performance, especially for smaller θ values. It is hypothesised that the low detection sensitivities for shortly dispersed simulations led to higher final epidemic sizes upon detection and this was the causal reason as to why Mastin *et al.*'s rule of thumb performed less well for shortly dispersed epidemics with low detection assay sensitivity. Notably, as the detection sensitivity decreased, the difference between the accuracy of the two models became more pronounced. These results emphasized the critical role of detection sensitivity in influencing the accuracy of plant pest detection models, where the choice of model played an important role in the achievement of precise predictions. Moreover, the findings underscored the need for careful consideration and calibration of detection sensitivity parameters when designing surveillance modelling approaches for effective plant pest management. The detection sensitivity of an assay is largely dependent on the unique characteristics of the assay, and the stage in which the pest has progressed within a population. With asymptomatic pests, early detection is often difficult and requires advanced technologies such as hyperspectral reflectance and machine learning algorithms for processing (Rumpf *et al*., 2010). With advances in technology, the sensitivity of assays are becoming increasingly accurate however classical methods such as visual inspection may still be deployed (Sankaran *et al*., 2010). Early detection of pest is paramount to effectively managing plant pest epidemics (Parnell *et al*., 2017). Having a

model that can include a variable sensitivity parameter is important to exploring the ongoing challenges of identifying and quantifying pest incursion during the early phase of establishment, when effective pest management is more feasible.

In conclusion, this chapter has provided valuable insights into the complex dynamics of plant pest detection modeling and the impact of changing surveillance dynamics and detection sensitivities on the accuracy of early detection models. The sensitivity of detection methods and the choice of early detection models were identified as key elements in influencing the accuracy of early pest prevalence prediction. My results emphasized the importance of careful calibration and evaluation of detection sensitivity parameters when designing an early detection surveillance modelling approach. Additionally, my chapter demonstrated the intricate interactions between epidemiological parameters, resource allocation, and the accuracy of detection models. It is noted that the more highly-dispersed virulent epidemics should require more surveillance resources to be allocated to ensure effective estimations when modelling prevalence at first detection. In addition, shortly-dispersed epidemics also produce less accurate estimations, regardless of virulence, and should be monitored carefully to facilitate accurate prediction. Though these pests produce less risk in becoming unmanageable. Consequentially, my research findings have practical implications for stakeholders involved in developing biosecurity strategies and predictive modeling approaches for plant pest management. By providing insights into the evaluation of early detection modelling techniques, this chapter contributes to enhancing the effectiveness of plant pest management practices. Further research is needed to explore and refine these modeling approaches, incorporating advancements in detection technologies and considering the specific characteristics of different plant pests.

6. Chapter six: A study of *Oak Processionary Moth* as an example of the application of an early detection surveillance model

6.1. Chapter Summary

Climate change and globalisation driven range expansion of *Oak Processionary Moth* (*Thaumetopoea processionea*; OPM) means it is an increasing problem in Northern Europe, causing significant defoliation of oak trees, potential allergic reactions in humans and animals, and increasing the risk of exposure to further pests within ecosystems. This chapter explores the effectiveness of surveillance strategies for early detection of invading OPM populations. It also explores to what extent simple spread models can accurately capture the performance of surveillance strategies for the pest. I use a spatially-explicit stochastic epidemiological model to simulate OPM dispersal and prevalence at the point of detection with a spatially-explicit stochastic surveillance model. I compare the output of this experiment with the Mastin *et al.* (2019) rule of thumb. My primary findings highlight that there is a correlation between lower levels of surveillance and a higher prevalence of OPM when it is first detected. This indicates that current surveillance may be missing early infestations, allowing the moths to establish more extensively before being found by surveillance. These results could inform plant health managers within NPPOs as to a more realistic surveillance protocol to manage emerging OPM populations, reducing the rate of defoliation and presentation of allergic reactions. OPM will continue to proliferate under favourable conditions, though effective early detection modelling will help to manage this proliferation. However, tools like the rule of thumb have limitations based on the assumptions they use and are not applicable when surveillance frequency is too high. This implies that there is a disconnect between epidemic speed and frequency of detection; if the

delay between sampling rounds is sufficiently wide enough, epidemics will be spreading too fast for representation by the rule of thumb.

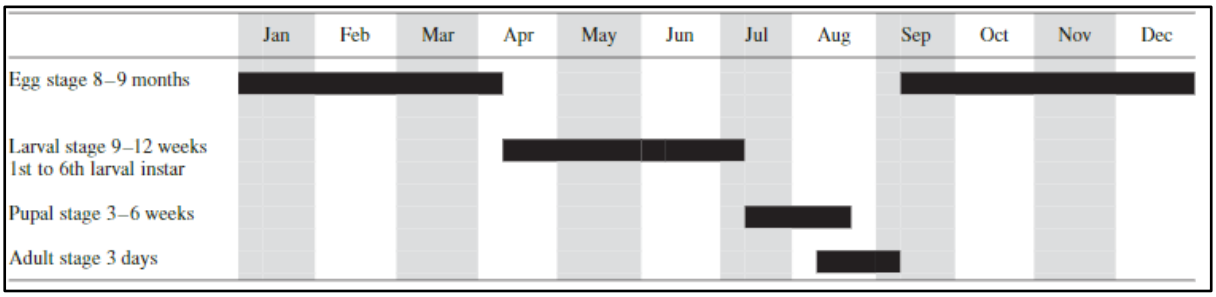
6.2. Introduction

The sharp increase in the major introduction of exotic plant pests to the United Kingdom over the past 100 years is indicative of the increase in the establishment of invasive species globally (Spence, 2020). The establishment of invasive species within the United Kingdom is a consequence of the development of global trade networks and the large-scale mixing of biota (Spence, 2020). The importance of managing plant diseases effectively is stated in Chapter 3. Therefore, it is crucial that the repertoire of epidemiological models available are tested and validated under different epidemic scenarios using simulation modelling before application in the real world to maximise their utility (Cunniffe *et al.*, 2016). If early detection models have not been validated prior to pest incursion, then predictions derived from the model may under-estimate or over-estimate the total presence of the pest, leading to the misallocation of resources which will further increase cost (Moffit *et al.*, 2006). Failure to validate models that have potential application during early pest incursion may result in epidemics that are costly, leading to a high burden on National Plant Protection Organisation (NPPO) resources (Cunniffe *et al.*, 2016; Carpenter *et al.*, 2011).

An example of such a pest that threatens species of Oak (*Quercus* spp.) within the UK is Oak Processionary Moth (*Thaumetopoea processionea*) (Godefroid *et al.*, 2020). Oak Processionary Moth (OPM) is an emerging plant pest threat to the UK that was introduced in 2005 via imported trade material (Forest Research, 2021). OPM is native to the Mediterranean basin but has expanded to many countries within Europe due to changes in climatic conditions and accidental transportation of infested Oak material (De boer, 2020; Groenen & Maurisse, 2012). OPM is a priority pest because as well as causing defoliation of

oak trees, leading to vulnerability to other stressors, the caterpillars of OPM produce urticating hairs which are major irritants to animals and people (Groenen & Maurisse, 2012; Suprunenko *et al.*, 2021; Rahlenbeck & Utikal, 2015). OPM is a univoltine species that feeds off oak trees, the epidemics are predominantly proliferated by the female moths that lay eggs in the upper canopy of oak trees (Suprunenko *et al.*, 2021). The eggs hatch through spring and go through six instars of development (Suprunenko *et al.*, 2021). The larvae pupate in nests and emerge as adults temporarily during mid-July to mid-September during which period the adults seek out new oak trees to establish upon (Suprunenko *et al.*, 2021). Gottschling & Meyer provide an excellent summary table of this development (Table 1) (Gottschling & Meyer, 2006). Dispersal only occurs through the movement of larvae, the movement of adult females and accidental human-mediated transportation (Suprunenko *et al.*, 2021). Since 2005, the expansion of OPM has increased rapidly within the UK (Suprunenko *et al.*, 2021). Research indicates that the distribution of OPM, initially introduced into the UK via contaminated trade goods from the Netherlands, has increased from the Greater London region to beyond the borders of Greater London (Suprunenko *et al.*, 2021). OPM initially expanded at a rate of 1.66km per year, but this increased to 6.17km per year in 2015. Fortunately, the UK government is responsible for reducing the overall size of this expansion via restricting trade of oak material surrounding the Greater London region, effectively creating a buffer zone in which material cannot be realistically transported across (Forestry Commission, 2022).

Table 6.1. Chart of the yearly Oak Processionary Moth life cycle. The life cycle of the univoltine OPM, with the length of bars representing the duration in months for each stage of the OPM life cycle. The bars also correspond to the specific months within the year in which each stage of the life cycle occurs. Taken from *An epidemic airborne disease caused by the Oak Processionary Caterpillar* (Gottschling & Meyer, 2006).



Similarly to the expansion of *Pine Processionary Moth* in central Europe, managing the expansion of OPM is contingent on sufficient data to determine whether the costs of short term management such as quarantine and horizon scanning is more or less expensive than managing the pest long term with surveillance and eradication (Moffit *et al.*, 2006; Welsh *et al.*, 2021; Cayuela *et al.*, 2011). Surveillance efforts are a determining factor in the prevalence of pest when first detected, which is important because early detection can significantly influence containment and mitigation strategies, allowing for more efficient allocation of resources and reducing the potential damage these pests can inflict (Cunniffe *et al.*, 2016; Parnell *et al.*, 2015). Therefore, utilizing modelling strategies, I explore a current *OPM* surveillance strategy within the UK and compare it against other possible surveillance strategies. Officially, it is the responsibility of official inspectors to inspect Oak trees for possible infestations using post planting inspection forms. However, to my knowledge, beyond the current recommended modelling strategy, there is no UK wide policy on the number of trees to be inspected at what interval for native forests as opposed to trade material (Forestry Commission & Defra, 2023).

Secondly, I utilize the early detection model established by Parnell *et al.*, 2012 and adapted by Mastin *et al.*, (2020) which was investigated in previous chapters (Parnell *et al.*, 2012; Mastin *et al.*, 2020). I seek to understand how the prediction would match surveillance efforts, and how the accuracy of the model would change with changing surveillance strategy. Therefore, the research questions in this chapter are:

- How well does the current OPM surveillance modelling strategy in the UK perform?
- How are early detection models sensitive to changing surveillance conditions?

6.3. Methods

6.3.1. Epidemiological model

The model tracked simulated host plants in a two compartmental model (Keeling & Rohani, 2007): susceptible (*S*) hosts were uninfected and infected (*I*) hosts were infectious. If host *i* was susceptible at time *t*, then the probability of infection of *i* was given by:

$$\varphi_i(t) = \beta \sum_j K(d_{ij}; \theta)$$

(Eqn 6.1)

$\phi_i(t)$ was the rate of infection at time *t*, β was the transmission coefficient that defined the rate for any given distance within the kernel, and the summation ran over all infectious hosts *j* at time *t*. The dispersal $K(d_{ij}; \theta)$ set the rate of disease transmission between a pair of hosts separated by distance d_{ij}, and was parameterised by the dispersal scale parameter θ. Therefore, β is a rate and θ has the units of distance. To allow robustness in dispersal mechanics if needed based on the dispersal kernel tail, I included the power exponential function dispersal kernel explored by previous authors (Bourhis *et al.*, 2019; Rieux *et al.*, 2014). I fixed the scale parameter θ as the thin-tailed exponential kernel in this study, and

then normalised so that the probability of infection translated to a total probability of 1 for the entire kernel at any given time:

$$\theta_{norm} = \frac{1}{2\pi\theta^2 \Gamma 2}$$

(Eqn 6.2)

In order to reduce the time taken to run epidemic simulations, I used the tau-leap Gillespie algorithm (Gillespie, 2001). This algorithm was defined as:

$$\delta I_T \approx Poisson(S_t \varphi_i(t) \tau)$$

(Eqn 6.3)

Where τ was the time leap between the initial state and updated state of the simulation process, ($T - t$). The time steps between OPM movement were restricted to 40 day intervals.

The mean dispersal distance for the exponential kernel was equal to twice the value of the θ parameter, as given by the equation stated by Bourhis *et al.* (2019) (Bourhis *et al.*, 2019):

$$\phi = \frac{\theta \Gamma(\frac{3}{b})}{\Gamma(\frac{2}{b})}$$

(Eqn 6.4)

Where θ was the dispersal parameter, b was the shape parameter, Γ was the gamma function and φ was the mean dispersal distance.

6.3.2. Estimation of the dispersal distance parameter

To calculate a realistic dispersal of a single OPM, I utilised the 95th percentile to convert expert knowledge collected from an OPM management team at Defra on OPM dispersal into

a defined dispersal distance parameter (θ) (Hoppit, 2023). The equation that defined the 95% of an exponential decay distribution was given by:

$$95^{th}\ percentile = \frac{-ln\ (0.05)}{\lambda}$$

(Eqn. 6.5)

Where λ was given by the inverse of the mean dispersal distance. 500 meters was given as the natural dispersal of a single OPM moth between oak trees, assuming no human-mediated or weather-mediated transport as provided by lead UK OPM government expert Andrew Hoppit. By rearranging this equation and using 500 meters as the maximum dispersal of OPM, I estimated mean dispersal distance of OPM as 167 meters (Hoppit, 2023). Using equation 4 above, I estimated θ as approximately 83.5.

6.3.3. Estimation of the transmission coefficient parameter

By using the spread model defined above in section 3.1., I recorded the flight distances of the simulated OPM from the point of infestation across 2000 simulations. For each simulation, the transmission coefficient (β) was considered for the above θ of 83.5. I chose the transmission coefficient that produced the 95% percentile of maximum distance that corresponded to the reported expansion rate upon establishment in London, UK in 2006 (Figure 6.1) (Suprunenko *et al.*, 2021).

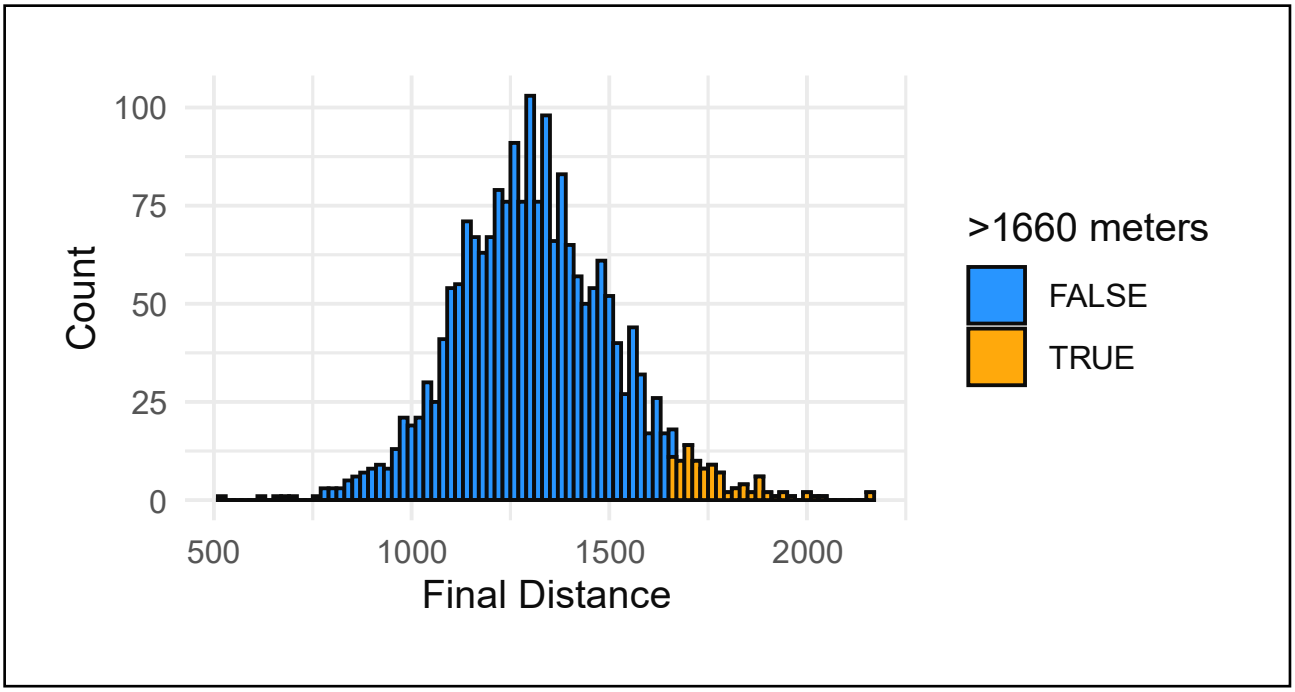


Figure 6.1. The simulated final distance from the source of infestation using the estimated OPM dispersal parameter. Based on expertise and a transmission coefficient that generates 5% total simulations above the reported annual expansion rate in London UK upon establishment in 2006 (Hoppit, 2023; Suprunenko *et al.*, 2021).

6.3.4. Landscape generation

I initially modelled the landscapes as 5km^2 areas containing 5000 hosts, with coordinates drawn from a uniform distribution, which reflected a realistic oak tree density according to literature; to visit the databases used to calculate this see Hill *et al.* (2017) (Hill *et al.*, 2017). These landscapes were generated in the R environment using the spatstat package (Baddeley & Turner, 2005). I initiated epidemics with 1 randomly selected host infection at time $t = 0$, and the initial host was re-selected for each epidemic run within the simulation, which continued until the end of one year. This method was chosen for pragmatic reasons; OPM epidemics spread rapidly. Time constraints were also a consideration. These experiments informed the selection of epidemiological parameters. I reduced the landscapes to 1km^2 whilst maintaining the same host density to reduce computational time.

6.3.5. Surveillance model

I developed a spatially-explicit stochastic surveillance model. This model was designed to consider the randomness inherent in a random surveillance protocol. Utilizing this model, I sampled hosts at consistent time intervals. This method was adopted to capture the temporal evolution of the epidemic. I relied on previously generated epidemic data as the basis for the sampling events. I stopped the surveillance model at the point at which the epidemic was detected for the first time, and total prevalence in the simulation run was recorded. I then calculated the average prevalence for the complete simulation set (2000 simulations). The initial sample size per sampling round was 24 hosts (N), with the interval between surveillance rounds (Δ) set to 182 days at a scale of $1km^2$. This reflects current surveillance allocation in the UK for the monitoring of new OPM populations in susceptible regions (Andrew Hoppit, 2023). I then fixed surveillance interval and sampling size respectively whilst changing the other surveillance parameter. The surveillance combinations were based on the current strategy reported by a specialist within the UK Government (Hoppit, 2023). Surveillance parameters of frequency interval and sample size were increased and decreased and prevalence at detection were recorded. The list of explored surveillance strategies is written below (Table 6.2).

Table 6.2. List of the surveillance protocols tested for chapter six. Surveillance strategy was initiated at a random value in between 0 and Δ.

Sampling Size (*N*)	**Sampling Interval (Δ)**
24	14
24	28
24	42
24	56
24	70
24	84
24	98
24	112
24	126
24	140
24	154
24	168
24	182
48	182
72	182
96	182
120	182
144	182
168	182
192	182
216	182
240	182

6.3.6. The rule of thumb model

Parnell *et al.*'s 2012 (Parnel *et al.*, 2012) rule of thumb was defined as:

$$q^* = \frac{r\Delta}{N}$$

(Eqn 6.6)

Where $q^*$was the theoretical detection prevalence which was determined by the relationship between the growth rate of the epidemic r and the surveillance protocol, Δ was the interval between surveillance rounds, and N was the sample size taken for inspection each surveillance round. The derivation of this equation was discovered as follows: by the assumption that if an epidemic grows exponentially, then the probability of detecting the disease at any given time was a simple interaction between rate of surveillance and the exponential growth of the epidemic (Parnell *et al.*, 2012). By solving this equation via integration, Parnell *et al.* (2012) arrived at the total probability of detection at the point of detection. Using a random variable transformation of the initial time with an exponential growth assumption, the prevalence at the time of detection was translated from this probability using the Jacobian of the transformation. A full derivation of the approach can be found in Parnell et al (2012; 2015). Thereby, the above rule of thumb states that given the epidemic growth parameter (r), the surveillance frequency (Δ) and the sample size (N), then the prevalence of disease at detection (q^*) is estimated as an approximation of the true prevalence of disease at detection. Key assumptions from the rule of thumb are exponential increase of the pathogen population and continuous monitoring of the host population with time.

Based on the previous research findings, I used a modified rule of thumb to account for sensitivity as derived by Mastin *et al.* (2019). This modified rule of thumb is defined as:

$$q^* = \frac{r\Delta}{N\sigma}$$

(Eqn 6.7)

Where σ is defined as the sensitivity of the detection method used.

I performed for every unique epidemiological parameter combination of θ and β, 2000 simulation runs, providing a unique set of a growth rate values (r). I estimated r by using a linear transformation on the initial growth phase of the epidemic, corresponding to the first quartile of transition events and measuring the mean gradient of the curve. The average of these 2000 r values was then used as input within the modified rule of thumb (Eqn 7). For N and Δ in the above equation, these values were taken from the surveillance model as defined above. Based on expertise this value for sensitivity of detection assay is defined as 0.8 for OPM surveillance (Hoppit, 2023). This is based on the visual inspection of OPM nests where 20% of inspections result in false negatives.

6.4. Results

In this section, I present the findings from my surveillance and modelling efforts for OPM. The results are categorised into two primary aims:

- The simulated detection prevalence of OPM according to varying surveillance strategies
- The performance of a simplistic epidemiological early detection modelling approach

6.4.1. The simulated detection prevalence of OPM according to varying surveillance strategies

Figure 6.2 represented the 2000 simulated epidemic growth curves for the OPM parameters (Figure 6.2). The growth curve was included derived from the average growth rate of the

2000 simulations. By day 100, 50% prevalence was nearly achieved. This aligned with my models projections based on the expansion behaviour of OPM (Suprunenko, 2021). When I modelled this growth, I took into account both the transmission coefficient and dispersal distance of OPM. Figure 6.3 illustrated four combinations of surveillance strategy with the recorded prevalence of pest at detection within the simulation model for the OPM parameters (Figure 6.3).

Below are described the outcomes of these combinations (Figure 6.3.):

a. The most intense surveillance sampling protocol.
b. The sample protocol with small but frequent sampling, most reported prevalence at detection were less than 0.1.
c. The sampling protocol with large but infrequent sampling, there is a divide between early detection and missing the epidemic completely.
d. The least intense sampling, the current protocol.

6.4.2. The performance of a simplistic epidemiological early detection modelling approach

Figures 6.4-6.9 were the simulated prevalences at detection versus the recorded rule of thumb prediction with a modifier to account for the sensitivity of sampling (β, rate of infection; θ, metres) (Mastin *et al.*, 2020). Figure 4.5 showed varying sample frequency with fixed sample size and varying sample size with fixed sample frequency. Figure 6.7 showed varying sample frequency with fixed sample size and varying sample size with fixed sample frequency respectively on a clustered type of landscape. Figure 6.8-6.9 showed varying sample frequency with fixed sample size and varying sample size with fixed sample frequency respectively on a partially clustered/random type of landscape.

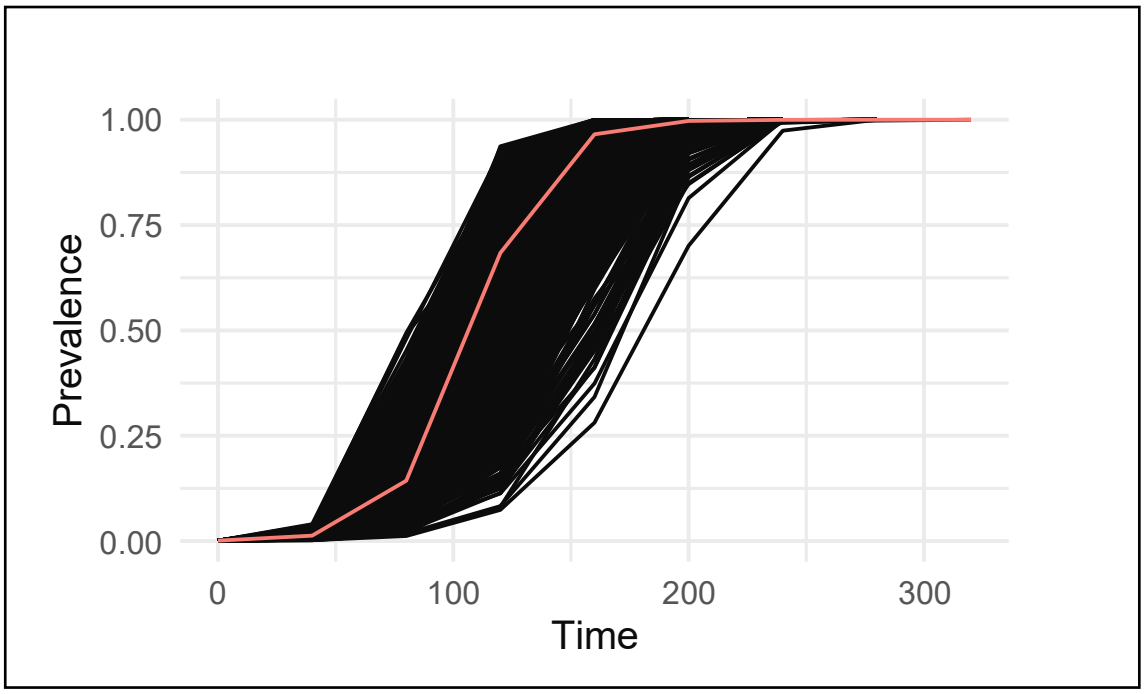


Figure 6.2. The epidemic growth curves for the OPM parameters. 2000 simulations are reported as the black lines (β = 315, θ = 83.45). The red line is the average growth curve (r = 0.063).

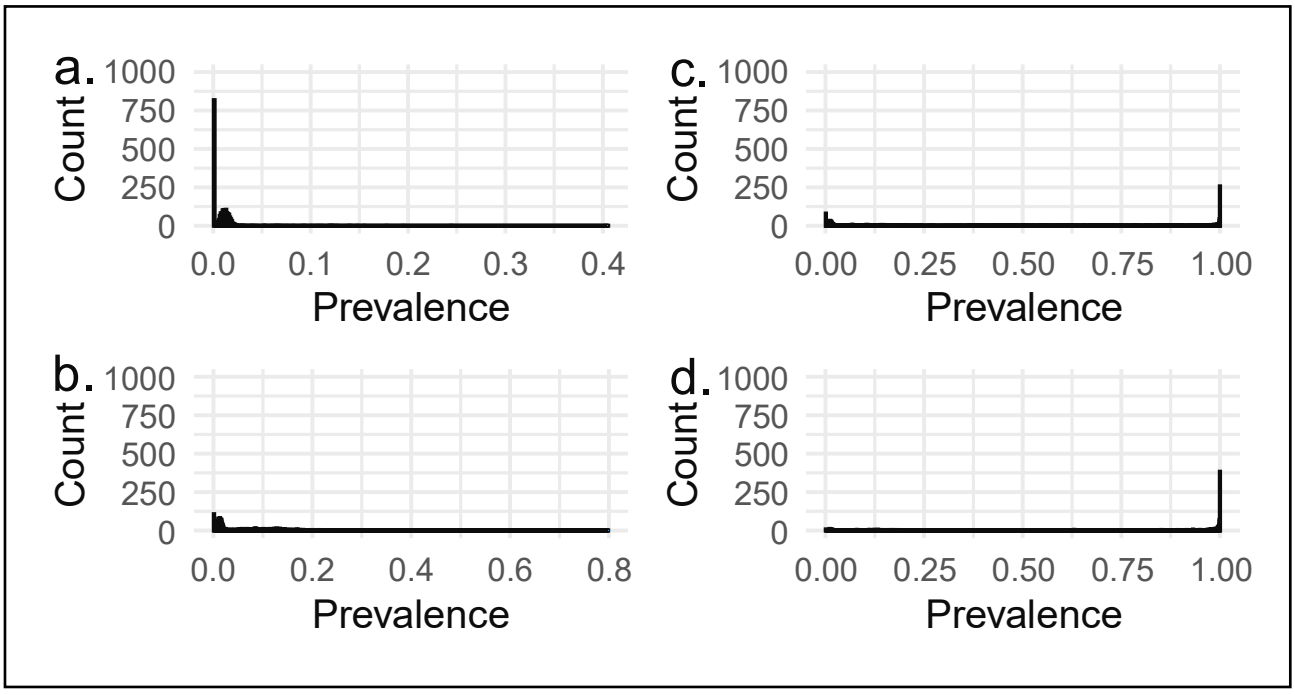


Figure 6.3. The distribution of detection prevalences of OPM for different surveillance strategies. a. N=240, Δ = 14 b. N=240, Δ = 182 c. N= 24, Δ = 14 d. N = 24, Δ = 182.

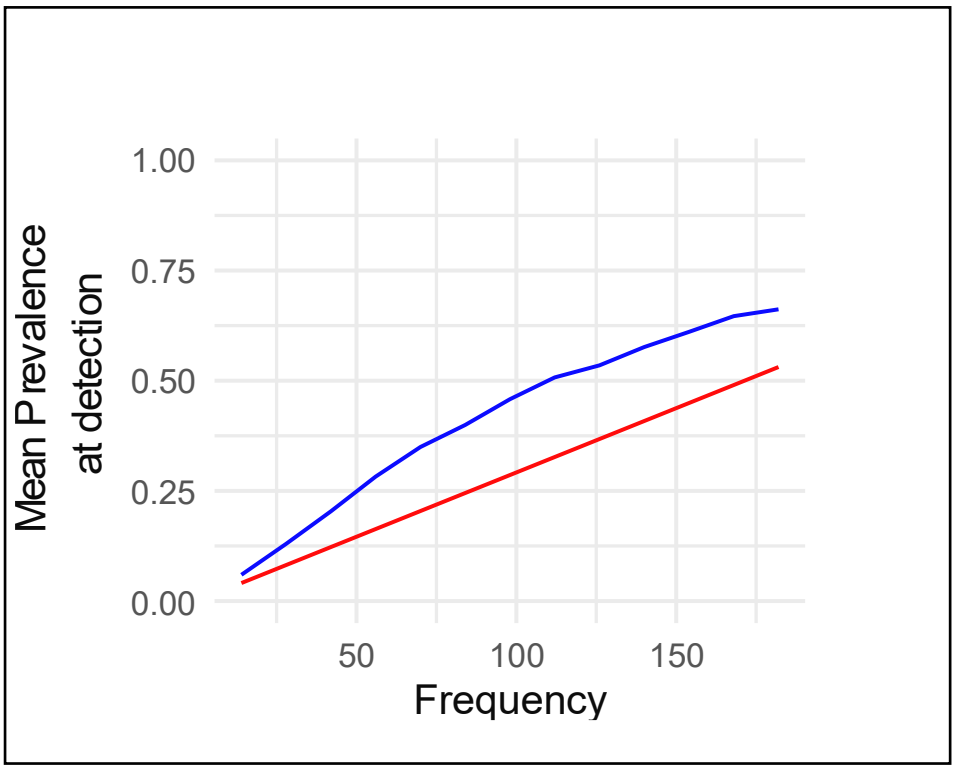


Figure 6.4. The relationship between the frequency of sampling and the prevalence of pest at detection for both the simulated model and the modified rule of thumb. Sample size is fixed at 24 per round. Landscape parameter set to random ($\xi = 1$). Blue indicates the simulated prevalence of pest at detection and red indicates the early detection model prediction.

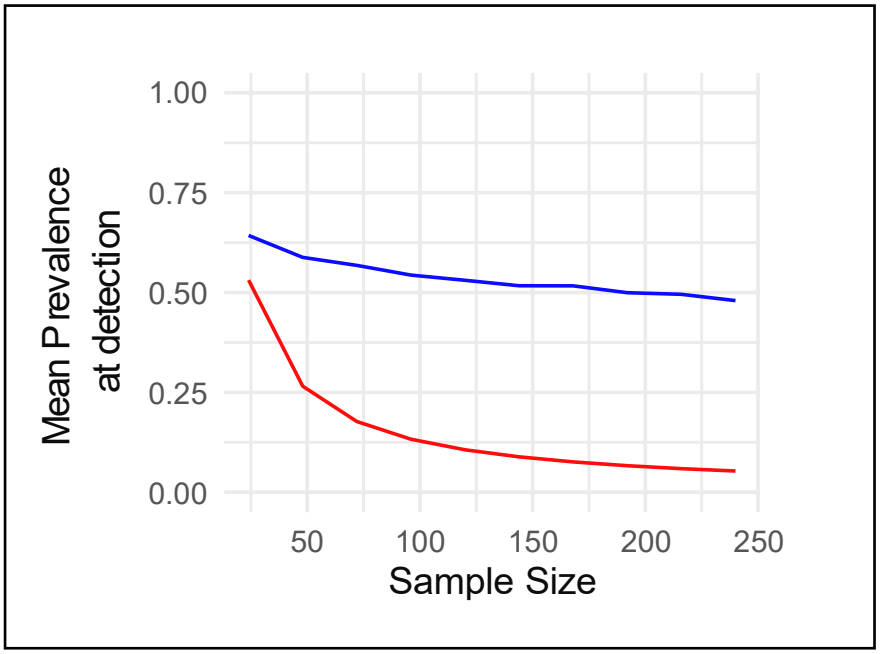


Figure 6.5. The relationship between the sample size per round and the prevalence of pest at detection for both the simulated model and the modified rule of thumb. Sample frequency is fixed at 182 days. Landscape parameter set to random ($\xi = 1$). Blue indicates the simulated prevalence of pest at detection and red indicates the early detection model prediction.

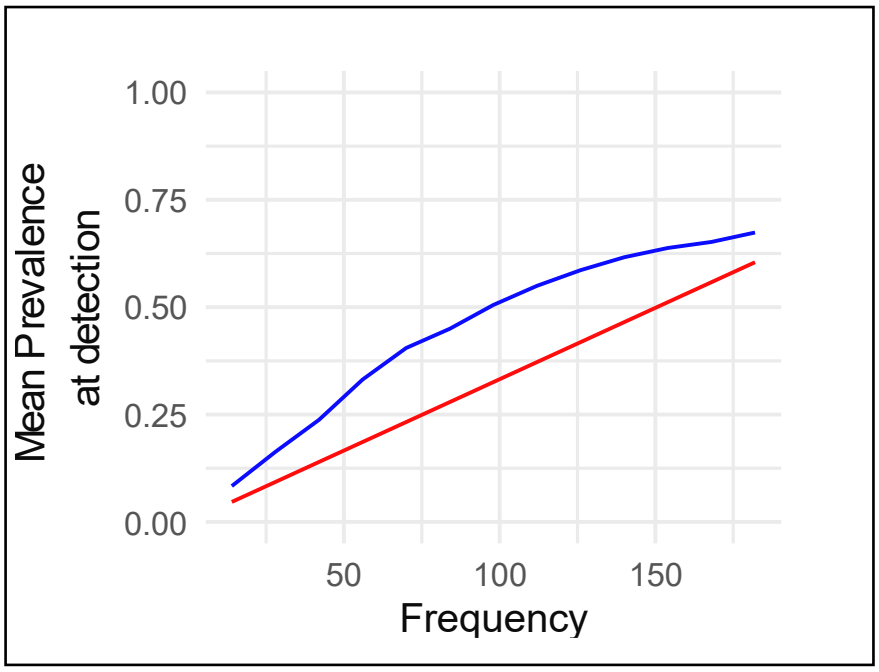


Figure 6.6. Showing the relationship between the frequency of sampling and the prevalence of pest at detection for both the simulated model and the modified rule of thumb. Sample size is fixed at 24 per round. Landscape parameter is set to clustered ($\xi = 0$). Blue indicates the simulated prevalence of pest at detection and red indicates the early detection model prediction.

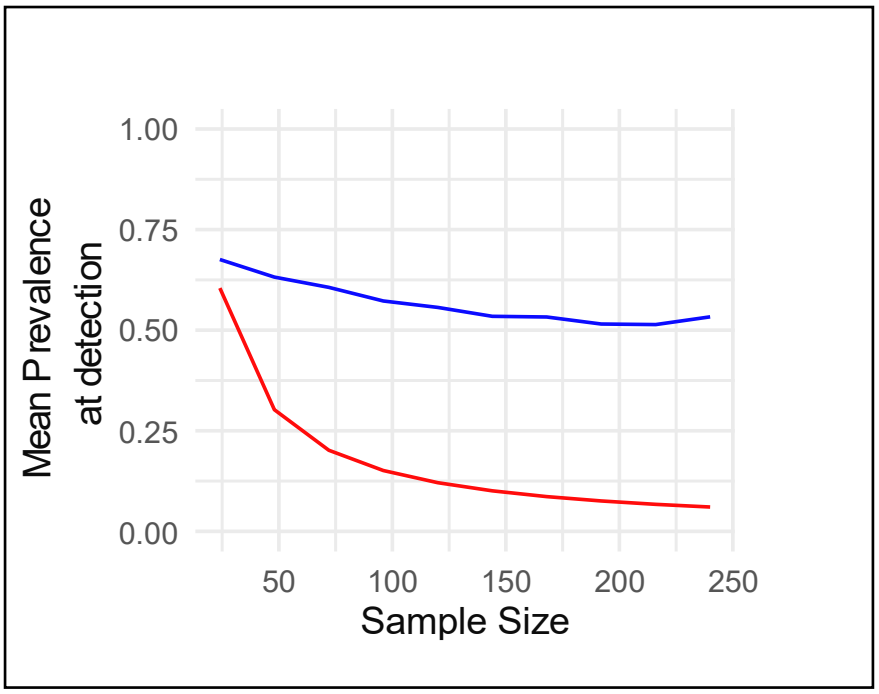


Figure 6.7. Showing the relationship between the sample size per round and the prevalence of pest at detection for both the simulated model and the modified rule of thumb. Sample frequency is fixed at 182 days. Landscape parameter is set to clustered ($\xi = 0$). Blue indicates the simulated prevalence of pest at detection and red indicates the early detection model prediction.

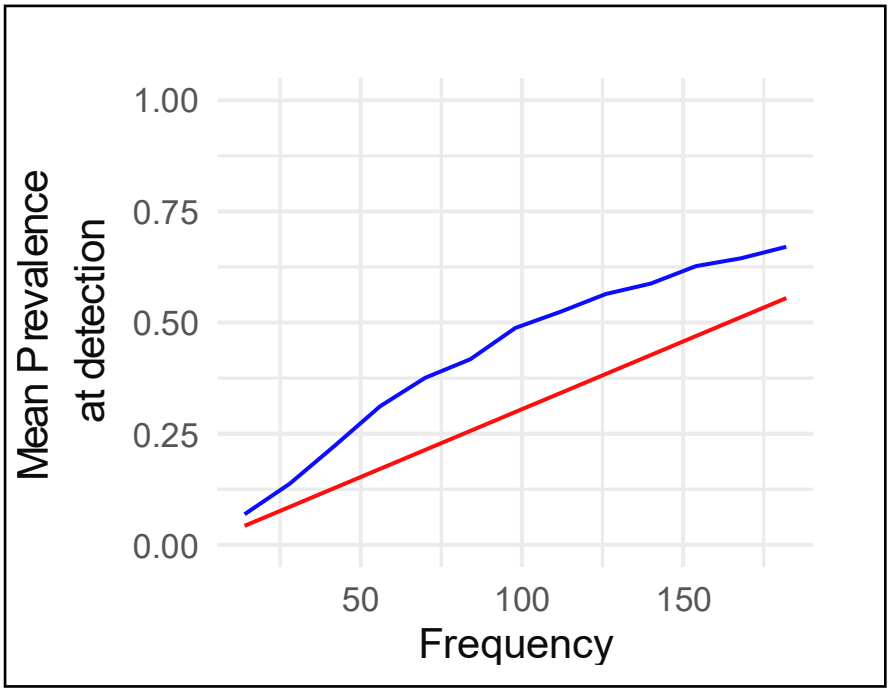


Figure 6.8. Showing the relationship between the frequency of sampling and the prevalence of pest at detection for both the simulated model and the modified rule of thumb. Sample size is fixed at 24 per round. Landscape parameter is set to clustered ($\xi = 0.5$). Blue indicates the simulated prevalence of pest at detection and red indicates the early detection model prediction.

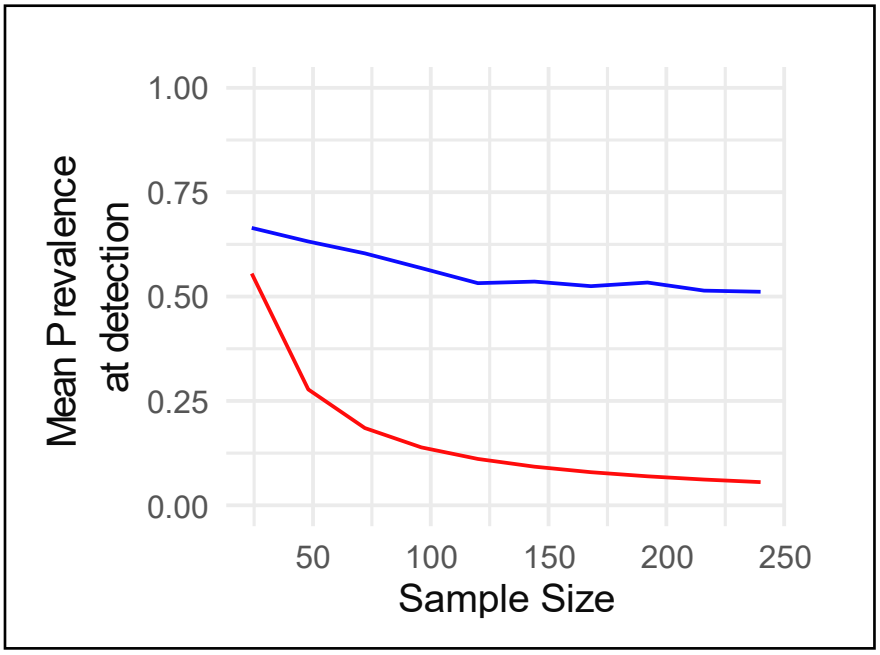


Figure 6.9. Showing the relationship between the sample size per round and the prevalence of pest at detection for both the simulated model and the modified rule of thumb. Sample frequency is fixed at 182 days. Landscape parameter is set to clustered ($\xi = 0.5$). Blue indicates the simulated prevalence of pest at detection and red indicates the early detection model prediction.

I estimated simulated prevalence of approximately 63% with the current surveillance strategy of 24 trees inspected twice a year (Figure 6.4 & 6.5). As sampling frequency was increased, both the simulated and predicted prevalence at detection reduced (Figure 6.4). This relationship between the simulated and predicted prevalence was relatively concordant, the prediction was useful in the regards that the output produced some degree of accuracy (see Figure 6.4). The range of simulated prevalence across the surveillance interval space was between 0.06 and 0.63.

The relationship between sample size and prevalence at detection indicated that the accuracy of the modified rule of thumb dropped off as sample size continued to grow (Figure 6.5). While the drop in accuracy was evident, it is important to evaluate further whether this decline was substantial enough to impact real-world application. The modified rule of thumb

demonstrated an exponential decrease in the predicted prevalence of pest at detection whilst the simulation model demonstrated a linear decrease in prevalence as sample size increased (Figure 6.5). The range of simulated prevalence across the surveillance sample size space ranged from 0.48 to 0.63. This range was much more limited that by the effects of changing the surveillance frequency.

Figure 6.9 was representative of changing surveillance sample size and surveillance frequency on more aggregated landscapes (Figure 6.9). The results from these landscapes indicated similar results to the homogenous landscapes of Figures 6.4-6.5 (Figures 6.4-6.5). Figures 6.6-6.9 were like Figures 6.4-6.5 in that the relationship between surveillance frequency and sample size and the rule of thumb accuracy was consistent. The surveillance frequency protocol changed to match the prevalence of pest at detection and the early detection model prediction with more accuracy than changing the surveillance sample size protocol. From this investigation, it appeared that even with different types of landscapes, varying from homogenous to heterogenous, that similar estimates of accuracy for the rule of thumb were generated.

6.5. Discussion

The fundamental message from the results of this chapter was that surveillance allocation according to recommended modelling structure and subsequent strategies is limiting the potential to manage the current expansion of OPM (Figures 6.4-6.9). Typical losses of around 0.05 are acceptable, and pest management strategies such as sanitation modelling are used to approach this value (Magarey & Sutton, 2007). Control methods for OPM include biological spraying and nest removal (Wadkin *et al.*, 2022). However, matching control methods to prevalence requires sufficient surveillance prior to unmanageable disease expansion.

The results demonstrated that improving the prevalence estimates of pest is directly related to increasing sampling intensity. Specifically, by increasing the number of trees sampled and the frequency in which trees are sampled. There was a disproportional impact of surveillance frequency as opposed to the impact of surveillance sampling size (Figures 6.4, 6.6 & 6.8). From this result, it can be inferred that increasing the sampling size of oak trees within a surveillance region in the UK will not lead to earlier or more accurate detection of OPM infestations. In fact, by only increasing sampling size, the prevalence of the pest upon detection remained relatively high. This is counterintuitive to the assumption that an increased surveillance sampling size protocol would yield a more comprehensive understanding of OPM prevalence.

The greater efficacy of enhanced surveillance frequency highlighted the importance of more frequent surveillance and timely assessment. More regular or frequent inspections may facilitate earlier detection of incursions and thereby increase the likelihood of successful intervention. From a practical perspective, these findings presented a significant implication for current OPM management strategies. Resources may be more effectively allocated to increasing the frequency of surveillance rather than expanding the total number of trees inspected. This shift could lead to more efficient use of resources, quicker detection, and more effective management of OPM outbreaks.

Moreover, it’s important to consider the constraints of this possible strategic shift. More frequent surveillance may require additional manpower or financial resources. Finding the balance between increasing surveillance frequency and the costs of delaying management activity is left to the discretion of stakeholders and NPPOs.

Figure 6.3 demonstrated the effects of changing the sampling protocol from the current strategy of 24 trees twice every year (Figure 6.3). Here, I found that increasing the frequency

of sampling decreased the prevalence of OPM at detection, concordant with the other results from this chapter. The distribution of output from the simulations indicated that increasing the sampling size of oak trees did reduce some prevalence of pest at detection, however, many of the simulations resulted in very large prevalence of pest at detection. This would explain why the increase in sampling did not reduce the prevalence of pest at detection by a large degree. Many of the infestations within the simulation were simply proliferating too quickly given the large interval between detection events, regardless of sampling size.

Currently, it is understood that for a region of approximately $1km^2$ the modelling strategy is the allocation of 24 tree inspections twice a year (Hoppit, 2023). Surveillance allocations should be informed by modelling approaches such as the use of the proportional odds model (POM) (Irvine & Rodhouse, 2010). The use of such models as the POM are the given benefits of knowing how much allocation is appropriate to detect trends in the changing prevalence of pest (Irvine & Rodhouse, 2010; Pocock *et al.*, 2017). The implications of the rule of thumb prescribe further methodology for choosing appropriate pest surveillance, from my results there was an impact in changing the surveillance sample size on the accuracy of the rule of thumb (Figures 6.5, 6.7 & 6.9).

The reduction of accuracy in the rule of thumb model with increasing sample size was attributed to the nature of the early epidemic expansion phase. The epidemiological simulation model assumed that epidemics would be initiated at some time point in between t_0 and the initial sampling protocol. With the relatively long-time interval of 182 days, there was sufficient time for the epidemic to initially proliferate without sampling occurring. This led to mathematical representations beyond the rule of thumb assumption of constant sampling from time point t_0 (Parnell *et al.*, 2015). Furthermore, the longer the first sampling effort was left whilst the epidemic proliferated, the less likely that the epidemic was

representative of a very small initial prevalence. This is hypothesised to have an impact on the resulting accuracy of the rule of thumb, though confirmation requires mathematical analysis and further experimentation. In effect, where sampling is infrequent (certainly when the beginning of an epidemic is unknown), the experimental results will reflect the lag between sampling points and the continuous growth of an epidemic. More frequent sampling will give more accurate prediction estimates because the epidemic will be detected before the exponential assumptions of the rule of thumb are not applicable. In addition, with finer sampling protocol, there are more opportunities for assessment and hence by the time the epidemic is detectable there will be a surveillance event within a short time frame of this occurrence, this in principle is a key assumption of the rule of thumb (continuous sampling).

The implication of this finding is that surveillance efforts should be largely focused on the repeated exploration of a region, as opposed to the allocation of more resources conditioned on the existing surveillance structures. To this end, to service the protection of Oak trees for horticultural purpose it is recommended that citizen scientists are actively recruited to survey possible OPM regions of establishment and encouraged to report sightings as regularly as possible (Pocock *et al.*, 2017).

Facilitating such activities requires the co-ordinated activity of NPPOs and stake holders. The benefit of the modeling process deployed in this chapter is based on the fundamental prediction knowledge acquired from the epidemiological simulation model itself. Clearly, prevalence of pest above 50% upon detection is challenging for stake holders to manage. The threshold for acceptable plant pest levels is usually defined as 0.05, though the broader definition of integrated pest management includes preventing damage to the extent that it causes significant economic damage (Ciancio & Mukerji, 2007; Nutter, 2007). Furthermore, evidence suggests that despite significant funding being allocated to the eradication of this

pest by Defra after legislation implementation by Forestry Commission, eradication of OPM has been a failed objective until now (Suprunenko, 2021). How much of this is dependent on early surveillance of the pest in 2005 when it was first identified? Recent research efforts into modelling this pest have struggled due to OPM's complicated aetiology despite an official pest status under the updated 10-GM, now 11-GM (Rahlenbeck & Utikal, 2015). Urticating hairs have the potential to spread several hundred meters; coupling the dispersal of OPM moths with the dispersal of their urticating hairs could be of benefit to scientists and stakeholders within the future.

I used an exponential decay dispersal kernel to shape the dispersal gradient of the pest with parameters calibrated to the 95% percentile of dispersal events occurring above 1.66 km (Suprunenko, 2021). This value reflected the early expansion phase of OPM, and may go to some way to explaining why expansion has increased since 2014. However, using this model to predict the behaviour of OPM has caveat assumptions. For example, the true nature of OPM dispersal has yet to be defined, with varying results in literature and expert opinion (Hoppit, 2023; Groenen & Meurisse, 2012). I used a modified gamma-function dispersal gradient but fixed this gradient to power-exponential dispersal for simplicity. The 95% percentile dispersal gradient was then calculated. Previous efforts have made excellent advancements in modelling the dispersal of urticating hairs in processionary moth species (Toffolo *et al.*, 2014). However, modelling efforts have not been made to match the dispersal of urticating hairs and the dispersal of the moth itself. Studies have tactically diverted attention to the establishment of nests, with various modelling approaches available for the consolidated understanding of OPM pest dynamics including geostatistical and bayesian approaches (Suprunenko *et al.*, 2021; Wadkin *et al.*, 2022).

As is the case for *Pine Processionary Moth* (PPM) in the Estemadura Province of Portugal, there is a noticeable lack of control of OPM around the greater London Metropolitan area due to the costs of managing a pest that is already firmly established (Gatto *et al.*, 2009). The short-term benefits of avoiding control measures may impact the long-term trends in market value of timber (Gatto *et al.*, 2009). Arguably, the prevalence estimates from this paper indicate that OPM early management strategies as it stands is not sufficiently efficient enough to remain below "economically acceptable thresholds" (Magarey & Sutton, 2007). The results of this paper indicate that surveillance efforts should be targeted towards the deployment of plant inspectors at more frequent intervals. Concurrently, promoting the issue of OPM to citizens is recommended to improve the rate at which trees are inspected. Efforts of co-ordinating citizen science efforts such as the Open-Air Laboratory are increasing, and these efforts could be potentially encouraged by concerned NPPOs (Brown & Williams, 2018; Pocock *et al.*, 2017).

It is hypothesised that the early detection model did not account for the early expansion phase of the epidemic, especially given the high transmission coefficient parameter value of OPM within my spatially-explicit stochastic epidemiological model. Similarly, pests with asymptomatic phases will present difficulties if the early detection model is to be applied.

To compare the effects of the proliferation of OPM between different landscape types, the simulation and surveillance models were run across different degrees of landscape clustering configurations, according to the reallocation algorithm described in the Methods of Chapter Four (See Methods, Chapter Four) (Figures 6.6-6.9). In concurrence with the findings of Chapter 2, changing the landscapes did have a small effect on the prevalence of pest, however, this was relatively small. It was inferred here that changing the landscape parameter did not have a large effect on the accuracy of the early detection model because overall the

dispersal parameter of OPM was large enough that the effects of changing landscape dynamics were effectively mitigated. My findings are a positive indicator that landscape heterogeneity is not a priority consideration when modelling OPM.

Overall, the efforts to minimise the damage caused by OPM could see significant improvement if surveillance were allocated much more frequently, ideally once a fortnight as opposed to once every half a year. It is also clear that developing models for the better understanding of OPM dynamics requires a synthesis of previously existing knowledge, though additional efforts to model key criteria in the management of the establishment of OPM in novel regions is necessary. In conclusion, the results of this thesis have been hopefully synthesised together in this chapter with a focus on OPM as a priority pest within the UK. It is hoped that the reader, having read the thesis, has understood as to why modelling the early detection of novel pests such as OPM given the current surveillance strategies is a challenge for several reasons; these reasons are the summation and integration of surveillance strategy, epidemiological behaviours, and landscape dynamics.

7. Chapter seven: Discussion

The overall objective of this thesis was to explore the extent to which simple models for complex epidemics can be used to inform early detection surveillance. I have investigated epidemiologically relevant parameters (including dispersal gradients), landscape heterogeneity, varying detection sensitivities and surveillance intensities. These variables have all been indicated to influence the accuracy of previous modelling approaches (Donatelli *et al.,* 2017; Carrasco *et al.,* 2010; Mastin *et al.,* 2020; Parnell *et al.,* 2015). In addition to this, I have applied the simple detection model to the case study of *Oak Processionary Moth* (OPM), demonstrating the insight such a model can provide and exploring implications for surveillance strategies. The purpose of these experiments was to better understand what drives epidemics in their early phases, and to use models to explore these dynamics and inform plant pest management with regards to surveillance. In addition, I have developed a useful methodology in validating simple epidemiological models with a stochastic spatially-explicit stochastic modelling approach.

I discovered that the dispersal distance of inoculum drastically affects the accuracy of the rule of thumb, and in doing so I have demonstrated how epidemiological parameters can affect the transferability (generality) of an epidemic model; the rule of thumb model explored in this thesis is most transferable to epidemics with dispersal distances greater than 60 meters, with smaller dispersal distances accuracy declines. However, the caveat is that the relative accuracy of the rule of thumb is influenced by the infection rate of the epidemic. Hence, whilst for two given epidemics with similar growth rates, a longer dispersal distance will provide greater accuracy; this information is contingent on knowing the general growth speed of the epidemic. How useful the rule of thumb potentially can be is contingent on the availability of data prior to epidemic growth, thus highlighting the importance of international co-operation and sharing of data.

In addition, with the introduction of spatial heterogeneity, I have demonstrated that early epidemic behaviours are influenced by the distribution of hosts within a range of epidemiological parameter values with regards to the dispersal distance of pests and the virulence of pests. I have shown that increased spatial heterogeneity of hosts reduces the transferability of the rule of thumb model and in doing so I have highlighted the importance of considering spatial heterogeneity when applying epidemiological models of this type. I have also showed that the accuracy of the rule of thumb reduced at lower detection sensitivities, showing how the transferability (generality) of epidemic models can be impacted by the efficacy of detection methods. However, detection sensitivities are often known, and I have shown that by including a modification which accounts for detection sensitivity, the accuracy of the rule of thumb model under different detection sensitivity scenarios, and thus the model's transferability, can be enhanced.

Together my findings, which are explored in more depth below, have added to our understanding of early detection models, their transferability (generality) and their utility for informing surveillance strategies.

7.1. Epidemiological parameters affect model transferability (generalisability)

In chapter three, I explored two epidemiological parameters associated with disease proliferation within the spatially-explicit stochastic model. These parameters were the transmission co-efficient of infection (virulence) and the dispersal distance of inoculum respectively. Here, I observed that by increasing the dispersal distance of inoculum, the accuracy of the rule of thumb increased. Furthermore, as dispersal distance increased the impact of the transmission co-efficient became less impactful on the accuracy of the rule thumb. It is important that an understanding of how epidemiological parameters affect model

transferability is integrated into considerations of the use of this early detection model. If a plant disease is constrained to dispersal distances less than 60 meters, the variability in the accuracy of the rule of thumb will be more contingent on the virulence of an epidemic. Epidemics that spread shorter distances will be more difficult to detect early on if surveillance is spread across the landscape (Meats *et al.*, 2007). If an epidemic spreads quickly before it has been detected, then more infection events have a chance to occur and stochastic forces will negatively impact the accuracy of the rule of thumb. This has the disadvantage of reducing the usefulness of the rule of thumb when a disease is potentially a higher risk for affecting the stakeholder i.e. if a disease is spreading with a very clear wave front quickly from a relatively small and localised source of inoculum. The rule of thumb will be less accurate when these short wavefront (where the "frontline" of newly infected plants are relatively constrained or compressed in terms of distance spread), highly virulent diseases are spreading, and consequently cannot be applied so confidently, reducing the overall generalisability of the rule of thumb model. However, the research of chapter three demonstrates that the impact of short-distance, highly virulent epidemics can be mitigated by increased allocation of surveillance resources. This indicates that the rule of thumb generalisability is also contingent on apriori knowledge of disease parameters, and the active role of the stake holders in accounting for these reductions in accuracy.

7.2. Distribution of hosts affects Epidemiological model transferability

Chapter four was an exploration into the effects of landscape heterogeneity. To my knowledge, this has not been explored previously. During this exploration, a notable finding was that clustering of hosts in the landscape led to a reduction in the accuracy of the chosen early detection model. This reduction can be traced to factors relevant to landscape epidemiology. Specifically, by incorporating patches into the landscape, disease progression

is slowed down by the emergence of edges, where disease inoculum has difficulty in transitioning across to other patches of susceptible hosts. The rule of thumb assumes the non-spatial exponential growth of the early phase of an epidemic. However, the results of this chapter indicate that the early detection model is compromised under realistic epidemic scenarios by landscape heterogeneity. This effect diminished with increasing scale, where the clustering effects, as depicted by the Matérn cluster process, are less pronounced with identically fixed parameters. Where there are more clusters in the landscape but an overall reduction in the total hosts per cluster, landscapes align closer to randomised landscapes and subsequently are less subject to reductions in accuracy. This finding is important, because it indicates that when applying early detection models in the future, landscape heterogeneity is going to heighten the reduction in accuracy of models that include epidemiological parameters such as dispersal distance and transmission coefficients. Which means that again for the short wavefront, highly virulent epidemics accounting for landscape heterogeneity through increased surveillance protocol would be necessary to maximise the generalisability of the rule of thumb model.

It was the objective of this chapter to explore how host density and landscape aggregation were related to dispersal. Fundamentally, it was important to explore the transmission coefficient of dispersal kernel models under changing scale and landscape conditions because the aim of this thesis was to explore the early expansion phase of mechanistic epidemic modelling in the context of surveillance. In retrospect, perhaps I could have verified whether the algorithm for generating landscapes was correctly programmed through select graphical plots. The findings from chapter three demonstrated that the transmission coefficient β influenced the accuracy of the rule of thumb model when large and also when the dispersal distance θ was short. Based on these findings, it is inferred that the transmission coefficient β will decrease accuracy of the output of the rule of thumb when applied to heterogenous

landscapes if sufficiently large enough. The aggregated effects of landscape heterogeneity and high virulence are expected to enhance the difference between the non-spatial assumptions of the early detection model and the simulated model with edge effects. The effects of each constituent of the spatially-explicit stochastic epidemiological model have influenced the accuracy of the rule of thumb. My findings suggest that with increasing complexity in epidemiological model simulations, the further accuracy deviations will occur for early detection models such as the rule of thumb. The implications of these findings indicate that surveillance modelling would benefit from further exploration.

Chapter four demonstrated that the effects of landscape aggregation impacted the accuracy of the rule of thumb, though again the dispersal parameter is the most defining parameter in the accuracy of the rule of thumb. This finding indicates that the rule of thumb is most generalisable for long distance dispersal events. As research indicates that climate change is leading to more extreme weather events, and that the movement of commodities across borders is at historic highs, the likelihood of long-distance dispersal events is hypothesised to be increasing (Defra, 2023). Of course, the focus of this thesis was not on these extreme movement events; the thesis does however indicate that the rule of thumb is most suited to large scale surveillance efforts. As with statistical methodology, the sample size of the subpopulation under question will only partially reflect the true total size of the population to some degree. In this sense, measuring the accuracy of the rule of thumb on $1km^2$ regions is a productive foundation to exploring the parameters associated with the accuracy of early detection model, but much benefit could be derived from expanding the modelling approach to the region or intercontinental scale.

7.3. The impact of detection sensitivity on transferability can be accounted for with a modifying parameter

In Chapter five, I explored the nature of sensitivity and myriad surveillance strategies as the focal point for measuring the association between early phase epidemic behaviour and the accuracy of the rule of thumb. I achieved this by varying the sensitivity of the detection likelihood parameter within the designed surveillance model whilst also exploring the frequency of sampling rounds and number of plants sampled within each round. This chapter was founded in the observation that surveillance methods vary in sensitivity, or the likelihood to detect disease when assayed. My findings indicated that the rule of thumb dropped in predictive accuracy when the sensitivity of detection was low. This is understandable, the rule of thumb assumed that detection events are occurring with 100% accuracy. I introduced a modifying parameter to account for the drop in accuracy and verified this by simulation to be more accurate than the rule of thumb alone. However, the models both performed best with 100% detection accuracy, which in reality is difficult to achieve. Fortunately, the modified rule of thumb's performance was more acceptable, and hence increased in theory the generalisability of this rule of thumb. The results indicated that for the shortest dispersal distance, relative accuracy of the Parnell *et al.*'s (2012) rule of thumb was actually improved by low sensitivity values. This result was counter intuitive, though it is hypothesised for low values of dispersal distance, a low accuracy in detection sensitivity will delay the time point of detection, allowing for the total prevalence at detection to increase and hence the total percentage difference between the rule of thumb prediction and simulated prevalence at detection to reduce. This is an important finding because it indicates that for shortwave front highly virulent epidemics, the selection of the rule of thumb model that will be most useful is dependent on the sensitivity of the assay choice. This has implications for generalisability of the rule of thumb model, as it is not perfectly clear which is the preferred model for these short wavefront epidemics.

Previous research has also confirmed that this modified early detection model works according to mathematical derivation (Mastin *et al.*, 2020). The modifier accounts for the entire range of sensitivities available to practitioners. Overall, this is a promising result for increasing the generalisability of the early detection modelling approach. The results of this chapter indicate that a modified early detection model could have potential beneficial application for real world scenarios where the detection assay sensitivity is less than 100% accurate and indicates how a model can be made more generalisable through scientific investigation.

7.4. Transferability appears independent of surveillance allocation (sampling intensity)

In addition, an investigation into the dynamics of surveillance modeling with fixed sampling intensity was conducted. My findings indicated that regardless of how sampling intensity was deployed, the accuracy of the rule of thumb model would remain consistent. This finding was promising, as it implied that surveillance operations have some degree of flexibility with regards to managing an early phase epidemic, assuming that successful management hinges on the reliability of the prediction of prevalence at first detection. Surveillance strategy is flexible and context-dependent. Therefore, having this degree of accuracy consistency independent of surveillance intensity could be of use for stakeholders. This implication is of course assuming that epidemiological parameters are within the range defined in this experiment. For example, by increasing the transmission coefficient β beyond the range explored here, epidemics may proliferate so fast that there is little allowance for flexibility in surveillance allocation. This may be the case with OPM as explored in chapter six. An additional experiment was designed to explore the effects of changing surveillance allocation and its impact on the accuracy of the early detection model. The findings of this experiment indicated that decreasing the sampling size and increasing the frequency interval of sampling

decreased the accuracy of the rule of thumb. This was indeed further corroborated by the findings for chapter six with the study of OPM. By parameterising OPM and applying the early detection model, I have demonstrated for the first time that the applicability of the rule of thumb for OPM was largely dependent on surveillance intensity. Returning to chapter five, across the parameter range this finding for changing surveillance intensity was consistent; this indicated that the accuracy of the early detection model explored in this chapter was largely dependent on the parameters of not only the epidemiological model but also the surveillance model. I recommend future research focuses on testing hypothetical areas of total sampling and total frequency and gradually reducing each parameter respectively to monitor how the growth curve deviates from perfect alignment. This iterative approach is one suggestion for how it would be possible to quantify the reduction in accuracy of the early detection model. Ultimately, the early detection model assumed constant sampling from t_0 which may have led to inaccuracies for realistic sampling protocol that used a delayed first response. This result is corroborated in chapter six, where the sample size had little impact on the prevalence of OPM when first detected because the early phase epidemic had expanded rapidly regardless of how much sampling was allocated. As for the transmission coefficient β and dispersal distance θ parameters, I observed an increase in the accuracy of the rule of thumb model with increasing dispersal distance, with the transmission coefficient causing the early detection model to be more sensitive on shorter dispersal distances. These results correspond to the results of chapter three and hence further validate the methodology of the spatially-explicit stochastic modelling approach deployed in this thesis. By gradually introducing variability in parameters across the thesis, I have demonstrated that the results of this thesis are consistent, which I hope increases the reader's confidence in the findings of the thesis.

7.5. Transferable epidemic models can enable better informed surveillance strategies

In Chapter six, I explored OPM dynamics in the context of early detection surveillance modelling. The results of this chapter indicated that the occurrence of OPM within the UK and its antecedent establishment in the Greater London region may be associated with a lack of sufficient surveillance resource deployment. Logistically, it is challenging to co-ordinate the activities of NPPOs and stakeholders, however, the cost of not deploying enough surveillance resources could outweigh the possible benefits of increasing surveillance intensity. Furthermore, the results indicated that allocating surveillance at more regular intervals as opposed to allocating more sampling on sites is a strategic decision that should be considered. There was a reduction in the accuracy of the rule of thumb when increasing surveillance intensity via sampling. This further corroborated the strategic decision to decrease the interval between sampling events. The reasoning behind the apparent reduction in the accuracy of the rule of thumb and inefficient means of reducing prevalence via increasing sampling was hypothesised to be due the early phase of expansion not being fully accounted for within the rule of thumb prediction within the context of the spatially-explicit stochastic epidemiological model. The expansion rate of OPM within the UK is rapidly increasing, the research here suggested that this expansion rate could be better managed if more surveillance were allocated sooner to intervene when prevalence within a novel region is low. Ultimately, the rule of thumb is only useful when surveillance frequency is fast enough to represent with confidence the underlying assumption of constant sampling (Parnell et al., 2012). By increasing the rate of surveillance frequency, as also observed in chapter five, it is possible to increase the accuracy of the rule of thumb. This is important, because stakeholders and NPPOs require informed decision-making to ensure that efficient plant disease management strategies can be deployed through the adequate allocation of resources such as bio-controls and biological pesticides.

7.6. Limitations

Firstly, how early detection models are expected to perform on differing scales with clustered dynamics and host densities similar to the $1km^2$ scale experiment designed has yet to be explored. Experimental evidence from field studies indicate that clustering slows down the rate of epidemic spread, however the expansion of this data analysis to various scales ranging from less than $1km^2$ to above $100km^2$ requires further investigation (Tortosa *et al.,* 2023). Secondly, the explorations of this thesis did not account for other relevant parameters to the expansion of plant pests and early detection modelling. There are numerous other compartmental models to account for, such as asymptomatic phases of disease progression that would significantly alter modelling approaches and disease dynamic behaviour (Ngah *et al.,* 2018; Cunniffe *et al.,* 2016). In addition, environmental considerations such as urbanisation (i.e. landscape channels for OPM forming indirectly through building clusters) are not accounted for in this thesis (Wang *et al.,* 2019). Ultimately, the disease triangle is the underpinning foundation to modeling plant pests, further investigation into its components is necessary to build upon the research established in this thesis.

7.7. Future works

The modelling approach taken in this thesis has been restricted to a SI compartmental model, though expansion into other models appears very feasible. By integrating other parameters into a spatially-explicit stochastic epidemiological model, it would be possible to explore other compartmental models in a similar fashion to the methods presented here to expand the knowledge of the rule of thumb's generalisability for myriad plant pests. Integrating other factors such as urbanisation, weather conditions and landscape topography I believe would present a deeper challenge to future scientists. By adjusting the rule of thumb to these other considerations or alternatively accounting for the reductions in accuracy of the rule of thumb

with these other considerations, the repertoire of scenarios known in which the rule of thumb could be applied would be expanded; hence, generalisability would be increased. These efforts I suspect would be of much benefit to further exploring how epidemiologically relevant parameters alter the predictive accuracy of early detection models. Inevitably, further models will be developed with better understanding of epidemic behaviours and how epidemics interact with monitoring efforts.

7.8. Conclusion

From this research, I hope to have presented to the reader a sufficient picture of the complexities inherent in generalisable modelling; specifically in the context of early plant pest surveillance. Early plant pest surveillance modelling presents a unique set of challenges: namely capturing the mechanistic processes associated not only with the plant disease triangle but also the surveillance efforts that must be included within the modelling process to describe even the simplest of surveillance modelling approaches. Surveillance is a fundamental aspect of early plant pest management, and from my findings I have concluded that more surveillance leads to more accurate prediction, and subsequently better management of plant pests. This is a double-edged conclusion, as an increased allocation of surveillance resources inevitably leads to higher costs of management. Thankfully, the results of this thesis indicated that it might be possible to prioritise pests according to their virulence and dispersal capabilities.

To my knowledge, the application of the rule of thumb to OPM is novel, and in addition the exploration of the parameters highlighted in this thesis in a synthesised systematic approach has yet been performed in the context of early detection modelling. The findings of this thesis are intricate, it is difficult to summarise any particular rules for early surveillance modelling without first investigating a pest on case by case basis. However, by synthesising a range of

realistic scenarios, generalised expectations of performance of the rule of thumb have been observed.

The results of this thesis have indicated that the rule of thumb model is more generalisable when stake-holders take an active role in determining a surveillance intensity which is most appropriate for the disease. Short wavefront highly virulent epidemics will decrease the accuracy of the rule of thumb more so than other types of epidemics, and to maximise the generalisability of the model, it is recommended that stakeholders use heavy surveillance intensity to account for this drop in accuracy. Furthermore, this reduction in accuracy is heightened by landscape heterogeneity, though only to a small degree which may be acceptable to stakeholders in terms of accuracy reduction. The application of the rule of thumb model and the spatially-explicit stochastic epidemiological model to OPM is novel and has highlighted that the prevalence of OPM currently expanding in the UK may be linked to a limited surveillance intensity. The OPM study also highlighted that surveillance frequency is a more important component than sample size in detecting a disease whilst it is still within early expansion phase. The accumulated efforts of this research project will hopefully inform future researchers in the field of early detection surveillance modelling as to the complexities inherent in discovering the realistic accuracy of early detection models.

8. References

Adams, M. W., Ellingboe, A. H., & Rossman, E. C. (1971). Biological uniformity and disease epidemics. *BioScience, 21*(21), 1067-1070. doi:10.2307/1295991

Agrios, G.N., (2005) *Plant Pathology (5th Edition).* Academic Press.

Ali, M. M., Bachik, N. A., Muhadi, N. A., Tuan Yusof, T. N., & Gomes, C. (2019). Non-destructive techniques of detecting plant diseases: A review. *Physiological and Molecular Plant Pathology*, *108*, 101426. https://doi.org/https://doi.org/10.1016/j.pmpp.2019.101426

Almeida, R. P. P. (2018). Emerging plant disease epidemics: Biological research is key but not enough. *PLoS Biology, 16*(8), e2007020. doi:10.1371/journal.pbio.2007020

Alonso Chavez, V., Parnell, S., & van den Bosch, F. (2016). Monitoring invasive pathogens in plant nurseries for early-detection and to minimise the probability of escape. *Journal of Theoretical Biology*, *407*, 290-302. Doi: https://doi.org/10.1016/j.jtbi.2016.07.041

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., & Daszak, P. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution, 19*(10), 535-544. doi:https://doi.org/10.1016/j.tree.2004.07.021

Antonelli, A., Hancock, S., Lennon, S., Simmonds, M., Smith, R.J. & Young, B. (2020) Protecting and sustainably using the world's plants and fungi. *Plants. People. Planet. 2*(5). 368-370. doi: https://doi.org/10.1002/ppp3.10150

Arias, J. H., Gómez-Gardeñes, J., Meloni, S., & Estrada, E. (2018). Epidemics on plants: Modeling long-range dispersal on spatially embedded networks. *Journal of Theoretical Biology*, *453*, 1-13. doi:10.1016/j.jtbi.2018.05.004

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 Bull, 42*. 515-551. doi: https://doi.org/10.1111/epp.2600

Baddeley, A., & Turner, R. (2005). spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software, 12*(6), 1–42. doi:10.18637/jss

Baldermann, S., Blagojević, L., Frede, K., Klopsch, R., Neugart, S., Neumann, A., Ngwene, B., Norkeweit, J., Schroter, D., Schroter, A., Schweigert, F.J., Wiesner, M. & Schreiner, M. (2016). Are neglected plants the food for the future? *Critical Reviews in Plant Sciences, 35*(2), 106-119. doi:10.1080/07352689.2016.1201399

Balodi, R., Bisht, S., Ghatak, A., & Rao, K. H. (2017). Plant disease diagnosis: technological advancements and challenges. *Indian Phytopathology,. 70*(3): 275-281.

Barnes, B., Giannini, F., Arthur, A., & Walker, J. (2019). Optimal allocation of limited resources to biosecurity surveillance using a portfolio theory methodology. *Ecological Economics, 161*, 153-162. https://doi.org/https://doi.org/10.1016/j.ecolecon.2019.03.012

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. doi: https://doi.org/10.1111/ppa.13231

Barrett, L. G., Kniskern, J. M., Bodenhausen, N., Zhang, W., & Bergelson, J. (2009). Continua of specificity and virulence in plant host–pathogen interactions: causes and consequences. *New Phytologist, 183*(3), 513-529. doi:10.1111/j.1469-8137.2009.02927.x

Battisti, A., Holm, G., Fagrell, B., & Larsson, S. (2011). Urticating hairs in arthropods: Their nature and medical significance. *Annual Review of Entomology, 56*(1), 203-220. https://doi.org/10.1146/annurev-ento-120709-144844

Bebber, D. P., & Gurr, S. J. (2015). Crop-destroying fungal and oomycete pathogens challenge food security. *Fungal Genetics and Biology, 74*, 62-64. doi:https://doi.org/10.1016/j.fgb.2014.10.012

Best, R. J. (1968). Tomato Spotted Wilt Virus. In K. M. Smith & M. A. Lauffer (Eds.), *Advances in Virus Research 13*, 65-146: Academic Press.

Bisin, A., & Moro, A. (2022). JUE insight: Learning epidemiology by doing: The empirical implications of a Spatial-SIR model with behavioural responses. *Journal of Urban Economics, 127*, 103368. doi:https://doi.org/10.1016/j.jue.2021.103368

Borer, E. T., Laine, A.-L., & Seabloom, E. W. (2016). A multiscale approach to plant disease using the metacommunity concept. *Annual Review of Phytopathology, 54*(1), 397-418. doi:10.1146/annurev-phyto-080615-095959

Bourhis, Y., Gottwald, T., & van den Bosch, F. (2019). Translating surveillance data into incidence estimates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 374*(1776), 20180262. doi:10.1098/rstb.2018.0262

Boyd, I. L., Freer-Smith, P. H., Gilligan, C. A., & Godfray, H. C. J. (2013). The Consequence of Tree Pests and Diseases for Ecosystem Services. *Science, 342*(6160), 1235773. doi:10.1126/science.1235773

Bradley, B.A., Blumenthal, D.M., Wilcove, D.S. & Ziska, L.H. (2010) Predicting plant invasions in an era of global change *Trends in Ecology & Evolution 25*(5), 310-318, doi: https://doi.org/10.1016/j.tree.2009.12.003

Brauer, F. (2017). Mathematical epidemiology: Past, present, and future. *Infectious Disease Modelling, 2*(2), 113-127. doi:10.1016/j.idm.2017.02.001

Bronzato Badial, A., Sherman, D., Stone, A., Gopakumar, A., Wilson, V., Schneider, W., & King, J. (2018). Nanopore sequencing as a surveillance tool for plant pathogens in plant and Insect tissues. *Plant Disease, 102*(8), 1648-1652. doi:10.1094/pdis-04-17-0488-re

Brown, D. & Bolker, B.M. (2004). The effects of disease dispersal and host clustering on the epidemic threshold in plants. *Bulletin of Mathematical Biology, 66*(2), 341-371. doi:10.1016/j.bulm.2003.08.006

Brown, J. K. M., Jessop, A. C., Rezanoor, N. H., & Hopwood, D. A. (1997). Genetic uniformity in barley and its powdery mildew pathogen. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 246*(1315), 83-90. doi:10.1098/rspb.1991.0128

Buja, I., Sabella, E., Monteduro, A. G., Chiriacò, M. S., De Bellis, L., Luvisi, A., & Maruccio, G. (2021). Advances in plant disease detection and monitoring: from traditional assays to in-field diagnostics. *Sensors, 21*(6): 2129. doi:10.3390/s21062129

Bullock, J. M., Mallada González, L., Tamme, R., Götzenberger, L., White, S. M., Pärtel, M., & Hooftman, D. A. P. (2017). A synthesis of empirical plant dispersal kernels. *Journal of Ecology, 105*(1), 6-19. doi:10.1111/1365-2745.12666

Burdon, J. J., Barrett, L. G., Yang, L.-N., He, D.-C., & Zhan, J. (2020). Maximizing world food production through disease control. *BioScience, 70*(2), 126-128. doi:10.1093/biosci/biz149

CABI. (2019). *Global Burden of Crop Loss*. "https://www.cabi.org/projects/global-burden-of-crop-loss/"

Cardillo, E., Acedo, A., & Abad, E. (2018). Topographic effects on dispersal patterns of Phytophthora cinnamomi at a stand scale in a Spanish heathland. *PLoS ONE, 13*(3), e0195060. doi:10.1371/journal.pone.0195060

Carpenter, T. E., O'Brien, J. M., Hagerman, A. D., & McCarl, B. A. (2011). Epidemic and economic impacts of delayed detection of Foot-And-Mouth disease: A case study of a simulated outbreak in california. *Journal of Veterinary Diagnostic Investigation, 23*(1), 26-33. https://doi.org/10.1177/104063871102300104

Carrasco, L. R., Mumford, J. D., MacLeod, A., Harwood, T., Grabenweger, G., Leach, A. W., . . . Baker, R. H. A. (2010). Unveiling human-assisted dispersal mechanisms in invasive alien insects: Integration of spatial stochastic simulation and phenology models. *Ecological Modelling, 221*(17), 2068-2075. https://doi.org/https://doi.org/10.1016/j.ecolmodel.2010.05.012

Cayuela, L., Hódar, J. A., & Zamora, R. (2011). Is insecticide spraying a viable and cost-efficient management practice to control pine processionary moth in Mediterranean woodlands? *Forest Ecology and Management, 261*(11), 1732-1737. https://doi.org/10.1016/j.foreco.2011.01.022

Chappell, T. M., Beaudoin, A. L. P., & Kennedy, G. G. (2013). Interacting virus abundance and transmission intensity underlie Tomato Spotted Wilt Virus incidence: An example

weather-based model for cultivated tobacco. *PLoS ONE, 8*(8), e73321. doi:10.1371/journal.pone.0073321

Chaudhary, V. B., Aguilar-Trigueros, C. A., Mansour, I., & Rillig, M. C. (2022). Fungal dispersal across spatial scales. *Annual Review of Ecology, Evolution, and Systematics, 53*(1), 69-85. doi:10.1146/annurev-ecolsys-012622-021604

Ciancio, A., & Mukerji, K. G. (2007). *General concepts in integrated pest and disease management* (1st ed.): Springer. https://doi.org/10.1007/978-1-4020-6061-8

Clarkson, J. P., Fawcett, L., Anthony, S. G., & Young, C. (2014). A model for Sclerotinia sclerotiorum infection and disease development in lettuce, based on the effects of temperature, relative humidity and ascospore density. *PLoS ONE, 9*(4), e94049. doi:10.1371/journal.pone.0094049

Coakley, S. M., Scherm, H., & Chakraborty, S. (1999). Climate change and plant disease management. *Annual Review of Phytopathology, 37*(1), 399-426. doi:10.1146/annurev.phyto.37.1.399

Cobos, A., Montes, N., López-Herranz, M., Gil-Valle, M., & Pagán, I. (2019). Within-host multiplication and speed of colonization as infection traits associated with plant virus vertical transmission. *Journal of Virology, 93*(23), doi:10.1128/jvi.01078-19

Conrad, A.O., Li, W., Lee D-W., Rogriguez-Saona L. & Bonello P. (2020) Machine Learning-Based Presymptomatic Detection of Rice Sheath Blight Using Spectral Profiles. *Plant Phenomics, 2020* doi: 10.34133/2020/8954085

Cruz, S. S., Roberts, A. G., Prior, D. A. M., Chapman, S., & Oparka, K. J. (1998). Cell-to-cell and phloem-mediated transport of Potato Virus X: The role of virions. *The Plant Cell, 10*(4), 495-510. doi:10.1105/tpc.10.4.495

Culbreath, A. K., & Srinivasan, R. (2011). Epidemiology of spotted wilt disease of peanut caused by Tomato spotted wilt virus in the southeastern U.S. *Virus Research, 159*(2), 101-109. doi: https://doi.org/10.1016/j.virusres.2011.04.014

Cunniffe, N. J., Cobb, R. C., Meentemeyer, R. K., Rizzo, D. M., & 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. Doi: https://doi.org/10.1073/pnas.1602153113

Cunniffe, N. J., Koskella, B., E. Metcalf, C. J., Parnell, S., Gottwald, T. R., & Gilligan, C. A. (2015). Thirteen challenges in modelling plant diseases. *Epidemics, 10*, 6-10. doi:10.1016/j.epidem.2014.06.002

De Luca, V., Salim, V., Atsumi, S. M., & Yu, F. (2012). Mining the biodiversity of plants: A revolution in the making. *Science, 336*(6089), 1658-1661. doi:10.1126/science.1217410

De Wolf, E. D., & Isard, S. A. (2007). Disease cycle approach to plant disease prediction. *Annual Review of Phytopathology, 45*(1), 203-220. doi:10.1146/annurev.phyto.44.070505.143329

Dean, R. A., Talbot, N. J., Ebbole, D. J., Farman, M. L., Mitchell, T. K., Orbach, M. J., Thon, M., Kulkarni, R., Xu, J-R, Pan, H, Read, N.D., Lee, Y.H, Carbone, I., Brown, D., Yee Oh, Y., Donofrio, N., Jeong, J.S., Soanes, D.M, Djonovic, S., Kolomiets, E., Rehmeyer, C., . . . Birren, B. W. (2005). The genome sequence of the rice blast fungus Magnaporthe grisea. *Nature, 434*(7036), 980-986. doi:10.1038/nature03449

Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., . . . Foster, G. D. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology, 13*(4), 414-430. doi:https://doi.org/10.1111/j.1364-3703.2011.00783.x

Department of Environment and Rural Affairs (Defra). (2014). *Protecting plant health: A plant biosecurity strategy for Great Britain.* Protecting Plant Health - A Plant Biosecurity Strategy for Great Britain (publishing.service.gov.uk)

Department of Environment and Rural Affairs (Defra). (2023). *Plant biosecurity strategy for Great Britain*. https://www.gov.uk/government/publications/plant-biosecurity-strategy-for-great-britain-2023-to-2028/plant-biosecurity-strategy-for-great-britain-2023-to-2028#:~:text=This%20strategy%20sets%20out%20the,trees%2C%20gardens%2C%20and%20countryside.

Department of Environment and Rural Affairs (Defra). (2023). *Defra's approach to pest risk analysis (PRA).* https://planthealthportal.defra.gov.uk/pests-and-diseases/pest-risk-analyses/defras-approach-to-pest-risk-analysis/#:~:text=Plant%20pests%20include%20insects%2C%20other,on%20them%20or%20causing%20disease.

Dickau, R. (2003). *Sierpinski Tetrahedron*. https://www.robertdickau.com/tetrahedron.html

De Boer, J., & Harvey, J. (2020). Range-expansion in processionary moths and biological control. *Insects, 11*(5), 267. https://doi.org/10.3390/insects11050267

Donatelli, M., Magarey, R. D., Bregaglio, S., Willocquet, L., Whish, J. P. M., & Savary, S. (2017). Modelling the impacts of pests and diseases on agricultural systems. *Agricultural Systems, 155*, 213-224. Doi: https://doi.org/10.1016/j.agsy.2017.01.019

Dukes, J.S., Pontius J., Orwig, D., Garnas, J.R., Rodgers, V.L., Brazee, N., Cooke, B., Theoharides, K.A., Stange, E.E., Harrington, R., Ehrenfield, J., Gurevitch, J., Lerdau, M., Stinson, K., Wick, R. & Ayres. M. (2009) Responses of insect pests, pathogens, and invasive plant species to climate change in the forests of northeastern North America: What can we predict? *Canadian Journal of Forest Research. 39*(2): 231-248. doi: https://doi.org/10.1139/X08-171

Dun-chun, H. Burdon, J. J., Xie, L.-h., & Zhan, J. (2021). Triple bottom-line consideration of sustainable plant disease management: From economic, sociological and ecological perspectives. *Journal of Integrative Agriculture, 20*(10), 2581-2591. doi:https://doi.org/10.1016/S2095-3119(21)63627-4

Epanchin-Niell, R. S., Brockerhoff, E. G., Kean, J. M., & Turner, J. A. (2014). Designing cost-efficient surveillance for early detection and control of multiple biological invaders. *Ecological Applications, 24*(6), 1258-1274. doi:https://doi.org/10.1890/13-1331.1

European Food Safety Authority, Lázaro, E., Parnell, S., Civera, A. V., Schans, J., Schenk, M., . . . Vos, S. (2020). General guidelines for statistically sound and risk-based surveys of plant pests. *EFSA Supporting Publications, 17*(9), 1919E. https://doi.org/https://doi.org/10.2903/sp.efsa.2020.EN-1919

EPPO (2021). *EPPO A1 and A2 lists of pests recommended for regulation as quarantine pests*. 21 Boulevard Richard Lenoir, 75011 Paris, France

Evans, H. C., & Waller, J. M. (2010). Globalisation and the Threat to Biosecurity. In R. N. Strange & M. L. Gullino (Eds.), *The role of plant pathology in food safety and food security*. (1st Ed., pp. 53-71) Dordrecht: Springer Netherlands.

Farber, D. H., De Leenheer, P., & Mundt, C. C. (2019). Dispersal kernels may be scalable: Implications from a plant pathogen. *Journal of Biogeography, 46*(9), 2042-2055. doi:10.1111/jbi.13642

Fabre, F., Coville, J. & Cunniffe, N.J. (2021). Optimising reactive disease management using spatially explicit models at the landscape scale. In: Scott, P., Strange, R., Korsten, L., Gullino, M.L. (eds) *Plant diseases and food security in the 21st century* (47-72). Springer, Cham. doi: htpps://doi.org/10.1007/978-3-030-57899-2_4

FAO. (2021). *Climate change fans spread of pests and threatens plants and crops, new FAO study*. https://www.fao.org/news/story/en/item/1402920/icode/

Filho, A. B., Inoue-Nagata, A. K., Bassanezi, R. B., Belasque Jr., J., Amorim, L., Macedo, M. A., Barbosa, J. C., Willocquet, L. & Savary, S. (2016). The importance of primary inoculum and area-wide disease management to crop health and food security. *Food Sec 8.* doi: 10.1007/s12571-015-0544-8

Filipe, J. A. N., & Maule, M. M. (2004). Effects of dispersal mechanisms on spatio-temporal development of epidemics. *Journal of Theoretical Biology, 226*(2), 125-141. doi:https://doi.org/10.1016/S0022-5193(03)00278-9

Filipe, J. A. N., Cobb, R. C., Meentemeyer, R. K., Lee, C. A., Valachovic, Y. S., Cook, A. R., Rizzo, D.M. & Gilligan, C. A. (2012). Landscape epidemiology and control of pathogens with cryptic and long-distance dispersal: Sudden Oak Death in Northern Californian forests. *PLoS Computational Biology, 8*(1), e1002328. doi:10.1371/journal.pcbi.1002328

Flood, J. (2010). The importance of plant health to food security. *Food Security 2*. 215–231. doi:10.1007/s12571-010-0072-5.

Francl, L.J., (2001) The disease triangle: a plant pathological paradigm revisited. *The plant health instructor, 1*. doi: 10.1094/PHI-T-2001-0517-01.

Futch, S., Weingarten, S., & Irey, M. (2009). Determining HLB infection levels using multiple survey methods in Florida Citrus. *Proceedings of Florida State Horticultural Society, 122*, 152-157.

Galperin, M., Graf, S., & Kenigsbuch, D. (2003). Seed treatment prevents vertical transmission of Fusarium moniliforme, making a significant contribution to disease control. *Phytoparasitica, 31*(4), 344-352. doi:10.1007/bf02979805

Gatto, P., Zocca, A., Battisti, A., Barrento, M. J., Branco, M., & Paiva, M. R. (2009). Economic assessment of managing processionary moth in pine forests: A case-study in Portugal. *Journal of Environmental Management, 90*(2), 683-691. https://doi.org/10.1016/j.jenvman.2008.01.007

Gaur, R. K., Ali, A., Cheng, X., Mäkinen, K., Agindotan, B., & Wang, X. (2021). Editorial: Plant Viruses, Volume II: Molecular plant virus epidemiology and its management. *Frontiers in Microbiology, 12*.

Gergerich, R. C., & Dolja, V. V. (2006). Introduction to plant viruses, the invisible foe. *The Plant Health Instructor.* DOI: 10.1094/PHI-I-2006-0414-01

Gilbert, G.S. & Parker, I. M. (2023) *The Evolutionary Ecology of Plant Disease.* Oxford University Press.

Global Plant Health Assessment & Savary, S. (2023). A global assessment of the state of plant health. *Plant Disease.* https://doi.org/10.1094/PDIS-01-23-0166-FE

Gogoi, A., Baruah, N., Poudel, M., Gupta, R., Baruah, G., & Kumar Borah, B. (2022). Parasitic plants as vectors for pathogens. *IntechOpen*. DOI: 10.5772/intechopen.100187

González-Domínguez, E., Fedele, G., Salinari, F. & Rossi, V. (2020). A general model for the effect of crop management on plant disease epidemics at different scales of complexity. *Agronomy, 10*(4), 462. doi: https://doi.org/10.3390/agronomy10040462

Goodman, C. A., & Hattingh, M. J. (1988). Mechanical transmission of Xanthomonas campestris pv. pruni in plum nursery trees. *Plant disease, 72*(7).

Goss, E.M., Tabima, J.F., Cooke, D.E.L. & Grünwald, N.J. (2014) The Irish potato famine pathogen *Phytophthora infestens* originated in central Mexico rather than the Andes. *PNAS Agricultural Sciences 111*(24). 8791-8796. doi: https://doi.org/10.1703/pnas.1401884111

Gottschling, S., & Meyer, S. (2006). An epidemic airborne disease caused by the Oak Processionary caterpillar. *Pediatric Dermatology, 23*(1), 64-66. https://doi.org/https://doi.org/10.1111/j.1525-1470.2006.00173.x

Gottwald, T. R., Graham, J. H., & Schubert, T. S. (2002). Citrus Canker: The pathogen and its impact. *Plant Health Progress, 3*(1), 15. doi:10.1094/PHP-2002-0812-01-RV

Gottwald, T. R. (2010). Current epidemiological understanding of citrus huanglongbing. *Annual Review of Phytopathology, 48*(1), 119-139. doi:10.1146/annurev-phyto-073009-114418

Gregory, P. H. (1968). Interpreting Plant disease dispersal gradients. *Annual Review of Phytopathology, 6*(1), 189-212. doi:10.1146/annurev.py.06.090168.001201

Grierson, C, Barnes, S.R., Chase, M.W., Clarke, M., Grierson, D., Edwards, K.J., Jellis, G.J., Jones, J.D., Knapp, S., Oldroyd, G., Poppy, G., Temple, P., Williams, R., Bastow, R. (2011).

One hundred important questions facing plant science research. *New Phytologist 192*(1), 6-12. doi: https://doi.org/10.1111/j.1469-8137.2011.03859.x

Groenen, F., & Meurisse, N. (2012). Historical distribution of the oak processionary moth Thaumetopoea processionea in Europe suggests recolonization instead of expansion. *Agricultural and Forest Entomology, 14*(2), 147-155. https://doi.org/https://doi.org/10.1111/j.1461-9563.2011.00552.x

Grosdidier, M., Scordia, T., Ioos, R., & Marçais, B. (2020). Landscape epidemiology of ash dieback. *Journal of Ecology, 108*(5), 1789-1799. https://doi.org/10.1111/1365-2745.13383

Hall, C., & Knuth, M. (2019). An update of the literature supporting the well-being benefits of plants: A review of the emotional and mental health benefits of plants. *Journal of Environmental Horticulture, 37*(1), 30-38. doi:10.24266/0738-2898-37.1.30

Hauser, C. E., & McCarthy, M. A. (2009). Streamlining 'search and destroy': cost-effective surveillance for invasive species management. *Ecology Letters, 12*(7), 683-692. doi:10.1111/j.1461-0248.2009.01323.x

He, D.-c., Zhan, J.-s., & Xie, L.-h. (2016). Problems, challenges and future of plant disease management: from an ecological point of view. *Journal of Integrative Agriculture, 15*(4), 705-715. doi:https://doi.org/10.1016/S2095-3119(15)61300-4

He, K., Lin, K., Ding, S., Wang, G., & Li, F. (2019). The vitellogenin receptor has an essential role in vertical transmission of Rice Stripe Virus during oogenesis in the small brown plant hopper. *Pest Management Science, 75*(5), 1370-1382. doi:10.1002/ps.5256

Hellmann, J.J., Byers, J.E., Bierwagen, B.G. & Dukes, J.S. (2007). Five potential consequences of climate change for invasive species. *Conservative Biology, 22*(3), 534-543. doi: 10.1111/j.1523-1739.2008.00951.x

Heil, M., & Bostock, R. M. (2002). Induced Systemic Resistance (ISR) against pathogens in the context of induced plant defences. *Annals of Botany, 89*(5), 503-512. doi:10.1093/aob/mcf076

Hill, L., Hector, A., Hemery, G., Smart, S., Tanadini, M., & Brown, N. (2017). Abundance distributions for tree species in Great Britain: A two-stage approach to modeling abundance using species distribution modeling and random forest. *Ecology and Evolution, 7*(4), 1043-1056. https://doi.org/10.1002/ece3.2661

Hill, L., Jones, G., Atkinson, N., Hector, A., Hemery, G., & Brown, N. (2019). The £15 billion cost of Ash Dieback in Britain. *Current Biology, 29*(9), R315-R316. doi:https://doi.org/10.1016/j.cub.2019.03.033

Holdenrieder, O., Pautasso, M., Weisberg, P. J., & Lonsdale, D. (2004). Tree diseases and landscape processes: the challenge of landscape pathology. *Trends in Ecology & Evolution, 19*(8), 446-452. doi:https://doi.org/10.1016/j.tree.2004.06.003

Horino, O., Mew, T. W., & Yamada, T. (1982). The effect of temperature on the development of bacterial leaf blight on rice. *Japanese Journal of Phytopathology*, *48*(1), 72-75.

Hywatt-Twynam, S.R., Parnell, S., Stutt, R.O.J.H., Gottwald, T.R., Gilligan, C.A. and Cunniffe, N.J. (2017). Risk-based management of invading plant disease. *New Phytol*, *214*. 1317-1329. doi: https://doi.org/10.1111/nph.14488

International Plant Protection Convention (IPPC). (1997) International plant protection convention. 1997 - International Plant Protection Convention (New Revised Text) - International Plant Protection Convention (ippc.int)

International Plant Protection Convention (IPPC) Secretariat, (2016). Surveillance guide – A guide to understand the principal requirements of surveillance programmes for national plant protection organizations. FAO.

International Plant Protection Convention (IPPC) Secretariat, (2021). 2021 IPPC annual report – Protecting the world's plant resources from pests. 2021 IPPC Annual Report - Protecting the world's plant resources from pests (fao.org)

Ioannidis, J. P. A., & Powis, S. H. (2022). COVID-19 models and expectations - Learning from the pandemic. *Advances in biological regulation*, *86*, 100922. doi: https://doi.org/10.1016/j.jbior.2022.100922

Irvine, K. M., & Rodhouse, T. J. (2010). Power analysis for trend in ordinal cover classes: implications for long-term vegetation monitoring. *Journal of Vegetation Science, 21*(6), 1152-1161. https://doi.org/10.1111/j.1654-1103.2010.01214.x

Jackson, S. L., & Bayliss, K. L. (2011). Spore traps need improvement to fulfil plant biosecurity requirements. *Plant Pathology, 60*(5), 801-810. doi:10.1111/j.1365-3059.2011.02445.x

Jacobs, D. F. (2007). Toward development of silvical strategies for forest restoration of American chestnut (Castanea dentata) using blight-resistant hybrids. *Biological Conservation, 137*(4), 497-506. doi:https://doi.org/10.1016/j.biocon.2007.03.013

Jakubowska A.K, Murillo R, Carballo A, Williams T, van Lent J.W.M, Caballero P, & S., H. (2016). Iflavirus increases its infectivity and physical stability in association with baculovirus. *PeerJ, 4*, e1687. doi:https://doi.org/10.7717/peerj.1687

James, W. C. (1974). Assessment of plant diseases and losses. *Annual Review of Phytopathology, 12*(1), 27-48. doi:10.1146/annurev.py.12.090174.000331

Jaspers, M.V., Falloon, P.G. & Pearson, M.N. (2015). Seed and pollen transmission of asparagus virus 2. *Eur J Plant Pathol* 142, 173–183. doi: https://doi.org/10.1007/s10658-015-0601-3

Jeger, M. J. (2000). Theory and plant epidemiology. *Plant Pathology, 49*(6), 651-658. doi:10.1046/j.1365-3059.2000.00522.x

Jetiyanon, K., & Kloepper, J. W. (2002). Mixtures of plant growth-promoting rhizobacteria for induction of systemic resistance against multiple plant diseases. *Biological Control, 24*(3), 285-291. doi:https://doi.org/10.1016/S1049-9644(02)00022-1

Jones, J. T., Haegeman, A., Danchin, E. G. J., Gaur, H. S., Helder, J., Jones, M. G. K., Kikuchi, T., Manzanilla-López, Palomares-Rius, J.E., Wesemael, W.M.L., Perry, R. N. (2013). Top 10 plant-parasitic nematodes in molecular plant pathology. *Molecular Plant Pathology, 14*(9), 946-961. doi:10.1111/mpp.12057

Johnson, B. L., & Haddad, N. M. (2011). Edge effects, not connectivity, determine the incidence and development of a foliar fungal plant disease. *Ecology, 92*(8), 1551-1558. doi:10.1890/10-1072.1

Jones, A.M. & Harrison, R.M. (2004) The effects of meteorological factors on atmospheric bioaerosol concentrations – a review. *Science of the total environment 326*(1-3), 151-180. doi: https://doi.org/10.1016/j.scitotenv.2003.11.021

Jung, T., Blaschke, H., & Oswald, W. (2000). Involvement of soilborne Phytophthora species in Central European oak decline and the effect of site factors on the disease. *Plant Pathology, 49*(6), 706-718. doi:https://doi.org/10.1046/j.1365-3059.2000.00521.x

Kalaris, T., Fieselmann, D., Magarey, R., Colunga-Garcia, M., Roda, A., Hardie, D., Cogger, N., Hammond, N., Tony Martin, P.A. & Whittle, P. (2013). *The role of surveillance methods and technologies in plant biosecurity*. 307-337. Springer, Dodrecht. doi: https://doi.org/10.1007/978-94-007-7365-3_11

Kamoun, S., Furzer, O., Jones, J.D.G., Judelson, H.S., Ali, G.S., Dalio, R.J.D, Roy, S.G., Schena, L., Zambounis, A., Panabières, F., Cahill, D., Ruocco, M., Figueiredo, A., Chen, X., Hulvey, J., Stam, R., Lamour, K., Gijzen, M., Tyler, B.M., Grünwald, N.J, Mukhtar, M.S., Tomé, D.F.A, Tör, M., van den Ackerveken, G., McDowell, J., Daayf, F., Fry, W.E., Lindqvist-Kreuze, H., Meijer, H.J.G., Petre, B., Ristaino, J., Yoshida, K., Birch, P.R.J & Govers, F. (2014) The top 10 oomycete pathogens in molecular plant pathology. *Molecular Plant Pathology*, *16*(4), 413-434. doi: https://doi.org/10.1111/mpp.12190

Kaniyassery, A., Thorat, S. A., Kiran, K. R., Murali, T. S., & Muthusamy, A. (2023). Fungal diseases of eggplant (Solanum melongena L.) and components of the disease triangle: A review. *Journal of Crop Improvement, 37*(4), 543-594. https://doi.org/10.1080/15427528.2022.2120145

Kannan, V. R., & Bastas, K. K. (2015). Sustainable approaches to controlling plant pathogenic bacteria (1st ed.): CRC Press. https://doi.org/10.1201/b18892

Kc, K. B., Dias, G. M., Veeramani, A., Swanton, C. J., Fraser, D., Steinke, D., Lee, E., Wittman, H., Farber, J.M., Dunfield, K., McCann, K., Anand, M., Campbell, M., Rooney, N., Raine, N.E., Van Acker, R., Hanner, R., Pascoal, S., Sharif, S., Benton, T., Fraser, E. D. G. (2018). When too much isn't enough: Does current food production meet global nutritional needs? *PLoS ONE, 13*(10), e0205683. doi:10.1371/journal.pone.0205683

Keeling, M. J., & Rohani, P. (2007). Modeling infectious diseases in humans and animals (1st ed.) : Princeton University Press. doi: https://doi.org/10.2307/j.ctvcm4gk0

Kim, E., Seo, Y., Su Kim, Y., Park, Y. & Kim, Y.H. (2017) Effects of soil texture on infectivity of Root-Knot nematodes on carrots. *The plant pathology journal 33*(1), 66-74. doi: 10.5423/PPJ.OA.07.2016.0155

Klosterman, S. J., Anchieta, A., McRoberts, N., Koike, S. T., Subbarao, K. V., Voglmayr, H., Choi, Y.J., Thines, M. & Martin, F. N. (2014). Coupling spore traps and quantitative PCR assays for detection of the Downy Mildew pathogens of spinach peronospora effuse and Beet P. schachtii. *Phytopathology, 104*(12), 1349-1359. doi:10.1094/phyto-02-14-0054-r

Kompas, T., Chu, L., Van Ha, P., & Spring, D. (2019). Budgeting and portfolio allocation for biosecurity measures. *Australian Journal of Agricultural and Resource Economics*, *63*: 412-438. doi:10.1111/1467-8489.12305

Kullman, L. (2006). Tree line population monitoring of *Pinus sylvestris* in the Swedish Scandes, 1973-2005: implications for tree line theory and climate change ecology. *Journal of Ecology (95)*1, 41-52. doi:https://doi.org/10.1111/j.1365-2745.2006.01190.x

Kumar, P. (2004). Valuation of medicinal plants for pharmaceutical uses. *Current Science, 86*(7), 930-937. http://www.jstor.org/stable/24109275.

Lagomarsino Oneto, D., Golan, J., Mazzino, A., Pringle, A., Seminara, A., 2020. Timing of fungal spore release dictates survival during atmospheric transport. *Proceedings of the National Academy of Sciences 117*, 5134–5143. https://doi.org/10.1073/pnas.1913752117

Lapin, D., & van den Ackerveken, G. (2013). Susceptibility to plant disease: more than a failure of host immunity. *Trends in Plant Science, 18*(10), 546-554. doi:10.1016/j.tplants.2013.05.005

Lamichhane, J. R., & Venturi, V. (2015). Synergisms between microbial pathogens in plant disease complexes: a growing trend [Review]. *Frontiers in Plant Science, 6.*

Laughlin, D. C. (2014). The intrinsic dimensionality of plant traits and its relevance to community assembly. *Journal of Ecology, 102*(1), 186-193. doi:https://doi.org/10.1111/1365-2745.12187

Lázaro, E., Sesé, M., López-Quílez, A., Conesa, D., Dalmau, V., Ferrer, A., & Vicent, A. (2021). Tracking the outbreak: an optimized sequential adaptive strategy for Xylella fastidiosa delimiting surveys. *Biological Invasions, 23*(10), 3243-3261. doi:10.1007/s10530-021-02572-x

Le, D. T., & Vu, N. T. (2017). Progress of loop-mediated isothermal amplification technique in molecular diagnosis of plant diseases. *Applied Biological Chemistry, 60*(2), 169-180. doi:10.1007/s13765-017-0267-y

Leclerc, M., Doré, T., Gilligan, C. A., Lucas, P., & Filipe, J. A. N. (2014). Estimating the delay between host infection and disease (incubation period) and assessing its significance to the epidemiology of plant diseases. *PLoS ONE, 9*(1), e86568. doi:10.1371/journal.pone.0086568

Legg, J. P., Jeremiah, S. C., Obiero, H. M., Maruthi, M. N., Ndyetabula, I., Okao-Okuja, G., Bouwmeester, H., Bigirimana, S., Tata-Hangy, W., Gashaka, G., Mkamilo, G., Alicai, T. & Lava Kumar, P. (2011). Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Research, 159*(2), 161-170. doi:https://doi.org/10.1016/j.virusres.2011.04.018

Liu, J., & Wang, X. (2020). Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model. *Plant Methods, 16*(1). doi:10.1186/s13007-020-00624-2

Lortie, C. J., Filazzola, A., Westphal, M., & Butterfield, H. S. (2022). Foundation plant species provide resilience and microclimatic heterogeneity in drylands. *Scientific Reports, 12*(1). doi:10.1038/s41598-022-22579-1

Madden, L. V., & Hughes, G. (1995). Plant disease incidence: distributions, heterogeneity, and temporal analysis. *Annual Review of Phytopathology, 33*(1), 529-564. doi:10.1146/annurev.py.33.090195.002525

Madden, L. V. (2006). Botanical epidemiology: Some key advances and its continuing role in disease management. *European Journal of Plant Pathology, 115*(1), 3-23. doi:10.1007/s10658-005-1229-5

Madden, L. V., Hughes, G. & van den Bosch, F. (2017) *The study of plant disease epidemics.* APS Publications.

Magarey, R. D., & Sutton, T. B. (2007). *General concepts in integrated pest and disease Management* (1st ed.): Springer Link. 3-25. https://doi.org/10.1007/978-1-4020-6061-8

Magnani, R., Sabin, K., Saidel, T., & Heckathorn, D. (2005). Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS*, 19 Suppl 2:S67-72. doi: 10.1097/01.aids.0000172879.20628.e1. PMID: 15930843.

Malavasi, M., Santoro, R., Cutini, M., Acosta, A. T. R., & Carranza, M. L. (2016). The impact of human pressure on landscape patterns and plant species richness in Mediterranean coastal dunes. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, *150*(1), 73-82. doi:10.1080/11263504.2014.913730

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., Toth, I., Salmond, G. & Foster, G. D. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology*, *13*(6), 614-629. doi:10.1111/j.1364-3703.2012.00804.x

Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L.R., Davis, C.E. & Dandekar, A. M. (2015). Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*, *35*(1), 1-25. doi:10.1007/s13593-014-0246-1

Martini, X., Malfa, K., Stelinski, L. L., Iriarte, F. B., & Paret, M. L. (2020). Distribution, phenology, and overwintering survival of Asian Citrus Psyllid (Hemiptera: Liviidae), in urban and grove habitats in North Florida. *Journal of Economic Entomology, 113*(3), 1080-1087. doi:10.1093/jee/toaa011

Mastin, A. J., van den Bosch, F., van den Berg, F., & Parnell, S. R. (2019). Quantifying the hidden costs of imperfect detection for early detection surveillance. *Philosophical Transactions of the Royal Society B: Biological Sciences, 374*(1776), 20180261.

Mastin, A. J., Gottwald, T. R., van den Bosch, F., Cunniffe, N. J., & Parnell, S. (2020). Optimising risk-based surveillance for early detection of invasive plant pathogens. *PLoS biology, 18*(10), e3000863-e3000863. doi:10.1371/journal.pbio.3000863

Matta, C. (2010) Spontaneous generation and disease causation: Anton de Bary's experiments with *Phytophthora infestans* and Late Blight of potato. *Journal of the History of Biology 43.* 459-491. doi: 10.1007/s10739-009-9220-1.

McMullan, M., Rafiqi, M., Kaithakottil, G., Clavijo, B.J., Bilham, L., Orton, E., Percival-Alwyn, L., Ward, B. J., Edwards, A., Saunders, D.G.O., Accinelli, G. G., Wright, J., Verweij, W., Koutsovoulos, G., Yoshida, K., Hosoya, T., Williamson, L., Jennings, P., Ioos, R., Husson, C., Hietala, A. M., Vivian-Smith, A., Solheim, H., MaClean, D., … Clark, M. The ash dieback invasion of Europe was founded by two genetically divergent individuals. *Nature Ecology & Evolution (2).* 1000-1008. doi: https://doi.org/10.1038/s41559-018-0548-9.

Meats, A., & Smallridge, C. J. (2007). Short- and long-range dispersal of medfly, Ceratitis capitata (Dipt., Tephritidae), and its invasive potential. *Journal of Applied Entomology, 131*(8), 518-523. Doi: https://doi.org/10.1111/j.1439-0418.2007.01168.x

Meentemeyer, R. K., Anacker, B. L., Mark, W., Rizzo, D. M. (2008). Early detection of emerging forest disease using dispersal estimation and ecological niche modelling. *Ecological Applications 18*(2). 377-390. doi: https://doi.org/10.1890/07-1150.1

Meentemeyer, R. K., Haas, S. E., & Václavík, T. (2012). Landscape epidemiology of emerging infectious diseases in natural and human-altered ecosystems. *Annual Review of Phytopathology, 50*(1), 379-402. doi:10.1146/annurev-phyto-081211-172938

Meredith, D. S. (1973). Significance of spore release and dispersal mechanisms in plant disease epidemiology. *Annual Review of Phytopathology, 11*(1), 313-342. doi:10.1146/annurev.py.11.090173.001525

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science, 7*. Doi: 10.3389/fpls.2016.01419

Moore, J. L., Rout, T. M., Hauser, C. E., Moro, D., Jones, M., Wilcox, C., & Possingham, H. P. (2010). Protecting islands from pest invasion: optimal allocation of biosecurity resources between quarantine and surveillance. *Biological Conservation, 143*(5): 1068-1078. https://doi.org/https://doi.org/10.1016/j.biocon.2010.01.019

Moreira, A.S., Stuchi, E.S., Silva, P.R.B. Bassanezi, R.B., Girardi, E.A. & Laranjeira, F.F. (2019). Could tree density play a role in managing Citrus Huanglongbing epidemics? *44*, 268-274. doi: https://doi.org/10.1007/s40858-019-00284-1

Mota, M. M., Futai, K., Vieira, P. (2009) *Integrated management of fruit crop nematodes.* Springer, Dordrecht. doi: https://doi.org/10.1007/978-1-4020-9858-1_11

Muhie, S.H., (2022) Novel approaches and practises to sustainable agriculture. *Journal of Agriculture and Food Research, 10.* Doi: https://doi.org/10.1016/j.jafr.2022.100446

Mukherjee, R., Gruszewski, H.A., Bilyeu, L.T., Schmale III, D.G. & Boreyko, J.B. (2021) Synergistic dispersal of plant pathogens spores by jumping-droplet condensation and wind. *Proceedings of the National Academy of Sciences of the United States of America. 118*(34), doi: https://doi.org.10.1073/pnas.2106938118

Mundt, C. C., Sackett, K. E., & Wallace, L. D. (2011). Landscape heterogeneity and disease spread: experimental approaches with a plant pathogen. *Ecological Applications, 21*(2), 321-328. doi:10.1890/10-1004.1

Mundt, C. C., & Sackett, K. E. (2012). Spatial scaling relationships for spread of disease caused by a wind-dispersed plant pathogen. *Ecosphere, 3*(3), art24. doi:10.1890/es11-00281.1

Murakishi, H. H., Hartmann, J. X., Beachy, R. N., & Pelcher, L. E. (1971). Growth curve and yield of tobacco mosaic virus in tobacco callus cells. *Virology, 43*(1), 62-68. doi: https://doi.org/10.1016/0042-6822(71)90224-8

Nagel, J.M., Huxman, T.E., Griffin, K.L. & Smith, S.D. (2004) CO_2 Enrichment reduces the energetic of biomass construction in an invasive desert grass. *Ecology, 85*(1), 100-106. doi: https://doi.org/10.1890/02-3005.

Nalley, L., Tsiboe, F., Durand-Morat, A., Shew, A., & Thoma, G. (2016). Economic and environmental impact of rice blast pathogen (Magnaporthe oryzae) Alleviation in the United States. *PLoS ONE, 11*(12), e0167295. doi:10.1371/journal.pone.0167295

Nazarov, P. A., Baleev, D. N., Ivanova, M. I., Sokolova, L. M., & Karakozova, M. V. (2020). Infectious plant diseases: etiology, current status, problems and prospects in plant protection. *Acta Naturae, 12*(3), 46-59. doi:10.32607/actanaturae.11026

Neyman, J., & Scott, E. L. (1958). Statistical Approach to Problems of Cosmology. *Journal of the Royal Statistical Society. Series B (Methodological), 20*(1), 1–43. http://www.jstor.org/stable/2983905 doi:10.1098/rstb.2018.0261

Newlands, N. K. (2018). Model-based forecasting of agricultural crop disease risk at the regional scale, integrating airborne inoculum, environmental, and satellite-based monitoring data. *Frontiers in Environmental Science, 6.* doi:10.3389/fenvs.2018.00063

Ngah, N., Thomas, R., Shaw, M., & Fellowes, M. (2018). Asymptomatic host plant infection by the widespread pathogen Botrytis cinerea alters the life histories, Behaviors, and Interactions of an Aphid and Its Natural Enemies. *Insects, 9*(3), 80. Doi: https://doi.org/10.3390/insects9030080

Norelli, J. L., Jones, A. L., & Aldwinckle, H. S. (2003). Fire Blight management in the twenty-first century: Using new technologies that enhance host resistance in apple. *Plant Disease, 87*(7), 756-765. doi:10.1094/pdis.2003.87.7.756

Nutter, F. F. W. (2007). The Role of plant disease epidemiology in developing successful integrated disease management programs. In A. Ciancio & K. G. Mukerji (Eds.), *General Concepts in Integrated Pest and Disease Management* (1st ed.) (pp. 45-79). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6061-8_3

Pagán, I., Montes, N., Milgroom, M.G. & García-Arenal, F. (2014) Vertical transmission selects for reduced virulence in a plant virus and for increased resistance in the host. *Plos Pathogens 10*(7). doi: https://doi.org/10.1371/journal.ppat.1004293

Parlevliet, J. E. (1979). Components of resistance that reduce the rate of epidemic development. *Annual Review of Phytopathology, 17*(1), 203-222. doi:10.1146/annurev.py.17.090179.001223

Parker, I. M., & Gilbert, G. S. (2004). The evolutionary ecology of novel plant-pathogen interactions. *Annual Review of Ecology, Evolution, and Systematics, 35*(1): 675-700. doi:10.1146/annurev.ecolsys.34.011802.132339

Parnell, S., Gottwald, T. R., Gilks, W. R., & van den Bosch, F. (2012). Estimating the incidence of an epidemic when it is first discovered and the design of early detection monitoring. *Journal of Theoretical Biology, 305*, 30-36. doi:10.1016/j.jtbi.2012.03.009

Parnell, S., Gottwald, T. R., Riley, T., & van den Bosch, F. (2014). A generic risk-based surveying method for invading plant pathogens. *Ecological Applications, 24*(4), 779-790. doi:https://doi.org/10.1890/13-0704.1

Parnell, S., Gottwald, T. R., Cunniffe, N. J., Alonso Chavez, V., & van den Bosch, F. (2015). Early detection surveillance for an emerging plant pathogen: a rule of thumb to predict prevalence at first discovery. *Proceedings of the Royal Society B: Biological Sciences, 282*(1814), 20151478. doi:10.1098/rspb.2015.1478

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*(1), 591-610. doi:10.1146/annurev-phyto-080516-035334

Parrella, G., Gognalons, P., Gebre-Selassiè, K., Vovlas, C., & Marchoux, G. (2003). An update of the host range of tomato spotted wilt virus. *Journal of Plant Pathology, 85*(4), 227-264.

Patil, S. B., & Bodhe, S. K. (2011). Leaf disease severity measurement using image processing. *International Journal of Engineering 3*(5).

Pautasso, M., Döring, T. F., Garbelotto, M., Pellis, L., & Jeger, M. J. (2012). Impacts of climate change on plant diseases—opinions and trends. *European Journal of Plant Pathology, 133*(1), 295-313. doi:10.1007/s10658-012-9936-1

Pegg, K. G., Coates, L. M., O'Neill, W. T., & Turner, D. W. (2019). The epidemiology of fusarium wilt of banana. *Frontiers in Plant Science, 10*. Doi=10.3389/fpls.2019.01395

Peng, Y., Dallas, M. M., Ascencio-Ibáñez, J. T., Hoyer, J. S., Legg, J., Hanley-Bowdoin, L., . . . Yin, H. (2022). Early detection of plant virus infection using multispectral imaging and spatial–spectral machine learning. *Scientific Reports, 12*(1). doi:10.1038/s41598-022-06372-8

Pfender (1981) Monocyclic and Polycyclic Root Diseases: Distinguishing Between the Nature of the Disease Cycle and the Shape of the Disease Progress Curve. *American Phytopathology Society.*

Plantegenest, M., Le May, C., & Fabre, F. (2007). Landscape epidemiology of plant diseases. *Journal of The Royal Society Interface, 4*(16), 963-972. doi:10.1098/rsif.2007.1114

Phillips, B., xa, L, Brown, G., P, Travis, J. & Shine, R. Symposium Editor: Trevor, D. P. (2008). Reid's paradox revisited: The evolution of dispersal kernels during range expansion. *The American Naturalist, 172*(S1), S34-S48. doi:10.1086/588255

Pilz, C. & Tabarelli, D., (2020) *Sustainable Plant Disease Protection Strategies to Keep Your Customers Crops Healthy.* https://www.precisionfarmingdealer.com/articles/4343-sustainable-plant-disease-protection-strategies-to-keep-your-customers-crops-healthy

Plantegenest, M., Le May, C., & Fabre, F. (2007). Landscape epidemiology of plant diseases. *Journal of The Royal Society Interface, 4*(16), 963-972. doi:10.1098/rsif.2007.1114

Ponnuvel, K. M., de Miranda, J. R., Terenius, O., Li, W., Ito, K., Khajje, D., Shamintha, G., Jagadish, A., Dubey, H., Mishra, R. K. (2022). Genetic characterisation of an Iflavirus associated with a vomiting disease in the Indian tropical tasar silkworm, Antheraea mylitta. *Virus Research, 311,* 198703. doi:https://doi.org/10.1016/j.virusres.2022.198703

Potts, J. M., Cox, M. J., Barkley, P., Christian, R., Telford, G., & Burgman, M. A. (2013). Model-based search strategies for plant diseases: a case study using citrus canker (Xanthomonas citri). *Diversity and Distributions, 19*(5-6), 590-602. doi:10.1111/ddi.12065

Ponge, J.-F. (2005). Emergent properties from organisms to ecosystems: towards a realistic approach. *Biological Reviews, 80*(03), 403. doi:10.1017/s146479310500672x

Potter, C., Harwood, T., Knight, J., & Tomlinson, I. (2011). Learning from history, predicting the future: the UK Dutch elm disease outbreak in relation to contemporary tree disease threats. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1573), 1966-1974. doi:10.1098/rstb.2010.0395

Prospero, J.M., Blades, E., Mathison, G. & Naidu, R. (2005) Interhemispheric transport of viable fungi and bacteria from Africa to the Caribbean with soil dust. *Aerobiologia* (21), 1–19. doi: https://doi.org/10.1007/s10453-004-5872-7

Richter-Heitmann, T., Hofner, B., Krah, F-S., Sikorski, J., Wüst, P.K., Bunk, B., Huang, S., Regan, K.M., Berner, D., Boeddinghaus, R.S., Marhan, S., Prati, D., Kandeler, E., Overmann, J., Friedrich, M.W. (2020) Stochastic dispersal rather than deterministic selection explains the spatio-temporal distribution of soil bacteria in a temperate grassland. *Frontiers in Microbiology 11*(1393). doi: 10.3389/fmicb.2020.01391

Rieux, A., Soubeyrand, S., Bonnot, F., Klein, E. K., Ngando, J. E., Mehl, A., Ravigne, V., Carlier, J. & De Lapeyre De Bellaire, L. (2014). Long-distance wind-dispersal of spores in a fungal plant pathogen: Estimation of anisotropic dispersal kernels from an extensive field experiment. *PLoS ONE, 9*(8), e103225. doi:10.1371/journal.pone.0103225

Riley, S., Eames, K., Isham, V., Mollison, D. & Trapman, P., (2015) Five challenges for spatial epidemic models. *Epidemics 10*. 68-71. doi: https://doi.org/10.1016/j.epidem.2014.07.001.

Rimbaud, L., Fabre, F., Papaïx, J., Moury, B., Lannou, C., Barret, L.G. & Thrall, P.H. (2021) Models of plant resistance deployment. *Annual review of phytopathology 59.* 125-152. doi: https://doi.org/10.1146/annurev-phyto-020620-122134

Ristaino, J. B., & Gumpertz, M. L. (2000). New Frontiers in the Study of Dispersal and Spatial Analysis of Epidemics Caused by Species in the Genus Phytophthora. *Annual Review of Phytopathology, 38*(1), 541-576. doi:10.1146/annurev.phyto.38.1.541

Ristaino J. B., Anderson P. K., Bebber D. P., Brauman K. A., Cunniffe N. J., Fedoroff N.V., Finegold, C., Garret, K.A., Gilligan, C., Jones, C.M., Martin, M.D., MacDonald, G.K., Neenan, P., Records, A., Schmale, D.G., Tateosian, L. & Wei, Q. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences, 118*(23), e2022239118. doi:10.1073/pnas.2022239118

Rizzo, D. M., Garbelotto, M., & Hansen, E. M. (2005). Phytophthora ramorum: Integrative research and management of an emerging pathogen in California and Oregon forests. *Annual Review of Phytopathology, 43*(1), 309-335. doi:10.1146/annurev.phyto.42.040803.140418

Robinet, C., Imbert, C.-E., Rousselet, J., Sauvard, D., Garcia, J., Goussard, F., & Roques, A. (2012). Human-mediated long-distance jumps of the pine processionary moth in Europe. *Biological Invasions, 14*(8), 1557-1569. doi:10.1007/s10530-011-9979-9

Roberts, M., Andreasen, V., Lloyd, A., & Pellis, L. (2015). Nine challenges for deterministic epidemic models. *Epidemics, 10*, 49-53. doi:https://doi.org/10.1016/j.epidem.2014.09.006

Roggero, P., Masenga, V., & Tavella, L. (2002). Field isolates of tomato spotted wilt virus overcoming resistance in pepper and their spread to other hosts in Italy. *Plant Disease, 86*(9), 950-954. doi:10.1094/pdis.2002.86.9.950

Roos, J., Hopkins, R., Kvarnheden, A., & Dixelius, C. (2011). The impact of global warming on plant diseases and insect vectors in Sweden. *European Journal of Plant Pathology, 129*(1), 9-19. doi:10.1007/s10658-010-9692-z

Rossi, V., Ravanetti, A., Pattori, E., & Giosuè, S. (2001). Influence of temperature and humidity on the infection of wheat spikes by some fungi causing Fusarium Head Blight. *Journal of Plant Pathology, 83*(3), 189-198.

Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., & Plümer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture, 74*(1), 91-99. doi:https://doi.org/10.1016/j.compag.2010.06.009

Russell, J. C., Binnie, H. R., Oh, J., Anderson, D. P., & Samaniego-Herrera, A. (2017). Optimizing confirmation of invasive species eradication with rapid eradication assessment. *Journal of Applied Ecology, 54*(1), 160-169. doi:10.1111/1365-2664.12753

Sacristán, S., & García-Arenal, F. (2008). The evolution of virulence and pathogenicity in plant pathogen populations. *Molecular Plant Pathology, 9*(3), 369-384. doi:10.1111/j.1364-3703.2007.00460.x

Salvalaggio, A. E., López Lambertini, P. M., Cendoya, G., & Huarte, M. A. (2017). Temporal and spatial dynamics of Tomato spotted wilt virus and its vector in a potato crop in Argentina. *Annals of Applied Biology, 171*(1), 5-14. doi:10.1111/aab.12357

Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture, 72*(1), 1-13. doi:https://doi.org/10.1016/j.compag.2010.02.007

Schaad, N. W., & Frederick, R. D. (2002). Real-time PCR and its application for rapid plant disease diagnostics. *Canadian Journal of Plant Pathology, 24*(3), 250-258. doi:10.1080/07060660209507006

Scholthof, K.-B. G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., Hemenway, C. & Foster, G. D. (2011). Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology, 12*(9), 938-954. doi:https://doi.org/10.1111/j.1364-3703.2011.00752.x

Severns, P. M., Sackett, K. E., Farber, D. H., & Mundt, C. C. (2019). Consequences of long-distance dispersal for epidemic spread: Patterns, scaling, and mitigation. *Plant Disease, 103*(2), 177-191. doi:10.1094/pdis-03-18-0505-fe

Sibylle, R., & Jochen, U. (2015). The oak processionary moth: a new health hazard? *British Journal of General Practice, 65*(637), 435. https://doi.org/10.3399/bjgp15X686341

Shevliakova, E., Stouffer, R. J., Malyshev, S., Krasting, J. P., Hurtt, G. C., & Pacala, S. W. (2013). Historical warming reduced due to enhanced land carbon uptake. *Proceedings of the National Academy of Sciences, 110*(42), 16730-16735. doi:10.1073/pnas.1314047110

Shiferaw, B., Prasanna, B. M., Hellin, J., & Bänziger, M. (2011). Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Security, 3*(3), 307. doi:10.1007/s12571-011-0140-5

Speer, C. A., Scott, M. C., Bannantine John, P., Waters, W. R., Mori, Y., Whitlock Robert, H., & Eda, S. (2006). A novel enzyme-linked immunosorbent assay for diagnosis of mycobacterium avium subsp. paratuberculosis infections (Johne's Disease) in cattle. *Clinical and Vaccine Immunology, 13*(5), 535-540. doi:10.1128/CVI.13.5.535-540.2006

Spence, N. (2020). Implementation of the GB plant health and biosecurity strategy 2014–2019 with foresight on a new strategy for 2020. *Outlook on Agriculture, 49*(1), 5-12. doi:10.1177/0030727020906831

Spence, N., Hill, L., & Morris, J. (2020). How the global threat of pests and diseases impacts plants, people, and the planet. *Plants, People, Planet, 2*(1), 5-13. doi:10.1002/ppp3.10088

Stack, J.P. (2008). Challenges to Crop Biosecurity. In: Gullino, M.L., Fletcher, J., Gamliel, A., Stack, J.P. *(eds) Crop Biosecurity.* NATO Science for Peace and Security Series C: Environmental Security. Springer, Dordrecht. doi: https://doi.org/10.1007/978-1-4020-8476-8_3

Somasekhar, N. & Prasad, J.S. (2011) Plant-nematode interactions: Consequences of climate change. *Crop stress and its management: Perspectives and strategies.* 547-564. Springer, Dordrecht. doi: https://doi.org/10.1007/978-94-007-2220-0_17

Stakman, E.C. & Harrar, J.G. (1957) *Principles of Plant Pathology* Ronald Press, New York.

Stagg, B. C., & Dillon, J. (2022). Plant awareness is linked to plant relevance: A review of educational and ethnobiological literature (1998–2020). *Plants, People, Planet, 4*(6), 579-592. doi:10.1002/ppp3.10323

Smith, S.D., Huxman, T.E., Zitzer, S.F., Charlet, T.N., Housman, D.C., Coleman, J.S., Fenstermaker, L.K., Seemann, J.R. & Nowak, R.S. (2000). Elevated CO_2 increases

productivity and invasive species success in an arid ecosystem. *Letters to Nature 408*(6808), 79-82. doi: https://doi.org/10.1038/35040544

Sundström, J. F., Albihn, A., Boqvist, S., Ljungvall, K., Marstorp, H., Martiin, C., Nyberg, K., Vågsholm, I., Yuen, J., Magnusson, U. (2014). Future threats to agricultural food production posed by environmental degradation, climate change, and animal and plant diseases – a risk analysis in three economic and climate settings. *Food Security, 6*(2), 201-215. doi:10.1007/s12571-014-0331-y

Sutherland, W. J., Bennett, C., Brotherton, P. N. M., Butterworth, H. M., Clout, M. N., Côté, I. M., . . . Thornton, A. (2023). A global biological conservation horizon scan of issues for 2023. *Trends in Ecology & Evolution*, *38*(1), 96-107. Doi: https://doi.org/10.1016/j.tree.2022.10.005

Thomas, M. B. (2020). Epidemics on the move: Climate change and infectious disease. *PLoS Biology*, 18(11), e3001013. https://doi.org/10.1371/journal.pbio.3001013

Thompson, S., Levin, S., & Rodriguez-Iturbe, I. (2013). Linking plant disease risk and precipitation drivers: A dynamical systems framework. *The American Naturalist, 181*(1), E1-E16. doi:10.1086/668572

Tortosa, A., Giffard, B., Sirami, C., Larrieu, L., Ladet, S., & Vialatte, A. (2023). Increasing landscape heterogeneity as a win–win solution to manage trade-offs in biological control of crop and woodland pests. *Scientific Reports*, *13*(1). Doi: https://doi.org/10.1038/s41598-023-40473-2

Tourinho, L., & Vale, M. M. (2023). Choosing among correlative, mechanistic, and hybrid models of species' niche and distribution. *Integrative Zoology*, *18*(1), 93-109. https://doi.org/https://doi.org/10.1111/1749-4877.12618

Truyens, S., Weyens, N., Cuypers, A., & Vangronsveld, J. (2015). Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environmental Microbiology Reports, 7*(1), 40-50. doi:10.1111/1758-2229.12181

Turner, M. G. (2005). Landscape Ecology: What Is the state of the science? *Annual review of ecology, evolution, and systematics, 36*(1), 319-344. doi:10.1146/annurev.ecolsys.36.102003.152614

Turner-Skoff, J. B., & Cavender, N. (2019). The benefits of trees for livable and sustainable communities. *Plants, People, Planet, 1*(4), 323-335. doi:https://doi.org/10.1002/ppp3.39

Uekoetter, F. (2011). The magic of one: Reflections on the pathologies of monoculture. *RCC Perspectives(2),* 1-20. doi.org/10.5282/rcc/5584.

Van Dijk, M., Morley, T., Rau, M. L., & Saghai, Y. (2021). A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nature Food, 2*(7), 494-501. doi:10.1038/s43016-021-00322-9

Van Maanen, A., & Xu, X. M. (2003). Modeling plant disease epidemics. *European Journal of Plant Pathology, 109*(7), 669-682. doi:10.1023/a:1026018005613

van der Plank, J. E. (1963). Plant Diseases: Epidemics and Control. In: New York and London (1st Ed): Academic Press. https://doi.org/10.1016/B978-0-12-711450-7.50004-4.

Velásquez, A. C., Castroverde, C. D. M., & He, S. Y. (2018). Plant-Pathogen Warfare under Changing Climate Conditions. *Current biology : CB, 28*(10), 619–634. doi: https://doi.org/10.1016/j.cub.2018.03.054

Volkovitsh, M. G., Bieńkowski, A. O., & Orlova-Bienkowskaja, M. J. (2021). Emerald Ash Borer approaches the borders of the European Union and Kazakhstan and is confirmed to infest European Ash. *Forests, 12*(6). doi:10.3390/f12060691

Wadkin, L. E., Branson, J., Hoppit, A., Parker, N. G., Golightly, A., & Baggaley, A. W. (2022). Inference for epidemic models with time-varying infection rates: Tracking the dynamics of oak processionary moth in the UK. *Ecology and Evolution, 12*(5), e8871. https://doi.org/https://doi.org/10.1002/ece3.8871

Wang, B., Tian, C., & Sun, J. (2019). Effects of landscape complexity and stand factors on arthropod communities in poplar forests. *Ecology and Evolution, 9*(12), 7143-7156. Doi: https://doi.org/10.1002/ece3.5285

Welsh, M. J., Turner, J. A., Epanchin-Niell, R. S., Monge, J. J., Soliman, T., Robinson, A. P., Kean, J.M, Phillips, C., Stringer, L.D., Vereijssen, J., Liebhold, A.M, Kompas, T., Ormsby, M. & Brockerhoff, E. G. (2021). Approaches for estimating benefits and costs of interventions in plant biosecurity across invasion phases. *Ecological Applications, 31*(5). https://doi.org/10.1002/eap.2319

Woodward, S., & Boa, E. (2013). Ash dieback in the UK: a wake-up call. *Molecular Plant Pathology, 14*(9), 856-860. doi:https://doi.org/10.1111/mpp.12084

Yang, X.B. (2006). Framework Development in Plant Disease Risk Assessment and its Application. *Eur J Plant Pathol 115*, 25–34. doi: https://doi.org/10.1007/s10658-005-4513-5

Yang, X & Guo, T. (2017) Machine learning in plant disease research. *Biomedical Research Journal 34* doi:10.18088/ejbmr.3.1.2017.pp6-9

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A.H., Bamford, A.J., Ban, S., Barbosa, A.M., Dormann, C.F., Elith, J., Embling, C.B., Ervin, G.N., Fischer, R., Gould, S., Graf, R.F., Gregr, E.J., Halpin, P.N., Heikkinen, R.K… Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology & Evolution*, *33*(10), 790-802. https://doi.org/10.1016/j.tree.2018.08.001

Yeh, K., Monagin, C., Fletcher, J., 2017. Promoting Scientific Transparency to Facilitate the Safe and Open International Exchange of Biological Materials and Electronic Data. *Tropical Medicine and Infectious Disease 2*, 57. https://doi.org/10.3390/tropicalmed2040057

Zabel, F., Delzeit, R., Schneider, J. M., Seppelt, R., Mauser, W., & Václavík, T. (2019). Global impacts of future cropland expansion and intensification on agricultural markets and biodiversity. *Nature communications*, *10*(1), 2844. doi: https://doi.org/10.1038/s41467-019-10775-z

Zadoks, J.C. (2001) Plant disease epidemiology in the twentieth century: A picture by means of selected controversies. *Plant Disease 85*(8). 808-816. doi: 10.1094/PDIS.2001.85.8.808.

Zaffaroni, M., Cunniffe, N.J., Bevacqua, D., 2020. An ecophysiological model of plant–pest interactions: the role of nutrient and water availability. *Journal of The Royal Society Interface 17*, 20200356. doi: https://doi.org/10.1098/rsif.2020.0356

Zeilinger, A.R., Rapacciuolo, G., Turek, D., Oboyski, P.T., Almeida, R.P.P. and Roderick, G.K. (2017). Museum specimen data reveal emergence of a plant disease may be linked to increases in the insect vector population. *Ecol Appl*, *27*. 1827-1837. doi: https://doi.org/10.1002/eap.1569

Zhang, J., Huang, Y., Pu, R., Gonzalez-Moreno, P., Yuan, L., Wu, K., & Huang, W. (2019). Monitoring plant diseases and pests through remote sensing technology: A review. *Computers and Electronics in Agriculture, 165*, 104943. doi:https://doi.org/10.1016/j.compag.2019.104943

9. Appendices

9.1. Chapter Three

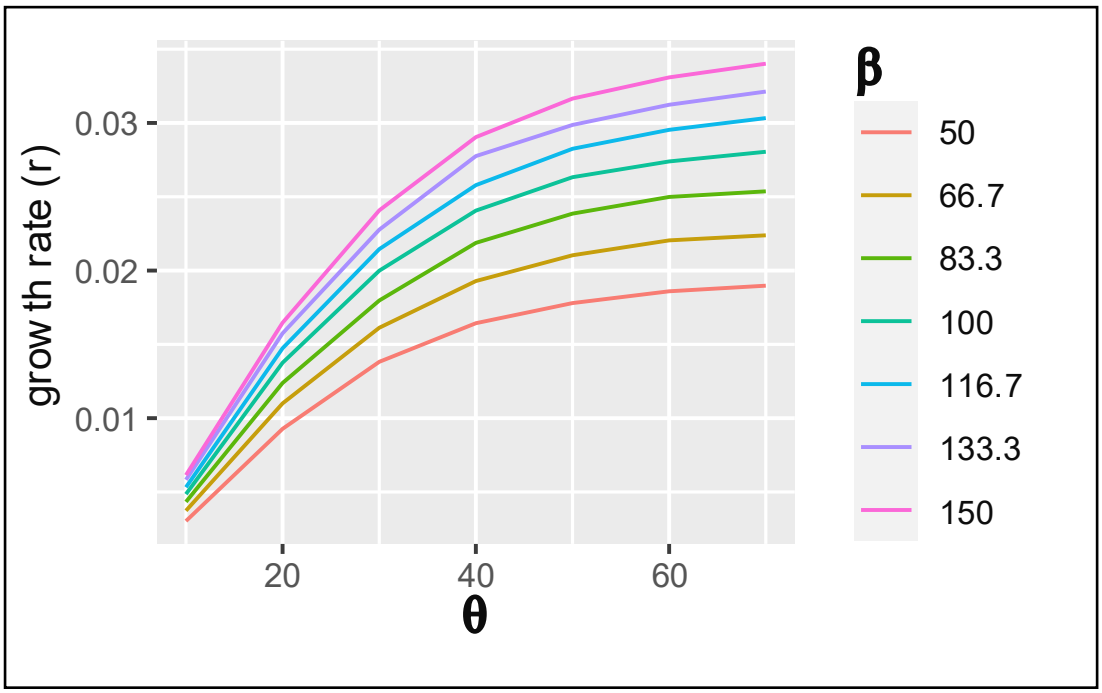


Figure 9.1. The growth rate values for the epidemic parameter combinations. Parameter values of transmission coefficient β and dispersal distance θ are shown (β = 10 to 70; θ=10 to 70). 2000 iterations per parameter combination were performed.

9.2. Chapter Four

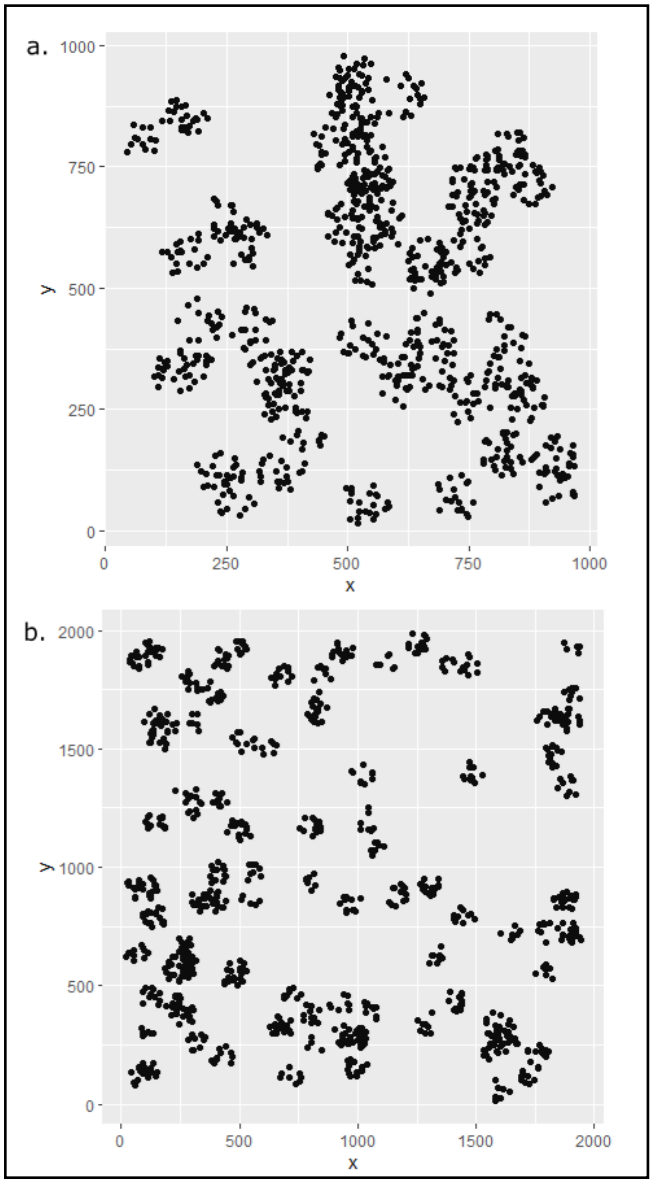


Figure 9.2. The distribution of hosts on a. 1 km^2 and b. 2 km^2. Cluster parameters are fixed on both landscapes. The scale on the axes is in meters.

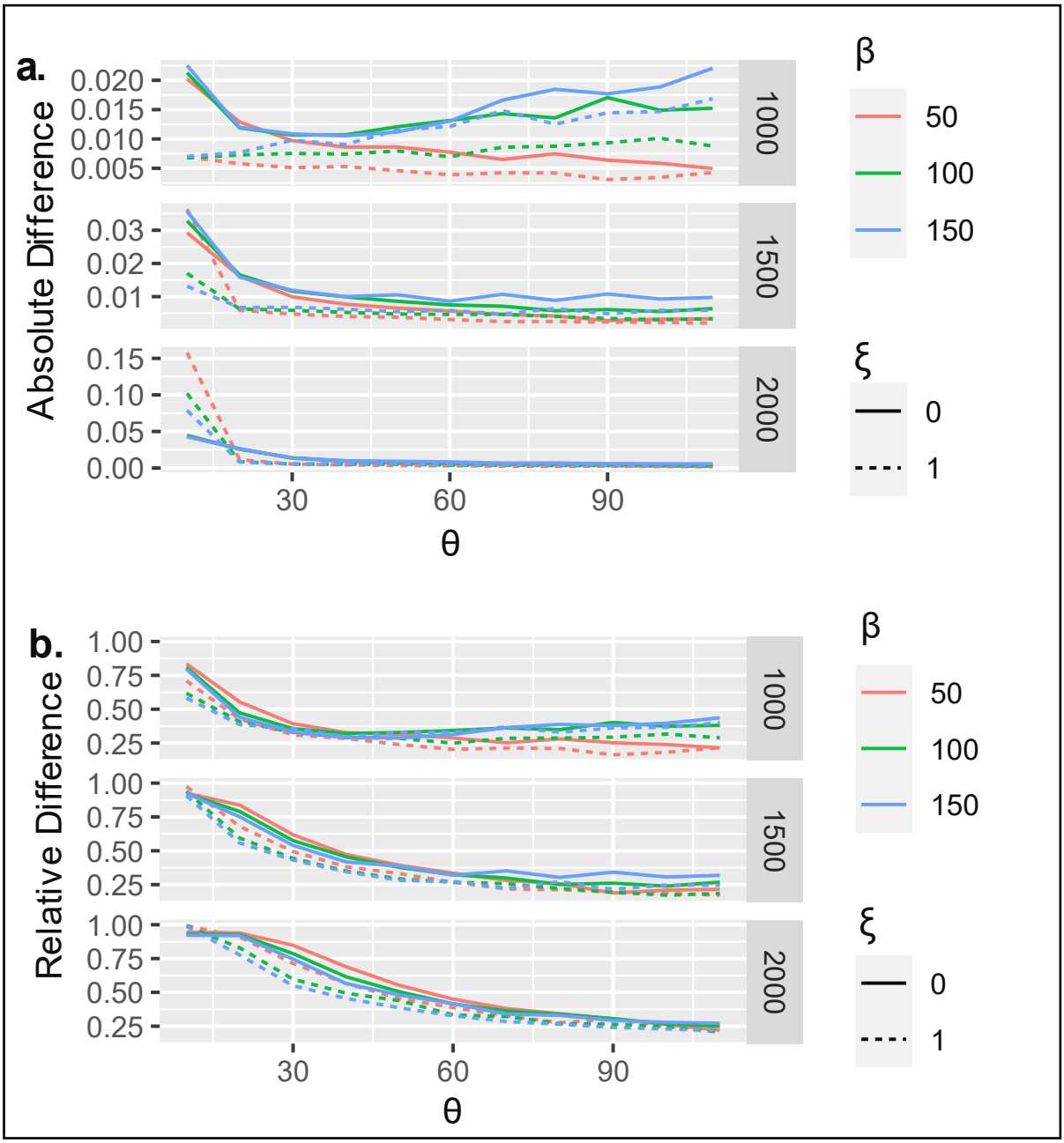


Figure 9.3. a. The absolute difference and b. The relative difference across a range of parameters for scales of 1000 m^2, 1500 m^2 and 2000 m^2. Note that the scale on the y axes of the absolute difference are unique to each scale within the plot.