A SPATIALLY-EXPLICIT EPIDEMIOLOGICAL MODEL FOR ARTHROPOD-VECTORED PLANT VIRUSES

by

ALLYSA CAMPBELL

(Under the Direction of Paul M. Severns)

ABSTRACT

In the face of a changing climate, food crops are increasingly threatened by arthropod-vectored viruses. The development of accessible predictive tools for these pathosystems has been overlooked. I developed a generalizable spatially-explicit simulation model for arthropod-vectored plant diseases that was parameterized for two plant virus pathosystems, Blueberry Necrotic Ring Blotch Virus (BNRBV) of Southern Highbush Blueberry and a Whitefly Transmitted Virus Complex (WTVC) of yellow squash, to replicate disease incidence field data for these systems. For the BNRBV pathosystem, the model performed well at replicating severity, location, and patchiness of disease. For the WTVC pathosystem, the model performed well at representing patchiness of disease but tended to overestimate whole-field disease incidence and misrepresent disease distribution in the field. Model results suggest that a general understanding of pathosystem dynamics coupled with demographic and behavioral data from vector literature can be sufficient to reasonably represent plant virus disease spread.

INDEX WORDS: Simulation models, compartmental models, plant epidemiology, plant viruses, Blueberry Necrotic Ring Blotch, *Bemisia tabaci*, eriophyid mites, *Calacarus* sp., *Crinivirus*, *Begomovirus*, Kitaviridae

A SPATIALLY-EXPLICIT EPIDEMIOLOGICAL MODEL FOR ARTHROPOD-VECTORED PLANT VIRUSES

by

ALLYSA CAMPBELL

B.S., University of Georgia, 2020

B.A., University of Georgia, 2020

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

2024

A SPATIALLY-EXPLICIT EPIDEMIOLOGICAL MODEL FOR ARTHROPOD-VECTORED PLANT VIRUSES

by

ALLYSA CAMPBELL

Major Professor:	Paul M. Severns
Committee:	Shavannor Smith
	Amy Baldwin
	Carl M. Deom

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2024

DEDICATION

The completion of this thesis was made possible by the copious support I received from my loving family, loyal friends, and steadfast fiancé.

To my parents, sisters, and grandparents: thank you for believing in me even when I struggled to believe in myself. You constantly remind me of who I am. You are the pillars of my life.

To my friends: thank you for staying up with me during the late nights, for helping me phrase and rephrase sentences, for reassuring me that my words made sense, for your encouragement and your kindness and your love. Some of you were alongside me for this process, some of you were many miles away – all of you carried me through it.

To my fiancé: thank you for being, as you have always been, my rock. It is difficult to overstate your role in this work. Your faith, your patience, and your selflessness have made this possible.

Pop-up, I never forgot what I came here for. I hope I made you proud.

ACKNOWLEDGEMENTS

Supported by USDA Project No. 6080-22000-030-001S. We would like to thank Nathan Schumaker[1] and Sydney Watkins[1,2] ([1]Environmental Protection Agency, [2]Oak Ridge National Laboratory) for their guidance and assistance with model development. We would also like to thank C. Mike Deom and Phil Brannen for providing raw field data used to evaluate model performance in Chapter 2. Finally, we would also like to thank Vishnu Ramachandran for his assistance with data processing that significantly increased the speed of our data analysis.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Climate change and a rapidly growing global population have elevated the importance, socio-economic value, and complexity of plant health maintenance worldwide. Global food demand is expected to increase significantly by 2050 (by some estimates, 50-60%), and the potential impact of climate change on disease risk for most crops remains unclear ("Plant Diseases and Food Security in the 21st Century", 2021, Falcon *et al.*, 2022). This situation is made more complex by climatically- and anthropogenically-driven emergence and introduction of novel and exotic pathogens. In the face of such challenges, tools that can predict when, where and how disease occurs under potential future conditions are advantageous in the control of plant disease epidemics. Epidemiological models facilitate projections of disease spread over a range of potential and known biological scenarios through statistical representations of important biological and environmental factors. Because the invasion of disease into a population of hosts is complex and empirical experimentation is necessarily focused on the limited number of epidemiological factors that can be directly measured, models are important tools to understand the dynamics of disease outbreaks and epidemics. Epidemiological models are an essential tool to explore and project potential outcomes from what is known about specific hostplant-pathogen-environmental interactions and while simultaneously accounting for the uncertainty in important epidemiological factors that are difficult, if not impossible, to study empirically. Mathematical modelling attempts to simplify biological events into their component parts in order to

understand the rules that govern them (Antoniouk & Melnik, 2013). When constructed from previously collected data, models allow us to extend the usefulness of that data, evaluate assumptions that govern system dynamics, and make projections about complex biological systems that might not be understood in the absence of a well-designed mathematical model (Brauer, 2008, Jeger, 2000).

Common predictive tools in plant epidemiology. Disease risk models are commonly used in plant pathology. These models usually serve one of two purposes: to determine when risk of an economically consequential infection (either initial infection or infection to an economically important plant part) is high (the most common form of disease risk model), or to provide an early warning of infection through detection of pre- or asymptomatically diseased plants (Fenu & Malloci, 2021, Hardwick, 2006, Miller & O'brien, 1952). In their calculation of disease risk, these models primarily focus on weather station sourced climate variables (temperature, humidity, precipitation) that are favorable for disease infection and intensification. Theoretically, this prevents unnecessary application of pesticides (primarily fungicides), because growers have the option to apply treatments to crops only when infection is predicted to be likely (Chappell *et al.*, 2020, Cooke *et al.*, 2006, Everts *et al.*, 2016, Garrett *et al.*, 2004, Jørgensen *et al.*, 2020, Smith *et al.*, 2018). To facilitate potential early disease detection, some models use remote sensing and multispectral imaging data (Cooke et al., 2006). Early detection allows for treatment of at-risk but potentially exposed and latently infected host plants while disease pressure is low. These models could benefit growers in that sufficiently early timing of treatments tends to enhance their effectiveness with a smaller footprint to control a localized disease outbreak (Severns & Mundt, 2022).

Disease risk models promote the proactive and reactive management of plant disease, but their use is restricted to specific types of well-documented pathosystems and rely heavily on climate envelopes as the sole means of prediction (*e.g.* these models do not integrate the current or future spatial distribution of disease). The utility of these models in terms of providing data to guide management decisions is clear, but these models do not have the flexibility to allow users to elucidate unknown elements of pathogen biology. Modern disease forecasting models often rely on machine learning analyses and surprisingly few of these models have been developed in the literature between 2010 and 2020, perhaps due to their intensive data requirements (Cooke et al., 2006).

The serviceability of existing phytopathological models is limited by their specificity. They are mostly created for one specific crop or even a single disease of one crop (Andrade-Piedra *et al.*, 2005, Bregaglio *et al.*, 2021, Chappell et al., 2020, Contreras-Medina, 2009, Coop, 2023, Everts et al., 2016, Fenu & Malloci, 2021, Ferriss & Berger, 1993, Jørgensen et al., 2020, Kim & Jung, 2020, Pietravalle *et al.*, 2003, Smith et al., 2018). Most often, the pathogens whose occurrences they are created to predict tend to be (mostly wind-dispersed) fungal and fungal-like pathogens and not the causal agents of arthropod-vectored diseases (Caubel *et al.*, 2012, Juroszek *et al.*, 2022, Juroszek *et al.*, 2020, McLeish *et al.*, 2020, Miller *et al.*, 2022, Pietravalle et al., 2003, Savary *et al.*, 2015, Severns, 2022). This homogeneity and lack of generalizability raises the barrier to entry for the creation of arthropod-vectored plant virus models. Extant phytopathological models also tend to be constructed using approaches that do not account for the spatial arrangement of hosts, a key component in the determination of plant disease outcomes – especially for viral plant diseases (Cunniffe *et al.*, 2015a, Fabre *et al.*, 2021, Ferriss & Berger, 1993, Jones, 2009, Ostfeld *et al.*, 2005, Zhang *et al.*, 2000).

Spatially-explicit models. Spatially-explicit models (SEMs) are simulation models that account for real spatial measurements. In general, SEMs have the power to more faithfully represent the epidemiology of and evaluate management practices for emerging diseases (DeAngelis & Yurek, 2017, Fabre et al., 2021). Phytopathological SEMs are usually constructed for aerially dispersed fungal pathogens (Beasley *et al.*, 2022, Calonnec *et al.*, 2008, Christopher P. Brooks *et al.*, 2008, Fabre et al., 2021, Lof *et al.*, 2017, Meentemeyer *et al.*, 2011, Severns & Mundt, 2022, Vinatier *et al.*, 2009). Those that are developed for arthropod-vectored bacterial and viral pathogens tend not to meaningfully account for vector behavior and/or demography (Cunniffe *et al.*, 2015b, Nguyen *et al.*, 2023, Parnell *et al.*, 2009, Picard *et al.*, 2017, Thierry *et al.*, 2021, White *et al.*, 2017). The relative paucity of SEMs for arthropod-vectored diseases is concerning, given that nearly half of all emerging plant diseases are caused by viruses and viral plant diseases are primarily vectored by arthropods (Anderson *et al.*, 2004, Ferriss & Berger, 1993). Existing SEMs in the field of plant pathology are almost never open access and tend to be constructed for and parameterized for only one pathosystem in ways that prevent the models from being modified to another related system (Beasley et al., 2022, Christopher P. Brooks et al., 2008, Contreras-Medina, 2009, Fabre et al., 2021, Lof et al., 2017, Meentemeyer et al., 2011, Vinatier et al., 2009). Generalizable plant disease SEMs can be parameterized for many similarly vectored plant diseases and making SEMs publicly available allows for widespread adoption and adaptation for various purposes. With this in mind, I developed a generalizable, open access SEM for arthropod-vectored plant viruses that accounted for some vector behaviors.

Most plant viruses are vectored by arthropods, which vary in their modes of viral transmission. It is the vector that differentiates plant viruses from passively dispersed

microscopic plant pathogens in a distinct way: the distribution of plant viruses is dependent upon vector behavior. Insect vectors are able to make decisions about which plants to feed and oviposit on and are influenced by factors such as host preference, host nutritional value, host physiological status, interactions among conspecifics, weather, and the presence of natural enemies (Avery *et al.,* 2015, Byrne & Bellows Jr, 1991, Costa *et al.,* 1991, Felicio *et al.,* 2019, Liu *et al.*, 2007, Mayer *et al.*, 2002, Milenovic *et al.,* 2019, Nam & Hardie, 2012, Rodelo-Urrego *et al.*, 2013, Shi *et al.*, 2018, Zhang *et al.*, 2019, Zhao *et al.*, 2021). Predicting the movement of a discerning organism requires more considerations than predicting that of a passively dispersed fungal spore spread along a gradient such as wind. For this reason, models predicting the movement of arthropod vectors must account for factors relating to landscape distribution of vector resources over space and time, vector behavioral ecology, vector demography and population dynamics, climatic conditions, host plant phenology, and transmission rates (Cunniffe et al., 2015a, Jeger *et al.*, 1998, Zhang et al., 2000).

Goals and objectives. In my thesis, I develop two spatially explicit disease models featuring arthropod vectors that serve as within-field base models for arthropod vectored plant diseases that are relatively limited in their dispersal capacity (e.g. mites) and those that are capable of more locally active and occasional long-distance dispersal (e.g. whiteflies). My models were designed to accommodate alterations to fit a variety of vector-virus pathosystems by adjusting the spatial data and the parameters describing the survival, reproduction, transmission, and dispersal parameters of the modelled invertebrate vector. The specific objectives were:

1. To develop a spatially explicit simulation model for the spread of Blueberry Necrotic Ring Blotch Virus (BNRBV) in Southern Highbush Blueberry (SHB)

fields that integrates both the demography and dispersal of its suspected eriophyid mite vector in the genus *Calacarus*.

2. To validate the BNRBV model by comparing its projections to previously collected field data.
3. To develop a spatially explicit simulation model for the spread of a Whitefly-Transmitted Virus Complex (WTVC) in yellow squash (*Cucurbita pepo*) fields that incorporates both the demography and dispersal of its vector, the silverleaf whitefly, *Bemisia tabaci*.

LITERATURE CITED

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R. and Daszak, P. (2004) Emerging Infectious Diseases of Plants: Pathogen Pollution, Climate Change and Agrotechnology Drivers. *Trends in Ecology & Evolution,* **19,** 535-544.

Andrade-Piedra, J. L., Forbes, G. A., Shtienberg, D., Grünwald, N. J., Chacón, M. G., Taipe, M. V., *et al.* (2005) Qualification of a Plant Disease Simulation Model: Performance of the Lateblight Model across a Broad Range of Environments. *Phytopathology®,* **95,** 1412-1422.

Antoniouk, A. V. and Melnik, R. (2013) *Mathematics and Life Sciences*. Berlin: De Gruyter.

Avery, P. B., Kumar, V., Simmonds, M. S. and Faull, J. (2015) Influence of Leaf Trichome Type and Density on the Host Plant Selection by the Greenhouse Whitefly, Trialeurodes Vaporariorum (Hemiptera: Aleyrodidae). *Applied entomology and zoology,* **50,** 79-87.

Beasley, E. M., Aristizábal, N., Bueno, E. M. and White, E. R. (2022) Spatially Explicit Models Predict Coffee Rust Spread in Fragmented Landscapes. *Landscape Ecology,* **37,** 2165-2178.

Brauer, F. (2008) Compartmental Models in Epidemiology. In: *Mathematical Epidemiology.* (Brauer, F., van den Driessche, P. and Wu, J., eds.). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 19-79.

Bregaglio, S., Willocquet, L., Kersebaum, K. C., Ferrise, R., Stella, T., Ferreira, T. B., *et al.* (2021) Comparing Process-Based Wheat Growth Models in Their Simulation of Yield Losses Caused by Plant Diseases. *Field Crops Research,* **265,** 108108.

Byrne, D. N. and Bellows Jr, T. S. (1991) Whitefly Biology. *Annual review of entomology,* **36,** 431-457.

Calonnec, A., Cartolaro, P., Naulin, J.-M., Bailey, D. and Langlais, M. (2008) A Host-Pathogen Simulation Model: Powdery Mildew of Grapevine. *Plant Pathology,* **57,** 493-508.

Caubel, J., Launay, M., Lannou, C. and Brisson, N. (2012) Generic Response Functions to Simulate Climate-Based Processes in Models for the Development of Airborne Fungal Crop Pathogens. *Ecological Modelling,* **242,** 92-104.

Chappell, T. M., Codod, C. B., Williams, B. W., Kemerait, R. C., Culbreath, A. K. and Kennedy, G. G. (2020) Adding Epidemiologically Important Meteorological Data to Peanut Rx, the Risk Assessment Framework for Spotted Wilt of Peanut. *Phytopathology®,* **110,** 1199-1207.

Christopher P. Brooks, Janis Antonovics and Timothy H. Keitt (2008) Spatial and Temporal Heterogeneity Explain Disease Dynamics in a Spatially Explicit Network Model. *The American Naturalist,* **172,** 149-159.

Contreras-Medina, L. M. (2009) Mathematical Modeling Tendencies in Plant Pathology. *African journal of biotechnology.,* **8,** 7399.

Cooke, B. M., Jones, D. G. and Kaye, B. (2006) *The Epidemiology of Plant Diseases*. Springer.

Coop, L. (2023) Technical Documentation and Links for Plant Disease Risk and Other Hourly Weather Driven Models. Oregon IPM Center: Oregon State University.

Costa, H. S., Brown, J. K. and Byrne, D. (1991) Host Plant Selection by the Whitefly, Bemisia Tabaci (Gennadius),(Hom., Aleyrodidae) under Greenhouse Conditions. *Journal of Applied Entomology,* **112,** 146-152.

Cunniffe, N. J., Koskella, B., E. Metcalf, C. J., Parnell, S., Gottwald, T. R. and Gilligan, C. A. (2015a) Thirteen Challenges in Modelling Plant Diseases. *Epidemics,* **10,** 6-10.

Cunniffe, N. J., Stutt, R. O. J. H., DeSimone, R. E., Gottwald, T. R. and Gilligan, C. A. (2015b) Optimising and Communicating Options for the Control of Invasive Plant Disease When There Is Epidemiological Uncertainty. *PLOS Computational Biology,* **11,** e1004211.

DeAngelis, D. L. and Yurek, S. (2017) Spatially Explicit Modeling in Ecology: A Review. *Ecosystems,* **20,** 284-300.

Everts, K. L., Korir, R. C. and Newark, M. J. (2016) Re-Evaluation of Melcast for Fungicide Scheduling in Mid-Atlantic Watermelon. *Plant Health Progress,* **17,** 51-52.

Fabre, F., Coville, J. and Cunniffe, N. J. (2021) Optimising Reactive Disease Management Using Spatially Explicit Models at the Landscape Scale. In: *Plant Diseases and Food Security in the 21st Century.* (Scott, P., Strange, R., Korsten, L. and Gullino, M. L., eds.). Cham: Springer International Publishing, pp. 47-72.

Falcon, W. P., Naylor, R. L. and Shankar, N. D. (2022) Rethinking Global Food Demand for 2050. *Population and Development Review,* **48,** 921-957.

Felicio, T. N. P., Costa, T. L., Sarmento, R. A., Ramos, R. S., Pereira, P. S., da Silva, R. S., *et al.* (2019) Surrounding Vegetation, Climatic Elements, and Predators Affect the Spatial Dynamics of Bemisia Tabaci (Hemiptera: Aleyrodidae) in Commercial Melon Fields. *Journal of Economic Entomology,* **112,** 2774+.

Fenu, G. and Malloci, F. M. (2021) Forecasting Plant and Crop Disease: An Explorative Study on Current Algorithms. *Big Data and Cognitive Computing,* **5,** 2.

Ferriss, R. and Berger, P. (1993) A Stochastic Simulation Model of Epidemics of Arthropod-Vectored Plant Viruses. *Phytopathology,* **83,** 1269-1278.

Garrett, K. A., Madden, L. V., Hughes, G. and Pfender, W. F. (2004) New Applications of Statistical Tools in Plant Pathology. *Phytopathology®,* **94,** 999-1003.

Hardwick, N. V. (2006) Disease Forecasting. In: *The Epidemiology of Plant Diseases.* Springer, pp. 239-267.

Jeger, M. J. (2000) Theory and Plant Epidemiology. *Plant Pathology,* **49,** 651-658.

Jeger, M. J., van den Bosch, F., Madden, L. V. and Holt, J. (1998) A Model for Analysing Plant-Virus Transmission Characteristics and Epidemic Development. *Mathematical Medicine and Biology: A Journal of the IMA,* **15,** 1-18.

Jones, R. A. C. (2009) Plant Virus Emergence and Evolution: Origins, New Encounter Scenarios, Factors Driving Emergence, Effects of Changing World Conditions, and Prospects for Control. *Virus Research,* **141,** 113-130.

Jørgensen, L. N., Matzen, N., Ficke, A., Nielsen, G. C., Jalli, M., Ronis, A., *et al.* (2020) Validation of Risk Models for Control of Leaf Blotch Diseases in Wheat in the Nordic and Baltic Countries. *European Journal of Plant Pathology,* **157,** 599-613.

Juroszek, P., Bartsch, L., Fontaine, J. F., Racca, P. and Kleinhenz, B. (2022) Summary of the Worldwide Available Crop Disease Risk Simulation Studies That Were Driven by Climate Change Scenarios and Published During the Past 20 Years. *Plant Pathology,* **71,** 1815-1838.

Juroszek, P., Racca, P., Link, S., Farhumand, J. and Kleinhenz, B. (2020) Overview on the Review Articles Published During the Past 30 Years Relating to the Potential Climate Change Effects on Plant Pathogens and Crop Disease Risks. *Plant Pathology,* **69,** 179-193.

Kim, K.-H. and Jung, I. (2020) Development of a Daily Epidemiological Model of Rice Blast Tailored for Seasonal Disease Early Warning in South Korea. *Plant Pathol J,* **36,** 406-417.

Liu, S.-S., De Barro, P., Xu, J., Luan, J.-B., Zang, L.-S., Ruan, Y.-M., *et al.* (2007) Asymmetric Mating Interactions Drive Widespread Invasion and Displacement in a Whitefly. *Science,* **318,** 1769-1772.

Lof, M. E., de Vallavieille-Pope, C. and van der Werf, W. (2017) Achieving Durable Resistance against Plant Diseases: Scenario Analyses with a National-Scale Spatially Explicit Model for a Wind-Dispersed Plant Pathogen. *Phytopathology®,* **107,** 580-589.

Mayer, R. T., Inbar, M., McKenzie, C., Shatters, R., Borowicz, V., Albrecht, U., *et al.* (2002) Multitrophic Interactions of the Silverleaf Whitefly, Host Plants, Competing Herbivores, and Phytopathogens. *Archives of Insect Biochemistry and Physiology: Published in Collaboration with the Entomological Society of America,* **51,** 151-169.

McLeish, M. J., Fraile, A. and García-Arenal, F. (2020) Trends and Gaps in Forecasting Plant Virus Disease Risk. *Annals of Applied Biology,* **176,** 102-108.

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 Heterogeneous Landscapes: Spread of Sudden Oak Death in California (1990–2030). *Ecosphere,* **2,** art17.

Milenovic, M., Wosula, E. N., Rapisarda, C. and Legg, J. P. (2019) Impact of Host Plant Species and Whitefly Species on Feeding Behavior of Bemisia Tabaci. *Frontiers in Plant Science,* **10,** 1.

Miller, I. F., Jiranek, J., Brownell, M., Coffey, S., Gray, B., Stahl, M., *et al.* (2022) Predicting the Effects of Climate Change on the Cross-Scale Epidemiological Dynamics of a Fungal Plant Pathogen. *Scientific Reports,* **12,** 14823.

Miller, P. R. and O'brien, M. (1952) Plant Disease Forecasting. *The Botanical Review,* **18,** 547-601.

Nam, K. J. and Hardie, J. (2012) Host Acceptance by Aphids: Probing and Larviposition Behaviour of the Bird Cherry-Oat Aphid, Rhopalosiphum Padi on Host and Non-Host Plants. *Journal of insect physiology,* **58,** 660-668.

Nguyen, V.-A., Bartels, D. W. and Gilligan, C. A. (2023) Modelling the Spread and Mitigation of an Emerging Vector-Borne Pathogen: Citrus Greening in the U.S. *PLOS Computational Biology,* **19,** e1010156.

Ostfeld, R. S., Glass, G. E. and Keesing, F. (2005) Spatial Epidemiology: An Emerging (or Re-Emerging) Discipline. *Trends in Ecology & Evolution,* **20,** 328-336.

Parnell, S., Gottwald, T. R., Bosch, F. v. d. and Gilligan, C. A. (2009) Optimal Strategies for the Eradication of Asiatic Citrus Canker in Heterogeneous Host Landscapes. *Phytopathology,* **99,** 1370-1376.

Picard, C., Rimbaud, L., Hendrikx, P., Soubeyrand, S., Jacquot, E. and Thébaud, G. (2017) Peso: A Modelling Framework to Help Improve Management Strategies for Epidemics – Application to Sharka. *EPPO Bulletin,* **47,** 231-236.

Pietravalle, S., Shaw, M. W., Parker, S. R. and van den Bosch, F. (2003) Modeling of Relationships between Weather and Septoria Tritici Epidemics on Winter Wheat: A Critical Approach. *Phytopathology®,* **93,** 1329-1339.

Plant Diseases and Food Security in the 21st Century (2021). (Scott, P., Strange, R., Korsten, L. and Gullino, M. L., eds.). Springer Cham.

Rodelo-Urrego, M., Pagán, I., González-Jara, P., Betancourt, M., Moreno-Letelier, A., Ayllón, M. A., *et al.* (2013) Landscape Heterogeneity Shapes Host-Parasite Interactions and Results in Apparent Plant–Virus Codivergence. *Molecular Ecology,* **22,** 2325-2340.

Savary, S., Stetkiewicz, S., Brun, F. and Willocquet, L. (2015) Modelling and Mapping Potential Epidemics of Wheat Diseases—Examples on Leaf Rust and Septoria Tritici Blotch Using Epiwheat. *European Journal of Plant Pathology,* **142,** 771-790.

Severns, P. and Mundt, C. (2022) Delays in Epidemic Outbreak Control Cost Disproportionately Large Treatment Footprints to Offset. *Pathogens,* **11,** 393.

Severns, P. M. (2022) Dispersal Kernel Type Highly Influences Projected Relationships for Plant Disease Epidemic Severity When Outbreak and at-Risk Populations Differ in Susceptibility. *Life,* **12,** 1727.

Shi, X., Chen, G., Pan, H., Xie, W., Wu, Q., Wang, S., *et al.* (2018) Plants Pre-Infested with Viruliferous Med/Q Cryptic Species Promotes Subsequent Bemisia Tabaci Infestation. *Frontiers in Microbiology,* **9**.

Smith, D. L., Kerns, J. P., Walker, N. R., Payne, A. F., Horvath, B., Inguagiato, J. C., *et al.* (2018) Development and Validation of a Weather-Based Warning System to Advise Fungicide Applications to Control Dollar Spot on Turfgrass. *PLOS ONE,* **13,** e0194216.

Thierry, H., Monteil, C., Parry, H. and Vialatte, A. (2021) Simulating Seasonal Drivers of Aphid Dynamics to Explore Agronomic Scenarios. *Ecosphere,* **12,** e03533.

Vinatier, F., Tixier, P., Le Page, C., Duyck, P.-F. and Lescourret, F. (2009) Cosmos, a Spatially Explicit Model to Simulate the Epidemiology of Cosmopolites Sordidus in Banana Fields. *Ecological Modelling,* **220,** 2244-2254.

White, S. M., Bullock, J. M., Hooftman, D. A. P. and Chapman, D. S. (2017) Modelling the Spread and Control of Xylella Fastidiosa in the Early Stages of Invasion in Apulia, Italy. *Biological Invasions,* **19,** 1825-1837.

Zhang, P.-J., Wei, J.-N., Zhao, C., Zhang, Y.-F., Li, C.-Y., Liu, S.-S., *et al.* (2019) Airborne Host–Plant Manipulation by Whiteflies Via an Inducible Blend of Plant Volatiles. *Proceedings of the National Academy of Sciences,* **116,** 7387-7396.

Zhang, X.-S., Holt, J. and Colvin, J. (2000) A General Model of Plant-Virus Disease Infection Incorporating Vector Aggregation. *Plant Pathology,* **49,** 435-444.

Zhao, P., Zhang, X., Gong, Y., Wang, D., Xu, D., Wang, N., *et al.* (2021) Red-Light Is an Environmental Effector for Mutualism between Begomovirus and Its Vector Whitefly. *PLOS Pathogens,* **17,** e1008770.

CHAPTER 2

A SPATIALLY-EXPLICIT EPIDEMIOLOGICAL MODEL FOR BLUEBERRY NECROTIC RING BLOTCH VIRUS AND ITS SUSPECTED VECTOR, A *CALACARUS* MITE[1]

[1] Campbell, A. To be submitted to a peer-reviewed journal.

ABSTRACT

Blueberry necrotic ring blotch virus (BNRBV) is the causal agent of an emerging disease of Southern Highbush Blueberry (SHB; *Vaccinium corymbosum*). BNRBV causes non-systemic, localized infections on leaf tissue. The vector of BNRBV has not been confirmed but is suspected to be an eriophyid mite (family: Eriophyidae). We used previously collected field data reporting per-plant disease severity of BNRBV on susceptible 'Star' SHB plants to help construct and evaluate a spatially-explicit simulation model for the spread of BNRBV, assuming the vector to be a mite in the genus *Calacarus*. This SEI epidemiological compartmental model, called SimpleMite, integrated vector dispersal behavior (derived from raw field data), vector demography, and runs in the open-access environment HexSim. I selected two fields from 2011 and 2012 field seasons for model development and evaluation as the observed disease levels were on the highest and lowest ends of observed disease severity. If the model is put together properly, it should reasonably project similar disease levels and spatial distribution over time comparable to the field data. For each virtual field, disease severity values from the first data collection date in 2011 and 2012 were used as the initial disease conditions to replicate the spatial disease severity patterns observed at the penultimate and ultimate data collection dates. Moran's I, a global index of data aggregation over space, mean disease levels, the statistical distribution of disease levels over space on two different time points, and infection saturation heat maps produced by SimpleMite overlapped with the field data on those target dates and fields. The model performed best when replicating disease outcomes when disease levels were high, and less accurately represented the spatial distribution or quantity of infection severity values when disease

levels were so low that they would likely have little influence on blueberry yields or plant vigor. Model outcomes were inaccurate without the inclusion of a variable accounting for infected vectors preferentially dispersing from already diseased plants. This suggests that vector dispersal behavior may be an important and overlooked aspect of invertebrate-vectored plant disease. SimpleMite was conceptualized to be a base model for any arthropod-vectored plant disease, and has built-in aspects that should further facilitate the epidemiological studies of viral and bacterial plant disease dynamics.

INTRODUCTION

The co-occurrence of increased international trade and climate change presents significant challenges to plant pathologists attempting to control disease outbreaks (Hulme, 2009). Many factors relevant to the management of plant disease are changing simultaneously and prediction of overall plant health trends sometimes yield conflicting results – for example, changes in precipitation patterns are predicted to negatively impact the growth of corn in some regions, but accompanying rises in CO_2 may temper the effects (Juroszek *et al.*, 2020). The flow of global commerce facilitates the introduction of pathogens into previously uncolonized areas, introducing further complexity into prediction of plant health outcomes (Fabre *et al.*, 2021, Hulme, 2009). Under these conditions, understanding host-pathogen interactions can be incredibly complicated, and further complexity is introduced with the consideration of vector species. For these reasons, predictive tools are increasingly valuable to the study and management of plant diseases, especially that of arthropod-vectored plant pathogens. The use of these models on smaller scales could elucidate the finer details of plant, pathogen, and vector biology that are needed for the construction of broader climate change models (Jones, 2016, Juroszek et al., 2020).

The Georgia blueberry industry is periodically threatened by Blueberry Necrotic Ring Blotch Virus (BNRBV; *Blunervirus vaccinium*), the causal agent of Blueberry Necrotic Ring Blotch Disease (BNRBD), an emerging disease of Southern Highbush Blueberry (SHB; *Vaccinium corymbosum*). The disease was first observed in Georgia in 2006 and is currently widespread across the southeastern U.S., appearing to be restricted to this region (Martin *et al.*, 2012, Martin & Tzanetakis, 2018, Ramos-González *et al.*,

2023). Because Georgia is typically one of the top three blueberry-producing states in the top blueberry-producing nation, BNRBD is an issue of both local and national importance (Upadhaya & Dwivedi, 2019). Consumer demand for blueberries and blueberry acreage are increasing, and BNRBD prevalence remains unpredictable within and between growing seasons (Fulcher *et al.,* 2015, Quito-Avila *et al.,* 2013). The development of predictive tools for the distribution and prevalence of this disease could potentially provide important insight to its superficially understood disease dynamics. A more holistic understanding of BNRBD dynamics may lead to more effective management practices when disease outbreaks occur.

BNRBV is in the genus *Blunervirus* in the family *Kitaviridae*. Kitaviruses are a newly recognized family of positive-sense single-stranded RNA plant viruses. The only known and documented vectors of kitaviruses are eriophyid mites (family: Eriophyidae) and mites in the genus *Brevipalpus* (family: Tenuipalpidae) (Ramos-González et al., 2023). Members of this virus family cause localized infections in plant foliar tissues and lack the capacity for systemic movement within their hosts, which is atypical of plant viruses (Agrios, 2005, Ramos-González et al., 2023). BNRBV symptoms present as irregularly shaped, darkly-discolored to necrotic leaf spots, some with green centers. Aggregations of these infections may coalesce into a larger, amorphous infected region, potentially triggering leaf abscission. Typical of kitaviruses, BNRBV appears to be localized to these necrotic spots, does not move systemically within its host, and consequently does not persist in host plants from one growing season to the next (Robinson *et al.,* 2016). The primary source of BNRBV in agricultural environments remains unknown as it appears to recolonize blueberry plantings each spring (Robinson et

al., 2016). These observations suggest that the virus possibly overwinters outside of blueberry fields (alternate hosts are presently unknown) and recolonizes from the field margins where the virus reservoirs and vector(s) exist. The impact of defoliation caused by severe cases of BNRBV on fruit fresh weight yield the following spring is currently unknown but a possibility on young plantings or if the disease pressures are high for consecutive growing seasons. Reduction in yield is physiologically plausible given that artificial defoliation has been shown to reduce the yields (sometimes by ~43% of grams of fresh weight per plant) of some SHB cultivars (Williamson & Miller, 2002).

To date, no organism has been definitively shown to vector BNRBV and the virus appears unable to spread through vegetative propagation or seed (Robinson et al., 2016). However, researchers suspect that an eriophyid mite in the genus *Calacarus* may be serving as a vector of BNRBV due to the occurrence of these mites in association with the virus (Burkle *et al.*, 2012, Cantu-Iris *et al.*, 2013, Ramos-González et al., 2023). Unpublished data reported by A. D. Tassi and D. Carrillo documents an association between BNRBV and the presence of both a *Calacarus* sp. and *Brevipalpus* sp., but the authors assert that between the two, only *Calacarus* was able to perform as a vector (Ramos-González et al., 2023). Though confirmation of the vector species for BNRBV is an area of active research, its genetic similarity to other Kitaviruses (many of which are vectored by eriophyid mites), clustered pattern of disease spread, and association with an eriophyid species suggests that BNRBV is also likely to be mite-transmitted (Robinson et al., 2016).

Eriophyid-vectored viruses are poorly understood. The family Eriophyidae belongs to Eriophyoidea: a superfamily of microscopic, four-legged, phytophagous mites.

These mites tend to have highly specialized, monophagous relationships with their host plants and several species are confirmed vectors of plant viruses (Baker & Wharton, 1952, Vacante, 2015, Westphal & Manson, 1996). Members of Eriophyoidea have been described since the mid-19th century, but much of what is known about them was discovered in recent history (Lindquist & Amrine, 1996). Though Eriophyidae is the second-most important family of mite plant pests after the Tetranychidae (superfamily: Tetranychoidea), the majority of eriophyoid literature has been published in the last fifty years (de Lillo & Skoracka, 2010, Lindquist & Amrine, 1996). There is a dearth of literature on both the behavior and molecular biology of eriophyoids in particular (de Lillo & Skoracka, 2010, Michalska *et al.*, 2010). As a result, the study of eriophyoid mite-vector pathosystems is slowly progressing but still in its infancy.

The objective of this study was to develop a spatially-explicit simulation model to predict BNRBD disease incidence based on reported life history, assumed mite vector, and a previous two-year field study of BNRBD spread in multiple south Georgia production blueberry fields. Much remains a mystery about BNRBD. In the absence of sufficient literature on the epidemiology of this disease, a simulation model acts as a bridge between what is relatively well-known and unknown about this pathosystem. This model allows for the evaluation of disease spread under various conditions based on our knowledge of the host, our current understanding of the virus, and the life history of the presumed vector. With these limitations in mind, this study provides a framework for the predictive modeling of arthropod-vectored plant disease that addresses uncertainty and environmental/epidemiological variation with stochastic processes embedded within the model as a parameter that can be varied. While my research can promote the collection of

high-quality life history data of eriophyid mites and further study of BNRBV to refine projections, it also provides plant pathologists with an open access base model for the spread of arthropod-vectored plant diseases. I have constructed this model with the vision that the parameter values can be readily changed to represent other invertebrate-vectored plant disease systems. There is no similar GPS enabled, spatially explicit, open access model for the study of invertebrate vectored plant diseases.

MATERIALS AND METHODS

Field data. This study utilized raw field data on BNRBD presented by Tanisha Robinson (2013). These raw data (hereafter collectively referred to as 'Robinson Field Data') list the per plant disease severity of BNRBD in several Georgia blueberry fields on the susceptible southern highbush blueberry (SHB) cultivar 'Star' and were recorded periodically across several weeks in 2011 and 2012 (see **Table 2.1**). Disease severity was reported as the percentage of leaves (up to a maximum of 100%) that displayed BNRBD symptoms on each observation date. 'Enigma (Site 1)' in Enigma, GA and 'Homerville' in Homerville, GA (hereafter referred to as 'Enigma' and 'Homerville') were selected for study because disease levels spanned the low and high ends of the recorded disease levels and they were consistent in their arrangement of alternating rows of 'Star' with rows of 'Emerald' – an SHB cultivar that is resistant to BNRBD. Blueberry plants in the selected field sites were spaced 1.2 m between plants within a row and 3.0 m between rows (Robinson, 2013).

Model overview. A simulation model for the spread and disease incidence of BNRBV, hereafter termed 'SimpleMite', was developed in the HexSim modeling

environment using HexSimPLE as a template which was modified to complement the BNRBV system. HexSim is a spatially-explicit simulation modelling environment for ecological scenarios and HexSimPLE is a females-only population model template constructed within it used for the analysis of spatial metapopulation dynamics (Schumaker, 2023). SimpleMite represents a population of vectors moving throughout a landscape of SHB plants and is adapted from the event sequence for HexSimPLE. See **Appendix A** for a summarized overview of HexSim or visit www.hexsim.net for a detailed description of the entire landscape event simulation environment. HexSimPLE (and by extension SimpleMite) is constructed of four basic parts:

1) Spatial data specifying the landscape in which the modeled populations of pathogen, host, and vector exist
2) Demographic and biological information on the pathogen, host, and vector
3) A set of traits and accumulator values that describe each individual vector and host in the model
4) An event sequence specifying all actions performed by the individuals (agents) in each population (pathogen, vector, host) and their interactions with the landscape

Spatial data.

SimpleMite requires the input of several spatial data called "Hexmaps". Given that HexSim is a spatially-explicit modeling environment, these Hexmaps are constructed of hexagons with a user-specified size that correspond to real life dimensions. Hexagons in SimpleMite have a width of 1.5 m, which is about the distance across each individual SHB plant canopy and half the distance between rows (Fang *et al.*, 2020, Robinson,

2013).The spatial data required for SimpleMite are mostly those used to populate the HexSimPLE routine, with the addition of an "Initial Disease Conditions Map". The maps required for the implementation of HexSimPLE are: Matrices Map, Habitat Map, Regions Map, Stress Map Fecundity, and Stress Map Survival. These maps are described extensively in **Appendix A** but are briefly described below. The Matrices Map determined which hexagons in the landscape are assigned a Leslie Matrix to estimate and project the mite population size. The Habitat Map determined the location of each host plant and border rows. The Regions Map determines which hexagons in the landscape are considered unique regions that should be subject to their own independent environmental stochasticity. Stress Map Fecundity and Stress Map Survival designate the hexagons in the landscape in which mites are subjected to specified/hypothesized reductions on their fecundity and survival, respectively. The Initial Disease Conditions Map represents the distribution of mites and disease at the beginning of the first time step of the model and is a recreation of the disease conditions from the first data collection date of Robinson Field Data.

Demographic information.

SimpleMite retains many of the population-focused qualities of HexSimPLE, with significant additions to represent BNRBV disease dynamics. The vector population was divided into three life stages: Stage 0 (eggs), Stage 1 (juvenile mites), and Stage 2 (adult mites). Survival and fecundity rates for each life stage were gathered from published life tables for *Calacarus* and other closely related, ecologically similar eriophyid mite species (**Table 2.2**). Population size was projected using a Leslie Matrix (**Appendix A**), a well-known demographic projection method, based on vital rates (**Table 2.2**) for each life

stage at each time step (a four-day interval) as part of the event sequence (Kajin *et al.*, 2012). In SimpleMite, Stages 0 and 1 are incapable of both dispersal and serving as vectors of BNRBV, which is biologically realistic as eriophyoids do not reach sexual maturity until the final molt and dispersal is generally only carried out by adult females (Alberti & Nuzzaci, 1996, Lindquist & Oldfield, 1996). Because HexSimPLE is a females-only model, demographic data in SimpleMite represent the fecundity and survival of and the per-plant carrying capacity for female *Calacarus* mites only. Because vagrant eriophyoid mites usually have a 50:50 sex ratio, carrying capacities of SHB for *Calacarus* mites estimated from the literature were halved (Sabelis & Bruin, 1996). To lessen the computation load on the model and encourage a more efficient run time, 1 mite in the model corresponds to 100 real mites. However, these ratios are variables in SimpleMite and could be changed depending on the acquisition of more refined information or the specific questions being evaluated.

Traits and accumulators.

Because characteristics of the hosts and vectors are required to model more biologically realistic disease outcomes, traits and accumulator values in SimpleMite describe individuals in both the host and vector populations. Traits describe a categorical characteristic of the population (e.g. yellow *vs.* blue), while accumulators describe some numerical characteristic of the population (e.g. the number of adult mites present on a host plant). Traits may, and often do, change throughout the model depending on the events in the simulation that the population members are subject to. Certain traits ("accumulated traits") may change as a result of the values of the accumulators associated with each individual. Therefore, specified events, called "accumulate events",

can change accumulated traits through changes in accumulators. Traits that are not associated with any accumulators are "probabilistic traits" and can only be altered by "transition events". Traits describing members of the vector population include: "Viruliferous" vs. "Non-Viruliferous". This trait is dependent upon whether the host plant the vector disperses from is BNRBV-Positive and can produce viruliferous (mite) vectors. An accumulator describing a member of the vector population is "Dispersal Distance". This accumulator determines how many hexagons away from the source hexagon (plant) a vector disperses. It is calculated based on a dispersal kernel for BNRBD derived from Robinson Field Data (see **Figure A.1** in **Appendix A**). This dispersal gradient was determined by identifying isolated instances of disease from Robinson Field Data to serve as "source" plants. The infection severity and distance from source plants was recorded for plants surrounding the source that became diseased. Model fitting for the dispersal kernel function was performed using the R package, drc (Ritz *et al.*, 2016).

Traits describing the individual host plants are: "BNRBV-Positive" vs. "BNRBV-Negative", "Viruliferous Vector Production – Yes" vs. "Viruliferous Vector Production – No", "Viruliferous Mites Acquired – Yes" vs. Viruliferous Mites Acquired – No", and "Lesion Category – Mild Lesions" vs. "Lesion Category – Moderate Lesions" vs. "Lesion Category – Severe Lesions". These sets of traits describe the disease status of the plant, whether the plant can produce viruliferous vectors, whether the plant has acquired viruliferous vectors within the time step, and the estimated number of lesions resulting from the BNRBV infection, respectively. Accumulators describing individual host plants include: "Infection Saturation", "Viruliferous Mites Acquired", and accumulators that

quantify the total number of mites and the number of mites of each life stage on a per-plant basis. "Infection Saturation" was calculated based on both the number of mites and the disease status of the plant. If a host plant has the "BNRBV-Positive" trait, the "Infection Saturation" accumulator increases with the number of Stage 2 mites on the plant, achieving a maximum of 100 when the number of Stage 2 mites reaches the carrying capacity (see **Equation 2.1**). "Infection Saturation" is scored out of 100 and determines whether host plant has the "Lesion Category – Mild" ($25 \leq$ Infection Saturation < 50), "Lesion Category – Moderate" ($50 \leq$ Infection Saturation < 75), or "Lesion Category – Severe" (Infection Saturation ≥ 75) trait. The "Viruliferous Mites Acquired" accumulator represents whether viruliferous mites were acquired by a host plant during the time step and it determines whether the host plant has the "Viruliferous Mites Acquired – Yes" trait (if accumulator is greater than zero) or the "Viruliferous Mites Acquired – No" trait (if the accumulator is not greater than zero).

Event sequence.

The SimpleMite event sequence was constructed of five core event groups: Initialize Simulation, Conduct Movement, Perform Demography, BNRBV Disease Dynamics, and Get Output Map Data. Each time that the sequence of events in the simulation was run from start to finish is called a "time step". In the SimpleMite model, each time step represented four days of real time. Some events and event groups from the HexSimPLE base model (none of which are listed or described here) that are specifically relevant for extracting data relevant to population studies were toggled off but retained in the model for potential future uses. Many events in the SimpleMite event sequence that carried over from the HexSimPLE base model were amended or altered to better

represent *Calacarus* mite biology, BNRBD disease dynamics, and/or to facilitate ease of reporting relevant data.

The Initialize Simulation event group determines the arrangement of host plants in the landscape, the starting disease conditions, and number of mites per plant. This event group runs only on the first time step. The Conduct Movement event group determined how many mites were dispersing from host plants and to which hexagons the mites dispersed. The Perform Demography event group determined how many mites were present in each hexagon of the landscape after reproduction and survival was executed for each time step. The BNRBV Disease Dynamics event group determines which host plants in the landscape have become infected with BNRBV and the severity of their infection. The BNRBV Disease Dynamics event group was triggered to run every other time step (starting with the second time step) to approximate a latent period of about 8 days to be a full multiple of the time step (4 days). When placed adjacent to diseased plants, virus-free SHB plants developed symptoms in 2-3 weeks, so we assumed a latent period for BNRBD of about ~1 week (Martin et al., 2012, Robinson, 2013). The Get Output Map Data event group calculates the population parameters of the mites on each host plant. In the original HexSimPLE model, this information would have been used in a subsequent event group to construct maps of the population in the simulated landscape, but this event group was unnecessary for the purposes of SimpleMite and was toggled off to conserve computing power and reduce simulation run time.

***Calacarus* mites as presumed vectors.** This model is constructed with the assumption that an eriophyid mite in the genus *Calacarus* is the only vector for BNRBV. Spatiotemporal disease patterns described by Robinson (2013) support the idea of a slow-

moving vector with limited capacity for long-distance dispersal. *Calacarus* also has a wide geographic distribution, having been found in all tropical regions of the world and several subtropical regions (Oldfield, 1996, Vacante, 2015). Some species of *Calacarus* are known to have wide host ranges, which is unusual for the Eriophyoidea, which tend to be host specialists (Oldfield, 1996). Mites of this genus are known to exist in Georgia on camellia (Keifer *et al.*, 1982). Although the role of a *Calacarus* sp. as a vector has not been definitively proven by Koch's postulates in the literature, *Calacarus* mites have been found in high numbers in association with BNRBD (Burkle et al., 2012, Cantu-Iris et al., 2013).

SimpleMite requires the input of demographic and life history information of the vector. These parameters were taken from the literature on eriophyid and *Calacarus* mites. A selection of the parameters used for the model is provided in **Table 2.2**. Immature stages of the mites were assigned a fecundity of zero, while adult mites were assigned a fecundity of 6 (T.J. Van der Merwe & Coates, 1965). Given that SimpleMite is based on a females-only model and has a four-day time-step, this value represents the number of eggs laid in a four-day period that will become female mites. This value is based on the mean ovipositional rate for a *Calacarus* species and assumes a 50:50 sex ratio. The per-plant carrying capacity was calculated by taking a maximum number of *Calacarus* mites per cm^2 from the literature (17.4 cm^2) and multiplying this by the average SHB bush size, estimating 1000 leaves per bush and obtaining average leaf size of 'Star' SHB plants from another study (Severns, unpublished) (Fournier *et al.*, 2004).

SimpleMite event sequence. The SimpleMite event sequence consists of the five aforementioned event groups in the following order: Initialize Simulation, Conduct

Movement, Perform Demography, BNRBV Disease Dynamics, and Get Output Map Data. The event sequence is best understood when conceptualized not as a list of events that occur but as a cycle that repeats with each time step. Some events contained within event groups seem out of place if the event sequence is not considered in this way. For a simplified representation of the disease cycle contained within the event sequence, see **Appendix A**, **Figure A.2**.

Initialize Simulation.

This event group sets up the simulation to run and is triggered to only run on the first time step. It assigns one individual to each of the cells in the Matrices Map. These individuals (collectively, "Hosts") are then labeled as having a "Host Type" of either "Gap" or a "Plant" depending on if the cell they occupy is marked with a 0 or a 1 on the Habitat Map, respectively. For both Homerville and Enigma, Gaps corresponded to the Robinson Field Data map that were not 'Star' SHB plants, and Plants represented susceptible 'Star' SHB plants. In this specific simulation, only Plants can support populations of mites *and* become infected with BNRBV. Also in this event group, the Regions Map sets the number of unique regions for which environmental stochasticity will be determined each time step. For all the simulations described herein, the entire virtual landscape was considered one region, the environmental conditions of which varied with each time step. This was a logical choice given that all Robinson field data maps corresponded to a small geographic area (a portion of one SHB field). A sub-event group called Set Initial Disease Conditions makes use of the Initial Disease Conditions Map. This map lists the per-plant infection severity of the first data collection date of Robinson Field Data (see **Table 2.1**).

Plants that corresponded to a 'Star' plant with an infection severity score greater than zero were assigned the "Viruliferous Mites Acquired – Yes" trait. The disease status trait of all Plants and Gaps was updated in a subsequent transition event. **Table 2.3** represents the decision table used for that transition event. As a result, all Plants that were designated as having acquired viruliferous mites were BNRBV-Positive. The value of the infection severity was also assigned to the "Infection Saturation" accumulator. In the SimpleMite model, BNRBV-Positive Plants can only produce viruliferous vectors when their "Infection Saturation" accumulator (a variable in the simulation) had a value greater than fifty (this is the "Viruliferous Vector Threshold" described in **Table 2.2**). This decision was made based on a sensitivity analysis (unpublished) that demonstrated that disease was substantially overestimated when all "BNRBV-Positive" Plants produced viruliferous vectors, but severely underestimated when an infection saturation of 100 was required to produce viruliferous vectors. After the BNRBV-Positive Plants are assigned an "Infection Saturation" accumulator value, they are then marked as "Viruliferous Vector Production – Yes" or "Viruliferous Vector Production – No" if the accumulator value is greater than 50 or not greater than 50, respectively.

Finally, the initial mite population was added to the model as part of this event group. Mite populations were only added to "BNRBV-Positive" Plants. The quantity of mites added to each infected plant was calculated based on the Infection Severity Value of that plant: the value of the "Infection Saturation" is multiplied by both 1/100 and the per-plant carrying capacity (see "Carrying Capacity [per-hexagon]" in **Table 2.2**). This calculation is an adaptation of **Equation 2.1** in order to keep the relationship between

Infection Saturation and the number of Stage 2 mites consistent. All mites introduced into the model during this event group are Stage 2 (adults).

Conduct Movement.

The event group begins by setting the "Viruliferous Mites Acquired" accumulator to zero for all Plants. The mite carrying capacity for the Hosts is determined by the Habitat Map, meaning that Plants had a carrying capacity of 4,210 simulated mites (with each simulated mite representing 100 real mites) and Gaps having a carrying capacity of zero. The number of mites that disperse from each Plant was dependent upon the carrying capacity for each Plant. Only the number of mites that exceeded the carrying capacity could disperse from each Plant and only Stage 2 mites (adults) can disperse from Plants. The default version of this event group in HexSimPLE allows for the dispersal of all three life stages. It assigns dispersal distances to the appropriate number of Stage 0, Stage 1, and Stage 2 individuals that will be dispersing. Because Stage 0 and Stage 1 mites do not disperse in SimpleMite (which is representative of the BNRBV pathosystem), the dispersal distances were set to zero for these immature mites and only calculated for Stage 2. Dispersing Stage 2 mites were either assigned the "Non-Viruliferous" or "Viruliferous" trait, depending on if they dispersed from a Plant that had the "Viruliferous Vector Production – No" or "Viruliferous Vector Production – Yes" trait, respectively. Once the dispersers moved their assigned distance and direction, the number of "Viruliferous" mites received by a Plant was marked on a map called "Viruliferous Stage 2".

Perform Demography.

Storing mite population sizes as accumulator values facilitates the use of Leslie Matrices to calculate population growth in this event group. For more information about Leslie Matrices and the calculation of population size, please reference **Appendix A** and Schumaker (2023). In this event group, the number of individuals in each mite life stage is determined after reproduction and survival for each Plant (a single hexagon in SimpleMite).

BNRBV Disease Dynamics.

In order to simulate a latent period of ~8 days, this event group only triggers every even time step. The event sequence begins by determining which Plants have acquired "Viruliferous" vectors by referencing the Viruliferous Stage 2 Map and assigning the number received by each Plant to its "Viruliferous Mites Acquired" accumulator, which updates the "Viruliferous Mites Acquired" trait. A transition event then determines the updated disease status of each Plant. The decision table used for this transition event is represented in **Table 2.4** and is distinct from the decision table used in the Initialize Simulation event group (**Table 2.3**) in that there is a 50% chance that a previously BNRBV-Negative Plant that has the "Viruliferous Mites Acquired – Yes" trait becomes BNRBV-Positive instead of a 100% chance. Preliminary data (unpublished) demonstrated that this 50% reduction in probability prevents the overestimation of "Infection Saturation" when used as a direct comparison to the reported infection severity in Robinson Field Data. Following this transition event, Infection Saturation was calculated for each BNRBV-Positive Plant using **Equation 2.1**. Because this "Infection Saturation" accumulator value is meant to serve as a direct comparison to the reported infection severity in Robinson Field Data, this equation sets the maximum value for

Infection Saturation at 100. As discussed previously in this chapter, the "Infection Saturation" accumulator value corresponds to the "Lesion Category" trait. Though no analyses were performed with this trait, it was included in the model to serve as a placeholder for a consequence of BNRBD such as percent defoliation or yield loss. To date, the relationship between BNRBD severity and these disease outcomes are unknown. Despite this, the inclusion of the trait provides the scaffolding for the representation of either of these relationships in the simulation, should they be elucidated, or the user wishes to evaluate a range of potential scenarios. The event group concludes by constructing two maps: one map of the locations of "BNRBV-Positive" Plants and another map that records the "Infection Saturation" accumulator value for each Plant to its location.

Get Output Map Data.

This event group calculates several values related to population size and dispersal that in HexSimPLE were used to construct demographic output maps in a subsequent event group. The construction of these maps has been toggled off in SimpleMite but this event group was retained to prevent the creation of unforeseen errors in the simulation. For more information on this event group, please reference Schumaker (2023).

Model output and analysis. I constructed virtual fields for both Enigma and Homerville in the HexSim modelling environment. Each 'Star' and 'Emerald' SHB plant from Robinson Field Data was represented as one hexagon, and two hexagons of spacing were included between SHB plant rows unless otherwise indicated. For each of these fields, infection severity data from the first collection date in Robinson Field Data from 2011 and 2012 was used for the Initial Disease Conditions Map (**Table 2.1**). This resulted

in four unique versions of the SimpleMite model. For each of these four versions, three hundred simulation replicates were generated for a total of 1,200 simulation results files. For each set of results files, the Infection Saturation Map was analyzed for the time steps corresponding to the penultimate and ultimate data collection dates - a grand total of 2,400 Infection saturation maps were produced and analyzed.

Some elements of the model vary stochastically with each time step. Each region in the Regions map (in the case of SimpleMite, there is only one) experiences regional effects on fecundity and survival (to represent environmental stochasticity; for more on this, see **Appendix A**). The degree to which the fecundity and survival values vary is determined by this environmental stochasticity and the Percent Stochasticity values for fecundity and survival (see **Table 2.2**; for more details, see **Appendix A**). This means that for each version of SimpleMite, all results are different from one another and this intentionally introduced stochasticity is a method to address the contribution of varying environmental scenarios on outcomes of the same outbreak. This approach is particularly useful to represent how explosive or predictable diseases can be when major epidemiological conditions randomly vary (as is often the case with disease epidemics).

Analyses of simulation outputs focus on the Infection Saturation Map generated in the BNRBV Disease Dynamics event group. The closest corresponding time step of the simulation was directly compared to the infection severity field data of the penultimate and ultimate collection dates for Robinson Field Data (see **Table 2.1**). Because each time step of the simulation corresponds to ~4 days of real time and because the BNRBV Disease Dynamics event group only runs every even time step, the closest corresponding simulation time step to any given data collection day was calculated by

dividing the data collection day number by four and choosing the closest even number in the simulation runs. HexSim produces output maps in a proprietary Hexmap format (".hxn"). All 1,200 Infection Saturation Map .hxn files were converted to CSV files for ease of analysis using a C program called "hxn2csv" available under "Accessories & Programs" on the "Resources" page of the HexSim website (www.hexsim.net).

Model Evaluation for Mean Infection Saturation.

The mean infection saturation (total of the "Infection Saturation" values of each Plant in the simulation divided by the number of Plants in the simulation) was calculated for each of the 300 simulation replicates of each field (Homerville or Enigma), year (2011 or 2012), and observation date (penultimate or ultimate). As a general summary of model projection performance, these mean values were compared with each other and to a notched boxplot which visually represented the breadth of infection severity for the penultimate and ultimate data collection dates. If the model performed well, I expected to observe an overlap in ranges of the notched boxplots for the simulation infection saturation mean values and the field infection severity data.

In a more refined comparison of the model projections and the field data, simulation infection saturation values were assigned to percentile groups based on their mean infection saturation value using the R package, dplyr (Wickham *et al.*, 2023). These percentile groups represented the extremes on the low and high disease severity projections while conserving the middle percentile values: the 0% to 10% percentile group encompassed the simulation runs in the 10^{th} percentile and below, the 10% to 50% percentile group encompassed all replicates between the 10^{th} percentile up to the 50^{th}, the 50% to 90% percentile group encompassed replicates between the 50^{th} up to the 90^{th}, and

the 90% to 100% percentile group encompassed replicates above the 90th percentile. Because these percentile groups were calculated based on the field, year, and collection date, the same simulation run could be grouped into two different percentile groups depending on the analysis. For example, replicate 207 for the Enigma 2012 simulation could be grouped into the 10% to 50% percentile group when grouping by the means of the time step corresponding to the penultimate collection date and be grouped into the 50% to 90% percentile group when grouping by the means of the time step corresponding to the ultimate collection date. Once the simulation replicates were separated into these groups, they were again visually compared with the breadth of infection saturation values reported for the corresponding field, year, and collection date combination with a notched boxplot generated using the R package ggplot2 (Wickham, 2016). Superimposed onto this boxplot were the maximum values for each simulation replicate, also organized by the mean-based percentile group.

Total Infection Saturation.

The total infection saturation was calculated for each of the 300 simulation replicates of each field, year, and corresponding collection date combination. Total infection saturation was calculated as the total of the "Infection Saturation" values of each Plant in the simulation. These values were separated into the mean-based percentile groups described previously in this chapter. A notched boxplot of the range of total infection saturation values was visually compared to the singular value for total infection severity obtained from Robinson Field Data. If the model performed well, I am expecting that the total infection saturation values overlap the total infection severity value for the field data. The notched boxplot was constructed in ggplot2 (Wickham, 2016).

Moran's I Analysis.

A subset of 20 simulation replicates was selected for each field, year, and corresponding collection date combination. Moran's I analysis was performed on the Infection Saturation Map of each of the subset of 20 simulation replicates and the corresponding Robinson Field Data infection severity map. I used Moran's I to determine similarities in spatial aggregation of disease between simulation outputs and field data. Moran's I is a commonly used test for global (as opposed to local) spatial autocorrelation (or the degree of aggregation over space), (Gedamu *et al.*, 2024) and compares the value of a particular attribute of any one location to that of its neighbors. Moran's I values range from -1 to 1, with a value of 1 indicating that data are perfectly aggregated in space, a zero indexing no aggregation, and a -1 indicating perfect repulsion of the data. Moran's I analysis was performed using the Moran.I() function of the R package, ape (Paradis & Schliep, 2019) on altered versions of each Infection Saturation Map that only contained the cells in which Plants were located and altered versions of the infection severity maps from Robinson Field Data that only contain 'Star' SHB plants.

To select a subset of 20 simulations, I first determined the average infection severity value for each map selected from Robinson Field Data and compared this value to the mean infection saturation values of the percentile groups for the corresponding simulations. I randomly selected twenty replicates from the percentile group that had a range of mean infection saturation values that were most similar to the mean infection severity calculated from Robinson Field Data infection severity map of interest. For example, if the mean infection severity obtained from Robinson Field Data fell within the range of the means of the replicates in the 90% to 100% percentile group, the subset of 20

simulation replicates would be selected from the 90% to 100% percentile group. If the mean obtained from Robinson field data did not fall within any of the percentile groups, the subset of 20 was selected from either the 0% to 10% percentile group, or the 90% to 100% percentile group, depending on if it the mean was lower or higher than all of the simulation means, respectively. The subset of 20 was randomly selected from the chosen percentile group using the R package, kimisc (Müller, 2017).

Infection Saturation Heatmaps.

To determine whether the simulations were generating reasonable outputs, I created infection saturation heatmaps to visualize and qualitatively compare field data to simulations. Infection saturation maps were constructed using ggplot2 for five replicates of each field, year, and collection date combination (Wickham, 2016). These heatmaps were arranged into one figure using the R package ggpubr (Kassambara, 2023). The five replicates were randomly selected from the subset of 20 replicates discussed previously in this chapter. This random selection was performed using the kimisc R package (Müller, 2017). The heatmaps for both the field and simulation data only represent the locations containing 'Star' SHB plants or simulated Plants, respectively.

RESULTS

Homerville 2011 – Penultimate.

When I plotted the field data from Homerville in 2011 from the penultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation values overlapped with the field data (**Figure 2.1**), indicating that despite a large number of stochastic

simulations the field data outcome was represented from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 2.9**). The maximum infection saturation values of the simulation percentile groups also overlapped with the fourth quartile (upper 25%) of the field data but the max outlier simulation values were not as great as the max field severity values (**Figure 2.9**). Total infection severity for field data fell within the range of total infection saturation values for the >90% percentile group (**Figure 2.17**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.25**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.25**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data and qualitatively the disease severity heatmaps based on field data and randomly selected simulation runs produced similar levels, locations, and patchiness of BNRBD (**Figure 2.29**).

Homerville 2011 – Ultimate.

When I plotted the field data from Homerville 2011 from the ultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation values overlapped with the field data (**Figure 2.2**), indicating that despite a large number of stochastic simulations the field data outcome was represented from the same initial disease configuration. When the infection saturation means were separated into percentile groups,

the range of all groups overlapped with the corresponding field data (**Figure 2.10**). The maximum infection saturation values of the simulation percentile groups also overlapped with the fourth quartile (upper 25%) of the field data and the max outlier simulation values overlapped with the field severity values (**Figure 2.10**). Total infection severity for field data fell within the range of total infection saturation values for the <10% percentile group (**Figure 2.18**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.25**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.25**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data and qualitatively the disease severity heatmaps based on field data and randomly selected simulation runs produced similar levels, locations, and patchiness of BNRBD (**Figure 2.30**).

Homerville 2012 – Penultimate.

When I plotted the field data from Homerville 2012 from the penultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation only overlapped with the field data outliers (**Figure 2.3**), indicating that the field data outcome was not represented from the same initial disease configuration.. When the infection saturation means were separated into percentile groups, the range of all groups only overlapped with some of the outliers from the corresponding field data (**Figure 2.11**). The maximum infection saturation values of the simulation percentile groups did not overlap with the

fourth quartile (upper 25%) of the field data, nor did the max outlier simulation values overlap with the field severity values (**Figure 2.11**). Total infection severity for field data fell within the range of total infection saturation values for the <10% percentile group (**Figure 2.19**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were negative and had a p-value that was not statistically significant ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.26**). The Moran's I value of the corresponding field data was also negative and had p-value that was not statistically significant p-value ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.26**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data, but p-values for both field and simulation data were not significant. Qualitatively, the disease severity heatmaps based on field data and randomly selected simulation runs produced similar locations and patchiness of BNRBD, though the field data showed more incidences of low disease levels (**Figure 2.31**).

Homerville 2012 – Ultimate.

When I plotted the field data from Homerville 2012 from the ultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation values overlapped with the field data (**Figure 2.4**), indicating that despite a large number of stochastic simulations the field data outcome was represented from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 2.12**). The maximum infection saturation values of the simulation percentile groups also overlapped with the max field severity values (**Figure 2.12**). Total infection severity for field data

exceeded and does not fall within the range of total infection saturation values for the percentile groups (**Figure 2.20**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were negative and had a p-value that was not statistically significant ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.26**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 2.5** and **Figure 2.26**). Qualitatively, the disease severity heatmaps based on field data and randomly selected simulation runs produced dissimilar levels, locations, and patchiness of BNRBD (**Figure 2.32**).

Enigma 2011 – Penultimate.

When I plotted the field data from Enigma in 2011 from the penultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation values overlapped with the field data (**Figure 2.5**), indicating that despite a large number of stochastic simulations the field data outcome was represented from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 2.13**). The maximum infection saturation values of the simulation percentile groups also overlapped with the fourth quartile (upper 25%) of the field data and the max field severity values (**Figure 2.13**). Total infection severity for field data exceeded and did not fall within the range of total infection saturation values for all percentile groups (**Figure 2.21**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.27**). The Moran's I value of the corresponding field data

was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.27**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data and qualitatively the disease severity heatmaps based on field data and randomly selected simulation runs produced similar levels, locations, and patchiness of BNRBD (**Figure 2.33**).

Enigma 2011 – Ultimate.

When I plotted the field data from Enigma 2011 from the ultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation values overlapped with the field data (**Figure 2.6**), indicating that despite a large number of stochastic simulations the field data outcome was represented from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 2.14**). The maximum infection saturation values of the simulation percentile groups also overlapped with the max field severity values (**Figure 2.14**). Total infection severity for field data fell within the range of total infection saturation values for the 50-90% percentile group (**Figure 2.22**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.27**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.27**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data and qualitatively the disease

severity heatmaps based on field data and randomly selected simulation runs produced similar levels, locations, and patchiness of BNRBD (**Figure 2.34**).

Enigma 2012 – Penultimate.

When I plotted the field data from Enigma in 2012 from the penultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation values overlapped with the field data (**Figure 2.7**), indicating that despite a large number of stochastic simulations the field data outcome was represented from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 2.15**). The maximum infection saturation values of the simulation percentile groups also overlapped with the max field severity values (**Figure 2.15**). Total infection severity for field data fell within the range of total infection saturation values for the <10% percentile group (**Figure 2.23**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.28**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.28**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data and qualitatively the disease severity heatmaps based on field data and randomly selected simulation runs produced similar levels, locations, and patchiness of BNRBD (**Figure 2.35**).

Enigma 2012 – Ultimate.

When I plotted the field data from Enigma in 2012 from the ultimate data collection date against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection saturation values overlapped with the field data (**Figure 2.8**), indicating that despite a large number of stochastic simulations the field data outcome was represented from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 2.16**). The maximum infection saturation values of the simulation percentile groups also overlapped with the fourth quartile (upper 25%) of the field data and the max field severity values (**Figure 2.16**). Total infection severity for field data fell within the range of total infection saturation values for the 10-50% percentile group (**Figure 2.24**). Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected virtual fields were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.28**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 2.6** and **Figure 2.28**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data and qualitatively the disease severity heatmaps based on field data and randomly selected simulation runs produced similar levels, locations, and patchiness of BNRBD (**Figure 2.36**).

Similarities and differences among simulations and field data. Overall, SimpleMite appeared to perform well in generating reasonable projections of BNRBD over space and time which overlapped with the field data at later dates given the same initial disease distribution. There were some exceptions to this outcome, however. In all

but one of eight instances (the penultimate collection date for Homerville in 2012) all percentile groups of simulation means overlapped with the range of the corresponding field data. In only one instance (the penultimate collection date of Homerville in 2012) did the simulation maximums of at least one percentile group not overlap with either the fourth quartile (upper 25%) or upper outliers of the corresponding field data. Total infection severity values for field data tended to fall within the range of total infection saturation values for corresponding simulation data, except for in two instances (the ultimate and penultimate collection dates for Homerville in 2012 and Enigma in 2011, respectively). SimpleMite also tended to generate higher levels of disease than the field data (e.g. the field data were not positioned in the middle of the projected disease levels) but these projected disease levels also did not lead to severe epidemics that far outstripped the field data. In almost all cases, the max disease levels in the field data were greater than the maximum disease levels projected by SimpleMite.

DISCUSSION

Accessible simulation models for arthropod-vectored viruses are still incredibly rare but should be increasingly valuable to the field of plant pathology as many of the emerging plant diseases around the world are viral/bacterial and likely to be arthropod vectored. There are different valid conceptual approaches to the creation of phytopathological simulations, some more data-intensive than others (González-Domínguez *et al.*, 2023, Savary *et al.*, 2018). SimpleMite was constructed with published mite demographic information, the epidemiological aspects of the virus, a dispersal gradient, and a variable for vector dispersal behavior. Despite the dearth of knowledge on

BNRBV and its relationship with its suspected vector, this published information combined with some simplifying assumptions appear to be sufficient to replicate disease patterns observed in field settings, demonstrating that the development of a simulation model for an arthropod-vectored virus is possible even for pathosystems for which data is sparse.

SimpleMite performed reasonably well in predicting the spatial extent and severity of disease outcomes, especially in years with high disease levels, despite having a high range of percent stochasticity (+/- 20%) in fecundity and survival to replicate environmental variation. Given that biological parameters remained consistent across all four versions of SimpleMite, this suggests that initial concentration of disease in space is an important factor in disease outcomes in years with high total disease levels (Severns & Mundt, 2022, Severns *et al.*, 2015). The SimpleMite model is populated with the estimated number of mites based on the initial disease severity measurements and it must be assumed that plants that are not BNRBV-Positive are not mite-infested. This assumption may or may not align with biological reality; it is unknown if mites are distributed widely across SHB plants in the field regardless of disease status. With this caveat in mind, the capacity of the model to reasonably project disease outcomes in years with high disease levels suggests one of two things: 1) that mite populations are omni-present in SHB fields and that those populations living on BNRBV-infected plants earlier in the season are more important drivers of disease outcomes than the whole field-distribution of mite populations, or 2) that BNRBV infection in the field follows the arrival of *Calacarus* mites to SHB fields and that these mites are not omnipresent on SHB plants.

There are potential reasons for the former and latter conclusions about the mite and disease distributions. Species of eriophyoid mites tend to have specialized relationships with one specific host plant species, which supports the idea that *Calacarus* mites would be omnipresent in SHB fields (Oldfield, 1996). However, some *Calacarus* species have been reported to have unusually wide host ranges spanning multiple plant families, meaning that mites could enter SHB fields after dispersing from a different host species (Li *et al.*, 2014, Oldfield, 1996, Vacante, 2015). Because BNRBV does not appear to persist in SHB plants from one year to the next (Robinson, 2013), the former idea suggests that the virus persists in the overwintering mite population and the latter suggests that a different *Calacarus* host plant serves as a reservoir for the virus. If the virus persists in the mite population from year to year, perhaps it is doing so within the bodies of deutogynes – special survival forms of female eriophyid mites meant to withstand harsh conditions (Manson & Oldfield, 1996). These two perspectives present different disease management approaches: one that would primarily focus on reducing the existing mite population in the field and another that would primarily focus on eradication of the virus reservoir. Additional research is necessary to determine which of these ideas is more representative of biological reality which could be integrated into a refined SimpleMite model to evaluate potential management scenarios before attempting them in the field.

In SimpleMite and in Robinson Field Data, BNRBD does not appear to get out of control. Robinson Field Data reports per-plant disease severity in the form of the proportion of leaves on the SHB plant displaying symptoms of BNRBD. Because more severe infections can lead to leaf abscission, this would mean that plants that are severely

infected may be reported in the data as having a low infection severity value (Robinson, 2013). Also, eriophyid mites are mild parasites that tend not to overcrowd or destroy their hosts and actively disperse when host quality decreases or mite populations climb (Lindquist & Oldfield, 1996, Vacante, 2015). Some elements of one or both of these phenomena appear to be reflected in the results for SimpleMite.

SimpleMite sometimes struggled to replicate observations from Homerville in 2012 – the only field and year combination for which disease levels were low. Robinson hypothesized that disease levels were lower due to increased rainfall in the early growing season in 2012 (2013). This is consistent with evidence that shows that rainfall reduces population sizes of other *Calacarus* mite species and that extreme climatic conditions can significantly alter the demographic parameters and life history of other eriophyoids (Vacante, 2015). Presumably, this represents a link between vector demography and environmental conditions that is not adequately represented by SimpleMite. Vector demography was an especially important element of this model given that the presumed vector is slow-moving and the disease is non-systemic. The number of mites in the simulation related directly to the disease severity of a given plant in the virtual landscape. Relationships between different climatic events and eriophyoid mite population dynamics are not straightforward and may be species-specific. These results highlight yet another area of eriophyoid mite biology that remains poorly understood. Because their feeding tends to produce little cellular damage and their host plant relationships tend to be highly specific, eriophyoid mites are in some ways uniquely positioned to serve as effective viral vectors (Lindquist & Oldfield, 1996). Though eriophyoid-vectored plant pathogens have been documented since the 1930s, little is understood about their transmission and the

few that are studied tend to have highly specific relationships with their vector (Oldfield & Proeseler, 1996, Vacante, 2015). It is difficult to overstate the global distribution of these mites – there are estimated to be at least 250,000 species distributed across 320,000 host plant species and we are only aware of a small percentage of their taxa (Lindquist *et al.*, 1996, Ozman-Sullivan & Sullivan, 2023). These mites are also notorious for colonizing new regions (Lindquist et al., 1996, Navia *et al.*, 2010). If eriophyoid-vectored viruses become more prevalent, this nescience could have significant consequences for the field of plant pathology.

Even the smallest of animals can exemplify complex behaviors that collectively have outsized impacts. Eriophyid mites can barely be seen with the human eye (~ 0.2 mm), and yet they can discern between host and non-host plants, evaluate host plant quality, engage in parental care behaviors, and make dispersal decisions based on their environment (Lindquist & Oldfield, 1996, Michalska et al., 2010, Vacante, 2015). Preliminary versions of SimpleMite (unpublished) failed to replicate disease patterns observed in Robinson Field Data without the inclusion of dispersal behavior patterns derived from field data. If this is true for eriophyoid mites that can only disperse through active ambulatory movement or passive aerial dispersal (Lindquist & Oldfield, 1996), the inclusion of vector dispersal behavior is likely even more critical for models of arthropods that conduct active flight. Perhaps if other elements of eriophyoid mite behavior were better understood, their incorporation into this model would improve its ability to predict spatial disease patterns in years with low-disease levels. Simulation models for arthropod-vectored viruses that do not treat vector behavior as a cornerstone of model development are likely inherently incomplete.

The evaluation of SimpleMite was limited by the lack of field studies on BNRBV. The study presented by Robinson (2013) is the only published research, to the best of our knowledge, that reports per plant disease severity for BNRBD and the creation of SimpleMite would not have been possible without this contribution. This is an illustration of the importance of detailed field study for plant diseases – the usefulness of the data collected may live far beyond the study itself. Future field studies on the spatial patterns of this disease would help to more robustly evaluate the SimpleMite model. We plan to make this model publicly available in the hopes that further studies are conducted by adapting its framework to other pathosystems.

LITERATURE CITED

Agrios, G. N. (2005) *Plant Pathology*. Burlington, MA, USA: Elsevier Academic Press.

Alberti, G. and Nuzzaci, G. (1996) Chapter 1.3 Morphogenesis and Cytogenetics 1.3.1 Oogenesis and Spermatogenesis. In: *World Crop Pests.* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 151-167.

Baker, E. W. and Wharton, G. W. (1952) *An Introduction to Acarology*. New York: Macmillan.

Burkle, C., Olmstead, J. W. and Harmon, P. F. (2012) A Potential Vector of Blueberry Necrotic Ring Blotch Virus and Symptoms on Various Host Genotypes. *PHYTOPATHOLOGY,* **102,** 17-17.

Cantu-Iris, M., Harmon, P. F., Londoño, A. and Polston, J. E. (2013) A Variant of Blueberry Necrotic Ring Blotch Virus Associated with Red Lesions in Blueberry. *Archives of Virology,* **158,** 2197-2200.

de Lillo, E. and Skoracka, A. (2010) What's "Cool" on Eriophyoid Mites? *Experimental and Applied Acarology,* **51,** 3-30.

Fabre, F., Coville, J. and Cunniffe, N. J. (2021) Optimising Reactive Disease Management Using Spatially Explicit Models at the Landscape Scale. In: *Plant Diseases and Food Security in the 21st Century.* (Scott, P., Strange, R., Korsten, L. and Gullino, M. L., eds.). Cham: Springer International Publishing, pp. 47-72.

Fang, Y., Nunez, G. H., Silva, M. N. d., Phillips, D. A. and Munoz, P. R. (2020) A Review for Southern Highbush Blueberry Alternative Production Systems. *Agronomy,* **10,** 1531.

Fournier, V., Rosenheim, J. A., Brodeur, J. and Johnson, M. W. (2004) Population Dynamics and within-Plant Distribution of the Mite Calacarus Flagelliseta (Acari: Eriophyidae) on Papaya in Hawaii. *Journal of Economic Entomology,* **97,** 1563-1569.

Fulcher, A., Gauthier, N. W., Klingeman, W. E., Hale, F. and White, S. A. (2015) Blueberry Culture and Pest, Disease, and Abiotic Disorder Management During Nursery Production in the Southeastern U.S.: A Review. *Journal of Environmental Horticulture,* **33,** 33-47.

Gedamu, W. T., Plank-Wiedenbeck, U. and Wodajo, B. T. (2024) A Spatial Autocorrelation Analysis of Road Traffic Crash by Severity Using Moran's I Spatial Statistics: A Comparative Study of Addis Ababa and Berlin Cities. *Accident Analysis & Prevention,* **200,** 107535.

González-Domínguez, E., Caffi, T., Rossi, V., Salotti, I. and Fedele, G. (2023) Plant Disease Models and Forecasting: Changes in Principles and Applications over the Last 50 Years. *Phytopathology®,* **113,** 678-693.

Hulme, P. E. (2009) Trade, Transport and Trouble: Managing Invasive Species Pathways in an Era of Globalization. *Journal of Applied Ecology,* **46,** 10-18.

Jones, R. A. C. (2016) Chapter Three - Future Scenarios for Plant Virus Pathogens as Climate Change Progresses. In: *Advances in Virus Research.* (Kielian, M., Maramorosch, K. and Mettenleiter, T. C., eds.). Academic Press, pp. 87-147.

Juroszek, P., Racca, P., Link, S., Farhumand, J. and Kleinhenz, B. (2020) Overview on the Review Articles Published During the Past 30 Years Relating to the Potential Climate Change Effects on Plant Pathogens and Crop Disease Risks. *Plant Pathology,* **69,** 179-193.

Kajin, M., Almeida, P. J., Vieira, M. V. and Cerqueira, R. (2012) The State of the Art of Population Projection Models: From the Leslie Matrix to Evolutionary Demography. *Oecologia Australis,* **16,** 13-22.

Kassambara, A. (2023) Ggpubr:'Ggplot2'based Publication Ready Plots. pp. 2.

Keifer, H. H., Baker, E. W., Kono, T., Delfinado, M. and Styler, W. E. (1982) *An Illustrated Guide to Plant Abnormalities Caused by Eriophyid Mites in North America*. U.S. Dept. of Agriculture, Agricultural Research Service.

Li, H.-S., Sun, J.-T., Liu, Y., Xue, X.-F. and Hong, X.-Y. (2014) Population Structures of Acaphylla Theae and Calacarus Carinatus (Acari: Eriophyoidea) from Different Tea-Producing Areas of China Reveal No Host-Associated Differentiation and Recent Colonization. *Systematic and Applied Acarology,* **19,** 113-122, 110.

Lindquist, E. E. and Amrine, J. W. (1996) 1.1.2 Systematics, Diagnoses for Major Taxa, and Keys to Families and Genera with Species on Plants of Economic Importance. In: *Eriophyoid Mites: Their Biology, Natural Enemies and Control (World Crop Pests).* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 33-87.

Lindquist, E. E. and Oldfield, G. N. (1996) Chapter 1.5 Evolution and Phylogeny 1.5.1 Evolution of Eriophyoid Mites in Relation to Their Host Plants. In: *Eriophyoid*

Mites: Their Biology, Natural Enemies and Control (World Crop Pests). (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 277-300.

Lindquist, E. E., Sabelis, M. W. and Bruin, J. (1996) Preface. In: *Eriophyoid Mites: Their Biology, Natural Enemies and Control (World Crop Pests).* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. v-xiii.

Manson, D. C. M. and Oldfield, G. N. (1996) Chapter 1.4 Biology and Ecology 1.4.1 Life Forms, Deuterogyny, Diapause and Seasonal Development. In: *World Crop Pests.* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 173-183.

Martin, R. R., Polashock, J. J. and Tzanetakis, I. E. (2012) New and Emerging Viruses of Blueberry and Cranberry. *Viruses,* **4,** 2831-2852.

Martin, R. R. and Tzanetakis, I. E. (2018) High Risk Blueberry Viruses by Region in North America; Implications for Certification, Nurseries, and Fruit Production. *Viruses,* **10,** 342.

Michalska, K., Skoracka, A., Navia, D. and Amrine, J. W. (2010) Behavioural Studies on Eriophyoid Mites: An Overview. *Experimental and Applied Acarology,* **51,** 31-59.

Müller, K. (2017) Kimisc: Kirill's Miscellaneous Functions. R Package Version 0.4.

Navia, D., Ochoa, R., Welbourn, C. and Ferragut, F. (2010) Adventive Eriophyoid Mites: A Global Review of Their Impact, Pathways, Prevention and Challenges. *Experimental and Applied Acarology,* **51,** 225-255.

Oldfield, G. N. (1996) 1.4.3 Diversity and Host Plant Specificity. In: *Eriophyoid Mites: Their Biology, Natural Enemies and Control (World Crop Pests).* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 199-216.

Oldfield, G. N. and Proeseler, G. (1996) 1.4.9 Eriophyoid Mites as Vectors of Plant Pathogens. In: *Eriophyoid Mites: Their Biology, Natural Enemies and Control (World Crop Pests).* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 259-275.

Ozman-Sullivan, S. and Sullivan, G. (2023) Global Patterns of the Species Richness and Distribution of Eriophyoid Mites: A Response to Li Et Al. 2023. *Journal of Biogeography,* **51**.

Paradis, E. and Schliep, K. (2019) Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R. *Bioinformatics,* **35,** 526-528.

Quito-Avila, D. F., Brannen, P. M., Cline, W. O., Harmon, P. F. and Martin, R. R. (2013) Genetic Characterization of Blueberry Necrotic Ring Blotch Virus, a Novel Rna Virus with Unique Genetic Features. *Journal of General Virology,* **94,** 1426-1434.

Ramos-González, P. L., Arena, G. D., Tassi, A. D., Chabi-Jesus, C., Kitajima, E. W. and Freitas-Astúa, J. (2023) Kitaviruses: A Window to Atypical Plant Viruses Causing Nonsystemic Diseases. *Annual Review of Phytopathology,* **61,** 97-118.

Ritz, C., Strebig, J. C. and Ritz, M. C. (2016) Package 'Drc'. *Creative Commons: Mountain View, CA, USA*.

Robinson, T. S. (2013) Epidemiology of Blueberry Necrotic Ring Blotch Virus of Southern

Highbush Blueberry in Georgia. In: *Department of Plant Pathology*. University of Georgia, pp. 108.

Robinson, T. S., Scherm, H., Brannen, P. M., Allen, R. and Deom, C. M. (2016) *Blueberry Necrotic Ring Blotch Virus* in Southern Highbush Blueberry: Insights into in Planta and in-Field Movement. *Plant Disease,* **100,** 1575-1579.

Sabelis, M. W. and Bruin, J. (1996) 1.5.3. Evolutionary Ecology: Life History Patterns, Food Plant Choice and Dispersal. In: *Eriophyoid Mites: Their Biology, Natural Enemies and Control (World Crop Pests).* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 329-366.

Savary, S., Nelson, A. D., Djurle, A., Esker, P. D., Sparks, A., Amorim, L., *et al.* (2018) Concepts, Approaches, and Avenues for Modelling Crop Health and Crop Losses. *European Journal of Agronomy,* **100,** 4-18.

Schumaker, N. (2023) Hexsimple.

Severns, P. and Mundt, C. (2022) Delays in Epidemic Outbreak Control Cost Disproportionately Large Treatment Footprints to Offset. *Pathogens,* **11,** 393.

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.

T.J. Van der Merwe, G. G. and Coates (1965) Biological Study of the Grey Mite Calacarus Citrifolii Keifer. *South African Journal of Agricultural Science,* **8,** 817-824.

Upadhaya, S. and Dwivedi, P. (2019) The Role and Potential of Blueberry in Increasing Deforestation in Southern Georgia, United States. *Agricultural Systems,* **173,** 39-48.

Vacante, V. (2015) *The Handbook of Mites of Economic Plants: Identification, Bio-Ecology and Control*. Cabi.

Westphal, E. and Manson, D. C. M. (1996) 1.4.6 Feeding Effects on Host Plants: Gall Formation and Other Distortions. In: *Eriophyoid Mites: Their Biology, Natural Enemies and Control (World Crop Pests).* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 231-242.

Wickham, H. (2016) *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, H., François, R., Henry, L., Müller, K. and Vaughan, D. (2023) Dplyr: A Grammar of Data Manipulation.

Williamson, J. G. and Miller, E. P. (2002) Early and Mid-Fall Defoliation Reduces Flower Bud Number and Yield of Southern Highbush Blueberry. *HortTechnology horttech*, **12,** 214-216.

Table 2.1. Data collection dates for each field site from Robinson Field Data.

YEAR	COLLECTION NUMBER	ENIGMA	HOMERVILLE	ROLE
2011	1	June 28th / Day 1	June 21st / Day 1	Initial Disease Conditions
	2	July 12th / Day 15	July 7th / Day 17	-
	3	July 26th / Day 29	July 19th / Day 29	Penultimate Date
	4	August 13th / Day 47	August 12th / Day 53	Ultimate Date
2012	1	June 12th / Day 1	June 6th / Day 1	Initial Disease Conditions
	2	June 26th / Day 15	June 25th / Day 20	-
	3	July 10th / Day 29	July 9th / Day 34	-
	4	July 23rd / Day 42	July 23rd / Day 48	-
	5	September 25th / Day 106	September 27th / Day 114	Penultimate Date
	6	October 16th / Day 127	October 18th / Day 135	Ultimate Date

Table 2.2. Select Global Variable values used in the SimpleMite model.

Variable Name	Variable Value	Citation(s)
Fecundity [stage 0]	0	-
Fecundity [stage 1]	0	-
Fecundity [stage 2]	6	T. J. Van der Merwe & Coates, 1965
Survival [stage 0]	0.95	Druciarek *et al.*, 2014
Survival [stage 1]	0.9	Druciarek *et al.*, 2014
Survival [stage 2]	0.3	Druciarek *et al.*, 2014
Percent Stochasticity [fecundity]	20	-
Percent Stochasticity [survival]	20	-
Carrying Capacity [per-hexagon]	4210	
Viruliferous Vector Threshold	50	-

Table 2.3. Decision table used to transition the disease status of Plants in the Initialize Simulation event group of SimpleMite.

Viruliferous Mites Acquired:	Gap or Plant:	Current Disease Status:	Probability of Becoming BNRBV-Positive:	Probability of Becoming BNRBV-Negative:
No	Plant	BNRBV-Negative	0	1
No	Plant	BNRBV-Positive	1	0
No	Gap	BNRBV-Negative	0	1
No	Gap	BNRBV-Positive	0	1
Yes	Plant	BNRBV-Negative	1	0
Yes	Plant	BNRBV-Positive	1	0
Yes	Gap	BNRBV-Negative	0	1
Yes	Gap	BNRBV-Positive	0	1

Table 2.4. Decision table used to transition the disease status of Plants in the BNRBV Disease Dynamics event group of SimpleMite.

Viruliferous Mites Acquired:	Gap or Plant:	Current Disease Status:	Probability of Becoming BNRBV-Positive:	Probability of Becoming BNRBV-Negative:
No	Plant	BNRBV-Negative	0	1
No	Plant	BNRBV-Positive	1	0
No	Gap	BNRBV-Negative	0	1
No	Gap	BNRBV-Positive	0	1
Yes	Plant	BNRBV-Negative	0.5	0.5
Yes	Plant	BNRBV-Positive	1	0
Yes	Gap	BNRBV-Negative	0	1
Yes	Gap	BNRBV-Positive	0	1

Table 2.5. Summary of Moran's I Analysis of field and SimpleMite simulation data for Homerville, GA in 2011 and 2012.

Year	Type	Collection Date	($\bar{x}$) Moran's I	Min. Moran's I	Max. Moran's I	$\bar{x}$ |(ΔMoran's I)|	Min. |(ΔMoran's I)|	Max. |(ΔMoran's I)|
2011	Field*	Penultimate	*0.050*	-	-	-	-	-
	Simulation	Penultimate	*0.103*	0.064	0.120	0.053	0.013	0.070
	Field*	Ultimate	*0.082*	-	-	-	-	-
	Simulation	Ultimate	*0.093*	0.027	0.120	0.024	$4.33*10^{-4}$	0.055
2012	Field**	Penultimate	$-4.28*10^{-4}$	-	-	-	-	-
	Simulation	Penultimate	-0.002	-0.002	-0.002	0.002	0.001	0.002
	Field**	Ultimate	*0.083*	-	-	-	-	-
	Simulation	Ultimate	-0.002	-0.002	-0.002	0.085	0.085	0.085

*Field data with high disease levels based on total infection severity.
**Field data with low disease levels based on total infection severity.
|(ΔMoran's I)| was calculated as the absolute value of the difference between the field data Moran's I value and the simulation replicate Moran's I value. Italicized values indicate a Moran's I value with a significant p-value or the average of 20 Moran's I values with significant p-values, for field data and simulation data, respectively ($\alpha = 0.05$).

Table 2.6. Summary of Moran's I Analysis of field and SimpleMite simulation data for Enigma, GA in 2011 and 2012.

Year	Type	Collection Date	($\bar{x}$) Moran's I	Min. Moran's I	Max. Moran's I	$\bar{x}$ |(ΔMoran's I)|	Min. |(ΔMoran's I)|	Max. |(ΔMoran's I)|
2011	Field*	Penultimate	*0.120*	-	-	-	-	-
	Simulation	Penultimate	*0.100*	0.092	0.108	0.019	0.012	0.028
	Field*	Ultimate	*0.181*	-	-	-	-	-
	Simulation	Ultimate	*0.113*	0.106	0.124	0.068	0.057	0.074
2012	Field*	Penultimate	*0.173*	-	-	-	-	-
	Simulation	Penultimate	*0.082*	0.063	0.115	0.091	0.058	0.110
	Field*	Ultimate	*0.173*	-	-	-	-	-
	Simulation	Ultimate	*0.077*	0.074	0.081	0.068	0.081	0.088

*Field data with high disease levels based on total infection severity.
|(ΔMoran's I)| was calculated as the absolute value of the difference between the field data Moran's I value and the simulation replicate Moran's I value. Italicized values indicate a Moran's I value with a significant p-value or the average of 20 Moran's I values with significant p-values, for field data and simulation data, respectively ($\alpha = 0.05$).

Equation 2.1. Equation for the calculation of "Infection Saturation" accumulator value in BNRBV Disease Dynamics event group of SimpleMite.

$$Infection\ Saturation = Min\ (100, (\frac{Number\ Stage\ 2}{Carrying\ Capacity\ [\ per - hexagon\ ]} * 100))$$

This equation follows the formula: $Min(x, y)$ which selects the smaller number between x and y.

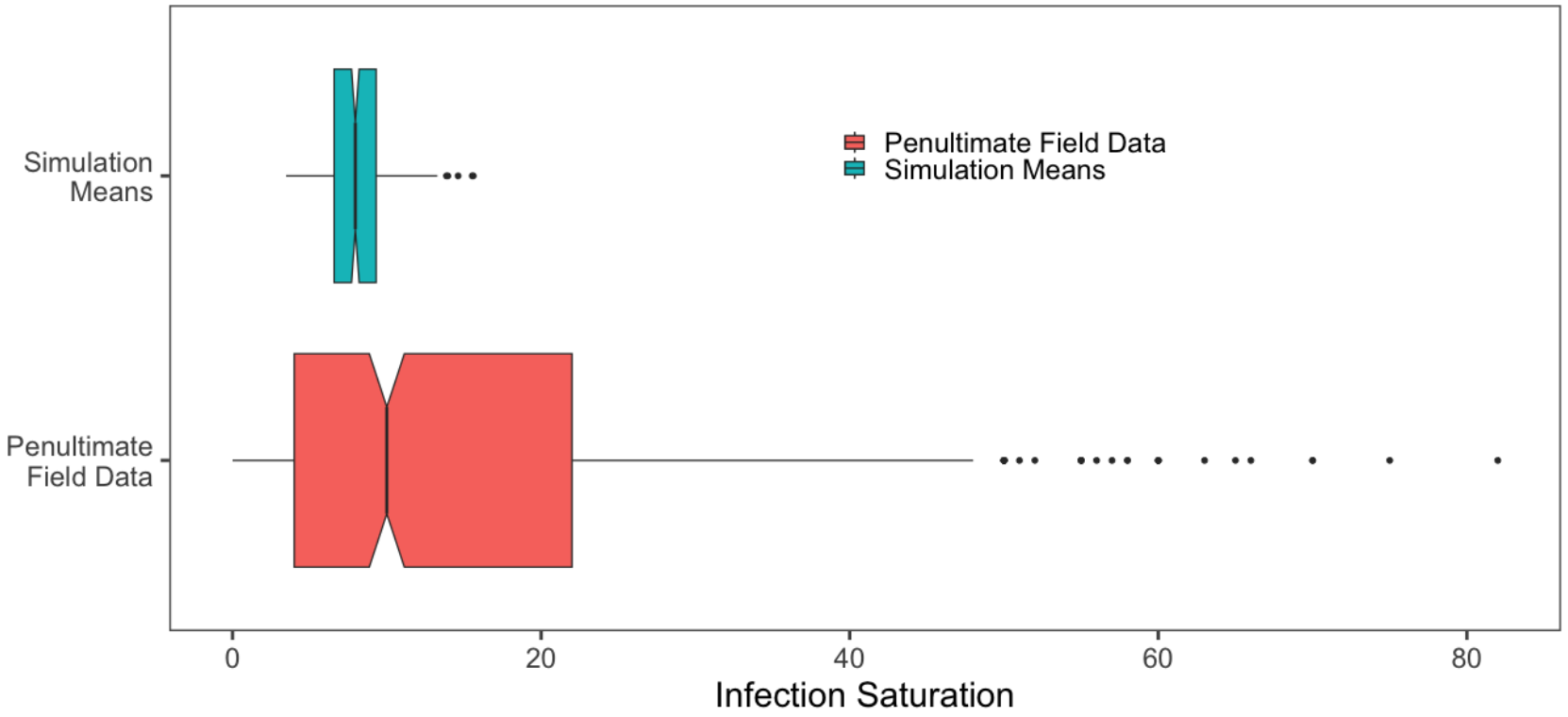

Figure 2.1. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA penultimate (day 29) field data from 2011 and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

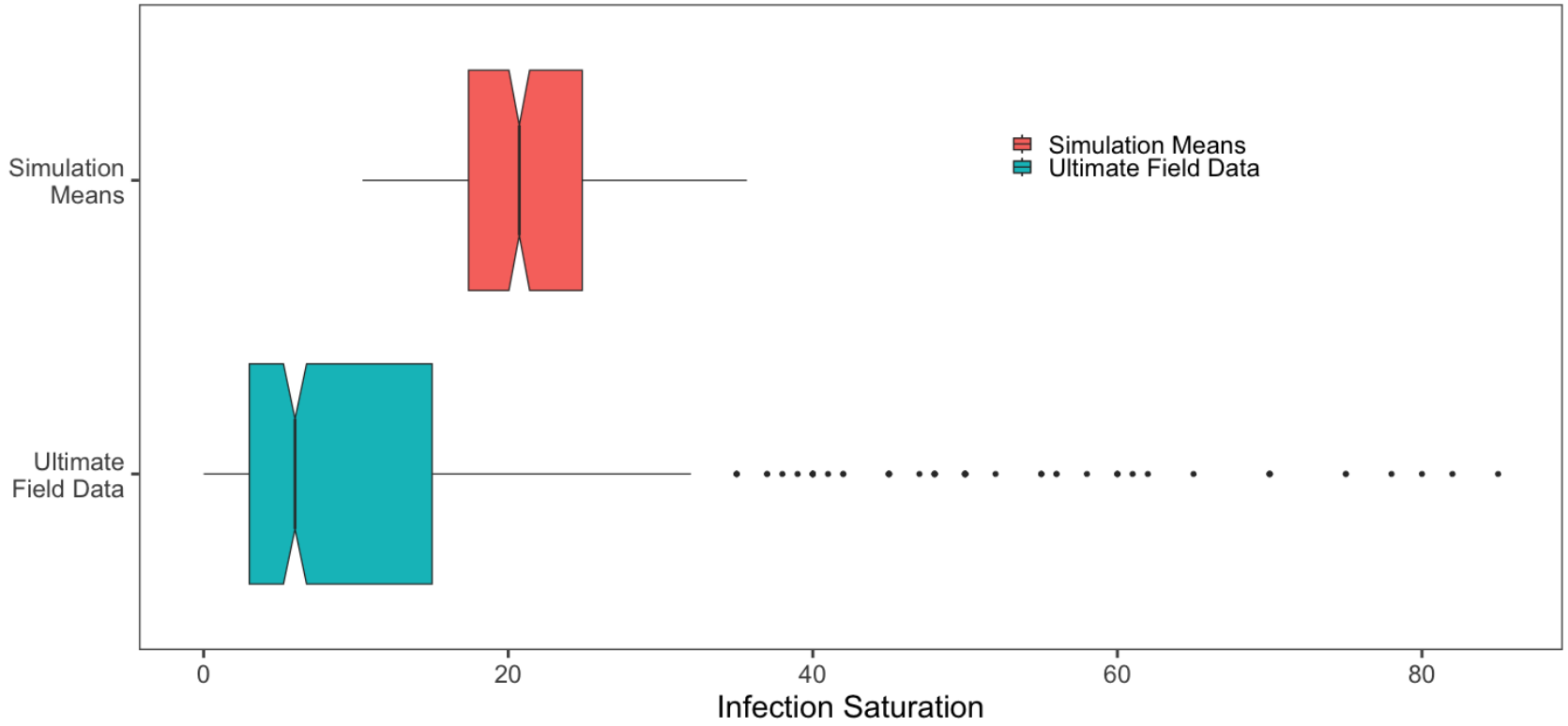

Figure 2.2. Comparison of field data to simulated results of the same time point. Notched boxplot of Homerville, GA 2011 ultimate (day 53) field data and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

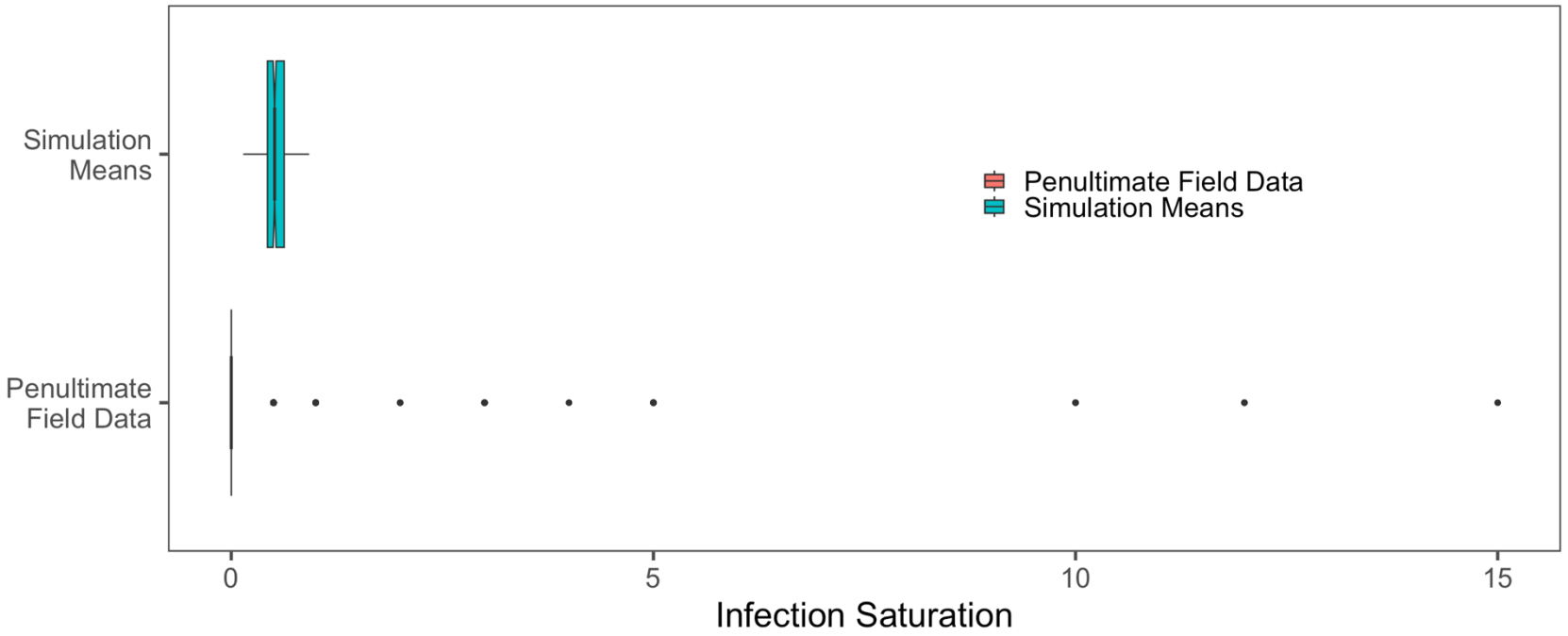

Figure 2.3. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2012 penultimate (day 114) field data and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

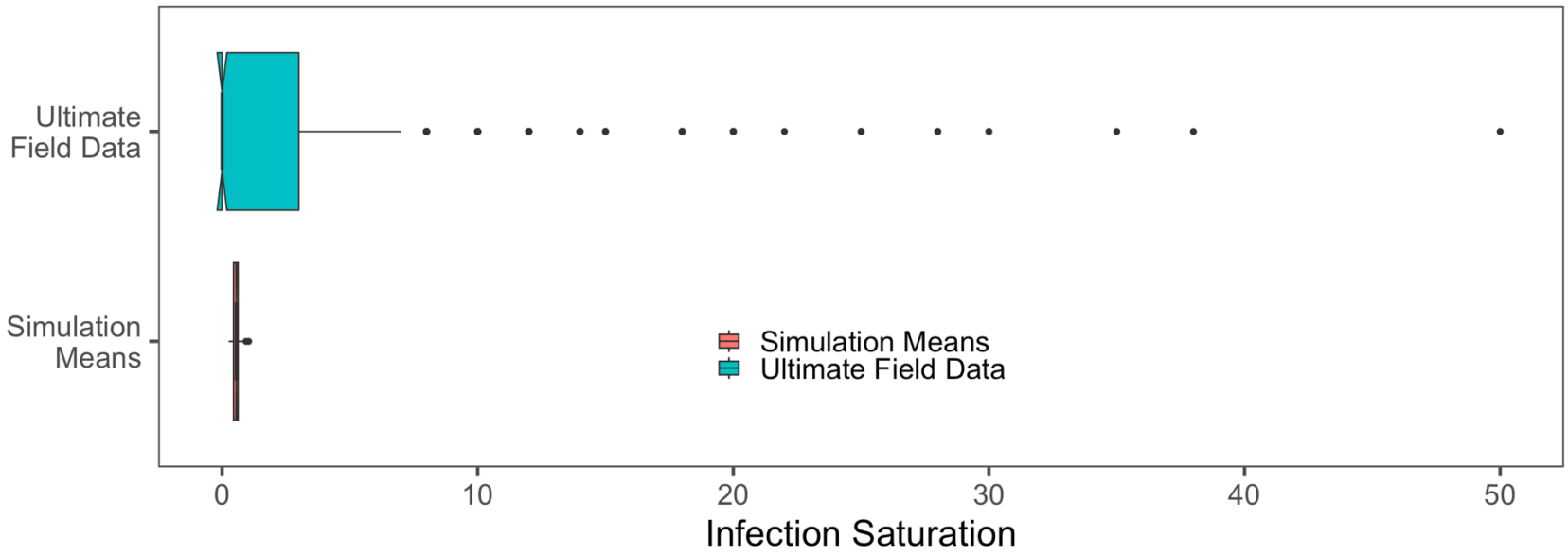

Figure 2.4. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2012 ultimate (day 135) field data and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

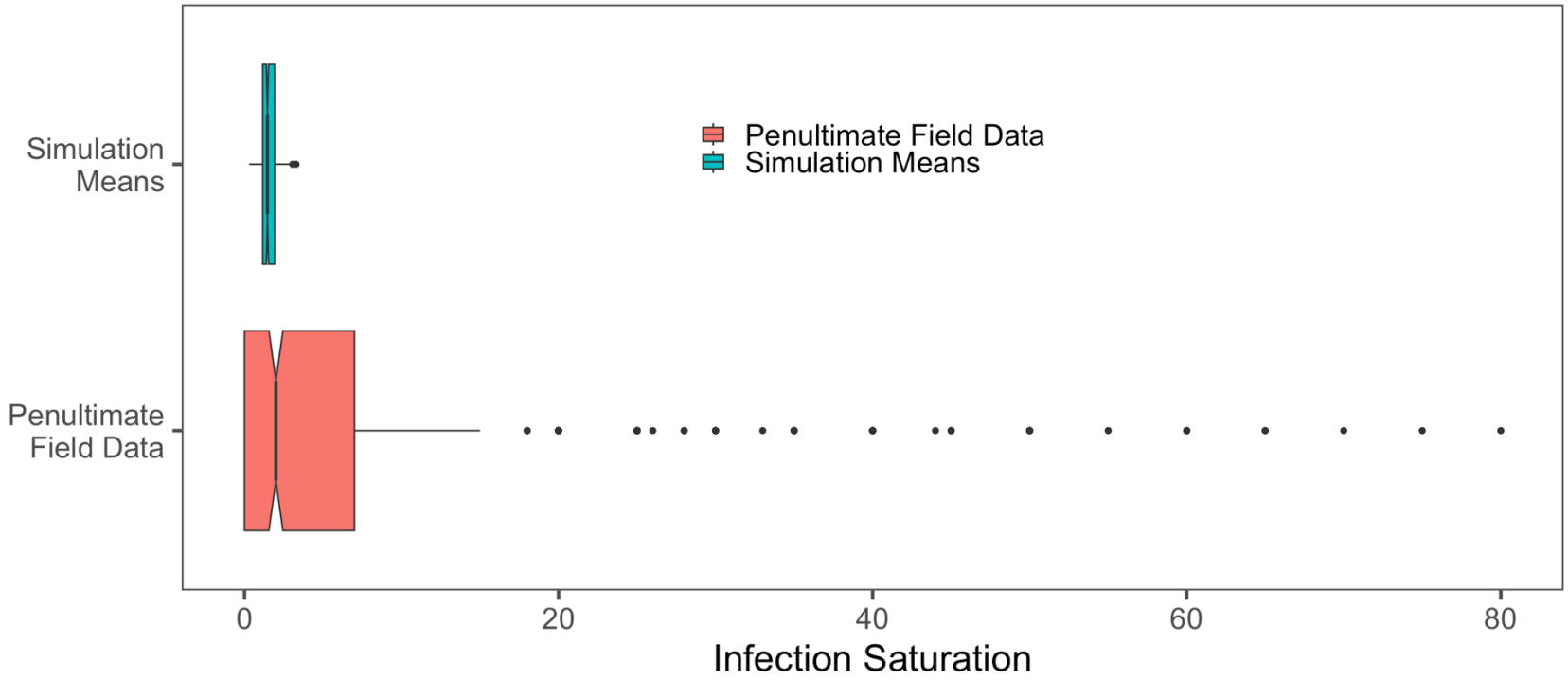

Figure 2.5. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2011 penultimate (day 29) field data and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

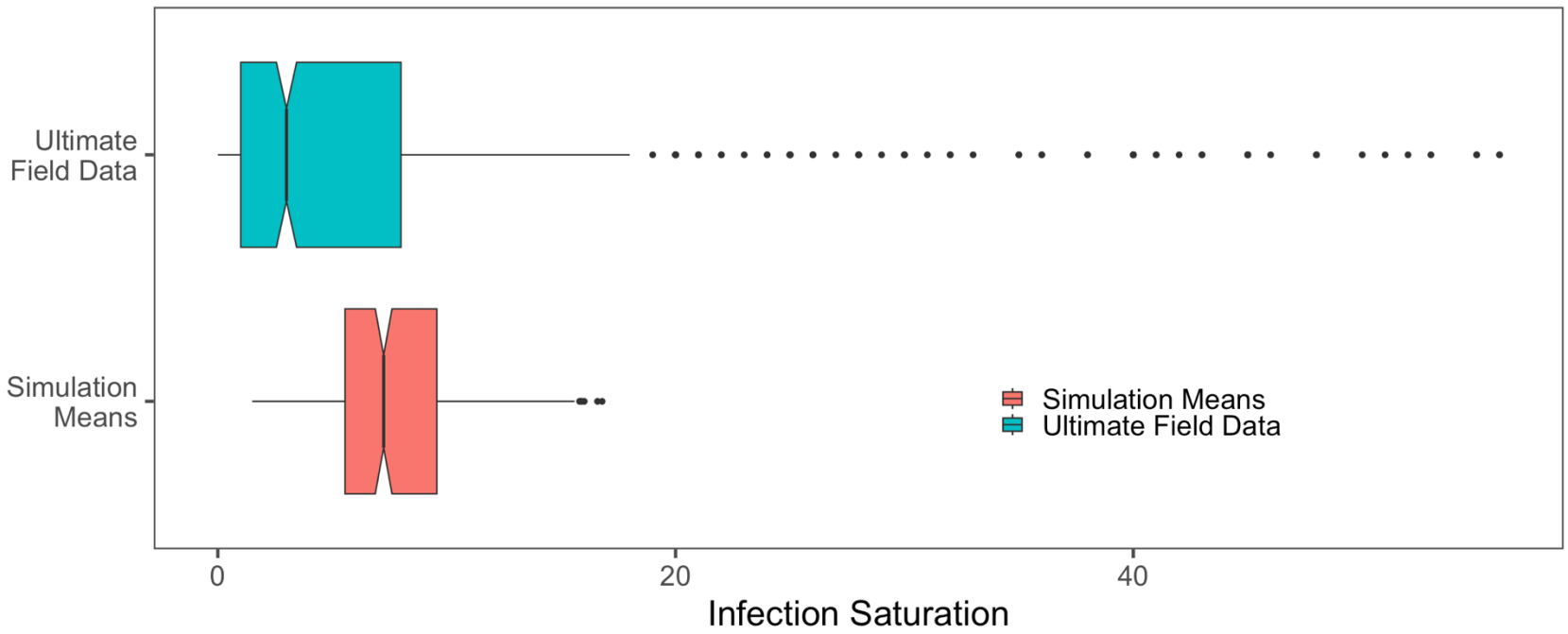

Figure 2.6. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2011 ultimate (day 47) field data and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

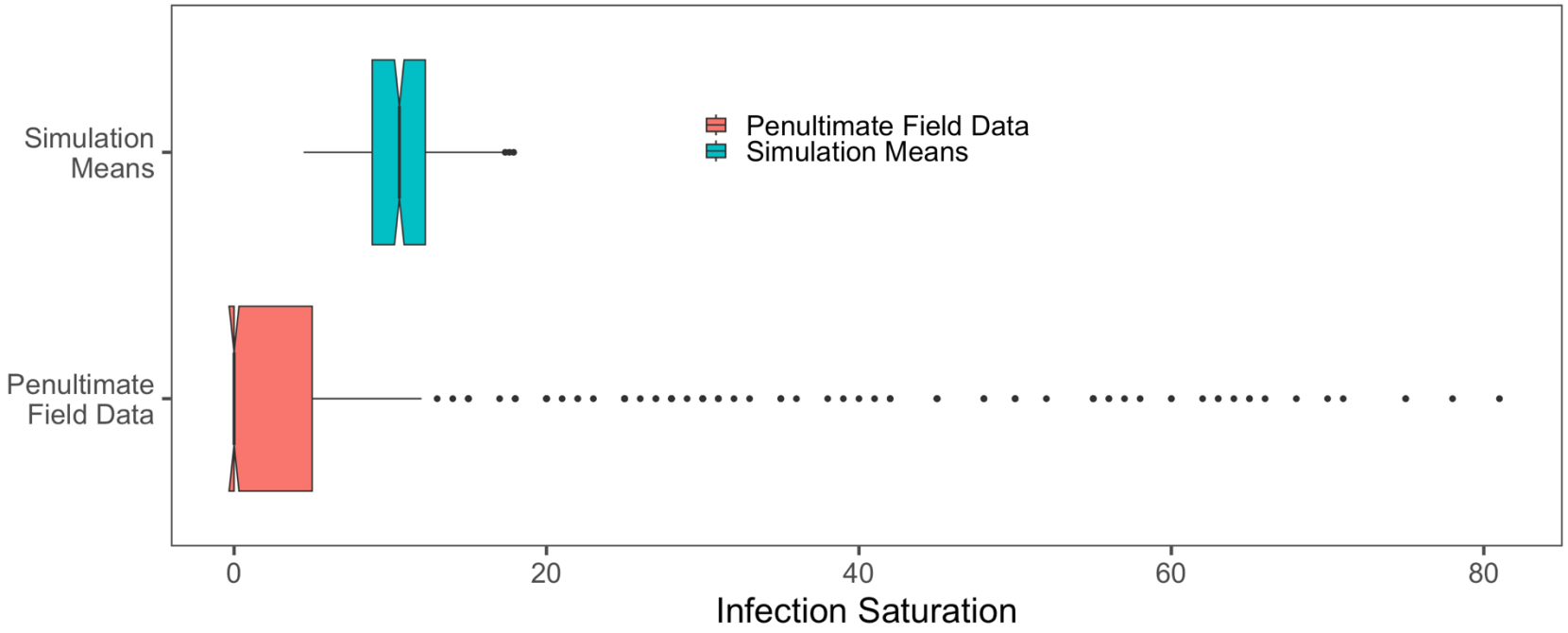

Figure 2.7. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2012 penultimate (day 106) field data and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

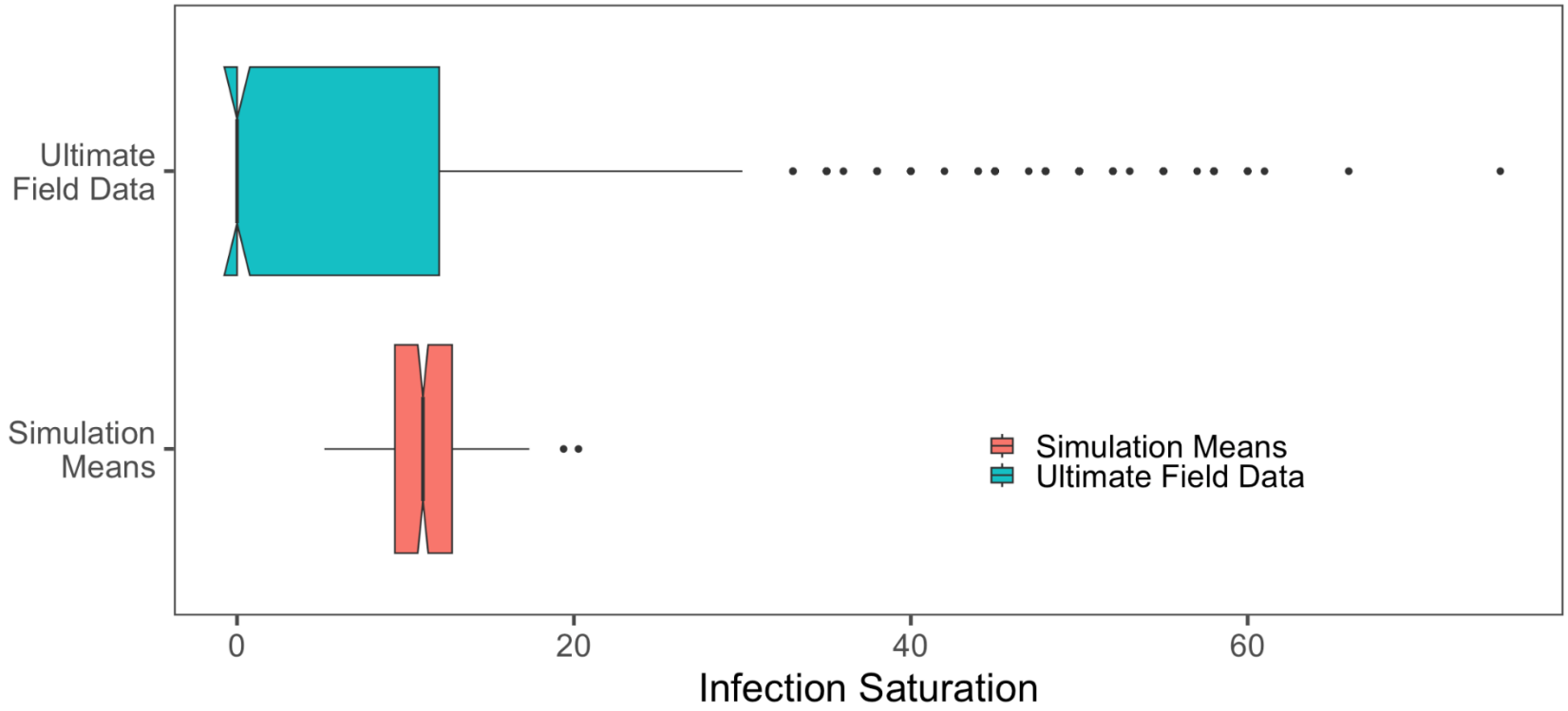

Figure 2.8. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2012 ultimate (day 127) field data and corresponding simulation means. Field data represents individual infection severity values of each SHB plant. Simulation means represent the mean infection saturation of all plants in the virtual field for 300 replicates.

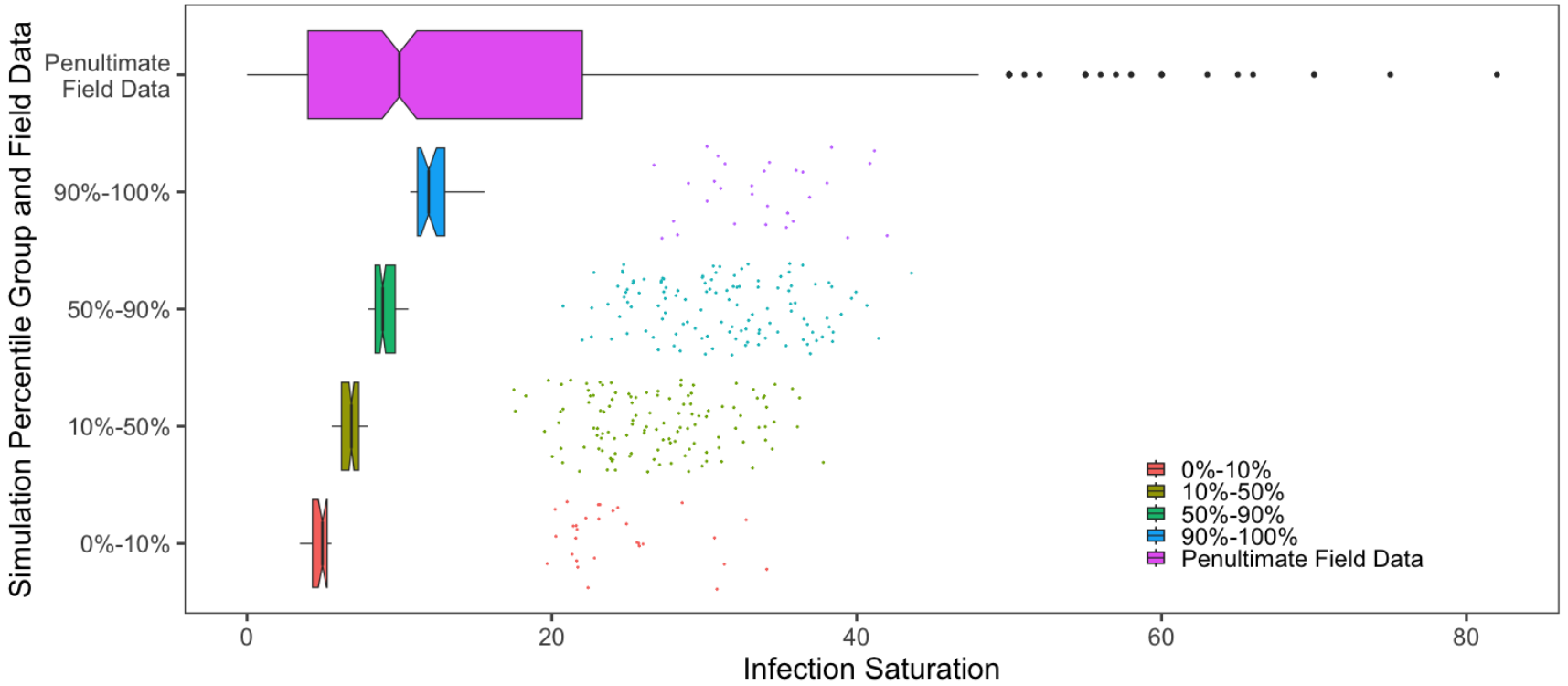

Figure 2.9. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2011 penultimate (day 29) field data and corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

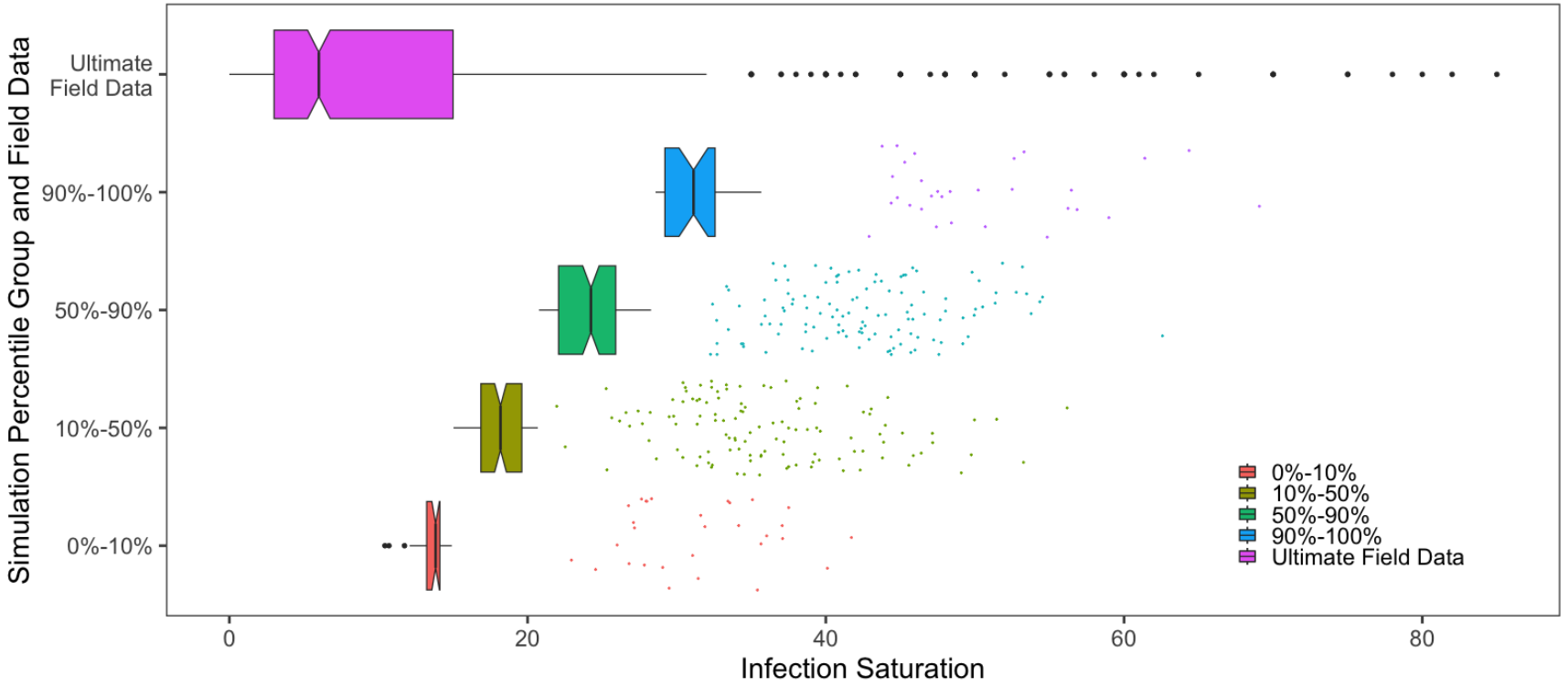

Figure 2.10. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2011 ultimate (day 53) field data and corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

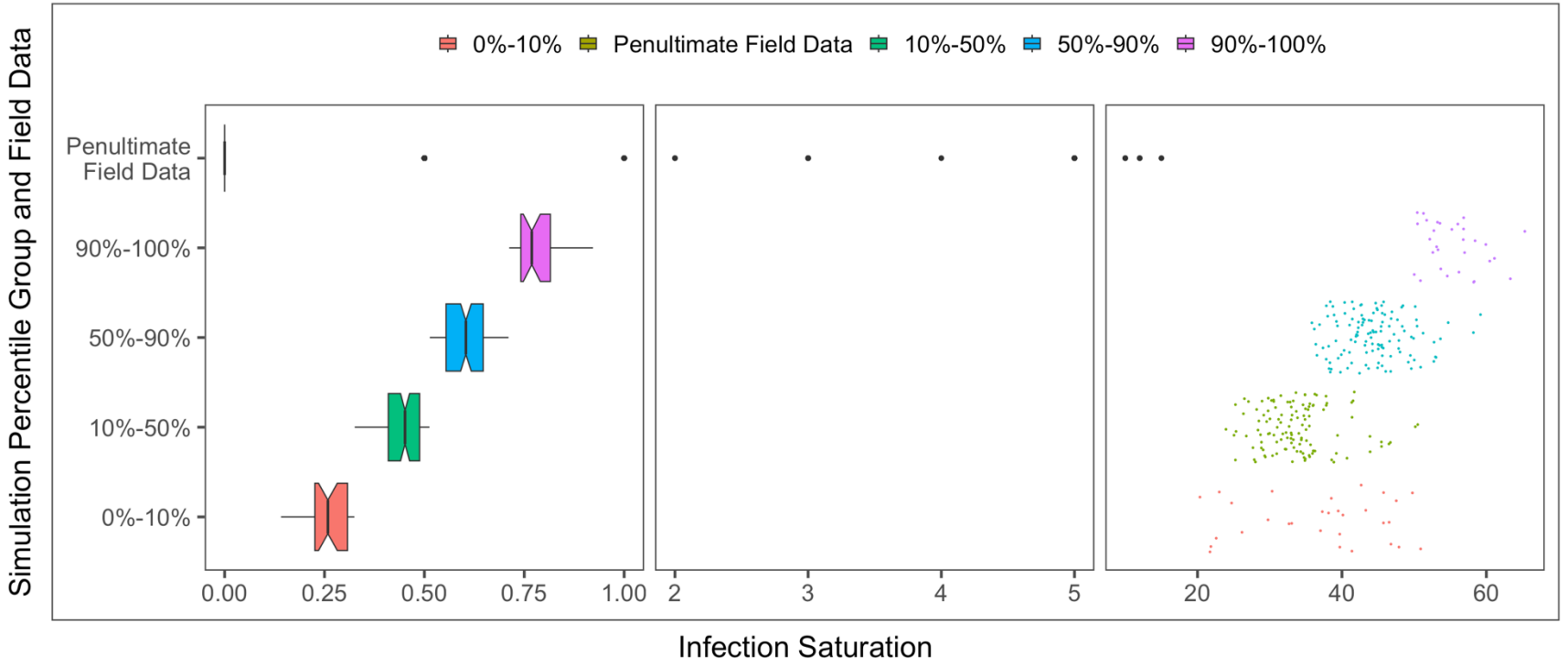

Figure 2.11. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2012 penultimate (day 114) field data and corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

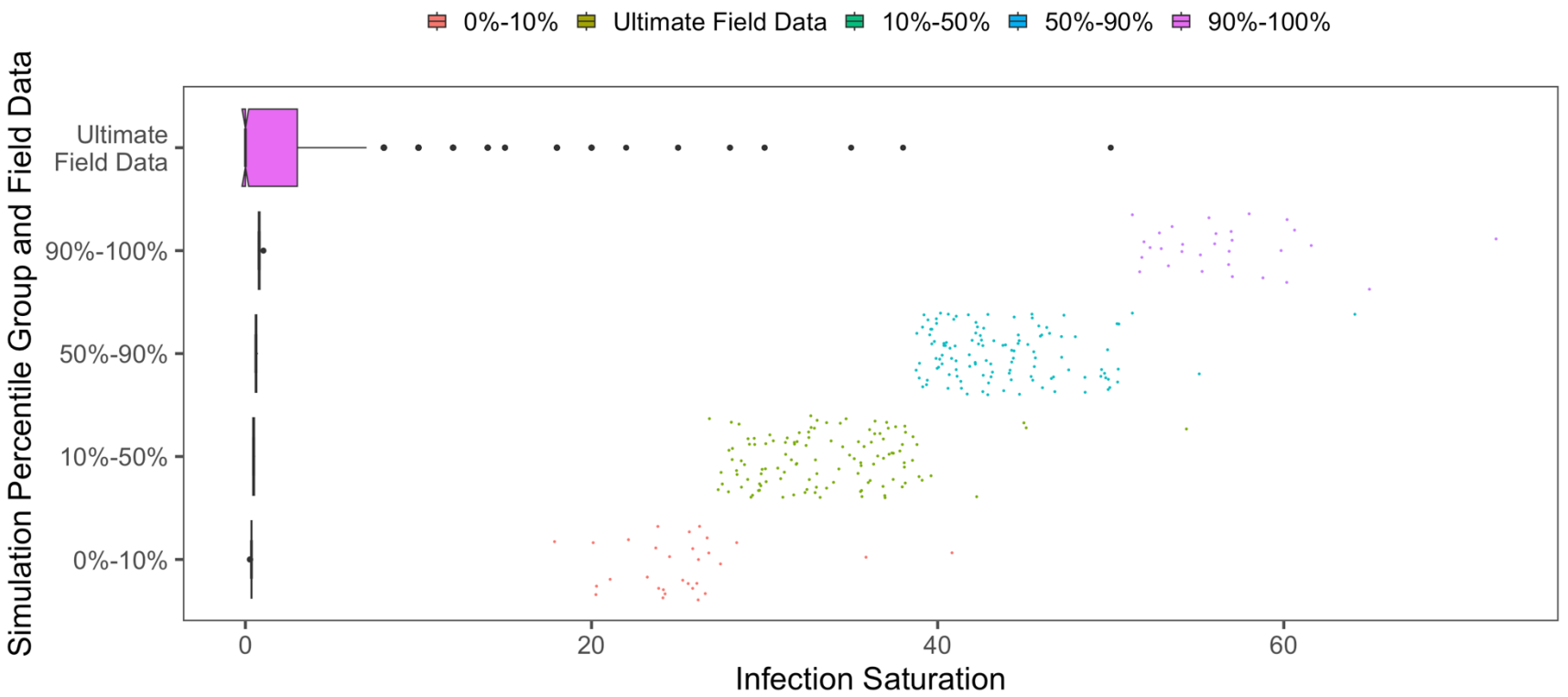

Figure 2.12. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2012 ultimate (day 135) field data and corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

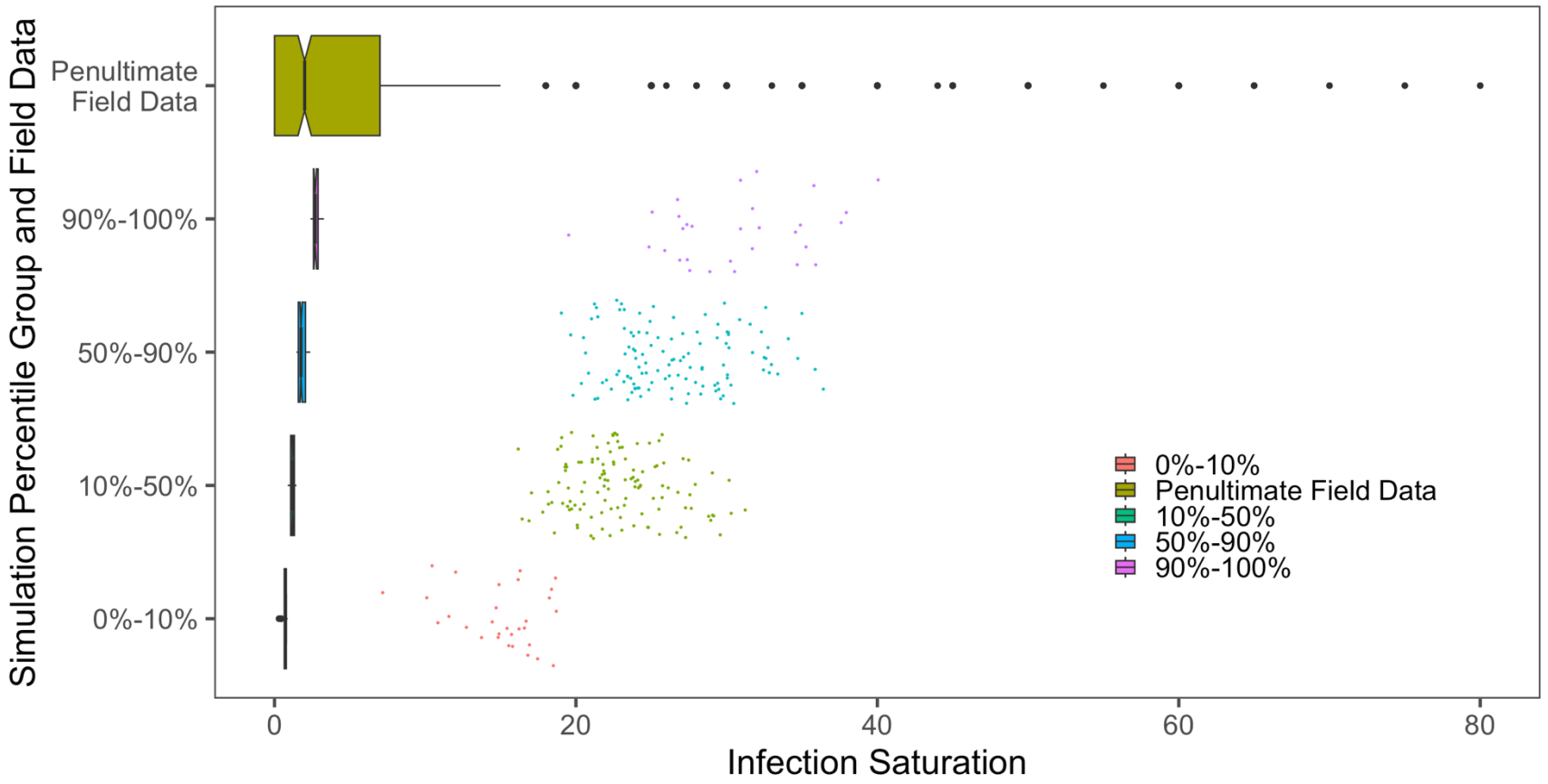

Figure 2.13. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2011 penultimate (day 29) field data to corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

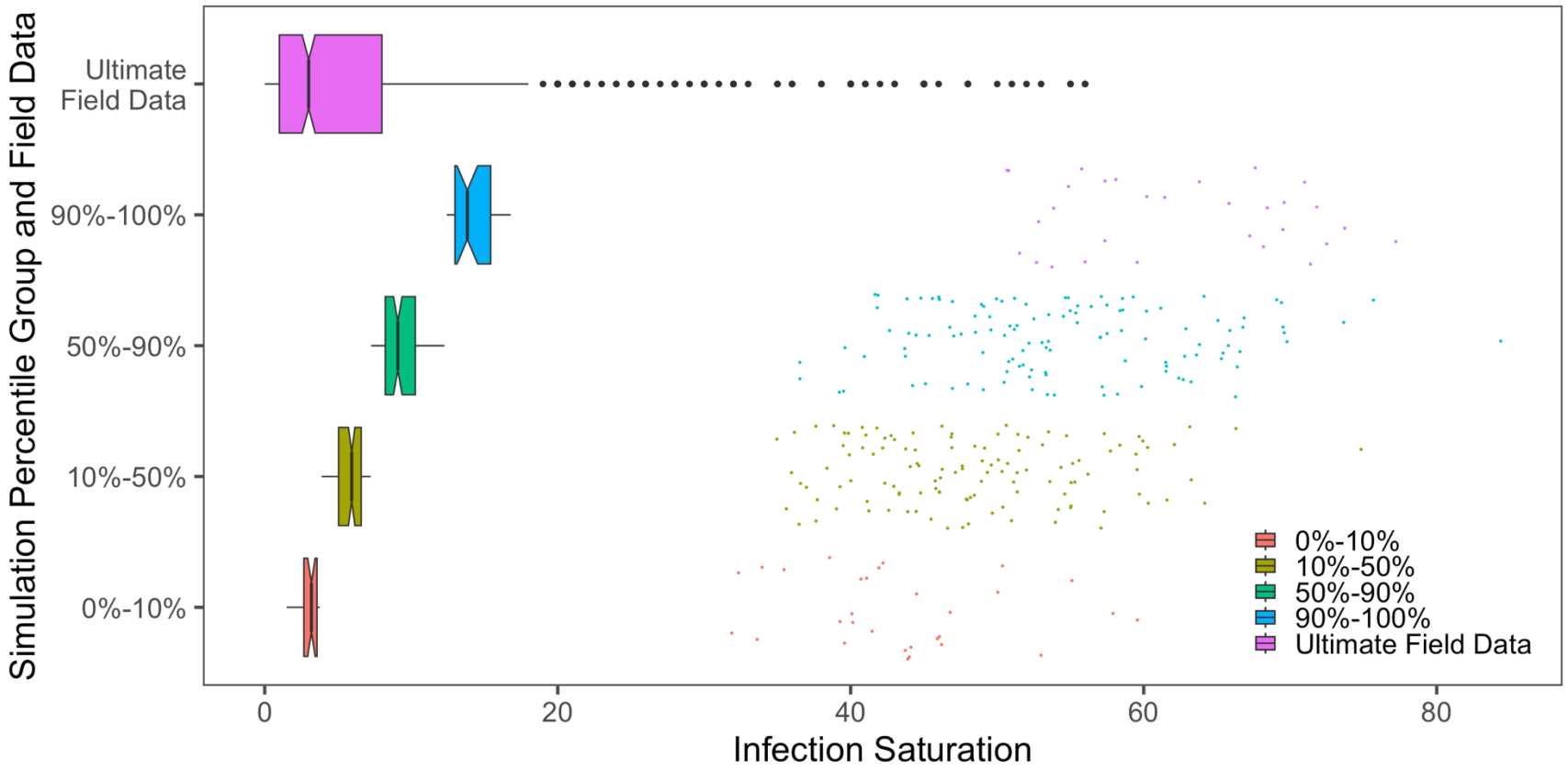

Figure 2.14. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2011 ultimate (day 47) field data to corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

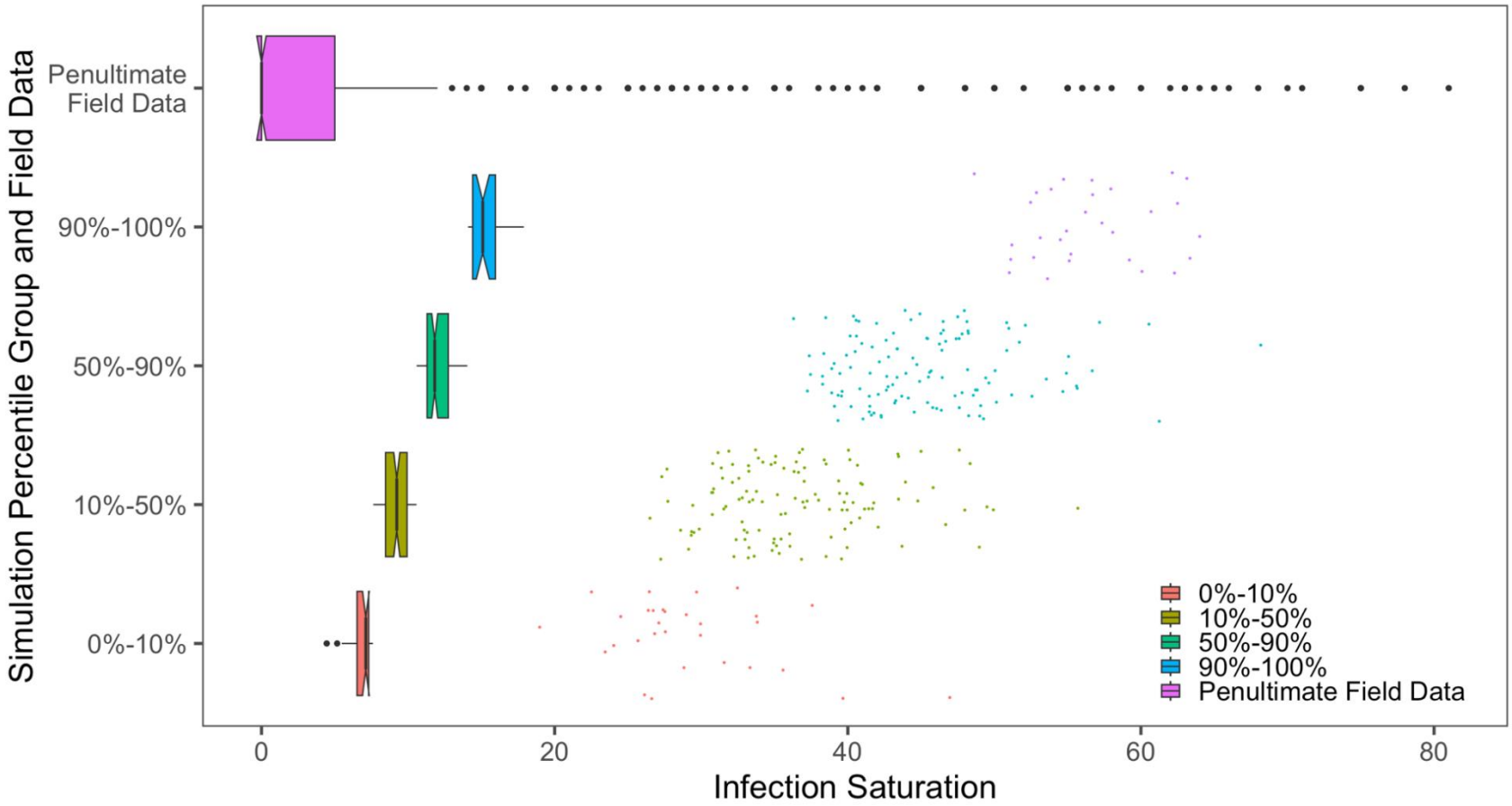

Figure 2.15. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2012 penultimate (day 106) field data and corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

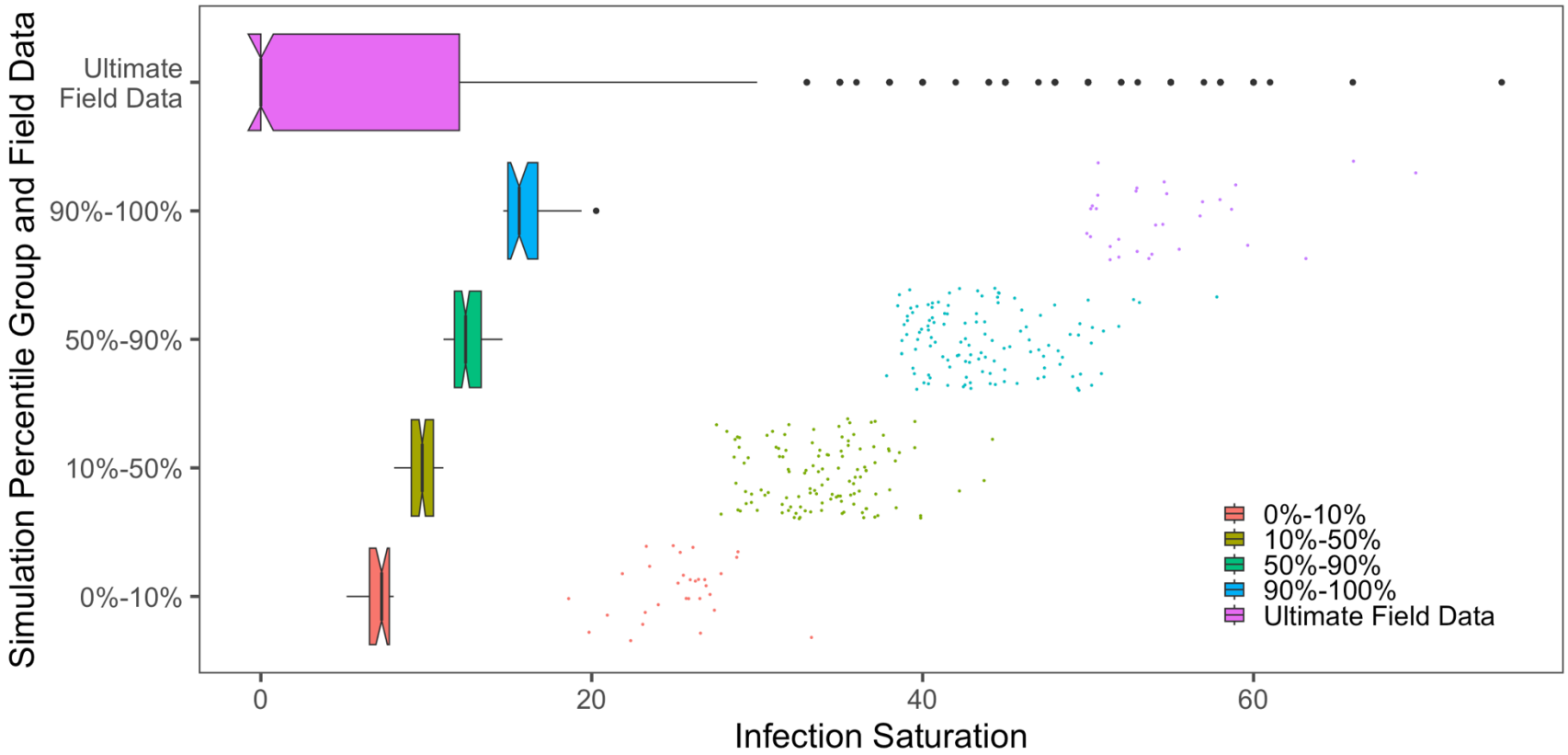

Figure 2.16. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2012 ultimate (day 127) field data and corresponding simulation means and maximums grouped by percentile. Field data represents individual infection severity values of each SHB plant. Simulation percentile group boxplots represent the mean infection saturations for each replicate in the percentile group, while points represent maximum infection saturation values for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

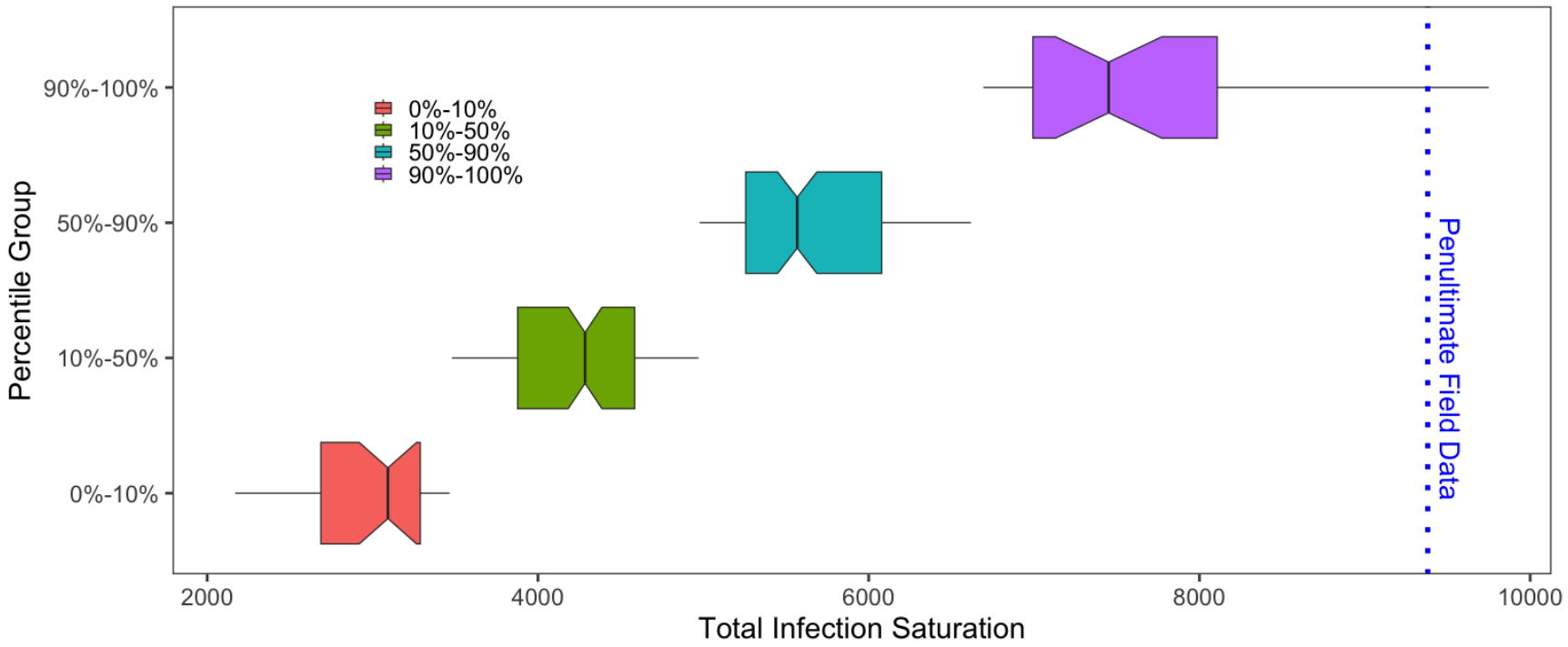

Figure 2.17. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2011 penultimate (day 29) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

Figure 2.18. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2011 ultimate (day 53) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

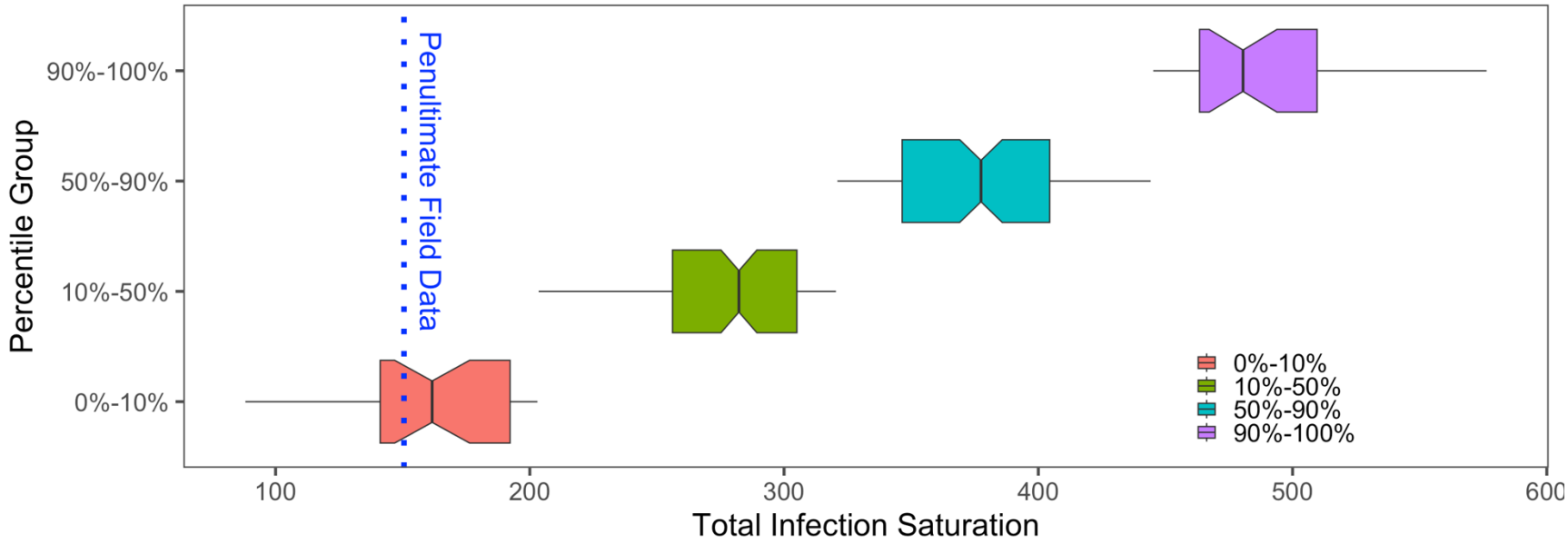

Figure 2.19. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2012 penultimate (day 114) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

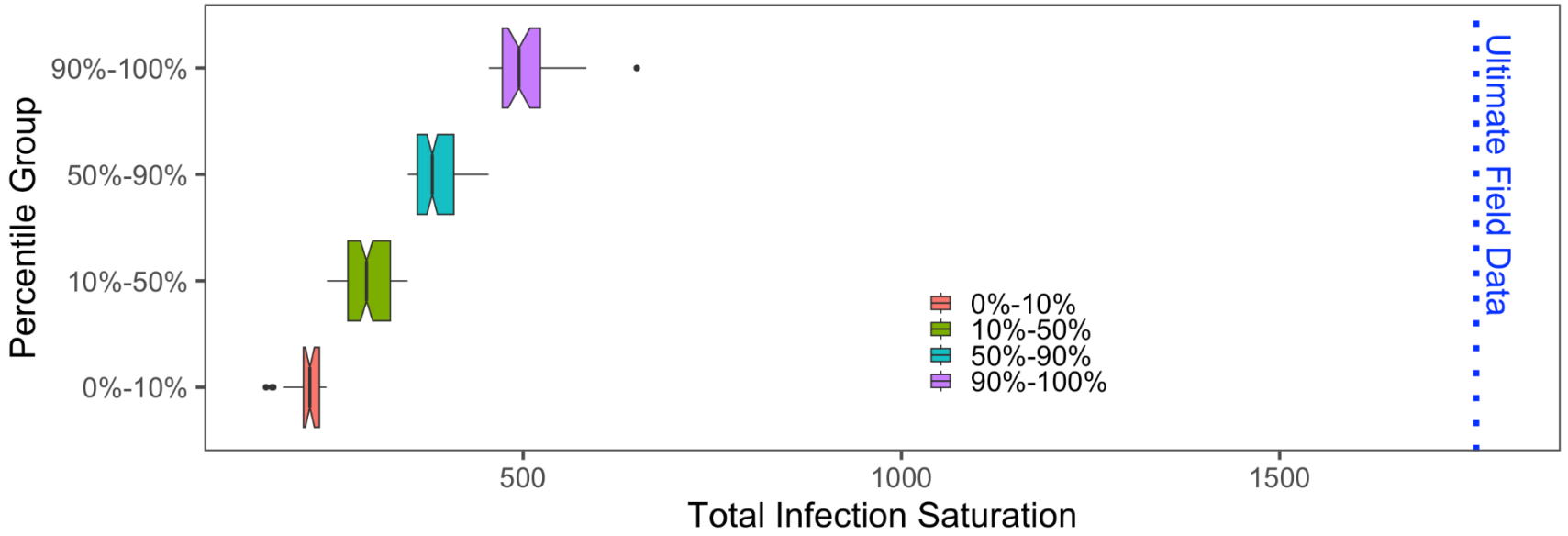

Figure 2.20. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Homerville, GA 2012 ultimate (day 135) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

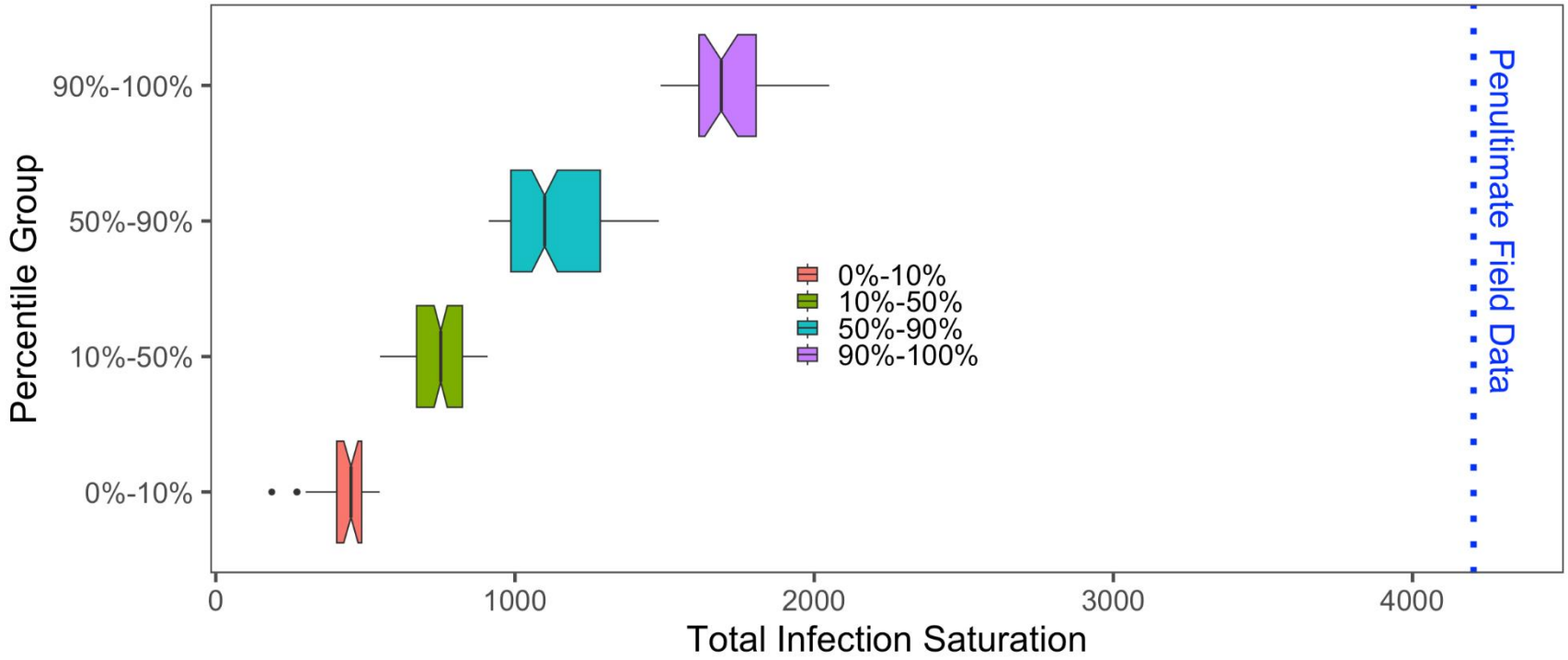

Figure 2.21. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2011 penultimate (day 29) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

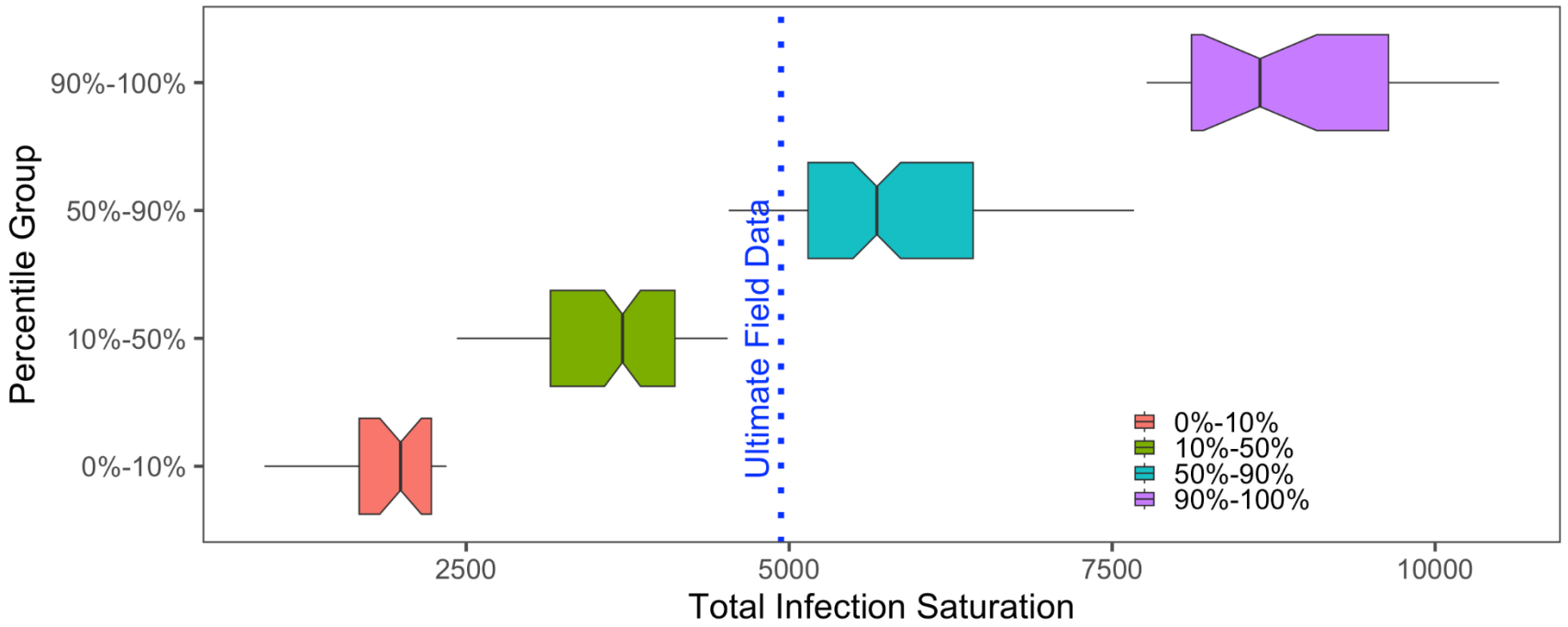

Figure 2.22. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2011 ultimate (day 53) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

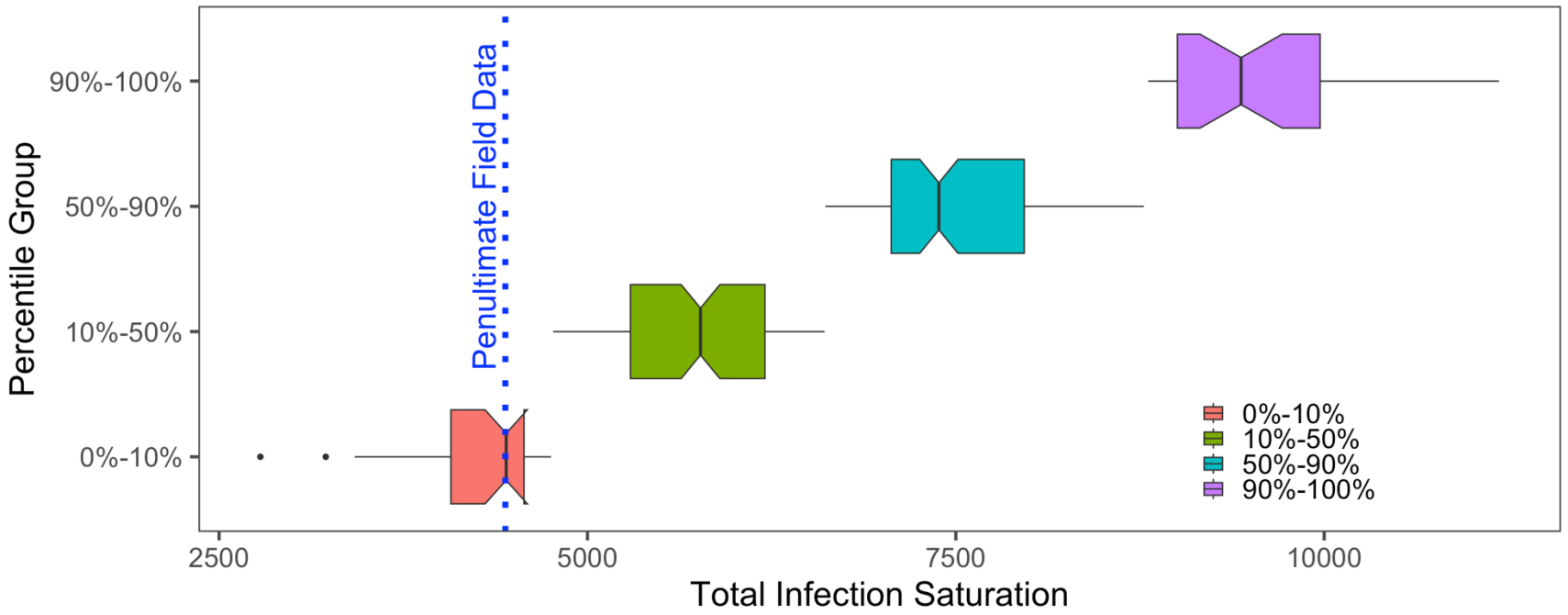

Figure 2.23. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2012 penultimate (day 106) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

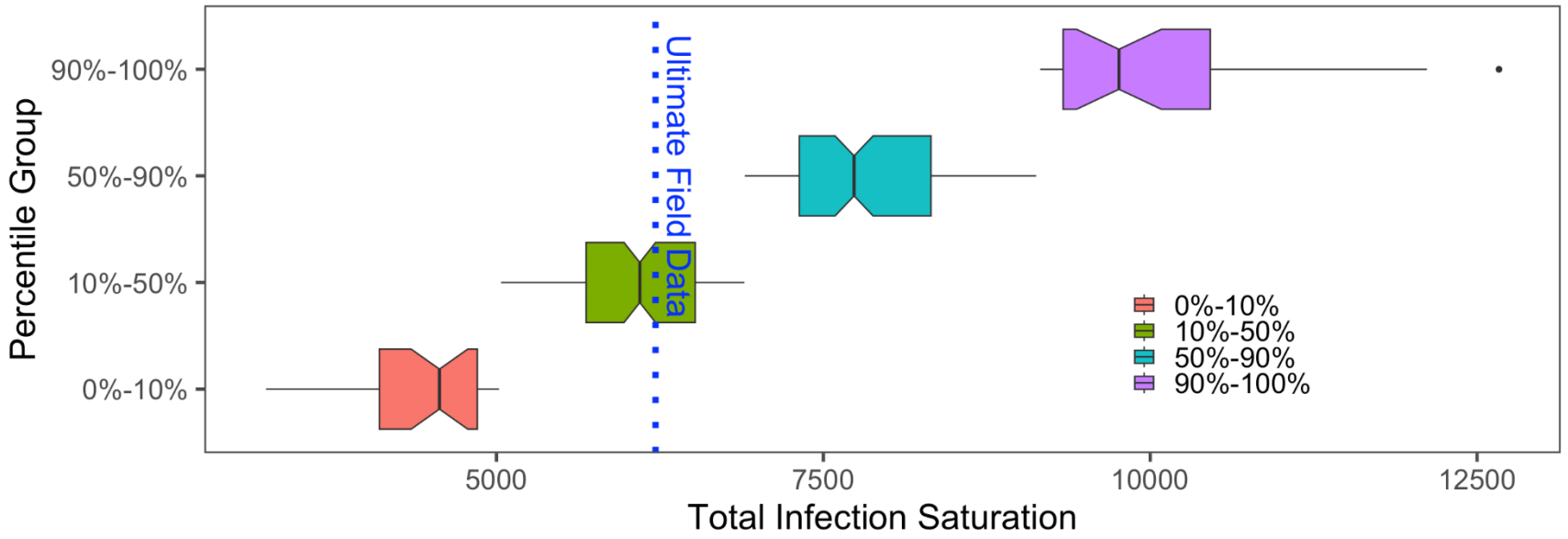

Figure 2.24. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of Enigma, GA 2012 ultimate (day 127) field data and corresponding simulation total infection saturation grouped by percentile. Field data, represented by dotted blue line, represents total infection severity of all SHB plants. Simulation percentile group boxplots represent the total infection saturations for each replicate in the percentile group. Percentile groups were determined based on mean field infection saturation.

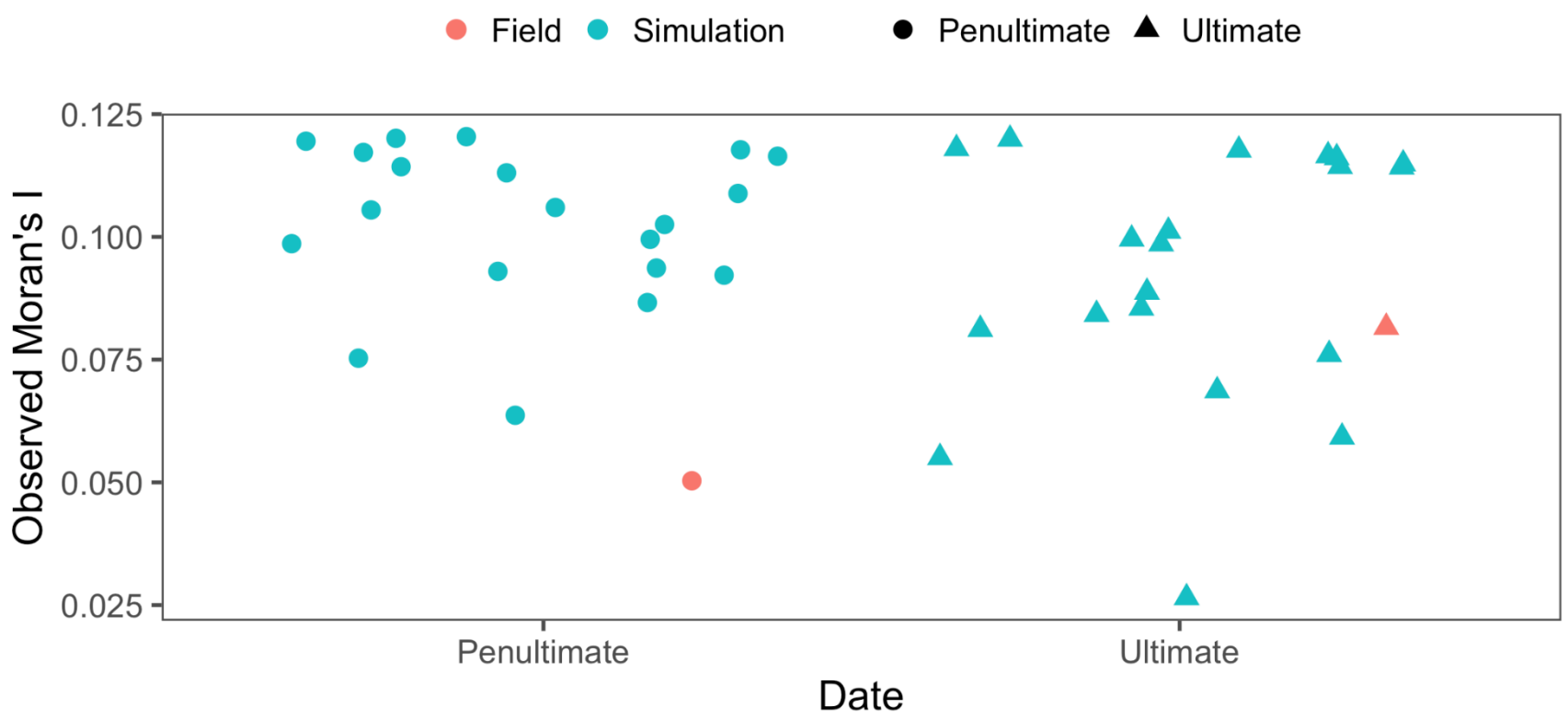

Figure 2.25. Degree of spatial aggregation for disease incidence field data and that of simulation results at the same time point. Comparison Moran's I value for Homerville, GA 2011 field data to Moran's I values of 20 randomly selected virtual fields from corresponding simulations. For penultimate and ultimate collection dates, 20 replicates each were selected from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean.

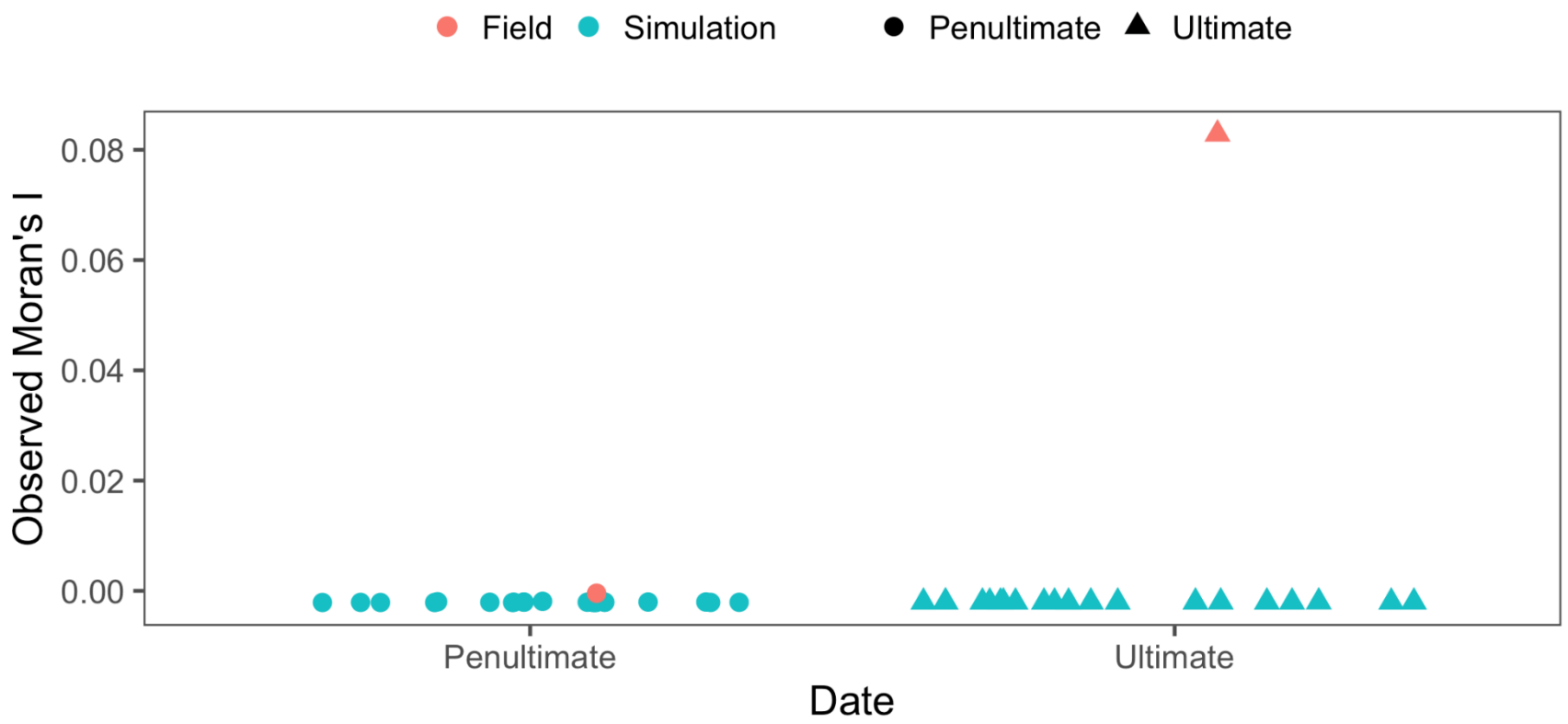

Figure 2.26. Degree of spatial aggregation for disease incidence field data and that of simulation results at the same time point. Comparison of Moran's I value for Homerville, GA 2012 field data to Moran's I values of 20 randomly selected virtual fields from corresponding simulations. For penultimate and ultimate collection dates, 20 replicates each were selected from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean.

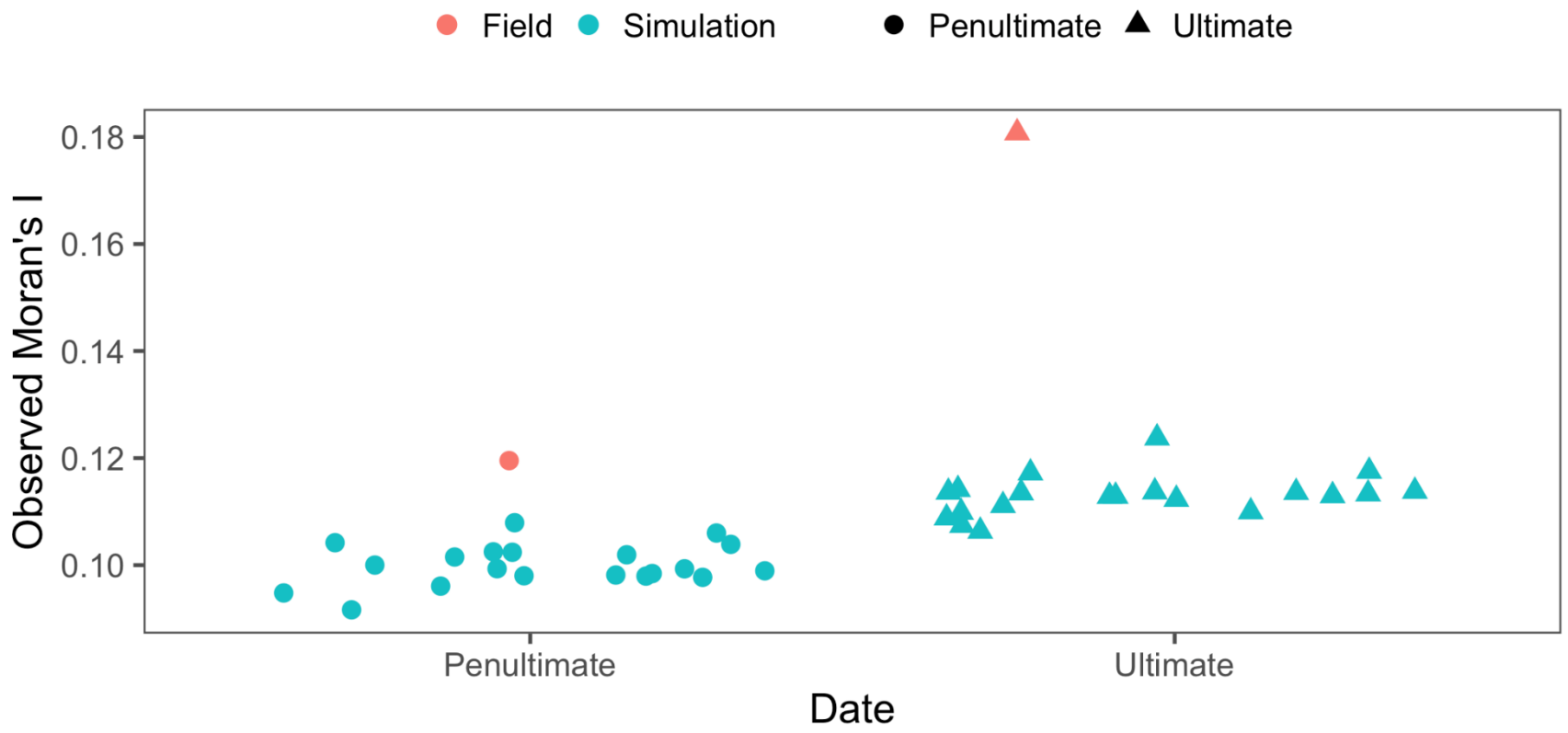

Figure 2.27. Degree of spatial aggregation for disease incidence field data and that of simulation results at the same time point. Comparison of Moran's I value for Enigma, GA 2011 field data to Moran's I values of 20 randomly selected virtual fields from corresponding simulations. For penultimate and ultimate collection dates, 20 replicates each were selected from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean.

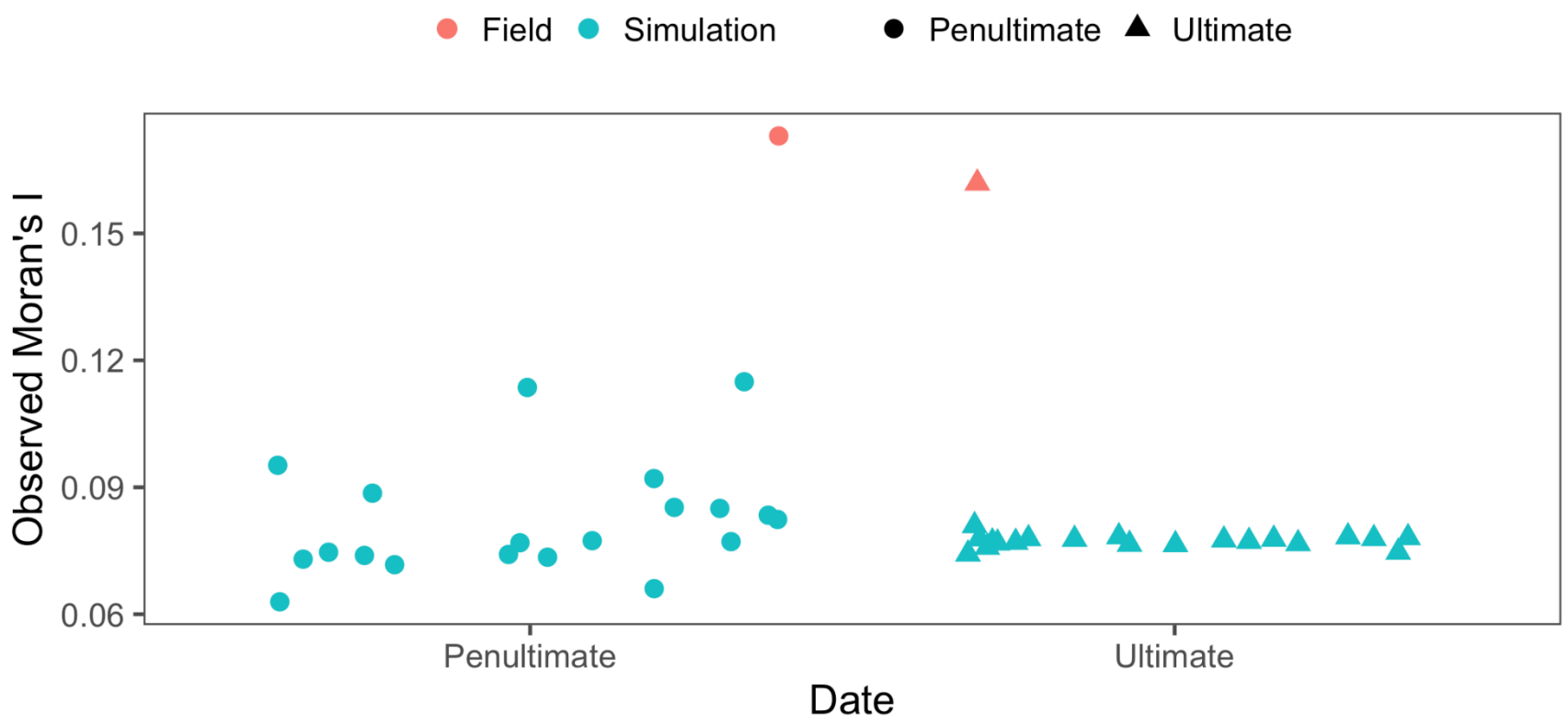

Figure 2.28. Degree of spatial aggregation for disease incidence field data and that of simulation results at the same time point. Comparison of Moran's I value for Enigma, GA 2012 field data to Moran's I values of 20 randomly selected virtual fields from corresponding simulations. For penultimate and ultimate collection dates, 20 replicates each were selected from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean.

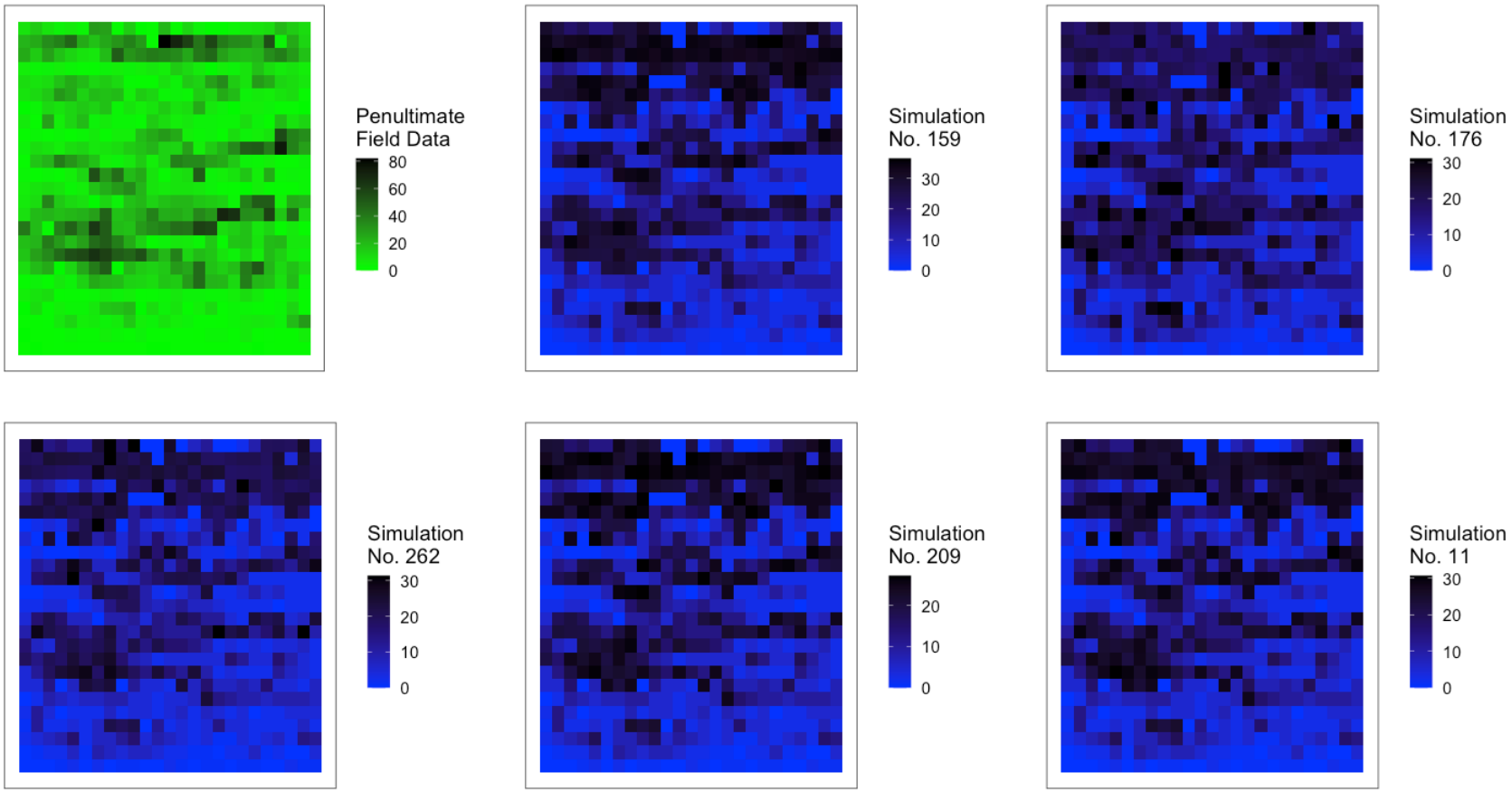

Figure 2.29. Heatmaps of field infection severity for Homerville, GA 2011 penultimate (day 29) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

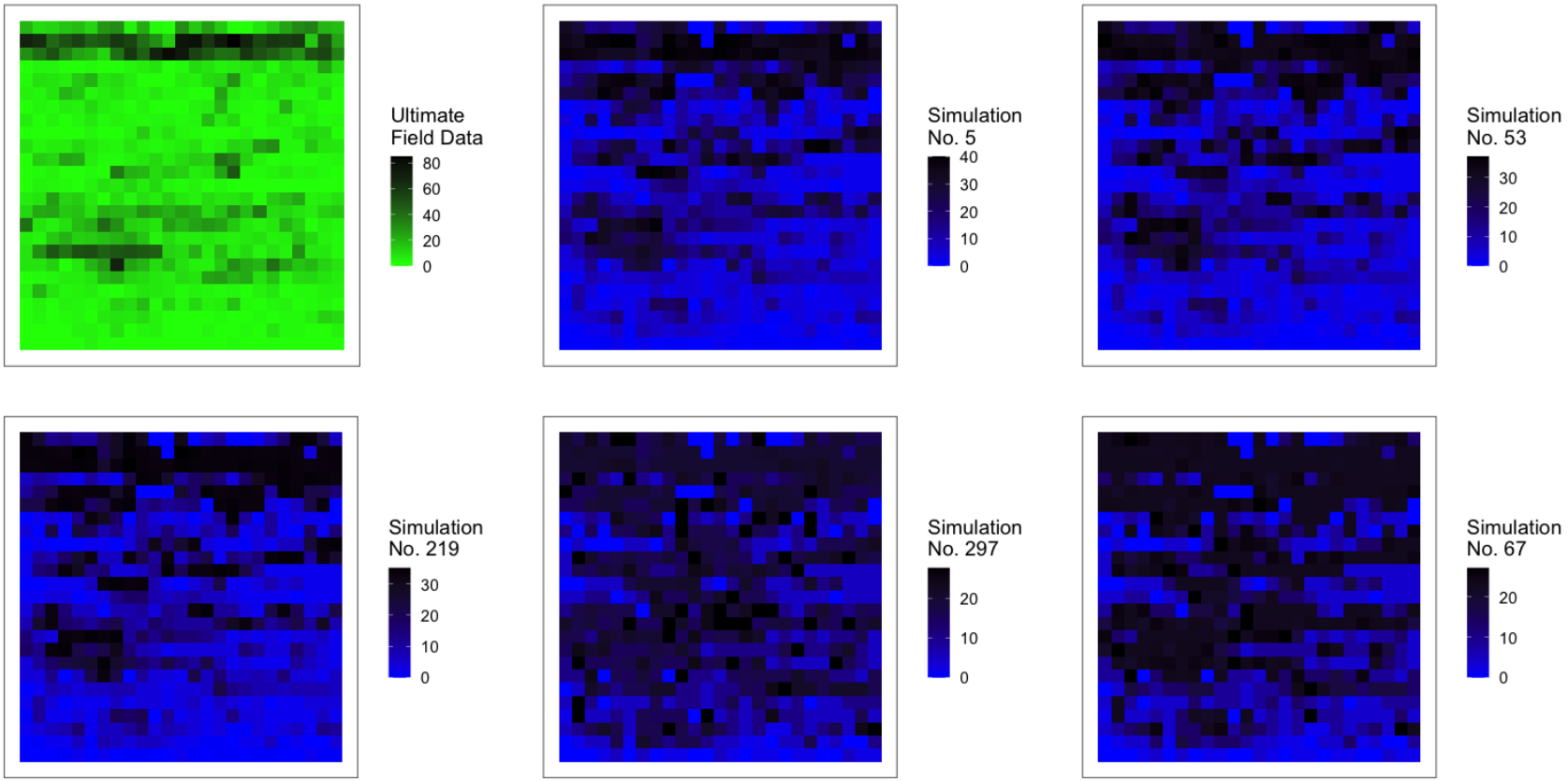

Figure 2.30. Heatmaps of field infection severity for Homerville, GA 2011 ultimate (day 53) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

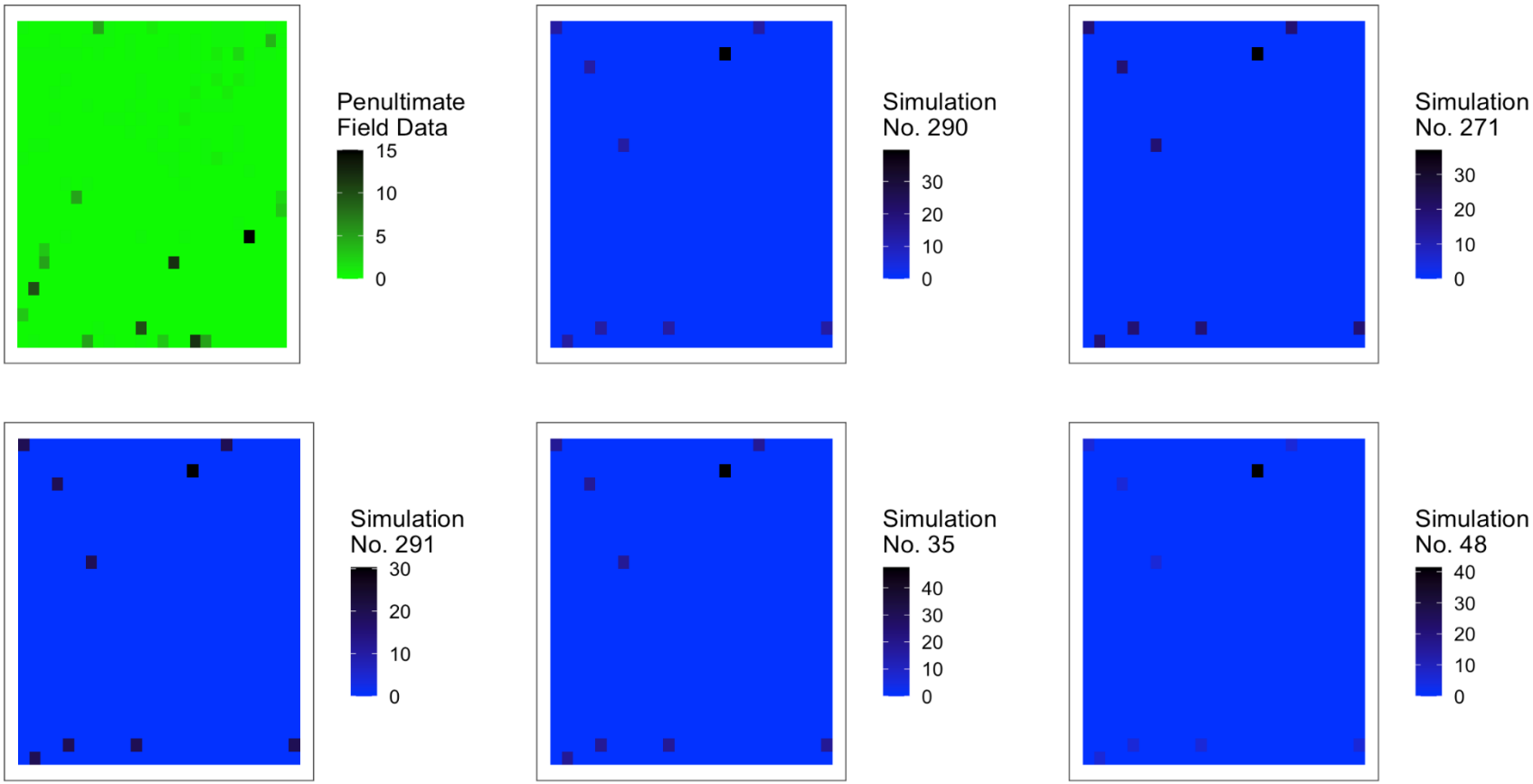

Figure 2.31. Heatmaps of field infection severity for Homerville, GA 2012 penultimate (day 114) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

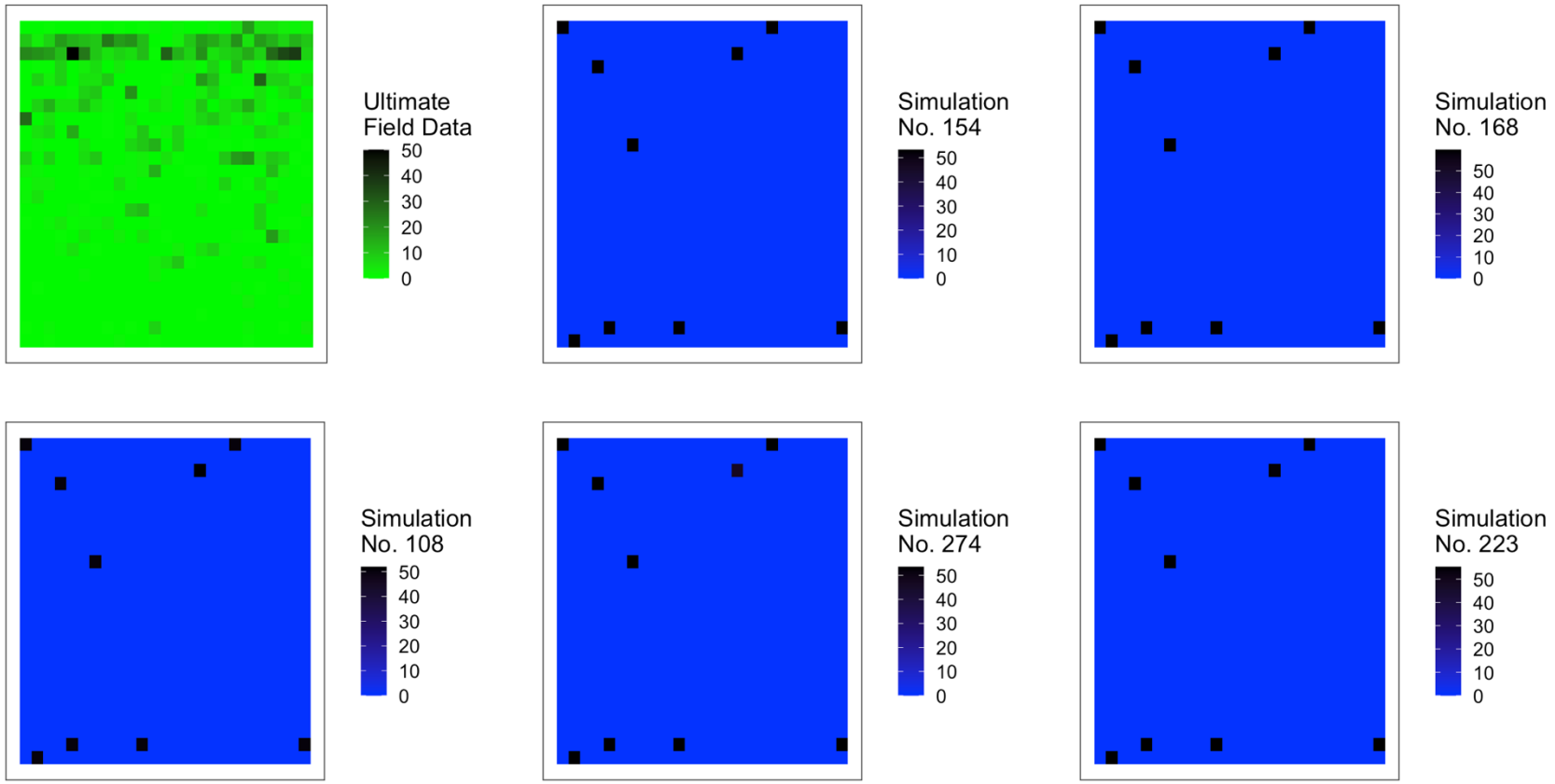

Figure 2.32. Heatmaps of field infection severity for Homerville, GA 2012 ultimate (day 135) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

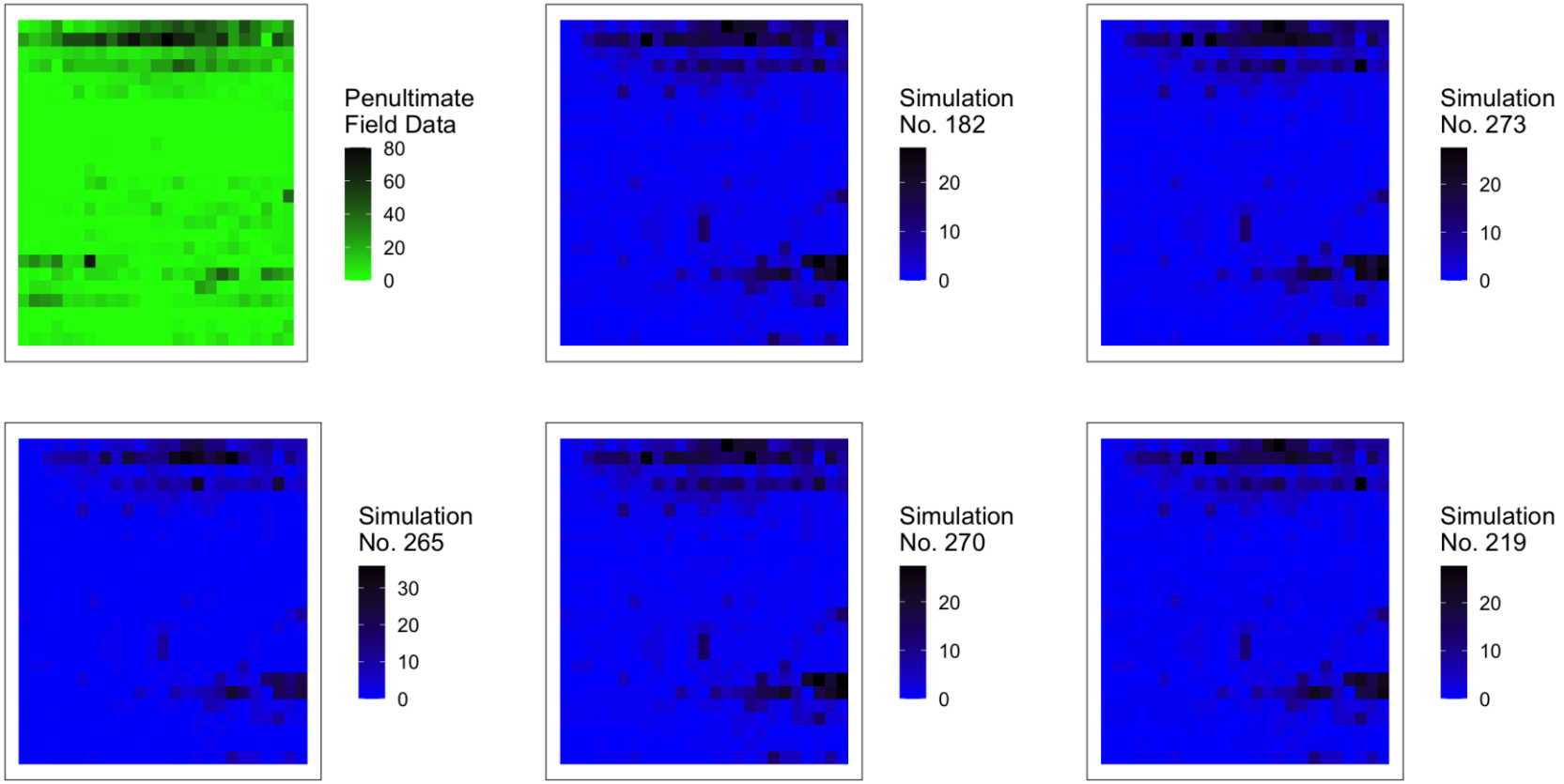

Figure 2.33. Heatmaps of field infection severity for Enigma, GA 2011 penultimate (day 29) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

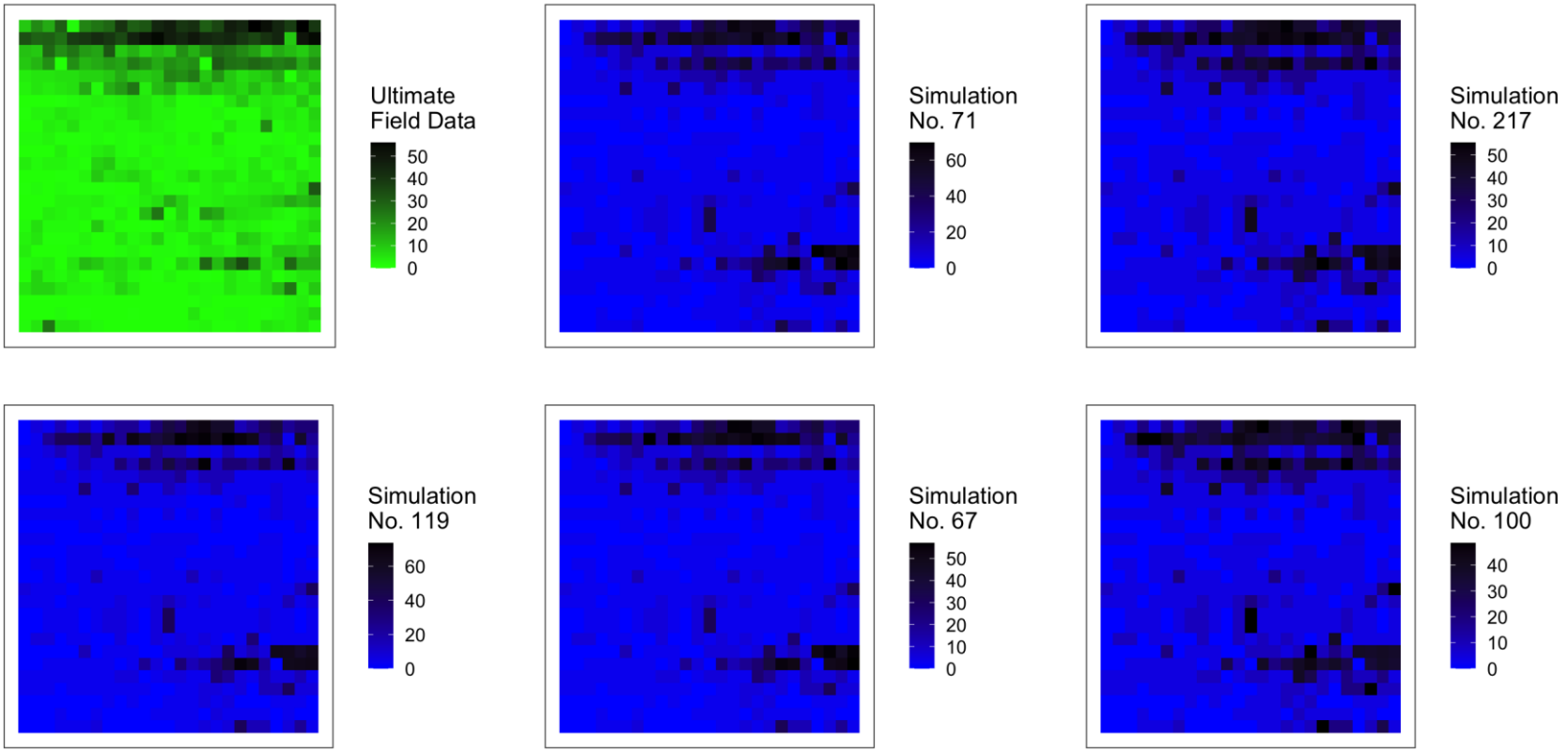

Figure 2.34. Heatmaps of field infection severity for Enigma, GA 2011 ultimate (day 47) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

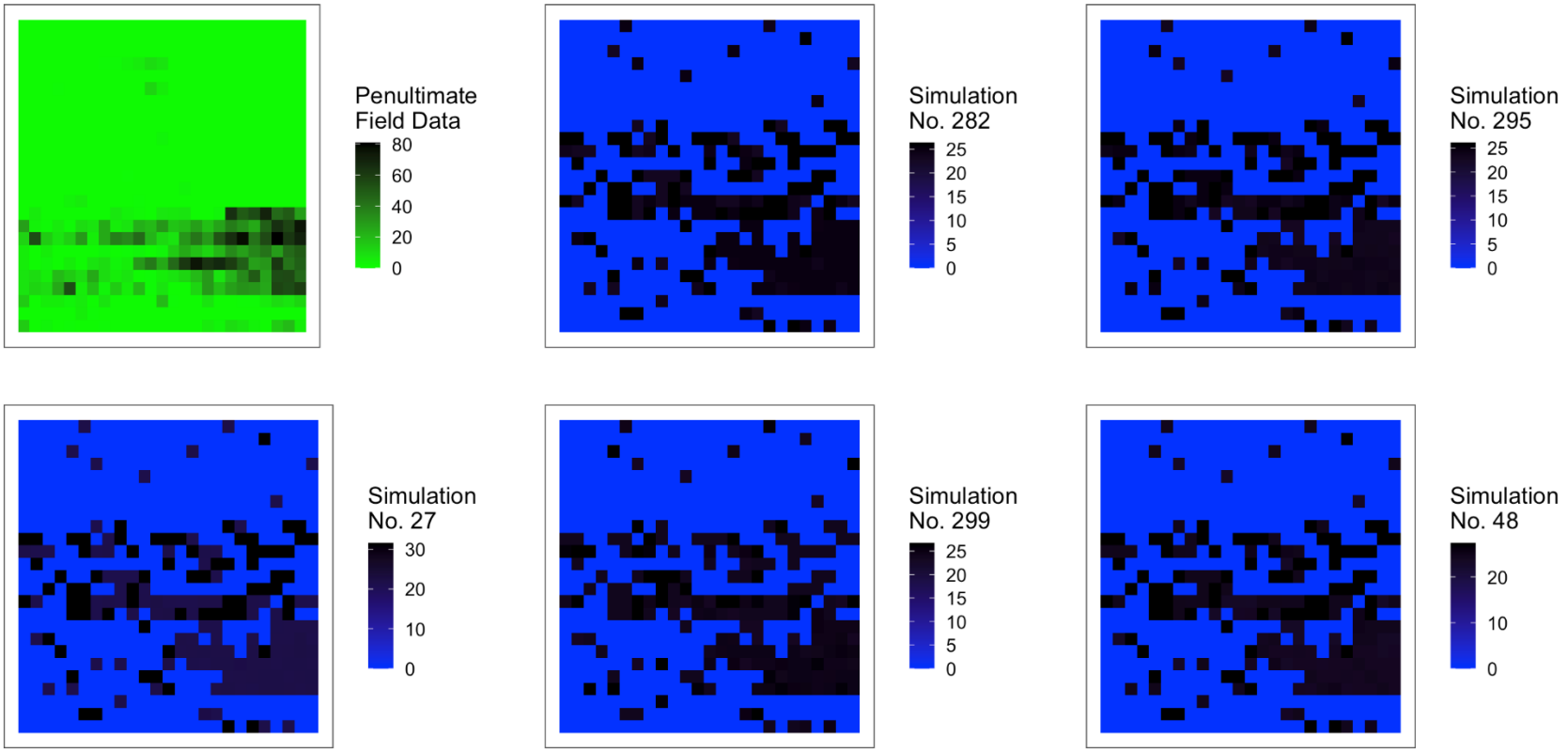

Figure 2.35. Heatmaps of field infection severity for Enigma, GA 2012 penultimate (day 106) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

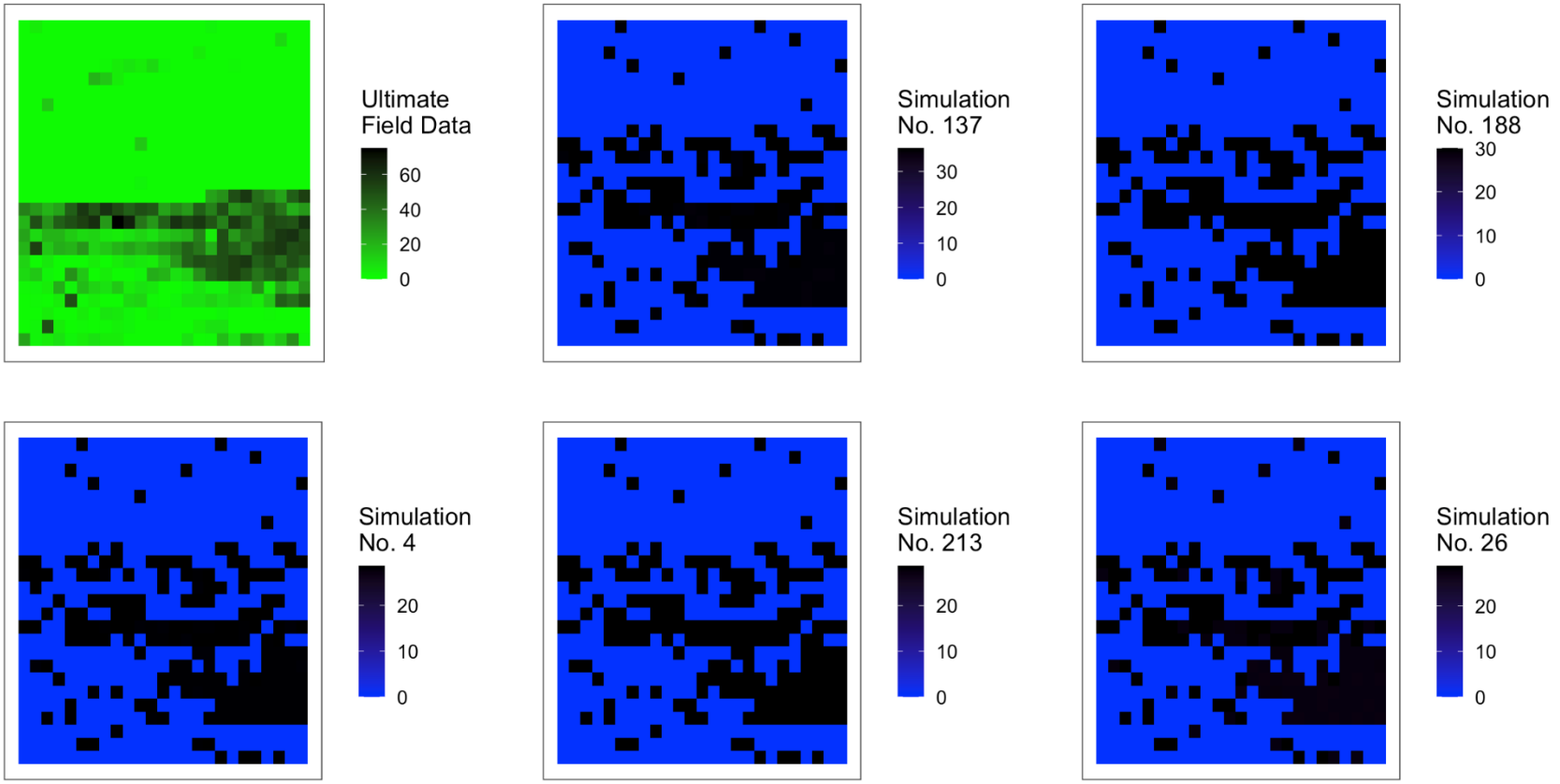

Figure 2.36. Heatmaps of field infection severity for Enigma, GA 2012 ultimate (day 127) field data and field infection saturation for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean. Rows not containing SHB plants are omitted from actual and virtual fields.

CHAPTER 3

A SPATIALLY-EXPLICIT EPIDEMIOLOGICAL MODEL FOR A WHITEFLY-TRANSMITTED VIRUS COMPLEX OF YELLOW SQUASH [2]

[2] Campbell, A. To be submitted to a peer-reviewed journal.

ABSTRACT

The spread of a whitefly-transmitted virus complex (WTVC) of cucurbits has been studied in yellow squash (*Cucurbita pepo*) for the last several years in the southeast US. This virus complex consists of three systemic diseases of cucurbits: Cucurbit Leaf Crumple Virus (CuLCrV), Cucurbit Chlorotic Yellows Virus (CCYV), and Cucurbit Yellow Stunting Disorder Virus (CYSDV). I used data from published studies of this virus complex that provided disease incidence values to parameterize and evaluate a simulation model for its spread that provides real-time yield loss estimates. This whitefly model was adapted from SimpleMite described in Chapter 2. Moran's I, a global index of data aggregation over space, mean disease levels, the statistical distribution of disease levels over space on two different time points, and infection incidence heat maps produced overlapped with the published whitefly field data. The whitefly model tended to overestimate the amount of disease compared with field data and also yielded some different spatial distributions of disease in the different stochastic runs. Additional information on the dispersal behavior of whiteflies, especially in relation to the WTVC-infected plants, would likely improve model accuracy. The use of the base model for two starkly different pathosystems lends credibility to the assertion that it can be used for a variety of arthropod-vectored viruses.

INTRODUCTION

For plant pathologists and those managing crop health, the threat of plant viruses is of particular and increasing concern. Although the number of documented plant viruses is only about 10% of the ~ 8000 species of plant pathogenic fungi (Westcott & Horst, 2008), plant viruses constitute 47% of all emerging plant diseases – a combination of newly discovered plant diseases, those that have jumped to novel host plants, and pathogens that have colonized new geographic regions. A similar estimate of emerging fungal plant pathogens is notably less at~ 30%, though there are considerably more plant pathogenic fungo than viruses (Anderson *et al.*, 2004) These differences suggest that viral plant diseases are being discovered at an increasing rate compared with fungal plant pathogens, viral diseases are proportionally increasing in prevalence more quickly than fungi and the significance of viral plant diseases to crop production is a burgeoning issue.

Arthropods are the most significant vectors of plant viruses (Luria & Darnell Jr., 1967). Among the most infamous of arthropod plant disease vectors is *Bemisia tabaci* Gennadius (Order Hemiptera, Family Aleyrodidae), also referred to as the whitefly, silverleaf whitefly, or the sweet potato whitefly. *Bemisia tabaci* is a world-wide pest of tremendous importance for agricultural production. This insect is currently distributed across all continents, except for Antarctica, and occurs throughout tropical and warm temperate regions of the world (Sani *et al.*, 2020). Even in the absence of plant viruses, whiteflies can cause substantial physiological damage to crops. *B. tabaci* negatively impacts plant health through phloem feeding and honeydew secretion, causing physiological disorders, irregular ripening, secondary infections, reduction in photosynthetic ability, reduction in fruit quality, and reduced yields (Baig *et al.*, 2015,

Barman *et al.*, 2022, Byrne & Bellows Jr, 1991, Dutta *et al.*, 2018, Jiménez *et al.*, 1995, Jones, 2003, Li *et al.*, 2021, Navas-Castillo *et al.*, 2011, Van Lenteren & Noldus, 1990). Additionally, *B. tabaci* is incredibly prolific as a vector of plant diseases, transmitting more than 150 different plant pathogens to a large number of different host species (Stansly & Naranjo, 2010). Combined with its vast geographic distribution and significant genetic diversity (as evidenced by its many biotypes discussed later in this section), *B. tabaci* presents considerable challenges to those seeking to combat its effects on plant health.

Damage from *B. tabaci* in regions in which insecticides are not widely available can be particularly devastating. For example, annual economic losses in Sub-Saharan Africa resulting from cassava mosaic disease (CMD) and cassava brown streak disease (CBSD), the causal agents of which are vectored by *B. tabaci*, are estimated at $1B USD (Parry *et al.*, 2020). Cassava's position as a culturally significant staple crop for the region means that this crop loss transcends economics (Falade & Akingbala, 2010, Montagnac *et al.*, 2009). Despite the widespread availability of insecticides, economically advanced nations are not immune to whitefly destruction. In 2017, damage to several crops resulting from whitefly infestations cost the Georgia (USA) vegetable industry $161.2M USD (Li et al., 2021).

Though the whitefly was first described in 1736, it was not considered an insect of agricultural concern until it began causing major damage to crops in the late 19th century (Byrne & Bellows Jr, 1991, Sani *et al.*, 2020). The study of *B. tabaci* has historically been ripe with controversy. While it was originally suggested that the insect originated in Asia, it is now believed and supported through phylogenetics, that its true

origin is in Sub-Saharan Africa (Boykin *et al.*, 2013, Byrne & Bellows Jr, 1991, Jones, 2003, Mugerwa *et al.*, 2018, Stansly & Naranjo, 2010). Adding further confusion, there has long been debate about whether *B. tabaci* is a "complex species or a species complex" (Xu *et al.*, 2010). The current understanding posits that *B. tabaci* is a cryptic species complex composed of at least 40 morphologically indistinct species (Li et al., 2021). These cryptic species (formerly termed biotypes) appear to be mostly reproductively and genetically isolated from one another. Some biotypes produce inter-specific hybrids showing a continuum of partial to complete pre-zygotic isolation and others produce a low percentage of offspring with potentially reduced reproductive capacity, suggesting that these biotypes can be considered as separate biological species (Coyne & Orr, 2004, Xu et al., 2010).

All whiteflies are classified in the family Aleyrodidae, the only group in the superfamily Aleyrodoidea, in the order Hemiptera (Foottit & Adler, 2018). Aleyrodidae is further divided into three subfamilies of extant whiteflies: Aleurodicinae, Aleyrodinae, and Udamoselinae. While there are more than 1500 species of whiteflies, *Bemisia* and *Trialeurodes,* are the only two whitefly genera that contain species known to transmit plant viruses, with *B. tabaci* serving as the vector for the vast majority of them (Jones, 2003, Li et al., 2021, Sani et al., 2020). *Bemisia tabaci* is also the exclusive vector of the plant virus genus *Begomovirus* (Rosen *et al.*, 2015, Valverde *et al.*, 2004).

Whiteflies are hemimetabolous insects. They have an egg stage, followed by four nymphal instar stages which culminate in the adult whitefly. Immature stages of the whitefly can acquire viral particles from phloem feeding but can only carry them into adulthood if the virus is persistently transmitted (Jones, 2003, Perilla-Henao & Casteel,

2016). The egg stage is entirely stationary and attaches to the underside of the leaf tissue by an egg pedicel, a structure through which the egg absorbs water from the leaf (Baig et al., 2015, Stansly & Naranjo, 2010). *B. tabaci* emerges from the egg chorion after hatching as a first-instar nymph with fully developed legs, an oval-shaped body, and no wings. This stage first undergoes a very brief "crawler" period in which the juvenile whitefly walks across the leaf surface in search of a minor leaf vein from which to feed. Crawlers usually settle near to the egg (a few centimeters) from which they hatched. If hatched in an unsuitable location, they can travel to an adjoining leaf, but many do not survive this, making oviposition site selection a significant determinant of first-instar mortality (Byrne & Bellows Jr, 1991, Stansly & Naranjo, 2010). After selecting a suitable feeding site, the crawler inserts its stylet in the leaf tissue and remains sessile throughout the following three nymphal stages (Butter & Dhawan, 2021, Stansly & Naranjo, 2010). Second-, third-, and fourth-instar nymphs are also oval-shaped and wingless but have underdeveloped legs in comparison to crawlers (Li et al., 2021). These legs primarily seem to serve the function of firmly attaching the instars to the leaf surface as they are non-functional for walking (Stansly & Naranjo, 2010). These three primary whitefly life stages meet their nutritional needs by continuing to feed on phloem in the same location in which they settled as a crawler.

Though the end of the fourth-instar nymphal stage is often referred to as the "pupal" stage, whiteflies are hemimetabolous and do not form true pupae (Byrne & Bellows Jr, 1991, Li et al., 2021, Stansly & Naranjo, 2010). The fourth-instar whitefly retracts its stylet from the leaf tissue prior to eclosion (Stansly & Naranjo, 2010). The adult whitefly then emerges, leaving the pupal case behind. It has a yellow body with

well-developed legs and hyaline wings that are coated in an epicuticular wax secreted by wax plates present on the abdomen (Baig et al., 2015, Byrne & Bellows Jr, 1991, Stansly & Naranjo, 2010). Adult whiteflies possess a pair of segmented antennae, two ocelli, and compound eyes (with low visual acuity) that consist of an upper eye and lower eye (Baig et al., 2015, Stansly & Naranjo, 2010). Copulation usually takes place within hours to days of adult emergence, depending on the season (Byrne & Bellows Jr, 1991). Adult *B. tabaci* can reproduce using a form of arrhenotokous parthenogenesis (also known as haplodiploidy) in which females are diploid and males are haploid (Byrne & Bellows Jr, 1991). In this parthenogenetic system, unfertilized eggs become haploid males and fertilized eggs become diploid females (Butter & Dhawan, 2021, Li et al., 2021, Normark & Kirkendall, 2009). However, they most often reproduce sexually by which they can produce males or females (Manzari & Fathipour, 2021).

Some *B. tabaci* species apparently present more of a threat to crop systems than others. Of particular concern to plant health are *B. tabaci* Middle East Asia Minor 1 (MEAM1) and *B. tabaci* Mediterranean (MED), formerly termed the B biotype and *B. argentifolii*, and Q biotype, respectively. In the late 20th century, MEAM1 began a global invasion that would result in the displacement of many of the indigenous whiteflies of the world (Liu *et al.*, 2007, McKenzie *et al.*, 2004, Xu et al., 2010). In the United States, MEAM1 arrived in the 1980s and promptly displaced *B. tabaci* New World (formerly termed the A biotype), which was not considered a major pest (Barman et al., 2022). In the early 21st century, MED invaded on a global scale and displaced indigenous whiteflies in many regions, though it has yet to displace MEAM1, perhaps due to the ability of MEAM1 to outcompete MED and prevent copulation (Liu et al., 2007, Pascual

& Callejas, 2004, Xu et al., 2010). To date, MEAM1 is the dominant whitefly in Georgia agricultural systems (Barman et al., 2022).

Though they are morphologically indistinct, *B. tabaci* cryptic species differ in ways that are relevant to disease outcomes and therefore range significantly in their ability to cause damage to plants. MEAM1 and MED are the most destructive of all the cryptic species (Rosen et al., 2015). Between the two of them, MEAM1 has a broader host range and can more efficiently transmit begomoviruses than MED and the other cryptic species (Fiallo-Olivé *et al.*, 2020, Rosen *et al.*, 2015). *Bemisia tabaci* can transmit several different virus genera, but of particular concern throughout the world are in impacts of begomoviruses and criniviruses on agricultural crops, including fruit and vegetable production in Georgia.

Three viruses of cucurbits occur in mixed infections in Georgia: Cucurbit Leaf Crumple Virus (CuLCrV), Cucurbit Chlorotic Yellows Virus (CCYV), and Cucurbit Yellow Stunting Disorder Virus (CYSDV) (Kavalappara *et al.*, 2021a, Kavalappara *et al.*, 2021b). These viruses are all vectored by *Bemisia tabaci*, the sweet potato whitefly. CYSDV and CCYV are in the genus *Crinivirus* and the family *Closteroviridae* and produce nearly identical symptoms: mottling, yellowing, and, in severe cases, stunting (Dutta et al., 2018, Kavalappara et al., 2021b, Wintermantel *et al.*, 2019). CuLCrV is in the genus *Begomovirus* and the family *Geminiviridae* and causes pale, curled, or crumpled leaves as well as stunting and growth distortions in severe cases (Dutta et al., 2018). Individually, CuLCrV is transmitted in a persistent manner and CCYV and CYSDV are transmitted in a semi-persistent manner. Little is known about how these viruses come to be found in mixed infections; it is unknown if whiteflies transmit more

than one of these viruses simultaneously or if host plants are individually infected with the complex. Given the frequency that these viruses to co-occur in a Whitefly-Transmitted Virus Complex (WTVC), an epidemiological simulation model of disease spread can range from a simple one-vector one-disease system to one vector with multiple pathogens that when combined cause different diseases (Codod *et al.*, 2022, Kavalappara et al., 2021a). The following model conceptualizes the WTVC as a three-virus complex in which all three viruses are transmitted by viruliferous whiteflies at one time.

Given its wide host breadth, global distribution, and devastating economic consequences there is a vast selection of scientific literature written on the *B. tabaci* species complex. There are, however, no whitefly specific simulation models that reasonably account for whitefly dispersal behaviors and disease spread. When insect-vectored plant disease simulation models are created, they are unfortunately often not open access, making it difficult for other users to improve upon a model or adapt it for different scenarios (Parry et al., 2020, Roy *et al.*, 2021). The simulation model described in this chapter is generally parameterized for the aforementioned WTVC in yellow squash (*Cucurbita pepo*) fields, but is generalizable to similar small, winged invertebrate plant disease vectors and will be made available to the public as a subroutine that runs within the HexSim 4.0 environment (which is open access). This will enable others access to and provide opportunities for refinement and adaptation to other similar insect vectored plant disease systems.

Arthropod-vectored plant disease models should include vector behavior, vector life history, and account for disease-relevant parameters such as acquisition access period

(AAP) and inoculation access period (IAP). In these ways, the simulation models described in Chapter 2 and this chapter are similar. These models differ in that the simulation described in Chapter 2 represents a non-systemic viral disease with a comparably sedentary vector (eriophyid mite) while the simulation in this chapter describes a systemic disease virus complex vectored by a considerably more active, winged invertebrate vector. Although SimpleMite served as the backbone for this model, there are significant departures between the models that specifically account for differing vector lifestyles and disease dynamics. Given the high degree of similarity between whitefly cryptic species, these systems can be easily represented by the following model with small alterations in biological parameters.

MATERIALS AND METHODS

Field data. This study utilized published field data of disease incidence of a Whitefly-Transmitted Virus Complex (WTVC; composed of CuLCrV, CCYV, and CYSDV) in Georgia yellow squash fields (Codod et al., 2022). In the field study (Codod et al., 2022), disease incidence was reported as the percentage of diseased plants in each ~3 x 1 m field quadrat. Each quadrat had an average of ~9 squash plants but the number of plants appeared to vary from 6 to 13 plants depending on specific locations (which were not provided in the publication). Field data (hereafter termed, 'WTVC data') was recorded weekly for a total of four weeks (Codod et al., 2022). Figure 1 from Codod *et al.*, 2022 was used to configure experimental squash fields in the HexSim modelling environment. The authors of the WTVC study recorded data for experimental fields for

2019 and 2020, but only data from the field in 2020 (Figure 4C from Codod *et al.*, 2020) was used for model performance comparisons in my study.

Model overview. A simulation model for the spread of a whitefly-transmitted virus complex (hereafter termed 'WTVC Model') was developed in the HexSim modelling environment by modifying SimpleMite (see **Chapter 2**). Most elements of SimpleMite were retained in the WTVC Model and only the differences between these models will be highlighted below. These changes were necessary to adapt the virtual landscape and epidemiological/vector parameters to the Codod et al. (2022) data and whitefly-virus pathosystem.

Spatial data. Maps were constructed to replicate squash fields as described by Codod et al. (2022) in Figure 1. Experimental fields from the WTVC data consisted of 30 rows by 10 columns of quadrats. Codod *et al.* reported that each quadrat contained on average 9 squash plants (Codod et al., 2022). The virtual landscape created to replicate the study was constructed of 1 m x 1 m hexagons, meaning that each ~ 3 x 1 m quadrat was represented by a line of three hexagons. Simulation disease incidence was measured on a per-quadrat basis. For ease of calculation and interpretation, and because the number of plants in each quadrat in the field study appears to be inconsistent (as discussed above), each quadrat in the simulation represented 10 yellow squash plants – a very slight departure from the reported field quadrat average of 9 plants. Because each simulated quadrat represented 10 squash plants, each hexagon in the simulation represented an average of 3 ⅓ squash plants. To distinguish between quadrats in the field study and quadrats in the virtual landscape of the model, they will hereafter be termed "field quadrats" and "simulated quadrats", respectively.

Demographic information. The vector population was again divided into three life stages: Stage 0 (whitefly eggs and first instar), Stage 0 (second through fourth whitefly instars), and Stage 2 (adult whiteflies). The life stages were divided in this manner because egg to adult development time tends to be ~16 days at 30 °C. temperatures like those found in Tifton, GA in September 2020, when the study began (Bayhan *et al.*, 2006, Butler *et al.*, 1983). This development time roughly aligns with the 7-day time step of the model in which egg to adult development (the time it takes for an individual in the model to progress from Stage 0 to Stage 2) takes two time steps, or 14 days. Also, adult whiteflies tend to live for at least 7 days (Butler et al., 1983). This time step was chosen because it reconciles both the whitefly life cycle and the data collection dates for the WTVC data, which occurred every 7 days.

Because eggs and nymphal stages of *B. tabaci* are incapable of dispersal, only Stage 2 whiteflies are capable of dispersal and virus transmission. Survival and fecundity rates were gathered from published life tables for *Bemisia tabaci* (**Table 3.1**). Population size was projected using a Leslie Matrix (**Appendix A**) for each simulated quadrat based on vital rates for each life stage at each time step. Fecundity values represent average egg production for one time step. Because WTVC is a females-only model, estimates for fecundity carrying capacity were halved, which assumes a 1:1 sex ratio. Sex ratios vary by cryptic species, but MEAM1 has been reported to have a sex ratio of around 1:1 (Ahmed *et al.*, 2014). Carrying capacity was calculated for adult whiteflies and roughly estimated at 50 adults per 100 cm^2 of leaf area and is based on the whiteflies per cm^2 reported by Zhang *et al.* (2014). At the beginning of the first time step, simulated quadrats were estimated to have a total leaf area of 3,000 cm^2 based on 10 squash plants

having individual leaf area of 300 cm^2 (Loy, 2004). Carrying capacity was halved to represent female whiteflies only and one whitefly in the WTVC model simulation represents 10 real whiteflies, leading to a carrying capacity of 75 adults in the model (see **Table 3.1**). Because cucurbits tend to undergo rapid leaf expansion, carrying capacity increased at each time step up to a value of 375. Incorporating leaf expansion into the calculation of carrying capacity was a simpler approach than altering the map. This increase in carrying capacity is an estimate derived from general trends in cucurbit leaf growth patterns (Loy, 2004), and host plant size data was not reported by Codod et al. (2022).

The WTVC model posed the additional challenge of representing whitefly dispersal, which can be challenging due to the capacity of *B. tabaci* for active flight and dispersal decisions. Isaacs and Byrne (1998) conducted a mark-recapture study on whiteflies in melon fields in which they described the relationship between the proportion of female whiteflies recaptured and the distance they traveled (Isaacs and Byrne, 1998 Fig. 4). Preliminary data (unpublished) demonstrated that this relationship alone misrepresented the spatial patterns of disease outcomes demonstrated in the WTVC data, perhaps because the authors recorded whitefly dispersal distance less than 24 hours after marking likely underestimating the distance a n individual whitefly travels in the course of 7 days (the length of the WTVC model time step) (Isaacs & Byrne, 1998). Isaacs and Byrne (1998) also excluded data from collection days on which less than 100 whiteflies were recaptured (Isaacs & Byrne, 1998). This dispersal kernel was retained in the model and a randomly varying integer under 10 was added to the resulting value to facilitate the dispersal of whiteflies to neighboring simulated quadrats.

WTVC Model event sequence.

The event sequence structure for WTVC is nearly identical to that of SimpleMite (see **Chapter 2**). Disease was introduced initially through the Initial Disease Conditions Map that mirrored the disease incidences (per simulated quadrat) represented in the initial data collection date (11 DAP as shown in Codod *et al.*, 2022 Figure. 4C) for WTVC data. An initial population of whiteflies was added to the model. Preliminary simulation results data (unpublished) demonstrated that disease outcomes were underestimated when only adult whiteflies were added to the model at this stage. Seven hundred whiteflies (translating to 70 virtual whiteflies) of each stage class were added to each simulated quadrat during the initialization of the simulation. Because some plants were already diseased at the time of the initial collection date (11 days after planting) of the field study, it seemed reasonable to assume that whiteflies were already present and had reproduces in the field. Equal numbers of each life stage were populated within each diseased simulated quadrat to reduce the influence of the starting number of whiteflies on the population dynamics.

In the WTVC model, disease status was determined on a per-simulated quadrat basis as opposed to a per-plant basis like in the SimpleMite model. Simulated quadrats that contain at least one diseased plant were considered WTVC-Positive. WTVC-Positive simulated quadrats always begot viruliferous vectors and WTVC-Negative simulated quadrats always begot non-viruliferous vectors. When viruliferous vectors dispersed and arrived at a WTVC-Negative simulated quadrat, their new location became exposed to the virus complex. In the following time step, simulated quadrats that were previously

exposed to the virus complex become WTVC-Positive, to represent a latent period of ~1 week.

As in the field study, an increase in the disease incidence of simulated quadrats represented an increase in the number of WTVC-Positive plants within it. WTVC-Positive simulated quadrats experienced an increase in their disease incidence upon receiving viruliferous mites and with each passing time step. To represent the increase in the number of plants diseased in each simulated quadrat, the accumulator value for disease incidence of a WTVC-Positive simulated quadrat increased by a value of 15 plus a randomly varying integer (the absolute value of the result of 20 multiplied by a random number from a normal distribution of mean 0 and standard deviation 1) every time that quadrat received viruliferous mites and with each passing time step. This is because viruliferous whiteflies entering a WTVC-Positive simulated quadrat presumably infect more plants upon entering, and viruliferous whiteflies within a WTVC-Positive quadrat likely move between plants in the quadrat, spreading disease.

There is not yet precise information on the relationship between infection severity, time since infection, whitefly pressure, and their impact on crop yields for this virus complex. However, this model includes a real-time yield loss sub model that predicts percentage yield loss based on these factors to serve as a framework for the incorporation of this information in the future. The yield loss subroutine calculates the whitefly pressure of each simulated quadrat as the proportion of the adult carrying capacity that has been reached and assigned it to one of three categories: Low (>0.25), Medium (>0.50), or High (<0.75). The model categorized the simulated quadrats into an "infection saturation class" in a very similar manner, based on their infection incidence.

These two factors were used in the Disease Dynamics event group to construct a yield loss map of the field that updates with each time step. Each simulated quadrat is given a two-part score: a score of 1, 2, or 3 for Low, Medium, or High whitefly pressure, respectively, and a score of 1, 2, or 3 for being in a Low, Medium, or High infection saturation class, respectively. The maximum score that any simulated quadrat can have at each time step is 6 (a score of 3 for both whitefly pressure and infection saturation class). For each time step, the proportion of the maximum score reached for each simulated quadrat is recorded and added to a "Cumulative Yield Loss Score" accumulator. If a simulated quadrat has a score of 6 for four time steps in a row (which is equivalent to the maximum Cumulative Yield Loss Score of 4), this would be considered 100% yield loss (or entirely non-marketable fruit). If a simulated quadrat has a score of 3 for three time steps in a row (which is equivalent to a Cumulative Yield Loss Score of 0.75), this would be considered a yield loss score of 18.75% (~ 19% reduction in fruit weight). Scoring yield loss in this way allows the incorporation of time since infection, infection incidence, and whitefly pressure as contributing factors. See **Appendix B** for an example of a sequence of heatmaps constructed from the real-yield loss maps.

RESULTS

Penultimate Data Collection Date.

When I plotted the field data from the penultimate date of the WTVC study against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection incidences overlapped with the field data (**Figure 3.1**). This suggested that despite a large number of widely varying stochastic

simulations (± 20% survival and fecundity of whiteflies) the field data outcome was projected by the model from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 3.3**). The maximum infection saturation values of the simulation percentile groups reached 100, as did the maximum of the field data (**Figure 3.3**). Field total infection incidence did not fall in the range of the total infection incidence values for the simulation (**Figure 3.5**), and the simulations produced notably lower disease levels when compared to the penultimate field data (**Table 3.2**).

Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected simulations were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 3.3** and **Figure 3.7**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 3.3** and **Figure 3.7**). Summary statistics and Moran's I results suggested that the model projections were within a reasonable range of disease distribution and patchiness, while the heatmaps indicated that the locations of disease in the simulations could vary substantially between simulation runs and when compared with the field data (**Figure 3.8**).

Ultimate Data Collection Date.

When I plotted the field data from the penultimate date of the WTVC study against the data for the nearest corresponding simulation time step, the distribution of all corresponding simulation mean infection incidences overlapped with the field data (**Figure 3.2**). This suggested that despite a large number of widely varying stochastic

simulations (± 20% survival and fecundity of whiteflies) the field data outcome was projected by the model from the same initial disease configuration. When the infection saturation means were separated into percentile groups, the range of all groups overlapped with the corresponding field data (**Figure 3.4**). The maximum infection saturation values of the simulation percentile groups reached 100, as did the maximum reported values in the field data (**Figure 3.4**). Total infection incidence did not fall in the range of the total infection incidence values for the simulation (**Figure 3.6**), but the simulation results were not notable outliers in terms of the percentage of maximum disease incidence (**Table 3.2**).

Moran's I values, an index of the global degree of aggregation of disease in space, of all 20 randomly selected simulations were positive and had statistically significant p-values ($\alpha < 0.05$) (**Table 3.3** and **Figure 3.7**). The Moran's I value of the corresponding field data was also positive and had a statistically significant p-value ($\alpha < 0.05$) (**Table 3.3** and **Figure 3.7**). Summary statistics and Moran's results suggested that the model projections were quantitatively comparable to the field data and qualitatively the disease severity heatmaps based on field data and randomly selected simulation runs produced similar levels and patchiness of WTVC, but WTVC locations were again variable (**Figure 3.9**).

DISCUSSION

The WTVC model more accurately predicted the disease outcomes for the ultimate data collection date from the WTVC field data than that of the penultimate collection date. While WTVC model overestimated the total infection incidence

demonstrated in the field data for both data collection dates (**Table 3.2**; **Figures 3.5** and **3.6**), this overestimation was greater for the penultimate data collection date. This difference in the degree of inaccuracy of the model projections suggests a potential shift in disease dynamics that occurs as fields become more infected with disease or as more of the whiteflies in the field become viruliferous. Such as shift in dynamics is plausible as whitefly-vectored viruses have been demonstrated to influence whitefly behavior (Domingo-Calap *et al.*, 2020, Moreno-Delafuente *et al.*, 2013, Shi *et al.*, 2018b). Additionally, we assumed a one week latent period for disease development. But, this maybe inaccurate, as WTVC is a complex with the molecular interactions related to disease expression unknown. If the latent period is longer than 7 days, this could explain the differences between penultimate and ultimate disease projections and the field data.

In general, comparable Moran's I values for field and simulation data indicated that the WTVC model represented the degree of disease aggregation in space reasonably well (**Table 3.3**; **Figure 3.7**). Total infection incidence values demonstrated that the model tends to overestimate total amount of disease in the field (**Table 3.2**; **Figures 3.5** and **3.6**). Visual assessment of infection incidence heatmaps (**Figures 3.8** and **3.9**) indicated that more areas of the field remain uninfected with WTVC in the simulation data than in the field data. There are three possible explanations for this phenomenon: 1) at the time of the first data collection by Codod et al. (2022), there could have been plants that were latently infected with WTVC, but had not yet shown symptoms and therefore were not captured in the data as being diseased, or 2) whitefly movement through the field was not accurately represented by the model, or 3) in the field data, whiteflies entering the field non-negligibly contribute to disease within the field, which is not

represented in the model. All of these explanations are plausible and they may occur in combination with each other. When collection of the WTVC data began, some plants were already diseased, meaning that the field had already been exposed to viruliferous whiteflies, making it probable that some latently infected squash plants were not accounted for. The WTVC model assumes that additional whiteflies do not enter the field throughout the simulation (*i.e.* the whiteflies exist in a closed population in which new individuals can only be introduced through reproduction). Given the potential capacity of *B. tabaci* for long-distance travel (Byrne, 1999, Franco De Oliveira *et al.*, 2023) and the breadth of weed hosts that can support *B. tabaci* populations (Barman et al., 2022), it is plausible that whiteflies may enter the field throughout the observation period in significant numbers. It is also possible that the model underestimated the distance (or variety of distances) that is typically traveled by dispersing whiteflies. Estimates of the degree of long-distance dispersal on whiteflies are difficult to obtain as they mostly rely on mark and recapture experiments in which only a small subset of the whiteflies released are recaptured (Byrne, 1999, Doukas & Payne, 2014, Franco De Oliveira et al., 2023, Isaacs & Byrne, 1998, Legarrea *et al.*, 2012). Mark and recapture studies tend to be particularly unhelpful for determining the degree of long-distance dispersal in a population because sampling efforts are usually biased towards the release point (Terui, 2020).The characterization of whitefly dispersal in the WTVC model is important because it directly relates to the resulting locations of disease. The inaccuracy of the model in projecting disease locations (**Figure 3.8** and **3.9**) based on initial conditions potentially suggests that the elements of the model relating to dispersal behaviors may be in need of refinement.

It is clear that whitefly dispersal behavior is complex and difficult to study due to their small size and potentially large population numbers. For example, vector behavior is typically separated into three behavioral categories: host-seeking, feeding, and dispersal (Fingu-Mabola & Francis, 2021). *B. tabaci* is known to respond to visual plant cues and are particularly attracted to green wavelengths of light wavelengths (Fereres & Moreno, 2009). There is also some evidence that whiteflies can be retracted or repelled by plant volatiles and that viral infection and previous visitation by other whiteflies may manipulate these volatiles potentially influencing feeding on uninfected plants when the whitefly is viruliferous (Shi *et al.*, 2018a, Zhao *et al.*, 2022). While some insect vectors of plant viruses land on plants and feed on several different plants before selecting one, whiteflies instead make this decision during the pre-alighting phase and begin feeding after landing on the selected plant (Costa *et al.*, 1991, Fereres & Moreno, 2009). The patterns of whitefly attraction to the cues of an infected plant are virus- and cryptic species-specific and are not accounted for in the WTVC model (Lu *et al.*, 2017, Maluta *et al.*, 2017, Moreno-Delafuente et al., 2013). The omission of these potential relationships could have impacted the accuracy of model spread, especially if there are attraction/repulsion interactions between viruliferous/non-viruliferous whiteflies and/or infected/uninfected host plants.

Because plant viruses rely on vectors for dispersal and transmission, the movement of competent vectors is directly related to the patterns of disease spread. Characterizing the movement patterns of *B. tabaci* has been a frequent subject of interest in published whitefly studies. Adult *B. tabaci* are winged insects that engage in powered flight. Because most of its dispersal occurs across short distances, *B. tabaci* is often

erroneously considered to be a weak flier. However, it has been demonstrated experimentally by Blackmer and Byrne that whiteflies can actively fly against a 15.6 ± 2 cm/s wind. Furthermore, 6% of the *B. tabaci* population under study were capable of sustaining flight in the wind chamber for over 15 minutes (~ 140 meters) and < 1% of individuals capable of the headwind flight for over 2 hours (~ 1.1 km) (Blackmer & Byrne, 1993a). This evidence supports the theory that within *B. tabaci* populations there are two different dispersal modes – one that actively disperses short distances (within fields) and another that disperses long distances (between fields and across a broader landscape), either with or without the aid of winds. This long-dispersal behavior is referred to as "migratory" and the other, short distance dispersal as "non-migratory" (Byrne, 1999, Wolfenbarger, 1946). Dispersal patterns of whiteflies observed in some field experiments were patchy and bimodal, further supporting this bimodal dispersal characteristic of individuals within populations (Byrne *et al.*, 1996). The presence of a migratory morph may explain the high degree of gene flow between populations in an agricultural region and sudden colonization of fields by whitefly vectored plant viruses (Crossley & Snyder, 2020). Though these dispersal abilities have been observed and demonstrated experimentally, there is mixed evidence for a potential underlying causative morphological or definitive genetic difference (Blackmer *et al.*, 1994, Blackmer & Byrne, 1993b, Byrne & Houck, 1990). The debate of the true existence of these dispersal morphs complicates their potential inclusion in a whitefly dispersal model.

Though *Bemisia tabaci* as an agricultural pest has long been under study, the variety and complexity of this cryptic species can make it difficult to discern the important elements of its role as a vector. Improved methods for the study of whitefly

dispersal behavior may significantly further our understanding of whitefly-vectored disease patterns. While there are potentially important elements of the WTVC pathosystem that could refine the model described in this chapter, the existing model provided reasonable replication of in-field disease patterns, demonstrating the ability of the SimpleMite model to serve as a base model for a variety of arthropod-vectored diseases.

To our knowledge, this is the only generalizable spatially-explicit arthropod-vectored plant disease model that includes real-time yield loss projections. Though the data required to parameterize the real-time yield loss subroutine does not yet exist, the framework was included in this model to encourage the collection of such data and to extend its future usefulness by incorporating what is perhaps the most important consequence of plant disease in commercial agriculture.

LITERATURE CITED

Ahmed, M. Z., Naveed, M., Noor ul Ane, M., Ren, S.-X., De Barro, P. and Qiu, B.-L. (2014) Host Suitability Comparison between the Meam1 and Asiaii 1 Cryptic Species of Bemisia Tabaci in Cotton-Growing Zones of Pakistan. *Pest Management Science*, **70,** 1531-1537.

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R. and Daszak, P. (2004) Emerging Infectious Diseases of Plants: Pathogen Pollution, Climate Change and Agrotechnology Drivers. *Trends in Ecology & Evolution,* **19,** 535-544.

Baig, M. M., Dubey, A. K. and Ramamurthy, V. V. (2015) Biology and Morphology of Life Stages of Three Species of Whiteflies (Hemiptera: Aleyrodidae) from India. *The Pan-Pacific Entomologist,* **91,** 168-183.

Barman, A. K., Roberts, P. M., Prostko, E. P. and Toews, M. D. (2022) Seasonal Occurrence and Reproductive Suitability of Weed Hosts for Sweetpotato Whitefly, Bemisia Tabaci (Hemiptera: Aleyrodidae), in South Georgia. *Journal of Entomological Science,* **57,** 1-11.

Bayhan, E., Ulusoy, M. R. and Brown, J. K. (2006) Effects of Different Cucurbit Species and Temperature on Selected Life History Traits of the 'B' Biotype Ofbemisia Tabaci. *Phytoparasitica,* **34,** 235-242.

Blackmer, J., Byrne, D. and Tu, Z. (1994) Behavioral, Morphological, and Physiological Traits Associated with Migratory Bemisia Tabaci (Homoptera: Aleyrodidae). *Journal of Insect Behavior,* **8,** 251-267.

Blackmer, J. L. and Byrne, D. N. (1993a) Environmental and Physiological Factors Influencing Phototactic Flight of Bemisia Tabaci. *Physiological Entomology,* **18,** 336-342.

Blackmer, J. L. and Byrne, D. N. (1993b) Flight Behaviour of Bemisia Tabaci in a Vertical Flight Chamber: Effect of Time of Day, Sex, Age and Host Quality. *Physiological Entomology,* **18,** 223-232.

Boykin, L. M., Bell, C. D., Evans, G., Small, I. and De Barro, P. J. (2013) Is Agriculture Driving the Diversification of the Bemisia Tabaci Species Complex (Hemiptera: Sternorrhyncha: Aleyrodidae)?: Dating, Diversification and Biogeographic Evidence Revealed. *BMC Evolutionary Biology,* **13,** 228.

Butler, G. D., Jr., Henneberry, T. J. and Clayton, T. E. (1983) Bemisia Tabaci (Homoptera: Aleyrodidae): Development, Oviposition, and Longevity in Relation to Temperature1. *Annals of the Entomological Society of America,* **76,** 310-313.

Butter, N. S. and Dhawan, A. K. (2021) *A Monograph on Whiteflies*. Boca Raton: CRC Press.

Byrne, D. N. (1999) Migration and Dispersal by the Sweet Potato Whitefly, Bemisia Tabaci. *Agricultural and forest meteorology,* **97,** 309-316.

Byrne, D. N. and Bellows Jr, T. S. (1991) Whitefly Biology. *Annual review of entomology,* **36,** 431-457.

Byrne, D. N. and Houck, M. A. (1990) Morphometric Identification of Wing Polymorphism in Bemisia Tabaci (Homoptera: Aleyrodidae). *Annals of the Entomological Society of America*, **83,** 487-493.

Byrne, D. N., Robin, J. R., Thomas, V. O. and John, C. P. (1996) Localized Migration and Dispersal by the Sweet Potato Whitefly, Bemisia Tabaci. *Oecologia,* **105,** 320-328.

Codod, C. B., Severns, P. M., Sparks, A. N., Srinivasan, R., Kemerait, R. C. and Dutta, B. (2022) Characterization of the Spatial Distribution of the Whitefly-Transmitted Virus Complex in Yellow Squash Fields in Southern Georgia, USA. *Frontiers in Agronomy,* **4**.

Costa, H. S., Brown, J. K. and Byrne, D. (1991) Host Plant Selection by the Whitefly, Bemisia Tabaci (Gennadius),(Hom., Aleyrodidae) under Greenhouse Conditions. *Journal of Applied Entomology,* **112,** 146-152.

Coyne, J. A. and Orr, H. A. (2004) *Speciation*. Sunderland, Mass: Sinauer Associates.

Crossley, M. S. and Snyder, W. E. (2020) What Is the Spatial Extent of a Bemisia Tabaci Population? *Insects,* **11,** 813.

Domingo-Calap, M. L., Moreno, A. B., Díaz Pendón, J. A., Moreno, A., Fereres, A. and López-Moya, J. J. (2020) Assessing the Impact on Virus Transmission and Insect Vector Behavior of a Viral Mixed Infection in Melon. *Phytopathology®,* **110,** 174-186.

Doukas, D. and Payne, C. C. (2014) Greenhouse Whitefly (Homoptera: Aleyrodidae) Dispersal under Different Uv-Light Environments. *Journal of Economic Entomology,* **100,** 389-397.

Dutta, B., Myers, B., Coolong, T., Srinivasan, B. and Sparks, A. (2018) Whitefly-Transmitted Plant Viruses in South Georgia. University of Georgia Cooperative Extension.

Falade, K. O. and Akingbala, J. O. (2010) Utilization of Cassava for Food. *Food Reviews International,* **27,** 51-83.

Fereres, A. and Moreno, A. (2009) Behavioural Aspects Influencing Plant Virus Transmission by Homopteran Insects. *Virus Research,* **141,** 158-168.

Fiallo-Olivé, E., Pan, L.-L., Liu, S.-S. and Navas-Castillo, J. (2020) Transmission of Begomoviruses and Other Whitefly-Borne Viruses: Dependence on the Vector Species. *Phytopathology®,* **110,** 10-17.

Fingu-Mabola, J. C. and Francis, F. (2021) Aphid–Plant–Phytovirus Pathosystems: Influencing Factors from Vector Behaviour to Virus Spread. *Agriculture,* **11,** 502.

Foottit, R. and Adler, P. H. (2018) *Insect Biodiversity : Science and Society. Volume Ii*. Hoboken, NJ: John Wiley & Sons.

Franco De Oliveira, F., Favara, G. M., Carmo-Sousa, M., Lopes, J. R. S. and Rezende, J. A. M. (2023) Evaluation of Local Movement of Bemisia Tabaci Meam1 and Tomato Severe Rugose Virus Transmission to Tomato Plants Using Marked and Unmarked Whiteflies. *Journal of Applied Entomology,* **147,** 1057-1066.

Isaacs, R. and Byrne, D. N. (1998) Aerial Distribution, Flight Behaviour and Eggload: Their Inter-Relationship During Dispersal by the Sweetpotato Whitefly. *Journal of Animal Ecology,* **67,** 741-750.

Jiménez, D., Yokomi, R., Mayer, R. and Shapiro, J. (1995) Cytology and Physiology of Silverleaf Whitefly-Induced Squash Silverleaf. *Physiological and Molecular Plant Pathology*, **46,** 227-242.

Jones, D. R. (2003) Seasonal Occurrence and Reproductive Suitability of Weed Hosts for Sweetpotato Whitefly, Bemisia Tabaci (Hemiptera: Aleyrodidae), in South Georgia. *European Journal of Plant Pathology,* **109,** 195-219.

Kavalappara, S. R., Milner, H., Konakalla, N. C., Morgan, K., Sparks, A. N., McGregor, C., *et al.* (2021a) High Throughput Sequencing-Aided Survey Reveals Widespread Mixed Infections of Whitefly-Transmitted Viruses in Cucurbits in Georgia, USA. *Viruses,* **13,** 988.

Kavalappara, S. R., Milner, H., Sparks, A., McGregor, C., Wintermantel, W. M. and Bag, S. (2021b) First Report of Cucurbit Chlorotic Yellows Virus in Association with Other Whitefly-Transmitted Viruses in Yellow Squash (Cucurbita Pepo) in Georgia, USA. *Plant Disease,* **105,** 1862.

Legarrea, S., Weintraub, P. G., Plaza, M., Viñuela, E. and Fereres, A. (2012) Dispersal of Aphids, Whiteflies and Their Natural Enemies under Photoselective Nets. *BioControl,* **57,** 523-532.

Li, Y., Mbata, G. N., Punnuri, S., Simmons, A. M. and Shapiro-Ilan, D. I. (2021) Bemisia Tabaci on Vegetables in the Southern United States: Incidence, Impact, and Management. *Insects,* **12,** 198.

Liu, S.-S., De Barro, P., Xu, J., Luan, J.-B., Zang, L.-S., Ruan, Y.-M., *et al.* (2007) Asymmetric Mating Interactions Drive Widespread Invasion and Displacement in a Whitefly. *Science,* **318,** 1769-1772.

Loy, J. B. (2004) Morpho-Physiological Aspects of Productivity and Quality in Squash and Pumpkins (Cucurbita Spp.). *Critical Reviews in Plant Sciences,* **23,** 337-363.

Lu, S., Li, J., Wang, X., Song, D., Bai, R., Shi, Y., *et al.* (2017) A Semipersistent Plant Virus Differentially Manipulates Feeding Behaviors of Different Sexes and Biotypes of Its Whitefly Vector. *Viruses,* **9,** 4.

Luria, S. E. and Darnell Jr., J. E. (1967) *General Virology*. John Wiley & Sons, Inc.

Maluta, N. K. P., Fereres, A. and Lopes, J. R. S. (2017) Settling Preferences of the Whitefly Vector Bemisia Tabaci on Infected Plants Varies with Virus Family and Transmission Mode. *Entomologia Experimentalis et Applicata,* **165,** 138-147.

Manzari, S. and Fathipour, Y. (2021) Whiteflies. In: *Polyphagous Pests of Crops.* (Omkar, ed.). Singapore: Springer Singapore, pp. 183-230.

McKenzie, C., Anderson, P. K. and Villarreal, N. (2004) An Extensive Survey of Bemisia Tabaci (Homoptera: Aleyrodidae) in Agricultural Ecosystems in Florida. *The Florida Entomologist,* **87,** 403-407.

Montagnac, J. A., Davis, C. R. and Tanumihardjo, S. A. (2009) Nutritional Value of Cassava for Use as a Staple Food and Recent Advances for Improvement. *Comprehensive Reviews in Food Science and Food Safety,* **8,** 181-194.

Moreno-Delafuente, A., Garzo, E., Moreno, A. and Fereres, A. (2013) A Plant Virus Manipulates the Behavior of Its Whitefly Vector to Enhance Its Transmission Efficiency and Spread. *PloS one,* **8,** e61543.

Mugerwa, H., Seal, S., Wang, H.-L., Patel, M. V., Kabaalu, R., Omongo, C. A., *et al.* (2018) African Ancestry of New World, Bemisia Tabaci-Whitefly Species. *Scientific Reports,* **8,** 2734.

Navas-Castillo, J., Fiallo-Olivé, E. and Sánchez-Campos, S. (2011) Emerging Virus Diseases Transmitted by Whiteflies. *Annual review of phytopathology,* **49,** 219-248.

Normark, B. B. and Kirkendall, L. R. (2009) Chapter 192 - Parthenogenesis in Insects and Mites. In: *Encyclopedia of Insects (Second Edition).* (Resh, V. H. and Cardé, R. T., eds.). San Diego: Academic Press, pp. 753-757.

Parry, H., Kalyebi, A., Bianchi, F., Sseruwagi, P., Colvin, J., Schellhorn, N., *et al.* (2020) Evaluation of Cultural Control and Resistance-Breeding Strategies for Suppression of Whitefly Infestation of Cassava at the Landscape Scale: A Simulation Modeling Approach. *Pest Management Science,* **76,** 2699-2710.

Pascual, S. and Callejas, C. (2004) Intra-and Interspecific Competition between Biotypes B and Q of Bemisia Tabaci (Hemiptera: Aleyrodidae) from Spain. *Bulletin of Entomological Research,* **94,** 369-375.

Perilla-Henao, L. M. and Casteel, C. L. (2016) Vector-Borne Bacterial Plant Pathogens: Interactions with Hemipteran Insects and Plants. *Frontiers in Plant Science,* **7**.

Rosen, R., Kanakala, S., Kliot, A., Pakkianathan, B. C., Farich, B. A., Santana-Magal, N., *et al.* (2015) Persistent, Circulative Transmission of Begomoviruses by Whitefly Vectors. *Current Opinion in Virology,* **15,** 1-8.

Roy, B., Dubey, S., Ghosh, A., Shukla, S. M., Mandal, B. and Sinha, P. (2021) Simulation of Leaf Curl Disease Dynamics in Chili for Strategic Management Options. *Scientific Reports,* **11**.

Sani, I., Ismail, S. I., Abdullah, S., Jalinas, J., Jamian, S. and Saad, N. (2020) A Review of the Biology and Control of Whitefly, Bemisia Tabaci (Hemiptera:

Aleyrodidae), with Special Reference to Biological Control Using Entomopathogenic Fungi. *Insects,* **11,** 619.

Shi, X., Chen, G., Pan, H., Xie, W., Wu, Q., Wang, S., *et al.* (2018a) Plants Pre-Infested with Viruliferous Med/Q Cryptic Species Promotes Subsequent Bemisia Tabaci Infestation. *Frontiers in Microbiology,* **9**.

Shi, X., Tang, X., Zhang, X., Zhang, D., Li, F., Yan, F., *et al.* (2018b) Transmission Efficiency, Preference and Behavior of Bemisia Tabaci Meam1 and Med under the Influence of Tomato Chlorosis Virus. *Frontiers in Plant Science,* **8**.

Stansly, P. A. and Naranjo, S. E. (2010) *Bemisia: Bionomics and Management of a Global Pest.* Springer.

Terui, A. (2020) Modeling Dispersal Using Capture–Recapture Data: A Comparison of Dispersal Models. *Ecological Research,* **35,** 686-699.

Valverde, R. A., Sim, J. and Lotrakul, P. (2004) Whitefly Transmission of Sweet Potato Viruses. *Virus research,* **100,** 123-128.

Van Lenteren, J. v. and Noldus, L. (1990) Whitefly-Plant Relationships: Behavioural and Ecological Aspects. *Whiteflies: their bionomics, pest status and management,* **47,** 49.

Westcott, C. and Horst, R. K. (2008) *Westcott's Plant Disease Handbook*. Berlin ;: Springer.

Wintermantel, W. M., Jenkins Hladky, L. L., Fashing, P., Ando, K. and McCreight, J. D. (2019) First Report of Cucurbit Chlorotic Yellows Virus Infecting Melon in the New World. *Plant Disease,* **103,** 778-778.

Wolfenbarger, D. O. (1946) Dispersion of Small Organisms. Distance Dispersion Rates of Bacteria, Spores, Seeds, Pollen, and Insects; Incidence Rates of Diseases and Injuries. *The American Midland Naturalist,* **35,** 1-152.

Xu, J., De Barro, P. and Liu, S. (2010) Reproductive Incompatibility among Genetic Groups of Bemisia Tabaci Supports the Proposition That the Whitefly Is a Cryptic Species Complex. *Bulletin of entomological research,* **100,** 359-366.

Zhang, X.-m., Yang, N.-w., Wan, F.-h. and Lövei, G. L. (2014) Density and Seasonal Dynamics of Bemisia Tabaci (Gennadius) Mediterranean on Common Crops and Weeds around Cotton Fields in Northern China. *Journal of Integrative Agriculture,* **13,** 2211-2220.

Zhao, K., Liu, S.-S., Wang, X.-W., Yang, J.-G. and Pan, L.-L. (2022) Manipulation of Whitefly Behavior by Plant Viruses. *Microorganisms,* **10,** 2410.

Table 3.1. Select Global Variable values used in the WTVC Model.

Variable Name	Variable Value	Citation(s)
Fecundity [stage 0]	0	-
Fecundity [stage 1]	0	-
Fecundity [stage 2]	23.1	Pascual & Callejas, 2004,
Survival [stage 0]	0.895	Ahmed, 2014
Survival [stage 1]	1	Ahmed, 2014
Survival [stage 2]	0.3	-
Percent Stochasticity [fecundity]	20	-
Percent Stochasticity [survival]	20	-
Carrying Capacity [per-hexagon]	75	Zhang, 2014

Table. 3.2. Summary of total infection saturation values for experimental field data from Codod, 2022 and corresponding simulations.

Year	Type	Collection Date	Percentile Group	($\bar{x}$) Total Infection Incidence	($\bar{x}$) Percent of Maximum Infection Incidence[b]
2020	Field	Penultimate	-	10,376[a]	34.6%
	Simulation	Penultimate	0 – 10%	17,810	59.4%
			10 – 50%	18,562	61.9%
			50 – 90%	19,327	64.4%
			90 – 100%	20,112	67.0%
	Field	Ultimate	-	19,184[a]	63.9%
	Simulation	Ultimate	0 – 10%	20,761	69.2%
			10 – 50%	22,746	75.82%
			50 – 90%	24,219	80.7%
			90 – 100%	25,121	83.7%

[a]Values calculated from Codod *et al.*, 2022, Figure 4C (25 and 32 DAP for penultimate and ultimate, respectively).
[b]Percentage calculated for a maximum infection incidence of 30,000 (300 quadrats at 100% disease incidence).

Table 3.3. Summary of Moran's I Analysis for experimental field data from Codod, 2022 and corresponding simulations.

Year	Type	Collection Date	($\bar{x}$) Moran's I	Min. Moran's I	Max. Moran's I	$\bar{x}$ |(ΔMoran's I)|	Min. |(ΔMoran's I)|	Max. |(ΔMoran's I)|
2020	Field	Penultimate	*0.12*[a]	-	-	-	-	-
	Simulation	Penultimate	*0.149*	0.111	0.183	0.030	0.002	0.063
	Field	Ultimate	*0.14*[a]	-	-	-	-	-
	Simulation	Ultimate	*0.160*	0.101	0.196	0.026	0.004	0.056

[a]Moran's I and p-value as reported in Codod, 2022.
|(ΔMoran's I)| was calculated as the absolute value of the difference between the field data Moran's I value and the simulation replicate Moran's I value. Italicized values indicate a Moran's I value with a significant p-value or the average of 20 Moran's I values with significant p-values, for field data and simulation data, respectively ($\alpha = 0.05$).

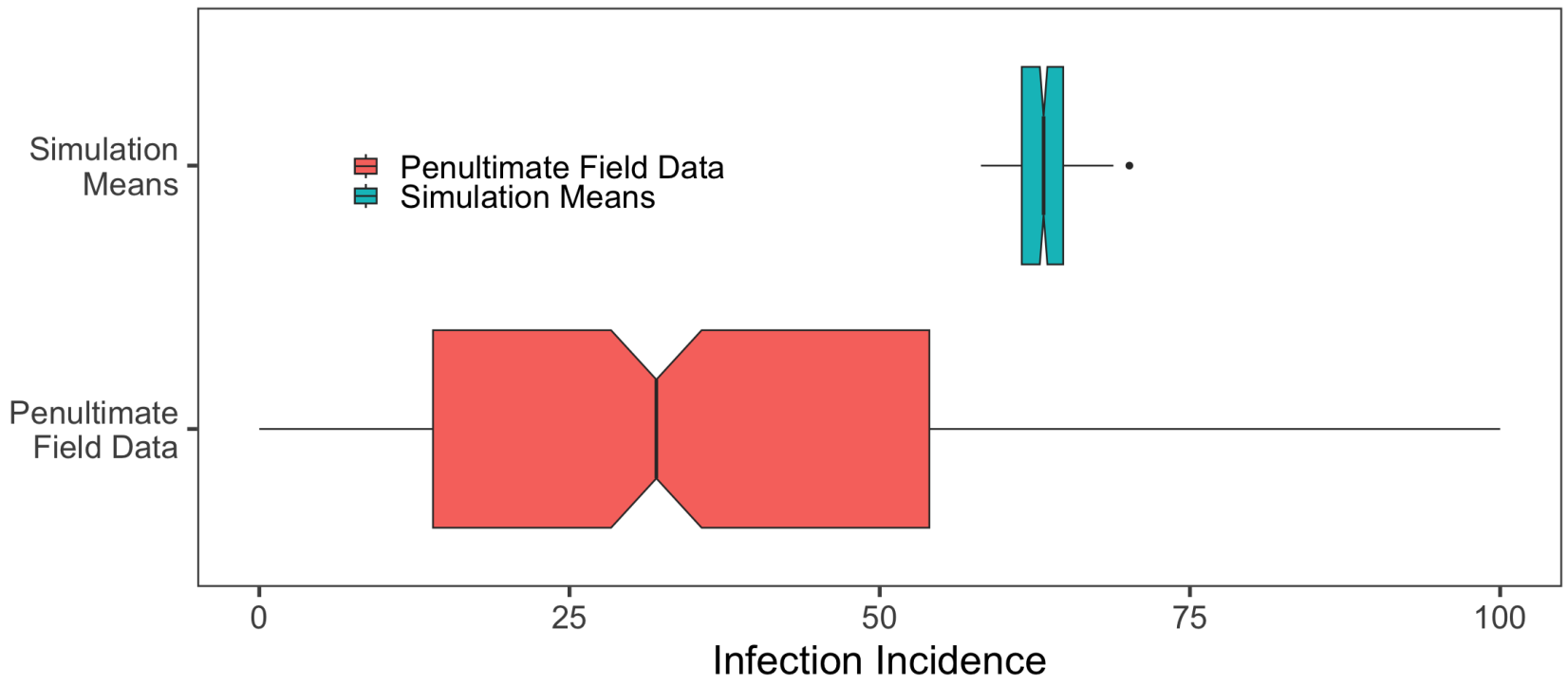

Figure 3.1. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of WTVC penultimate date field data and corresponding simulation means. Field data represents infection incidence values of each squash-planted field quadrat. Simulation means represent the mean infection incidence of all simulated quadrats in the virtual field for 300 replicates.

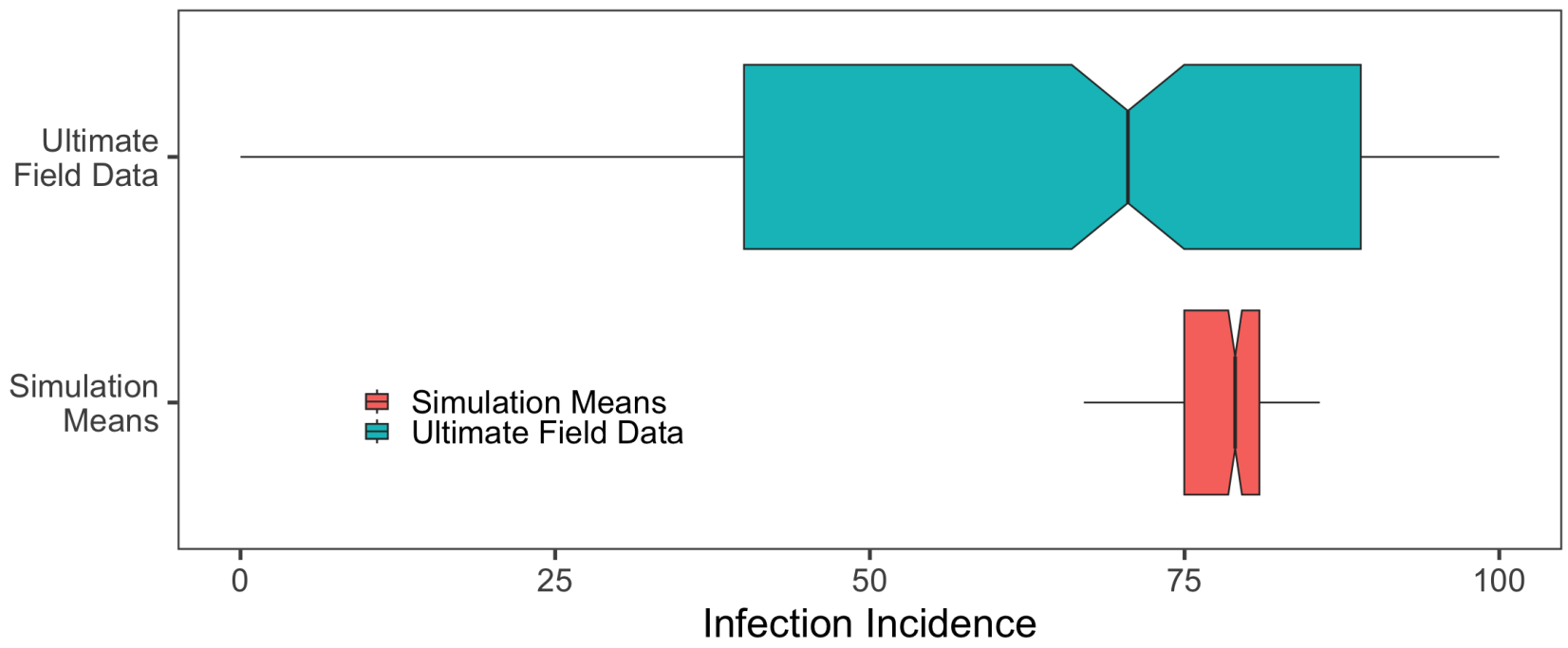

Figure 3.2. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of WTVC ultimate date field data and corresponding simulation means. Field data represents infection incidence values of each squash-planted field quadrat. Simulation means represent the mean infection incidence of all simulated quadrats in the virtual field for 300 replicates.

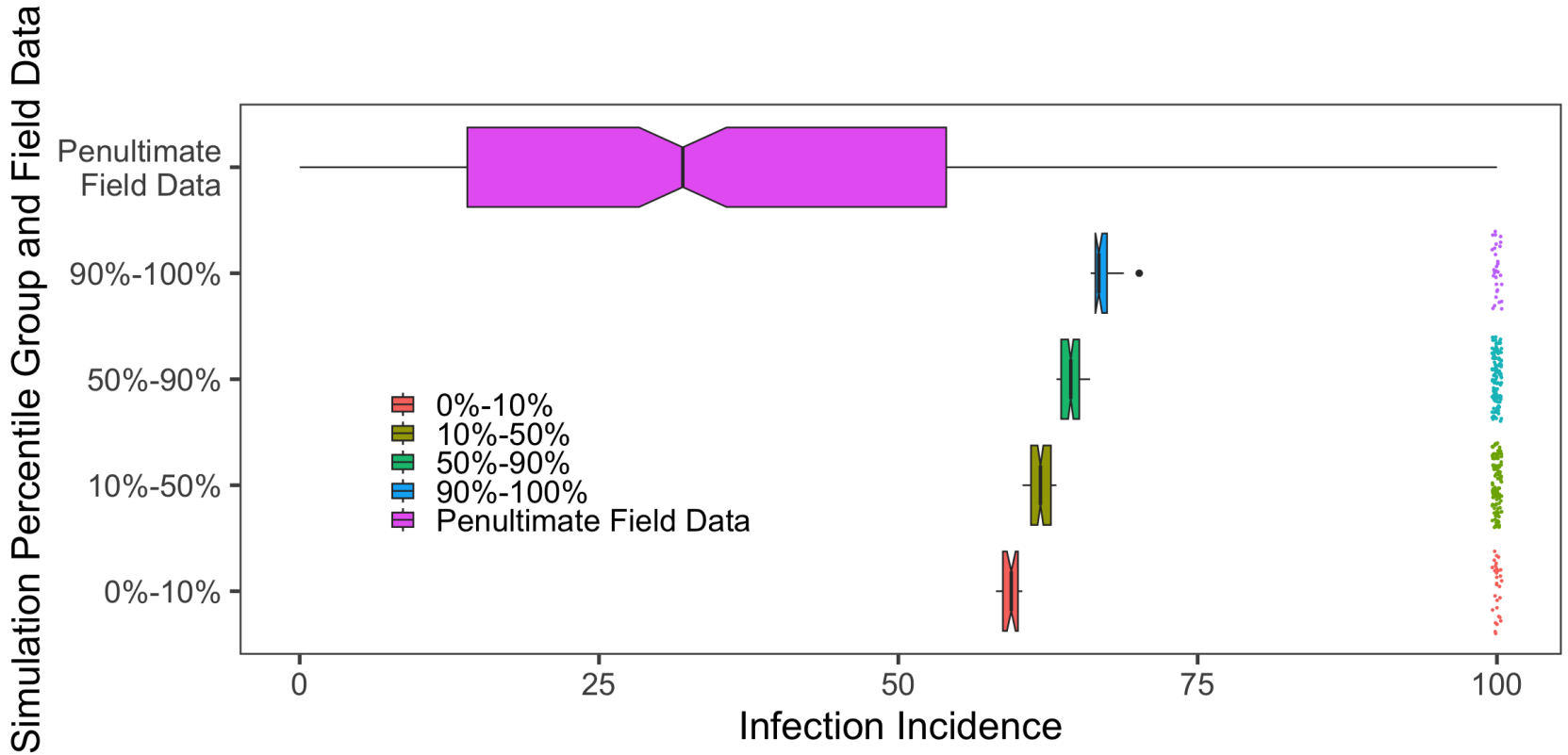

Figure 3.3. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of WTVC ultimate date field data and corresponding simulation means and maximums grouped by percentile. Field data represents infection incidence values of each squash-planted field quadrat. Simulation means represent the mean infection incidence of all simulated quadrats in for 300 replicates. Simulation percentile group boxplots represent the mean infection incidence for each replicate in the percentile group, while points represent maximum infection incidence values for each replicate in the percentile group. Percentile groups were determined based on mean field infection incidence.

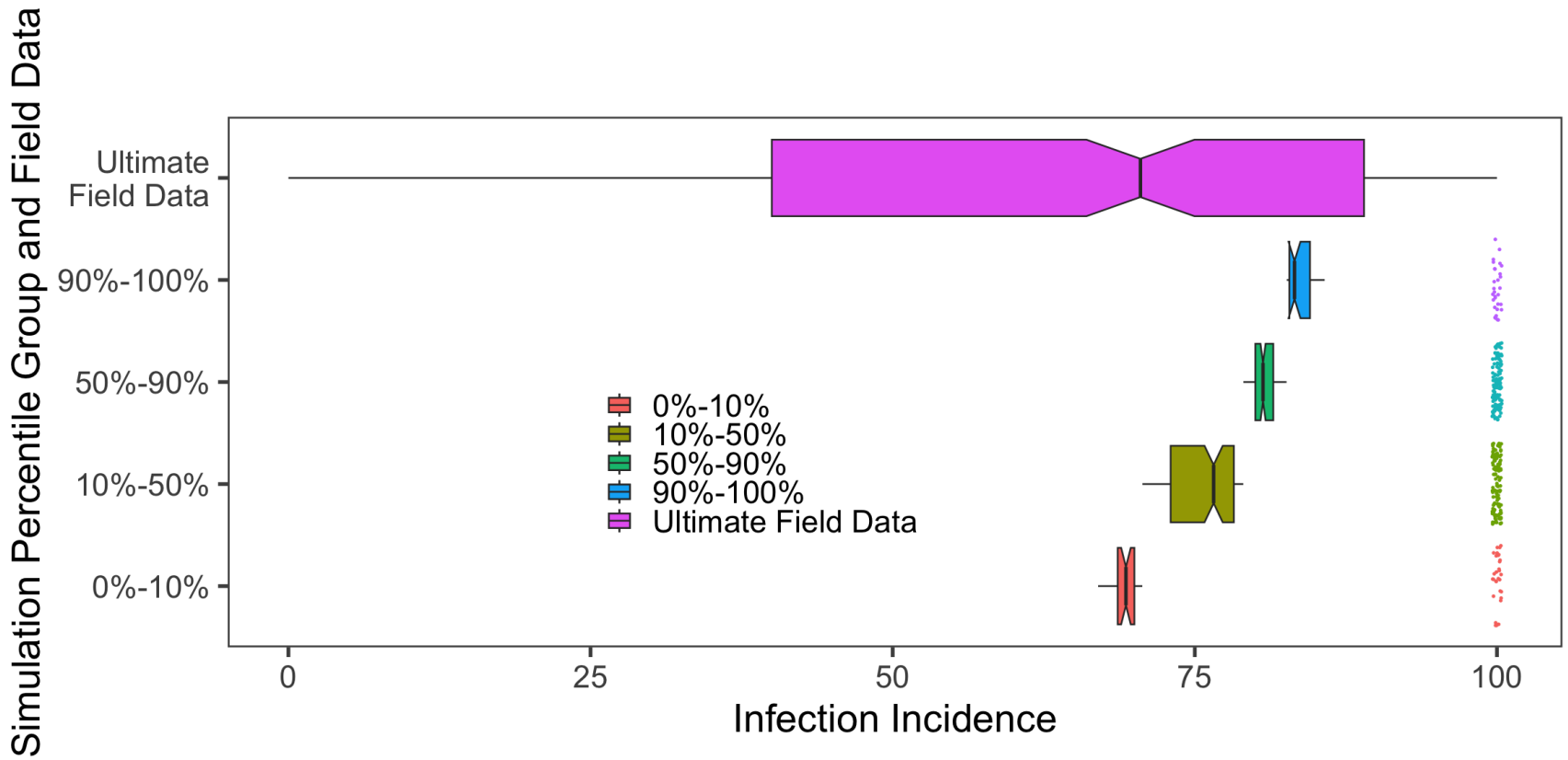

Figure 3.4. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of WTVC ultimate date field data and corresponding simulation means and maximums grouped by percentile. Field data represents infection incidence values of each squash-planted field quadrat. Simulation means represent the mean infection incidence of all simulated quadrats for 300 replicates. Simulation percentile group boxplots represent the mean infection incidence for each replicate in the percentile group, while points represent maximum infection incidence values for each replicate in the percentile group. Percentile groups were determined based on mean field infection incidence.

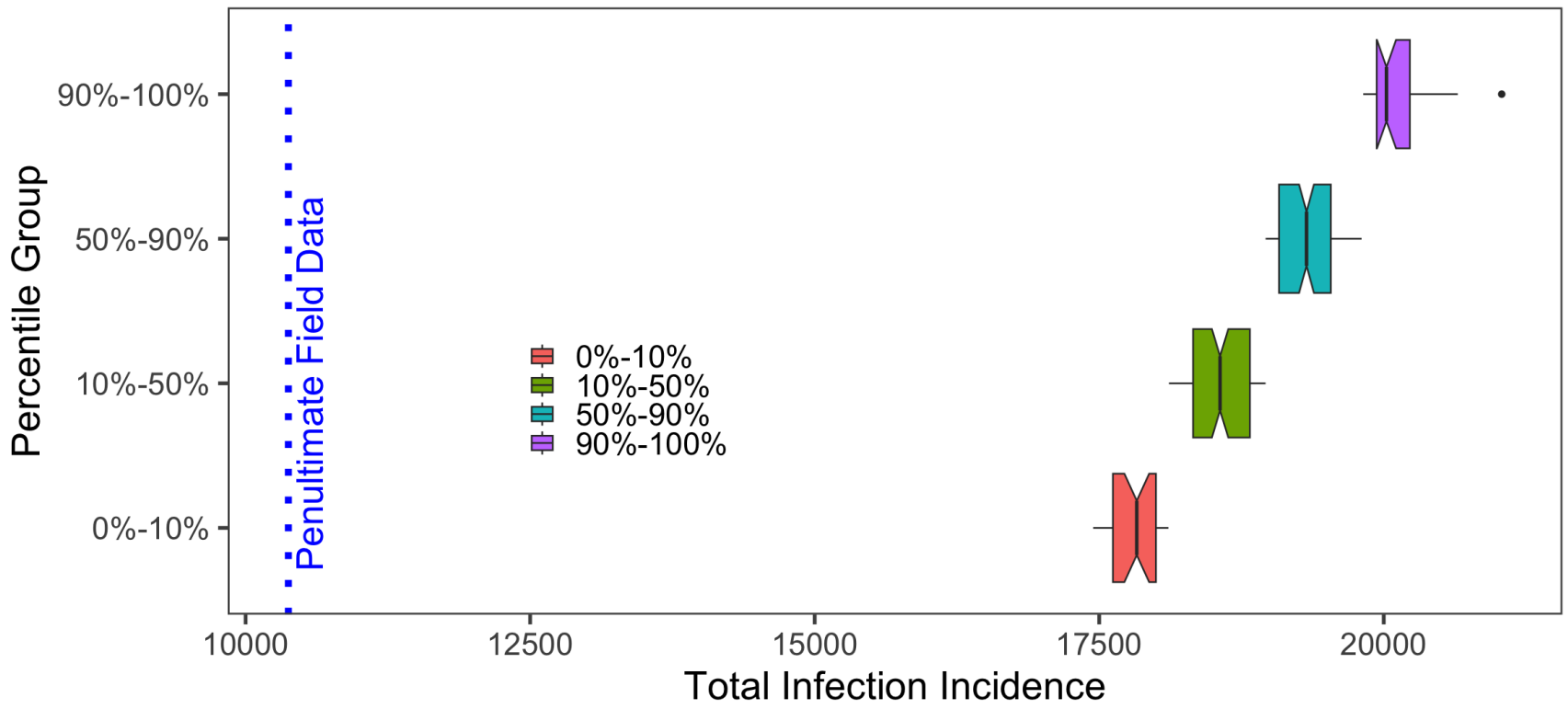

Figure 3.5. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of WTVC penultimate date field data and corresponding simulation total infection incidence grouped by percentile. Field data, represented by dotted blue line, represents total infection incidence of all squash-planted field quadrats. Simulation percentile group boxplots represent the total infection incidence for each replicate in the percentile group. The highest attainable value for total infection incidence is 30,000. Percentile groups were determined based on mean field infection incidence.

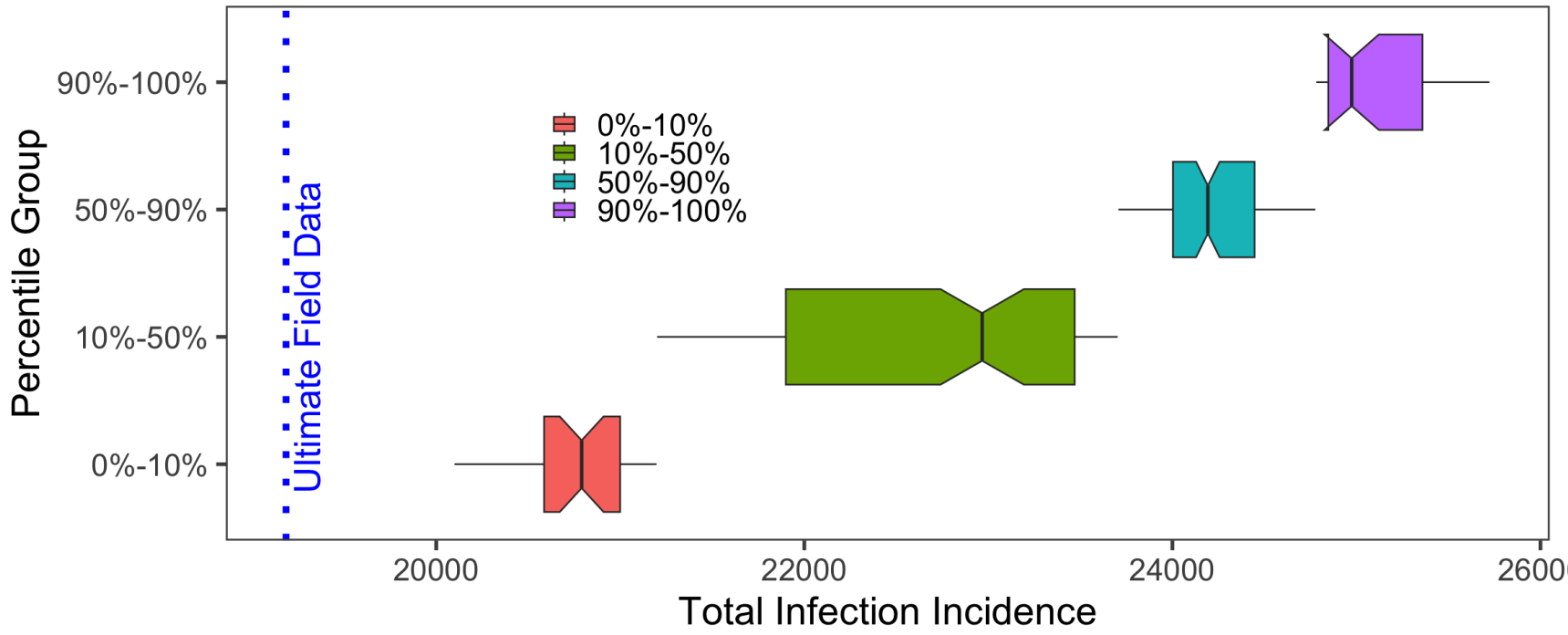

Figure 3.6. Comparison of disease incidence field data to simulation results at the same time point. Notched boxplot of WTVC penultimate date field data and corresponding simulation total infection incidence grouped by percentile. Field data, represented by dotted blue line, represents total infection incidence of all squash-planted field quadrats. Simulation percentile group boxplots represent the total infection incidence for each replicate in the percentile group. The highest attainable value for total infection incidence is 30,000. Percentile groups were determined based on mean field infection incidence.

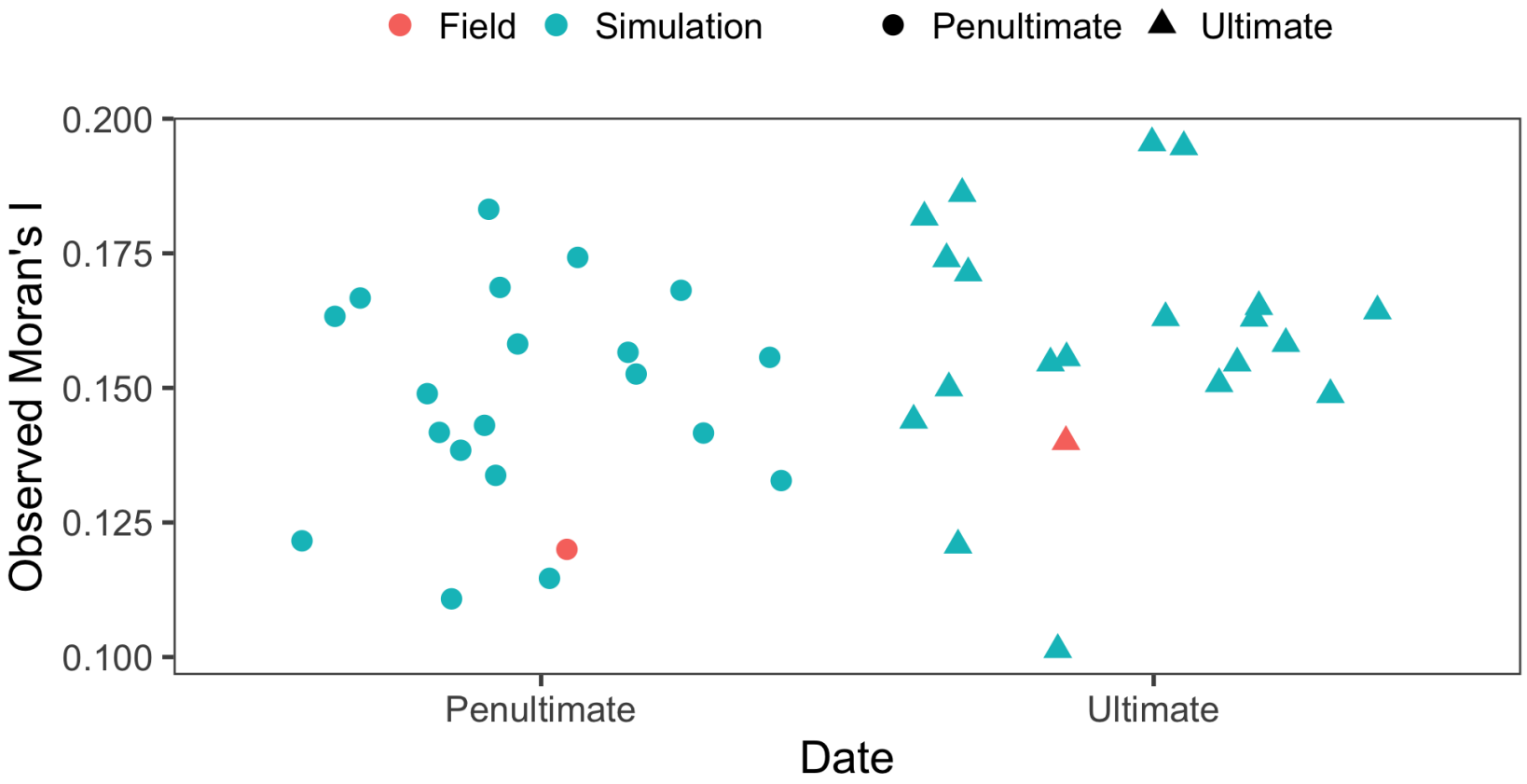

Figure 3.7. Degree of spatial aggregation for disease incidence field data and that of simulation results at the same time point. Comparison of Moran's I value for WTVC field data to Moran's I values of 20 randomly selected virtual fields from corresponding simulations. For penultimate and ultimate collection dates, 20 replicates each were selected from the percentile group most similar to field data. Percentile groups were determined based on mean field infection incidence and similarity to field data was determined based on field data mean.

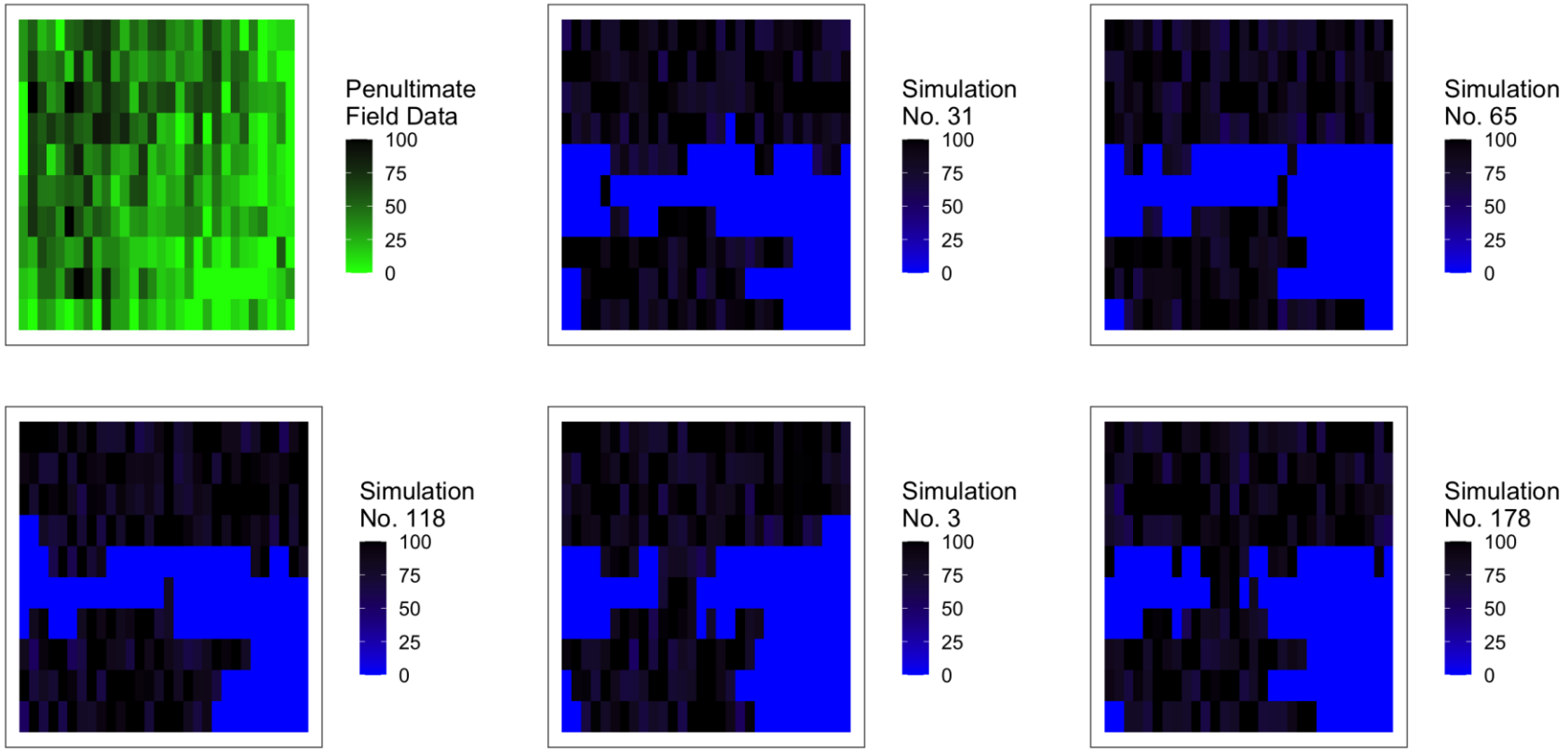

Figure 3.8. Heatmaps of field infection incidence for WTVC penultimate date field data and field infection incidence for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean.

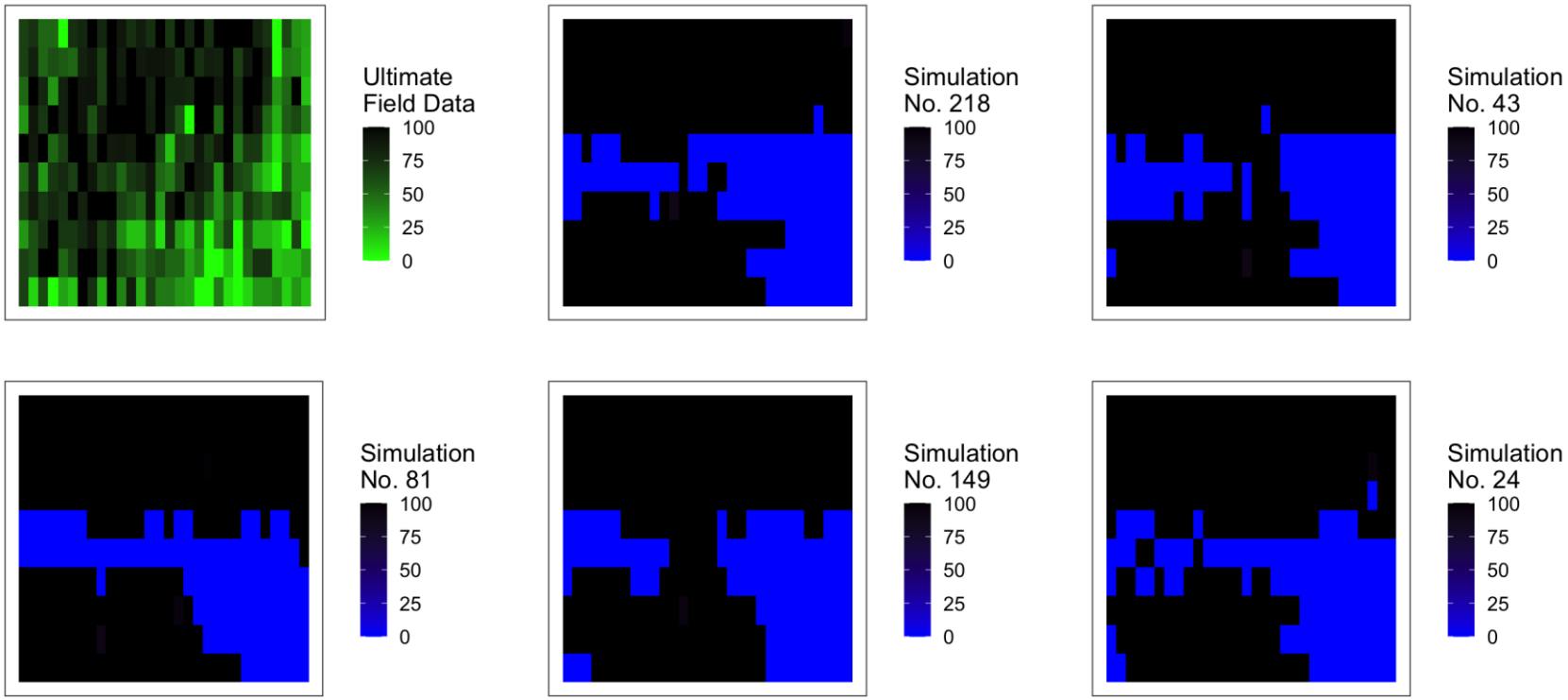

Figure 3.9. Heatmaps of field infection incidence for WTVC ultimate date field data and field infection incidence for five randomly selected replicates from corresponding simulation. Five replicates were randomly selected from a subset of 20 randomly selected simulation replicates from the percentile group most similar to field data. Percentile groups were determined based on mean field infection saturation and similarity to field data was determined based on field data mean.

CHAPTER 4

CONCLUSIONS

The vast majority of novel and exotic viral plant pathogens appear to be (re-)emerging due to anthropogenically-mediated introduction events (Anderson *et al.*, 2004, Gibbs *et al.*, 2010, Jones, 2009, Webster *et al.*, 2007). Because arthropods serve as vectors for most of these pathogens, arthropod invasions likely play a significant role in the management of emerging plant diseases. The simulation models described in the previous two chapters involve vectors that are cases in point. The suspected vector of BNRBV is a mite in the superfamily Eriophyoidea – a group of organisms that has been recognized as having a high propensity for invading new areas (Navia *et al.*, 2010). *Bemisia tabaci* is the vector of a virus complex that is threatening squash production in South Georgia and crops around the world. Two *Bemisia tabaci* cryptic species were part of a relatively recent, infamous world-wide biological invasion that displaced many less damaging indigenous whiteflies and led to the proliferation of agriculturally damaging plant diseases caused by begomoviruses, criniviruses, and ipomoviruses (Navas-Castillo *et al.*, 2011). The development of effective management practices for these recent plant diseases is hindered due to a lack of understanding how these arthropod-vectored pathosystems function. One way that simulation models can help plant pathologists with these diseases is to develop models based on what is known and suspected to evaluate the

contribution of environmental, behavioral and epidemiological variables in a complex and dynamic plant pathosystem (González-Domínguez *et al.*, 2023).

The models described in **Chapter 2** and **Chapter 3** present arthropod-vectored pathosystems that integrate vector demographics and dispersal behavior using a contemporary compartmental epidemiological model approach. Vector demography and behavior are essential components to understanding any arthropod-vectored plant disease (Fontenille *et al.*, 2020). Surprisingly little is known about BNRBV and its presumed vector, a *Calacarus* mite, and although comparatively more is known about *Bemisia tabaci,* its behavioral and dispersal ecology is only superficially established. Despite information deficiencies, my epidemiological models could capture known patterns with respect to field data over space and time. With more phenomenological information and directed empirical studies, these models could be refined and become more informative. Additional field studies on the spatial spread of both BNRBV and the WTVC would provide further points of comparison to evaluate the performance of SimpleMite and the WTVC Model, respectively.

The previous chapters present models for two very different pathosystems. BNRBV causes local infections on plant leaves while the viruses that compose the WTVC systemically infect their hosts. Severe BNRBV infections cause leaf abscission with unknown consequences for subsequent growing seasons, while severe infections of the WTVC cause plant death (Codod *et al.*, 2022, Robinson, 2013). The suspected vector of BNRBV disperses actively by crawling across leaf surfaces or passively on air currents, while *B. tabaci* is capable of sustained flights against a wind current of 4.0 cm/s (Byrne, 1999, Lindquist & Oldfield, 1996). Regardless of these important differences in

modes of dispersal, the same base model was used for both simulations, demonstrating that these models have potential use far beyond the two pathosystems through a straightforward change of parameter values to match similar a similar pathosystem (e.g. aphid vectored viruses). While there is a rich history of the use of predictive tools in plant pathology (Charaya *et al.*, 2021), most of the modern options available are not appropriate and complementary to the tools used by epidemiologists outside of the plant sciences. The development of modern epidemiological analyses has been hindered by a lack of available data (Cooke *et al.*, 2006) and availability of analytical tools that are currently used by epidemiologists in animal and human systems. There appears to be no such comparable and publicly available epidemiological spatially explicit simulation models for arthropod-vectored plant viruses other than those presented in my thesis. My hope is that these models will serve as templates for the development of more phytopathological SEMs and encourage other plant pathologists to learn, refine, and use the computationally more flexible and appropriate epidemiological analyses to understand the complex nature of arthropod vectored plant diseases.

LITERATURE CITED

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R. and Daszak, P. (2004) Emerging Infectious Diseases of Plants: Pathogen Pollution, Climate Change and Agrotechnology Drivers. *Trends in Ecology & Evolution,* **19,** 535-544.

Byrne, D. N. (1999) Migration and Dispersal by the Sweet Potato Whitefly, Bemisia Tabaci. *Agricultural and forest meteorology,* **97,** 309-316.

Charaya, M. U., Upadhyay, A., Bhati, H. P. and Kumar, A. (2021) Plant Disease Forecasting: Past Practices to Emerging Technologies. *Plant Disease: Management Strategies; Nehra, S., Ed.; Agrobios Research: Rajasthan, India***,** 1-30.

Codod, C. B., Severns, P. M., Sparks, A. N., Srinivasan, R., Kemerait, R. C. and Dutta, B. (2022) Characterization of the Spatial Distribution of the Whitefly-Transmitted Virus Complex in Yellow Squash Fields in Southern Georgia, USA. *Frontiers in Agronomy,* **4**.

Cooke, B. M., Jones, D. G. and Kaye, B. (2006) *The Epidemiology of Plant Diseases*. Springer.

Fontenille, D., Cruaud, A., Vial, L. and Garros, C. (2020) Understanding the Role of Arthropod Vectors in the Emergence and Spread of Plant, Animal and Human Diseases. A Chronicle of Epidemics Foretold in South of France. *Comptes Rendus. Biologies,* **343,** 311-344.

Gibbs, A. J., Fargette, D., García-Arenal, F. and Gibbs, M. J. (2010) Time – the Emerging Dimension of Plant Virus Studies. *Journal of General Virology,* **91,** 13-22.

González-Domínguez, E., Caffi, T., Rossi, V., Salotti, I. and Fedele, G. (2023) Plant Disease Models and Forecasting: Changes in Principles and Applications over the Last 50 Years. *Phytopathology®,* **113,** 678-693.

Jones, R. A. C. (2009) Plant Virus Emergence and Evolution: Origins, New Encounter Scenarios, Factors Driving Emergence, Effects of Changing World Conditions, and Prospects for Control. *Virus Research,* **141,** 113-130.

Lindquist, E. E. and Oldfield, G. N. (1996) Chapter 1.5 Evolution and Phylogeny 1.5.1 Evolution of Eriophyoid Mites in Relation to Their Host Plants. In: *Eriophyoid Mites: Their Biology, Natural Enemies and Control (World Crop Pests).* (Lindquist, E. E., Sabelis, M. W. and Bruin, J., eds.). Elsevier, pp. 277-300.

Navas-Castillo, J., Fiallo-Olivé, E. and Sánchez-Campos, S. (2011) Emerging Virus Diseases Transmitted by Whiteflies. *Annual Review of Phytopathology,* **49,** 219-248.

Navia, D., Ochoa, R., Welbourn, C. and Ferragut, F. (2010) Adventive Eriophyoid Mites: A Global Review of Their Impact, Pathways, Prevention and Challenges. *Experimental and Applied Acarology,* **51,** 225-255.

Robinson, T. S. (2013) Epidemiology of Blueberry Necrotic Ring Blotch Virus of Southern

Highbush Blueberry in Georgia. In: *Department of Plant Pathology.* University of Georgia, pp. 108.

Webster, C. G., Coutts, B. A., Jones, R. A. C., Jones, M. G. K. and Wylie, S. J. (2007) Virus Impact at the Interface of an Ancient Ecosystem and a Recent Agroecosystem: Studies on Three Legume-Infecting Potyviruses in the Southwest Australian Floristic Region. *Plant Pathology,* **56,** 729-742.

APPENDIX A

AN OVERVIEW OF HEXSIM AND RELEVANT METHODS FOR BIOLOGICAL, POPULATION, AND DISEASE MODELS

Mathematical models for biological scenarios. Mathematical models are heuristic tools used to represent biological systems of which the mechanisms are not fully understood (Antoniouk & Melnik, 2013). Modelling can solve some of the challenges inherent within the study of epidemiology. Within any discipline, causative conclusions are generated from manipulative experiments, but these experiments can be challenging to perform in the field of epidemiology because disease events often occur across large spatial scales that are difficult to replicate within experimental field settings (Ostfeld *et al.*, 2005). Furthermore, supposing the resources to replicate these events are available, it may be impossible to fully contain the pathogens to an experimental population, posing a danger to susceptible hosts outside of the experiment. These challenges impede the ability of epidemiologists to empirically test important hypotheses. However, well-designed and validated epidemiological models do provide a means to evaluate disease spread under varying scenarios, including some conditions that were known to exist, some that are thought to exist but with little empirical study to substantiate, and hypothetical situations that may exist in the future (*e.g.* climate change and the evolution of virulence). Model outcomes can be analyzed to determine if they agree with existing

hypotheses about disease patterns and used to make projections under scenarios of interest (Brauer, 2008, Jeger, 2000).

Compartmental models. Compartmental disease models are commonly used as the primary analysis for contemporary epidemiological models other than plant diseases. In these models, individuals in populations are divided into "compartments" that describe their status relative to the disease under study. The most common type of compartmental model is the SEIR model in which the compartments are Susceptible (S), Exposed (E), Infected (I), and Recovered (R). This model is used for systems where designated organisms are susceptible to disease and there is a latent period between exposure to the pathogen and infection leading to detectable symptom/ infection confirmation through serological tests. Organisms in the S compartment are at risk of encountering the pathogen and are **s**usceptible to the disease. Organisms in the E compartment have been **e**xposed to the pathogen but have not yet become infected or infectious. Organisms in the I compartment are **i**nfected and can spread the disease to other members in the population following a defined latent period. Organisms in the R compartment have **r**ecovered from the infection (or are **r**emoved from the population) and are no longer contagious. In the standard SEIR model, organisms in the R compartment cannot become reinfected, at least for a period, because the infection conferred immunity. In the context of viral plant diseases, plants often do not recover from infection. In this case, the R compartment could also represent plants that were rogued or have died.

In compartmental landscape event simulation models, every organism in the population exists within one compartment (a cell in a grid of cells) at any given time.

Because of this compartmentalization, which can represent actual space, spatially-explicit compartmental models (like those used in landscape ecology) are a natural complement to modeling disease spread over space and time. Organisms can move from one compartment to another, or across many compartments, the likelihood of which is determined by transfer rates, which in aggregate are used to describe a dispersal gradient (a.k.a. dispersal kernel, disease gradient, contact gradient). Rates of transfer between disease compartments are determined by the contact frequency between members of the population, the latent period of the disease, and the rate of recovery or death. For example, the rate of transfer for an organism within the population from the S to the E compartment is a function of the transmissibility of the disease and the probability of a random contact between a susceptible individual and an infected individual (Brauer, 2008, Tolles & Luong, 2020). These combined factors are represented by the effective contact rate, β. The transfer rate from the E to the I compartment is dependent upon the latent period of the disease. The average number of daughter infections caused by a single infective individual (or single infection in the case of a polycyclic fungal foliar plant disease) is known as the basic reproductive number, or R_0. The transfer rate from the I to the R compartment is dependent on the rate of recovery from the infection, often represented as γ. Because the number of individuals within any one compartment is a function of time, these mathematical models were traditionally formulated as a series of differential equations with respect to time (Brauer, 2008). Though the SEIR model is commonly used, there are many variations of disease compartmental models (e.g. SIR, SI, SEI) that are adapted to fit a range of disease scenarios and account for human interventions.

Although compartmental disease models are often used for animal species in which members of the population spread the disease upon encountering one another, this framework must be conceptualized differently for invertebrate-vectored plant diseases. The first difference to consider is that the host population is not moving – planted crops are sessile. Second, invertebrates are the vectors that transmit pathogens between plants, which have behaviors and their own demographic population vital rates, as opposed to wind or plant-plant contact. Some epidemiologists have approached this challenge by using a dual compartmental model: an SEI model to represent transmission from the vectors and an SEIR model to represent the disease dynamics of the host plant (Jeger *et al.*, 1998). Another difference between the standard compartmental disease model and invertebrate vectored plant diseases is that vectors such as whiteflies, may be biased in selecting host plants to settle on, feed from and oviposit on, potentially resulting in positive or negative spatial autocorrelation patterns of disease that might not occur in wind vectored plant diseases. In invertebrate vectored plant diseases, space may play an important part in plant disease dynamics that is distinctly different from that of a passively dispersed plant pathogen.

Representing demography in my models through Leslie Matrices. To make decisions for the future it is often necessary to use data from the past. Potential future scenarios constructed of assumptions from previously collected data are called projections. Projections are not necessarily the same as predictions or forecasts, which try to accurately portray a real future. Instead, projections are constructed of if-then statements that apply assumptions about a previous state of a system to determine the

potential future state of that system, given that the assumptions are true (Keyfitz & Caswell, 2005). Projections can be predictions if the assumptions and data used to construct them are realistic. Real-life biological systems are almost always more complicated than can be represented in projection models. While the "simplifying assumptions" made by those who study populations reduce their exactitude, they allow them to be more easily replicated mathematically (Vandermeer & Goldberg, 2013). These mathematical representations distill information about the population demographic parameters and estimate their possible population growth responses to variation in life history traits (Kajin *et al.*, 2012). While these methods are often applied by conservation biologists to manage rare focal species, they have also been used in population ecology for decades to project the growth of any stage- or age-structured population (Caswell and Goldberg, 2013).

Most often, projections attempt to represent and generally summarize the size and structure of imagined populations. In mathematical biology, population projection models (PPMs) are tools used to project the population dynamics of a particular organism. PPMs use information about the vital rates and life histories of organisms to determine population dynamics. One of the most widely used methods in population ecology is the Leslie matrix (Nichols *et al.*, 1992). Leslie matrices are a specific form of population projection model used to project a population of stage/age-structured organisms as it progresses through multiple generations using matrix algebra. They are often used for organisms that have life cycles divided into distinct life stages (and occasionally age-structured populations). Organisms are categorized into different, demographically important, developmental stages (e.g. egg, juvenile and adult) and transition rates

represent the likelihood that an organism moves from one stage to the next taking into account biological rules (*e.g.* eggs cannot reproduce, individuals in all stages can die or survive to transition to the next developmental stage, organisms cannot become younger). Ultimately, population growth is determined by the rates of mortality and the fecundity of each life stage. A key assumption of this model is that members of the population within the same life stage have similar demographic parameters and that the initial conditions (from which the transition rate information is collected) remains the same for any future projections (Kajin et al., 2012). However, there are methods to induce variation into demographic projections through elasticities or introducing stochastic parameter value selection from a pool of values given parameter ranges (which is computationally more advanced).

HexSim and the HexSimPLE template. HexSim is a spatially-explicit modelling environment developed for the simulation of ecological scenarios (Schumaker & Brookes, 2018). Its graphical user interface enables users to create a simulation for their event sequence of choice without having to write code, lowering the barrier to entry for its use. The interface allows the user to customize three main components of the simulation: the populations involved, the space in which the simulation takes place, and the events that occur in the simulation. There are extensive options provided to users to describe the structure, resource needs, and behavior of their population(s) of interest. Additionally, individual population members can acquire new **accumulated traits** as a result of their life history.

HexSim requires the input of spatial data in the form of Hexmaps and barrier maps. **Hexmaps** are grids that define the space in which the populations can exist and can contain biologically relevant information such as the resource value of the environment to members of the population. **Barrier maps** determine linear boundaries on Hexmaps that population members cannot cross. Both kinds of maps are inherently time series maps but can be used as static maps by only assigning a map to the first time-step of the series. A **time step** is one complete instance of the **event sequence** which can be customized to match the biological system being modeled.

The event sequence allows for population members to interact with and within the space. The event sequence determines what life history events the members of the population experience. Any action performed by or upon population members is described by the events in the event sequence. Survival, reproduction, population member interactions, aging, feeding, and migration are dictated by the event sequence. The event sequence inherently determines the relationship between one time step and real time. For example, if the event sequence includes the relevant life history events that occur in two weeks of the life of the organism of interest, then one time step in the simulation is equivalent to two weeks of real time. If the event sequence includes the relevant life history events that occur in one day of the life of the organism of interest, then the time step would only represent one day of real time. The user determines how many time steps occur in one run of the simulation, which is to say, how much real time one run of the simulation represents.

My simulations use a HexSim template called **HexSimPLE**. This template was created for users that are particularly interested in studying metapopulation dynamics of

age- or stage-stratified population (Schumaker, 2023). HexSimPLE requires the input of demographic parameters for each life stage of the population under study and uses Leslie matrices to perform matrix algebra that determines the structure of the population. I created a modified version of the HexSim template that, in addition to being customized to the demographic parameters of the vector, involved interactions between vectors and hosts that result in the transmission of viral diseases. The focal points of these models are the consequences of vector population dynamics on disease outcomes, making this modified version of HexSimPLE ideal for the study.

Representing Space in HexSimPLE

Within spatially-explicit models, the user specifies meaningful dimensions that directly relate to the real world. The two-dimensional space within these is often broken up into a grid pattern. The user can specify, that each cell in the grid measures 1 meter by 1 meter, for example. This specificity allows the user to both represent the actual size of real objects or habitats in the model and obtain model outputs that provide information about the way that organisms behave or interact in two-dimensional space. HexSim differs from most spatially explicit models in that two-dimensional space is assembled of a framework of hexagons instead of rectangles or squares. This shape was chosen because the probability of traveling from the middle of a hexagon to any edge is more symmetrical than distributing from the center of a square (all other landscape event simulators) to the edge or a corner (much greater distance to the corner but assume the same probability, which is incorrect). Additionally, similar shapes can be made with hexagons (with a slightly ragged edge) as can be made with squares (a circle would leave gaps between the compartments which would be a spatial and probabilistic black hole). The additional

surface area achieved by using hexagonal cells allows for more detailed and refined movement estimates within the HexSim environment.

Multiple maps of the same space can be supplied to the simulation for different purposes. HexSimPLE requires the use of both a Habitat Map and a Matrices Map. The **Habitat Map** communicates resource quality by scoring each hexagon from 0.0 (non-habitat) to 1.0 (ideal habitat). This score determines both the carrying capacity of the cell and the vital rates of individuals existing within it, a topic discussed in the following section. This map can be a time series or a static map. In a host-pathogen model, adjusting the habitat quality could be used to represent the susceptibility, resistance, or relative attractiveness of the host to an invertebrate vector. The **Matrices Map** is constructed of a collection of **patches** made up of one or more hexagons. These patches can be contiguous or non-contiguous and the sizes of the patches can vary. A Leslie matrix is assigned to each patch in this Hexmap. Adjusting the boundaries of these patches allows the user to adjust the granularity of the simulated population, allowing a more detailed or general view of population dynamics. This map cannot be a time series.

In order for the HexSimPLE template to function, the aforementioned Hexmaps must contain entered parameter values. HexSimPLE technically also requires the user to supply four other maps, but they can optionally contain information. There are three required Hexmaps called **Regions Map**, **Stress Map Fecundity**, and **Stress Map Survival**. These Hexmaps contain patches of hexagons that should be identical in terms of size and location to those in the Matrices Map. There is one optionally meaningful barrier map called **Movement Barriers**. If the user does not wish to use these Hexmaps, the value of all hexagons contained within a patch can be set to 1, and all others set to

zero. If the user does not wish to introduce any barriers, the Movement Barriers map should not contain any barriers.

The **Regions Map** allows the user to introduce stochasticity into the environment. The landscape can be broken into patches of hexagons that will be assigned a different environmental quality value during the event sequence. This map cannot be a time series. Users that do not want to introduce multiple regions should start with a map identical to the **Matrices Map** and then assign to all non-zero hexagons a value of 1 (Schumaker, 2023). This results in a map with one region. **Stress Map Fecundity** and **Stress Map Survival** are collectively referred to as Stress Maps. These maps allow the user to introduce location-specific effects on fecundity and survival, either as separate or interacting effects. Patches of cells in this map are assigned a score from 0 (maximal stress) to 1 (no stress). These stress scores are used "to extract coefficients that are subsequently used to multiply the individual Leslie matrix survival and fecundity rates", respectively (Schumaker, 2023). In a viral plant disease model, stress maps could perhaps be used to represent the effects of an insecticide on the vector population, either temporarily or permanently. If the user does not want to introduce stress into the model, all patch stress scores should be set to 1. The Stress Maps can be time series. Because most of the Hexmaps used in this template are divided into patches, barriers created using the **Movement Barriers** map should not intersect these patches and should instead border the edges of the patch. Otherwise, "individuals who emigrate into a sub-divided patch will effectively cross the barrier when they become absorbed into the patch's population vector" (Schumaker, 2023).

Representing Populations in HexSimPLE

The primary purpose of the HexSimPLE routine is to obtain information about the population of interest, and the user is required to provide information about the demographic parameters of the focal organism. These parameters are collectively called **Global Variables**. The required parameters are Fecundity, Survival Rate, Habitat Exponent [fecundity], Habitat Exponent [survival], Percent Stochasticity [fecundity], Percent Stochasticity [survival], Dispersal Path Length (minimum), Dispersal Path Length (maximum), Per-Hexagon Carrying Capacity, Initial Population Size, and Data Sampling Delay.

Because HexSimPLE was constructed with age- or stage-stratified populations in mind, the **Fecundity** and **Survival** Global Variables require the input of survival and fecundity rates for three life stages of the organism: stage 0, stage 1, and stage 2. This approach is complementary to the biology of both plant pathogen vectors and fungal plant pathogens. For the models proposed here, these three stages correspond to the egg, juvenile, and adult stages, respectively. These rates, as well as the habitat quality of the patch supplied in the **Habitat Map** and the **Habitat Exponent** Global Variables, are combined to create patch-specific vital rates for fecundity and survival. The formula used to determine these rates is described in *Equation A.1*. Adjusting the Habitat Exponent for fecundity or survival allows the user to set "a habitat quality below which survival or fecundity rates begin dropping rapidly" (Schumaker, 2023). Incorporating the habitat quality scores allows the user to represent a scenario in which the environment impacts the reproduction and/or survival of the population members. The incorporation of the Habitat Exponent allows the user to control the degree to which these parameters respond to habitat quality.

The fecundity and survival rates used in individual Leslie matrices are also impacted by the **Percent Stochasticity** Global Variables. The user can enter a value for both **Percent Stochasticity [fecundity]** and **Percent Stochasticity [survival]**. This allows for the introduction of environmental stochasticity into the model, if desired. Each region specified in the Regions Map will be assigned a different environmental condition at each time step. The value of fecundity and survival for that region will vary randomly by the percentage specified in this Global Variable. For example, if the Percent Stochasticity [fecundity] Global Variable is set to 10 (= ±10% of the global variable value) and the Fecundity Global Variable is set to 20, the fecundity for any given region at any time step will be randomly selected from a range of 18 to 22. The relationship between the Percent Stochasticity [survival] and Survival Global Variables functions identically. One or both Percent Stochasticity Global Variables can be set to zero to "turn off" environmental stochasticity.

The **Dispersal Minimum [number of hexagons in a straight line]** and **Dispersal Maximum [number of hexagons in a straight line]** Global Variables determine the boundaries of how far individuals can move during the Conduct Movement Event Group (discussed in detail in the following section). These Global Variables specify the minimum and maximum length of their dispersal path. The **Carrying Capacity [per-hexagon]** Global Variable sets the maximum number of individuals that one hexagon of ideal habitat (score of 1.0 on the Habitat Map) can support. The **Initial Population Size [per matrix]** Global Variable sets the initial number of individuals added to each matrix in the Matrices Map. The **Begin Sampling [cumulative hexmaps]** Global Variable determines the time step at which HexSim begins constructing the

Cumulative Population Size, Cumulative Productivity, and Cumulative Dispersal Generated Hexmaps. These Generated Hexmaps will be discussed in further detail in the following section. These Generated Hexmaps require several time steps of data before they become meaningful, which is why they are not collected from the beginning of the simulation.

Equation A.1:

$$r_q = r_o [\, 1 - (1 - q)^{\eta} \,]$$

Equation for patch-specific vital rates, where r_q is the patch-specific survival or fecundity rate, r_o is the survival or fecundity rate provided as a Global Variable, q is the mean habitat quality of the patch obtained from the values in the Habitat Map, and η is the Habitat Exponent for either fecundity or survival. Adapted from Schumaker (2023).

The HexSimPLE Event Sequence and Outputs

There are five main **Event Groups** in the default HexSimPLE event sequence: Initialize Simulation, Conduct Movement, Perform Demography, Get Output Map Data, and Construct Output Maps. Events connect the individuals in the model to the space that they occupy and allow them to interact with the hexagons, taking into account the Global Variables and characteristics of the population and individuals of interest. Event Groups are collections of related events that determine any action that occurs within both the HexSim program and within the simulation. Event Groups can contain individual events and/or other Event Groups. Individual events can be *toggled on* or *off* by the user.

The **Initialize Simulation** Event Group only occurs on the first time step of the simulation. It introduces the population into the map. The **Conduct Movement** Event Group determines which individuals of each stage class will disperse, if any. Whether

individuals disperse is dependent upon the distribution of habitat quality across the map – individuals are attracted to patches with higher habitat quality. To represent populations in which certain stage classes do not disperse, the user can toggle off the individual events that correspond to the movement of those stage classes. The **Perform Demography** Event Group performs the matrix algebra that calculated the survival and fecundity rates of each patch in the Matrices Map. It contains two sub-Event Groups: **Set Vital Rates** and **Matrix Multiplication**. The Set Vital Rates Event Group calculates the survival and fecundity rates for each patch in the Matrices Map. These vital rates are dependent on both the Survival and Fecundity Global Variables and the many factors discussed previously in this section, e.g. habitat quality and environmental stochasticity. The Matrix Multiplication Event Group performs the matrix algebra that determines the population size for the following time step.

HexSim can summarize and present spatial data to visualize the results of a simulation run. This spatial data is called a **Generated Hexmap**. The **Get Output Map Data** Event Group contains events that compile data for the six default Generated Hexmaps produced by HexSimPLE: Population size, Cumulative population size, Population density, Cumulative population density, Cumulative productivity (births – deaths), and Cumulative productivity (emigration – immigration). Finally, the **Construct Output Maps** Event Group complies and presents the data of the aforementioned Generated Hexmaps. These maps are arranged in a time series that begin after the time step specified in the **Begin Sampling** Global Variable and can also be visualized within HexSim for individual time steps or viewed as a "movie".

LITERATURE CITED

Antoniouk, A. V. and Melnik, R. (2013) *Mathematics and Life Sciences*. Berlin: De Gruyter.

Brauer, F. (2008) Compartmental Models in Epidemiology. In: *Mathematical Epidemiology.* (Brauer, F., van den Driessche, P. and Wu, J., eds.). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 19-79.

Jeger, M. J. (2000) Theory and Plant Epidemiology. *Plant Pathology,* **49,** 651-658.

Jeger, M. J., van den Bosch, F., Madden, L. V. and Holt, J. (1998) A Model for Analysing Plant-Virus Transmission Characteristics and Epidemic Development. *Mathematical Medicine and Biology: A Journal of the IMA,* **15,** 1-18.

Kajin, M., Almeida, P. J., Vieira, M. V. and Cerqueira, R. (2012) The State of the Art of Population Projection Models: From the Leslie Matrix to Evolutionary Demography. *Oecologia Australis,* **16,** 13-22.

Keyfitz, N. and Caswell, H. (2005) *Applied Mathematical Demography*. New York, NY: Springer.

Nichols, J. D., Sauer, J. R., Pollock, K. H. and Hestbeck, J. B. (1992) Estimating Transition Probabilities for Stage-Based Population Projection Matrices Using Capture-Recapture Data. *Ecology,* **73,** 306-312.

Ostfeld, R. S., Glass, G. E. and Keesing, F. (2005) Spatial Epidemiology: An Emerging (or Re-Emerging) Discipline. *Trends in Ecology & Evolution,* **20,** 328-336.

Schumaker, N. (2023) Hexsimple.

Schumaker, N. H. and Brookes, A. (2018) Hexsim: A Modeling Environment for Ecology and Conservation. *Landscape Ecology*, **33,** 197-211.

Tolles, J. and Luong, T. (2020) Modeling Epidemics with Compartmental Models. *JAMA*, **323,** 2515-2516.

Vandermeer, J. H. and Goldberg, D. E. (2013) *Population Ecology : First Principles - Second Edition*. Princeton, UNITED STATES: Princeton University Press.

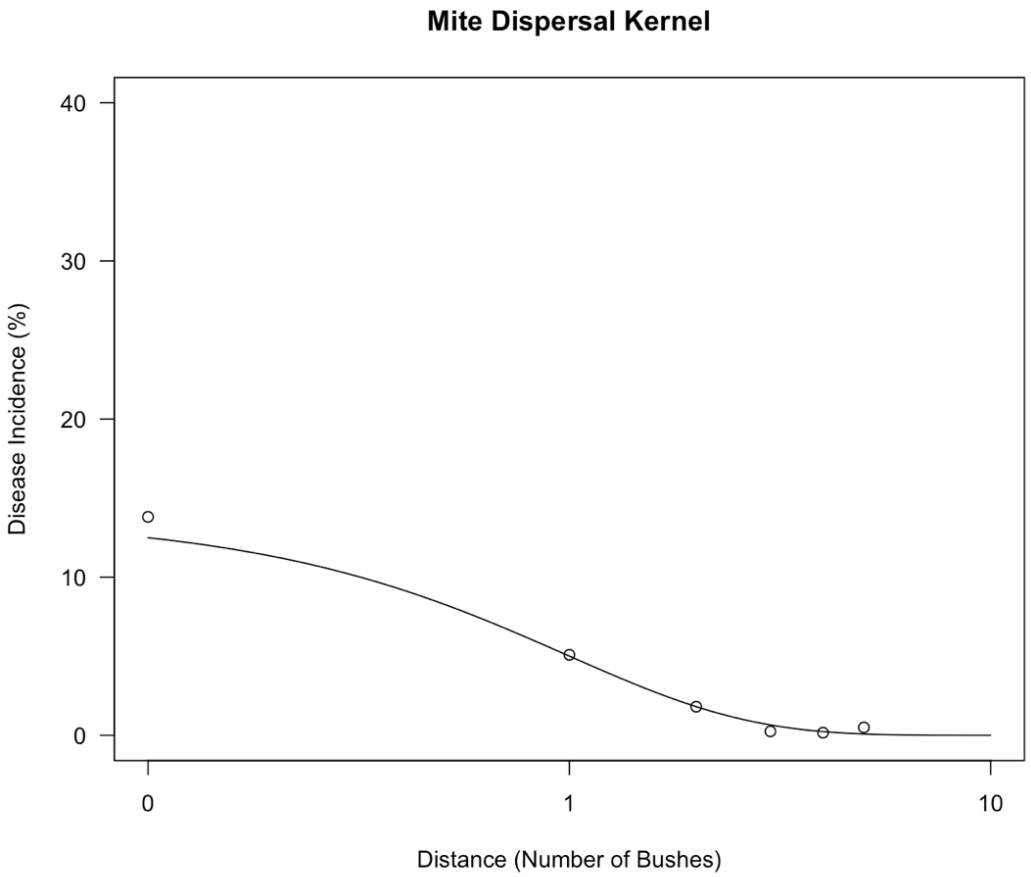

Figure A.1. Dispersal kernel constructed for dispersal of suspected vector of BNRBV based on per-plant diseases incidence patterns observed from Enigma in 2012. Kernel corresponds to equation:

$$y = 13.83^{-\frac{x}{0.985}}$$

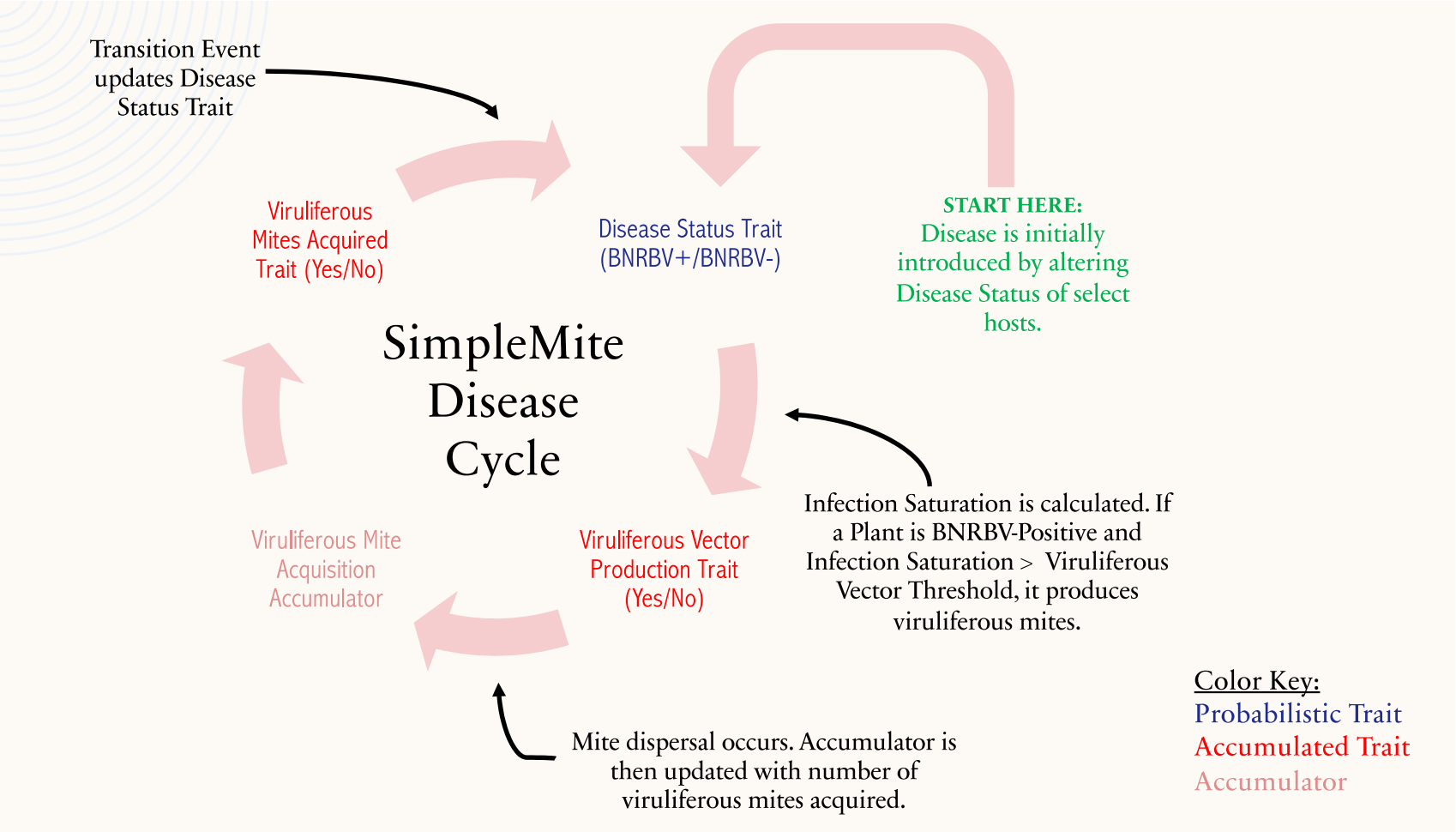

Figure A.2. A diagram of the simplified SimpleMite model disease cycle, describing how disease moves through the model.

APPENDIX B

EXAMPLE REAL-TIME YIELD LOSS HEATMAPS DERIVED FROM WTVC MODEL OUTPUTS

The WTVC model described in **Chapter 3** produces real-time yield loss maps in addition to infection incidence maps. A description of the calculations used to create these maps can be found in the Methods of **Chapter 3**. An example of a sequence of heatmaps derived from these real-time yield loss maps can be found below in **Figure B.1**.

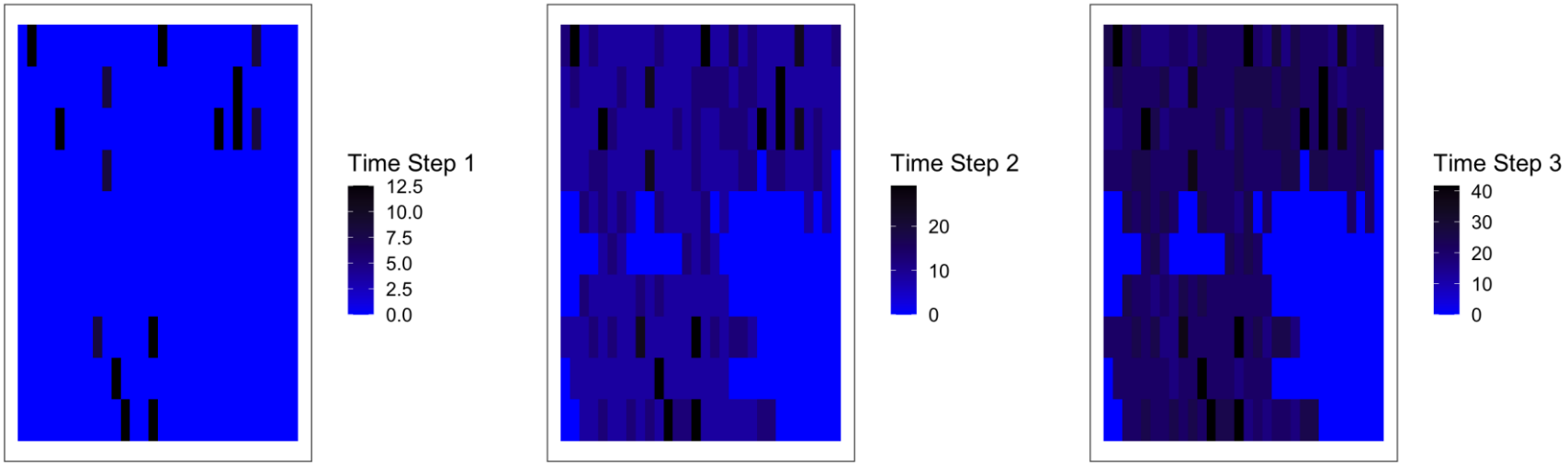

Figure B.1. Example sequence of real-time yield loss data produced by the WTVC Model described in **Chapter 3**.