WIDESPREAD TRUNCATION OF PLANT DISEASE GRADIENTS LEADS TO UNDER-PREDICTIONS OF DISEASE DISPERSAL AND SEVERITY

by

EMILY M. SYKES

(Under the Direction of Paul M. Severns)

ABSTRACT

Epidemiological models are important for understanding both the underpinnings of disease spread and projecting epidemic magnitude. Disease gradients are difficult to accurately measure in the field – particularly the distance disease should be tracked from a source to sufficiently represent the shape of the dispersal gradient and accurately reflect the expanse to which disease is dispersed. Truncated dispersal gradients could severely underestimate the extent and magnitude of epidemics, leading to deceptive projections and misinformed management recommendations. I reviewed dispersal gradient figures in journal articles published from 1950–2020 and found that truncated gradients were common (71.2% of studies) and occurred across all surveyed dispersal vector categories. Wheat stripe rust disease spread simulations confirmed that dispersal kernel truncation generates substantially misleading projections of disease when host plants are continuous (agricultural fields) or patchily distributed (natural ecosystems). Future studies of plant disease spread should attempt to more accurately represent the dispersal gradient.

INDEX WORDS: plant disease epidemiology, dispersal gradient, truncated, ecology

WIDESPREAD TRUNCATION OF PLANT DISEASE GRADIENTS LEADS TO UNDER-PREDICTIONS OF DISEASE DISPERSAL AND SEVERITY

by

EMILY M. SYKES

B.S., Oregon State University, 2015

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2022

WIDESPREAD TRUNCATION OF PLANT DISEASE GRADIENTS LEADS TO UNDER-PREDICTIONS OF DISEASE DISPERSAL AND SEVERITY

by

EMILY M. SYKES

Major Professor: Paul M. Severns

Committee: Marin T. Brewer
Philip M. Brannen

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2022

ACKNOWLEDGEMENTS

I would like to thank the Department of Plant Pathology for providing funding for this research. I would also like to thank my major advisor, Dr. Paul Severns, for his support and guidance throughout this project and for his mentorship during my time at the University of Georgia. I also need to extend my gratitude to the members of my committee, Dr. Marin Brewer and Dr. Philip Brannen, for their invaluable advice and comments. Finally, I wish to give a very big thank you to my friends and family for supporting me, whether in Athens or from across the country.

It takes a village to support a graduate student, and it is due to all of you that I was able to complete a project I can be proud of.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

One of the fundamental challenges in plant pathology is understanding the spread of plant diseases in time and space. Understanding the spatiotemporal spread of epidemics has important implications for predicting and controlling disease, because fast-developing and far-reaching epidemics are more difficult to combat than slow-developing, local ones. Epidemiological models based on field-measured disease data can be used to project how much disease may occur and where it occurs as an epidemic develops. Epidemiological models integrate time (infectious and latent periods), dispersal (dispersal kernel), infection probability (host susceptibility and population size), and demographic information (initial pathogen infection level, reproductive rate). As with many predictive statistical models, their projections improve as parameters more closely match the biology of the system being represented; for example, weather predictions have improved drastically over the last two decades as the quantity and quality of phenomenological data have improved (NOAA 2016).

Accurately accounting for dispersal is critical to understanding the spatial patterns of disease spread; if the parameters statistically representing disease dispersal are not measured or estimated correctly, model outputs may not provide realistic projections of disease occurrence and abundance over space and time. A dispersal gradient, or the decline of pathogen propagules from a source (Gregory 1968), measures a pathogen's dispersal potential. These gradients are synonymous with dispersal kernels, the probability density functions (PDFs) that describe the distribution of a population relative to their source (Nathan et al. 2012). Dispersal gradients are

extrapolated from both observational and experimental data (Nathan et al. 2012). In agricultural systems, tracking pathogen movement begins by identifying an outbreak area or establishing disease within a field (Sackett and Mundt 2005; Severns et al. 2019). Measurements of a pathogen's effective dispersal, often captured as disease severity, are then taken at set sampling distances in straight lines from the focus (a spatially distinct cluster of infected individuals) and may be repeated over time. These disease measurements are then graphed on a plot of disease severity (y-axis) over distance (x-axis) and a function is fit to the data points, ultimately forming the dispersal gradient that is used to model disease spread (Figure 1.1). Seemingly subtle changes in the shape of the dispersal gradient can have a large effect on the predicted pattern of disease spread (Kot et al. 1996; Clark et al. 2001; Hastings et al. 2005). For example, not accounting for dispersal events in the tail of a function could result in a model that underestimates epidemic velocity and magnitude, so accurate measurements and accurate equation fitting are critical.

Plant pathogens can be dispersed in a variety of ways – cedar-apple rust is moved across the landscape primarily via wind, eriophyid mites carry Rose rosette virus through nurseries, and root knot nematodes are spread through fields by tools and tractors (Agrios 2014). These dispersal types encompass complex biological processes, but they can be unified by one crucial element – long distance dispersal (LDD). While there is no strict definition for what constitutes long distance dispersal, it is generally defined by extreme dispersal events (< 1% of all propagules) or by dispersal beyond distance thresholds given a species propensity for dispersal in specific environments (Nathan et al. 2008b). While LDD events are infrequent, they are integral to population connectivity, colonization across landscapes, and the persistence of organisms despite local extinction or control measures (Traketenbrot et al. 2005; Baguette and Schtickzelle 2006; Ronce 2007; Schloss et al. 2012). LDD events are characterized by dispersal gradients

with “fat tails” – a leptokurtic probability distribution with considerably longer tails than a normal distribution, resulting in a greater probability of rare or extreme outcomes (Figure 1.1).

Some fat-tailed dispersal gradients are not exponentially bound and are often described using various inverse power law functions (Gisiger 2001; Mundt et al. 2009; Tildesley et al. 2012; Mundt et al. 2013; Ojiambo et al. 2017) but other fat-tailed dispersal gradients are described by functions that are bounded by an exponential (Kot et al. 1996). The degree of kurtosis in the long tails of the disease gradient is associated with the rate of organism spread. In the case of the “shorter” tailed distributions (those that are bound by an exponential), organisms reach a maximum, constant rate of spread as long as host availability and space are available. However, when the tails are sufficiently long (those that are not bound by an exponential function), the organisms spread with increasing velocity, as long as host availability and space are available (Ferrandino 1993; Kot et al. 1996; Cowger et al. 2005; Mundt et al. 2009a, b; Pybus et al. 2012). Historically, dispersal models predicted spatiotemporal movement to occur at a constant velocity – a “traveling wave”. Traveling waves are an emergent characteristic based on random walk and diffusion models and assume that the decline in propagules over distance is described by an exponentially bound function (Mollison 1977). Pathogens that spread exclusively through direct or close contact between hosts are likely to be well-represented by negative exponential functions and move at a constant velocity as a traveling wave (Grenfell et al. 2001; Cummings et al. 2004). While traveling wave models have been applied to plant disease dispersal (Minogue and Fry 1983a, b; Heesterbeek and Zadoks 1987; van den Bosch et al. 1988a,b,c; Viruega et al. 2013; Prussin et al. 2015), use of these models for wind dispersed pathogens have been called into question. Studies have demonstrated that thin-tailed functions, which produce invasion fronts with a constant rate of spread, cannot explain the distance traveled

and dispersal rate of organisms displaying patterns of long-distance dispersal (Mollison 1991; Pybus et al. 2012; Nathan et al. 2012), such as wind dispersed plant pathogens, leading to underestimation of a species' projected range and movement capabilities.

Observational and theoretical studies have demonstrated that the dispersal gradients of wind dispersed pathogens consistently exceed the limits of exponential functions (Ferrandino 1993; Zadocks and van den Bosch 1994; Scherm 1996; Nathan et al. 2012, Ojiambo et al. 2017). Inverse power law functions have been used since the beginning of plant epidemiology to describe various LDD pathogens (Gregory 1968) and can account for the "fat-tailed" disease gradient generated by them (Brown and Hovmoller 2002). Notably, these power law functions can generate epidemics that increase in velocity as disease spreads (Ferrandino 1993; Cowger et al. 2005; Mundt et al. 2009a, b; Pybus et al. 2012) – a "dispersive wave". Epidemics with increasing velocities can quickly overwhelm on-the-ground disease management attempts based on diffusion (traveling wave) models, infecting hosts both more quickly and further away than predicted (Keeling et al. 2001; Chis Ster and Ferguson 2007; Peterson et al. 2015).

Disease gradients can be challenging to measure in the field, particularly how far from a source disease should be tracked to sufficiently represent the shape and tail of the dispersal gradient. The large amounts of land required, strong initial disease sources, and the risk of contamination by outside inoculum can complicate attempts to measure the low levels of disease in the dispersal tail where occasional presences and far more absences of disease are recorded (Sackett and Mundt 2005). Wind vectored plant diseases are known to travel great distances, potentially traveling kilometers away from their source (Farber et al. 2019); this scale of movement cannot be captured by most empirical plant epidemiology studies, which are often constrained by field size and low levels of disease that are difficult to detect. For novel diseases,

there may be little to no understanding of pathogen movement, making it nearly impossible to set up a field study that completely accounts for all dispersal. It is easy to artificially truncate disease gradients because spatially limited or data deficient studies present disease dispersed relatively close to the source, which favors Gaussian or negative exponential functions. If the researcher uses a thin-tailed function (Gaussian or exponential) to describe the dispersal gradient, they can only model disease spread that occurs at a constant rate following an ephemeral period of acceleration (Frantzen and van den Bosch 2000). If the disease being studied truly abides by these thin-tailed functions, then a diffusion rate is a biologically accurate representation. However, studies have demonstrated that observed dispersal velocity and disease magnitude cannot be explained by these simple diffusion models (Kot et al. 1996; Nathan et al. 2003; Tackenberg 2003; Lietch et al. 2021). Even when disease dispersal data is gathered at a sufficient distance from the source, an ill-fitting function predicting a constant rate of spread may also underestimate epidemic magnitude and how far disease is likely to spread over time (Sackett and Mundt 2005); this phenomenon is well-illustrated in Figure 1.2, which shows how the improper function does not sufficiently describe dispersal close to the source, in the middle, or further away in the tails.

Building a model with a traveling wave is relatively straightforward as the dispersal rate is constant. However, because velocity is always increasing and the rate of spread is impacted by other variables (time, pathogen reproductive rates, host properties, etc.), dispersive waves are not conveniently modeled. Simulations are the only way to project the effects of dispersive waves because their increasing rate of spread cannot be accommodated by linear models without underestimating the arrival time and spatial expanse of organism spread.

Research Goals

A well-fitting function is key to understanding and projecting the patterns of plant disease spread, just as it is for understanding the patterns of any dispersing or invading organism (Nathan et al. 2012). However, while the issue of dispersal kernel truncation has been explored for plants and animals (Nathan et al. 2012; Rogers et al. 2019), it is unclear how often disease gradient truncation is occurring in the field of plant pathology. Chapter 2 provides a review of the plant epidemiological literature to answer the following questions: (1) How common are truncated plant disease gradients in the published plant pathology literature, and if truncated disease gradients occur, were they as common in the past as they are presently, and (2), are the major forms of propagule dispersal (wind, splash, nematode, insects) for plant pathogens represented by accurately fit dispersal gradients and, if not, which modes of dispersal are under-represented? In Chapter 3, I examine the effect truncation has on models of disease. I use wheat stripe rust, one of the best-described LDD plant pathogen systems, to parameterize a model and simulate the differences in disease incidence and abundance between exponential (truncated) and fat-tailed (confirmed system) dispersal gradients. I additionally evaluate the effect of host distribution on epidemic extent and magnitude by simulating across a series of susceptible and resistant host landscapes, ranging from 100% to 5% host availability. In Chapter 4, I synthesize the results from Chapters 2 and 3, culminating in research recommendations for future disease spread studies.

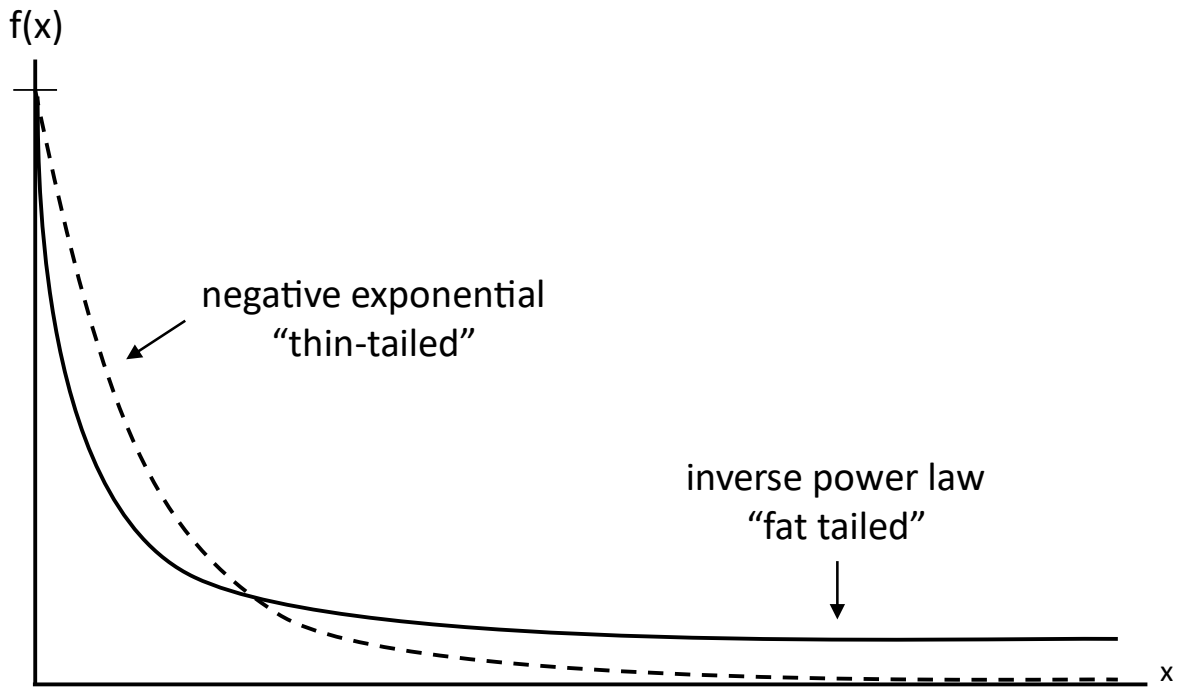


Figure 1.1. Difference in shape between negative exponential (dashed line) and power law (solid line) functions. LDD events are characterized by dispersal gradients with "fat tails" – a leptokurtic probability distribution with thicker tails than a normal distribution, resulting in a greater probability of rare or extreme outcomes. Fat -tailed dispersal gradients are not exponentially bound and are often described using various inverse power law functions. While these two functions may appear similar, the patterns of disease spread described by them are drastically different.

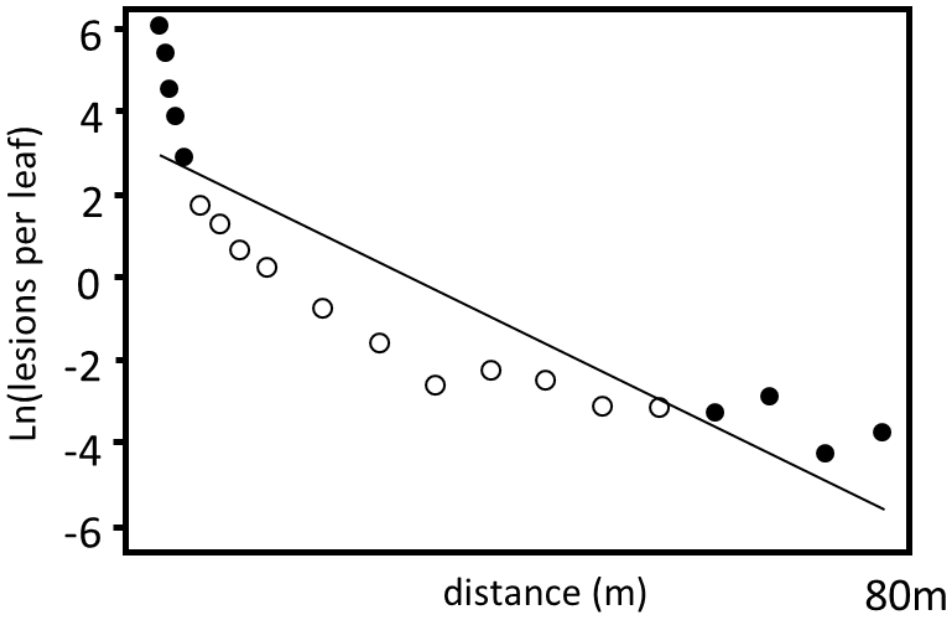


Figure 1.2. Field data of WSR lesions per leaf fitted to a natural log transformed exponential function. Demonstrates the improper fit of the exponential function to WSR data, which underpredicts disease levels near the source and in the tails (filled circles), and overpredicts disease levels at intermediate distances (open circles). Adapted from Sackett and Mundt (2005).

Literature Cited

Agrios, G. N. 2005. *Plant Pathology*. 5th ed. Amsterdam: Elsevier Academic Press.

Baguette, M., and Schtickzelle, N. 2006. Negative relationship between dispersal distance and demography in butterfly metapopulations. Ecology. 87:648–654.

Brown, J. K., and Hovmoller, M. S. 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. Science. 297:537–541.

Chis Ster, I., and Ferguson, N. M. 2007. Transmission parameters of the 2001 foot and mouth epidemic in Great Britain. PLoS ONE. 2:e502.

Clark, J. S., Lewis, M., and Horvath, L. 2001. Invasion by extremes: Population spread with variation in dispersal and reproduction. The American Naturalist. 157:537–554.

Cowger, C., Wallace, L. D., and Mundt, C. C. 2005. Velocity of spread of wheat stripe rust epidemics. Phytopathology. 95:972–982.

Cummings, D. A. T., Irizarry, R. A., Huang, N. E., Endy, T. P., Nisalak, A., Ungchusak, K., et al. 2004. Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature. 427:344–347.

Farber, D., De Leenheer, P., and Mundt, C. C. 2019. Dispersal kernels may be scalable: implications from a plant pathogen. Journal of Biogeography 46:2042-2055.

Ferrandino, F. J. 1993. Dispersive epidemic waves: I. Focus expansion within a linear planting. Phytopathology. 83:795–802.

Frantzen, J., and van den Bosch, F. 2000. Spread of organisms: can travelling and dispersive waves be distinguished? Basic and Applied Ecology. 1:83–91.

Fryxell, J. M. 2008. Predictive modeling of patch use by terrestrial herbivores. In *Resource Ecology*, Dordrecht: Springer, p. 105–123.

Gisiger, T. 2001. Scale invariance in biology: coincidence or footprint of a universal mechanism? Biological Reviews. 76:161–209.

Gregory, P. H. 1968. Interpreting plant disease dispersal gradients. Annual Review of Phytopathology. 6:161–209.

Grenfell, B. T., Bjornstad, O. N., and Knappey, J. 2001. Travelling waves and spatial hierarchies in measles epidemics. Nature. 414:716–723.

Hastings, A., Cuddington, K., Davies, K. F., Dugaw, C. J., Elmendorf, S., Freestone, A., et al. 2005. The spatial spread of invasions: new developments in theory and evidence. Ecology Letters. 8:91–101.

Heesterbeek, J. A. P., and Zadoks, J. C. 1987. Modelling pandemics of quarantine pests and diseases: problems and perspectives. Crop Protection. 6:211–221.

Isard, S. A., and Gage, S. H. 2001. *Flow of Life in the Atmosphere: An Airscape Approach to Understanding Invasive Organisms*. East Lansing, MI: Michigan State University Press.

Keeling, M. J., Woolhouse, M. E. J., Shaw, D. J., Matthews, L., Chase-Topping, M., Haydon, D. T., et al. 2001. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science. 294:813–817.

Kot, M., Lewis, M. A., and van den Driessche, P. 1996. Dispersal data and the spread of invading organisms. Ecology. 77:2027–2042.

Lietch, K. J. 2021. The long-distance flight behavior of *Drosophila* supports an agent-based model for wind-assisted dispersal in insects. Proceedings of the National Academy of Sciences. 118:1–9.

Minogue, K. P., and Fry, W. E. 1983a. Models for the spread of disease: model description. Phytopathology. 73:1168–1173.

Minogue, K. P., and Fry, W. E. 1983b. Models for the spread of plant disease: some experimental results. Phytopathology. 73:1173–1176.

Mollison, D. 1977. Spatial contact models for ecological and epidemic spread. Journal of the Royal Statistical Society: Series B. 39:283–313.

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

Mundt, C. C., Sackett, K. E., Wallace, L. D., Cowger, C., and Dudley, J. P. 2009b. Aerial dispersal and multiple-scale spread of epidemic disease. EcoHealth. 6:546–552.

Mundt, C. C., Wallace, L. D., Allen, T. W., Hollier, C. A., Kemerait, R. C., and Sikora, E. J. 2013. Initial epidemic area is strongly associated with the yearly extent of soybean rust spread in North America. Biological Invasions. 15:1431–1438.

Nathan, R., Perry, G., Cronin, J. T., Strand, A. E., and Cain, M. L. 2003. Methods for estimating long-distance dispersal. Oikos. 103:261–273.

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., et al. 2008a. A movement ecology paradigm for unifying organismal movement research. Proceedings of the National Academy of Sciences. 105:19052–19059.

Nathan, R., Schurr, F. M., Spiegel, O., Steinitz, O., Trakhtenbrot, A., and Tsoar, A. 2008b. Mechanisms of long-distance seed dispersal. Trends in Ecology and Evolution. 23:638–647.

Nathan, R., Klein, E., Robledo-Arnuncio, J. J., and Revilla, E. 2012. Dispersal kernels: review. In *Dispersal Ecology and Evolution* (J. Clobert, J. M. Bullock, T. G. Benton, and M. Baguette, Eds.), Oxford, UK: Oxford University Press.

NOAA. Improving weather forecasts. 2016. National Ocean and Atmospheric Administration. Available at: https://www.noaa.gov/explainers/improving-weather-forecasts [Accessed July 25, 2022].

Ojiambo, P. S., Gent, D. H., Mehra, L. K., Christie, D., and Margarey, R. 2017. Focus expansion and stability of the spread parameter estimate of the power law model for dispersal gradients. PeerJ. 5:e3465.

Peterson, E. K., Hansen, E. M., and Kanaskie, A. 2015. Temporal epidemiology of sudden oak death in Oregon. Phytopathology. 105:937–946.

Prussin, A. J., Marr, L. C., Schmale, D. G., Stoll, R., and Ross. 2015. Experimental validation of long-distance transport model for plant pathogens: application to *Fusarium graminearum*. Agricultural and Forest Meteorology. 203:118–130.

Pybus, O. G., Suchard, M. A., Lemey, P., Bernardin, F. J., Rambaut, A., Crawford, F. W., et al. 2012. Unifying the spatial epidemiology and molecular evolution of emerging epidemics. Proceedings of the National Academy of Sciences. 109:15066–15071.

Rogers, H. S., Beckman, N. G., Hartig, F., Johnson, J. S., Pufal, G., Shea, K., et al. 2019. The total dispersal kernel: a review and future directions. AoB PLANTS. 11:plz042.

Ronce, O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. Annual Review of Ecology, Evolution, and Systematics. 38:231–253.

Sackett, K. E., and Mundt, C. C. 2005. Primary disease gradients of wheat stripe rust in large field plots. Phytopathology. 95:983–991.

Scherm, H. 1996. On the velocity of epidemic waves in model plant disease epidemics. Ecological Modeling. 87:217–222.

Schloss, C. A., Nuñez, T. A., and Lawler, J. 2012. Dispersal will limit ability of mammals to track climate change in the Western Hemisphere. Proceedings of the National Academy of Sciences. 109:8606–8611.

Tackenberg, O. 2003. Modeling long-distance dispersal of plant diaspores by wind. Ecological Monographs. 72:173–189.

Tildesley, M. J., Smith, G., and Keeling, M. J. 2012. Modeling the spread and control of foot-and-mouth disease in Pennsylvania following its discovery and options for control. Preventive Veterinary Medicine. 104:224–239.

Trakhtenbrot, A., Nathan, R., Perry, G., and Richardson, D. M. 2005. The importance of long-distance dispersal in biodiversity conservation. Diversity and Distributions. 11:173–181.

van den Bosch, F., Zadoks, J. C., and Metz, J. A. J. 1988a. Focus expansion in plant disease. I: The constant rate of focus expansion. Phytopathology. 78:54–58.

van den Bosch, F., Zadoks, J. C., and Metz, J. A. J. 1988b. Focus expansion in plant disease. II. Realistic parameter-sparse models. Phytopathology. 78:59–64.

van den Bosch, F., Frinking, H. D., Metz, J. A. J., and Zadoks, J. C. 1988c. Focus expansion in plant disease. III: Two experimental examples. Phytopathology. 78:919–925.

Viruega, J. R., Moral, J., Roca, L. F., Navarro, N., and Trapero, A. 2013. *Spilocaea oleagina* in olive groves of southern Spain: survival, inoculum production, and dispersal. Plant Disease. 97:1549–1556.

Zadoks, J. C., and van den Bosch, F. 1994. On the spread of plant disease: A theory on foci. Annual Review of Phytopathology. 32:503–521.

CHAPTER 2

CRITICAL REVIEW OF PLANT PATHOGEN DISPERSAL GRADIENTS

Introduction

Statistical models are important for our understanding of the natural world (Pazhamala et al. 2020; Tredennick et al. 2021). Biologists rely on phenomenological information to create mechanistic and predictive models; however, performance is contingent on the accuracy of input data, and poorly parameterized models will not reflect reality. For models of organism occurrence, movement and distribution forecasting, the dispersal kernel is central to understanding and accounting for dispersal processes and spatial patterns, so it critical that this parameter is accurately statistically described (Nathan et al. 2012).

The implications of dispersal function truncation are substantial – many studies have demonstrated that thin-tailed dispersal kernels, which generate a constant rate of spread, cannot sufficiently explain either the observed distance traveled or observed dispersal rates of organisms characterized by long-distance dispersal (Mollison 1991; Pybus et al. 2012; Nathan et al. 2012), leading to underestimation of species range and movement capabilities. This shift in understanding about dispersal from travelling (constant velocity) to dispersive waves (accelerating velocity), was heralded by two landmark papers (Ferrandino 1993; Kot et al. 1996) which were among the first to demonstrate that observed population spread rates cannot be explained by simple exponentially bound dispersal gradients and are best fit by fat-tailed ones. These two papers represent a critical shift in dispersal ecology philosophy and mark the

beginning of a movement to improve the measurement of long-distance dispersal, critical to capturing the accelerating rate of invasions over space.

The first influential paper was written by Ferrandino (1993), a plant pathologist, who specifically tested whether the dispersal of aerial spores in turbulent conditions could be modeled by a traveling wave (the standing theory at the time). Ferrandino compared reported field data for bean rust (*Uromyces phaseoli*) and wheat and barley stem rust (*Puccinia recondita*) to several statistical models to determine whether the models effectively explained actual dispersal. In this work, Ferrandino noted that previous spatial models had ignored spore movement in the vertical direction and sought to capture the importance of uplift on spore deposition rate moving towards a more holistic dispersal kernel. He noted that while most spores were deposited close to the infection source, some spores were lifted out of the host canopy by wind turbulence and carried over great distances before being deposited far from the source. In his models, Ferrandino found that these turbulent dispersal events resulted in epidemics characterized by an increase in velocity with distance from the source: "One generation's turbulently dispersed propagules 'pioneer' new territory, and the succeeding generations' locally dispersed spores complete the process of colonization." (Ferrandino 1993).

The second influential study focused on biological invasions (a broad conceptual framework that encompasses plant diseases epidemics). Kot et al. 1996, modeled the spread of invading organisms – in this case, genetically unique *Drosophila* genotypes to evaluate which dispersal kernels may best describe their movement ecology. In their introduction, Kot et al. (1996) reference both dispersal ecology and plant pathology studies presenting dispersal kernels, and they found leptokurtic distributions of propagules from the source but note that these patterns were often ignored by the researchers. In these instances, the observed data extended

farther from the source than the selected statistical function characterizing the dispersal gradient. Instead, "…individuals are assumed to disperse in each direction with equal probability and with the dispersal distance *normally* (rather than leptokurtically) distributed." The authors had three major findings from their evaluation of the dispersal models against one another: 1) the speed of an invading population is remarkably sensitive to the shape of the dispersal gradient, especially in how far the tails of the distribution extend; 2) fat-tailed gradients are able to generate the accelerating fronts of biological invasions in the *Drosophila* data set; and 3) normal or diffusion (exponential family) dispersal gradients may "grossly underestimate" the rate of spread of an invading population compared to more realistic leptokurtic distributions which had been more accurately fit to empirically gathered data.

Inappropriately characterizing the dispersal gradient has a significant biological impact on either projecting where the organism would go or whether the conservation or intervention measures were compromised. In van Houtan et al. (2007), the authors tested the dispersal of tropical birds before and after habitat fragmentation. Previous studies suggested that tropical birds were poor dispersers and were therefore unable to travel between habitats in a fragmented landscape. After collecting dispersal data on the bird populations, the authors tested both thin- (Rayleigh and gamma) and fat-tailed functions (log-hyperbolic secant) and found the fat-tailed distribution best fit the data. The fat-tailed distribution also provided the best conceptual description of tropical bird movement in continuous and fragmented forests – while birds were less mobile within fragmented landscape, when they did cross gaps, they dispersed farther than they did before fragmentation occurred. Thin-tailed gradients were not able to replicate the rare, LDD (long-distance dispersal) events and if used in conservation modeling context, would under-predict the presence of tropical birds across the landscape. In their study of pollen

dispersal, Klein et al. (2006) estimated individual dispersal functions for transgenic oilseed rape. After allowing a point source of transgenic oilseed rape (homozygous transgene conferring oxynil herbicide resistance) to pollenate a gridded field of male sterile plants, the resulting seeds were sown in individual plots according to the previous grid location. Each plot was sprayed with oxynil herbicide, and the proportion of resistant versus susceptible seedlings per grid location were collected as the cross-pollination rate, providing a pollen dispersal gradient. The authors tested exponential-like functions and fat-tailed functions against the resulting pollen dispersal gradient and found that the exponential family of functions had the poorest fit, likely due to, "…an underestimate of the proportions of resistant seedlings at short and long distances from the central point." Fat-tailed functions, especially the power law functions, provided the best fit to the data. Thin-tailed gradients were unable to capture the full extent and amount of pollen dispersed in the study, and if used in a predictive model, would underestimate the dispersal capability of transgenic oilseed rape – a considerable error when attempting to study the ecological impacts of transgenic crops on both surrounding conventional crops and the environment.

A well-fitting dispersal kernel function is key to understanding and projecting the patterns of plant disease spread, just as it is for understanding the patterns of any dispersing or invading organism. While the issue of dispersal gradient truncation has been explored for plants and animals (Robledo-Arnuncio and Gil 2005; Klein et al. 2006; Bullock and Clarke 2000; Fujiwara et al. 2006; Paradis et al. 2002; van Houtan et al. 2007; Leitch et al. 2021), it is unclear how often disease gradient truncation is occurring in the plant pathology literature. Here, I review the plant epidemiological literature to determine whether truncated plant disease gradients are common. Other questions of interest are the following: (1) based on the papers of

Ferrandino (1993) and Kot et al. (1996), are there any differences in patterns of truncation before and after the shift in statistical characterization of the disease dispersal gradient concerning travelling versus dispersive waves, and (2) are the major forms of propagule dispersal (wind, splash, nematodes, insects) for plant pathogens represented accurately by fit dispersal gradients and, if not, which modes of dispersal are under-represented?

Methods

Literature Review

To find plant disease gradients in the plant pathology literature for review, I searched Web of Science, Google Scholar, and GALILEO databases (through the UGA library site). There are a variety of terms equivalent to dispersal gradient in the literature, including dispersal kernel or contact distribution (Tufto et al. 1997). To address this potential issue, I used a variety of search terms and keywords to find disease gradient studies (Table 2.1). Once I identified potential studies that could be considered for the literature review, I used the following criteria to determine whether papers would be included: (1) Did the study measure plant pathogen dispersal?; (2) Did the study include a figured dispersal gradient (scatter plot data with function plotted necessary to evaluate whether the dispersal gradient was likely to be reasonably fit or truncated)? If the studies met these criteria, I then recorded the information on the article itself, including the pathogen, measured dispersal distance, and dispersal function(s) presented by the authors (Table 2.2).

Organism & Dispersal Type

To assess which pathogens were represented in the plant disease dispersal literature, I recorded and tallied each represented pathogen, its type (fungus, bacteria, virus, nematode), and its primary vector (defined here as the pathogen's method of dispersal – wind, splash, insect, flowing water, or human). Even though most plant pathogens can be dispersed by a combination of vectors (Agrios 2004), I selected the primary vector as described by the author for their specific study system. For the purposes of this review, I define each of the primary vector types in separate paragraphs below:

Wind

Propagules of most plant pathogens are disseminated by air currents, particularly spores of most oomycetes and fungi (Agrios 2004). Depending on air turbulence and velocity, air currents may carry propagules upward or horizontally over short and long distances (Agrios 2004).

Splash

Splash dispersal occurs when propagules are washed or splashed onto plants or soil by overhead irrigation or rain (includes all bacterial and fungal spores in a sticky medium) (Agrios 2004). Splash dispersal does not make as significant a contribution towards long-distance dispersal compared to wind, which can carry splash-dispersed plant pathogens farther in more severe weather conditions (Agrios 2004). However, the more localized infections caused by splash dispersal may generate short-term and rapid disease intensification because the water associated with dispersal coincides in space and

time with the water necessary for the pathogen to infect the plant (Agrios 2004). It is important to note that splash dispersed studies I encountered in the published literature specifically excluded wind as a factor and experiments took place in controlled outdoor environments or in a greenhouse.

Flowing Water

Distinct from splash dispersal, flowing water from rain and/or ground irrigation moves pathogens present in the soil across or through the soil (via bulk water flow). This type of dispersal is common for bacteria, nematodes, spores, and mycelial fragments – oomycetes in particular (Agrios 2004).

Insects

Insects can travel over short or long distances, depending on the type of insect, relationship between pathogen and insect vector, and prevailing weather conditions – especially wind, which may carry insects over distances much further than they could move on their own (Agrios 2004). Insects may transfer certain fungal, bacterial, and nematode pathogens – depending on the insect-pathogen association (Agrios 2004).

Humans

Human-mediated dispersal can include a large swath of activities – from planting contaminated seeds or plant material to transporting infected plants across state lines (Agrios 2004). For the purposes of my literature review, however, I designated human dispersal as either mechanical transmission and movement of contaminated soil, by either

people or equipment moving through the field; this type of dispersal was only included when explicitly measured.

Classification of Dispersal Gradients: Truncated *vs.* Reasonably-Well Fit

To determine the degree of dispersal gradient truncation in the plant disease dispersal literature, I assessed and categorized all figures that visually depicted dispersal gradients (hereafter called "figured gradients"). The gradients without a figure were unable to be assessed, as I could not visually inspect the fit of measured data to the reported dispersal function. If figures were presented, the dispersal kernels were evaluated for truncation, and placed into one of three categories: reasonably fit, truncated, or indeterminate (Figure 2.1). **Reasonably fit gradients** displayed data points that were fit well close to the disease source and fit between data points in the tails (the predicted function did not end prior to or below points in the long-tailed part of the distribution, nor substantially above the data points). **Truncated gradients** could be represented by either spatially limited data, where no measurements were plotted yet the dispersal gradient clearly has not flattened into the tail of the distribution, or where the calculated dispersal gradient function approached or crossed the x-axis prior to the termination of the recorded data points. Finally, **indeterminate gradients** were those instances where dispersal functions either lacked sufficient data points for a reasonable function fit, those which only displayed either the dispersal function or the measured dispersal values (not both), or those which had a visually ambiguous representation where a function was presented as a figure but the data fit could not be reasonably assessed (this was uncommon).

When authors presented multiple figured gradients, all figured gradients were assessed and the overall trend for being reasonably fit, truncated, or indeterminable was assigned to the

study. When there was a choice between multiple figured gradients, I chose to represent the longest non-truncated gradient to give the study the best chance of meeting the reasonably fit criteria. If the authors presented multiple gradients for the same field over the course of their study (a chronosequence), I chose the last disease gradient in the timeseries to maintain consistency with studies that only presented a single end-of-season gradient. Finally, if the authors had multiple different gradients measuring different field sites out to the same dispersal distance, I arbitrarily selected the gradient that was presented first in the study.

Are Plant Disease Gradients Becoming More Appropriately Described With Time?

To assess patterns of truncation with respect to time, I recorded the years that each reviewed study was published. As mentioned previously, the studies of Ferrandino (1993) and Kot et al. (1996) represent a shift in dispersal ecology – invading organisms characterized by long-distance dispersal move as a dispersive wave, not a traveling wave. I wanted to determine if this philosophical shift in understanding was reflected in the reviewed plant pathology studies, particularly whether the frequency of disease gradient truncation decreased following these two important studies. The year 2000 was selected as the cutoff point for studies completed before and after these landmark studies, because statistical packages around this time were more diverse and widely available for function fitting algorithms and there was ~ 5 years following these publications for plant pathologists to become aware of and consider dispersive waves.

To determine the differences between truncation levels in my review pre- and post-2000, 1 ran a Z-proportions test on the truncated and reasonably fit gradients to determine whether the proportion of truncated studies statistically differed before and after the year 2000.

Results

I found over 200 papers through my initial database screening, but most of these manuscripts did not include figured dispersal gradients and instead presented tables of functions with model evaluation statistics and functions. For my purposes, the majority of these studies did not allow me to evaluate whether the dispersal function was reasonably well fit or truncated. In total, I reviewed 104 papers that met the requirements for assessment, and these studies were published between 1958 and 2021.

Fungi were commonly represented in the plant pathology literature and were the focus of 67.3% of dispersal studies in the reviewed manuscript ensemble. The next largest studied groups of plant pathogens were oomycetes and bacteria at 14.4% and 9.6%, respectively. Finally, only a handful of studies covered viruses or nematodes – just 5.8% (viruses) and 2.9% (nematodes) each (Figure 2.2).

Wind was by far the most common dispersal vector in the reviewed manuscripts, representing 64.4% of studies (Figure 2.3). Splash dispersal was represented by 20.2% of the reviewed studies, insects 9.6%, and flowing water and humans accounted for 3.9% and 1.9% of the remaining studies, respectively. Notably, all four flowing water studies measured oomycete (all *Phytophthora* spp.) dispersal in irrigation systems, and both human dispersal studies measured nematode dispersal via tractor plowing/disking activity.

As wind was the most common dispersal vector in my review (67 studies) and is well-known for dispersing plant pathogen propagules over long distances from a source, I wanted to understand the potential relationship between observation distances and whether the dispersal gradient was truncated within wind vectored plant diseases (Figure 2.4). I found that ~75% (50) of the studies tracked disease dispersal less than 100 m from the source, while only 5 studies

followed disease > 100 m (up to 250 m) from the source. The remainder of the wind vectored plant disease studies, 8 (~12%) attempted to account for disease dispersed over 500 m from the source. While I observed a wide range of measured dispersal distances of wind vectored plant diseases, over half (36 of 50) of the studies that tracked disease < 100 m from the source did so over distances < 50 m from the source (Figure 2.5). Of the studies that measured wind-dispersed pathogen dispersal < 50 m from the source (36 studies), more than half of those only tracked disease to no more than 15 m from the source (23 studies, 63.9%) (Figure 2.5). In grand total, 35% of the reviewed studies investigating wind vectored disease presented a disease gradient from data gathered within no more than 15 meters of the disease source.

Truncated gradients, which appeared from the presentation of limited data and/or poorly fit models, were common, comprising 63.8% of all figured gradients, while unclear and reasonably fit comprised 13.8% and 22.2% of all figured gradients, respectively.

There were some authors who acknowledged they were not intentionally measuring the full extent of their pathogen's dispersal, and knowingly truncated their disease gradients. Most of these studies were performed in an isolated environment, like a greenhouse, and were focused on understanding the mechanics of splash dispersal in the absence of wind. However, while many of these gradients were well-fit to their dispersal data, their relevance for extrapolation to an outdoor environment is limited. After removing these intentionally truncated gradients (n = 21) from the data set, I found that 71.2% of the remaining gradients (n = 73) were still truncated and only 17.8% were reasonably fit, the remainder (11.6%) of the reviewed studies were indeterminable.

The Z-proportions test indicated that there was a statistically significant difference between dispersal gradient truncation levels in the pre- and post-2000 eras (Table 2.4). Although

there was significantly less truncation in the post-2000 era, the number of studies truncating a dispersal gradient was still high (62.2% of the assessed gradients).

Discussion

In my review, I found relatively few studies of vectors of common viral and bacterial plant diseases (splash, insects, nematodes). Splash dispersal was the most studied of these vectors, making up ~20% of studies in the review, but insect and nematode vectors combined comprised less than 15% of the reviewed manuscripts. Although fungi are the most common plant pathogens, making up 85% of all plant pathogenic organisms, viral and bacterial diseases are extremely detrimental to high-value crops like fruit and vegetables. Additionally, human dispersed organisms – representing soilborne diseases and nematodes in my review – were barely accounted for in the literature and represented < 2% of studies. A lack of dispersal-focused studies on these organisms demonstrates a critical gap in plant pathogen epidemiological knowledge.

I found strong evidence that studies of wind dispersal are artificially and regularly truncated across the plant pathology literature. As seen in Figure 2.4, I found that over one third of studies of LDD, wind dispersed organisms never tracked disease further than 15 m from the source. These extremely short measurements indicate nearly all these studies will have truncated disease gradients, as LDD organisms are capable of dispersing hundreds of meters in a single generation (Farber et al. 2019) and hundreds of kilometers over a field season (Mundt et al. 2009; Mundt et al. 2013).

In my investigation of truncation levels in the literature pre- and post-2000, I found a decrease in dispersal kernel truncation in the post-2000 period; however, it is important to note

that truncation in this period was still common and was observed in over 60% of the studies in my review. The issue of continued dispersal truncation is not just an issue in plant pathology – it echoes problems throughout dispersal ecology in general (Clobert et al. 2012). While dispersal occurs over multiple spatial scales, there has been a movement to increase the spatial and temporal scales of dispersal studies to better capture LDD events (Nathan et al. 2006; Nathan et al. 2011b).

Given the significance of the length of dispersal gradient tails to accurately projecting and modeling plant disease spread, both collecting and properly accounting for long-distance dispersal from the epidemic source is essential. Artificially truncating disease gradients not only impacts the projections of how far disease may disperse, but also the rate at which it spreads. LDD organisms generate epidemics that accelerate at the invasion edge, such rapidly spreading epidemics quickly become difficult to suppress as they gain momentum and easily overcome dispersal barriers (Severns and Mundt 2022).

Table 2.1. List of all search terms and keywords used to find figured dispersal kernels in the plant pathology literature.

Search Terms
Plant disease/pathogen gradient
Plant disease/ pathogen contact distribution
Plant disease/pathogen dispersal kernel
Plant disease/pathogen distribution
Plant disease/pathogen dispersal distance

Table 2.2. List of all recorded information from each article in the literature review.

Article	Publication year	e.g. 2001
	Journal	e.g. Phytopathology
Pathogen	Host plant	e.g. Barley
	Organism	e.g. *Phytophthora infestans*
	Organism type	Fungus, oomycete, bacterium, virus, nematode
	Reproductive type	Sexual, asexual
	Dispersal type	wind, splash, insect, flowing water, humans
Experiment	Sampling scale from source	e.g. 50 m
	Inoculum type	Natural, artificial
Dispersal Function	Function	e.g. Gaussian, negative exponential, inverse power law
	Function type	thin-tailed, fat-tailed
	Funtion shown in figure	Yes, no
	Truncated	Yes, no, unclear
	Intentionally truncated	Yes, no

Table 2.3. Breakdown of literature by organism studied (a) and dispersal vector (b).

a)

Organism	Percent	Papers
Fungus	67.3%	70
Oomycete	14.4%	15
Bacterium	9.6%	10
Virus	5.8%	6
Nematode	2.9%	3

b)

Vector	Percent	Papers
Wind	64.4%	67
Splash	20.2%	21
Insect	9.6%	10
Flowing Water	3.9%	4
Humans	1.9%	2

Table 2.4. Z-proportions test for differences in truncation pre- and post-2000.

Publication Date	Total Truncated	Proportion Truncated	Not Truncated	Total Assessed
Pre-2000	42	0.91	4	46
Post-2000	23	0.62	14	37

Significance	0.05
Z-value	3.2
p-value (two-tailed)	0.0013

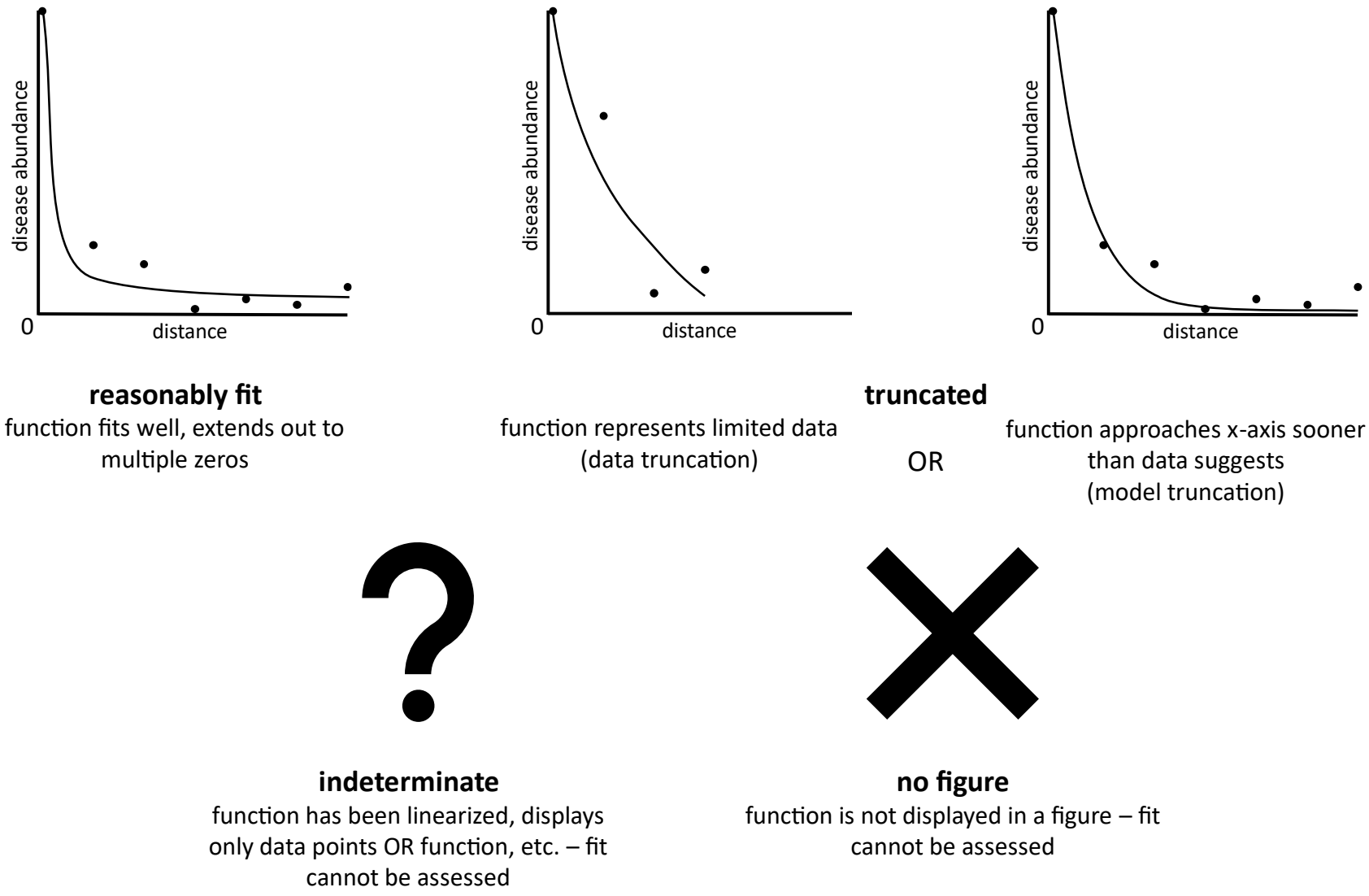


Figure 2.1. Graphs representing examples of the different categories of dispersal gradient function fit.

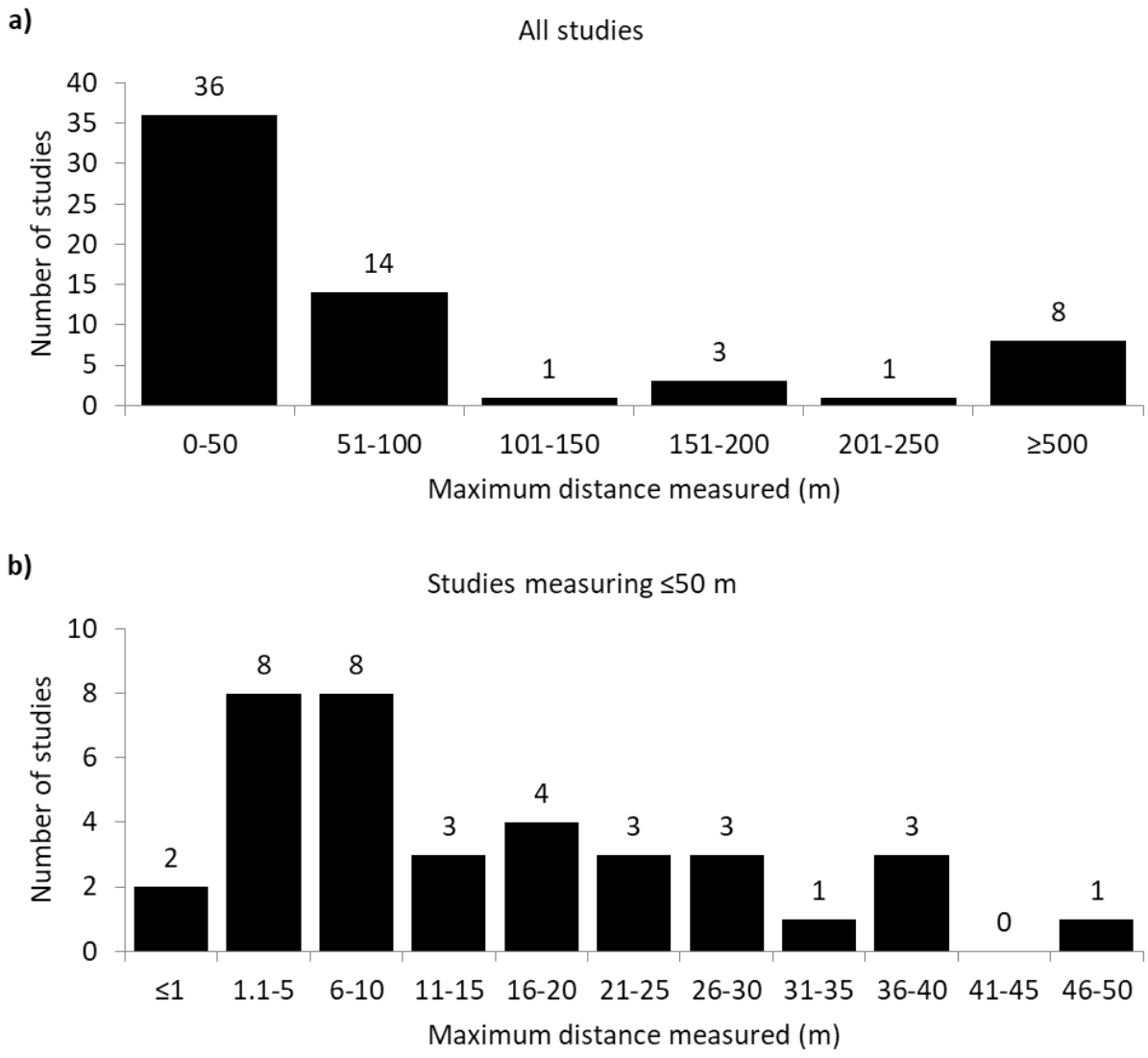


Figure 2.2. Histograms of dispersal distances for all wind dispersed pathogens (a) and all wind dispersed pathogens measured less than 50 m from their source (b).

Literature Cited

Agrios, G. N. 2005. *Plant Pathology*. 5th ed. Amsterdam: Elsevier Academic Press.

Bullock, J. M., and Clarke, R. T. 2000. Long distance seed dispersal by wind: measuring and modelling the tail of the curve. Oecologia. 124:506–521.

Clark, J. S. 1998. Why trees migrate so fast: confronting theory with dispersal biology and paleo-record. The American Naturalist. 152:204–224.

Clark, J. S., Silman, M., Kern, R., Macklin, E., and HilleRisLambers, J. 1999. Seed dispersal near and far: patterns across temperate and tropical forests. Ecology. 80:1475–1494.

Ferrandino, F. J. 1993. Dispersive epidemic waves: I. Focus expansion within a linear planting. Phytopathology. 83:795–802.

Fujiwara, M., Anderson, K. E., Neubert, M. G., and Caswell, H. 2006. On the estimation of dispersal kernels from individual mark-recapture data. Environmental and Ecological Statistics. 13:183–197.

Kable, P. F., Fried, P. M., and MacKenzie, D. R. 1980. The spread of a powdery mildew of peach. Phytopathology. 76:601–604.

Klein, E. K., Lavigne, C., Herve, P., Renard, M., and Gouyon, P.-H. 2006. Pollen dispersal of oilseed rape: estimation of the dispersal function and effects of field dimension. Journal of Applied Ecology. 43:141–151.

Kot, M., Lewis, M. A., and van den Driessche, P. 1996. Dispersal data and the spread of invading organisms. Ecology. 77:2027–2042.

Lietch, K. J., Ponce, F. V., Dickson, W. B., van Bruegel, F., and Dickinson, M. H. 2021. The long-distance flight behavior of *Drosophila* supports an agent-based model for wind-assisted dispersal in insects. Proceedings of the National Academy of Sciences. 118:1–9.

Mollison, D. 1977. Spatial contact models for ecological and epidemic spread. Journal of the Royal Statistical Society: Series B. 39:283–313.

Mollison, D. 1991. Dependence of epidemic and population velocities on basic parameters. Mathematical Biosciences. 107:255–287.

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

Mundt, C. C., Wallace, L. D., Allen, T. W., Hollier, C. A., Kemerait, R. C., and Sikora, E. J. 2013. Initial epidemic area is strongly associated with the yearly extent of soybean rust spread in North America. Biological Invasions. 15:1431–1438.

Nathan, R. 2006. Long-distance dispersal of plants. Science. 313:786–788.

Nathan, R., Katul, G. G., Bohrer, G., Kuparinen, A., Soons, M. B., Thompson, S. E., et al. 2011. Mechanistic models of seed dispersal by wind. Theoretical Ecology. 4:113–132.

Nathan, R., Klein, E., Robledo-Arnuncio, J. J., and Revilla, E. 2012. Dispersal kernels: review. In *Dispersal Ecology and Evolution* (J. Clobert, J. M. Bullock, T. G. Benton, and M. Baguette, Eds.), Oxford, UK: Oxford University Press.

Paradis, E., Baillie, S. R., and Sutherland, W. J. 2002. Modeling large-scale dispersal distances. Ecological Modelling. 151:279–292.

Pazhamala, L. T., Kudapa, H., Weckwerth, W., Millar, A. H., and Varshney, R. K. 2020. Systems biology for crop improvement. The Plant Genome. 14:e20098.

Pybus, O. G., Suchard, M. A., Lemey, P., Bernardin, F. J., Rambaut, A., Crawford, F. W., et al. 2012. Unifying the spatial epidemiology and molecular evolution of emerging epidemics. Proceedings of the National Academy of Sciences. 109:15066–15071.

Robledo-Arnuncio, J. J., and Gil, L. 2005. Patterns of pollen dispersal in a small population of *Pinus sylvestris L.* revealed by total-exclusion paternity analysis. Heredity. 94:13–22.

Roche, B. M., Alexander, H. M., and Maltby, A. D. 1995. Dispersal and disease gradients of anther-smut infection of *Silene alba* at different life stages. Ecology. 76:1863–1871.

Tredennick, A. T., Hooker, G., Ellner, S. P., and Adler, P. B. 2021. A practical guide to selecting models for exploration, inference, and prediction in ecology. Ecology. 102:e03336.

Tufto, J., Engen, S., and Hindar, K. 1997. Stochastic dispersal processes in plant populations. Theoretical Population Biology. 52:16–26.

van Houtan, K. S., Pimm, S. L., Halley, J. M., Bierregaard Jr., R. O., and Lovejoy, T. E. 2007. Dispersal of Amazonian birds in continuous and fragmented forest. Ecology Letters. 10:219–229.

Zwankhuizen, M. J., Govers, F., and Zadoks, J. C. Development of potato late blight epidemics: disease foci, disease gradients, and infection sources. Phytopathology. 88:754–763.

CHAPTER 3

SIMULATING TRUNCATED DISEASE DISPERSAL GRADIENTS IN EPIMUL

Introduction

In Chapter 2, I found that 64% of the published, peer-reviewed figured gradients in the review were truncated. Once disease gradients measured under greenhouse conditions were removed from consideration, the percentage of studies with truncated dispersal gradients increased to 71%. Although there was a significant decrease in the number of truncated studies after the year 2000, the proportion of truncated post-2000 published studies only decreased to 62%, still well over half of all studies with figured gradients. The number of studies including figures of their dispersal gradients had not changed over time, a z-proportions test found no statistically significant difference between pre- (25% unfigured) and post-2000 (21% unfigured) publications. Studies post-2000 also still commonly used traditional means of fitting (see Appendix A), such as the maximum likelihood or least squares method using log-transformed data fit to a linear model (Madden et al. 2017). When this function fitting approach is applied to truncated data, they will yield thin-tailed, exponentially bound functions that assumes pathogens disperse at a constant velocity as a traveling wave (Grenfell et al. 2001; Cummings et al. 2004). In comparison, fat-tailed functions like the power law predict epidemics that increase in velocity over space and time (dispersive wave) until hosts or the environment limit disease.

In this chapter, I am interested in understanding the effect truncation has on model projections of disease spread. I use wheat stripe rust (WSR, *Puccinia striiformis*) to parameterize an epidemiological model and simulate the spread of WSR in various landscape configurations

of susceptible host plants to the understand how projections from a poorly-fit, intentionally truncated WSR dispersal gradient differs from a properly and well-fit power law WSR dispersal gradient. The WSR dispersal gradient is likely the most well-described of all aerially dispersed plant pathogens and it has an accurately and statistically well-fit, fat-tailed dispersal gradient (Sackett & Mundt 2005a; Farber et al. 2019). The long tails of the WSR dispersal gradient are best fit to a power law function (Sackett and Mundt 2005a), and this function is commonly used to more accurately model the movement of long-distance dispersed organisms (Nathan et al. 2012).

I use computer simulation models to evaluate potential impacts of truncation on plant disease spread for two main reasons. First, to run a field experiment at the scale necessary to address the research questions would be incredibly costly, time consuming and potentially impossible. Second, modeling is relatively quick to perform and allows me to evaluate more scenarios than one would be able to assess in field experiments. For example, to complete a thorough study of dispersal, similar field studies required over 15 hectares of land, multiple years of inoculum production, and hundreds of hours of data collection – even with these resources, these studies could only assess a couple different questions with a small number of replicates (Sackett and Mundt 2005; Cowger et al. 2005; Severns et al. 2015). Disease dispersal is a complex process affected by variable and difficult to estimate factors, so modeling, although requiring simplifying assumptions and imperfect process, much like many other scientific methodologies, provides a means to potentially understand important aspects of complex systems. There are three main uncertainties in predictive models of dispersal: model, parameter, and inherent uncertainty (Clark et al. 2003; Higgins et al. 2003; Nathan et al. 2012). However, despite the abstractions and simplifications, models still provide useful information, especially

when they are well-parameterized and more faithfully represent biological realities. Disease dispersal models have been successfully used to characterize landscape level spread of human, animal, and plant long-distance dispersal (LDD) epidemics (Gisiger 2001; Mundt et al. 2009; Mundt et al. 2013; Tildesley et al. 2012; Rieux et al. 2014; Ojiambo et al. 2017).

Studies have shown that varying levels of susceptible host availability, connectivity, and host patchiness alter disease progression (Mundt et al. 2009; Meentemeyer et al. 2012; Burdon & Thrall 2014; Severns et al. 2014; Zhang et al. 2019; Iwamura et al. 2020). Landscape connectivity is defined as the degree to which the landscape facilitates or impedes movement among resource patches (Taylor et al. 1993). For pathogens, connectivity is straightforward – if pathogens can find susceptible hosts, they can colonize and subsequently reproduce on that host. However, when host plants are sparse or patchily distributed, their colonization probability is reduced relative to more continuously distributed hosts (Ruxton and Schaefer 2012; Wilson et al. 2016; Ismail et al. 2017). In Lindström et al. (2011), the authors modeled potential disease dispersal for Dutch elm disease and livestock disease based on the observed spatial distribution of oak and elm trees and pig and cattle farms in Sweden. These landscapes varied in both host patch connectivity and density. After testing both thin- and fat-tailed dispersal functions on each landscape, the authors found that both effective dispersal and invasion speed is dependent on the spatial host arrangement. Importantly, the authors noted that the differences between invasion success between thin- and fat-tailed gradients was most pronounced in landscapes with low connectivity. In these scenarios, fat-tailed models were able to infect distant patches, albeit at a slower rate, while thin-tailed models almost never colonized host patches isolated by the greatest distances between host plants. This colonization and establishment phenomenon in which the dispersal kernel interacts with colonization and organism establishment has been observed in

other systems as well, including invasive plants (Minor and Gardner 2011; McConkey et al. 2012) and animals (Nesslage et al. 2007; Bridgman et al. 2012).

Given the intuitive influence of landscape on disease dispersal, I was interested in understanding how different landscape arrangements of susceptible hosts interact with thin- and fat-tailed gradients to influence disease progression. In this chapter, I evaluated the *in silico* patterns of disease spread projected from thin-tailed (exponentially bound, travelling wave) and fat-tailed (non-exponentially bound and dispersive wave) disease gradients across different landscapes of host plant distribution. To understand how dispersal kernels and the host landscape interact to influence epidemic spread, I used the same WSR dispersal data set which was properly and accurately fit to a fat-tailed function (modified inverse power law) and improperly fit to a truncated (exponential) dispersal gradient. I evaluated homogenous, slightly patchy, and highly patchy fields of susceptible and resistant host plants to evaluate how the degree of host distribution impacts epidemic extent and magnitude for the different disease gradients. I expected to see dispersal gradient truncation limit disease dispersal in landscapes with limited host availability in comparison to landscapes with unlimited host availability and to miss-project the amount of disease after multiple generations of disease intensification and spread.

Methods

EPIMUL

The EPIMUL simulation model was created by Kampmeijer & Zadoks (1977) to investigate the effectiveness of host mixtures and different pathogen races in wheat stripe rust. Although based off the original EPIMUL program, the version used in this study has been highly modified, making it broadly applicable to fungal plant pathogens in general. Modifications

include the addition of alternate dispersal gradients (Mundt & Leonard 1985; Sackett & Mundt 2005b), improved spore production functions, removal of obsolete sub-routines, stochastic disease dispersal gradients (Severns et al. 2015), and inclusion of a spray out function (not used in this study) (Severns and Mundt 2022).

EPIMUL is a flexible, compartmental model simulation program that can vary pathogen, host, and dispersal parameters. Most importantly, all parameters in EPIMUL are independent of one another, allowing for the precise control of the variables, including basic reproductive number (R_0), latent period, infectious period, and the initial infection level (P_0). Independent control of disease parameters allows EPIMUL users to explore a wider range of scenarios when compared with other epidemiological models like the well-known and population foot and mouth disease models of animals (Keeling et al. 2001; Chis Ster & Ferguson 2007) which automatically link R_0 to farm size, limiting the user's ability to decouple farm size from the basic reproductive number. Such tradeoffs are important to evaluate as host population size may or may not be tightly linked to the basic reproductive number if susceptibility or resistance differs between hosts and locations.

EPIMUL represents the host plant population as a two-dimensional spatial grid. Each cell (a compartment in compartmental models) in this grid may represent an individual, groups, or even fields of host plants – making it flexible to accommodate different conditions and spatial scales of study. Users may also specify host arrangement and level of susceptibility, as well as the locations and quantity of initial pathogen inoculum. Each compartment contains a user-defined number of infectible compartments (lesion sites) which exist in one of four states: healthy, latent (infected but not sporulating), sporulating, or removed (host death or recovery based on the specific SIR conceptual model). When a single cell equals a single plant, healthy

cells represent uninfected hosts, latent cells represent infected hosts that are generating lesions, sporulating cells represent infected hosts that are releasing infective inoculum, and removed cells represent infected hosts that have died and can no longer produce inoculum. When a single cell represents multiple plants, as in my simulations, the proportion of uninfected/ latent/ sporulating/ removed hosts is tracked for each day for the entire number of days in a simulated "epidemic".

Effective spores (those that produce an infection) can be dispersed via one of three dispersal gradient functions: the modified power law (Mundt & Leonard 1985),

$$y = a(x + c)^{-b}$$

an exponential function (Kiyosawa & Shiyomi 1972),

$$y = a\exp(-bx)$$

or the Lambert model (Lambert et al. 1980).

$$y = a\exp(-bx^n)$$

In each of three equations above, *y* represents the number of effective spores (those spores that generate infections) dispersed to *x* distance from the source lesion, *b* controls gradient steepness, and *a* is proportional to the amount of inoculum produced at the source (this parameter has no effect on gradient shape and is only used to set the proportions of asymmetrical upwind and downwind gradients). In the modified power law model, *c* is an offset parameter that allows the power law to have a finite value at $x = 0$. Finally, *n* controls gradient shape in the Lambert model (Sackett and Mundt 2005b).

In EPIMUL, the latent period begins after dispersal, and new infections are initiated in the destination cell in proportion to the number of healthy sites. After the latent period ends, the plant becomes infectious and lesions release spores, and are considered "removed" once the infectious period ends (*e.g.*, cannot infect or become infected). Spore production, infectiousness,

and loss to the environment are combined into a single parameter – the number of "effective spores" created per lesion per day (when multiplied by the number of days in the latent period which yields the basic reproductive number, R_0 the number of effective daughter lesions per parent lesion).

Simulation Parameter Values

To model disease progression, we used a field size of 200 × 200 cells in our simulations. While we used larger field sizes in our preliminary investigations, 1000 × 1000 cells (the maximum field size in EPIMUL) and know that the power law model distributes effective spores beyond this range, the 200 × 200 cell field size was selected for multiple reasons: (1) it captured more than the full extent of exponential dispersal after 5 disease generations, and (2) it allowed me to see patterns near the focus more clearly, and (3) drastically reduced the computational time. Each cell in the virtual fields were 1.52 m × 1.52 m, which was used by Sackett & Mundt (2005b) to parameterize their dispersal gradient equations and it is also the width of a wheat planter. Carrying capacity was set to 200,000 sites per cell (100 individual wheat plants which is typical production planting density) in accordance with previous studies (Severns & Mundt 2022). Disease outbreaks were set to begin in a single cell (a focus) in the center of the field to visualize disease progression in all directions from the source. Outbreak disease levels (P_0) play an important role in final epidemic magnitude, and seemingly small increases in the amount of initial disease can drastically increase final disease severity and extent (Severns et al. 2015; Severns and Mundt 2022). To account for the influence of P_0, the outbreak cell was inoculated at two different rates, depending on the scenario – low (200 spores, 0.1% initial infection) and high (2,000 spores, 1% initial infection) outbreak levels. These two initial disease levels represent low

naturally occurring disease levels and a higher level more likely in an agricultural environment (Antonovics and Alexander 1992; Zadoks and van den Bosch 1994).

To directly evaluate the impacts of a truncated disease gradient we parameterized exponential and power law dispersal gradient equations using the same WSR field dataset. For ease of interpretation, we made both dispersal gradients used in our simulations symmetrical and representative of downwind dispersal – the farthest and fastest available – giving both equations the best chance possible to represent long-distance dispersal. Modified power law parameters were derived from Farber et al. (2019), which characterized the rust dispersal gradient over 10 km from the source and is probably the best-known plant pathogen dispersal gradient available. Parameter values for the modified power law were as follows: a = 451, b = 2.28, and c = 0.23. Using the same dataset from Farber et al. (2019), an exponential function was fit via the least squares method, which is still used in the plant pathology literature to fit dispersal functions (Maffia and Berger 1999; Travadon et al. 2007; Qandah and Del Rio Mendoza 2012; Madden et al. 2017). Parameter values for the negative exponential were as follows: where a = 1, b = 0.1903, and c = 1. Previous studies have shown that truncating the dispersal gradient by using a thin-tailed function, like the exponential, misrepresents the distribution of disease – overpredicting disease levels at intermediate levels, while underpredicting disease levels at the focus and in the tails, and ultimately failing to capture LDD events (Sackett and Mundt 2005a; Nathan et al. 2012) (Figure 3.1).

We used two types of simulations in our study, deterministic and stochastic. Deterministic simulations represent systems with no variation – all values are fixed, and the model will output the same result every time it runs. Deterministic simulations were used in our study to provide a general prediction of epidemic magnitude and extent. In EPIMUL the average

disease projection values taken from large numbers of stochastic runs are comparable to values predicted by a single deterministic simulation (Severns et al. 2015). Stochastic simulations were used to account for variation that would occur in independent outbreaks and the variable nature of host-pathogen interactions. Stochasticity in EPIMUL is generated by resampling and redistributing the daily whole effective spores from the dispersal function (power law or exponential) with a Poisson distribution. I ran one deterministic and 100 stochastic simulations for each scenario, resulting in 2,020 total stochastic simulations (Table 3.1). The average runtime was about 30 minutes per simulation which, including deterministic simulations, amounts to over 1,000 hours (42 days) of runtime.

Empirical studies have shown that the spore production profile of WSR is not consistent throughout the latent period, and in fact peaks at 4 days post infection (Papastamati and van den Bosch 2007). While EPIMUL is not able to model WSR's variable spore production, it ultimately does not affect our simulations as R_0 is steady throughout the latent period; this is because the probability of effective spores is proportionally decreased as disease increases and available infection sites decrease. Additionally, I was focused on understanding the patterns between entire disease generations, not specific days or partial generations, so precision in this parameter was not required given the goals of this study.

The latent period of WSR is dependent on temperature, host resistance, and abiotic conditions. In cooler climates, the latent period for WSR can last weeks and shorten as temperatures increase, resulting in an average latent period of 17 days (Zadoks 1961; Tollenaar and Houston 1967; Shaner and Powelson 1971; Sackett and Mundt 2005b). Latent periods as short as 10 days have been reported in warmer wheat growing regions, where increased temperature and humidity rapidly accelerate fungal growth (Chen et al. 2002; Li and Zeng 2002;

Wan et al. 2004; Papastamati and van den Bosch 2007; Wellings 2007; Chen et al. 2009). I used a latent period of 10 days in this study, as it is still biologically observed and using a shorter latent period decreases computational time (Table 3.2). Infectious period is also dependent on host resistance. Studies have reported infectious periods of 6-7 days for slow-rusting varieties, while susceptible varieties have been observed producing spores for up to 14 days (van den Bosch et al. 1988; Luo & Zeng 1995). As my simulated fields contain a mixture of susceptible and resistant hosts, I set the latent period at an average value of 10 days (Table 3.2).

The length of WSR epidemics is controlled primarily by growing season, weather, and host availability. In cooler wheat growing climates, WSR epidemics typically only persist for three to four latent periods as they are bound by short growing seasons – unfavorable cool, spring weather early in the year and harvest early in the fall. In warmer wheat growing climates, however, longer growing seasons with favorable warm, humid weather mean that WSR epidemics can last up to 10 latent periods (Chen et al. 2002; Li and Zeng 2002; Wan et al. 2004; Wellings 2007). I settled on an epidemic length of 5 generations in this study (Table 3.2) as a middle-ground between cold regions (3 generations) and warm regions (10 generations).

The basic reproduction number (R_0) represents the mean number of effective daughter lesions per parent lesion; this number is obtained by multiplying the latent period by the number of effective spores produced per lesion, per day. According to previous studies, the mean R_0 value for a WSR epidemic is ~70 (Papastamati & van den Bosch 2007; Sackett & Mundt 2005b; Severns et al. 2019; Severns et al. 2022) and this value was used in my simulations.

Susceptible Host Landscapes

To understand how the degree of susceptible host plant distribution interacts with projections of disease spread and epidemic magnitude when the disease dispersal gradient is properly described (as a power law function) or artificially truncated (as an exponential function) I simulated dispersal across five different host arrangements. Hosts in the 5%, 25%, and 50% susceptible fields (Figure 3.2) were assigned to compartments at random and represent patchy environments with low to high levels of host availability, common for naturally occurring epidemics (Keeling et al. 2001; Parnell et al. 2009; Meentemeyer et al. 2012; Cunniffe et al. 2016; Severns et al. 2019). The 50% susceptible (rows) field and 100% susceptible fields (Figure 3.2) represent agricultural environments with homogenous or consistently alternating patterns of host resistance, such as variety mixtures or intercropping with different species.

Analysis

To prepare the simulation data for analysis, all simulation files (100 stochastic and 1 deterministic) were grouped by landscape, initial disease value, and dispersal gradient (*i.e.,* 5% susceptible, 0.1% initial disease, power law). Results files were extracted from each individual simulation folder, and the data in each cell was gathered and assigned to each compartment's location for the last day of the epidemic using a Python script (Appendix B).

To quantify the effects of disease gradient truncation on the extent of disease spread, I compared statistics from the deterministic simulations for each combination of landscape, initial disease value, and dispersal gradient. First, I converted the results data from percent infected per compartment to number of infections per compartment. Then I calculated the percent infected cells (disease extent, Figure 3.3) and percent infected compartments (disease severity, Figure

3.4) for each deterministic simulation, and these data were then plotted on a bar graphs to visualize the differences between all simulations.

To provide a visual representation of the differences in amount of disease and distance traveled by the pathogen, I made field maps of the final day of each simulation using percent infection in each compartment and indexed the relative amount of disease as a heat map. To create these images, I used a Python script to pull the results files from all 100 stochastic runs for each simulation, extracted data for the last day of the epidemic, concatenated all data and found the mean, median, and maximum values for each simulated field cell across all 100 stochastic runs (Appendix B). These results were then visualized by simulation scenario (*e.g.*, 5% susceptible, 0.1% initial disease, power law, Table 3.1), with median values displayed for the entire field, and minimum, median, and maximum values displayed as 50 x 50 cell insets (Figures 3.5–3.14).

To visualize the effects of disease gradient truncation on disease progression throughout the simulated epidemic, I visualized the epidemic from beginning to end. For simplicity, I only chose to represent 100% susceptible simulations. To create these images, one of the 100 stochastic runs was selected from each simulation, and results data was extracted for every half generation (5 days) of the simulated epidemic (Appendix B). These time series data were then visualized as whole fields and grouped by initial disease level and dispersal function (Figures 3.15–3.18).

Results

Exponential dispersal gradient simulations

Exponential gradient simulations showed that disease extent (the percent of cells containing at least one infection) increased with increases in initial disease levels and host availability (Figure 3.3). Disease extent for 1% initial disease fields was 25-30% greater than 0.1% initial disease fields for all the different susceptible host arrangements (5%, 25%, 50%, 100% of the cells). Disease extent also increased as host availability increased, with levels almost doubling between 5% to 100% in susceptible host fields within the different outbreak disease levels.

Exponential gradient simulations showed that percent infection (percent compartments filled) increased with both increases in initial disease levels and host availability (Figure 3.4). The percent compartments filled was an order of magnitude greater across all host susceptibility levels from 0.1% to 1% initial disease fields. As expected, percent infection increased as host susceptibility increased. Visual patterns of the disease spread at 5-day time intervals in the 100% susceptible field strongly suggested disease expansion at an approximate steady rate of increase, which is the travelling wave produced by an exponential dispersal gradient (Figures 3.15 & 3.17).

Power Law

In contrast to the exponential dispersal gradient simulations, the power law gradient simulations saturated the entire field for all initial disease levels and host availability scenarios (Figure 3.3). WSR was quickly dispersed over all compartments in the largest field that EPIMUL memory would allow (1000 × 1000 cells) when using the power law disease gradient in

preliminary exploratory simulations. This confirmed that I could not account for all WSR infections by the end of the 5-generation time frame which was readily accomplished in a 200 × 200 compartment field using the exponential dispersal gradient. Because the power law dispersal gradient dispersed disease over a far greater area than I could account for with EPIMUL, the percent infection was 100% for all combinations of parameter values and host availability scenarios (Figure 3.4). Like the exponential fields, infection levels for 1% initial disease fields were consistently an order of magnitude greater than fields with 0.1% initial disease. However, there were negligible differences in percent compartments filled across all host availability types, with within cell levels averaging differences oft ~0.3%. Analysis of the power law disease gradients over time in the 100% susceptible fields showed an increasing rate of dispersal over time, with total disease distributed to all compartments in the field by generation 2.5 (Figures 3.16 & 3.18), which is 1.5 generations of disease dispersal. Additionally, I observed the development of sub-foci using the power law dispersal gradient, which are frequently observed in wind dispersed plant disease epidemics (Zadoks and van den Bosch 1994; Severns et al. 2019). While these sub-foci were subtle in appearance (Figures 3.16 & 3.18), it is important to note that they were not observed in any of the exponential dispersal gradient simulations.

Discussion

Given the significance of the type of disease gradient used to project plant disease spread, collecting and accurately representing dispersal far from the epidemic source is essential. When I intentionally truncated the well-described power law dispersal gradient for wheat stripe rust to an exponential disease gradient (using the still common approach published in the plant disease literature Maffia and Berger 1999; Travadon et al. 2007; Qandah and Del Rio Mendoza 2012;

Madden et al. 2017), I documented two obvious and important outcomes. First, the disease dispersal distances projected from the exponential model was drastically underestimated compared to the power law. The power law (fat-tailed) disease gradient consistently reached the edges of the field across every combination of outbreak disease levels and all different landscape distribution of susceptible host plants (Figures 3.5–3.14). Even in 1,000 x 1,000 cell field simulations (Figures 3.19 & 3.20), the power law gradient distributed disease to the field edges for all field types. Second, my simulations provide evidence that increased host availability amplifies inaccuracies in epidemic magnitude created by dispersal truncation. Previous studies (Sackett and Mundt 2005a, b) have shown that when the exponential is used to represent WSR dispersal, the amount of disease is misrepresented, with a tendency to underestimate the amount of disease up close to the focus and at long distances away from the focus, while substantially overestimating the amount of disease at intermediate distances (Figure 3.1). While host availability does not appear to impact the disease levels when hosts are sparse, predictions by the exponential increasingly depart from what is likely a more biologically accurate projection as hosts become more abundant and continuously distributed. Additionally, while I observed the development of sub-foci in the power law simulations, I did not observe them emerge, even subtly, in the exponential simulations. Sub-foci are highly localized infections that develop via longer distance dispersal and stochastic events from the primary disease source, and act as an additional, more distant inoculum source as an epidemic develops (Zadoks and van den Bosch 1994; Severns et al. 2019). These sub-foci are commonly observed in plant epidemics (Zadoks and van den Bosch 1994) and are characteristic for LDD plant pathogens.

In my simulations, exponential disease gradient simulations manifested greater levels of whole field disease than the power law simulations in fields with 50% or greater host density.

However, when I compared the ratio of disease in 100% susceptible fields for both 200 x 200 and 1,000 x 1,000 cell fields, I observed this difference between the two dispersal gradient types notably diminish. While I was not able to run simulations in fields larger than 1000 × 1000 cells due to memory restrictions within the EPIMUL program code, it is highly likely that if the virtual field was large enough to follow the majority of dispersed disease by the power law gradient the same increases in extent with increases in initial disease and host availability seen in the exponential gradient simulations would have been observed. Power law functions culminate in a more expansive distribution of disease which, when aggregated over large areas, would be likely to surpass the aggregate disease levels projected by a truncated (exponential) disease gradient.

Although the exponential model may be able to coarsely project disease close to the focus, studies of WSR in the field have shown that they are unable to accurately represent disease in the tails and will overestimate the amount of disease up close to the outbreak as disease builds. In Sackett and Mundt (2005a), they measured WSR primary dispersal gradient (the disease dispersed from solely the first generation) over 75 m from the disease focus and fit different functions to their field data. While they found that power law functions provided the best fit to their complete data sets, when they mimicked a truncated field dataset by fitting models to only the first 20 m of their data, they found that log-transformed exponential equations provided an adequate fit to their truncated data; however, when the data were back transformed, they found the exponential overestimated dispersal up close to the outbreak. These results provide an excellent demonstration of model mis-fit due to an LDD organism due to truncated data.

Previous WSR studies have shown that exponential functions do not accurately represent patterns seen in the field. In a figure adapted from Sackett & Mundt (2005a), empirical measurements of lesions per leaf are fitted to a natural log transformed exponential function (Figure 3.1). In this figure, it is clear that the improperly fit exponential function estimates far too much disease at intermediate distances and, because WSR is polycyclic, this misrepresentation is compounded over multiple generations and results in highly inflated disease levels near and at intermediate distances from the outbreak. Most critically, however, the exponential function moves at a constant rate over time (Figures 3.15 & 3.17) and never reaches the field edge. This is in stark contrast to the power law function, which moves at an increasing rate and sends infections to the field edge within 21 days (1% initial disease) and 25 days (0.1% initial disease), or about 2.5 generations (Figures 3.16 & 3.18). In fact, in a large field simulation of 1000 × 1000 cells (over 1.5 km^2) of the power law function, I could not come close to representing its full dispersal expanse over 5 generations. This has important implications for disease management. Disease management strategies based on an assumption of constant rate of dispersal (a traveling wave generated through an exponential function) will underestimate the location of the disease front (Severns & Mundt 2021), resulting in delayed treatments that do not account for the full extent of disease required to eradicate the disease outbreak.

Even outside plant epidemiology, studies comparing the fit of fat- and thin-tailed dispersal kernels to empirical dispersal data have also concluded fat-tailed kernels outperform thin-tailed ones when the goal is to accurately describe and project the movement of LDD organisms. These include dispersal studies of pine and oilseed rape pollen (Robledo-Arnuncio and Gil 2005; Klein et al. 2006), heather seeds (Bullock and Clarke 2000), brown trout (Fujiwara et al. 2006), and British tits and Amazonian birds (Paradis et al. 2002; van Houtan et al. 2007).

Studies have demonstrated that limiting the susceptible host availability and connectivity will slow disease progression (Mundt et al. 2009; Meentemeyer et al. 2012; Burdon & Thrall 2014; Severns et al. 2014). In the exponential model simulations, my results showed a clear reduction in disease extent and severity as the number of available hosts decreased. However, in the power law model simulations, I found that disease extent was unaffected and disease severity was only slightly higher in the 100% susceptible fields compared to the 5% susceptible fields. As discussed earlier, the lack of change in epidemic extent between the different levels of host susceptibility for the power law simulations is likely due to limits of the simulation field size – even in 1000 × 1000 cell simulations, I was far from representing the full extent of the power law dispersal kernel.

The 50% susceptible landscapes (as patched and alternating rows), which represent agricultural mixtures, had a 32% and 15% decrease from the 100% fields for the exponential and power law fields, respectively. In other studies of disease in agricultural mixtures, authors found similar disease reductions. Finckh et al. (1999) found powdery mildew incidence was reduced by 33–71% in barley variety mixtures when compared to pure stands of the same varieties. Similarly, Garrett and Mundt's (2000) study of potato late blight disease reduction in genotype mixtures with qualitative disease resistance found close to 40% disease reduction on susceptible plants in their 2:1 mixture of susceptible and resistant plants. Cowger et al. (2005) found that 1:1 or 1:4 mixtures of susceptible and resistant varieties of wheat decreased apparent wheat stripe rust infection rates in all but one location-year combination. In their study on genetic diversity and rice blast disease control, Zhu et al. (2000), found that planting a resistant variety of rice in between susceptible rows resulted in 94% less disease than susceptible plants grown in monoculture. While intercropping and mixtures can reduce the total amount of disease, my

power law simulations suggest that the expanse of disease dispersal is unlikely to be limited by host patchiness in the same way that the exponential simulations were limited by susceptible host distribution.

The 25% susceptible landscapes (patchy and rows), which represent a dominant plant species in a natural landscape, had a 91% and 25% decrease in disease levels from the 100% fields for the exponential and power law simulations, respectively. In their study of the effects of landscape heterogeneity on sudden oak death (*Phytophthora ramorum*) in California, Condeso and Meentemeyer (2007) found that alternative host and canopy cover were consistently significant plot-level covariates to disease severity. The authors argued that denser woodlands likely meant an increase in alternative hosts, leading to larger reservoirs of disease and higher overall infection levels. They suggest these large, dense patches of woodland may become pathogen reservoirs for sudden oak death, containing more disease than smaller patches. Conversely, they note increased woodland fragmentation is likely to reduce the rate of less mobile and more environmentally vulnerable pathogens, a common theoretical paradigm in the field of landscape epidemiology (Burdon et al. 1989; Ostfeld et al. 2005). Meentemeyer et al. (2011) reported similar results in their study of the same sudden oak death pathosystem. The authors found that, while most disease dispersal occurred locally (< 250 m), infrequent LDD events accelerated disease epidemics in regions with high host availability (Meentemeyer et al. 2011).

The 5% susceptible landscapes (patchy and rows), meant to represent the distribution of a common plant species in a natural landscape, had a 99% and 38% decrease in whole-field disease levels from the 100% susceptible landscape for the exponential and power law fields, respectively. In their study the epidemiology of anther-smut infection in *Silene alba*, Antonovics

and Alexander (1992) measured spore deposition on host flowers in natural and artificially induced epidemics. The authors found that the rate of spore deposition increased with the frequency of diseased individuals but saw no change in deposition when there was an increase in the density of diseased individuals. These results complement my simulation results, which show that, regardless of host density, LDD pathogens have a high probability of dispersing, whereas this phenomenon would not be realized if the dispersal gradient is truncated to an exponential function.

My results confirm that accurately characterizing the dispersal gradient is critical to understanding and projecting how disease is distributed and epidemic progresses over time and space. I found that truncated gradients not only underestimate epidemic extent, but also epidemic velocity and magnitude, and this effect is magnified by the degree of host plant scarcity. While the penalty for truncating an LDD dispersal gradient with an exponential function is large and notable in monocultures and mixtures, it generates considerably more biologically inaccurate projections when susceptible host plants become increasingly rare. These results likely explain why LDD plant diseases persist in natural systems – even though the majority of propagule dispersal occurs close to the disease source, a small number of pathogens propagules can travel to distant host patches and establish new infections. This may, in turn, explain why LDD pathogens are so difficult to control and often escape area-based containment measures (Cunniffe et al. 2015; Peterson et al. 2015; Keeling et al. 2003; Dybiec et al. 2005; Parnell et al. 2009; Brooks-Pollock et al. 2015). These results have important implications for the timing of disease detection and treatments, which must occur earlier and at a larger scale than predicted by traveling wave models (Severns and Mundt 2022). I encourage plant pathologists to make the most of their time and effort and consider designing epidemiological field experiments that

measure disease dispersal over greater distances, include more sampling locations, and use methods that can detect low levels of disease to more accurately represent the LDD gradient expected from many plant pathogens.

Table 3.1. List of all simulation types.

Model	Susceptible Hosts	Arrangement	Initial Inoculum Level
modified power law	5%	random	0.1% 1%
	25%	random	0.1% 1%
	50%	random	0.1% 1%
	50%	alternating rows	0.1% 1%
	100%	fill	0.1% 1%
exponential	5%	random	0.1% 1%
	25%	random	0.1% 1%
	50%	random	0.1% 1%
	50%	alternating rows	0.1% 1%
	100%	fill	0.1% 1%

Table 3.2 List of all simulation parameter values.

Parameter	Value
Field size	200 x 200 cells
Compartment size	1.52 x 1.52 m
Epidemic legnth	5 generations
Sites/compartment	200,000
Spore placement	cell [100, 100]
Latent period	10 days
Infectious period	10 days
Daily multiplication factor	7
R_o	70
Host varieties	2 (100% susceptible, 100% resistant)

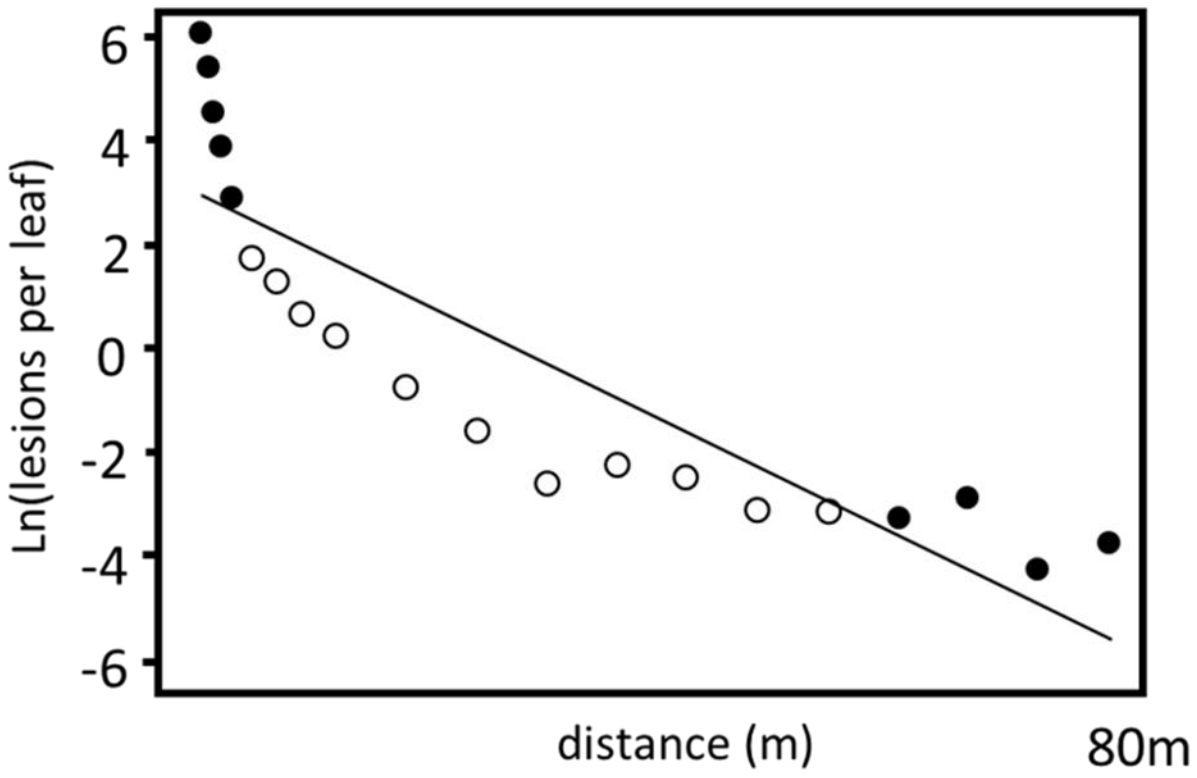


Figure 3.1. Field data of WSR lesions per leaf fitted to a natural log transformed exponential function. Demonstrates the improper fit of the exponential function to WSR data, which underpredicts disease levels near the source and in the tails (filled circles), and overpredicts disease levels at intermediate distances (open circles). Adapted from Sackett and Mundt (2005a).

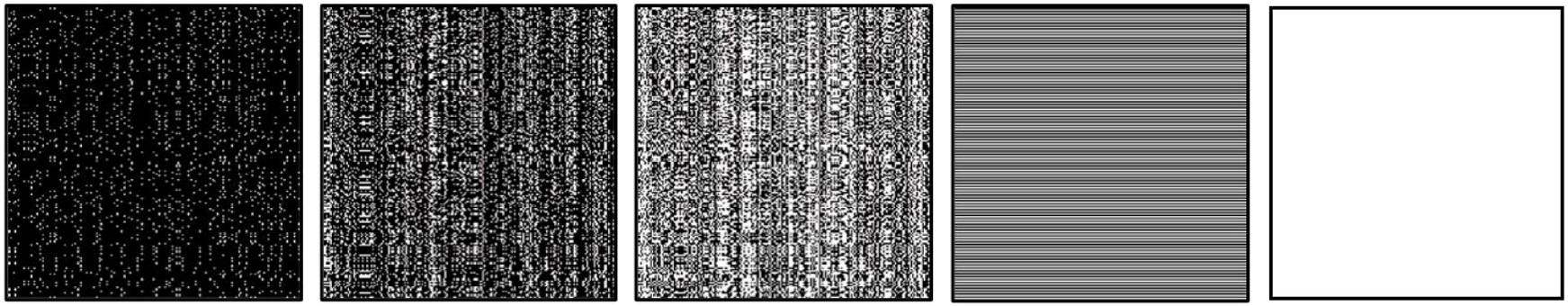


Figure 3.2. Randomly generated simulation field layouts. Black represents fully resistant hosts white represents fully susceptible hosts. From left to right: 5% susceptible (patchy), 25% susceptible (patchy), 50% susceptible (patchy), 50% susceptible (rows), 100% susceptible.

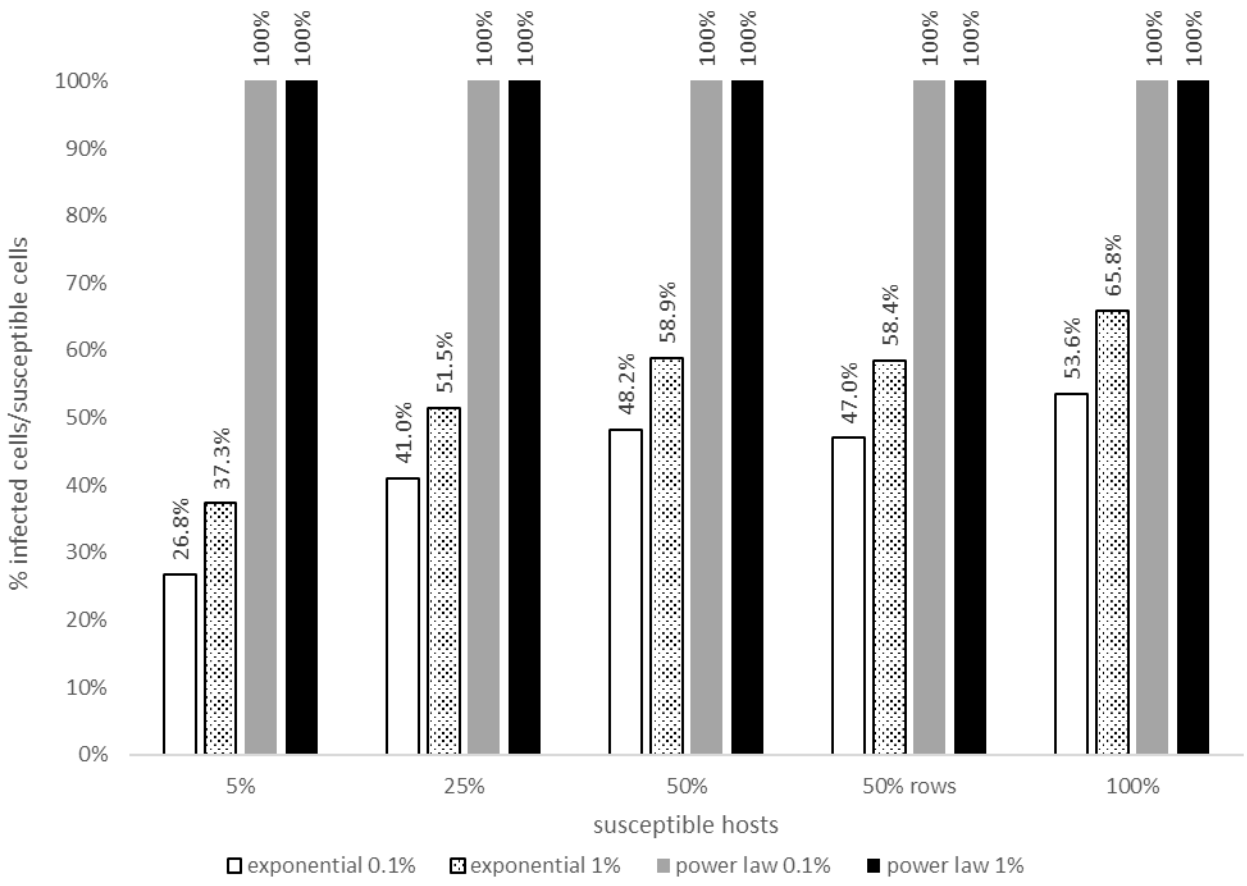


Figure 3.3. Bar chart of percent infected cells per total infectible cells for each field type and initial disease level.

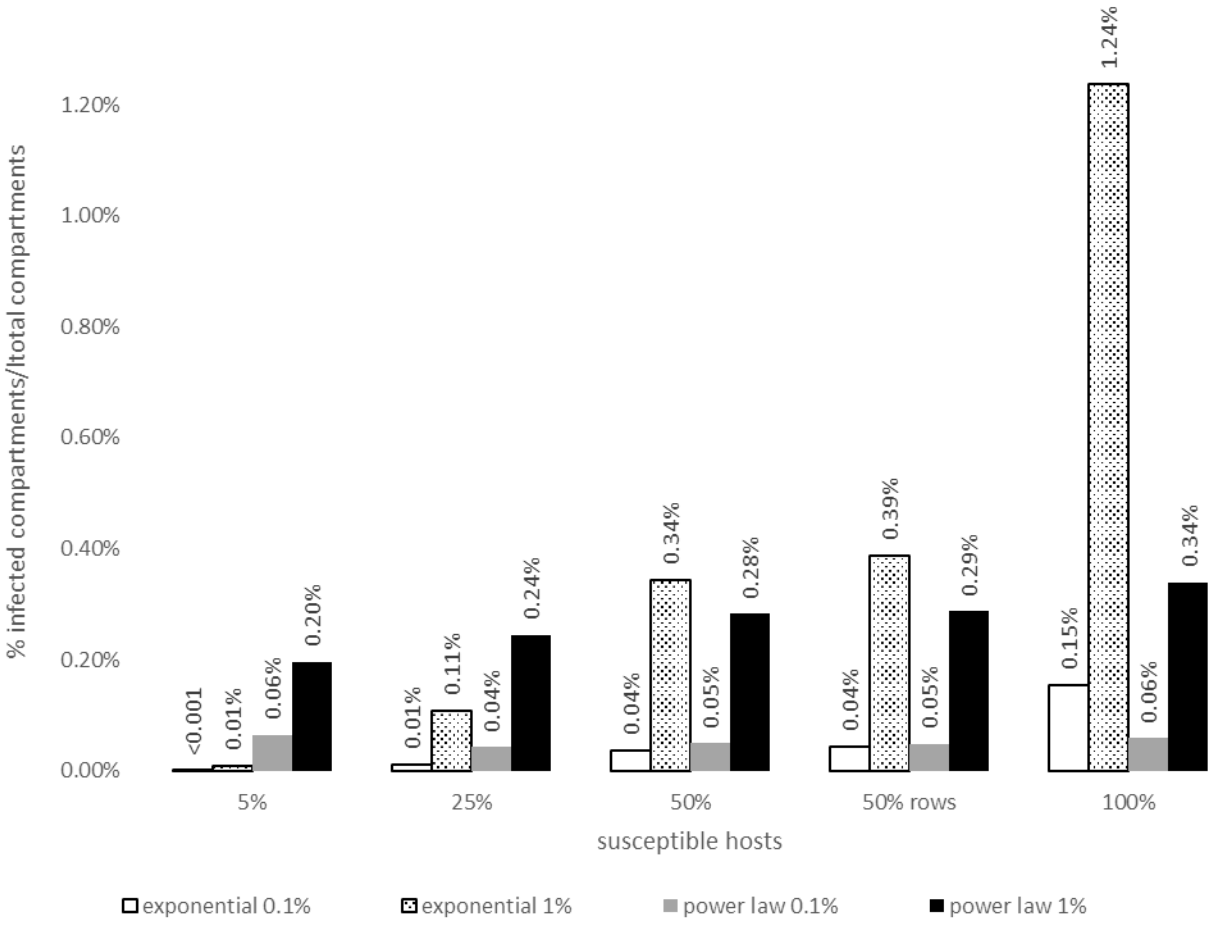


Figure 3.4. Bar chart of percent infected compartments per total infectible compartments (AUDG) for each field type and initial disease level.

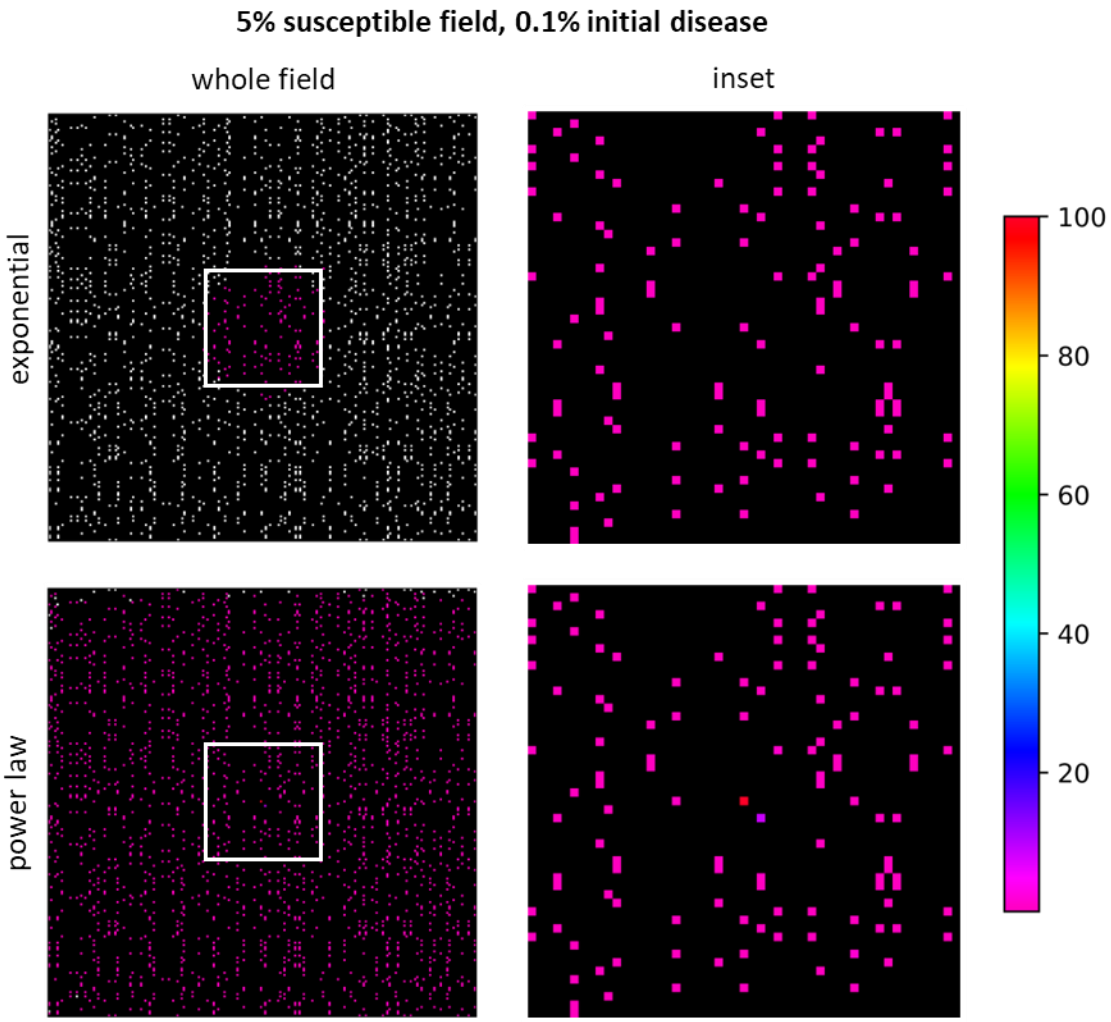


Figure 3.5. 200 x 200 cell field visualizing simulation results for the final day of the 5% susceptible field, 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

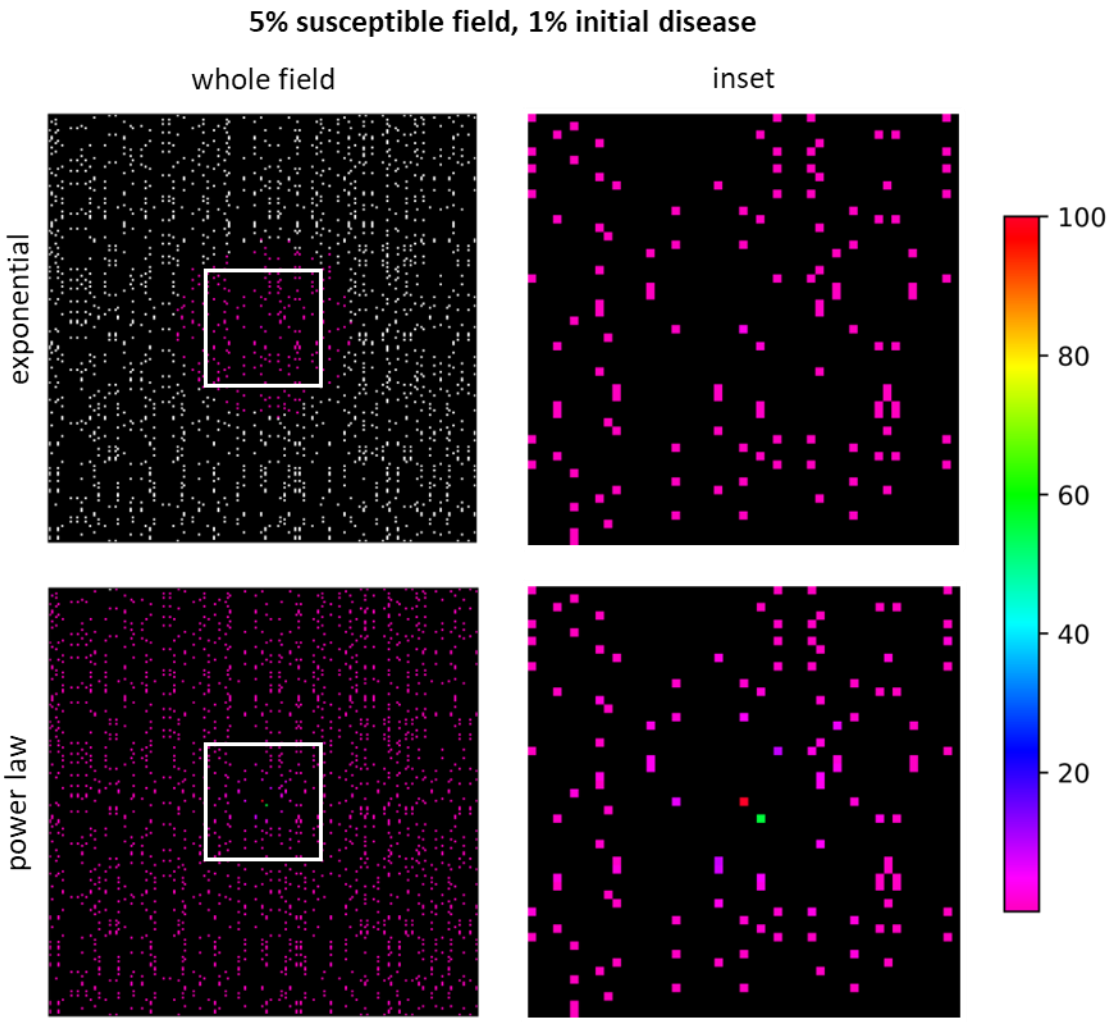


Figure 3.6. 200 x 200 cell field visualizing simulation results for the final day of the 5% susceptible field, 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

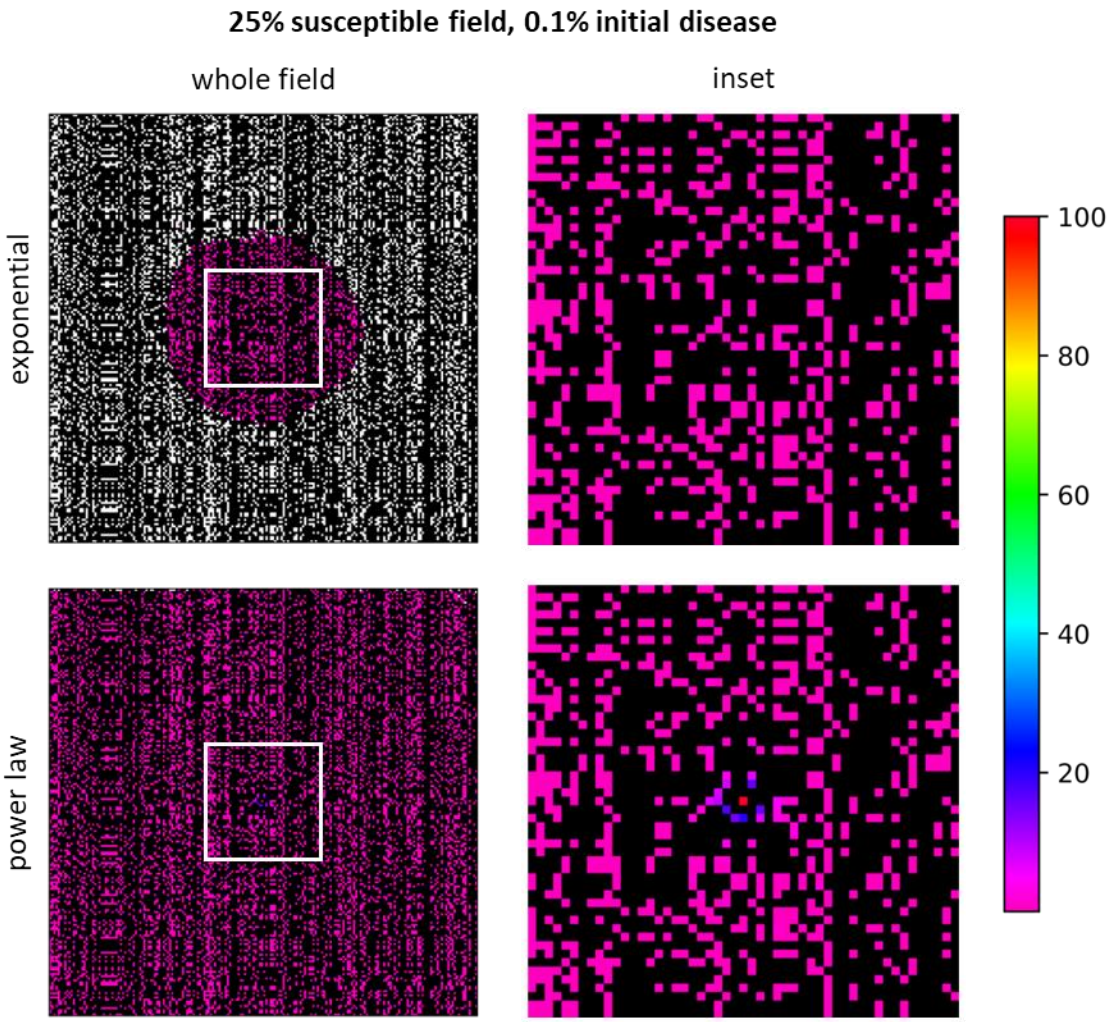


Figure 3.7. 200 x 200 cell field visualizing simulation results for the final day of the 25% susceptible field, 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

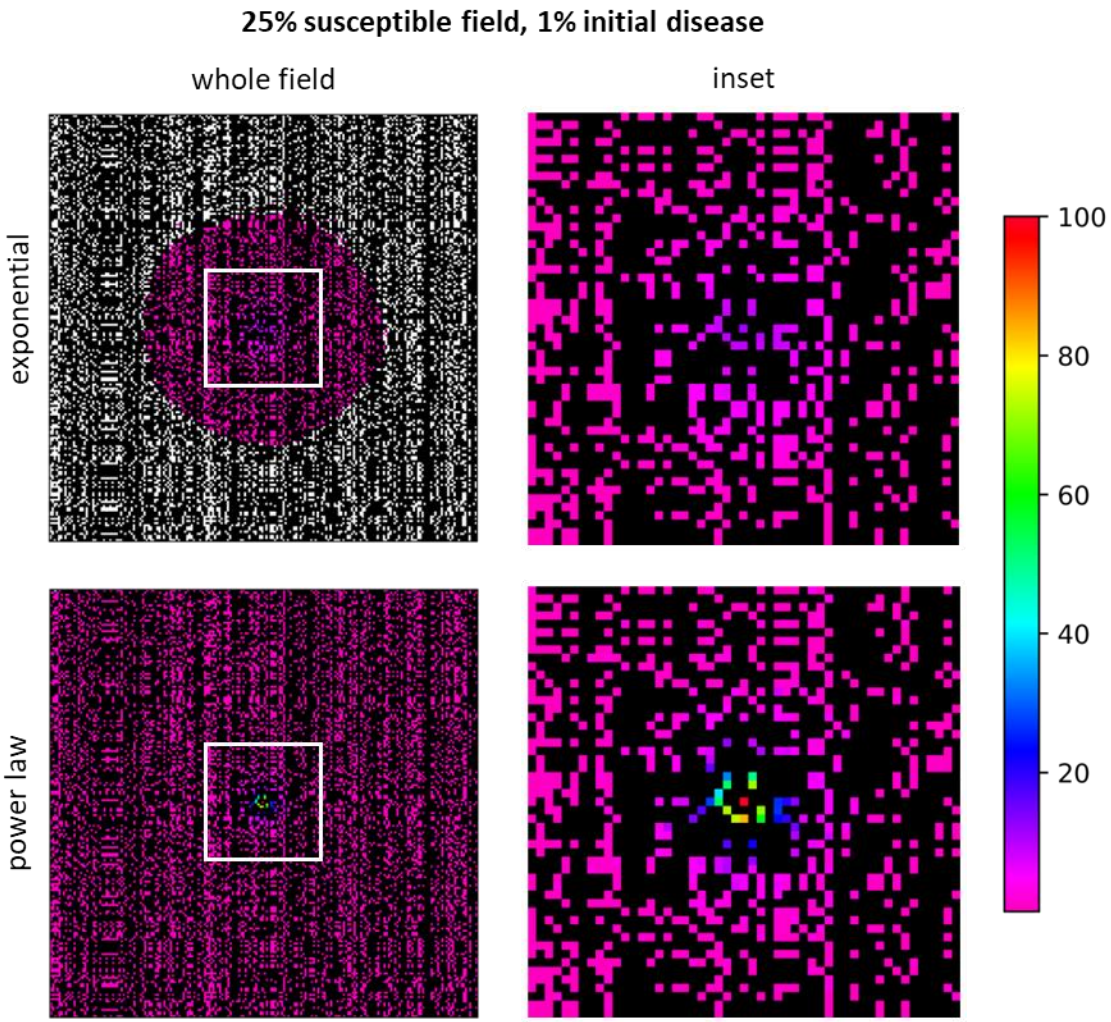


Figure 3.8. 200 x 200 cell field visualizing simulation results for the final day of the 25% susceptible field, 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

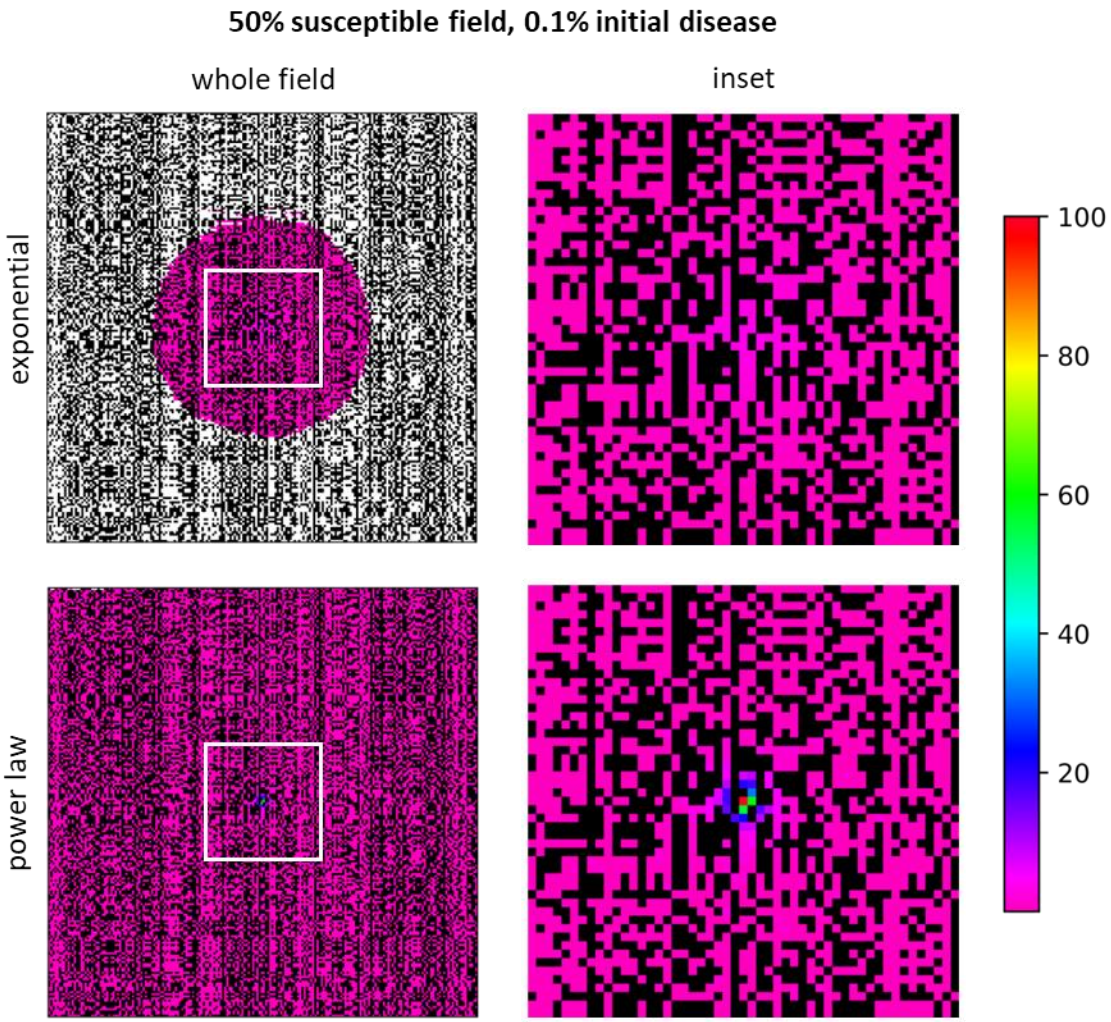


Figure 3.9. 200 x 200 cell field visualizing simulation results for the final day of the 50% susceptible field, 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

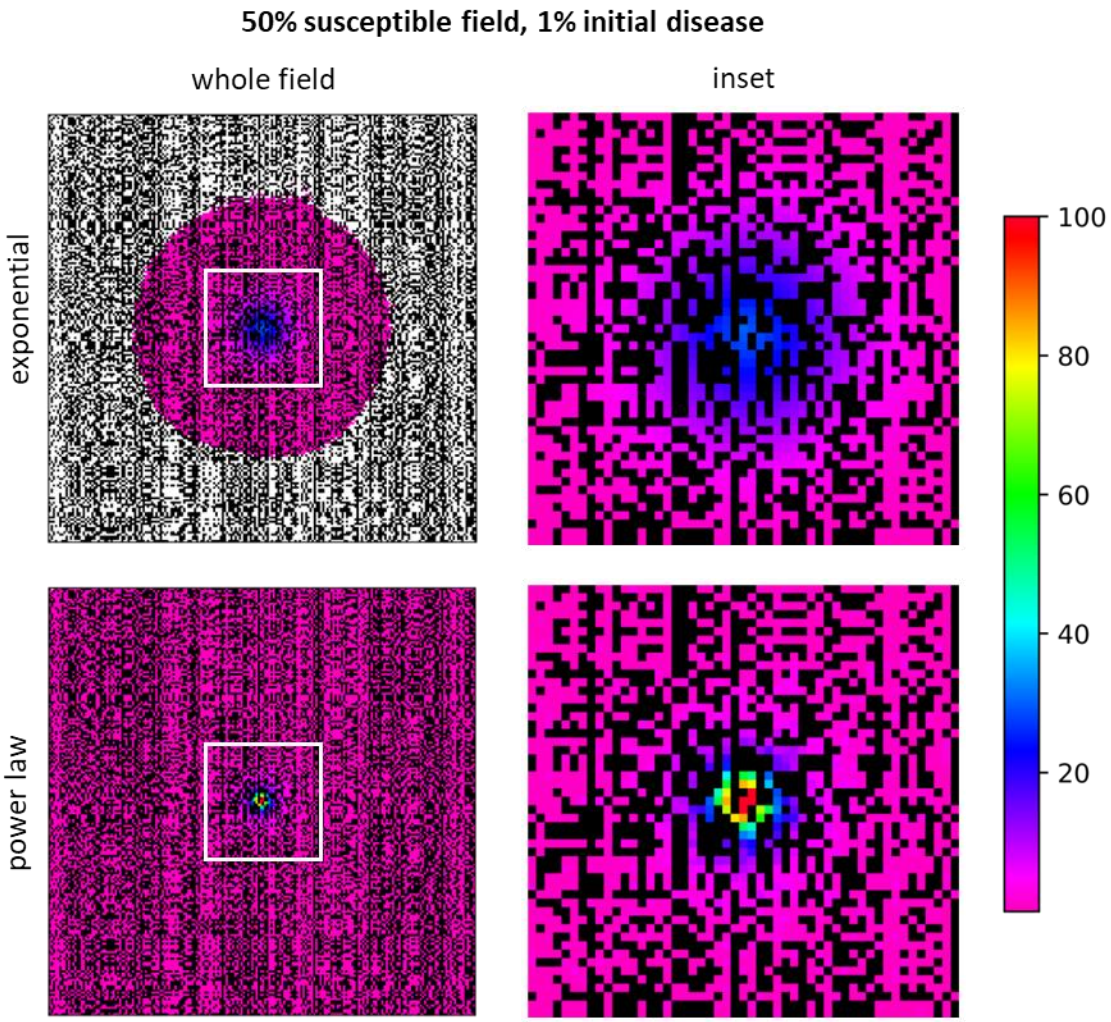


Figure 3.10. 200 x 200 cell field visualizing simulation results for the final day of the 50% susceptible field, 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

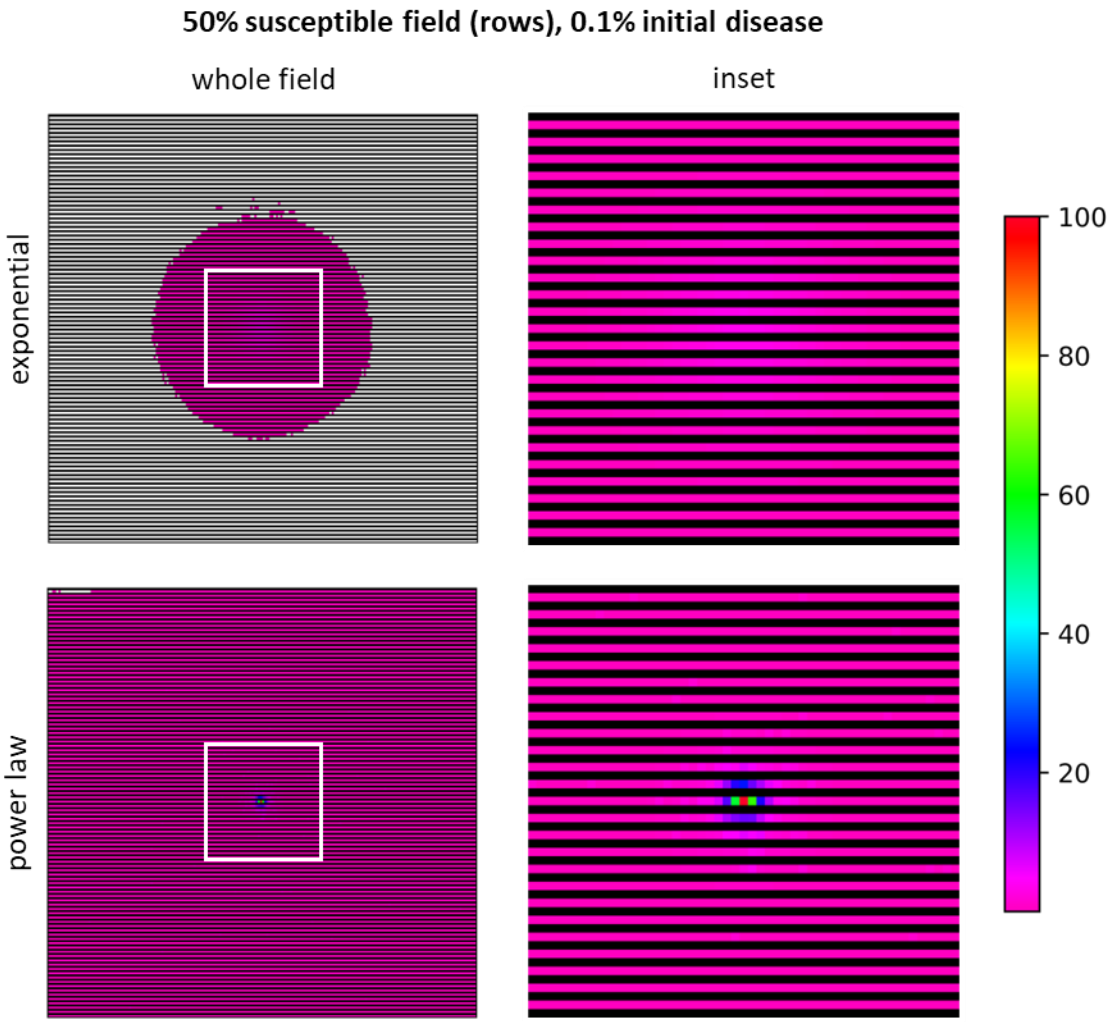


Figure 3.11. 200 x 200 cell field visualizing simulation results for the final day of the 50% susceptible field (rows), 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

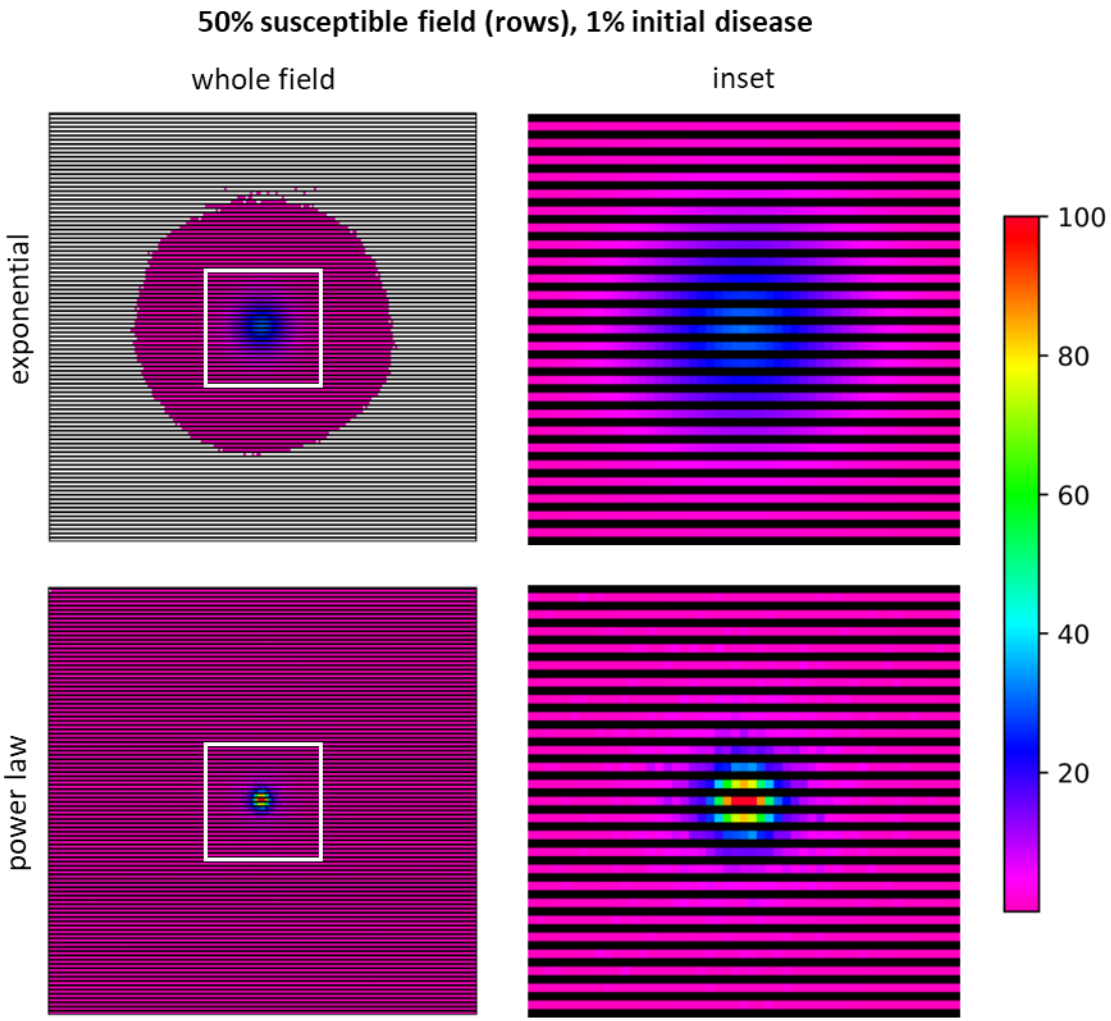


Figure 3.12. 200 x 200 cell field visualizing simulation results for the final day of the 50% susceptible field (rows), 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero, while black cells represent ND (non-host cell).

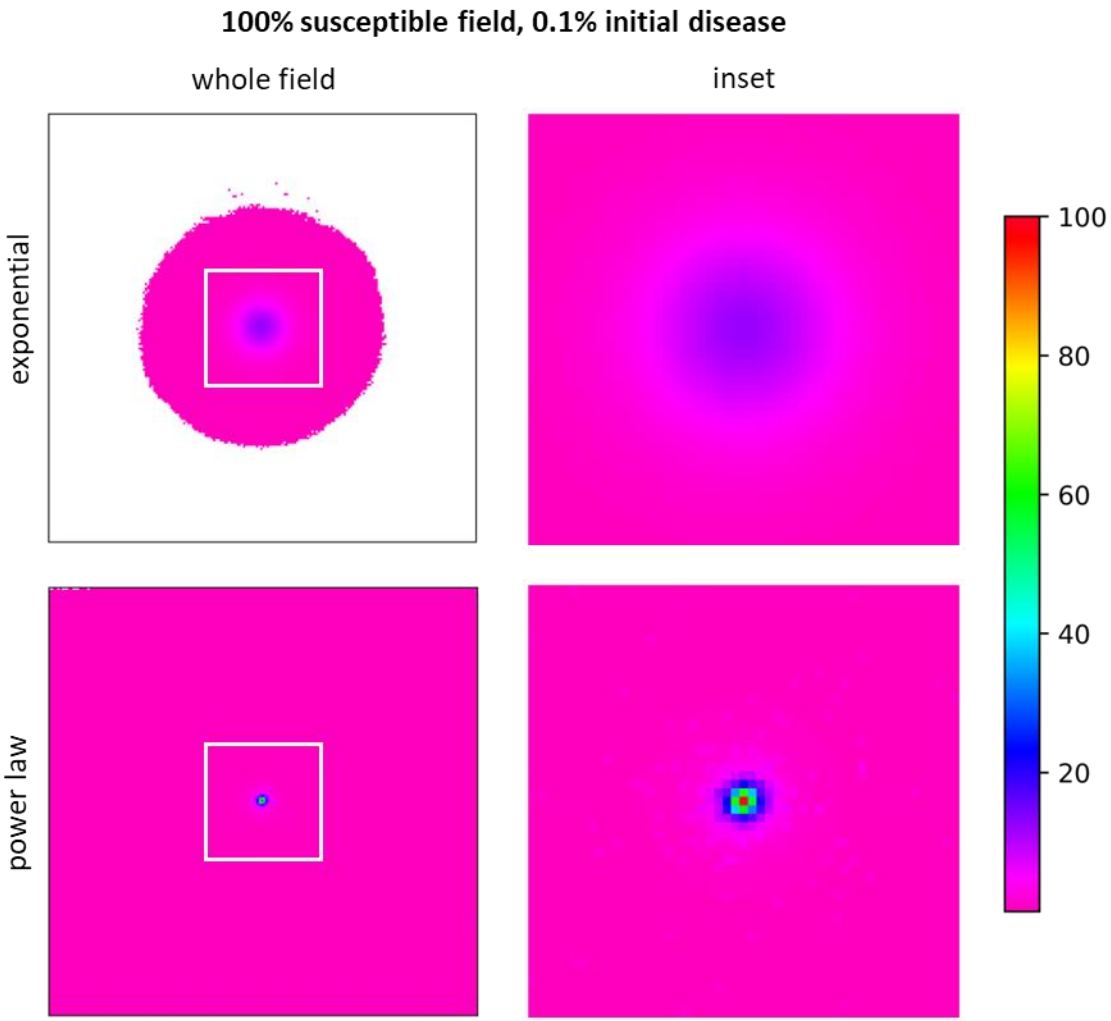


Figure 3.13. 200 x 200 cell field visualizing simulation results for the final day of the 100% susceptible field, 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero.

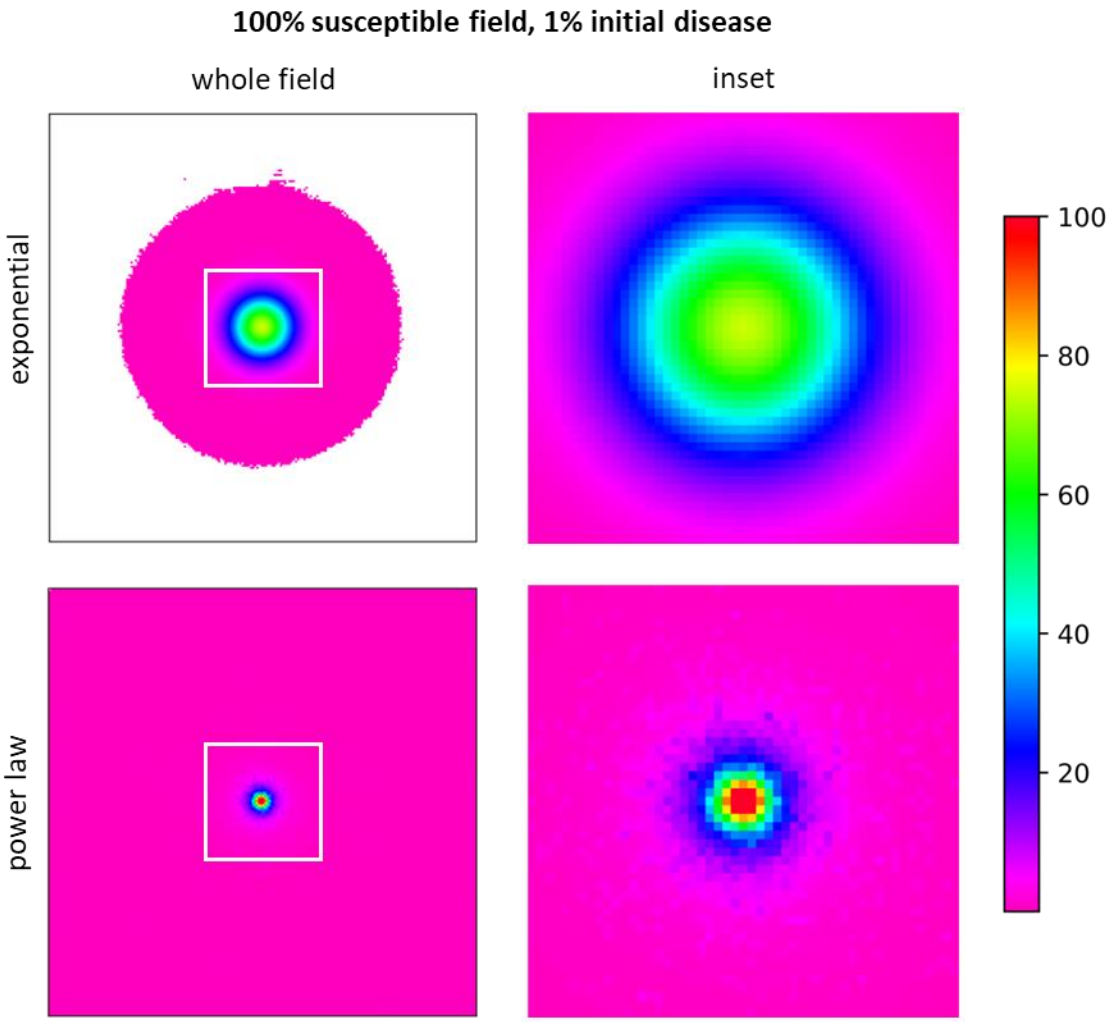


Figure 3.14. 200 x 200 cell field visualizing simulation results for the final day of the 100% susceptible field, 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the median values across all 100 stochastic simulations for exponential (top row) and power law (bottom row) dispersal gradients – minimum and maximum values were negligibly different from the median and were not included. White boxes in whole field images represent inset extent, highlighted to show differences in infection at the focus. White cells represent a value of zero.

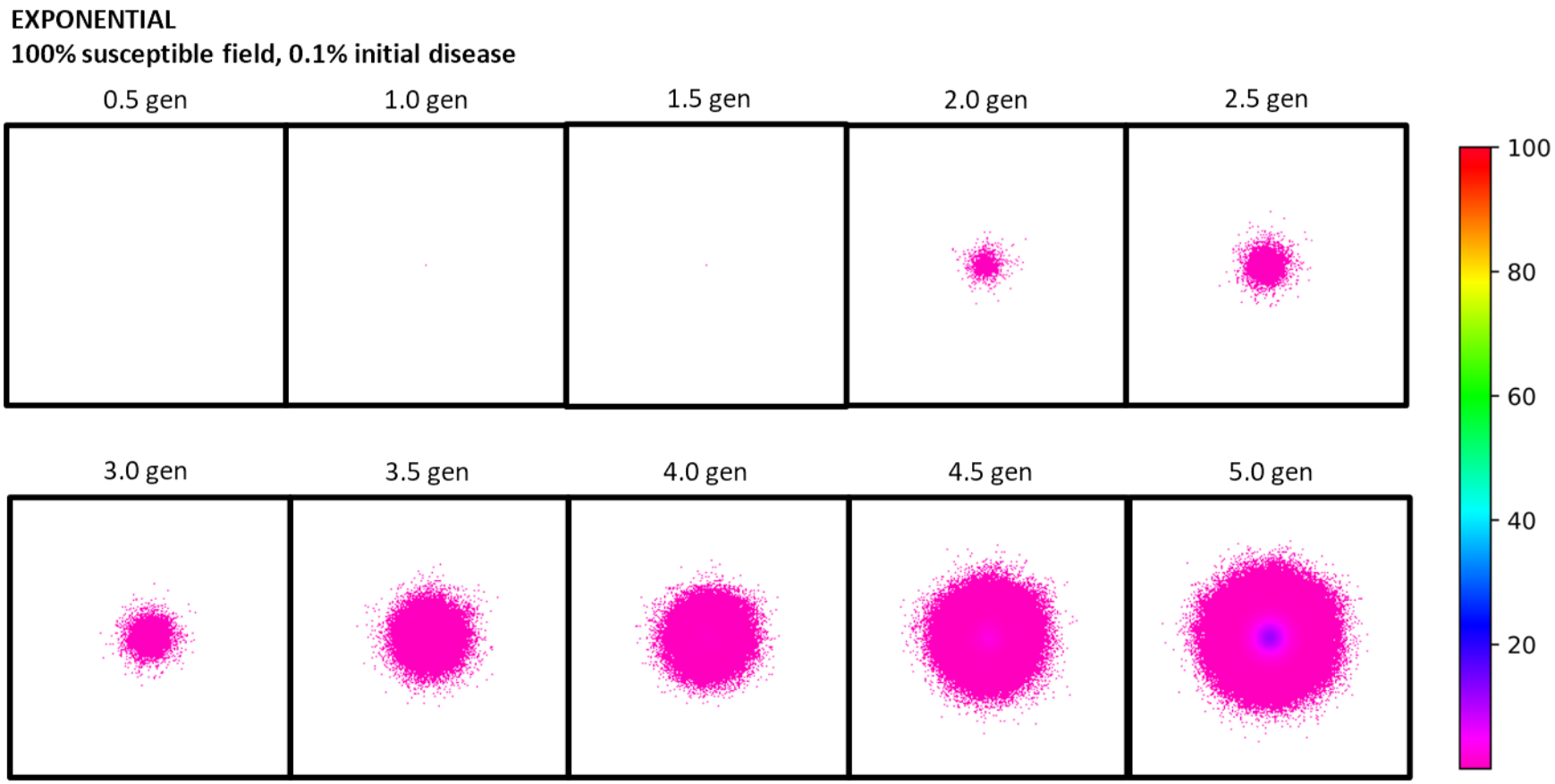


Figure 3.15. 200 x 200 cell field visualizing timeseries simulation results for the final day of the exponential, 100% susceptible field, 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the progression of epidemic extent and severity for each 0.5 generation (5 days) of the simulation (5 generations, 51 days).

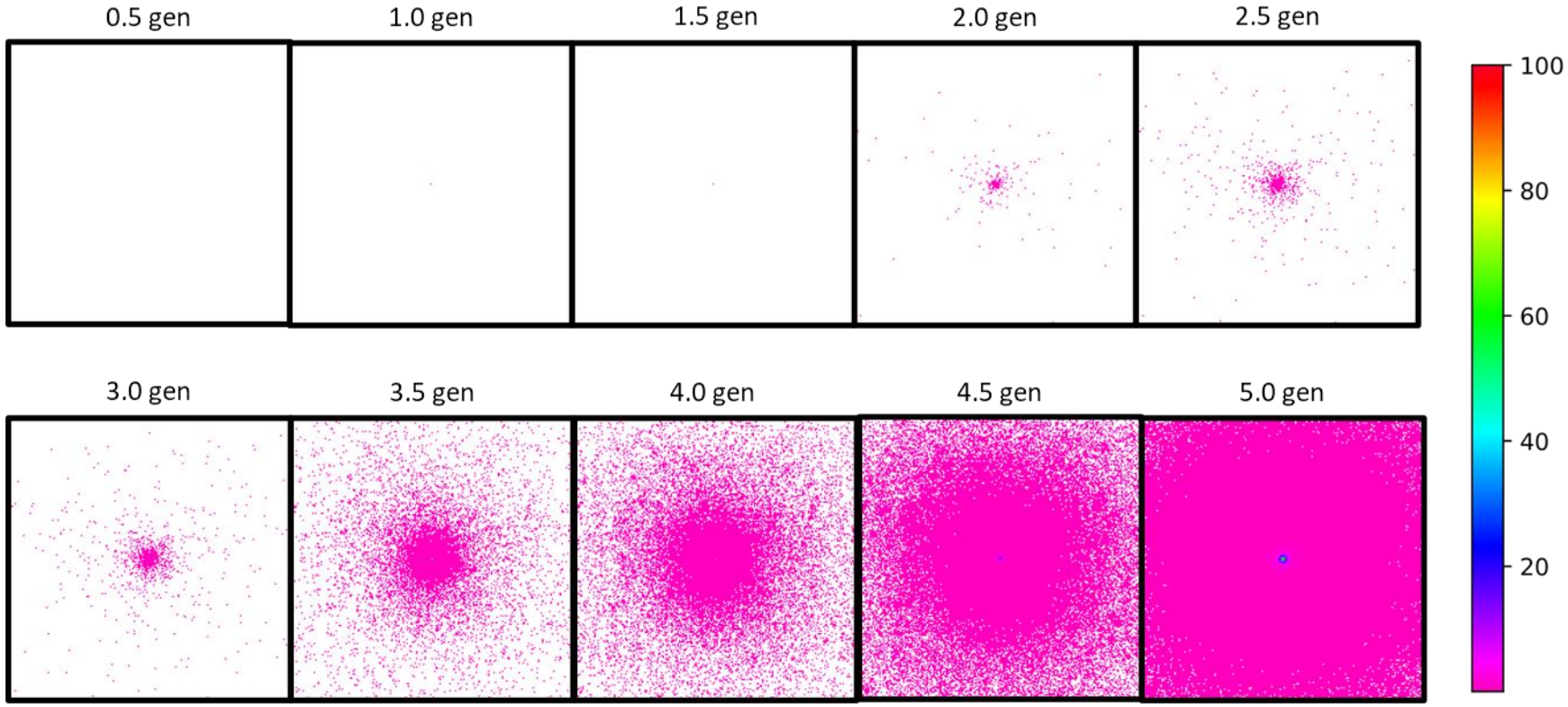


Figure 3.16. 200 x 200 cell field visualizing timeseries simulation results for the final day of the power law, 100% susceptible field, 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the progression of epidemic extent and severity for each 0.5 generation (5 days) of the simulation (5 generations, 51 days).

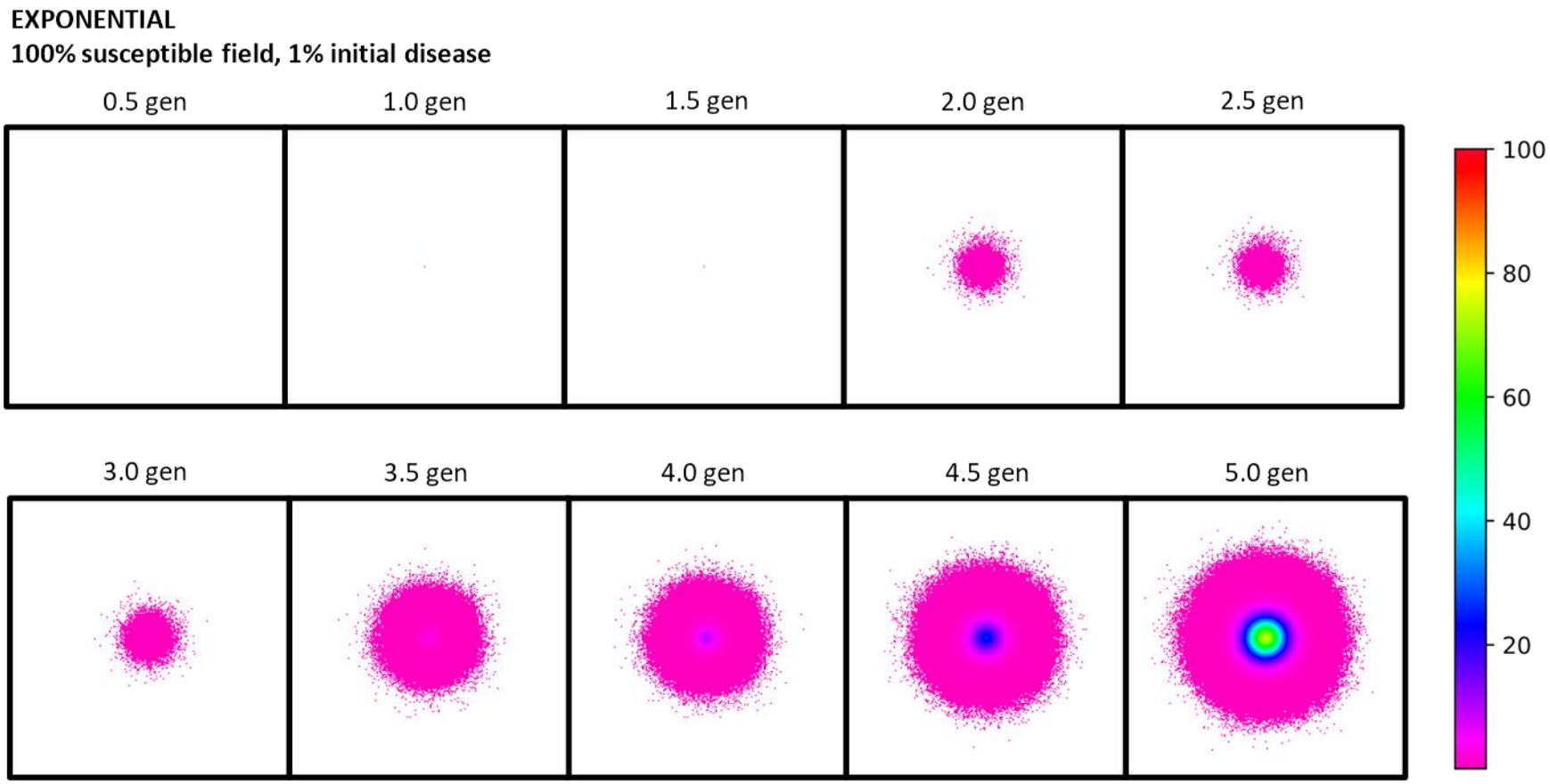


Figure 3.17. 200 x 200 cell field visualizing simulation results for the final day of the exponential, 100% susceptible field, 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the progression of epidemic extent and severity for each 0.5 generation (5 days) of the simulation (5 generations, 51 days).

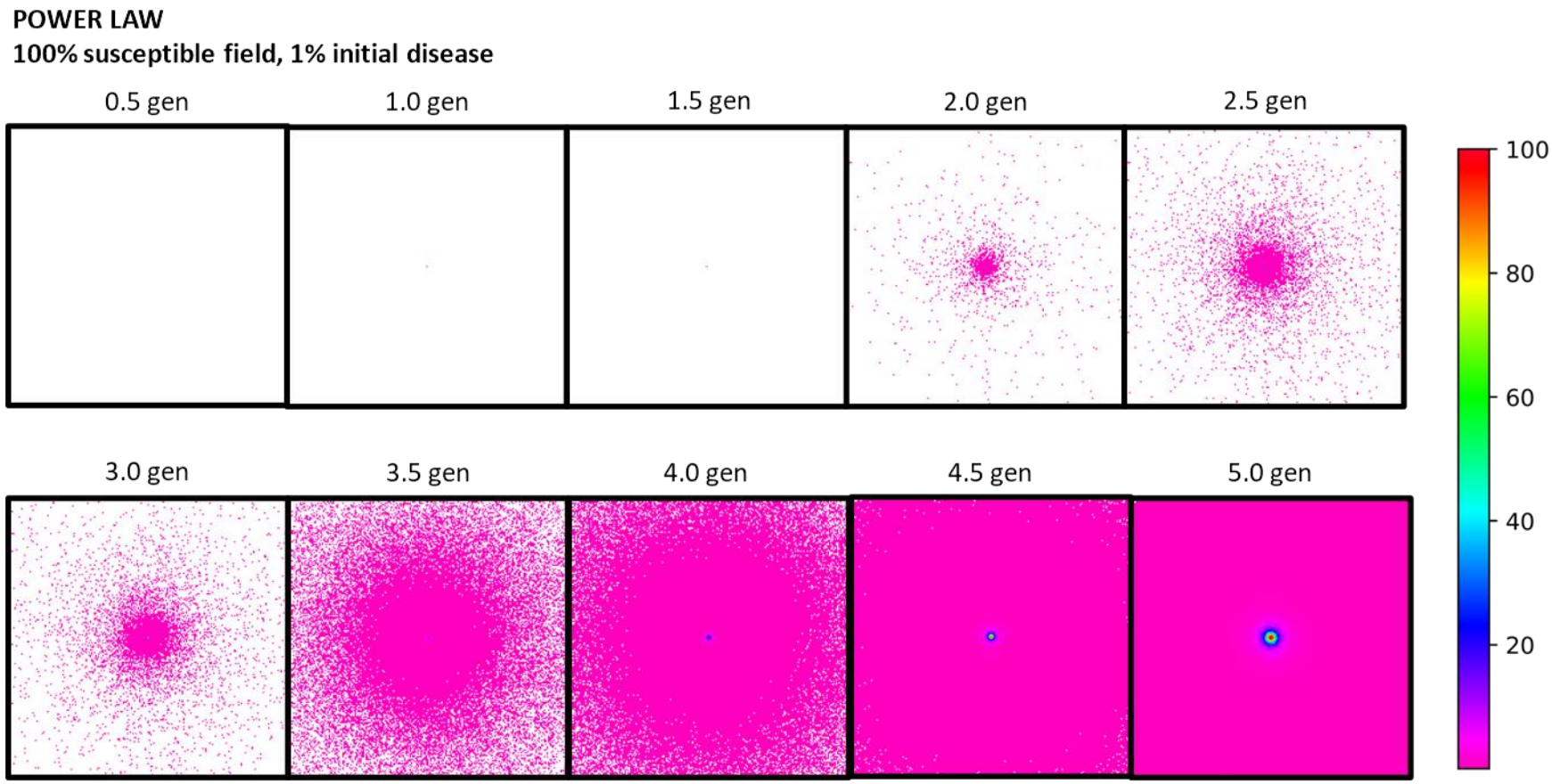


Figure 3.18. 200 x 200 cell field visualizing timeseries simulation results for the final day of the power law, 100% susceptible field, 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the progression of epidemic extent and severity for each 0.5 generation (5 days) of the simulation (5 generations, 51 days).

100% susceptible field, 0.1% initial disease

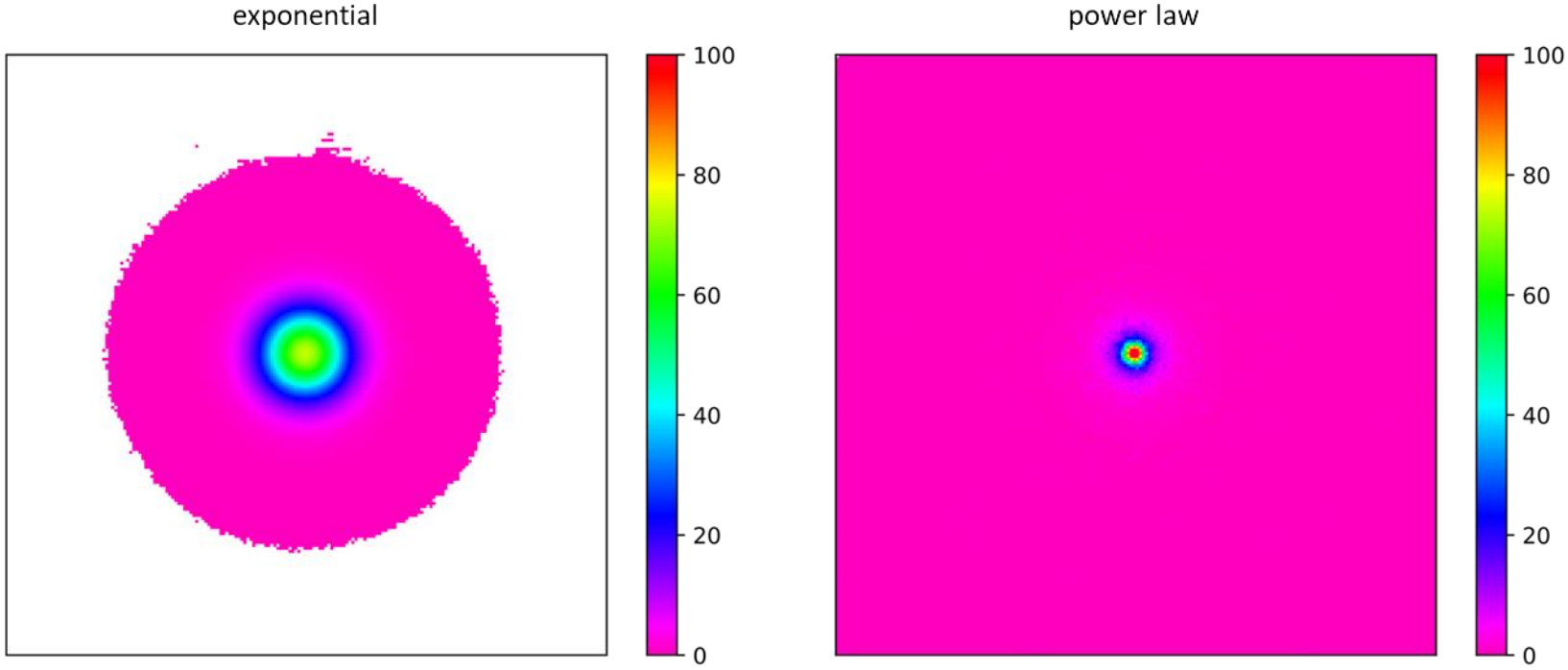


Figure 3.19. 1,000 x 1,000 cell field visualizing simulation results for the final day of the 100% susceptible field, 0.1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the results of deterministic simulations for exponential (left) and power law (right) dispersal gradients. White cells represent a value of 0.

100% susceptible field, 1% initial disease

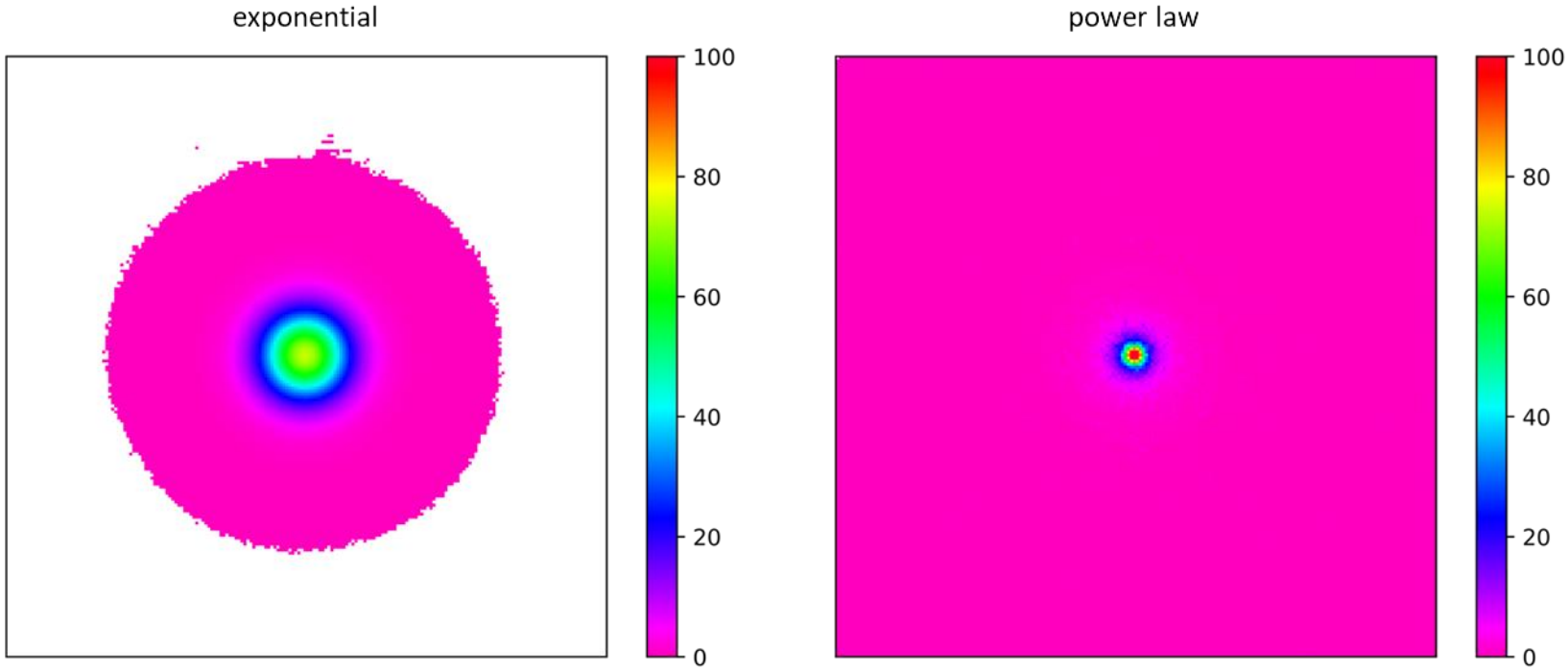


Figure 3.20. 1,000 x 1,000 cell field visualizing simulation results for the final day of the 100% susceptible field, 1% initial disease. The scale bar represents % infection per cell from >0 to 100%. Images represent the results of deterministic simulations for exponential (left) and power law (right) dispersal gradients. White cells represent a value of 0.

Literature Cited

Antonovics, J., and Alexander, H. M. 1992. Epidemiology of anther-smut infection of *Silene alba (= S. latifolia)* caused by *Ustilago violacea*: patterns of spore deposition in experimental populations. Proceedings of the Royal Society B. 250:157–163.

Bridgman, L. J., Benitez, V. V., Graña Grilli, M., Mufato, N., Acosta, D., and Guichón, M. L. 2012. Short perceptual range and yet successful invasion of a fragmented landscape: the case of the red-bellied tree squirrel (*Callosciurus erythraeus*) in Argentina. Landscape Ecology. 27:633–640.

Brooks-Pollock, E., Roberts, G. O., and Keeling, M. J. 2015. A dynamic model of bovine tuberculosis spread and control in Great Britain. Nature. 511:228–231.

Bullock, J. M., and Clarke, R. T. 2000. Long distance seed dispersal by wind: measuring and modelling the tail of the curve. Oecologia. 124:506–521.

Burdon, J. J., and Thrall, P. H. 2014. What have we learned from studies of wild plant-pathogen associations? – the dynamic interplay of time, space and life-history. European Journal of Plant Pathology. 138:417–429.

Burdon, J. J., Jarosz, A. M., and Kirby, G. C. 1989. Pattern and patchiness in plant-pathogen interactions – causes and consequences. Annual Review of Ecology and Systematics. 20:119–136.

Chen, X., Moore, M., Milus, E. A., Long, D. L., Line, R. F., Marshall, D., et al. 2002. Wheat stripe rust epidemics and races of *Puccinia striiformis* f. sp. *tritici* in the United States in 2000. Plant Disease. 86:39–46.

Chen, W. Q., Wu, L. R., Liu, T. G., Xu, S. C., Jin, S. L., Peng, Y. L., et al. 2009. Race dynamics, diversity, and virulence evolution in *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust in China from 2003 to 2007. Plant Disease. 93:1093–1101.

Chis Ster, I., and Ferguson, N. M. 2007. Transmission parameters of the 2001 foot and mouth epidemic in Great Britain. PLoS ONE. 2:e502.

Clark, J. S., Lewis, M., and Horvath, L. 2001. Invasion by extremes: population spread with variation in dispersal and reproduction. The American Naturalist. 157:537–554.

Clark, J. S., Lewis, M., McLachlan, J. S., and HilleRisLambers, J. 2003. Estimating population spread: what can we forecast and how well? Ecology. 84:1979–1988.

Clark, C. J., Poulsen, J. R., Bolker, B. M., Connor, E. F., and Parker, V. T. 2005. Comparative seed shadows of bird, monkey, and wind-dispersed trees. Ecology. 86:1684–2694.

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

Cowger, C., Wallace, L. D., and Mundt, C. C. 2005. Velocity of spread of wheat stripe rust epidemics. Phytopathology. 95:972–982.

Cummings, D. A. T., Irizarry, R. A., Huang, N. E., Endy, T. P., Nisalak, A., Ungchusak, K., et al. 2004. Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature. 427:344–347.

Cunniffe, N. J., Stutt, R. O. J. H., DeSimone, R. E., Gottwald, T. R., and Gilligan, C. A. 2015. Optimizising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. PLoS Computational Biology. 11:e1004211.

Dybiec, B., Kleczkowski, A., and Gilligan, C. A. 2005. Optimising control of disease spread on networks. Acta Physica Polonica. B. 36.

Emge, R. G., Kingsolver, C. H., and Johnson, D. R. 1975. Growth of the sporulating zone of *Puccinia striiformis* and its relationship to stripe rust epiphytology. Phytopathology. 65:679–681.

Finckh, M. R., Gacek, E. S., Czembor, H. J., and Wolfe, M. S. 1999. Host frequency and density effects on powdery mildew and yield in mixtures of barley cultivars. Plant Pathology. 48:807–816.

Frantzen, J., and van den Bosch, F. 2000. Spread of organisms: can travelling and dispersive waves be distinguished? Basic and Applied Ecology. 1:83–91.

Fujiwara, M., Anderson, K. E., Neubert, M. G., and Caswell, H. 2006. On the estimation of dispersal kernels from individual mark-recapture data. Environmental and Ecological Statistics. 13:183–197.

Garrett, K. A., and Mundt, C. C. 2000. Host diversity can reduce potato late blight severity for focal and general patterns of primary inoculum. Phytopathology. 90:1307–1312.

Gisiger, T. 2001. Scale invariance in biology: coincidence or footprint of a universal mechanism? Biological Reviews. 76:161–209.

Grenfell, B. T., Bjornstad, O. N., and Knappey, J. 2001. Travelling waves and spatial hierarchies in measles epidemics. Nature. 414:716–723.

Higgins, S. I., Clark, J. S., Nathan, R., Hovestadt, T., Schurr, F., Fragoso, J. M. V., et al. 2003. Forecasting plant migration rates: managing uncertainty for risk assessment. Journal of Ecology. 91:341–347.

Ismail, S. A., Ghazoul, J., Ravikanth, G., Kushalappa, C. G., Uma Shaanker, R., and Kettle, C. J. 2017. Evaluating realized seed dispersal across fragmented tropical landscapes: a two-fold approach using percentage analysis and the neighborhood model. New Phytologist. 214:1307–1316.

Iwamura, T., Guzman-Holst, A., and Murray, K. A. 2020. Accelerating invasion potential of disease vector *Aedes aegypti* under climate change. Nature Communications. 11:1–10.

Kampmeijer, P. and Zadoks, J. 1977. EPIMUL, a simulator of foci and epidemics in mixtures of resistant and susceptible plants, mosaics and multilines. Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands.

Keeling, M. J., Woolhouse, M. E. J., Shaw, D. J., Matthews, L., Chase-Topping, M., Haydon, D. T., et al. 2001. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science. 294:813–817.

Kiyosawa, S., and Shiyomi, M. 1972. A theoretical evaluation of the effect of mixing resistant variety with susceptible variety for controlling plant diseases. Japanese Journal of Phytopathology. 38:41–51.

Klein, E. K., Lavigne, C., Herve, P., Renard, M., and Gouyon, P.-H. 2006. Pollen dispersal of oilseed rape: estimation of the dispersal function and effects of field dimension. Journal of Applied Ecology. 43:141–151.

Lambert, D. H., Villareal, R. L., and Mackenzie, D. R. 1980. A general model for gradient analysis. Phytopathology. 98:150–154.

Lannou, C., De Vallavieille-Pope, C., and Goyeau, H. 1994. Host mixture efficacy in disease control: effects of lesion growth analyzed through computer-simulated epidemics. Plant Pathology. 43:651–662.

Li, Z. and Zeng, S. 2002. Wheat rusts in China. *China Agricultural, Beijing* (In Mandarin).

Lindström, T., Håkansson, N., and Wennergren, U. 2011. The shape of the spatial kernel and its implications for biological invasions in patchy environments. Proceedings of the Royal Society B. 278:1564–1571.

Luo, Y., and Zeng, S. M. 1995. Simulation studies on epidemics of wheat stripe rust (*Puccinia striiformis*) on slow-rusting cultivars and analysis of effects of resistance components. Plant Pathology. 44:340–349.

Madden, L. V., Hughes, G., and van den Bosch, F. 2017. Spatial aspects of epidemics – I: Pathogen dispersal and disease gradients. In The Study of Plant Disease Epidemics, p. 173–209.

Maffia, L. A., and Berger, R. D. 1999. Models of plant disease epidemics. II: Gradients of bean rust. Journal of Phytopathology. 147:199–206.

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

McConkey, K. R., Prasad, S., Corlett, R. T., Campos-Aceiz, A., Brodie, J. F., Rogers, H., et al. 2012. Seed dispersal in changing landscapes. Biological Conservation. 146:1–13.

Meentemeyer, R. K., Cunniffe, N. J., Cook, A. R., Filipe, J. A. N., Hunter, R. D., Rizzo, D. M., et al. 2011. Epidemiological modeling of invasion in heterogenous landscapes: spread of sudden oak death in California (1990-2030). Ecosphere. 2:1–24.

Meentemeyer, R. K., Haas, S. E., and Václavík, T. 2012. Landscape epidemiology of emerging infectious diseases in natural and human-altered ecosystems. Annual Review of Phytopathology. 50:379–402.

Minor, E. S., and Gardner, R. H. 2011. Landscape connectivity and seed dispersal characteristics inform the best management strategy for exotic plants. Ecological Applications. 21:739–749.

Mundt, C. C., and Leonard, K. J. 1985. A modification of Gregory's model for describing plant disease gradients. Phytopathology. 75:930–935.

Mundt, C. C., Sackett, K. E., Wallace, L. D., Cowger, C., and Dudley, J. P. 2009. Aerial dispersal and multiple-scale spread of epidemic disease. EcoHealth. 6:546–552.

Mundt, C. C., Sackett, K. E., and Wallace, L. D. 2011. Landscape heterogeneity and disease spread: experimental approaches with a plant pathogen. Ecological Applications. 21:321–328.

Mundt, C. C., Wallace, L. D., Allen, T. W., Hollier, C. A., Kemerait, R. C., and Sikora, E. J. 2013. Initial epidemic area is strongly associated with the yearly extent of soybean rust spread in North America. Biological Invasions. 15:1431–1438.

Nathan, R., Klein, E., Robledo-Arnuncio, J. J., and Revilla, E. 2012. Dispersal kernels: review. In *Dispersal Ecology and Evolution* (J. Clobert, J. M. Bullock, T. G. Benton, and M. Baguette, Eds.), Oxford, UK: Oxford University Press.

Nesslage, G. M., Maurer, B. A., and Gage, S. H. 2007. Gypsy moth response to landscape structure differs from neutral model predictions: implications for invasion and monitoring. Biological Invasions. 9:585–595.

Ojiambo, P. S., Gent, D. H., Mehra, L. K., Christie, D., and Margarey, R. 2017. Focus expansion and stability of the spread parameter estimate of the power law model for dispersal gradients. PeerJ. 5:e3465.

Papastamati, K., and van den Bosch, F. 2007. The sensitivity of epidemic growth rate to weather variables, with an application to yellow rust on wheat. Phytopathology. 97:202–210.

Paradis, E., Baillie, S. R., and Sutherland, W. J. 2002. Modeling large-scale dispersal distances. Ecological Modelling. 151:279–292.

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

Peterson, E. K., Hansen, E. M., and Kanaskie, A. 2015. Temporal epidemiology of sudden oak death in Oregon. Phytopathology. 105:937–946.

Qandah, I. S., and Del Rio Mendoza, L. E. 2012. Modelling inoculum dispersal and Sclerotinia stem rot gradients in canola fields. Canadian Journal of Plant Pathology. 34:390–400.

Rieux, A., Soubeyrand, S., Bonnot, F., Klein, E. K., Ngando, J. E., Mehl, A., et al. 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:e103225.

Robledo-Arnuncio, J. J., and Gil, L. 2005. Patterns of pollen dispersal in a small population of *Pinus sylvestris L.* revealed by total-exclusion paternity analysis. Heredity. 94:13–22.

Ruxton, G. D., and Schaefer, H. M. 2012. The conservation physiology of seed dispersal. Philosophical Transactions of the Royal Society B: Biological Sciences. 367:1708–1718.

Sackett, K. and Mundt, C. 2005a. Primary disease gradients of wheat stripe rust in large field plots. Phytopathology 95: 983-991.

Sackett, K. E., and Mundt, C. C. 2005b. The effects of dispersal gradient and pathogen life cycle components on epidemic velocity in computer simulations. Phytopathology. 95:992–1000.

Severns, P. M., and Mundt, C. C. 2022. Delays in epidemic outbreak control cost disproportionately large treatment footprints to offset. Pathogens. 11:393.

Severns, P. M., Estep, L. K., Sackett, K. E., and Mundt, C. C. 2014. Degree of host susceptibility in the initial disease outbreak influences subsequent epidemic spread. Journal of Applied Ecology. 51:1622–1630.

Severns, P. M., Sackett, K. E., and Mundt, C. C. 2015. Outbreak propagule pressure influences the landscape spread of a wind-dispersed, epidemic-causing, plant pathogen. Landscape Ecology. 30:2111–2119.

Severns, P. M., Sackett, K. E., Farber, D. H., and Mundt, C. C. 2019. Consequences of long-distance dispersal for epidemic spread: patterns, scaling, and mitigation. Plant Disease. 103:177–191.

Shaner, G., and Powelson, R. L. 1971. Epidemiology of stripe rust of wheat, 1961-1968. Oregon Agricultural Experimental Station Bulletin. 117.

Skelsey, P., With, K. A., and Garrett, K. A. 2013. Why dispersal should be maximized at intermediate scales of heterogeneity. Theoretical Ecology. 6:203–211.

Taylor, P. D., Fahrig, L., Henein, K., and Merriam, G. 1993. Connectivity is a vital element of landscape structure. Oikos. 68:571–573.

Tildesley, M. J., Smith, G., and Keeling, M. J. 2012. Modeling the spread and control of foot-and-mouth disease in Pennsylvania following its discovery and options for control. Preventive Veterinary Medicine. 104:224–239.

Tollenaar, H., and Houston, B. R. 1967. A study on the epidemiology of stripe rust, *Puccinia striiformis* West., in California. Canadian Journal of Botany. 45:291–307.

Travadon, R., Bousset, L., Saint-Jean, S., Brun, H., and Sache, I. 2007. Splash dispersal of *Leptosphaeria maculans* pycnidiospores and the spread of blackleg on oilseed rape. Plant Pathology. 56:595–603.

van den Bosch, F., Frinking, H. D., Metz, J. A. J., and Zadoks, J. C. 1988. Focus expansion in plant disease. III: Two experimental examples. Phytopathology 78:919-925.

van Houtan, K. S., Pimm, S. L., Halley, J. M., Bierregaard Jr., R. O., and Lovejoy, T. E. 2007. Dispersal of Amazonian birds in continuous and fragmented forest. Ecology Letters. 10:219–229.

Wan, A., Zhao, Z., Chen, X., He, Z., Jin, S., Jia, Q., et al. 2004. Wheat stripe rust epidemic and virulence of *Puccinia striiformis* f. sp. *tritici* in China in 2002. Plant Disease. 88:896–904.

Wellings, C. R. 2007. *Puccinia striiformis* in Australia: a review of the incursion, evolution and adaptation of stripe rust in the period 1979-2006. Australian Journal of Agricultural Research. 58:567–575.

Wilson, M. C., Chen, X.-Y., Corlett, R. T., Didham, R. K., Ding, P., Holt, R. D., et al. 2016. Habitat fragmentation and biodiversity conservation: key findings and future challenges. Landscape Ecology. 31:219–227.

Wolfe, M. S. 1992. Barley diseases: maintaining the value of our varieties. Sixth International Barley Genetics Symposium. 2:1055–1067.

Zadoks, J. C., and van den Bosch, F. 1994. On the spread of plant disease: a theory on foci. Annual Review of Phytopathology. 32:503–521.

Zadoks, J. C. 1961. Yellow rust on wheat, studies in epidemiology and physiologic specialization. Tijdschrift over Plantenziekten. 67:69–256.

Zhang, C., Dong, Y., Tang, L., Zheng, Y., Makowski, D., Yu, Y., et al. 2019. Intercropping cereals with faba bean reduces plant disease incidence regardless of fertilizer input; a meta-analysis. European Journal of Plant Pathology. 154:931–942.

Zhu, Y., and Chen, H. 2000. Genetic diversity and disease control in rice. Nature. 406:718–722.

CHAPTER 4

CONCLUSIONS

In Chapter 2, I completed a literature review of the plant epidemiological literature to determine how often plant disease gradients were truncated. I was interested to see if there were any patterns changes in the published literature whether plant disease studies adopted methods to more accurately fit disease gradients if truncation was common following the publication of two influential studies in the mid-1990s. I was also interested in understanding whether disease dispersal studies were representative of major plant pathogen dispersal mechanisms or if some were underrepresented. I selected 104 published papers displaying figured dispersal gradients that were suitable for evaluation from more than 200 studies that reported disease gradients but did not present figured data. Of these 104 papers, 10 had to be excluded from formal review because their dispersal gradient figures were not interpretable for my purposes. In total, I found almost three quarters of the published gradients in the review were truncated. When I investigated patterns of truncation based on the year of publication, I found that while there was a decrease in the proportion of truncated studies after the year 2000, about five years following the publication of studies presenting strong evidence for dispersive waves of organism spread (Ferrandino 1993; Kot et al. 1996). However, the proportion of studies presenting truncated dispersal gradients for long-distance dispersed organisms was still close to two-thirds. Nearly all these truncated dispersal gradients were for wind dispersed fungal pathogens which are characterized by long-distance dispersal and consequently dispersive waves (Ferrandino 1993). As presented in these studies with truncated dispersal gradients, the dispersal gradient function

would inaccurately produce traveling waves which would underestimate the speed and distance of plant disease spread.

Splash dispersal was the next most common mode of dispersal, making up a quarter of the selected studies in the reviewed literature, and was likely well-characterized by exponentially bound dispersal gradients, but only in the absence of wind. Splash dispersal alone would generate a traveling wave (a constant rate) of disease spread, however in nature rain is often accompanied by wind. Splash dispersal studies were performed in an isolated environment, like a greenhouse, and the authors measured the effects of rainfall over certain time periods, raindrops of certain diameters, etc. and specifically excluded the interactions of other environmental factors such as wind. These studies often measured dispersal within 1-2 m of the inoculum source, which appeared sufficient to capture dispersal exclusively from water droplet impacts. While these studies do provide a greater understanding the dispersal physics associated with rain droplet splash, they do not have a larger utility in describing splash dispersal as it occurs in nature – a mode of dispersal highly linked to and primarily driven by wind.

Insects, flowing water, and humans accounted for the remaining dispersal gradients in my review. All three of these vectors are characterized by long-distance dispersal which, as with wind dispersal, will likely generate dispersive waves. While some of these insect, flowing water, and human vector studies tracked plant dispersal tens of meters from the source, most of the studies only measured dispersal within 10 m of the source, which is insufficient to capture LDD events.

The Willocquet et al. (2008) study of powdery mildew dispersal on strawberries in a glasshouse (closed) and tunnel grow systems (partially closed) was particularly notable for their detailed and reasonable characterization of the dispersal gradient. In their study, the authors

included and measured the temperature, humidity, and relative wind speed at each of their disease monitoring locations – thus including the influence of roof vents and the fan cooling system in their closed greenhouse setting. Willocquet et al. (2008) reported that while exponential and power law functions could represent their measured disease gradients, the power law provided a superior fit, particular at larger distances even within a greenhouse. The authors noted their results agreed with previous studies of aerially dispersed plant pathogens, which were best fit by power law gradients. Willocquet et al. (2008) also noted that the lack of epidemic development in their glasshouse plots was likely due to the reduced wind speeds (ten times less than in the tunnel greenhouses which favor a unidirectional flow of air). While in my review I categorized their gradients as truncated according to my categorization criteria, the authors still acknowledge the importance of wind – even in their protected system – and selected a fat-tailed dispersal gradient.

In Chapter 3, I evaluated the potential effects of dispersal gradient truncation using a spatially explicit, compartmental model of aerially dispersed fungal plant pathogens. I parameterized the EPIMUL simulation model to represent wheat stripe rust (WSR), a plant pathogen with well-studied and thoroughly characterized dispersal gradients and compared the simulated epidemics produced by exponential (thin-tailed and intentionally truncated) and power law (fat-tailed and appropriately fit) disease gradients in the exact same suite of virtual landscapes. The exponential dispersal gradients generated a travelling wave (as expected) while the power law gradients generated a dispersive wave (as expected). Truncation of the WSR dispersal gradient provided biologically unreasonable projections of disease spread when it is well-known that WSR is biologically and statistically more appropriately described by a power law dispersal gradient. However, the degree of susceptible host plant patchiness in the simulated

landscape over which disease was dispersed magnified this already profound effect of dispersal gradient truncation. Compared to the reasonably fit power law gradient, the truncated exponential predicted both a more restricted extent of disease dispersal and reduced amount of disease further away from the outbreak source; this effect was magnified as susceptible hosts became rarer across the virtual landscapes.

Recommendations

Decision trees and charts can be helpful in making management decisions of all kinds. In healthcare, decision trees are commonly used to determine the best treatment for clinical patients, ensuring evidence-based and reproducible decision-making by physicians (Bae 2014; Podgorelec et al. 2022). In plant pathology, decision support systems (DSS) are often used by extension agents, growers, and other plant disease managers to determine best pesticide application strategies, prioritize resource allocation, and/or improve disease and insect control (Trapman and Polfliet 1997; Morgan et al. 2000; Gent et al. 2013; Wallhead and Zhu 2017). Examples of computer-based DSS that integrate weather data include Ag-Radar (https://extension.umaine.edu/ipm/ag-radar-apple-sites/), the Network for Environment and Weather Applications (NEWA, http://newa.cornell.edu), and RIMpro (http://www.rimpro.eu/). There are also simpler decision-making systems like PeanutRx (https://peanuts.caes.uga.edu/extension/peanut-rx.html).

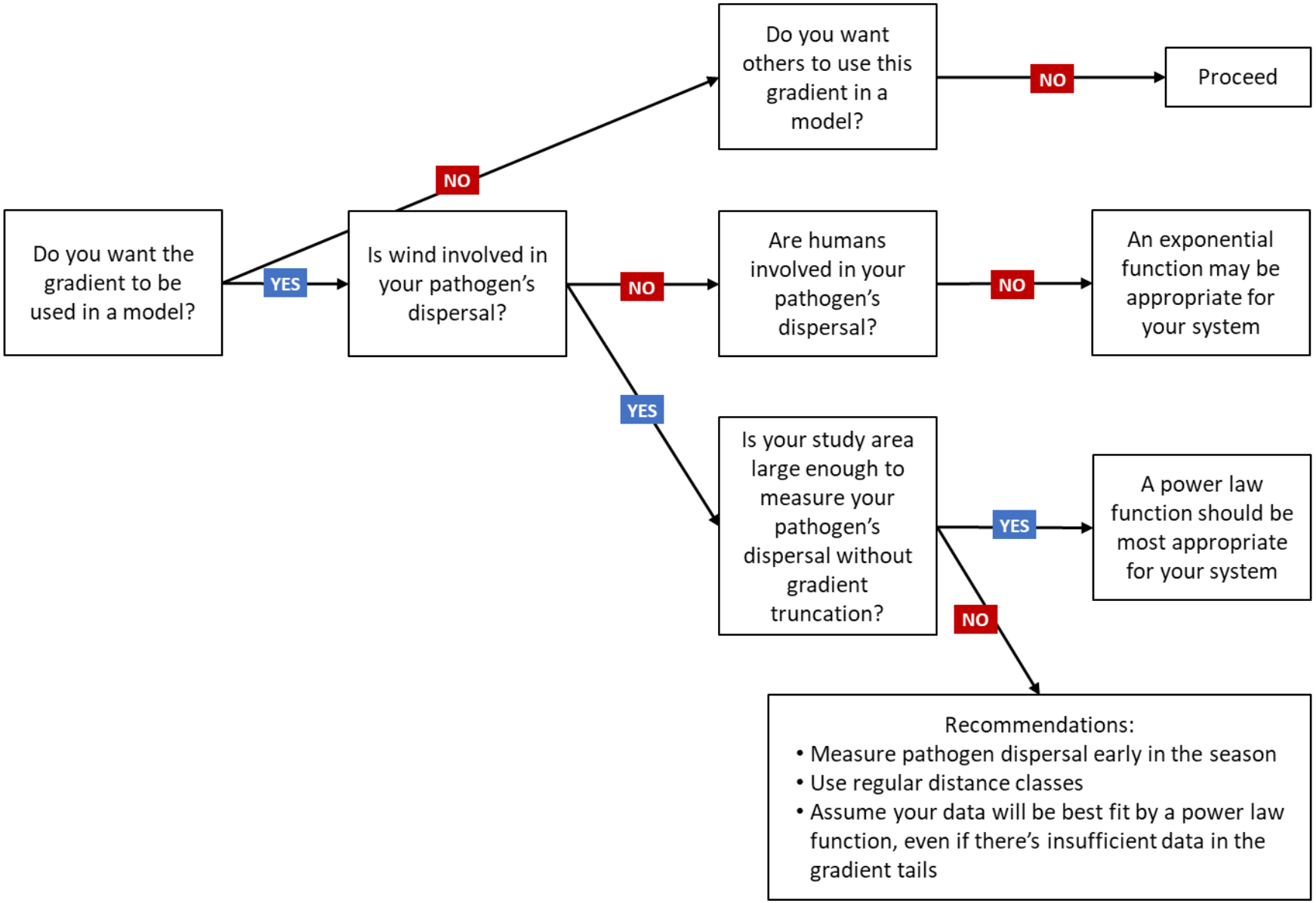
To help plant pathologists prevent dispersal gradient truncation, I have created a decision chart with recommendations (Figure 4.1). The first step of the decision chart asks if the dispersal gradient will be used in a model – if this is the case, it is critical that the gradient is measured correctly, as model performance is dependent on parameter accuracy. The next layers of the

decision chart ask about vector type; it is critical that researchers acknowledge any LDD vectors present in their system (*e.g.*, wind or human vectors), because if they are ignored, studies may not account for disease dispersed long distances from the source and mis-apply a truncated dispersal gradient. The final step in the decision tree asks if there is sufficient room to measure dispersal without truncation. As in the previous step, studies that are unable capture LDD events far from the source may mis-apply a truncated dispersal gradient to their system.

I have provided a series of recommendations to prevent the misapplication of a thin-tailed function. First, I encourage researchers to include points further from the source, especially to distances with zero disease observations because: 1) if disease is not followed to these distances the researcher is implicitly assuming that LDD does not exist and will bias descriptions towards a truncated disease gradient, and 2) by definition, LDD dispersed organisms sporadically occur in the tails of their distributions (Nathan et al. 2012). In my review, there was strong evidence that studies which had sampling locations that measured out to zero dispersal values were more reasonably fit than those that did not track disease farther from the source and were obviously truncated. In the scenario that a pre-planned study is constrained by field size, with limited distance to measure an LDD pathogen, I recommend tracking disease dispersal early in the outbreak when disease has not progressed as far. This allows researchers to use smaller areas for their experiments; however, even these early dispersal studies need to represent the tails (zero dispersal events) as a safeguard against truncating the dispersal gradient. I also recommend using regular distance classes to sample disease from the source to prevent spatial bias in the data – this was a common occurrence in the reviewed literature that led to improperly fit disease gradients. Finally, if the pathogen is wind vectored or invertebrate vectored that can be influenced by wind, then fit the dispersal gradient data to a distribution that is not bound by an

exponential (*i.e.* a power law function), even if there's limited data in the tails, as LDD pathogens will be underestimated by an exponential disease gradient.

Figure 4.1. Decision chart of recommendations for dispersal gradient studies.



Literature Cited

Bae, J.-M. 2014. The clinical decision analysis using decision tree. Epidemiology and Health. 36:e2014025.

Ferrandino, F. J. 1993. Dispersive epidemic waves: I. Focus expansion within a linear planting. Phytopathology. 83:795–802.

Gent, D. H., Mahaffee, W. F., McRoberts, N., and Pfender, W. F. 2013. The use and role of predictive systems in disease management. Annual Review of Phytopathology. 51:267–289.

Kot, M., Lewis, M. A., and van den Driessche, P. 1996. Dispersal data and the spread of invading organisms. Ecology. 77:2027–2042.

Morgan, D., Walters, K. F. A., Oakley, J. N., and Lane, A. 2000. An internet-based decision-support system for pests of rape. EPPO Bulletin. 30:155–158.

Podgorelec, V., Kokol, P., Stiglic, B., and Rozman, I. 2002. Decision trees: an overview and their use in medicine. Journal of Medical Systems. 26:445–463.

Trapman, M., and Polfliet, M. 1997. Management of primary infections of apple scab with the simulation program RIMpro: review of four years of field trials. IOBC-WPRS Bulletin. 20:241–250.

Wallhead, M., and Zhu, H. 2017. Decision support systems for plant disease and insect management in commercial nurseries in the Midwest: a perspective review. Journal of Environmental Horticulture. 35:84–92.

APPENDIX A

LITERATURE REVIEW SPREADSHEET

Appendix A contains a table of the literature reviewed in Chapter 2. Originally there were multiple for most of the reviewed articles, but they were reduced to a single entry per article for simplicity. As it is too large to fit on a single printed page, it has been divided by columns into four sections.

Table A.1. Literature Review: Title – Journal (1 of 2)

ID #	title	author(s)	year	journal
1	Spore dispersal and disease gradients in strawberry...	Willocquet et al.	2008	Canadian Journal of Plant Pathology
2	Development of potato late blight epidemics: disease...	Zwankhuizen et al.	1998	Phytopathology
3	Ascospore gradients of *Gibberella zeae* from overwintered...	de Luna et al.	2002	Canadian Journal of Plant Pathology
4	Estimation of dispersal gradients of S- and P-type...	Moykkynen et al.	1997	European Journal of forest pathology
5	Regional spatial-temporal spread of citrus huanglongbing...	Shimwela et al.	2018	Phytopathology
6	Primary disease gradients of wheat stripe rust in large...	Sackett & Mundt	2005	Phytopathology
7	Distribution and gradient analysis of Ink disease...	Vannini et al.	2010	Forest Pathology
8	Dispersal of *Venturia inaequalis* ascospores and disease...	Holb et al.	2004	Phytopathology
9	Models of plant disease epidemics II: gradients of bean rust	Maffia & Berger	1999	Phytopathology
10	Primary disease gradients of bacterial blight of rice	Mundt et al.	1998	Phytopathology
11	Dispersal and disease gradients of anther-smut infection...	Roche et al.	1995	Ecology
12	Effect of simulated wind-driven rain on duration...	Bock et al.	2005	Plant Disease
13	Modelling inoculum dispersal and Sclerotinia stem rot...	Qandah & Mendoza	2012	Canadian Journal of Plant Pathology
14	Long-distance wind-dispersal of spores in a fungal plant...	Rieux et al.	2014	PLoS ONE
15	Splash dispersal of *Leptosphaeria maculans* pycnidiospores...	Travadon et al.	2007	Plant Pathology
16	Population dynamics and dispersal of *Aphelenchoides*...	Kohl et al.	2010	Journal of Nematology
17	Periodicity and gradients in dispersal of *Alternaria*...	Vloutoglou et al.	1995	European Journal of Plant Pathology
18	Dispersal of beet yellows and beet mosaic viruses...	Shepherd & Hills	1970	Phytopathology
19	Effect of pea canopy architecture on splash dispersal...	Schoeny et al.	2008	Plant Pathology
20	Dispersal of *Phytophthora parasitica* in tomato fields...	Neher & Duniway	1992	Plant Disease
21	Spatial patterns of Pierce's disease in the Napa Valley	Purcell	1974	American Journal of Enology and Viticulture
22	Dissemination and survival of *Pseudomonas alboprecipitans*...	Gitaitis et al.	1978	Phytopathology
23	Initial inoculum and spatial dispersal of *Colletotrichum*...	Rahman et al.	2015	Plant Disease
24	Epizootiology of a nuclear polyhedrosis virus (*Baculoviridae*)...	Entwistle et al.	1983	Journal of Applied Ecology
25	Dispersal of *Phytophthora capsici* and *P. parasitica*...	Café Filho & Duniway	1995	Plant Pathology
26	Head blight gradients caused by *Gibberella zeae* from...	Fernando et al.	1997	Phytopathology
27	Dispersal gradients of conidia of the butternut canker...	Tisserat & Kuntz	1983	Canadian Journal of Forest Research
28	Rain dispersal of *Colletotrichhum gloeosporioides*...	Yang & TeBeest	1992	Phytopathology
29	Analysis of spore dispersal gradients of *Botrytis cinerea*...	Johnson & Powelson	1983	Phytopathology
30	Splash dispersal of conidia of *Fusarium culmorum*...	Jenkinson & Parry	1994	Mycological Research
31	Dispersal of conidia of *Botrytis cinerea* in tomato fields	Chastagner et al.	1978	Phytopathology
32	Production of sori and dispersal teliospores of *Ustilago*...	Hoy et al.	1991	Phytopathology
33	Relative abundance and deposition gradients of clusters...	Ferrandino & Aylor	1987	Phytopathology
34	Analysis of disease progress of citrus canker in nurseries...	Gottwald et al.	1989	Phytopathology
35	Splash dispersal of *Phytophthora cactorum* from infected...	Grove at al.	1985	Phytopathology
36	Tracking the invasion: dispersal of *Hymenoscuphys fraxinus*...	Grosdidier et al.	2018	FEMS Microbiology Ecology
37	Epidemiology of *Phytophthora ramorum* in Oregon...	Hansen et al.	2008	Canadian Journal of Forest Research
38	Gradients of Ascochyta blight in Saskachewan lentil crops	Pedersen et al.	1993	Plant Disease
39	Disease gradients of late blight of potato from infrared...	Farber et al.	2020	American Journal of Potato Research
40	Influence of simulated rain on dispersal of rust spores...	Geagea et al.	2000	Agricultural and Forest Meterology
41	Asiatic citrus canker: Spatial and temporal spread...	Gottwald et al.	1988	Phytopathology
42	Monitoring the long-distance transport of *Fusarium*...	Prussin et al.	2014	Plant Disease
43	Epidemiology of Cercosporella footrot of wheat...	Rowe & Powelson	1973	Phytopathology
44	Disease gradients of fusiform rust on oak seedlings...	Schmidt et al.	1982	Phytopathology
45	Patterns of unobstructed splash dispersal	Stedman	1979	Annals of Applied Biology
46	Splash droplet and spore dispersal studies in field beans	Stedman	1980	Agricultural Meteorology
47	Effects of surface topography and rain intensity on splash...	Yang et al.	1990	Phytopathology
48	Spatial patten analysis of epidemics of citrus bacterial...	Gottwald & Graham	1990	Phytopathology
49	Spatial and temporal analysis of spread of late leaf...	Alderman et al.	1989	Phytopathology
50	Focus expansion of bean rust in cultivar mixtures	Assefa et al.	1995	Plant Pathology
51	Sptial and temporal spread of oat crown rust	Berger & Luke	1979	Phytopathology
52	Local dispersal of *Puccinia striiformis* f. sp. *tritici* from...	Farber et al.	2016	Plant Pathology

Table A.1. Literature Review: Title – Journal (2 of 2)

ID #	title	author(s)	year	journal
53	Local dispersal of *Puccinia triticina* and wheat...	Frezal et al.	2009	Phytopathology
54	Models for the spread of plant disease: some experimental...	Minogue & Fry	1983	Phytopathology
55	Use of the modified Gregory model to describe primary...	Mundt	1989	Phytopathology
56	Effect of crop age on primary gradients of late leaf spot...	Savary & Santen	1992	Plant Pathology
57	Characterization of wheat leaf rust epidemics...	Subba Rao et al.	1990	Phytopathology
58	Quantificiation of within-season focal spread of wheat...	Willocquet et al.	2008	Plant Pathology
59	Factors affecting infection gradients from a point source...	Cammack	1958	Annals of Applied Biology
60	Dipsersal of *Phytopthora capsici* and *P. parasitica*...	Filho & Duniway	1995	Plant Pathology
61	Dispersal gradients from a point source of *Erysiphe*...	Fried et al.	1979	Phytopathology
62	Disease spread of non-specialised fungal pathogens...	Jeger et al.	1983	Annals of Applied Biology
63	The spread of a powdery mildew of peach	Kable et al.	1980	Phytopathology
64	Transmission and dispersal of *Sphaeceloma arachidis* ...	Kearney et al.	2002	Peanut Science
65	Disease severity gradient of soybean downy mildew...	Lim	1978	Phytopathology
66	Studies on the spread of the olive scab pathogen...	Lops et al.	1993	EEPO Bulletin
67	Deposition gradients near to a point source in...	McCartney & Bainbridge	1984	Phytopathology
68	Peach rust spot epidemiology: incidence as affected...	Ries & Royse	1978	Phytopathology
69	Spread of sclerotinia stem rot of soybean from are...	Wegulo et al.	2000	Canadian Journal of Plant Science
70	Spatial and temporal spread of soybean stem canker...	Damicone et al.	1990	Phytopathology
71	Dispersal of *Rhynchosporium secalis* conidia from...	Fitt et al.	1988	Annals of Applied Biology
72	Epidemiology of tomato spotted wilt in pepper...	Gitaitis et al.	1998	Plant Disease
73	Effect of host and inoculum patterns on take-all...	Gosme & Lucas	2011	European Journal of Plant Pathology
74	Spatial and temporal development of common rust...	Headrick & Pataky	1988	Phytopathology
75	Influence of leafhopper (Homoptera: Cicadelliadae)...	Hunt et al.	1993	Environmental Entomology
76	Disease gradients and small plot experiments on...	Jenkyn & Bainbridge	1974	Annals of Applied Biology
77	Effect of placement of *Geaeumannomyces graminis*...	Kabbage & Bockus	2002	Plant Disease
78	Studies on the spread of Gaeumannomyces gramini...	Prew	1980	Annals of Applied Biology
79	Spore dispersal gradients and disease gradients...	Blenis & Chang	1993	Canadian Journal of Forest Research
80	Mechanisms of alteration in bean rust epidemiology...	Boudreau & Mundt	1992	Phytopathology
81	Splash dispesal of conidia of *Mycocentrospora acerina*...	Evenhuis et al.	1997	Plant Pathology
82	Epidemology of vascular-streak dieback of cocoa	Keane	1981	Annals of Applied Biology
83	Gradients in horizontal dispersal of cereal rust uredospores	Roelfs	1972	Phytopathology
84	Effect and underlying mechanisms of pea-cereal...	Schoeny et al.	2010	European Journal of Plant Pathology
85	A spatiotemporal analysis and dispersal patterns...	Contina et al.	2020	Phytopathology
86	Diffusion model for describing the regional spread...	Flores-Sanchez et al.	2017	Plant Disease
87	Spatial analysis of *Heterodera glycines* populations...	Francl	1986	Journal of Nematology
88	Local distance of wheat spike infection by released...	Keller et al.	2010	Plant Disease
89	Spatial scaling relationships for spread of disease...	Mundt & Sackett	2012	Ecosphere
90	Landscape heterogeneity and disease spread...	Mundt et al.	2011	Ecological Applications
91	Source or sink? The role of soil and water borne...	Peterson et al.	2014	Forest Ecology & Management
92	Temporal epidemiology of sudden oak death in Oregon	Peterson et al.	2015	Phytopathology
93	Experimental validation of a long-distance transport...	Prussin et al.	2015	Agricultural & Forest Meteorology
94	Predicting regional-scale spread of ascospores of...	Salam et al.	2011	Australasian Plant Pathology
95	*Spilocaea oleagina* in olive groves of southern Spain...	Viruega et al.	2013	Plant Disease
96	Wheat streak mosaic: a classic case of plant disease...	Workneh et al.	2010	Plant Disease
97	Dispersal of *Philaenus spumarius* (Hemiptera: Aphrophoridae)...	Bodino et al.	2021	Environmental Entomology
98	Spatiotemporal spread of grapevine red blotch-associated...	Cieniewicz et al.	2017	Virus Research
99	Measuring splash-dispersal of a major wheat pathogen...	Karisto et al.	2021	bioRxiv
100	Effects of single-drop impaction and natural and simulated...	Ahimera et al.	2004	Phytopathology
101	Effect of strawberry density on dispersal of *Colletotrichum*...	Boudreau & Madden	1995	Phytopathology
102	Relationship between pathogen splash dispersal...	Saint-Jean et al.	2006	Agricultural & Forest Meteorology
103	Experiments on the dispersion of *Lycopodium* and...	Sreeramulu & Ramalingam	1961	Annals of Applied Biology
104	Application of the diffusion equation for modelling...	Yang et al.	1991	New Phytologist

Table A.2. Literature Review: Host – Experiment Duration (1 of 2)

ID #	host	organism	type	sexual	asexual	dispersal syndrome	experiment duration
1	strawberry	powdery mildew (*Podosphaera aphanis*)	fungus	n	y	wind	4 weeks
2	potato	potato late blight (*Phytopthora infestans*)	oomycete	y	y	wind	growing season
3	wheat	fusarium head blight (*Gibberella zeae*)	fungus	y	n	wind	May-August
4	spruce	root & butt rot (*Heterobasidion annosum*)	fungus	y	n	wind	24 hrs
5	citrus	citrus huanglongbing (*Candidatus Liberibacter asiaticus*)	bacterium	n	y	insect	2005-2012
6	wheat	wheat stripe rust (*Puccinia striiformis*)	fungus	n	y	wind	growing season
7	chestnut	Ink disease (*Phytophthora cambivora*)	oomycete	n	y	flowing water	2002-2003
8	apple	apple scab (*Venturia inaequalis*)	fungus	y	n	wind	~ 1 month
9	bean	bean rust (*Uromyces appendiculatus*)	fungus	n	y	wind	30 days
10	rice	bacterial blight of rice (*Xanthomonas oryzae* pv. *oryzae*)	bacterium	n	y	wind/splash	30 days
11	white campion	anther-smut (*Ustilago violacea*)	fungus	y	n	insect/wind	~ 1 month
12	citrus	citrus canker (*Xanthomonas axonopodis* pv. *citri*)	bacterium	n	y	wind/splash	1 day
13	canola	stem rot (*Sclerotinia sclerotiorum*)	fungus	y	n	wind/splash	3 hours
14	banana	black leaf streak disease (*Mycospharella fijiensis*)	fungus	n	y	wind/splash	30 days
15	oilseed rape	blackleg/phoma stem canker (*Leptosphaeria maculans*)	fungus	n	y	splash	21 days
16	lantana	foliar nematode (*Aphelenchoides fragariae*)	nematode	y	n	splash	2 months
17	linseed	(*Alternaria linicola*)	fungus	n	y	wind	4 months
18	sugarbeets	beet yellows, etc. (vectored by green peach aphid)	virus	n	n	insect	3 months
19	winter pea	aschochyta blight (*Mycosphaerella pinodes*)	fungus	n	y	splash	5 days
20	tomato	buckeye rot (*Phytophthora parasitica*)	fungus	n	y	flowing water	2 months
21	wine grape	Pierce's disease (vectored by sharpshooters)	bacterium	n	y	insect	growing season
22	sweet corn	bacterial leaf blight of corn (*Pseudomonas alboprecipitans*)	bacterium	n	y	wind/splash	growing season
23	strawberry	strawberry anthracnose crown rot (*Colletotrichum gloeosporioides*)	fungus	n	y	wind/splash	60 days
24	spruce	GHNPV (*Gilpinia hercyniae* nuicear polyhedrosis virus)	virus	n	y	insect	summer
25	bell pepper, tomato, squash	(*Phytophthora capsici* & *Phytophthora parasitica*)	oomycete	n	y	flowing water	75 days
26	wheat	fusarium head blight (*Gibberella zeae*)	fungus	n	y	wind/splash	1 month
27	butternut	butternut canker (*Sirococcus clavigignenti-juglandacearum*)	fungus	n	y	wind/splash	summer
28	rice	anthracnose (*Colletotrichum gloeosporioides*)	fungus	n	y	splash	10min rain, 4 day inc.
29	snap beans	gray mold (*Botrytis cinerea*)	fungus	n	y	wind/splash	summer
30	wheat	fusarium ear blight/scab (*Fusarium culmorum* & *avenaceum*)	fungus	n	y	splash	48 hours
31	tomato	gray mold (*Botrytis cinerea*)	fungus	n	y	wind/splash	summer
32	sugarcane	sugarcane smut (*Ustilago scitaminea*)	fungus	y	n	wind	1 week
33	snap beans	bean rust (*Uromyces phaseoli*)	fungus	n	y	wind	1 day
34	grapefruit, orange, citrumelo	citrus canker (*Xanthomonas campestris* pv. *citri*)	bacterium	n	y	wind/splash	1 year
35	strawberry	leather rot (*Phytophthora cactorum*)	oomycete	n	y	splash	15 min
36	ash	ash dieback (*Hymenoscyphus fraxineus*)	fungus	y	n	wind/splash	15 days
37	Oregon tanoak	(*Phytophthora ramorum*)	oomycete	n	y	wind/splash	
38	lentil	aschochyta blight (*Ascochyta fabae* f. sp. *lentis*)	fungus	n	y	splash	2 months
39	potato	late blight (*Phytophthora infestans*)	oomycete	y	y	wind/splash	growing season
40	wheat	brown & yellow rust (*Puccinia recondita* f. sp. *tritici* & *P. striiformis*)	fungus	n	y	splash	10 min
41	grapefruit, orange	citrus canker (*Xanthomoas campestris* pv. *citri*)	bacterium	n	y	wind/splash	220 days
42	wheat	Fusarium head blight (*Fusarium graminearum*)	fungus	n	y	wind	4 hrs
43	wheat	footrot of wheat (*Cercosperella herpotrichoides*)	fungus	n	y	splash	growing season
44	oak	fusiform rust (*Cronartium quercuum* f.sp. *fusiforme*)	fungus	n	y	wind	1 month
45	grass	club moss (*Lycopodicum*)	NA	NA	NA	splash	10 min
46	field beans	club moss (Lycopodicum)	NA	NA	NA	splash	10 min
47	NA	NA	NA	NA	NA	splash	10min
48	grapefruit, orange	citrus canker (*Xanthomoas campestris* pv. *citrumelo*)	bacterium	n	y	wind/splash	461 days
49	peanut	late leaf spot (*Cercosporidium personatum*)	fungus	n	y	wind/splash	2 months
50	bean	bean rust (*Uromyces appendiculatus*)	fungus	n	y	wind/splash	20 days
51	oat	oat crown rust (*Puccinia coronata*)	fungus	n	y	wind/splash	50 days
52	wheat	stripe rust (*Puccinia striiformis* f. sp. *tritici*)	fungus	n	y	wind/splash	~20 days (<2 generations)

Table A.2. Literature Review: Host – Experiment Duration (2 of 2)

ID #	host	organism	type	sexual	asexual	dispersal syndrome	experiment duration
53	wheat	(*Puccinia triticina*)	fungus	n	y	wind/splash	1 month
54	potato	potato late blight (*Phytopthora infestans*)	oomycete	y	y	wind	25 days
55	wheat	leaf rust (*Puccinia recondita*)	fungus	n	y	wind	1 month
56	groundnut	late leaf spot (*Cercosporidium personatum*)	fungus	n	y	wind	34 days (1.5 latent periods)
57	wheat	leaf rust (*Puccinia recondita*)	fungus	n	y	wind	60 days
58	wheat	take-all (*Gaeumannomyces graminis* var. *tritici*)	fungus	n	y	splash	120 days
59	corn	corn rust (*Puccinia polysora*)	fungus	n	y	wind	30 days
60	tomato, bell pepper, squash	(*Phytopthora capsici & parasitica*)	oomycete	y	y	flowing water	2 weeks
61	wheat	(*Erysiphe graminis* f. sp. *tritici*)	fungus	n	y	wind	23 days
62	wheat, barley	(*Septoria nodorum, Rhychosporium secalis*)	fungus	n	y	wind/splash	growing season
63	peach	(*Oidium* sp.)	fungus	n	y	wind/splash	growing season
64	peanut	peanut scab (*Sphaceloma arachidis*)	fungus	n	y	wind/splash	growing season
65	soybean	downy mildew (*Peronospora manshurica*)	oomycete	y	y	wind/splash	72 days
66	olive	olive spot (*Spilocaea oleagina*)	fungus	n	y	wind	20h
67	barley	na (chemical tracer)	NA	NA	NA	wind/splash	minutes
68	peach	rusty spot (*Podosphaera leucotricha*)	fungus	n	y	wind	2 months
69	soybean	stem rot (*Sclerotinia sclerotiorum*)	fungus	y	n	wind	growing season
70	soybean	stem canker (*Diaporthe phaseolorum* var. *caulivora*)	fungus	y	y	wind/splash	growing season
71	barley	leaf blotch (*Rhynchosporium secalis*)	fungus	n	y	splash	15 min
72	tomato and pepper	tomato spotted wilt virus	virus	n	y	insect	5 weeks
73	wheat	take-all (Gaeumannomyces graminis var. tritici)	fungus	n	y	splash	growing season
74	corn	common maize rust (*Puccinia sorghi*)	fungus	n	y	wind/splash	50 days
75	corn	maize chlorotic dwarf machlovirus	virus	n	y	insect	12 days
76	barley	cereal powdery mildew (*Erysiphe graminis*)	fungus	n	y	wind/splash	3 months
77	wheat	take-all (*Gaeumannomyces graminis* var. *tritici*)	fungus	n	y	splash	growing season
78	wheat	take-all (*Gaeumannomyces graminis* var. *tritici*)	fungus	n	y	splash	5wks
79	pine	western gall rust (*Endoconartium harkenssii*)	fungus	n	y	wind	33 days
80	bean	bean rust (*Uromyces appendiculatus*)	fungus	n	y	wind/splash	70 days, 10hrs
81	caraway	(*Mycocentrospora acerina*)	fungus	n	y	wind	2wks
82	cocoa	vascular-streak dieback (*Oncobasidium theobromae*)	fungus	n	y	wind	4 hours
83	wheat	(*Puccinia graminis & P. recondita*)	fungus	n	y	wind	growing season
84	field pea	aschochyta blight (*Ascochyta fabae* f. sp. *lentis*)	fungus	n	y	splash	3min
85	potato	potato cyst nematode (*Globodera pallida*)	nematode	y	n	humans	6 years
86	citrus	Huanglongbing (HLB)	bacterium	n	y	insect	1 year
87	soybean	soybean cyst nematode (*Heterodera glycines*)	nematode	y	n	humans	1 month
88	corn	fusarium head blight (*Gibberella zeae*)	fungus	y	y	wind/splash	2 months
89	wheat	wheat stripe rust (*Puccinia striiformis*)	fungus	n	y	wind/splash	growing season
90	wheat	wheat stripe rust (*Puccinia striiformis*)	fungus	n	y	wind/splash	growing season
91	oak	sudden oak death (*Phytophthora ramorum*)	oomycete	y	y	wind/splash	4 months
92	oak	sudden oak death (*Phytophthora ramorum*)	oomycete	y	y	wind/splash	longterm
93	wheat	fusarium head blight (*Fusarium graminearum*)	fungus	n	y	wind	14 days
94	field pea	ascochyta blight (*Didymella pinodes*)	fungus	y	n	wind	growing season
95	olive	olive scab (*Spilocaea oleagina*)	fungus	n	y	wind	2 months
96	wheat	wheat streak mosaic virus	virus	n	y	insect	3 years
97	olive	(*Xylella fastidiosa*)	bacterium	n	y	insect	210 days
98	grape	grapevine red blotch-assoiated virus	virus	n	y	insect	3 years
99	wheat	(*Zymoseptoria tritici*)	fungus	n	y	splash	growing season
100	pistachio	(*Botryosphaeria dothidea*)	fungus	n	y	splash	growing season
101	strawberry	(*Colletotrichum acutatum*)	fungus	n	y	splash	45min
102	tomato	(*Phytophthora infestans*)	oomycete	y	y	splash	minutes
103	NA	(*Lycopodium, Podaxis pistillaris*)	fungus	n	y	wind	1 day
104	strawberry	(*Colletotrichum acutatum*)	fungus	n	y	splash	25min

Table A.3. Literature Review: Sampling Scale – R^2 (1 of 2)

ID#	sampling scale (m)	intentionally truncated	landscape level	inoculum source	argicultural system	figure label	figure data description	function type	function family	R^2
1	1.50	n	n	artificial	y	Fig. 3a/Table 2	glasshouse, yr1 (dashed line)	power law	power law	0.991
2	3500.00	n	y	natural	y	Fig. 5a	Region-VS	exponential	exponential	0.44
3	50.00	n	n	natural	y	Table 3	Macdonald	general/Lambert	exponential	NA
4	1000.00	n	n	natural	n	Table 3	P-type, August, pooled	power law	power law	0.96
5	NA (65,000 m2)	n	y	natural	y&n	NA	NA	NA	NA	NA
6	80.00	n	n	artificial	y	Fig. 3a/Table 1	Hermiston 2002, Downwind	modified power law	power law	0.987
7	50.00	n	y	natural	n	Fig. 3/Table 2	severity	exponential	exponential	0.34
8	57.00	n	n	natural	y	Fig.3	Jonagold 1999	exponential	exponential	0.87
9	8.50	n	n	artificial	y	Fig. 1	Day 29	exponential	exponential	0.983
10	0.70	y	n	artificial	y	Fig. 1a	1994	modified power law	power law	0.948
11	12.00	n	n	artificial	n	Fig. 2	mean number of spores per flower	power law	power law	NA
12	12.00	n	n	artificial	y	Fig. 6a	18 September	power law	power law	0.77
13	66.00	n	n	natural	y	Fig. 2b	Cando 2007	exponential	exponential	NA
14	1000.00	n	n	artificial	y	Fig. 2a	ascospores, density of lesions	NA	NA	NA
15	0.17	y	n	artificial	y	Fig. 6	pycnidiospores 2.8 mm drop	exponential	exponential	0.94
16	1.00	y	n	artificial	y	Fig. 4a	2007	linear regression	regression	0.9517
17	60.00	n	n	artificial	y	Fig. 6/Table 1	1993 - 17 September	exponential	exponential	0.96
18	1609.34	n	y	natural	y	Fig. 1	early yellows	linear regression	regression	0.98
19	1.00	y	n	artificial	y	Fig. 5/Table 3	D81S6	exponential	exponential	0.89
20	68.00	n	n	artificial	y	Fig. 4	NA	log-linear regression	regression	0.88
21	198.12	n	n	natural	y	Table 2	Flora	linear regression	regression	0.953
22	82.30	n	n	natural	y	Fig. 2	NA	exponential	exponential	NA
23	4.00	y	n	artificial	y	Fig. 3b/Table 3	2007	modified power law	power law	0.89
24	27000.00	n	y	artificial	n	Fig. 5	450 days	NA	NA	NA
25	70.00	n	n	artificial	y	Fig. 1	*P. parasitica*	linear regression	regression	0.94
26	25.00	n	n	artificial	y	Table 2	Site 1 (4B) 50x50m, upwind	exponential	exponential	>0.85
27	40.00	n	n	natural	n	Table 2	Pooled, Plot 1	exponential	exponential	0.94
28	1.80	y	n	artificial	y	Fig. 1/Table 2	concrete	power law	power law	0.77
29	8.92	n	n	artificial	y	Fig. 2e	experiment 1, harvest	log-linear	regression	NA
30	1.00	y	n	artificial	y	Fig. 1	*Fusarium culmorum*	NA	NA	NA
31	60.00	n	n	artificial	y	Fig. 4c	incidence of plant killed	regression	regression	NA
32	12.60	n	n	natrual	y	Fig. 6	average smut teliospore deposition	NA	NA	NA
33	15.00	n	n	artificial	y	Fig. 3	turbulent dispersal	power law	power law	NA
34	7.00	n	n	artificial	y	Fig. 5a	SWN disease gradient	Gompertz	exponential	NA
35	1.00	y	n	artificial	y	Fig. 1	+4cm	exponential	exponential	NA
36	140000.00	n	y	natural	n	Fig. 4a	NA	NA	NA	NA
37	4000.00	n	y	natural	n	Fig. 4	NA	NA	NA	NA
38	200.00	n	n	natural	y	Fig. 2	Zealandia 1989 (check)	power law	power law	0.79
39	230.00	n	n	natural	y	Fig. 9b	Fig1CMain	exponential	exponential	NA
40	0.84	y	n	artificial	y	Fig. 2/Table 2	Yellow 2.5	exponential	exponential	0.91
41	15.80	n	n	artificial	y	Fig. 3a	OCP (220 days)	Gompertz	exponential	NA
42	1000.00	n	n	artificial	y	Fig. 3	recovery rate	log-regression	regression	NA
43	1.22	y	n	artificial	y	Fig. 2	July 11	power law	power law	NA
44	152.40	n	y	natural	n	Fig. 4	average disease gradient	exponential	exponential	0.94
45	10.00	n	n	artificial	n	Fig. 6	cylinders (unstained)	regression	regression	NA
46	8.00	n	n	artificial	y	Fig. 9a	dense crop (solid line), slides (unstained)	regression	regression	NA
47	0.80	y	n	artificial	y	Fig. 4	plastic	regression	regression	NA
48	123.00	n	n	artificial	y	Fig. 3/Table 5	Venice citrus nursery	regression	regression	NA
49	9.00	n	n	artificial	y	Fig. 1/Table 1	% late spot incidence (mean values)	logit vs linear model	regression	NA
50	1.80	y	n	artificial	y	Fig. 1	31 DAI, day 21	exponential	exponential	n
51	4.80	n	n	artificial	y	Fig. 3a	Fulghum	regression	regression	n
52	1.50	y	n	artificial	y	Fig. 6/7	NA	power law	power law	0.83

Table A.3. Literature Review: Sampling Scale – R^2 (2 of 2)

ID#	sampling scale (m)	intentionally truncated	landscape level	inoculum source	argicultural system	figure label	figure data description	function type	function family	R^2
53	0.68	y	n	artificial	y	Fig. 4	L1	power law	power law	NA
54	26.50	n	n	artificial	y	Fig. 2	unsprayed Katahdin - 17 August (open circles)	regression	regression	NA
55	25.6	n	n	artificial	y	Fig. 2/Table 3	Corvallis A	power law	power law	NA
56	3.20	n	n	artificial	y	Fig. 2	7 weeks	power law	power law	NA
57	6.00	n	n	artificial	y	Fig. 2	NA	power law	power law	NA
58	0.45	n	n	artificial	y	Fig. 4	NA	NA	NA	NA
59	40.00	n	n	artificial	y	Fig. 2	30 days	NA	NA	NA
60	70.00	n	n	artificial	y	Fig. 1/Table 2	*P. capsici*	regression	regression	0.83
61	8.00	n	n	artificial	y	Fig. 1	NA	power law	power law	0.937
62	2.00	n	n	artificial	y	Fig. 1	NA	NA	NA	NA
63	160.00	n	y	natural	Y	Fig. 1	NA	power law	power law	NA
64	25.00	n	n	artificial	y	Fig. 3	NA	regression	regression	0.964
65	27.30	n	n	artificial	y	Fig. 2	NA	regression	regression	NA
66	20.00	n	n	artificial	y	NA	NA	exponential	exponential	NA
67	1.50	n	n	artificial	y	Fig. 2/Table 4	mid-crop height, short open crop	exponential	exponential	NA
68	74.00	n	y	natural	y	Fig. 3	percent peach fruit with rust spot lesions	regression	regression	NA
69	88.00	n	n	artificial	y	Fig. 9a	incidence of Sclerotinia stem rot (Transect 4)	NA	NA	NA
70	4.50	n	n	artificial	y	Fig. 3a	DPL 105 - 1986	exponential	exponential	NA
71	0.40	y	n	artificial	y	Fig. 1b	number of conidia (10cm drop height)	exponential	exponential	NA
72	320.00	n	n	natural	y	Fig. 1b	% disease incidence (Cook Site)	regression	regression	0.85
73	0.50	n	n	artificial	y	Fig. 2	NA	sigmoid		NA
74	10.00	n	n	artificial	y	Fig. 2	NA	Gompertz	exponential	NA
75	12.00	n	n	artificial	y	Fig. 2	NA	log-logistic	regression	NA
76	83.00	n	n	natural	y	Fig. 2	% leaf area affected (Julia)	NA	NA	NA
77	0.10	n	n	artificial	y	Fig. 2	% yield loss (GH experiment 1, below)	regression	regression	NA
78	0.89	n	n	artificial	y	Fig. 2	% sampling units with take-all	NA	NA	NA
79	170.00	n	n	natural	y	Fig. 2i	% infection (more west)	power law	power law	NA
80	10.00	n	n	artificial	y	Fig. 3a	monocrop (1989)	modified Gregory	power law	NA
81	16.00	n	n	artificial	y	Fig. 4	mean % infected trap plants	exponential	exponential	NA
82	80.00	n	n	artificial	y	Fig. 1	% seedlings infected (Block B)	power law	power law	NA
83	72.00	n	n	artificial	y	Fig. 2	stem rust	power law	power law	NA
84	0.58	y	n	artificial	y	Fig. 4	S5-7 G81	exponential	exponential	0.88
85	1000.00	n	y	natural	y	Fig. 8	Bin025	power law	power law	NA
86	135000.00	n	y	natural	y	Fig. 3b	Qroo-1 North-South	diffusion model		NA
87	0.50	n	n	natural	y	Fig. 1	NA	NA	NA	NA
88	24.00	n	n	artificial	y	Fig. 1	% spikes infected (Aurora, NY 2007)	NA	NA	NA
89	60.00	n	n	artificial	y	Fig. 2/3	% prevalence (2006, large plot)	power law	power law	NA
90	60.00	n	n	artificial	y	Fig. 2	2006	NA	NA	NA
91	20.00	n	y	natural	n	Fig. 4	NA	spline		NA
92	19400.00	n	y	natural	n	Fig. 3	NA	NA	NA	NA
93	1000.00	n	n	artificial	y	Fig. 3	idp (2012)	Gaussian	normal	NA
94	100.00	n	y	artificial	y	Fig. 2	Week 21	half-Cauchy	exponential	NA
95	40.00	n	n	natural	y	Fig. 4	conidia density	exponential	exponential	0.989
96	700.00	n	n	natural	y	Fig. 1a	% soil water content (transect 1, 2006)	regression	regression	NA
97	1400.00	n	n	artificial	y	Fig. 4a	olive orchard	Gaussian	normal	NA
98	225.00	n	n	natural	y	Fig. 2b	spatial pattern of diseased vines (2014-2016)	NA	NA	NA
99	1.20	y	n	artificial	y	Fig. 3c	intensity (pycnidia/leaf)	exponential	exponential	NA
100	0.30	y	n	artificial	y	Fig. 2a	number of droplets	exponential	exponential	NA
101	1.00	y	n	artificial	y	Fig. 2a	plant density (25% max density)	exponential	exponential	NA
102	0.15	y	n	artificial	y	Fig. 1	1334	power law	power law	NA
103	35.00	n	n	artificial	y	Fig. 1	Experiment VII, L	Sutton	power law	NA
104	1.20	y	n	artificial	y	Fig. 5	plastic, t=5 min	used their own model	normal	NA

Table A.3. Literature Review: Comparing Models – Truncated (1 of 2)

ID #	comparing models	author's "best fit"	timeseries	last in timeseries	linearized fxn	figure of gradient	function displayed	data points displayed	truncated
1	y	y	n	NA	n	y	y	y	y
2	n	NA	n	NA	n	y	y	y	n
3	y	y	n	NA	y	n	NA	NA	NA
4	n	NA	y	y	y	n	NA	NA	NA
5	n	NA	n	NA	NA	NA	NA	NA	NA
6	y	y	n	NA	n	y	y	y	y
7	n	NA	n	NA	n	y	y	y	y
8	y	y	n	NA	n	y	y	y	y
9	y	y	y	y	n	y	y	y	y
10	y	y	n	NA	n	y	y	y	y
11	y	y	n	NA	y	y	n	y	y
12	y	y	y	y	y	y	y	y	y
13	y	y	n	NA	n	y	y	y	y
14	y	NA	n	NA	NA	y	n	y	y
15	n	NA	n	NA	n	y	y	y	y
16	n	NA	n	NA	NA	y	y	y	NA
17	y	y	y	y	y	y	y	y	NA
18	n	NA	y	y	NA	y	y	y	y
19	n	NA	n	NA	n	y	y	y	n
20	n	NA	n	NA	NA	y	y	y	y
21	n	NA	n	NA	NA	n	NA	NA	NA
22	n	NA	n	NA	n	y	y	y	y
23	y	y	n	NA	y	y	y	y	NA
24	n	NA	y	y	NA	y	n	y	y
25	n	NA	n	NA	NA	y	y	y	y
26	n	NA	n	NA	y	n	NA	NA	NA
27	n	NA	y	y	y	n	NA	NA	NA
28	y	y	n	NA	y	y	n	y	y
29	n	NA	y	y	NA	y	y	y	y
30	n	NA	n	NA	NA	y	n	n	y
31	n	NA	n	NA	NA	y	y	y	y
32	NA	NA	n	NA	NA	y	n	y	y
33	y	y	n	NA	n	y	y	y	y
34	y	y	y	y	n	y	n	y	y
35	n	NA	n	NA	y	y	y	y	NA
36	y	NA	n	NA	NA	y	n	y	n
37	n	NA	n	NA	NA	y	n	y	n
38	y	y	n	NA	n	y	n	y	n
39	y	y	n	NA	n	y	y	y	n
40	n	NA	n	NA	n	y	n	y	y
41	n	NA	y	y	y	y	y	n	y
42	n	NA	n	NA	NA	y	y	y	y
43	n	NA	y	y	y	y	y	n	NA
44	n	NA	n	NA	n	y	y	y	y
45	n	NA	n	NA	NA	y	y	y	y
46	n	NA	n	NA	NA	y	y	y	y
47	n	NA	n	NA	NA	y	y	y	y
48	n	NA	n	NA	NA	y	n	y	y
49	n	NA	y	y	NA	y	n	y	y
50	n	NA	y	y	y	y	y	y	NA
51	n	NA	y	y	NA	y	y	y	NA
52	y	y	n	NA	n	y	y	y	n

Table A.3. Literature Review: Comparing Models – Truncated (2 of 2)

ID #	comparing models	author's "best fit"	timeseries	last in timeseries	linearized fxn	figure of gradient	function displayed	data points displayed	truncated
53	y	y	n	NA	n	y	y	y	n
54	n	NA	y	y	y	y	y	y	NA
55	NA	NA	NA	NA	y	y	n	NA	n
56	y	y	y	y	n	y	n	y	y
57	y	y	y	y	y	y	y	y	y
58	n	NA	y	y	NA	y	n	y	n
59	n	NA	y	y	NA	y	n	y	y
60	n	NA	n	NA	n	y	y	y	y
61	y	y	n	NA	y	y	y	y	NA
62	n	NA	n	NA	NA	y	n	y	y
63	y	y	n	NA	NA	y	n	y	n
64	n	NA	y	y	NA	y	n	y	y
65	n	NA	y	y	y	y	y	y	y
66	n	NA	n	NA	NA	n	NA	NA	NA
67	y	y	y	y	y	n	NA	NA	NA
68	n	NA	y	y	NA	y	y	n	NA
69	n	NA	y	y	NA	y	n	y	y
70	n	NA	n	NA	n	y	y	y	y
71	n	NA	n	NA	y	y	n	y	y
72	n	NA	n	NA	NA	y	y	y	y
73	n	NA	y	y	n	y	n	y	n
74	n	NA	y	y	y	y	y	y	y
75	n	NA	n	NA	y	y	y	y	n
76	n	NA	y	y	NA	y	n	y	y
77	n	NA	n	NA	n	y	y	y	y
78	NA	NA	y	y	NA	y	n	y	y
79	n	NA	n	NA	NA	y	n	y	y
80	n	NA	n	NA	y	y	y	y	y
81	y	y	n	NA	n	y	y	y	y
82	n	NA	n	NA	y	y	y	y	y
83	n	NA	n	NA	y	y	y	y	y
84	n	NA	n	NA	n	y	y	y	n
85	n	NA	n	NA	n	y	y	y	n
86	y	y	y	y	n	y	n	y	n
87	n	NA	n	NA	NA	y	n	y	NA
88	n	NA	n	NA	NA	y	n	y	y
89	n	NA	n	NA	y	y	y	y	y
90	n	NA	y	y	NA	y	n	y	y
91	n	NA	n	NA	n	y	n	y	y
92	n	NA	n	NA	NA	y	n	y	n
93	y	y	n	NA	NA	y	y	y	y
94	n	NA	y	y	n	y	y	y	y
95	n	NA	n	NA	n	y	y	y	n
96	n	NA	n	NA	n	y	y	y	y
97	n	NA	y	y	n	y	y	n	NA
98	n	NA	n	NA	NA	y	n	y	y
99	n	NA	n	NA	n	y	y	y	y
100	y	y	n	NA	n	y	n	y	n
101	n	NA	n	NA	NA	y	n	y	y
102	y	y	n	NA	NA	y	n	y	y
103	n	NA	y	y	n	y	n	y	y
104	n	NA	y	y	n	y	y	y	y

APPENDIX B

PYTHON CODE TO MANAGE & VISUALIZE EPIMUL SIMULATION RESULTS FILES

Appendix B contains examples of the Python code used in Chapter 3 to manage and visualize EPIMUL simulation results files.

Step 1: Copy and rename EPIMUL simulation folders

```
# Inspiration Code:
# https://automatetheboringstuff.com/chapter9/

import shutil, os

for i in range(2, 101):
    shutil.copytree(r'C:\Users\es26255\Desktop\5% susceptible\5% susceptible
- 1% initial disease - power
law\5%susceptible_1%initialdisease_powerlaw_stoch_1',
                    r'C:\Users\es26255\Desktop\5% susceptible\5% susceptible -
1% initial disease - power
law\5%susceptible_1%initialdisease_powerlaw_stoch_{}'.format(i))
```

Step 2: Copy and rename ‘spot’ results files

```
#Inspiration Code:
# https://www.includehelp.com/python/copy-and-rename-files.aspx
# https://stackoverflow.com/questions/15942564/copying-files-based-on-file-name-rather-than-data-type-in-python

import os
import shutil
import glob
import pandas as pd

path = r'C:\Users\es26255\Desktop\5% susceptible (SIMULATIONS COMPLETE 10-6-2021)\1% initial disease - power law'

os.chdir(path)

#create a dir where we want to copy and rename
dest_dir = os.mkdir('5% susceptible - 1% disease - power law - results')
dest_dir = path+"/5% susceptible - 1% disease - power law - results"

for path, sub_directories, files in os.walk(path):
    directory_name = os.path.split(path)[1] #provides names for all the subdirectories

    for file in files:
        if file.startswith("spot"): #swap out for "file.endswith('.txt') to grab all text files instead
            source = os.path.join(path, file) #provides locations for all target files
            destination = os.path.join(dest_dir, directory_name + "_" + file)
            shutil.copy(source, destination)

for textfile in glob.iglob('**/*_spot.txt', recursive = True):
    with open(textfile, 'r') as old:
        lines = old.readlines()[10156:]
        clean = [l.replace('race = 1;    iday = 51\n','') for l in lines] #remove string, carriage return & convert list to string so can use 'replace'

    with open(textfile, 'w') as new:
        new.writelines(clean)

    read_file = pd.read_csv(textfile, delimiter=' ', header=None)
    read_file = read_file.dropna(axis='columns', how='any') #drop blank columns
    read_file.to_csv(textfile, header=None, index=False)

    pre, ext = os.path.splitext(textfile) #split into filename and extension
    os.rename(textfile, pre + ".csv")
```

Step 3: Matrix creation and summary statistics for all stochastic simulations

```
#Inspiration Code:
# https://blog.softhints.com/how-to-merge-multiple-csv-files-with-python/

import pandas as pd
import os
import glob

path = r'C:\Users\es26255\Desktop\5% susceptible (SIMULATIONS COMPLETE 10-6-
2021)\1% initial disease - power law\5% susceptible - 1% disease - power law
- results'

os.chdir(path)

all_files = glob.glob('**/*stoch*.csv', recursive=True)

df_all = (pd.read_csv(f, header=None) for f in all_files)

df_merged = pd.concat(df_all)

# df_merged.to_csv('df_merged.csv', header=None, index=False)

df_mean = df_merged.groupby(level=0).mean()

df_median = df_merged.groupby(level=0).median()

df_min = df_merged.groupby(level=0).min()

df_max = df_merged.groupby(level=0).max()

df_mean.to_csv('5%susceptible_1%disease_plaw_stoch_MEAN.csv', header=None,
index=False)

df_median.to_csv('5%susceptible_1%disease_plaw_stoch_MEDIAN.csv',
header=None, index=False)

df_min.to_csv('5%susceptible_1%disease_plaw_stoch_MIN.csv', header=None,
index=False)

df_max.to_csv('5%susceptible_1%disease_plaw_stoch_MAX.csv', header=None,
index=False)
```

Step 4: Create simulation field figure with center-field inset

```
#Inspiration code
# https://stackoverflow.com/questions/2578752/how-can-i-plot-nan-values-as-a-special-color-with-imshow-in-matplotlib
#
https://matplotlib.org/stable/gallery/lines_bars_and_markers/masked_demo.html

import os
import glob
import shutil
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

path = r'C:\Users\emssy\OneDrive - University of Georgia\UGA Work\Research\MS thesis\simulation results\5% susceptible'

os.chdir(path)

field_layout = pd.read_csv(r'C:\Users\emssy\OneDrive - University of Georgia\UGA Work\Research\MS thesis\simulation results\5% susceptible\field layout 5% susceptible.csv', header=None)

for file in glob.iglob('**/*MEDIAN.csv', recursive=True):

    whole_field_file = pd.read_csv(file, header=None)
    whole_field_data = np.array(np.where(field_layout==0, np.nan, whole_field_file)) #apply NaN to non-host cells
    nonzero_data = np.ma.masked_where(whole_field_data==0, whole_field_data) #apply color to non-zero data only

    whole_field_cmap = plt.cm.get_cmap('Greys').copy() #have to make a copy of cmap in order to change it
    whole_field_cmap.set_bad(color='black') #set NAN to display as black

    whole_field_map = plt.imshow(whole_field_data, cmap=whole_field_cmap) #display values as white (bottom of Grey colormap)
    nonzero_data = plt.imshow(nonzero_data, cmap='gist_rainbow_r', vmin=0, vmax=100) #needs to come after whole field map so color displays on top
    whole_field_map.axes.xaxis.set_ticks([])
    whole_field_map.axes.yaxis.set_ticks([])
    filename = os.path.splitext(file)[0]
    # plt.title(filename) #filename.replace("_"," ") - can add this if want to remove underscores
    plt.colorbar()
    plt.savefig(filename+'_nolabel_staticscale.png', dpi=800, bbox_inches='tight', transparent=True, pad_inches=0)
    plt.clf() #clears plot from current window, leaves window open for creating addtl plots (so I don't get 1000 colorbars on my figures lol)


for file in glob.iglob('**/*_stoch_M*.csv', recursive=True):

    csv_file = pd.read_csv(file, header=None)
    map_data = np.array(np.where(field_layout==0, np.nan, csv_file))
```

```
    inset = list(range(0,74)) + list(range(125,200))
    inset_data = np.delete(map_data, inset, 1)
    final_inset_data = np.delete(inset_data, inset, 0)

    inset_field_cmap = plt.cm.get_cmap('gist_rainbow_r').copy() #have to make
a copy of cmap in order to change it
    inset_field_cmap.set_bad(color='black') #set NAN to display as black

    inset_field_map = plt.imshow(final_inset_data, cmap=inset_field_cmap,
vmin=0, vmax=100)
    inset_field_map.axes.xaxis.set_ticks([])
    inset_field_map.axes.yaxis.set_ticks([])
    filename = os.path.splitext(file)[0]
    # plt.title(filename) #filename.replace("_"," ") - can add this if want
to remove underscores
    # plt.colorbar()
    plt.axis('off')
    plt.savefig(filename+'_inset_nolabel_staticscale_nobar.png', dpi=800,
bbox_inches='tight', transparent=True, pad_inches=0)
    plt.clf() #clears plot from current window, leaves window open for
creating addtl plots (so I don't get 1000 colorbars on my figures lol)
```

APPENDIX C

DISEASE DENSITY SIMULATIONS

Appendix C contains results files for seven different disease density simulations. We ran deterministic simulations with 0.1% initial disease spread across different numbers of cells at different densities to determine the impact of initial disease density on final epidemic magnitude.

Table C.1. AUDG for each disease density scenario.

equation	size	type	AUDG raw	AUDG # lesions	% infected cmpt	fill-corners	cells >0	% cells filled
power law	50x50	corners	21,472	42,943,353	0.060%	0.096%	360,000	100.00%
		fill	55,971	111,941,721	0.155%		360,000	100.00%
	30x30	corners	21,444	42,887,765	0.060%	0.081%	360,000	100.00%
		fill	50,511	101,022,232	0.140%		360,000	100.00%
	10x10	corners	21,017	42,033,839	0.058%	0.018%	360,000	100.00%
		fill	27,446	54,891,589	0.076%		360,000	100.00%
	1x1	center	16,267	32,534,957	0.045%		360,000	100.00%
exponential	50x50	corners	58,068	116,135,773	0.161%	-0.001%	42,596	11.83%
		fill	57,789	115,577,260	0.161%		36,404	10.11%
	30x30	corners	57,349	114,698,365	0.159%	-0.007%	34,420	9.56%
		fill	54,824	109,647,929	0.152%		30,672	8.52%
	10x10	corners	51,652	103,304,997	0.143%	-0.003%	27,616	7.67%
		fill	50,475	100,949,818	0.140%		26,892	7.47%
	1x1	center	49,522	99,044,439	0.138%		26,325	7.31%

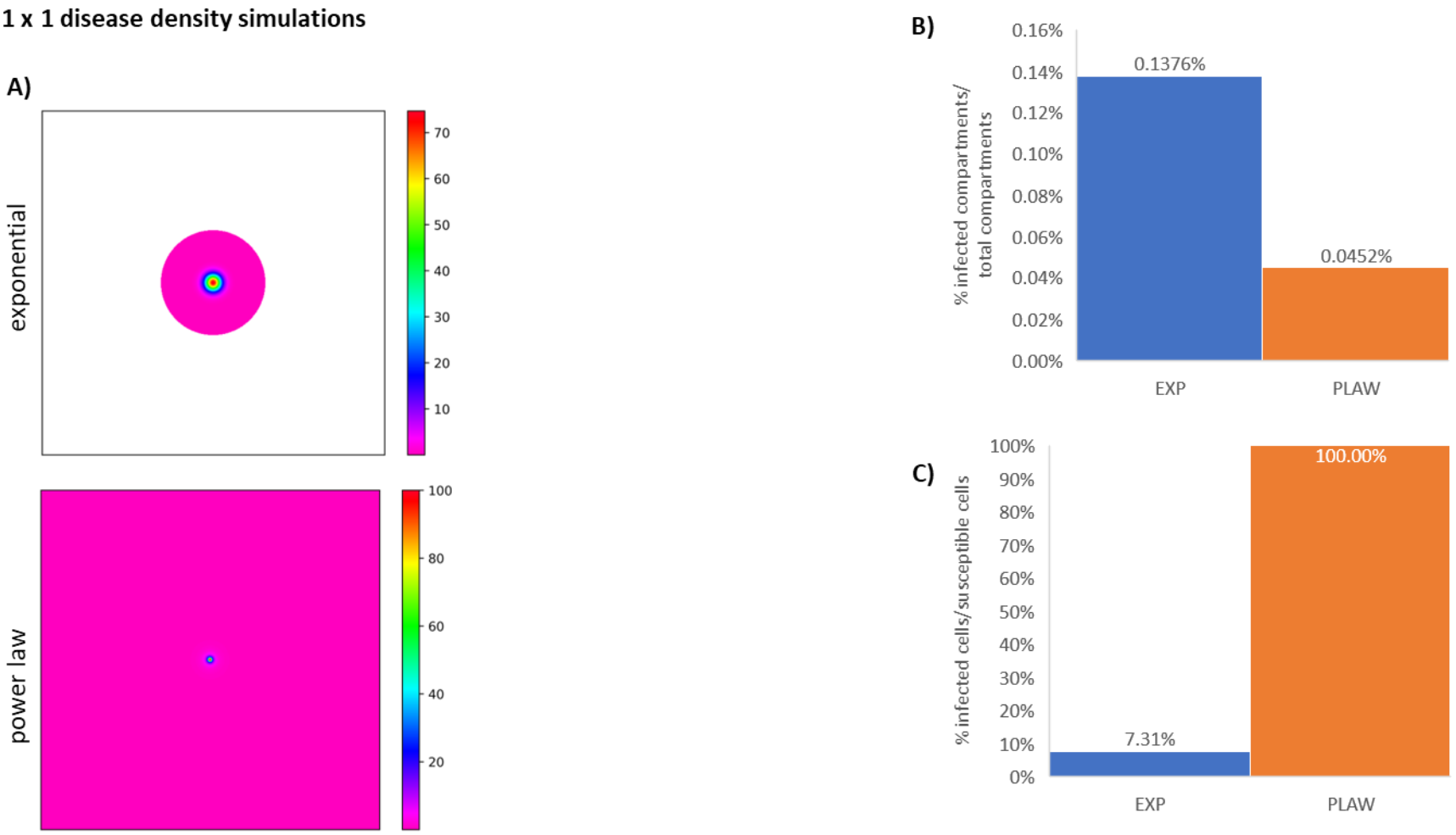


Figure C.1. Results for the final day of the 1 x 1 cell disease density simulation. a) 600 x 600 cell field visualizing results. The scale bar represents % infection per cell, white cells represent a value of 0.

Bar charts show b) percent of infected cells per total infectible cells and c) percent infected compartments per total infectible compartments (AUDG) for each field type and equation.

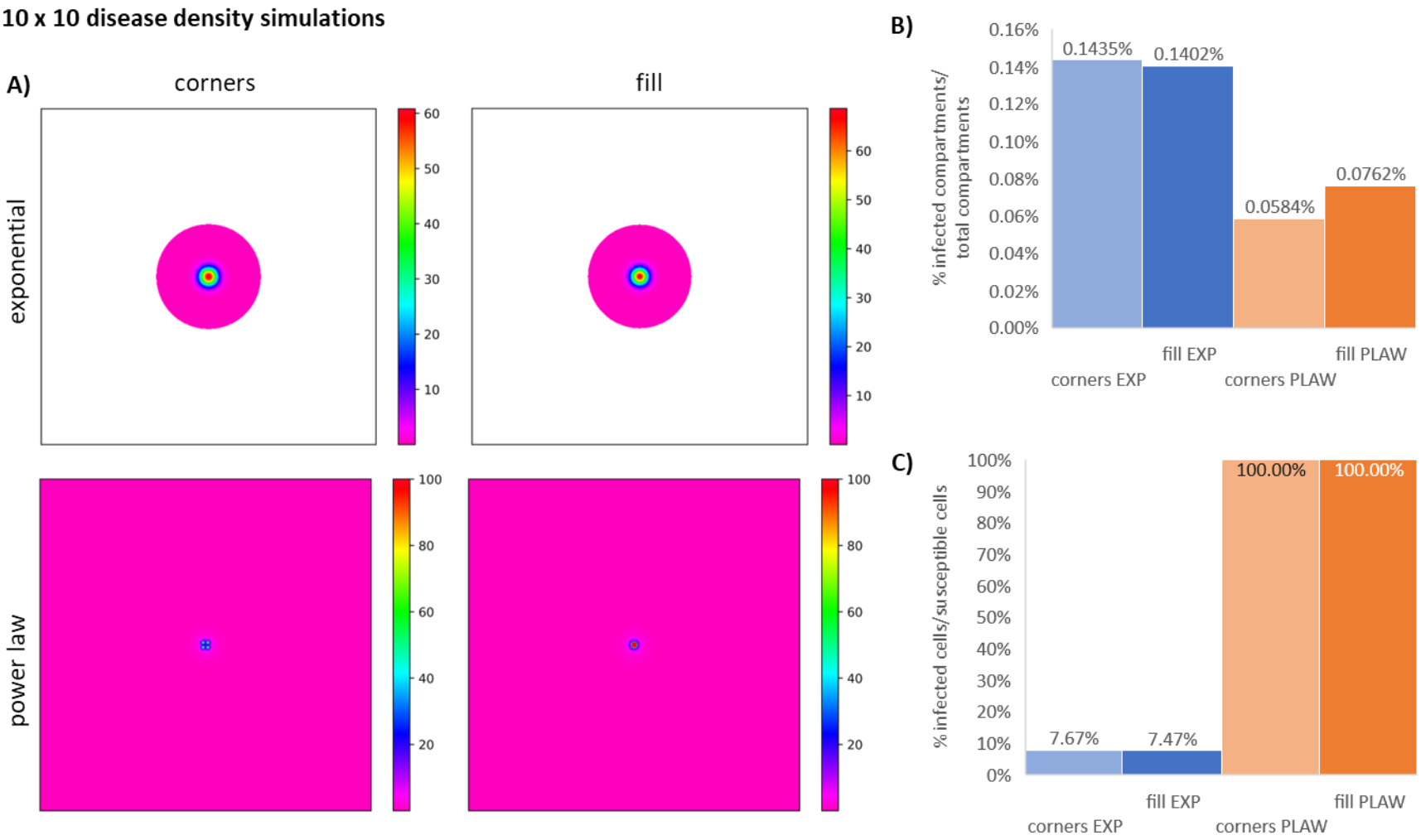


Figure C.3. Results for the final day of the 10 x 10 cell disease density simulation. a) 600 x 600 cell field visualizing results. The scale bar represents % infection per cell, white cells represent a value of 0.

Bar charts show b) percent of infected cells per total infectible cells and c) percent infected compartments per total infectible compartments (AUDG) for each field type and equation.

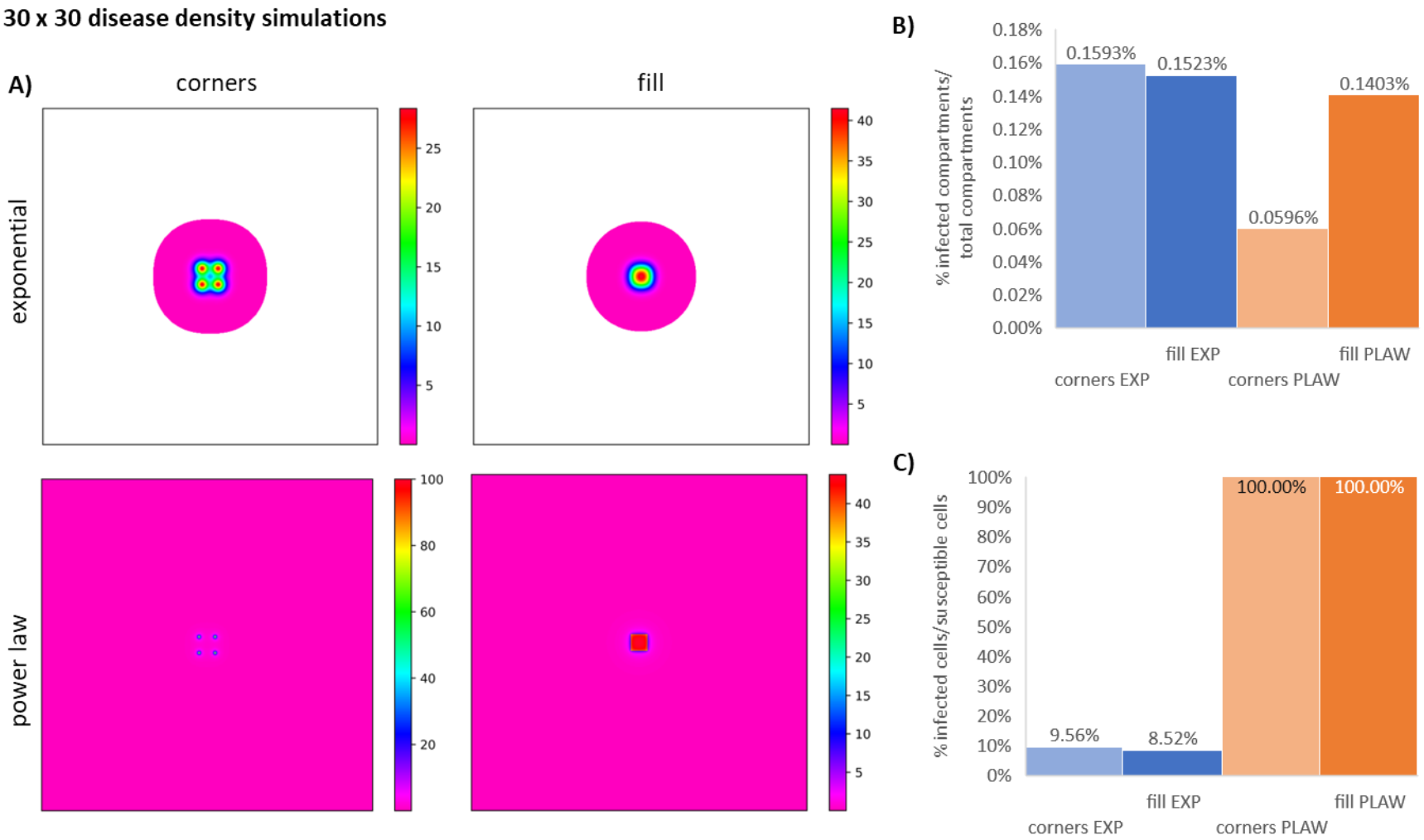


Figure C.3. Results for the final day of the 30 x 30 cell disease density simulation. a) 600 x 600 cell field visualizing results. The scale bar represents % infection per cell, white cells represent a value of 0.

Bar charts show b) percent of infected cells per total infectible cells and c) percent infected compartments per total infectible compartments (AUDG) for each field type and equation.

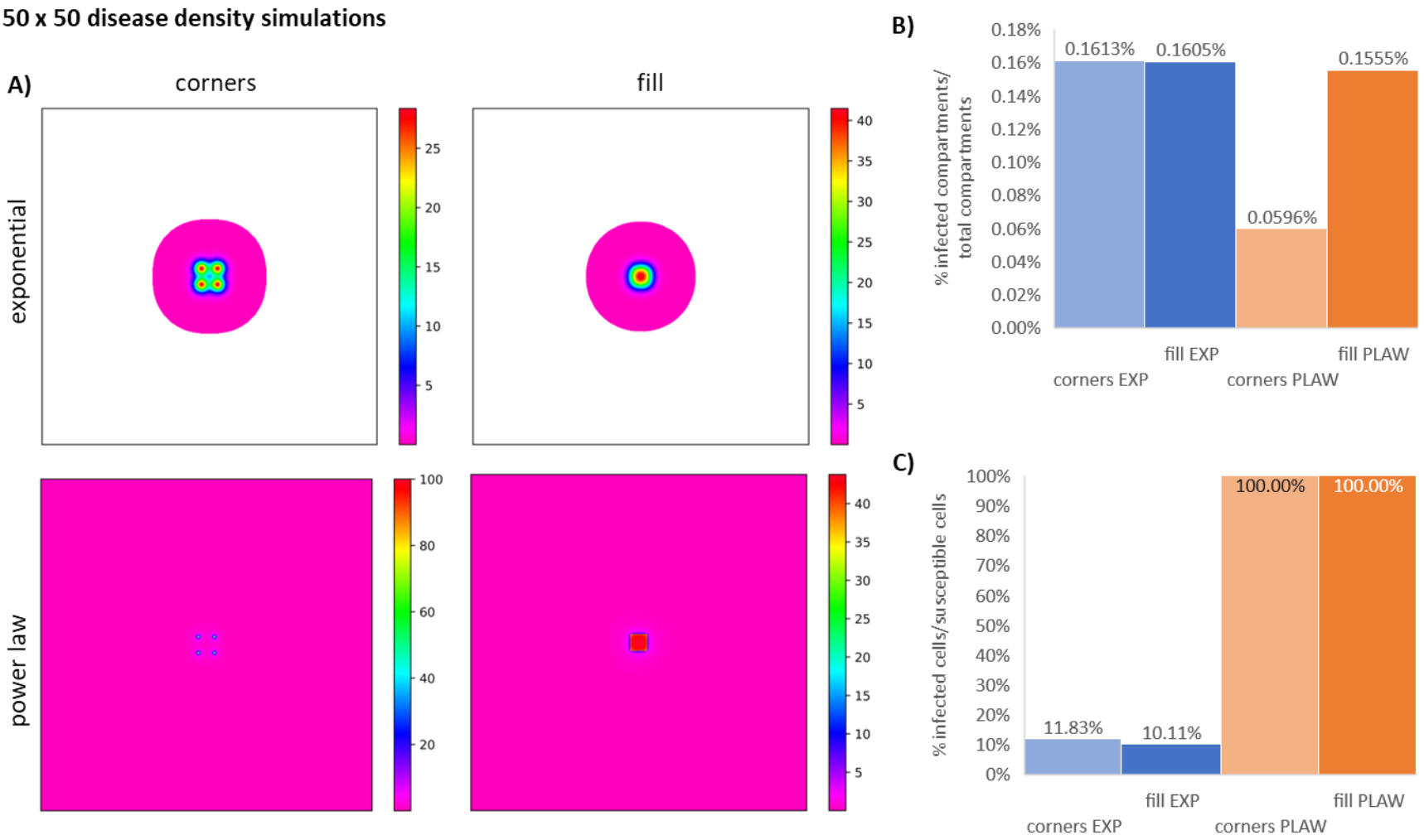


Figure C.4. Results for the final day of the 50 x 50 cell disease density simulation. a) 600 x 600 cell field visualizing results. The scale bar represents % infection per cell, white cells represent a value of 0.

Bar charts show b) percent of infected cells per total infectible cells and c) percent infected compartments per total infectible compartments (AUDG) for each field type and equation.