2019

On The Application Of Computational Modeling To Complex Food Systems Issues

Serge William Wiltshire
University of Vermont

ON THE APPLICATION OF COMPUTATIONAL MODELING
TO COMPLEX FOOD SYSTEMS ISSUES

A Dissertation Presented

by

Serge W. Wiltshire

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Food Systems

May, 2019

Defense Date: March 18, 2019
Dissertation Examination Committee:

Meredith Niles, Ph.D., Advisor
David Novak, Ph.D., Chairperson
Christopher Koliba, Ph.D.
Margaret J. Eppstein, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

Transdisciplinary food systems research aims to merge insights from multiple fields, often revealing confounding, complex interactions. Computational modeling offers a means to discover patterns and formulate novel solutions to such systems-level problems. The best models serve as hubs—or boundary objects—which ground and unify a collaborative, iterative, and transdisciplinary process of stakeholder engagement. This dissertation demonstrates the application of agent-based modeling, network analytics, and evolutionary computational optimization to the pressing food systems problem areas of livestock epidemiology and global food security. It is comprised of a methodological introduction, an executive summary, three journal-article formatted chapters, and an overarching discussion section.

Chapter One employs an agent-based computer model (RUSH-PNBM v.0.8) developed to study the potential impact of the trend toward increased producer specialization on resilience to catastrophic epidemics within livestock production chains. In each run, an infection is introduced and may spread according to probabilities associated with the various modes of contact between hog producer, feed mill, and slaughter plant agents. Experimental data reveal that more-specialized systems are vulnerable to outbreaks at lower spatial densities, have more abrupt percolation transitions, and are characterized by less-predictable outcomes; suggesting that reworking network structures may represent a viable means to increase biosecurity.

Chapter Two uses a calibrated, spatially-explicit version of RUSH-PNBM (v.1.2) to model the hog production chains within three U.S. states. Key metrics are calculated after each run, some of which pertain to overall network structures, while others describe each actor's positionality within the network. A genetic programming algorithm is then employed to search for mathematical relationships between multiple individual indicators that effectively predict each node's vulnerability. This "meta-metric" approach could be applied to aid livestock epidemiologists in the targeting of biosecurity interventions and may also be useful to study a wide range of complex network phenomena.

Chapter Three focuses on food insecurity resulting from the projected gap between global food supply and demand over the coming decades. While no single solution has been identified, scholars suggest that investments into multiple interventions may stack together to solve the problem. However, formulating an effective plan of action requires knowledge about the level of change resulting from a given investment into each wedge, the time before that effect unfolds, the expected baseline change, and the maximum possible level of change. This chapter details an evolutionary-computational algorithm to optimize investment schedules according to the twin goals of maximizing global food security and minimizing cost. Future work will involve parameterizing the model through an expert informant advisory process to develop the existing framework into a practicable food policy decision-support tool.

CITATIONS

Chapter One of this dissertation has been published in the following form:

Wiltshire, S.W.. (2018). Using an agent-based model to evaluate the effect of producer specialization on the epidemiological resilience of livestock production networks. PloS one, 13(3), e0194013.

Chapter Two of this dissertation has been published in the following form:

Wiltshire, S.W., Zia, A., Koliba, C., Bucini, G., Clark, E., Merrill, S., Smith, J., and Moegenburg, S.. (2019). Network Meta-Metrics: Using Evolutionary Computation to Identify Effective Indicators of Epidemiological Vulnerability in a Livestock Production System Model. Journal of Artificial Societies and Social Simulation, 22(2), 10.18564/jasss.3991.

ACKNOWLEDGEMENTS

Langevin have also been important influences and contributors to this project, with the latter two serving as co-authors on a preliminary version of Chapter Three.

Finally, none of this would have been possible without the ongoing support of my family and friends, who have put up with me, even when I was cranky from sitting behind a computer debugging code for hours on end. Sarah O Donnell, Karen Neeson, William Wiltshire (II and III), Gillian McKay, and others—including Bruce Willis the cat—thank you!

TABLE OF CONTENTS

Page

xi

LIST OF TABLES

LIST OF FIGURES

xvii

xix

INTRODUCTION

1.1 Problem Statement and Overview

Given the complex, interconnected nature of food production, distribution, and consumption, the shortcomings of traditional methodological approaches—i.e. those rooted within single academic disciplines—have become increasingly evident. For example, while statistical survey research, interviews and other qualitative methods, supply chain and network modeling, public policy analysis, and other methodologies can provide valuable clues, they often struggle to capture overarching patterns that are not immediately evident through these field-specific analytical techniques. Recent transdisciplinary food systems studies have shown that, when diverse research efforts are integrated under a single framework, the sum can be bigger than its parts (Deppisch and Hasibovic 2013). However, even when guided by systems thinking (Meadows 2008), amalgamating research findings across multiple fields often quickly becomes a "big data" problem, with complex, non-linear interaction effects making it difficult or impossible to simply intuit high-level patterns.

Computational modeling and data science offer a unique opportunity to parse this complexity and discover patterns that lie at the nexus of disciplinary boundaries, leading to new intuitions about how to address systems-level problems. Often the goal is to create a Digital Decision Support Systems (DDSS) that can help to inform and guide policymaking and other strategic decisions. The best models integrate existing findings from key disciplinary fields, available datasets compiled by governmental and

other large-scale agencies, as well as an iterative and collaborative process of original

qualitative and/or quantitative research to solicit expert stakeholder input and advice

(Alavi and Leidner 2001). The diverse stakeholder group involved in parameterizing

and ground-truthing this kind of model throughout its development can also serve as an

outlet to disseminate model results to a non-academic audience for maximum possible

real-world effect.

This dissertation includes three chapters, each of which provides an example of

the ways in which complex systems computational methods—including agent-based

modeling, computational optimization, and supply chain network analytics—can be

applied to the critical food systems problem areas of livestock disease prevention

(Chapters 1 and 2) and global food security (Chapter 3). The introduction section

provides an overview of previous research in the areas of food systems and complex

systems scholarship. Following that is an executive summary section focused on

research motivations and key takeaways from this dissertation. After the three chapters

are presented, a summary section reflects on lessons learned through the

transdisciplinary research process (specifically as pertains to the application of

computer modeling), suggests directions for future research, and highlights key policy

implications.

1.2 Food Systems Research

Food systems is a relatively novel field of study that has gained considerable

traction in recent years. It is rooted in an inter/trans-disciplinary ethos which aims to

bridge across multiple traditional avenues of academic study to inform a holistic understanding that's bigger than the sum of its parts. While they share similar motivations, a distinction may be drawn between the interdisciplinary and transdisciplinary approach. Interdisciplinary research involves a shared central core set of research questions, with scholars in various disciplines contributing to the overall project according to their own areas of expertise. In contrast, transdisciplinary research establishes a collaborative research agenda from the outset that is explicitly aimed at discovering new understandings that lie between traditional disciplinary bounds (Hammond and Dubé 2012, Méndez et al. 2013). Along with inter/trans-disciplinarity, a coequal core tenet of food systems research is that it should reach beyond the academy to inform and drive real-world positive change, for example by communicating key findings in clear language that can be understood by practitioners, rather than focusing solely on academic publications.

Breaking down the title itself, "food systems" is first and foremost focused on the multiple and varied implications of the interactions between human society and the food we eat. However, rather than focusing on just one component of food as traditional disciplines have done (e.g. nutrition, food science, agronomy), food systems works across multiple interrelated domains including agricultural production, ecological systems, pests and disease, labor justice, economics, distribution chains, consumption choices, preparation practices, human health, and waste processing.

Also inherent in the food systems approach is the concept of systems thinking. Viewing the world through a systems lens requires scholars to transcend the simple cause-and-effect reasoning that is so common in academia. Instead of, for example, chasing *p*-values of bivariate correlations, systems thinkers must accept complexity and wrestle with the interactions between the multiple and varied moving parts that make up our food systems, accounting for factors such as feedback loops, nonlinearities, time delays, leverage points, and unintended consequences (Meadows 2008).

The goal of food systems research is not to precisely predict outcomes, but instead to embrace the inherent complexity associated with our field of study. As food systems scholars, we should be up front about the limitations of our predictions and proposed solutions. However, we should still recognize the potential of our research to inform valuable new intuitions about how to tackle complex challenges. In the words of pioneering systems thinker Donella Meadows (2001):

> The future can't be predicted, but it can be envisioned and brought lovingly into being. Systems can't be controlled, but they can be designed and redesigned. We can't surge forward with certainty into a world of no surprises, but we can expect surprises and learn from them and even profit from them. We can't impose our will upon a system. We can listen to what the system tells us, and discover how its properties and our values can work together to bring forth something much better.

1.2.1 Development of Food Systems Frameworks

Given the diverse disciplines and components of food systems, as well as the varying scales across which it operates, many efforts have been made to formalize the complex set of interactions between key food systems actors and processes. Sobal et al. (1998) were among the first to propose an integrated conceptual model of the overall food system. The authors leverage existing academic understandings of individual disciplinary sub-systems, coupled with new qualitative research, to inform a synthesized, transdisciplinary framework linking production, consumption, nutrition, geophysical resources, social structures, policymaking, and other factors (Figure 1). Through such an integrated framework, the authors encourage future food systems researchers to anticipate how their own findings may resonate through the system and affect outcomes across interconnected domains.

THE FOOD AND NUTRITION SYSTEM
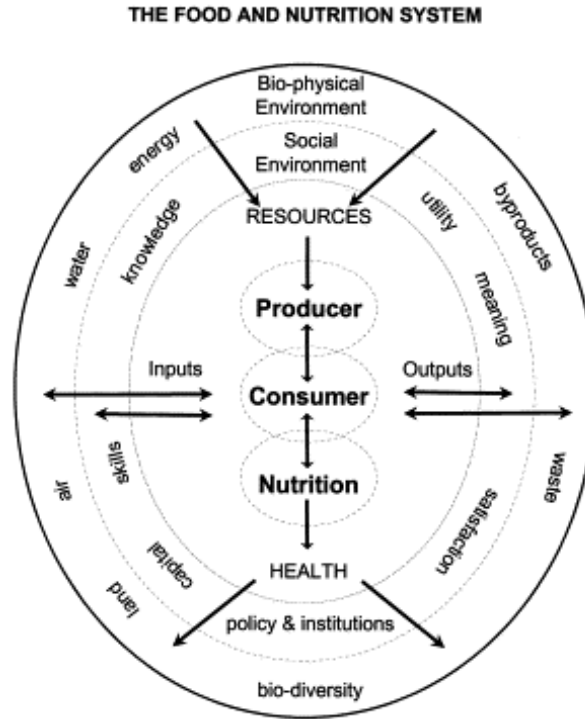
Figure 1: Food system interactions and feedbacks proposed by Sobal et al. (1998)

Ericksen (2008) follows up with a framework that delves deeper into the complex interactions between food systems structures, global environmental change, and resultant societal outcomes; including food security, ecosystem services, and social welfare (Figure 2). The authors aim to bridge findings from the natural sciences and social sciences to analyze how factors previously conceived as disparate may interact through the multi-scale interventions, feedback loops, and tradeoffs described by systems theorists. While they recognize that their model remains theoretical, they suggest that additional empirical research may one day flesh it out such that it can be used to guide real-world decision-making.

6

a

Environmental feedbacks
e.g. water quality, GHGs

GEC DRIVERS
Changes in:
Land cover & soils, Atmospheric
Comp., Climate variability & means,
Water availability & quality,
Nutrient availability & cycling,
Biodiversity, Sea currents
& salinity, Sea level

'Natural' DRIVERS
e.g. Volcanoes
Solar cycles

Food System ACTIVITIES
Producing
Processing & Packaging
Distributing & Retailing
Consuming

DRIVERS' Interactions

Food System OUTCOMES
Contributing to: Food Security, Environmental
Security, and other Societal Interests

Food Utilisation
Food Availability
Social Welfare
Food Access
Environ Capital

Socioeconomic DRIVERS
Changes in:
Demographics, Economics,
Socio-political context,
Cultural context
Science & Technology

Socioeconomic feedbacks
e.g. livelihoods, consumer advocacy

b

Food System ACTIVITIES
Producing food: *natural resources, inputs, technology, …*
Processing & packaging food: *raw materials, standards, storage life, …*
Distributing & retailing food: *transport, marketing, advertising,…*
Consuming food: *ac□uisition, preparation, sociali□ing,…*

Food System OUTCOMES Contributing to:

Social Welfare
• Income
• Employment
• Wealth
• Social & political capital
• Human capital

Food Security

FOOD UTILISATION
• *Nutritional Value*
• *Social Value*
• *Food Safety*

FOOD ACCESS
• *Affordability*
• *Allocation*
• *Preference*

FOOD AVAILABILITY
• *Production*
• *Distribution*
• *Exchange*

Environmental Security / Natural Capital
• Ecosystems stocks, flows
• Ecosystem services
• Access to natural capital

Figure 2: Theoretical framework showing major drivers of food systems outcomes (a), and primary food systems actors and activities (b), from Ericksen (2008).

As the field has grown, a more recent trend in theoretical food systems research aims to compare and contrast the multiple existent frameworks that have been proposed under the food systems umbrella. Yu at al. (2012) distinguish between several divergent perspectives through which scholars have traditionally addressed food security, for example crop yields (especially with respect to climate change), versus food distribution and allocation. The authors categorize these research approaches into spatial levels (field, regional, and global), as well as their specific research domains. They go on to synthesize these studies and develop an integrated framework through which food security policy can be evaluated, concluding that data from long-term *in situ* field experiments, regional monitoring and simulation, and global scenario assessments must be joined to inform effective solutions.

Similarly, Foran et al. (2014) use a "conceptual triangulation" approach to assess four different frameworks, focusing on the causes of food insecurity from the perspective of either agroecology, agricultural innovation, social-ecological systems, or political ecology. The authors identify divergent findings in some regards, but synergies in others, leading them to propose public policy interventions that can help to address food security from a systems perspective.

Ruben et al. (2019) also attempt to operationalize efforts to conceptualize a food systems framework toward the goal of improving nutrition outcomes. The authors present the problem through the so-called "3I approach," which considers (a)

intersections with other systems, (b) interactions within the food system, and (c) incentives for future innovations. They analyze various propositions aimed at improving nutrition, finding a diversity of proposed solutions, each largely situated within disciplinary bounds. The authors call for an integrated approach that accounts for dynamic feedbacks between actions at different food systems levels, more research toward identifying effective leverage points and external drivers of system dynamics, and the identification of tradeoffs between proposed interventions. By strategically considering how interventions in seemingly-disparate parts of the food system may impinge upon nutrition outcomes, it becomes possible to identify leverage points that can efficiently improve nutrition outcomes, while also potentially leading to positive change in other food systems problem areas.

1.2.2 Applications of the Food Systems Methodology

While serving as important foundational roadmaps, generalized frameworks representing interactions between the full range of food system actors are not generally sufficiently-detailed to guide on-the-ground action. To focus research efforts on real-world problem-solving, applied food systems research is generally tailored to address a subset of problem areas, albeit still accounting for exogenous factors stemming from other changes to the system.

For example, Born and Purcell (2006) analyze the validity of the commonly-held assumption that local food systems are necessarily more ecologically-sustainable, nutritious, and socially-just by examining case-study evidence. The authors conclude

9

that this assumption—which they call the "local trap"—is not necessarily valid due to heterogeneities in production scale, labor practices, transportation, and distribution independent of the physical proximity between producer and consumer.

Agroecology is another active area of transdisciplinary food systems research which aims to integrate key findings related to agricultural production, economics, environmental impacts, and social wellbeing (Méndez et al. 2013). Francis et al. (2007) criticize the "siloed" academic discourse in these areas, arguing that a key solution rests in the intentional development of effective pedagogical landscapes that enable students to think critically, adapt to change, cross disciplinary boundaries, and participate with community stakeholders to effect on-the-ground results.

Thompson and Scoones (2009) focus on agricultural policy, criticizing the hegemonic neoliberal approach for its focus solely on technology and production efficiency, and pointing out its failure to adequately grapple with issues of distributional equity and poverty. The authors propose that an interdisciplinary research agenda focusing on social, environmental, and economic factors at the local level can be far more effective than analyzing aggregate statistical averages, as it would allow for deeper understandings of the ways in which dynamic interactions between these factors can help inform more resilient food systems.

Ingram (2011) uses the food systems methodology to analyze the interactions between global climate change and projected food security over the coming decades. By integrating projected effects across all major food systems activities (e.g.

production, processing, consumption, etc.), the author makes the case that a systems-thinking approach can help to identify critical stakeholders, effective leverage points, likely environmental externalities, and research agendas which will help to produce and equitably-distribute sufficient quantities of food to feed our growing population.

Hammond and Dubé (2012) also peer through a systems lens to develop a transdisciplinary framework encapsulating the precursors underlying nutrition and food security. The authors point to three interacting systems—agri-food, health, and environment—as the primary drivers. They suggest that computational simulation models which integrate insights from multiple domains may be a valuable tool to better understand these complex interactions and to broaden the scope of policy intervention proposals designed to curb food insecurity.

Vermeulen et al. (2012) use interdisciplinary research methods to study the relationships between food systems and anthropogenic climate change, accounting for factors from multiple disciplines including agricultural yield trends, economic demand, supply chain logistics, and food safety. Based on computational integrative assessment modeling, the authors analyze the likely strengths and weaknesses associated with interventions designed to intensify agricultural production and/or reduce food waste.

Foran (2015) leverages the food systems approach to analyze the water-energy-food nexus in southeast Asia. She focuses on the interactions between three key study domains—energy efficiency, fisheries, and smallholder agriculture—to develop a

11

holistic model which identifies specific policy interventions aimed at the nexus between the three.

Despite the demonstrated potential of trans / interdisciplinary, systems-oriented research to identify new leverage points to solve food systems problems, there is often a knowledge gap between research focusing on high-level theoretical frameworks and research aimed at identifying workable solutions (Sobal et al. 1998, Ericksen 2008, Thornton et al. 2018). The crux of this shortcoming in food systems praxis stems from the inherent difficulty of distilling the complex interactions between diverse multi-scale interventions into efficient, concrete action steps. In the following sections we argue that recent advances in complex systems scholarship—which relies on computational science to discover the links between individual human actions and macro-level consequences—represent an effective means to integrate systems thinking with the goal of real-world positive change within our food systems.

1.3 Complex Systems Research

The genesis of complexity theory stems from the realization that many real-world problems and observed phenomena are governed by multi-scale, interconnected, adaptive, self-organizing systems (Boccara 2010). Complex systems has been called "the essential 21st Century science" due to its focus on producing radical new ways of understanding these phenomena (ISTFETC Action 2006). Complex systems research often leverages humanity's ever-growing mountain of data, applying novel

computational methods to explore the complex relationships and interactions that lead to emergent, sometimes-unanticipated systems-level outcomes.

Dynamic natural systems are often characterized as "robust yet fragile," meaning that—while they are generally balanced, flexible, and resilient to perturbations—under the right circumstances, they are prone to sudden changes, potentially leading to catastrophic outcomes from which recovery is difficult or impossible (Carlson and Doyle 2000, 2002, Newman et al. 2002). For example, positive feedback loops may accelerate the pace of change as some system parameter nears a point of criticality, potentially shifting the system toward a new equilibrium, or even a fundamentally-different dynamical regime, such as periodic fluctuations, or chaotic, unpredictable outcomes. A system which has reached such a "tipping point" may also experience hysteresis, meaning that simply reverting the key parameter back toward its original value will fail to bring the system back to its original equilibrium (Ball 2004, Sornette 2006).

For example, climate scientists predict that increasing arctic temperatures will result in increased snow-melt, which—because the bright surface of snow reflects a great deal of heat energy away from the earth—forms a positive feedback loop, leading to runaway increases in global temperature. There may well be a critical point—which global society had better hope we haven't already surpassed—beyond which simply reducing global temperatures, for example by ceasing to burn carbon fuels, will no

longer be sufficient to bring the system back to its original, comfortable equilibrium state (Lenton et al. 2008).

A related concept from complex systems science is percolation theory, which describes how incrementally increasing density or connectivity within a system can lead to nonlinear scaling in the size of a spreading event (Stauffer and Ammon 2014). Once the density or connectivity surpasses some "percolation threshold," the expected magnitude of spread in the system begins to skyrocket. This behavior has been observed across many systems, ranging from the impact of tree density upon forest fire size, to the impact of internet connectivity upon resilience to outages, to the impact of connectivity patterns within social, transportation, and trade networks upon the likelihood of catastrophic disease outbreaks (Grassberger 1983, Ohtsuki and Keys 1986, Moore and Newman 2000, Cohen et al. 2000, Sander et al. 2002, Serrano and Boguná 2006, Stauffer and Ammon 2014).

1.3.1 Modeling Complex Systems

Computational modeling represents a potent tool to explore the mechanisms underlying complex, emergent, systems-level outcomes. Systems composed of large numbers of heterogenous interacting agents do not lend themselves to analytical (i.e. formal mathematical) analysis. Instead, the fundamental properties of these systems have historically been obtained through numerical analysis (i.e. data-driven statistical approaches) of computer model outputs (Boccara 2010). The computer-modeling-plus-

numerical-analysis approach has led to many key insights into the dynamics of both conceptual and real-world complex systems.

Echoing Meadows' (2001) argument against the expectation of precise prediction and control, complexity theorists have philosophized about the common presumptions surrounding the computer modeling project. While it is generally assumed that prediction should be the ultimate goal of any computer model (as it is in, for example, weather forecasting), this is not the only valuable result that has come from complex systems modeling endeavors. Even where prediction is unfeasible—due, for example, to a lack of sufficiently-precise parameterization data—models can still illuminate complex phenomena, identify the core dynamics of a system, highlight uncertainties, raise new questions, and serve as touchstones which ground and verify the diverse assumptions of an interdisciplinary research team (Epstein 2008).

Due to the diversity of the complex systems research community—which is made up of computer scientists, physical scientists, natural scientists, social scientists, policymakers, and more—computational models have taken many forms. Some models—often referred to as "toy" models—are explicitly designed to be as simple as possible, and to be generalizable across multiple domains, serving as thought experiments rather than attempting to accurately represent any real-world system (Schelling 1978, Boccara 2010). For example, the Lotka-Volterra predator-prey model (Lotka 1925, Volterra 1926)—which is based on two simple differential equations—is classically taught in undergraduate ecology courses. Despite being empirically

"wrong" when applied to any real-world system, it remains valuable because it helps students to formulate intuitions about the core dynamics at play in predator-prey cycles more generally. Similarly, the Kermack-McKendrick equations, which form the basis of the Susceptible Infective Recovered (SIR) model, offer fundamental insights into the nonlinear threshold nature of disease spread, despite being too simplistic to predict the course of any real-world epidemic (Kermack and McKendrick 1927, Epstein 2008). Despite their limitations as predictive decision support tools, the "stylized facts" generated from this class of models can be academically-valuable in their own right.

Other modelers aim to apply the qualitative insights and technical methods of complex systems modeling to better understand, interpret, and perhaps even forecast real-world outcomes. These context-specific models are often intended to inform high-level decision-making (Miller and Page 2009, Squazzoni and Boero 2010). Whereas some have claimed that increased complexity muddies the waters when formulating concrete plans of action, others embrace the inherent complexity associated with virtually-all important decisions. By leveraging computational models, policymakers can begin to think beyond simple cause and effect logic, leading to improved considerations of uncertainty, probability, and potential unintended consequences (Bankes 2002).

1.4 Application of Complex Systems Modeling to Food Systems

Based on the above, it is clear that important synergies exist between food systems and complex systems research. Both are rooted in systems thinking and seek

to understand how the interactions between a heterogenous set of agents operating across multiple scales and in varying roles produce emergent macro-level outcomes (Meadows 2001). While complex systems science has been applied widely—exploring both "toy" models and context-specific ones—in food systems, we aim to apply insights from foregoing complexity research to effect positive change in real-world problem areas.

The food system encompasses many sub-disciplines, and as such, computational modeling has been applied broadly. Much as applied food systems research must focus on a limited set of problems, the application of complex systems modeling is generally narrowed around one of the key food systems problem areas, including human nutritional wellbeing, agricultural economic viability, food availability and security, disease resilience, ecological externalities, and more.

Perrot et al. (2011) lay out a comprehensive framework establishing the applicability and usefulness of complex systems science to tackle major food systems problems. The authors classify modeling approaches based on their scale (macro, meso, and micro). They also distinguish between empirical models in which all relevant parameterization data are available, versus those in which the general structure is prescribed by the modeling team, while experimental and statistical data are used to fill in missing knowledge wherever possible.

Hammond (2009) considers public policies designed to address the obesity epidemic that is currently sweeping the globe. He notes that the challenge of solving

the problem stems from the fact that it operates within a complex adaptive system, being characterized by multiple scales, heterogeneous actors, uncertainty as to effective leverage points, tipping points, and emergent systems-level properties. The author concludes that effective models of the obesity epidemic should capture these complexities, accounting for multiple theories of change across varied actors operating at different scales.

Van der Vorst et al. (2000) focus on the food supply chain, using discrete-event simulation to evaluate how supply chain management (SCM) decisions impinge upon the efficient and safe transportation of chilled food products from producer to consumer. Despite the fact that SCM is typically approached using analytical models, the authors suggest that the multi-echelon, heterogeneous nature of managerial decision-making necessitates complex dynamic modeling, and thus numerical analysis. Results suggest that a focus on delivery ordering, frequency, lead times, and new informational systems are most likely to increase supply chain throughput.

Bryceson and Smith (2008) also study decision-making within agri-food chains, once again identifying the fundamental characteristics of complex systems within the food system. In this case, the authors point out how firms from different industry sectors (for example grain, beef, dairy, and wool) must collaborate to form both inter-organizational and cross-organizational decision-making networks. Focusing on the Australian beef production chain, the authors compare and contrast Bayesian and agent-

based modeling approaches, concluding that—given sufficient computational power and data availability—both represent valuable means to study such systems.

Finally, a great deal of food systems modeling research has focused on the impact of agricultural production and land-use practices on nonpoint-source pollution. This problem is particularly "wicked," since it occurs within a coupled human and natural-system. Each piece of the puzzle is itself complex, with nutrient flows being governed by geophysical dynamics, and agricultural production practices by human decision-making. Young et al. (1989) were among the first to develop a computer model to explore how nonpoint-source pollution spreads across land cells in a watershed, accounting for factors including runoff vs. retention, nutrient transport through different soils, site elevation, surrounding geography, nutrient solubility, and discharge dynamics. More recently, models focused on human factors have been developed, such as Tsai et al.'s (2015) Land Use Transition Agent-Based Model (LUTABM), which accounts for each land manager's expected utility, socio-economic trends, and forestation / deforestation dynamics to forecast regional land-use changes. Zia et al. (2016) expand on this work, coupling results from LUTABM with nutrient loading models to develop an integrated assessment model which considers the potential effects of both climate change and land use decisions on agricultural runoff and water quality.

Upon first glance, it would seem that the primary purpose of computational modeling for food systems research rests in the final product, i.e., the DDSS that is

ultimately developed and can be leveraged to guide and inform policy decisions. While this is certainly a critical goal, it is merely the endpoint of a lengthy process. In this dissertation, I argue that the computational modeling process itself can be a valuable transdisciplinary research tool: success often derives from the journey, and not solely the destination. Computational model development can serve as a framework to ground a transdisciplinary research endeavor. With a collaboratively-developed model as the backbone, experts from multiple disciplines—both inside and outside academia—can provide input based on their own lived experiences. Their diverse hypotheses can be formally evaluated through the modeling process, and the iterative development of the model can serve as a sounding board through which consensus can be established. This companion modeling process can indeed produce effective DDSS tools, but it can also serve as an ecosystem which promotes transdisciplinary communication and builds mutual trust and understanding.

1.5 References

ACTION, I. (2006). LIVING ROADMAP FOR COMPLEX SYSTEMS SCIENCE.

ALAVI, M., & LEIDNER, D. E. (2001). KNOWLEDGE MANAGEMENT AND KNOWLEDGE MANAGEMENT SYSTEMS: CONCEPTUAL FOUNDATIONS AND RESEARCH ISSUES. MIS QUARTERLY, 107-136.

BALL, P. (2004). CRITICAL MASS: HOW ONE THING LEADS TO ANOTHER: MACMILLAN.

BANKES, S. C. (2002). TOOLS AND TECHNIQUES FOR DEVELOPING POLICIES FOR COMPLEX AND UNCERTAIN SYSTEMS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 99(SUPPL 3), 7263-7266.

BOCCARA, N. (2010). MODELING COMPLEX SYSTEMS: SPRINGER SCIENCE & BUSINESS MEDIA.

BORN, B., & PURCELL, M. (2006). AVOIDING THE LOCAL TRAP: SCALE AND FOOD SYSTEMS IN PLANNING RESEARCH. JOURNAL OF PLANNING EDUCATION AND RESEARCH, 26(2), 195-207.

BRYCESON, K. P., & SMITH, C. S. (2008). ABSTRACTION AND MODELLING OF AGRI-FOOD CHAINS AS COMPLEX DECISION MAKING SYSTEMS. RETRIEVED FROM

CARLSON, J. M., & DOYLE, J. (2000). HIGHLY OPTIMIZED TOLERANCE: ROBUSTNESS AND DESIGN IN COMPLEX SYSTEMS. PHYSICAL REVIEW LETTERS, 84(11), 2529.

CARLSON, J. M., & DOYLE, J. (2002). COMPLEXITY AND ROBUSTNESS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 99(SUPPL 1), 2538-2545.

COHEN, R., EREZ, K., BEN-AVRAHAM, D., & HAVLIN, S. (2000). RESILIENCE OF THE INTERNET TO RANDOM BREAKDOWNS. PHYSICAL REVIEW LETTERS, 85(21), 4626.

DEPPISCH, S., & HASIBOVIC, S. (2013). SOCIAL-ECOLOGICAL RESILIENCE THINKING AS A BRIDGING CONCEPT IN TRANSDISCIPLINARY RESEARCH ON CLIMATE-CHANGE ADAPTATION. NATURAL HAZARDS, 67(1), 117-127.

EPSTEIN, J. M. (2008). WHY MODEL? JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 11(4), 12.

ERICKSEN, P. J. (2008). CONCEPTUALIZING FOOD SYSTEMS FOR GLOBAL ENVIRONMENTAL CHANGE RESEARCH. GLOBAL ENVIRONMENTAL CHANGE, 18(1), 234-245.

FORAN, T. (2015). NODE AND REGIME: INTERDISCIPLINARY ANALYSIS OF WATER-ENERGY-FOOD NEXUS IN THE MEKONG REGION. WATER ALTERNATIVES, 8(1).

FORAN, T., BUTLER, J. R., WILLIAMS, L. J., WANJURA, W. J., HALL, A., CARTER, L., & CARBERRY, P. S. (2014). TAKING COMPLEXITY IN FOOD SYSTEMS SERIOUSLY: AN INTERDISCIPLINARY ANALYSIS. WORLD DEVELOPMENT, 61, 85-101.

FRANCIS, C., LIEBLEIN, G., BRELAND, T., SALOMONSSON, L., GEBER, U., SRISKANDARAJAH, N., & LANGER, V. (2008). TRANSDISCIPLINARY RESEARCH FOR A SUSTAINABLE AGRICULTURE AND FOOD SECTOR. AGRONOMY JOURNAL, 100(3), 771-776.

GRASSBERGER, P. (1983). ON THE CRITICAL BEHAVIOR OF THE GENERAL EPIDEMIC PROCESS AND DYNAMICAL PERCOLATION. MATHEMATICAL BIOSCIENCES, 63(2), 157-172.

HAMMOND, R. A., & DUBÉ, L. (2012). A SYSTEMS SCIENCE PERSPECTIVE AND TRANSDISCIPLINARY MODELS FOR FOOD AND NUTRITION SECURITY. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 109(31), 12356-12363.

INGRAM, J. (2011). A FOOD SYSTEMS APPROACH TO RESEARCHING FOOD SECURITY AND ITS INTERACTIONS WITH GLOBAL ENVIRONMENTAL CHANGE. FOOD SECURITY, 3(4), 417-431.

KERMACK, W. O., & MCKENDRICK, A. G. (1927). A CONTRIBUTION TO THE MATHEMATICAL THEORY OF EPIDEMICS. PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON. SERIES A, CONTAINING PAPERS OF A MATHEMATICAL AND PHYSICAL CHARACTER, 115(772), 700-721.

LENTON, T. M., HELD, H., KRIEGLER, E., HALL, J. W., LUCHT, W., RAHMSTORF, S., & SCHELLNHUBER, H. J. (2008). TIPPING ELEMENTS IN THE EARTH'S CLIMATE SYSTEM. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 105(6), 1786-1793.

LOTKA, A. J. (1925). ELEMENTS OF PHYSICAL BIOLOGY. BALTIMORE: WILLIAMS AND WILKINS: NEW YORK: DOVER PUBLICATIONS.

MEADOWS, D. (2001). DANCING WITH SYSTEMS. WHOLE EARTH, 106, 58-63.

MEADOWS, D. H. (2008). THINKING IN SYSTEMS: A PRIMER: CHELSEA GREEN PUBLISHING.

MÉNDEZ, V. E., BACON, C. M., & COHEN, R. (2013). AGROECOLOGY AS A TRANSDISCIPLINARY, PARTICIPATORY, AND ACTION-ORIENTED APPROACH. AGROECOLOGY AND SUSTAINABLE FOOD SYSTEMS, 37(1), 3-18.

MILLER, J. H., & PAGE, S. E. (2009). COMPLEX ADAPTIVE SYSTEMS: AN INTRODUCTION TO COMPUTATIONAL MODELS OF SOCIAL LIFE (VOL. 17): PRINCETON UNIVERSITY PRESS.

MOORE, C., & NEWMAN, M. E. (2000). EPIDEMICS AND PERCOLATION IN SMALL-WORLD NETWORKS. PHYSICAL REVIEW E, 61(5), 5678.

NEWMAN, M. E., GIRVAN, M., & FARMER, J. D. (2002). OPTIMAL DESIGN, ROBUSTNESS, AND RISK AVERSION. PHYSICAL REVIEW LETTERS, 89(2), 028301.

OHTSUKI, T., & KEYES, T. (1986). KINETIC GROWTH PERCOLATION: EPIDEMIC PROCESSES WITH IMMUNIZATION. PHYSICAL REVIEW A, 33(2), 1223.

PERROT, N., TRELEA, I.-C., BAUDRIT, C., TRYSTRAM, G., & BOURGINE, P. (2011). MODELLING AND ANALYSIS OF COMPLEX FOOD SYSTEMS: STATE OF THE ART AND NEW TRENDS. TRENDS IN FOOD SCIENCE & TECHNOLOGY, 22(6), 304-314.

RUBEN, R., VERHAGEN, J., & PLAISIER, C. (2019). THE CHALLENGE OF FOOD SYSTEMS RESEARCH: WHAT DIFFERENCE DOES IT MAKE? SUSTAINABILITY, 11(1), 171.

SANDER, L., WARREN, C., SOKOLOV, I., SIMON, C., & KOOPMAN, J. (2002). PERCOLATION ON HETEROGENEOUS NETWORKS AS A MODEL FOR EPIDEMICS. MATHEMATICAL BIOSCIENCES, 180(1-2), 293-305.

SCHELLING, T. C. (2006). MICROMOTIVES AND MACROBEHAVIOR: WW NORTON & COMPANY.

SERRANO, M. Á., & BOGUÑÁ, M. (2006). PERCOLATION AND EPIDEMIC THRESHOLDS IN CLUSTERED NETWORKS. PHYSICAL REVIEW LETTERS, 97(8), 088701.

SOBAL, J., KHAN, L. K., & BISOGNI, C. (1998). A CONCEPTUAL MODEL OF THE FOOD AND NUTRITION SYSTEM. SOCIAL SCIENCE & MEDICINE, 47(7), 853-863.

SORNETTE, D. (2006). CRITICAL PHENOMENA IN NATURAL SCIENCES: CHAOS, FRACTALS, SELFORGANIZATION AND DISORDER: CONCEPTS AND TOOLS: SPRINGER SCIENCE & BUSINESS

SQUAZZONI, F., & BOERO, R. (2010). COMPLEXITY-FRIENDLY POLICY MODELLING. INNOVATION IN COMPLEX SYSTEMS. LONDON: ROUTLEDGE.

STAUFFER, D., & AHARONY, A. (2014). INTRODUCTION TO PERCOLATION THEORY: TAYLOR & FRANCIS.

THOMPSON, J., & SCOONES, I. (2009). ADDRESSING THE DYNAMICS OF AGRI-FOOD SYSTEMS: AN EMERGING AGENDA FOR SOCIAL SCIENCE RESEARCH. ENVIRONMENTAL SCIENCE & POLICY, 12(4), 386-397.

THORNTON, P. K., WHITBREAD, A., BAEDEKER, T., CAIRNS, J., CLAESSENS, L., BAETHGEN, W., . . . HERRERO, M. (2018). A FRAMEWORK FOR PRIORITY-SETTING IN CLIMATE SMART AGRICULTURE RESEARCH. AGRICULTURAL SYSTEMS, 167, 161-175.

TSAI, Y., ZIA, A., KOLIBA, C., BUCINI, G., GUILBERT, J., & BECKAGE, B. (2015). AN INTERACTIVE LAND USE TRANSITION AGENT-BASED MODEL (ILUTABM): ENDOGENIZING HUMAN-ENVIRONMENT INTERACTIONS IN THE WESTERN MISSISQUOI WATERSHED. LAND USE POLICY, 49, 161-176.

VAN DER VORST, J. G., BEULENS, A. J., & VAN BEEK, P. (2000). MODELLING AND SIMULATING MULTI-ECHELON FOOD SYSTEMS. EUROPEAN JOURNAL OF OPERATIONAL RESEARCH, 122(2), 354-366.

VERMEULEN, S. J., CAMPBELL, B. M., & INGRAM, J. S. (2012). CLIMATE CHANGE AND FOOD SYSTEMS. ANNUAL REVIEW OF ENVIRONMENT AND RESOURCES, 37, 195-222.

VOLTERRA, V. (1927). VARIAZIONI E FLUTTUAZIONI DEL NUMERO D'INDIVIDUI IN SPECIE ANIMALI CONVIVENTI: C. FERRARI.

WASSERMAN, S., & FAUST, K. (1994). SOCIAL NETWORK ANALYSIS: METHODS AND APPLICATIONS (VOL. 8): CAMBRIDGE UNIVERSITY PRESS.

YOUNG, R., ONSTAD, C., BOSCH, D., & ANDERSON, W. (1989). AGNPS: A NONPOINT-SOURCE POLLUTION MODEL FOR EVALUATING AGRICULTURAL WATERSHEDS. JOURNAL OF SOIL AND WATER CONSERVATION, 44(2), 168-173.

YU, Q., WU, W., YANG, P., LI, Z., XIONG, W., & TANG, H. (2012). PROPOSING AN INTERDISCIPLINARY AND CROSS-SCALE FRAMEWORK FOR

GLOBAL CHANGE AND FOOD SECURITY RESEARCHES. AGRICULTURE, ECOSYSTEMS & ENVIRONMENT, 156, 57-71.

ZIA, A., BOMBLIES, A., SCHROTH, A. W., KOLIBA, C., ISLES, P. D., TSAI, Y., . . . TURNBULL, S. (2016). COUPLED IMPACTS OF CLIMATE AND LAND USE CHANGE ACROSS A RIVER–LAKE CONTINUUM: INSIGHTS FROM AN INTEGRATED ASSESSMENT MODEL OF LAKE CHAMPLAIN'S MISSISQUOI BASIN, 2000–2040. ENVIRONMENTAL RESEARCH LETTERS, 11(11), 114026.

BACKGROUND AND EXECUTIVE SUMMARIES BY CHAPTER

2.1 Chapters 1 and 2 Background, Methodology, and Overarching Goals

2.1.1 Background

Chapters One and Two report on two experiments based on the Regional U.S. Hog Production Network Biosecurity Model (RUSH-PNBM), the development and parameterization of which has been undertaken in collaboration with a nationwide team of livestock industry and veterinary professionals as part of the Animal Disease Biosecurity Coordinated Agricultural Project (ADB-CAP; PI: Julie Smith, DVM/Ph.D.). RUSH-PNBM is a Susceptible / Infective (SI) agent-based model (ABM) that simulates interactions between hog industry actors operating within regional networks. As producer, feed mill, and slaughter plant agents trade livestock and feed according to behavioral heuristics developed in consultation with industry experts, they create chances for disease spread. I have worked on this model for the past three years, in partnership with the UVM Social Ecological Gaming and Simulation (SEGS) lab.

ABMs have been used extensively to analyze complex phenomena that emerge from the relatively-simple actions of a cohort of interacting individuals (Granovetter 1978, Axelrod and Cohen 2000, Janssen and Ostrom 2006). Using the 2013 PEDv outbreak as a case study, two key mechanisms facilitating disease transmission in the hog industry were identified: (a) the transfer of infected animals between premises, and (b) the contamination of livestock and feed transportation equipment (Lowe et al. 2014,

Bowman et al. 2015). Contact between producers, slaughter plants, and feed mills was found to be largely responsible for spreading the disease (O'dea et al. 2015). In RUSH-PNBM, agents are placed realistically in space according to empirically-observed patterns. Decision heuristics developed in consultation with industry experts define how and when agents come into contact, potentially spreading the infection.

Two articles are presented, each of which evaluates the impact of network features on epidemiological resilience through model-based experiments. The key research questions explored in these articles are (a) to what extent does the ongoing trend toward increased livestock producer specialization influence disease risk, and (b) are there network metrics that are consistent predictors of epidemiological vulnerability, either at the agent level, or at the whole-network level?

2.1.2 Methodological Grounding

2.1.2.1 Epidemiological modeling

Many mathematical and computational frameworks have been put forward to understand and forecast the dynamics of disease spread. At the core of many of these approaches is the susceptible / infective / recovered (SIR) framework, under which an infective individual may spread a disease to susceptible individuals for a given period of time, after which the infective individual transitions to the recovered state and cannot be infected again (Kermack and McKendrick 1927). Several SIR derivatives have also been developed, such as the SI framework, which allows for repeated infections. SIR-class models based on differential equations—despite being relatively simple—

successfully replicate many of the complex temporal patterns typical of epidemic events (Heesterbeek 2000, Hethcote 2000, Harko et al. 2014). While numerous differential equation-based SIR variants—both deterministic and stochastic—have been developed, the real-world predictive power of this approach has been criticized, as it implicitly posits both a homogenous population and complete mixing (Anderson et al. 1992).

To overcome these limitations, computer simulations that rely upon numerical rather than analytical methods have increasingly been applied. Cellular Automata (CA) SIR models conceive of individuals as grid cells, with contact mediated by spatial neighborhoods. CA SIR models successfully reproduce complex phenomena such as waves of infection radiating outward from a source (Cliff 1981, Fuentes and Kuperman 1999), providing insights into the importance of agents' relative spatial positions (Keeling 1999, Pfeifer et al. 2008). Phase transitions have also been replicated using CA models (Kinzel 1985). Further developments have examined the impact of heterogenous susceptibility, transmission rate, and infectious period parameters (Keeling et al. 2003), as well as modulating contact rates and transmission probabilities as a function of spatial proximity (Mikler et al. 2005, White et al. 2007).

Realizing that real contagions occur within populations of agents whose relationships are not defined solely by spatial proximity, and whose differences do not simply come down to differential parameter values, in recent decades, researchers have begun to unpack the critical relationships between network structures, interaction

contexts, and contagion dynamics. Bailey (1975) suggests that, in global human epidemics, the majority of spreading occurs within small network clusters, and is only transmitted beyond those local contexts by individuals who occasionally bridge between them. Similar patterns have been empirically observed in livestock epidemics. Firestone et al. (2011) evaluated the spatial distribution and contact patterns of infected premises during the 2007 Equine Influenza outbreak in Australia, finding strong evidence that the movement of infected horses between spatially-clustered groups of premises was correlated with the spread of the disease. Similarly, Fournié et al. (2013) analyzed the network connectivity patterns of agents involved in Vietnam's live bird markets to assess the role of network structure on the spread of H5N1 influenza. The team found that individual market locations served as hubs, linked by a few bridging individuals, and concluded that the contact network may be largely "disconnected" by focusing on disinfection of transportation equipment at a few of the large hubs to prevent spread between localized clusters. In the 2013 U.S. hog industry PEDv epidemic, the structure of transportation networks for feed and livestock was found to be the primary predictive factor underlying observed disease spread patterns (O'dea et al. 2015).

In light of such observations, recent epidemiological models have increasingly investigated the role of the structural topology of "mixing networks" on disease percolation (Keeling and Eames 2005). In models using simple heterogeneously degree-distributed networks—including scale-free topologies—epidemics have been

shown to cascade from highly-connected hubs through smaller degree classes (Barthélemy et al. 2004), although the sharp percolation threshold associated with homogeneous models is not observed (Pastor-Satorras and Vespignani 2001, Vespignani 2012). In metapopulation (or subpopulation) models, however, critical values for the rate of spreading between otherwise-isolated groups has been found to induce percolation, shedding light on the empirical observation that epidemic events tend to be either quite local, or nearly systemic (Bailey 1975, Bajardi et al. 2011). Watts et al. (2005) used a metapopulation SIR model to show that a simple hierarchical system of interaction contexts can lead to multiscale epidemic size distributions, further explaining the role of network structure on the phase change between localized and global epidemics.

Coelho et al. (2008) characterize the complexification of epidemiological simulations over time as a shift between "strategic models" that explore the fundamental features of epidemics, to "tactical models" that mirror the conditions within which a real-world epidemic may unfold. Agent-based SIR models (ABMs) have proven especially useful as "tactical models" due to their ability to generate empirically-calibrated, spatially-explicit networks of interacting agents that are heterogeneous not only in their parameter values, but also in the behavioral heuristics that govern how and when contact occurs (Rahmandad and Sterman 2008). ABMs simulating disease spread have utilized spatial technologies such as GIS (Perez and Dragicevic 2009) and have incorporated real-world network data—such as the structure

of airline routes, or telephone contact records—that have empirically been found to correlate with disease spread (Colizza et al. 2006, Frias-Martinez et al. 2011). Similar approaches have been applied to analyze disease spread in livestock production networks, incorporating factors such as the life cycle stages of animals, operational details of farms, and different disease transmission vectors (Bagni et al. 2002, Bradhurst et al. 2013).

Another class of ABMs lets network structures emerge organically during each model run as a result of heuristics governing agents' decision-making. Using this methodology, Ghani and Garnett (2000) found network centrality measures that predicted an agent's chance of either getting or spreading a sexually-transmitted disease. Eubank et al. (2004) developed an ABM that utilized heuristics parameterized from large-scale datasets to generate realistic urban social contact networks in order to assess epidemiological vulnerabilities. The ABM developed for this experiment is of this latter class, using empirical data to generate realistic hog production systems, and behavioral heuristics specific to each agent type to generate the contact networks that underlie disease spread in real time as the model runs.

2.1.2.2 Overview of RUSH-PNBM

RUSH-PNBM may be classified as an agent-based, susceptible / infective (SI) epidemiological model, with high context specificity. Specialized versions of the model, v.0.8 and v.1.2, were developed for each experiment. The two versions are fundamentally similar in operation, but in v.0.8, agents are simply placed in a

continuous two-dimensional space, whereas in v.1.2, agents exist in a GIS-based spatial framework within study areas defined by three prominent hog-producing U.S. states (IL, IA, and NC), and have their locations, sizes, and typologies initialized using USDA, geospatial, and proprietary hog industry datasets.

Features shared between the two versions include multiple agent classes, encompassing six types of hog producers differing by industry role, in addition to feed mills and slaughter plants. Agents' decision rules govern how and when they interact within contact networks according to heuristics developed out of a collaborative and iterative expert informant advisory process. As batches of pigs reach the appropriate age, producer agents transfer them to trading partners within a certain radius, of the appropriate industry role, and with sufficient excess capacity. Feed mill agents periodically generate delivery routes encompassing a subset of producers in their service areas. The model is calibrated to reflect the transmission of a fecal-oral disease like Porcine Epidemic Diarrhea virus (PEDv), meaning that various modes of contact between agents—either in the form of contaminated feed, infected animals, or contaminated transportation equipment—each carry specific probabilistic risks of infection spread (O'dea et al. 2015). For full implementation details, see Appendices 1 and 2.

Data output by RUSH-PNBM allow for analyses of how the structure of livestock production networks can influence epidemiological resilience in the face of regional disease threats. Key parameter values and epidemiological statistics are output

after each run, in addition to adjacency lists encoding inter-agent contact patterns, which can be post-processed to analyze network features. By calculating key network metrics for these graphs, we can evaluate how agents' network positionalities, along with overall network structures, can make either individual agents, or the entire regional production networks, differentially-vulnerable to catastrophic disease risks.

2.1.3 High-Level Objectives

Food animal disease presents a threat to the economic wellbeing of industry actors, and in some cases—for example where quarantines or stop-motion orders are called for—such events can impinge upon the food security of members of the public, driving up food prices, and thus limiting access to key food products. Whereas the threat of such catastrophic events is well understood by stakeholders—and biosecurity efforts are continually promoted and implemented across all levels of livestock production systems—RUSH-PNBM facilitates more precise decision-making to combat disease threats, allowing funds and interventions to be allocated strategically for maximum effect.

By better understanding how patterns of trade and interaction between the diverse entities involved in complex livestock production networks translate into epidemiological risk factors, government regulatory agencies, private livestock production system managers, veterinarians, and other stakeholders can focus their preventive efforts on the nodes, links, and structures of production networks that are most likely to promote catastrophic epidemiological events. Such biosecurity

33

interventions could take the form of discrete physical measures, for example strategically-placed truck wash facilities; human behavioral approaches, such as means to promote farmer adoption of shower-in-shower-out protocols or the development of better educational resources used by university extension services; or systemic changes to industry network structures, for example efforts to limit disease spread by removing risky network edges, thus isolating potential outbreaks.

2.2 Chapter 1 Executive Summary

Chapter One uses RUSH-PNBM v.0.8, which focuses on network structure rather than context specificity. The goal of chapter one is to use RUSH-PNBM to study the interplay between two ongoing network-structural trends and the epidemiological resilience of hog industry production networks. The first of these trends is the growing size and spatial density of networks, and the second is increasing producer specialization, for example from single-phase "farrow to finish" farms to multi-phase production systems in which each premises focuses on a specific life cycle stage.

Whereas in the past it was standard practice for a single producer to take a pig "from farrow to finish," it is now increasingly common for livestock to be housed at two or, more recently, three different producer operations throughout their lives, with each operation specializing in a specific life cycle stage (USDA APHIS 2006). While there are efficiency advantages associated with increased specialization, its effect on disease spread is not fully understood.

34

2.2.1 Methods

To evaluate the interplay between these effects, three experimental treatments are used. In the 1-phase treatment, all producer agents are farrow to finish producers. In the 2-phase treatment, half are farrow to feeder, and half are feeder to finish. In the 3-phase treatment—the highest level of producer specialization—one third are farrow to wean, one third wean to feeder, and one third feeder to finish. All other parameters, such as the timing of agents' decision rules, and probabilities of infection, are held constant between treatments.

For each treatment, a parameter sweep experiment is conducted, varying the number of producers in the network—which corresponds to their spatial density—from 10 to 1500. 15,000 model runs are conducted for each treatment, and results are evaluated according to producer density level. The dependent variables chosen to indicate the severity of an infection are (a) the overall infection duration in the system in model days, and (b) the proportion of agents infected.

The purpose of the parameter sweep is to evaluate percolation dynamics in the networks. Percolation is a mathematical phase-change that occurs when the density of entities in a system becomes sufficient that the magnitude of spreading events no longer scales linearly with each incremental increase in density, but instead accelerates rapidly toward near-complete spreading. The point at which this transition occurs—the percolation threshold—is defined as the point at which, in an infinite network, the expected size of the "giant component"—i.e. the subset of nodes encountering the

35

spreading event—is also infinite (Stauffer and Aharony 1994, Grimmett 1999, Albert et al. 2000, Carlson and Doyle 2000, 2002, Newman et al. 2002).

Disease epidemics can be aptly evaluated through the lens of percolation theory. In general, pathogens are regularly introduced to a system with little consequence beyond relatively-minor localized outbreaks. However, due to stochasticity and internal heterogeneities in contact network structures, occasionally a single pathogen can ignite a widespread epidemic, a property referred to as "robust yet fragile" (Grassberger 1983, Ohtsuki and Keyes 1986, Moore and Newman 2000, Sander et al. 2002, Serrano and Boguná 2006, Sornette 2006).

As more producers are added to the system, the network's average degree should increase. With more edges in the system, we can anticipate that percolation behavior should emerge after a critical number of producers are added. Further, higher producer specialization requires more transfers of animals over their life cycle, adding contact network edges where previously there were none. Following the work of both epidemiological modelers and empirical analysts, a greater probability of large-scale "global" epidemics should be observed wherever the network typology is such that, without additional edges acting as bridges, the contagion would have been isolated within localized clusters. Here we investigate the extent to which increased producer specialization may create such bridging edges, negatively impacting the overall disease resilience of the simulated networks by shifting the percolation threshold downward.

36

Since the model is too complex to calculate expected percolation thresholds analytically, a numerical method was devised to approximate them. This method uses a LOESS-smoothing approach, with the point at which the slope of the LOESS-smoothed curve is maximized defining the approximate percolation threshold. This smoothing approach allows for a meaningful comparison between treatments and their effects on disease spread risk.

2.2.2 Key Takeaways

Results indicate that producer specialization is a very important factor underlying epidemic percolation in livestock production networks. Specifically, single-phase networks have significantly higher percolation thresholds than the more specialized treatments, meaning that these systems can withstand a higher density of producers in a given area before systemic epidemics become likely. Producer specialization also positively correlates with outcome variability once the threshold is surpassed, meaning that, for high-specialization networks, there is much more uncertainty as to the ultimate size and duration of an epidemic in the system.

As a secondary analysis, adjacency lists defining the inter-agent contact networks are output after each model run. Network analysis tools are then used to analyze the graphs, allowing us to examine the interplay between several network typology indicators and the percolation dynamics we identify.

Overall, our experiment reveals that both hog industry network trends—increased spatial density and increased specialization—negatively impact average

network-level disease resilience. Policy recommendations based on these findings include an additional focus on network configurations—rather than biosecurity interventions aimed at individual agents—when formulating strategies to reduce disease risk in livestock production chains.

2.3 Chapter 2 Executive Summary

Chapter two uses RUSH-PNBM v.1.2, which relies on empirical data to position agents with realistic characteristics at realistic locations within a GIS framework defined by the U.S. states of North Carolina, Iowa, or Illinois. The goal of this experiment is to assess how formal properties of the production systems in each state may impinge upon the disease risk for both (a) the overall hog production systems within the state, and (b) the individual agents within those networks. We conduct the assessment using network analysis algorithms and employ both traditional statistical techniques and a genetic programming algorithm to ascertain the degree to which either single network metrics or mathematical combinations of metrics ("meta-metrics") correlate with disease risk.

2.3.1 Methods

The GIS-enabled RUSH-PCBM v.1.2 model was executed 50 times for each study area, maintaining all parameters constant across all repetitions, aside from the numbers, spatial distributions, and pool of potential trading partners for each type of agent, which were hard coded for each state based on USDA NASS and other parameterization data, together with proprietary initialization functions.

Throughout each run, a contact network was built up by tracking the number of times each agent contacted another agent—as a result of either delivering or receiving livestock or feed—with the final edge weights being equal to the number of contacts between connected agents. In a similar fashion, an infectivitity network was constructed, with edge weights representing the number of times an agent transmitted the infection through inter-agent contact. The networks generated during the experiment, along with key statistics associated with each agent, were output after the conclusion of each run.

We use a twofold approach to analyze the relationship between network structure and epidemic risk factors in the RUSH-PNBM output data. First, we examine the relationship between network metrics that capture overarching features of the study area networks as a whole, and the average number of times an agent in these networks received or spread the infection. Second, we examine the relationships between node-level metrics that capture an individual agent's positionality in the network, and that individual agent's risk of receiving or spreading the infection.

Whole-network metrics include average degree, k-core size and order, average assortativity, and average clustering. Node-level metrics include in- and out-degrees and several measures of centrality. These metrics are statistically evaluated against disease vulnerability to determine the extent to which network structure and node positionality indicators may be utilized to predict infection risk.

2.3.2 Key Takeaways

Our experimental results show that, although the set of probabilities governing whether a given agent will spread or receive an infection upon contact were the same across all three study areas, there is nevertheless a great deal of variability in the scale and duration of epidemic events between study areas. Since the spread probabilities were constant, and nodes' spatial positions and trading partner pools were maintained across all runs within each state, the observed disease risk discrepancies must result from differences in the network topologies associated with each study area.

Overall, we find that no single network metric is sufficient to predict the vulnerability—in terms of infection frequency—of a given node in our model results across study areas. None of the global metrics explains more than 5% of the variability in the vulnerability of individual agents. The best node-level metric, weighted in-degree, explains 30% of the variability, but this result is self-evident, and still does not serve as an effective indicator of risk.

Our data show clear trends in some cases, but confusing non-linearities in others. Single metrics were often a strong predictor of node-level infection risk in one study area, but not in the others. We believe that this is due to complex interactions between network topology and node positionality indicators in complex ways. With this in mind, we present a method to incorporate both overall network structure as well as node positionality into a single indicator to predict epidemiological risk.

Using a Genetic Programming (GP) algorithm (Schmidt and Lipson 2013), we were indeed able to discover formulae combining multiple metrics—both node-level and global—to make more-precise predictions as to the infection risk of a given node. Using this novel methodology, which we're calling "meta-metrics," we were able to identify a node's vulnerability with up to 91% accuracy across all three study areas.

2.4 Chapter 3 Executive Summary

Chapter Three describes a procedure to optimize public investments into interventions aimed at bolstering future global food supplies, with the goal of improving food security over the decades to come according to UN Sustainable Development Goal Two. This work picks up on the "food wedge" framework originally conceived by Keating et al. (2014), which describes how multiple interventions aimed at either reducing demand or increasing supply may stack together to fill the large food production deficit that is projected to face global society by 2050. The overarching goal of Chapter Three is to use computational optimization to expand the theoretical food wedge model into a practicable decision support tool.

Owing to inherent heterogeneities between wedges—for example the fact that some wedges are long-term while some are short-term, some multiply existing levels while some simply add or subtract, some have a limited upside potential, and so on—as well as the fact that the effect of any wedge intervention depends in part on previous interventions into both that wedge and possibly other wedges, it becomes impossible to simply intuit one's way to an investment schedule that optimally allocates resources.

41

Further, the solution space is too large and complex to rely upon brute-force computational methods (i.e. simply trying all options). Complex tradeoff environments like this are prime candidates for computational optimization. We employ a multi-objective evolutionary algorithm to search for solutions that both minimize the supply / demand gap and minimize total cost.

2.4.1 Background

Due to factors such as a rising population and the shift toward more resource-intensive diets, global food demand is projected to increase significantly by 2050 (Brown and Cameron 2000, United Nations 2015). In light of production-side stressors such as soil / water degradation and climate change, scholars question whether our current food systems can meet future food demand without significant intervention-driven change (Eswaran et al. 2001, Keating et al. 2014).

Although a single solution has not been identified, it is suggested that public investments into multiple "food wedges," aimed at either reducing calorie demand or increasing supply, may stack together to close projected supply / demand gaps (Keating et al. 2014). The idea of the food wedge model—which is based on the Pacala and Socolow (2004) climate stabilization wedge model—is to identify a diverse portfolio of public interventions targeting either supply-increasing or demand-reducing measures (e.g. sustainable agricultural intensification, reducing consumption of meat, avoiding soil degradation, etc.) can each play a part in the overall solution. Keating et al. (2014) used an expert informant advisory process to arrive at ballpark estimates of the percent

of necessary demand-reduction or supply-increase that each wedge could theoretically contribute. This work builds upon the basic food security wedge framework by employing computational and mathematical methods to lend depth and precision to the wedge model's value as a decision support tool.

The original climate change wedge model (Pacala and Socolow 2004), along with its subsequent application to global food security (Keating et al. 2014), were meant primarily as high-level strategic concepts. Although the wedge framework serves as a valuable tool to conceptualize multiple interacting solutions to a large-scale problem, it was never intended to serve as a decision support tool, or to evaluate how resources may be allocated most efficiently. Examining Keating's framework with a critical eye reveals several implicit presuppositions that cast serious doubt on the model's real-world efficacy for practitioners in the field.

First, the authors assume that the effect of each wedge will unfold linearly at a constant rate (thus making the wedges truly "wedge"-shaped). In other words, each wedge is assumed to continue to effect the same annual quantity of change (at least out to 2050), without any acceleration or deceleration. In reality, intervention-driven change is likely to exhibit neither an immediate effect nor a steady trajectory, instead generally being characterized by a sigmoid curve, in which the annual level of change first slowly ramps up, and later gradually tapers off. Such a pattern has been commonly observed in the diffusion of innovations (Rogers 2010). For a food systems planner, it is critical to understand not only the ultimate quantity of change that can be expected

43

from each wedge, but also the temporal trajectory that characterizes the way in which that change will ultimately be realized.

The second implicit assumption is that changes resulting from investments into each wedge will all begin simultaneously (apparently in 2010). Realistically, only a certain quantity of funding is available during any given time period, so investments will need to be strategically staggered, with the highest-priority interventions receiving funding first. A wedge model that could serve as a decision support tool would need to include recommendations as to the timeframe by which investments into each wedge should optimally be scheduled.

We must also consider the scale at which action is feasible. While international organizations—such as the UN FAO—to some extent shape global food policy, concrete policy interventions are actualized at the national, regional, or even local scale. Thus, while global policy—as is considered in the Keating et al. (2014) wedge model— may be valuable from a high-level strategic perspective, it would also be useful to consider the finer-grained details of policy implementation by, for example, reshaping the model to focus on the goals and theories of change associated with an existing regional food policy organization.

The fourth assumption relates to how the baseline curves for supply and demand are conceived. The demand curve is pinned to world population growth projections, but demand projections should encompass other factors as well, such as trends in diets toward more meat consumption as people in developing countries gain wealth. The

authors discuss how interventions in these areas can bend the curve, but it is unclear whether the baseline curve includes "natural" changes; that is, those that would occur without any intervention. On the supply side, the authors break the baseline simply into "avoiding losses" and "filling the production gap," with the overall supply baseline (accounting for both factors) apparently projected to take a sharp and steady turn for the worse starting in 2010. In reality, the supply baseline—like the demand baseline— should encompass projected trends in each supply wedge, accounting for both positive trends (such as the ongoing spread of yield-enhancing agricultural innovations), as well as negative trends (such as climate change pressures), with the final curve being defined as the summation of all relevant factors.

The above presuppositions limit the usefulness of the Keating model for real-world strategists making high-level decisions. In light of these limitations, yet also recognizing the overall potential of the interacting-wedge framework as a guide for tackling complex, systemic food systems problems, the aim of this project is to build the simple Keating wedge model into a decision support tool that captures the nuances required for food systems leaders to begin to thoughtfully chip away at the looming global food supply/demand gap.

2.4.2 Methods

The concept of our wedge investment optimization procedure is to first

formalize the interacting dynamics of each wedge mathematically, and then use

computational optimization to hone in on an investment schedule that fulfils the desired

objectives. This procedure is based on the wedge concept, yet addresses the

shortcomings described above.

As a proof of concept, we selected six wedges from Keating et al. (2014)

identified as having large potential effects, and also added one wedge not included in

Keating (birth control). The functions associated with each were roughly

parameterized through a survey of the primary literature.

Following researchers such as Kummu et al. (2017), a technical insight that

adds depth to the original Keating work is the realization that some wedges act to

increment or decrement supply or demand (for example adding new land), while others

serve to multiply or divide the existing level of production or consumption (for example

increasing agricultural efficiency). Wedges may be further classified based on whether

they reduce demand or increase supply, yielding four high-level wedge categories.

Once the wedges are assigned to their appropriate categories, their dynamics

can be encoded formally. As in Keating et al. (2014), we use total kilocalories (kcal)

per year as a common metric of impact. Obviously, the model requires a function

describing how investment in a wedge will play out as supply or demand change over

time. However, as discussed above, each wedge also contributes to the baseline in

some way, requiring another function. Finally, owing to the realization that, for example, there is only so much land available for agriculture, we also needed a way to define how investment may become less effective over time. Thus, each wedge can be defined as a system of three interacting equations, these being:

1. Baseline change: What will happen if we do nothing?

2. Intervention-driven change: What will happen if we invest 100% of one year's food security policy budget into a wedge?

3. Diminishing returns: How will the effectiveness of additional investments in a wedge diminish once a certain level of change is achieved?

More specifically, the baseline change function defines the level of expected change in a wedge in the absence of any intervention, as well as the rapidity and trajectory with which that change may be expected to play out. Similarly, the intervention-driven change function defines the quantity, timeframe, and temporal dynamics of positive change that may be expected to result, going forward from the investment date. This function is calibrated for 100% annual budget investment, but its magnitude can be scaled down proportional to the investment quantity. Finally, the diminishing returns function defines how the efficacy of additional investments into a wedge can be expected to make less and less difference after a certain level of change from that wedge has been achieved.

In light of the above, it becomes apparent that—since the result of any given wedge investment depends in part on (a) the baseline curves, (b) previous investments

into the wedge, and (c) interactions between multiplying/dividing and adding/subtracting wedge interventions—stacking the effects of multiple wedge interventions is quite complex. To accomplish this, I developed an objective function that begins by establishing the overall baseline curves, and then uses a genome vector representing the schedule of wedge interventions to mathematically 'stack' the impact of each intervention—according to their function parameters—to evaluate the overall summed effect from all wedge interventions on the global food supply and demand curves. The function then calculates the total food supply/demand gap over the study period (which we call SD fitness), and also the level of public investment—relative to the total available public and private funding that may theoretically be allocated to food security—required to meet the intervention schedule (cost fitness).

We then employ this objective function within an evolutionary algorithm known as Multi-Objective Differential Evolution (MODE) to optimize the wedge intervention schedule (Storn and Price 1997). The algorithm generates a series of solutions along a Pareto front defined by SD fitness and cost fitness, allowing the user to evaluate tradeoffs between efficacy and cost. Solutions take the form of investment schedules that specify the level of funding devoted to each wedge in each investment year.

The primary goal of Chapter Three is an experiment to hone the parameter values of the MODE algorithm such that it functions best on this optimization task. We evaluate the effect of seven key parameters, accounting for both solution quality and execution time, to determine effective baselines. We also analyze the preliminary

48

solution set in several ways to explore patterns and tradeoffs in food wedge investment prioritization.

2.4.3 Key Takeaways

Despite the model being a proof of concept, initial results are promising, with the optimization procedure successfully discovering roadmaps toward closing the global food supply / demand gap, while also accounting for cost. We find that our baseline parameterization optimizes the effectiveness of the algorithm for this class of problems on both solution efficacy and computation time.

In general, results indicate that early investments should be made in long-term demand-reducing interventions such as birth control, as well as in interventions that bolster long-term supply-side productivity, like closing yield gaps and adapting agricultural systems to climate change. Another key takeaway is that, in general, if more funding is put into the problem in the near term, less overall funding will be required as we look out to 2050 and beyond.

A limitation of the study is that, at present, wedge parameterization was accomplished using a relatively-cursory survey of the primary literature. This yielded good estimates for some wedge function parameters, but also left several knowledge gaps that were simply filled by the best estimates of our modeling team. In light of such uncertainty, the current parameterization is meant mainly as a proof of concept. Further work on this project will endeavor to flesh out these known unknows. Much empirical as well as model-based research has been conducted in the area of global

food policy (e.g. Ryan and Gross 1943, Ellison and Fudenberg 1993, Fischer et al. 1996, Baerenklau 2005, Rogers 2010). We plan to use qualitative methods to solicit input from a cohort of experts in the various fields underlying each wedge. We have already developed a survey instrument which we intend to distribute to domain-level experts to parameterize the wedge functions. An iterative Delphi procedure may also be utilized strategically to resolve conflicting projections and converge upon realistic parameter values (Rescher 1998). We may also conduct follow-up interviews where called for.

Overall, our findings suggest that, once precisely calibrated, our method should be very useful as a decision support tool to guide high-level public policies aimed at reducing the threat of hunger stemming from future gaps between global food production and consumption. The model must navigate a highly-complex "fitness landscape," simultaneously weighing factors such as the relative influence of the spread of agricultural innovations in the near term vs. productivity-killing environmental degradation in the long term; the impact of previous interventions upon future ones; and tradeoffs between investing in solutions that have small but fast-acting impacts vs. those that take many years but ultimately have large effects. Because humans cannot simply intuit their way to good solutions within such tradeoff-laden, systems-level problem domains, computational decision support tools will likely become increasingly important. Computational methods such as the one demonstrated here can help planners and policymakers to analyze the effects of many compounding factors in

aggregate, allowing for more informed, evidenced-based policymaking that directs

funding to solutions that promise the maximum long-run positive impact.

2.5 References

ADGER, W. N., EAKIN, H., & WINKELS, A. (2009). NESTED AND TELECONNECTED VULNERABILITIES TO ENVIRONMENTAL CHANGE. FRONTIERS IN ECOLOGY AND THE ENVIRONMENT, 7(3), 150-157. DOI:10.1890/070148

ALBERT, R., JEONG, H., & BARABÁSI, A.-L. (2000). ERROR AND ATTACK TOLERANCE OF COMPLEX NETWORKS. NATURE, 406(6794), 378-382.

AMINE, E., BABA, N., BELHADJ, M., DEURENBERY-YAP, M., DJAZAYERY, A., FORRESTER, T., . . . MBUYAMBA, J. (2002). DIET, NUTRITION AND THE PREVENTION OF CHRONIC DISEASES: REPORT OF A JOINT WHO/FAO EXPERT CONSULTATION: WORLD HEALTH ORGANIZATION.

ANDERSON, R. M., MAY, R. M., & ANDERSON, B. (1992). INFECTIOUS DISEASES OF HUMANS: DYNAMICS AND CONTROL (VOL. 28): WILEY ONLINE LIBRARY.

AXELROD, R. M., AXELROD, R., & COHEN, M. D. (2000). HARNESSING COMPLEXITY: ORGANIZATIONAL IMPLICATIONS OF A SCIENTIFIC FRONTIER: BASIC BOOKS.

AZZARRI, C., BACOU, M., COX, C. M., GUO, Z., & KOO, J. (2016). SUBNATIONAL SOCIO-ECONOMIC DATASET AVAILABILITY. NATURE CLIM. CHANGE, 6(2), 115-116. DOI:10.1038/NCLIMATE2842

BAERENKLAU, K. A. (2005). TOWARD AN UNDERSTANDING OF TECHNOLOGY ADOPTION: RISK, LEARNING, AND NEIGHBORHOOD EFFECTS. LAND ECONOMICS, 81(1), 1-19.

BAGNI, R., BERCHI, R., & CARIELLO, P. (2002). A COMPARISON OF SIMULATION MODELS APPLIED TO EPIDEMICS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 5(3).

BAILEY, N. T. (1975). THE MATHEMATICAL THEORY OF INFECTIOUS DISEASES AND ITS APPLICATIONS: CHARLES GRIFFIN & COMPANY LTD, 5A CRENDON STREET, HIGH WYCOMBE, BUCKS HP13 6LE.

BAJARDI, P., POLETTO, C., RAMASCO, J. J., TIZZONI, M., COLIZZA, V., & VESPIGNANI, A. (2011). HUMAN MOBILITY NETWORKS, TRAVEL RESTRICTIONS, AND THE GLOBAL SPREAD OF 2009 H1N1 PANDEMIC. PLOS ONE, 6(1), E16591.

BAKKER, K. (2012). WATER SECURITY: RESEARCH CHALLENGES AND OPPORTUNITIES. SCIENCE, 337(6097), 914-915. DOI:10.1126/SCIENCE.1226337

BARRETT, C. B. (2013). FOOD SECURITY AND SOCIOPOLITICAL STABILITY: OUP OXFORD.

BARTHÉLEMY, M., BARRAT, A., PASTOR-SATORRAS, R., & VESPIGNANI, A. (2004). VELOCITY AND HIERARCHICAL SPREAD OF EPIDEMIC OUTBREAKS IN SCALE-FREE NETWORKS. PHYSICAL REVIEW LETTERS, 92(17), 178701.

BECK, U. (2010). REMAPPING SOCIAL INEQUALITIES IN AN AGE OF CLIMATE CHANGE: FOR A COSMOPOLITAN RENEWAL OF SOCIOLOGY*. GLOBAL NETWORKS, 10(2), 165-181. DOI:10.1111/J.1471-0374.2010.00281.X

BECK, U., & GRANDE, E. (2010). VARIETIES OF SECOND MODERNITY: THE COSMOPOLITAN TURN IN SOCIAL AND POLITICAL THEORY AND RESEARCH. THE BRITISH JOURNAL OF SOCIOLOGY, 61(3), 409-443. DOI:10.1111/J.1468-4446.2010.01320.X

BECKER, A., INOUE, S., FISCHER, M., & SCHWEGLER, B. (2012). CLIMATE CHANGE IMPACTS ON INTERNATIONAL SEAPORTS: KNOWLEDGE, PERCEPTIONS, AND PLANNING EFFORTS AMONG PORT ADMINISTRATORS. CLIMATIC CHANGE, 110(1), 5-29.

BOWMAN, A. S., KROGWOLD, R. A., PRICE, T., DAVIS, M., & MOELLER, S. J. (2015). INVESTIGATING THE INTRODUCTION OF PORCINE EPIDEMIC DIARRHEA VIRUS INTO AN OHIO SWINE OPERATION. BMC VETERINARY RESEARCH, 11(1), 38.

BRADHURST, R., ROCHE, S., GARNER, M., SAJEEV, A., & KWAN, P. (2013). MODELLING THE SPREAD OF LIVESTOCK DISEASE ON A NATIONAL

SCALE: THE CASE FOR A HYBRID APPROACH. PAPER PRESENTED AT THE PROCEEDINGS OF THE 20TH INTERNATIONAL CONGRESS ON MODELLING AND SIMULATION (MODSIM2013), MODELLING AND SIMULATION SOCIETY OF AUSTRALIA AND NEW ZEALAND.

BRANDES, U., & FLEISCHER, D. (2005). CENTRALITY MEASURES BASED ON CURRENT FLOW. PAPER PRESENTED AT THE STACS.

BROWN, P. M., & CAMERON, L. D. (2000). WHAT CAN BE DONE TO REDUCE OVERCONSUMPTION? ECOLOGICAL ECONOMICS, 32(1), 27-41.

CARLSON, J. M., & DOYLE, J. (2000). HIGHLY OPTIMIZED TOLERANCE: ROBUSTNESS AND DESIGN IN COMPLEX SYSTEMS. PHYSICAL REVIEW LETTERS, 84(11), 2529.

CARLSON, J. M., & DOYLE, J. (2002). COMPLEXITY AND ROBUSTNESS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 99(SUPPL 1), 2538-2545.

CGIAR. (2014). CLIMATE CHANGE, AGRICULTURE AND FOOD SECURITY (CCAFS) DATASETS. RETRIEVED FROM: HTTPS://DATAVERSE.HARVARD.EDU/DATAVERSE/CCAFSBASELINE

CHALLINOR, A. J., ADGER, W. N., & BENTON, T. G. (2017). CLIMATE RISKS ACROSS BORDERS AND SCALES. NATURE CLIM. CHANGE, 7(9), 621-623. DOI:10.1038/NCLIMATE3380

CHALLINOR, A. J., KOEHLER, A. K., RAMIREZ-VILLEGAS, J., WHITFIELD, S., & DAS, B. (2016). CURRENT WARMING WILL REDUCE YIELDS UNLESS MAIZE BREEDING AND SEED SYSTEMS ADAPT IMMEDIATELY. NATURE CLIM. CHANGE, 6(10), 954-958. DOI:10.1038/NCLIMATE3061

CLIFF, A. D. (1981). SPATIAL DIFFUSION: AN HISTORICAL GEOGRAPHY OF EPIDEMICS IN AN ISLAND COMMUNITY: CUP ARCHIVE.

COELHO, F. C., CRUZ, O. G., & CODEÇO, C. T. (2008). EPIGRASS: A TOOL TO STUDY DISEASE SPREAD IN COMPLEX NETWORKS. SOURCE CODE FOR BIOLOGY AND MEDICINE, 3(1), 3. DOI:10.1186/1751-0473-3-3

COLIZZA, V., BARRAT, A., BARTHÉLEMY, M., & VESPIGNANI, A. (2006). THE ROLE OF THE AIRLINE TRANSPORTATION NETWORK IN THE PREDICTION AND PREDICTABILITY OF GLOBAL EPIDEMICS. PROCEEDINGS

OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 103(7), 2015-2020.

DUBOIS, O. (2011). THE STATE OF THE WORLD'S LAND AND WATER RESOURCES FOR FOOD AND AGRICULTURE: MANAGING SYSTEMS AT RISK: EARTHSCAN.

ELLISON, G., & FUDENBERG, D. (1993). RULES OF THUMB FOR SOCIAL LEARNING. JOURNAL OF POLITICAL ECONOMY, 101(4), 612-643.

ESWARAN, H., LAL, R., & REICH, P. (2001). LAND DEGRADATION: AN OVERVIEW. RESPONSES TO LAND DEGRADATION, 20-35.

EUBANK, S., GUCLU, H., KUMAR, V. A., & MARATHE, M. V. (2004). MODELLING DISEASE OUTBREAKS IN REALISTIC URBAN SOCIAL NETWORKS. NATURE, 429(6988), 180.

FIRESTONE, S. M., WARD, M. P., CHRISTLEY, R. M., & DHAND, N. K. (2011). THE IMPORTANCE OF LOCATION IN CONTACT NETWORKS: DESCRIBING EARLY EPIDEMIC SPREAD USING SPATIAL SOCIAL NETWORK ANALYSIS. PREVENTIVE VETERINARY MEDICINE, 102(3), 185-195.

FISCHER, A., ARNOLD, A., & GIBBS, M. (1996). INFORMATION AND THE SPEED OF INNOVATION ADOPTION. AMERICAN JOURNAL OF AGRICULTURAL ECONOMICS, 78(4), 1073-1081.

FOURNIÉ, G., GUITIAN, J., DESVAUX, S., CUONG, V. C., PFEIFFER, D. U., MANGTANI, P., & GHANI, A. C. (2013). INTERVENTIONS FOR AVIAN INFLUENZA A (H5N1) RISK MANAGEMENT IN LIVE BIRD MARKET NETWORKS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 110(22), 9177-9182.

FRAITURE, C. D., WICHELNS, D., ROCKSTRÖM, J., KEMP-BENEDICT, E., ERIYAGAMA, N., GORDON, L. J., . . . KARLBERG, L. (2007). LOOKING AHEAD TO 2050: SCENARIOS OF ALTERNATIVE INVESTMENT APPROACHES.

FRIAS-MARTINEZ, E., WILLIAMSON, G., & FRIAS-MARTINEZ, V. (2011). AN AGENT-BASED MODEL OF EPIDEMIC SPREAD USING HUMAN MOBILITY AND SOCIAL NETWORK INFORMATION. PAPER PRESENTED AT THE PRIVACY, SECURITY, RISK AND TRUST (PASSAT) AND 2011 IEEE

THIRD INERNATIONAL CONFERENCE ON SOCIAL COMPUTING (SOCIALCOM), 2011 IEEE THIRD INTERNATIONAL CONFERENCE ON.

FUENTES, M., & KUPERMAN, M. (1999). CELLULAR AUTOMATA AND EPIDEMIOLOGICAL MODELS WITH SPATIAL DEPENDENCE. PHYSICA A: STATISTICAL MECHANICS AND ITS APPLICATIONS, 267(3), 471-486.

GASSERT, F., M. LANDIS, M. LUCK, P. REIG, AND T. SHIAO. (2014). AQUEDUCT GLOBAL MAPS 2.1. RETRIEVED FROM: HTTP://WWW.WRI.ORG/PUBLICATION/AQUEDUCT-METADATA-GLOBAL

GHANI, A. C., & GARNETT, G. P. (2000). RISKS OF ACQUIRING AND TRANSMITTING SEXUALLY TRANSMITTED DISEASES IN SEXUAL PARTNER NETWORKS. SEXUALLY TRANSMITTED DISEASES, 27(10), 579-587.

GRANOVETTER, M. (1978). THRESHOLD MODELS OF COLLECTIVE BEHAVIOR. AMERICAN JOURNAL OF SOCIOLOGY, 1420-1443.

GRASSBERGER, P. (1983). ON THE CRITICAL BEHAVIOR OF THE GENERAL EPIDEMIC PROCESS AND DYNAMICAL PERCOLATION. MATHEMATICAL BIOSCIENCES, 63(2), 157-172.

GRIMMETT, G. (1999). WHAT IS PERCOLATION? : SPRINGER.

HANSON, S., NICHOLLS, R., RANGER, N., HALLEGATTE, S., CORFEE-MORLOT, J., HERWEIJER, C., & CHATEAU, J. (2011). A GLOBAL RANKING OF PORT CITIES WITH HIGH EXPOSURE TO CLIMATE EXTREMES. CLIMATIC CHANGE, 104(1), 89-111.

HARKO, T., LOBO, F. S., & MAK, M. (2014). EXACT ANALYTICAL SOLUTIONS OF THE SUSCEPTIBLE-INFECTED-RECOVERED (SIR) EPIDEMIC MODEL AND OF THE SIR MODEL WITH EQUAL DEATH AND BIRTH RATES. APPLIED MATHEMATICS AND COMPUTATION, 236, 184-194.

HEESTERBEEK, J. (2000). MATHEMATICAL EPIDEMIOLOGY OF INFECTIOUS DISEASES: MODEL BUILDING, ANALYSIS AND INTERPRETATION (VOL. 5): JOHN WILEY & SONS.

HEIKKINEN, N. (2014). BIRTH CONTROL COULD HELP THE ENVIRONMENT, BUT NOT QUICKLY. SCIENTIFIC AMERICAN.

HERRERO, M., HAVLÍK, P., VALIN, H., NOTENBAERT, A., RUFINO, M. C., THORNTON, P. K., . . . OBERSTEINER, M. (2013). BIOMASS USE,

PRODUCTION, FEED EFFICIENCIES, AND GREENHOUSE GAS EMISSIONS FROM GLOBAL LIVESTOCK SYSTEMS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 110(52), 20888-20893.

HETHCOTE, H. W. (2000). THE MATHEMATICS OF INFECTIOUS DISEASES. SIAM REVIEW, 42(4), 599-653.

IPCC. (2014). CLIMATE CHANGE 2014: IMPACTS, ADAPTATION, AND VULNERABILITY. PART A: GLOBAL AND SECTORAL ASPECTS. CONTRIBUTION OF WORKING GROUP II TO THE FIFTH ASSESSMENT REPORT OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. (V. R. B. C.B. FIELD, D.J. DOKKEN, K.J. MACH, M.D. MASTRANDREA, T.E. BILIR, M. CHATTERJEE, K.L. EBI, Y.O. ESTRADA, R.C. GENOVA, B. GIRMA, E.S. KISSEL, A.N. LEVY, S. MACCRACKEN, P.R. MASTRANDREA, L.L. WHITE ED.). CAMBRIDGE, UNITED KINGDOM AND NEW YORK, NY, USA: CAMBRIDGE UNIVERSITY PRESS.

JANSSEN, M. A., & OSTROM, E. (2006). EMPIRICALLY BASED, AGENT-BASED MODELS. ECOLOGY AND SOCIETY, 11(2), 37.

KEATING, B. A., HERRERO, M., CARBERRY, P. S., GARDNER, J., & COLE, M. B. (2014). FOOD WEDGES: FRAMING THE GLOBAL FOOD DEMAND AND SUPPLY CHALLENGE TOWARDS 2050. GLOBAL FOOD SECURITY, 3(3), 125-132.

KEELING, M., WOOLHOUSE, M., MAY, R., DAVIES, G., & GRENFELL, B. (2003). MODELLING VACCINATION STRATEGIES AGAINST FOOT-AND-MOUTH DISEASE. NATURE, 421(6919), 136-142.

KEELING, M. J. (1999). THE EFFECTS OF LOCAL SPATIAL STRUCTURE ON EPIDEMIOLOGICAL INVASIONS. PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON B: BIOLOGICAL SCIENCES, 266(1421), 859-867.

KEELING, M. J., & EAMES, K. T. (2005). NETWORKS AND EPIDEMIC MODELS. JOURNAL OF THE ROYAL SOCIETY INTERFACE, 2(4), 295-307.

KERMACK, W. O., & MCKENDRICK, A. G. (1927). A CONTRIBUTION TO THE MATHEMATICAL THEORY OF EPIDEMICS. PAPER PRESENTED AT THE PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON A: MATHEMATICAL, PHYSICAL AND ENGINEERING SCIENCES.

Kinzel, W. (1985). Phase transitions of cellular automata. Zeitschrift für Physik B Condensed Matter, 58(3), 229-244.

Koetse, M. J., & Rietveld, P. (2009). The impact of climate change and weather on transport: An overview of empirical findings. Transportation Research Part D: Transport and Environment, 14(3), 205-221.

Konar, M., Dalin, C., Suweis, S., Hanasaki, N., Rinaldo, A., & Rodriguez-Iturbe, I. (2011). Water for food: The global virtual water trade network. Water Resources Research, 47(5), n/a-n/a. doi:10.1029/2010WR010307

Kummu, M., Fader, M., Gerten, D., Guillaume, J. H., Jalava, M., Jägermeyr, J., . . . Varis, O. (2017). Bringing it all together: Linking measures to secure nations' food supply. Current opinion in environmental sustainability, 29, 98-117.

Lemmen, D. S., & Warren, F. J. (2004). Climate change impacts and adaptation: a Canadian perspective.

Liu, J., Mooney, H., Hull, V., Davis, S. J., Gaskell, J., Hertel, T., . . . Li, S. (2015). Systems integration for global sustainability. Science, 347(6225). doi:10.1126/science.1258832

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate Trends and Global Crop Production Since 1980. Science, 333(6042), 616-620. doi:10.1126/science.1204531

Lowe, J., Gauger, P., Harmon, K., Zhang, J., Connor, J., Yeske, P., . . . Main, R. (2014). Role of transportation in spread of porcine epidemic diarrhea virus infection, United States. Emerging infectious diseases, 20(5), 872.

Melillo, J. M., Richmond, T., & Yohe, G. (2014). Climate change impacts in the United States. Third National Climate Assessment.

Montgomery, M. R. (2008). The Urban Transformation of the Developing World. Science, 319(5864), 761-764. doi:10.1126/science.1153012

MOORE, C., & NEWMAN, M. E. (2000). EPIDEMICS AND PERCOLATION IN SMALL-WORLD NETWORKS. PHYSICAL REVIEW E, 61(5), 5678.

NACHTERGAELE, F., PETRI, M., BIANCALANI, R., VAN LYNDEN, G., VAN VELTHUIZEN, H., & BLOISE, M. (2011). AN INFORMATION DATABASE FOR LAND DEGRADATION ASSESSMENT AT GLOBAL LEVEL. GLOBAL LAND DEGRADATION INFORMATION SYSTEM (GLADIS).

NASA. (2017). NASA EARTH EXCHANGE GLOBAL DAILY DOWNSCALED PROJECTIONS (NEX-GDDP). RETRIEVED FROM: HTTPS://CDS.NCCS.NASA.GOV/NEX-GDDP/

NATIONS, U. (2015). WORLD POPULATION PROSPECTS: KEY FINDINGS AND ADVANCE TABLES. RETRIEVED FROM HTTPS://ESA.UN.ORG/UNPD/WPP/PUBLICATIONS/FILES/KEY_FINDINGS_ WPP_2015.PDF

NELSON, G. C., ROSEGRANT, M. W., KOO, J., ROBERTSON, R., SULSER, T., ZHU, T., . . . BATKA, M. (2009). CLIMATE CHANGE: IMPACT ON AGRICULTURE AND COSTS OF ADAPTATION (VOL. 21): INTL FOOD POLICY RES INST.

NEWMAN, M. E. J., GIRVAN, M., & DOYNE FARMER, J. (2002). OPTIMAL DESIGN, ROBUSTNESS, AND RISK AVERSION. PHYSICAL REVIEW LETTERS, 89(2), 028301.

O'DEA, E. B., SNELSON, H., & BANSAL, S. (2015). STATE-LEVEL TRANSPORT FLOWS ARE PREDICTIVE OF THE DYNAMICS OF PORCINE EPIDEMIC DIARRHEA VIRUS. BIORXIV, 017178.

OHTSUKI, T., & KEYES, T. (1986). KINETIC GROWTH PERCOLATION: EPIDEMIC PROCESSES WITH IMMUNIZATION. PHYSICAL REVIEW A, 33(2), 1223.

OSTROM, E. (2010). POLYCENTRIC SYSTEMS FOR COPING WITH COLLECTIVE ACTION AND GLOBAL ENVIRONMENTAL CHANGE. GLOBAL ENVIRONMENTAL CHANGE, 20(4), 550-557. DOI:HTTPS://DOI.ORG/10.1016/J.GLOENVCHA.2010.07.004

OTTO, I. M., BIEWALD, A., COUMOU, D., FEULNER, G., KOHLER, C., NOCKE, T., . . . BECK, U. (2015). SOCIO-ECONOMIC DATA FOR GLOBAL

ENVIRONMENTAL CHANGE RESEARCH. NATURE CLIM. CHANGE, 5(6), 503-506. DOI:10.1038/NCLIMATE2593

PACALA, S., & SOCOLOW, R. (2004). STABILIZATION WEDGES: SOLVING THE CLIMATE PROBLEM FOR THE NEXT 50 YEARS WITH CURRENT TECHNOLOGIES. SCIENCE, 305(5686), 968-972.

PARRY, M., ROSENZWEIG, C., IGLESIAS, A., FISCHER, G., & LIVERMORE, M. (1999). CLIMATE CHANGE AND WORLD FOOD SECURITY: A NEW ASSESSMENT. GLOBAL ENVIRONMENTAL CHANGE, 9(SUPPLEMENT 1), S51-S67. DOI:HTTPS://DOI.ORG/10.1016/S0959-3780(99)00018-7

PARRY, M., ROSENZWEIG, C., & LIVERMORE, M. (2005). CLIMATE CHANGE, GLOBAL FOOD SUPPLY AND RISK OF HUNGER. PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B: BIOLOGICAL SCIENCES, 360(1463), 2125-2138. DOI:10.1098/RSTB.2005.1751

PASTOR-SATORRAS, R., & VESPIGNANI, A. (2001). EPIDEMIC SPREADING IN SCALE-FREE NETWORKS. PHYSICAL REVIEW LETTERS, 86(14), 3200.

PEREZ, L., & DRAGICEVIC, S. (2009). AN AGENT-BASED APPROACH FOR MODELING DYNAMICS OF CONTAGIOUS DISEASE SPREAD. INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS, 8(1), 50.

PFEIFER, B., KUGLER, K., TEJADA, M. M., BAUMGARTNER, C., SEGER, M., OSL, M., . . . WURZ, M. (2008). A CELLULAR AUTOMATON FRAMEWORK FOR INFECTIOUS DISEASE SPREAD SIMULATION. THE OPEN MEDICAL INFORMATICS JOURNAL, 2, 70.

PUGA, D., & VENABLES, A. J. (1996). THE SPREAD OF INDUSTRY: SPATIAL AGGLOMERATION IN ECONOMIC DEVELOPMENT. JOURNAL OF THE JAPANESE AND INTERNATIONAL ECONOMIES, 10(4), 440-464.

RAHMANDAD, H., & STERMAN, J. (2008). HETEROGENEITY AND NETWORK STRUCTURE IN THE DYNAMICS OF DIFFUSION: COMPARING AGENT-BASED AND DIFFERENTIAL EQUATION MODELS. MANAGEMENT SCIENCE, 54(5), 998-1014.

RANGANATHAN, J., VENNARD, D., WAITE, R., DUMAS, P., LIPINSKI, B., & SEARCHINGER, T. (2016). SHIFTING DIETS FOR A SUSTAINABLE FOOD FUTURE. INSTALLMENT 11 OF "CREATING A SUSTAINABLE FOOD FUTURE".

RESCHER, N. (1998). PREDICTING THE FUTURE: AN INTRODUCTION TO THE THEORY OF FORECASTING: SUNY PRESS.

ROGERS, E. M. (2010). DIFFUSION OF INNOVATIONS: SIMON AND SCHUSTER.

ROSENZWEIG, C., & PARRY, M. L. (1994). POTENTIAL IMPACT OF CLIMATE CHANGE ON WORLD FOOD SUPPLY. NATURE, 367(6459), 133-138.

RYAN, B., & GROSS, N. C. (1943). THE DIFFUSION OF HYBRID SEED CORN IN TWO IOWA COMMUNITIES. RURAL SOCIOLOGY, 8(1), 15.

SADRAS, V. (2015). YIELD GAP ANALYSIS OF FIELD CROPS: METHODS AND CASE STUDIES.

SANDER, L., WARREN, C., SOKOLOV, I., SIMON, C., & KOOPMAN, J. (2002). PERCOLATION ON HETEROGENEOUS NETWORKS AS A MODEL FOR EPIDEMICS. MATHEMATICAL BIOSCIENCES, 180(1), 293-305.

SCHMIDT, M., & LIPSON, H. (2013). EUREQA (VERSION 1.24)[SOFTWARE]. NUTONIAN, SOMERVILLE, MASS, USA.

SCOTT, D. M., NOVAK, D. C., AULTMAN-HALL, L., & GUO, F. (2006). NETWORK ROBUSTNESS INDEX: A NEW METHOD FOR IDENTIFYING CRITICAL LINKS AND EVALUATING THE PERFORMANCE OF TRANSPORTATION NETWORKS. JOURNAL OF TRANSPORT GEOGRAPHY, 14(3), 215-227.

SERRANO, M. Á., & BOGUNÁ, M. (2006). PERCOLATION AND EPIDEMIC THRESHOLDS IN CLUSTERED NETWORKS. PHYSICAL REVIEW LETTERS, 97(8), 088701.

SORNETTE, D. (2006). CRITICAL PHENOMENA IN NATURAL SCIENCES: CHAOS, FRACTALS, SELFORGANIZATION AND DISORDER: CONCEPTS AND TOOLS: SPRINGER.

STAUFFER, D., & AHARONY, A. (1994). INTRODUCTION TO PERCOLATION THEORY: CRC PRESS.

STORN, R., & PRICE, K. (1997). DIFFERENTIAL EVOLUTION–A SIMPLE AND EFFICIENT HEURISTIC FOR GLOBAL OPTIMIZATION OVER CONTINUOUS SPACES. JOURNAL OF GLOBAL OPTIMIZATION, 11(4), 341-359.

SULLIVAN, J., AULTMAN-HALL, L., & NOVAK, D. (2009). A REVIEW OF CURRENT PRACTICE IN NETWORK DISRUPTION ANALYSIS AND AN

ASSESSMENT OF THE ABILITY TO ACCOUNT FOR ISOLATING LINKS IN TRANSPORTATION NETWORKS. TRANSPORTATION LETTERS, 1(4), 271-280.

UK COMMITTEE ON CLIMATE CHANGE. (2017). UK CLIMATE CHANGE RISK ASSESSMENT 2017. RETRIEVED FROM HTTPS://WWW.THECCC.ORG.UK/TACKLING-CLIMATE-CHANGE/PREPARING-FOR-CLIMATE-CHANGE/UK-CLIMATE-CHANGE-RISK-ASSESSMENT-2017/.

UNFAO. (2017). AGRO-MAPS. RETRIEVED FROM: HTTP://KIDS.FAO.ORG/AGROMAPS/

UNFAO. (2017). FOOD BALANCE SHEETS DATA. RETRIEVED FROM: HTTP://WWW.FAO.ORG/FAOSTAT/EN/#DATA/FBS

UNFAO. (2017). GLOBAL INFORMATION AND EARLY WARNING SYSTEM, FOOD PRICE MONITORING AND ANALYSIS TOOL. RETRIEVED FROM: HTTP://WWW.FAO.ORG/GIEWS/FOOD-PRICES/TOOL/PUBLIC/#/HOME

UNFAO LADA. (2017). GLOBAL LAND DEGRADATION INFORMATION SYSTEM (GLADIS) DATABASE. RETRIEVED FROM: HTTP://WWW.FAO.ORG/NR/LADA/GLADIS/GLADIS_DB/

UNHCR. (2012). THE STATE OF THE WORLD'S REFUGEES: IN SEARCH OF SOLIDARITY. OXFORD UNIVERSITY PRESS RETRIEVED FROM HTTP://WWW.UNHCR.ORG/EN-US/PUBLICATIONS/SOWR/4FC5CECA9/STATE-WORLDS-REFUGEES-2012-SEARCH-SOLIDARITY.HTML.

UNITED STATES DEPARTMENT OF AGRICULTURE: ANIMAL AND PLANT HEALTH INSPECTION SERVICE. (2006). REFERENCE OF SWINE HEALTH AND HEALTH MANAGEMENT PRACTICES IN THE UNITED STATES.

US CENSUS BUREAU. (2012). COMMODITY FLOW SURVEY, 2012 DATA. RETRIEVED FROM: HTTPS://WWW.CENSUS.GOV/ECON/CFS/

USAID. (2016). DEMOGRAPHIC AND HEALTH SURVEYS (DHS) DATA. RETRIEVED FROM: HTTP://DHSPROGRAM.COM/DATA/AVAILABLE-DATASETS.CFM

VAN ITTERSUM, M. K., CASSMAN, K. G., GRASSINI, P., WOLF, J., TITTONELL, P., & HOCHMAN, Z. (2013). YIELD GAP ANALYSIS WITH LOCAL TO GLOBAL RELEVANCE—A REVIEW. FIELD CROPS RESEARCH, 143, 4-17.

VESPIGNANI, A. (2012). MODELLING DYNAMICAL PROCESSES IN COMPLEX SOCIO-TECHNICAL SYSTEMS. NATURE PHYSICS, 8(1), 32.

WATTS, D. J., MUHAMAD, R., MEDINA, D. C., & DODDS, P. S. (2005). MULTISCALE, RESURGENT EPIDEMICS IN A HIERARCHICAL METAPOPULATION MODEL. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 102(32), 11157-11162.

WEST, P. C., GERBER, J. S., ENGSTROM, P. M., MUELLER, N. D., BRAUMAN, K. A., CARLSON, K. M., . . . SIEBERT, S. (2014). LEVERAGE POINTS FOR IMPROVING GLOBAL FOOD SECURITY AND THE ENVIRONMENT. SCIENCE, 345(6194), 325-328. DOI:10.1126/SCIENCE.1246067

WHEELER, T., & VON BRAUN, J. (2013). CLIMATE CHANGE IMPACTS ON GLOBAL FOOD SECURITY. SCIENCE, 341(6145), 508-513. DOI:10.1126/SCIENCE.1239402

WILSON, K. J., FALKINGHAM, J., MELLING, H., & DE ABREU, R. (2004). SHIPPING IN THE CANADIAN ARCTIC: OTHER POSSIBLE CLIMATE CHANGE SCENARIOS. PAPER PRESENTED AT THE GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 2004. IGARSS'04. PROCEEDINGS. 2004 IEEE INTERNATIONAL.

WORLD BANK. (2016). SUBNATIONAL POPULATION DATABASE. RETRIEVED FROM: HTTPS://DATA.WORLDBANK.ORG/DATA-CATALOG/SUBNATIONAL-POPULATION

WU, T., BLACKHURST, J., & O'GRADY, P. (2007). METHODOLOGY FOR SUPPLY CHAIN DISRUPTION ANALYSIS. INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH, 45(7), 1665-1682.

CHAPTER 1: USING AN AGENT-BASED MODEL TO EVALUATE THE EFFECT OF PRODUCER SPECIALIZATION ON THE EPIDEMIOLOGICAL RESILIENCE OF LIVESTOCK PRODUCTION NETWORKS

3.1 Abstract

An agent-based computer model that builds representative regional U.S. hog production networks was developed and employed to assess the potential impact of the ongoing trend towards increased producer specialization upon network-level resilience to catastrophic disease outbreaks. Empirical analyses suggest that the spatial distribution and connectivity patterns of contact networks often predict epidemic spreading dynamics. Our model heuristically generates realistic systems composed of hog producer, feed mill, and slaughter plant agents. Network edges are added during each run as agents exchange livestock and feed. The heuristics governing agents' contact patterns account for factors including their industry roles, physical proximities, and the age of their livestock. In each run, an infection is introduced, and may spread according to probabilities associated with the various modes of contact. For each of three treatments—defined by one-phase, two-phase, and three-phase production systems—a parameter variation experiment examines the impact of the spatial density of producer agents in the system upon the length and size of disease outbreaks. Resulting data show phase transitions whereby, above some density threshold, systemic outbreaks become possible, echoing findings from percolation theory. Data analysis reveals that multi-phase production systems are vulnerable to catastrophic outbreaks at lower spatial densities, have more abrupt percolation transitions, and are characterized

by less-predictable outbreak scales and durations. Key differences in network-level metrics shed light on these results, suggesting that the absence of potentially-bridging producer–producer edges may be largely responsible for the superior disease resilience of single-phase "farrow to finish" production systems.

3.2 Introduction

There is widespread agreement among livestock veterinarians and epidemiologists that mitigating disease outbreaks is critical to promote food safety, maintain food availability, and reduce economic risk in the marketplace. Livestock epidemiologists commonly focus on promoting the adoption of discrete biosecurity measures, such as truck wash facilities and biocontainment procedures at individual premises [1]. However, empirical research has increasingly revealed the importance of understanding how the structures of trade and transportation networks can aid in predicting and preventing outbreaks [2-4]. In light of these observations, more work is clearly needed to understand how livestock biosecurity may be bolstered from a systems perspective.

Recent years have seen significant structural changes within the U.S. hog industry, with a marked trend toward increased producer specialization. Whereas in the past it was standard practice for a single producer to take a pig "from farrow to finish," it is now increasingly common for livestock to be housed at two or, more recently, three different producer operations throughout their lives, with each operation specializing in

a specific life cycle stage [5]. While operational efficiency advantages may be gained through increased specialization, its effect on disease spread is not fully understood.

To address this question, we developed an agent-based susceptible / infective (SI) computer model to simulate epidemiological events in hog production networks. Agent-based models (ABMs) have been used extensively to analyze complex phenomena that emerge from the relatively-simple actions of a cohort of interacting individuals [6-8]. Using the 2013 PEDv outbreak as a case study, three key mechanisms facilitating disease transmission in the hog industry were identified: (a) the transfer of infected animals between premises, (b) deliveries of contaminated feed, and (c) contaminated livestock and feed transportation equipment [9, 10]. Contact between producers, slaughter plants, and feed mills was found to be largely responsible for spreading the virus. In our ABM, structured populations of these three agent types are placed in the simulation, an infection is introduced randomly, and decision heuristics define how and when agents come into contact, potentially transmitting the infection. Using this model, we report on a series of parameter variation experiments that investigate the epidemic spread characteristics resulting from varied levels of producer specialization and numbers of producers in the system, finding evidence of percolation dynamics, with increased specialization leading to significantly diminished epidemiological resilience.

3.2.1 Percolation Theory

The "robust yet fragile" nature that describes a diversity of complex systems offers a useful framework to understand the spread of diseases through networks. Pathogens are regularly introduced with little consequence, but due to stochasticity and internal heterogeneities in contact network structures, a single pathogen can occasionally ignite a widespread epidemic [11-16]. Percolation is the mathematical phase-change that occurs when the density of entities in a system becomes sufficient that the expected outbreak magnitude no longer scales linearly with each added node, but instead accelerates rapidly toward near-complete spreading. The point at which this transition occurs—the percolation threshold—is defined as the density at which, in an infinite network, the expected size of the "giant component" is also infinite [17-22]. While much of the work in this area has concerned itself with analytically-formalizing percolation behavior in relatively-simple systems, the insights gained through such investigations are relevant for understanding dynamical regimes in complex real-world systems as well. To investigate percolation in our experimental results, we numerically assess how the size and duration of epidemic events in a series of model runs scale with the addition of producer nodes.

3.2.2 Models of Epidemic Spreading

At the core of many model-based inquiries into disease spread is the susceptible / infective / recovered (SIR) framework. In SIR models, an infective individual may spread a disease to susceptible individuals for a period of time, after which the infective

individual transitions to the recovered state and cannot be reinfected [23]. The SI

framework, a common SIR derivative, allows for repeated infections. SIR models based

on differential equations (DEs) have successfully replicated many of the complex

temporal patterns typical of epidemics [24-26]. However, the DE approach has been

criticized, as it implicitly posits both a homogeneous population and complete

mixing [27]. Structured population models partially overcome these shortfalls, defining

heterogeneous distributions for parameters such as age, size, spatial position, and

movement [28, 29]. Cellular Automata (CA) SIR models additionally reproduce spatial

phenomena such as waves of infection radiating outward from a source [30, 31], and

can also produce percolation-type phase transitions [32], providing insights into the

impact of agents' relative spatial positions on spreading dynamics [33, 34]. More

advanced CA models examine the impact of heterogeneous susceptibilities,

transmission rates, and infectious periods [35], as well as modulating parameters as a

function of spatial proximity [36, 37].

In recent years, epidemiological modelers have increasingly investigated the

role of the structural topology of "mixing networks" on disease percolation [38]. SIR

simulation studies on complex networks demonstrate the impact of degree distribution

on the speed, size, and variability of epidemic events, with more heterogeneously-

distributed networks pushing the percolation threshold towards zero as $N \rightarrow \infty$ [39].

Epidemics on scale-free networks cascade from highly-connected hubs through smaller

degree classes [40], although a sharp percolation threshold is not observed [41, 42]. In

metapopulation (or subpopulation) networks, surpassing critical values for the rate of spreading between subpopulations can trigger percolation, shedding light on the mechanism behind the "robust yet fragile" nature of these systems [43-45].

Coelho, Cruz, and Codeço [46] characterize the complexification of epidemiological simulations over time as a shift between "strategic models" that explore the fundamental features of epidemics, to "tactical models" that mirror the conditions within which a real-world epidemic may unfold. Agent-based models are often examples of the latter, generating empirically-calibrated networks of interacting agents that are heterogeneous not only in their parameter values, but also in the behavioral heuristics that govern how and when contact occurs [47]. Modelers can hard-code agents' positions or spatial distributions using a GIS framework [48], and may incorporate empirical data—such as airline routes or telephone records [49, 50]; or, in the case of livestock epidemics, the operational details and locations of farms [51, 52]—that have been shown to correlate with outbreak patterns. Other modelers let networks emerge organically during each model run as a result of agents' decision-making heuristics. Using the latter methodology, Ghani & Garnett [53] found network centrality measures that predicted an agent's chance of either getting or spreading a sexually-transmitted disease. Eubank et al. [54] developed an ABM that utilized heuristics parameterized from large-scale datasets to generate realistic urban social contact networks and identified resulting epidemiological vulnerabilities. Gojovic et al. [55] implemented a model to evaluate optimal immunization strategies during the

2009 H1N1 pandemic, using demographic and employment records to assign agent

parameters, and incorporating differential transmission probabilities for multiple

contexts. Keeling et al. [56] developed a model of U.K. farms—parameterized via

census data—and performed a Monte Carlo simulation to understand how factors

including agent heterogeneities and movement restrictions explain the observed spread

of the 2001 UK Foot and Mouth epidemic. The ABM we have developed for this

experiment builds on prior work in this area, leveraging empirical data to heuristically

generate hog production systems that are structurally-parallel to real-world examples,

and encoding behavioral rules in collaboration with industry experts that allow the

contact networks underlying disease spread to emerge organically in each model run.

3.2.3 Network Analytics and Epidemiological Vulnerability

Bailey [43] was among the first to suggest that, while contagions are generally

confined to small network clusters, global epidemics may result when an edge forms a

bridge between clusters. This pattern has been empirically observed in both human and

livestock epidemics. Firestone et al. [2] analyzed infected premises during the 2007

Equine Influenza outbreak in Australia, finding strong evidence that the movement of

infected horses between spatially-clustered groups of premises correlated with the

spread of the disease. Fournié et al. [3] investigated the network connectivity patterns

of agents involved in Vietnam's live bird markets during an H5N1 influenza outbreak,

concluding that the contact network could have been largely "disconnected" by

focusing on disinfection of transportation equipment at a few of the large hubs. The

structure of transportation networks for feed and livestock was found to be the primary factor underlying the 2013 U.S. hog industry PEDv epidemic, with slaughter plants and feed mills serving as the primary hubs [4].

Network theorists have conducted several investigations into the relationships between metrics describing a network's structure, and the propensity of that network to promote or inhibit spreading [57, 58]. Experiment two below evaluates six network metrics that may impact epidemiological resilience. As a baseline, we investigate whether—as would be expected—hog production networks with higher average degree $\langle k \rangle$ promote disease spread. Small-world graphs—a class into which many real-world livestock production networks fall [59-61]—are characterized by a smaller mean shortest path length $\langle l \rangle$ and greater clustering C than corresponding random graphs with equivalent $\langle k \rangle$ [62]. Research shows that spreading in small-world networks (versus corresponding random graphs) proceeds more rapidly but results in fewer infected nodes, and also that small-world networks exhibit significantly higher k-core densities [63]. Other studies have found that k-core boundaries often define the part of a graph in which a spreading event is more likely to persist [64]. Experiment two thus analyzes the role of k-core size S_{kc}, in addition to k-core order O_{kc}, and number of k-cores N_{kc}. While there is debate over whether weighted versus unweighted—as well as directed versus symmetrized—versions of network metrics are more appropriate, at least in some contexts, metrics calculated on unweighted, symmetrized graphs best

70

predict epidemiological vulnerability [65]. In light of this, we opt to binarize and symmetrize the graphs in our analysis.

3.2.4 Research Questions and Hypotheses

This study uses an agent-based model to investigate the impact of two ongoing network-structural trends in the U.S. hog industry upon system-level epidemiological resilience. The first of these trends is the growing spatial density of networks. With more potential trading partners from which to choose, the average degree of the network will tend to increase, which should correlate positively with outbreak severity. The second trend, increasing producer specialization, will necessarily add producer–producer edges where previously there were none. Following the empirical and computational studies cited above, we can hypothesize that a greater probability of large-scale "global" epidemics should be observed wherever the network typology is such that, without these additional edges acting as bridges, the contagion would have been isolated within localized clusters. By systematically varying both the spatial density of the simulated networks, along with the level of producer specialization, the experiments described below investigate how the interplay between these two factors may render a system more or less vulnerable to catastrophic disease outbreaks.

3.3 Materials and Methods

3.3.1 Model Design Concepts

The Regional U.S. Hog Production Network Biosecurity Model (RUSH-PNBM) v.0.8 generates simulated hog production systems composed of producer,

slaughter plant, and feed mill agents. Heterogeneous parameter values, multiple interaction contexts, and spatial proximity considerations, are incorporated into agents' decision heuristics, and determine contact patterns and infection spread potentials. The epidemiological submodel is of the SI type, since, in the case of PEDv, reinfections of the same premises have been reported [66].

RUSH-PNBM is a "tactical" model, in that it is empirically-calibrated to mirror a real-world system, but it also aims to avoid being overly context-specific, leaving sufficient flexibility to analyze a variety of scenarios [46]. To this end, elements that were deemed significant facilitators of disease spread by a cohort of industry consultants were included in the model, while many extraneous and/or uncertain details were bracketed [67]. The baseline parameterization—although ground-truthed by our advisors as well as several datasets—is not meant to project the course of a specific infectious agent through any real-world production network, but rather to facilitate a workable and reasonably-realistic simulation useful for understanding the network trends important for this experiment.

The model was developed using AnyLogic 7 software, with all functions written in Java. The sections below provide an overview of initialization procedures, agents' behavioral heuristics, and parameter calibration methods. For full implementation details, see S1 Protocol.

3.3.2 Agent Initialization

At model initialization, all agents are assigned a fixed location stochastically within a continuous two-dimensional spatial framework defined by an 880 x 490 unit rectangle, with units representing kilometers. Producer agents are assigned one of five industry roles (see Figure 3) according to distributions corresponding to the treatment scenario (see Experimental design section). The livestock capacity of each producer agent is assigned by drawing from a normal distribution (Table 1 gives parameter values). Producer agents may have one or several batches of pigs, with each batch considered to be the same age. Each producer begins at full capacity, with the age of pig batches drawn from a uniform distribution corresponding to the industry role of the producer.
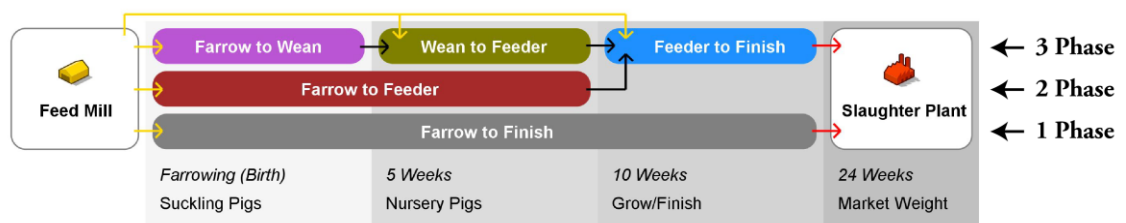
Figure 3: Structure of agent connections in the model. Shows 1-phase (low-specialization), 2-phase (mid-specialization), and 3-phase (high-specialization) connectivity heuristics. Also indicates livestock transfer age conditions.

Table 1: Parameter values used in the experiment.

Parameter	Value(s) Explored
Network parameters	
Area of network region (km2)	431200
Number of producer agents in the model	[10:10:1500]
Number of livestock per producer [normal distribution; rounded to integer]	μ=1,000; σ=300; $x{\geq}50$
Number of producer production phases	[1, 2, 3]
Number of slaughter plant agents in the model	3
Number of feed mill agents in the model	10
Epidemiological parameters	
Suckling pig mortality rate	0.95
Nursery pig mortality rate	0.6
Grow/finish hog mortality rate	0.1
Length of producer infection (days) [triangular distribution]	μ=30; $0{\leq}x{\leq}60$
Length of slaughter plant contamination (days) [triangular distribution]	μ=5; $0{\leq}x{\leq}10$
Farrowing parameters	
Frequency of farrowing (days)	30
Minimum farrowing quantity as a proportion of producer capacity	0.25
Producer–producer parameters	
Maximum producer to producer connection distance (km)	100
Minimum transfer quantity as a proportion of transferee capacity	0.25
Prob. of infection via trailer returning from infected transferee	0.15
Feed mill–producer parameters	
Frequency of feed distribution trips (days)	1
Percent of producers in feed mill service area visited per trip	15
Probability that truck will be contaminated upon visiting an infected producer	0.15
Probability that contaminated truck will infect subsequent producers on route	0.15
Producer–slaughter plant parameters	
Probability that infected hogs will contaminate slaughter plant receiving area	0.75
Probability of infection via truck returning from infected slaughter plant	0.15

3.3.3 Network Initialization

Network edges in the model are subdivided into three contexts: (a) *producer–producer*, (b) *producer–slaughter plant*, and (c) *feed mill–producer*. The basic structure of connections between agents of each industry role is visualized in Figure 3. A network initialization function generates a set of potential trading partners for each agent, defining the possible edges across which contact may occur during the remainder of the model run. Producer agents are each assigned to their most proximal feed mill, and finishing producers are also assigned to their most proximal slaughter plant. Each non-finishing producer is assigned a pool of potential transferee producers of the appropriate industry role for outgoing shipments (see Figure 3), and within a maximum distance of 100km. Figure 4 shows a sample network as displayed on the model dashboard, and briefly describes the heuristics associated with each agent class.

Feed mill agents:
- Periodically generate delivery routes
- Light yellow edges represent service areas
- Dark yellow edges represent active *feed mill–producer* delivery routes

Producer agents:
- Transfer livestock to another producer or slaughter plant at appropriate age
- Icon color corresponds to industry role
- Icon size corresponds to livestock capacity
- Gray edges represent possible trading partners
- Dark black edges represent active *producer–producer* livestock transfers

Slaughter plant agents:
- Periodically receive market-weight hogs
- Light red edges represent service areas
- Dark red edges represent active *producer–slaughter plant* livestock transfers

Red outer circle indicates infection

Figure 4: Sample network map as displayed on the model dashboard. Shows agents as nodes and inter-agent contacts (both potential and active) as edges. Key provides an overview of connectivity heuristics for each agent type.

3.3.4 Initial Infection

The initial infection event is triggered after one model year, to skip the transient period and allow the simulation to stabilize. At this point, a single producer agent is selected randomly and transitioned to the infected state.

3.3.5 Behavioral Heuristics

The major functions controlling agent behavior, network connectivity, and infection transmissibility, are detailed below. Specific parameter values appear in Table 1.

.

3.3.5.1 Farrowing

Farrowing producers (where piglets are born) periodically replenish their inventories by generating new pig batches. New batches are the size of the producer's

76

spare capacity. A minimum farrowing quantity parameter ensures reasonably-sized batches.

3.3.5.2 Producer–producer livestock transfers

Non-finishing producers transfer pig batches that have reached the transfer age corresponding to their industry role to an appropriate transferee (see Figure 3). A minimum transfer size requirement ensures realistically-sized shipments between producers. Transferee producers are sequentially evaluated in order of proximity until a producer with sufficient excess livestock capacity is identified, at which point the pig batch meeting the transfer age requirement is deleted from the transferring producer's stock and added to the transferee's. If the transferring producer is infected, the infected livestock will spread the disease to the transferee. If the transferee producer is infected but the transferring producer is not, the returning "delivery truck" may become contaminated and infect the transferring producer according to a set probability. If a producer becomes infected, the size of each of its pig batches is diminished by the mortality rate associated with the batch's age. Producers remain infected for a duration determined by a triangular distribution—an intuitive and reliable proxy for the beta distribution [68]—with a mean of 30 days.

3.3.5.3 Producer–slaughter plant livestock transfers

Finishing producers ship livestock to their slaughter plant as soon as a pig batch reaches the designated age. If the transferring producer is infected, the receiving area of the slaughter plant may become contaminated according to a set probability. If infected,

a slaughter plant will remain infected for a duration determined by a triangular distribution with a mean of 5 days. If the receiving area of the slaughter plant is already contaminated when a shipment arrives, the returning "delivery truck" may infect the shipping producer according to a set probability.

3.3.5.4 Feed mill–producer delivery routes

Every model day, each feed mill generates a delivery route by first selecting a producer agent within its service area at random. From this location, the nearest producer within the feed mill's service area that has not been visited becomes the next stop on the route, and this process is repeated until the "delivery truck" has visited the designated number of farms. Should the truck encounter an infected producer premises on its route, the truck may become contaminated according to a set probability. Once a truck is contaminated, the infection may spread to subsequent producers on the route according to a set probability.

3.3.6 Parameter Calibration

The structural makeup and contact patterns of the simulated hog industry network are based on several statistical datasets, as well as qualitative input from a cohort of experts including veterinarians, epidemiologists, and hog industry analysts. Distributions of producer livestock capacity and spatial density were generalized from USDA Census of Agriculture data [69], while slaughter plant density was generalized from USDA Food Safety and Inspection Service data [70]. Feed mill data proved more elusive, so parameterization was based primarily on expert estimates. Temporal aspects

of the simulation, such as the frequency with which contact events occur, were

generalized from a search of the primary literature coupled with industry expert

consultations.

Porcine Epidemic Diarrhea virus (PEDv) is a disease that swept through the U.S. hog

industry starting in 2013, causing widespread mortality and morbidity among

livestock [71]. This outbreak was used as a case study to calibrate the epidemiological

parameters of the model. In consultation with livestock veterinary professionals,

reasonable baseline values for parameters such as mortality rates for animals of different

ages were chosen. A series of parameter variation experiments were used to hone in on

baseline parameter values for infection probabilities and durations such that the infection

within the model spread in a manner similar to the patterns observed in the real-world

PEDv outbreak, for which tracking data are available [66]. These values were then

exogenized as baseline parameters that remained fixed across all experimental runs

(Table 1).

3.3.7 Experimental Design

Using the model detailed above, two experiments were performed. The first

explored disease percolation by varying the number of producers in the model. The

second explored the relationship between key network metrics and epidemiological

resilience.

The treatments differed according to the distribution of industry roles assigned

to producer agents. Aside from these producer classification assignments, all

parameters remained constant across all runs. The treatments are defined as follows

(see also Figure 3):

1. High specialization

    - Three-phase production system

    - Equal numbers of Farrow to Wean, Wean to Feeder, and Feeder to Finish

      producers

2. Medium specialization

    - Two-phase production system

    - Equal numbers of Farrow to Feeder and Feeder to Finish producers

3. Low specialization

    - One-phase production system

    - Farrow to Finish producers only

3.3.7.1 Experiment 1

Experiment one was a parameter variation experiment in which, for each

treatment scenario, the number of producer agents in the model (N_p) was varied

between 10 and 1,500 in increments of 10. Since the network region area is fixed,

varying N_p corresponds to a change in the spatial density of producers. For each

treatment, the model was executed 100 times at each of the 150 N_p values, for a total of

15,000 model runs for each specialization level, or 45,000 runs overall. In each run, the

model was stopped 4,135 model days after the initial infection; sufficient time for the

infection to either die out or become systemic. Each run generated two dependent

variable datapoints: (a) the overall duration of the infection event within the network as a whole, and (b) the proportion of agents that became infected at least once during the model run.

3.3.7.2 Experiment 2

In experiment two, features of production networks resulting from differential producer specialization were quantified and analyzed. Contact network data from 150 model runs for each specialization level—with N_p=100:100:1500, and 10 repetitions per parameterization—were exported from the model and analyzed as unweighted, undirected graphs. A similar edge list containing only the subset of nodes that had been infected during each model run was also stored. For both the contact network and the infected component network, six metrics were calculated using functions from the Python NetworkX library [72], with values for each metric plotted against N_p for each treatment. Table 2 gives Python and NetworkX code used in the analyses.

Table 2: Python / NetworkX Code Used in Experiment Two.

Network Metric	Python / NetworkX Code
Average Degree $\langle k \rangle$	sum(G.degree().values()) / len(G.nodes())
Average Shortest Path Length $\langle l \rangle$	nx.average_shortest_path_length(G)
Clustering Coefficient C	nx.average_clustering(G)
k-core Order O_{kc}	nx.degree(k_core(G), nbunch = k_core(G).nodes()[1])
k-core Size S_{kc}	len(nx.k_core(G).nodes())
Number of k-cores N_{kc}	nx.number_connected_components(nx.k_core(G))

Note: Python 2.7 and NetworkX 2.0 were used for all analyses. The prefix "nx.*" indicates a NetworkX function. "G" represents the NetworkX graph object to be analyzed.

3.4 Results

3.4.1 Experiment 1

As an initial step to examine the model output data, histograms were produced to visualize the distribution of outbreak sizes (the proportion of agents infected) and overall infection durations (the time between the initial infection and the last agent recovering from infection, in model days); with data stratified into three N_p ranges (Figure 5). These plots show that the distribution of infection severity—especially in the high-N_p runs—is bimodally-distributed. Kolmogorov-Smirnov normality tests using N_p as the theoretical distribution confirm that, overall, the data are not normally-distributed (for proportion infected $D = -9.1205$, $p = 0.000$; for infection duration $D = -9.1981$, $p = 0.000$). This finding would appear to mirror the literature on epidemic size

distributions, suggesting that infection events in the model generally remain within a local cluster, but sometimes explode in scale due to bridging links [2-4, 43-45].
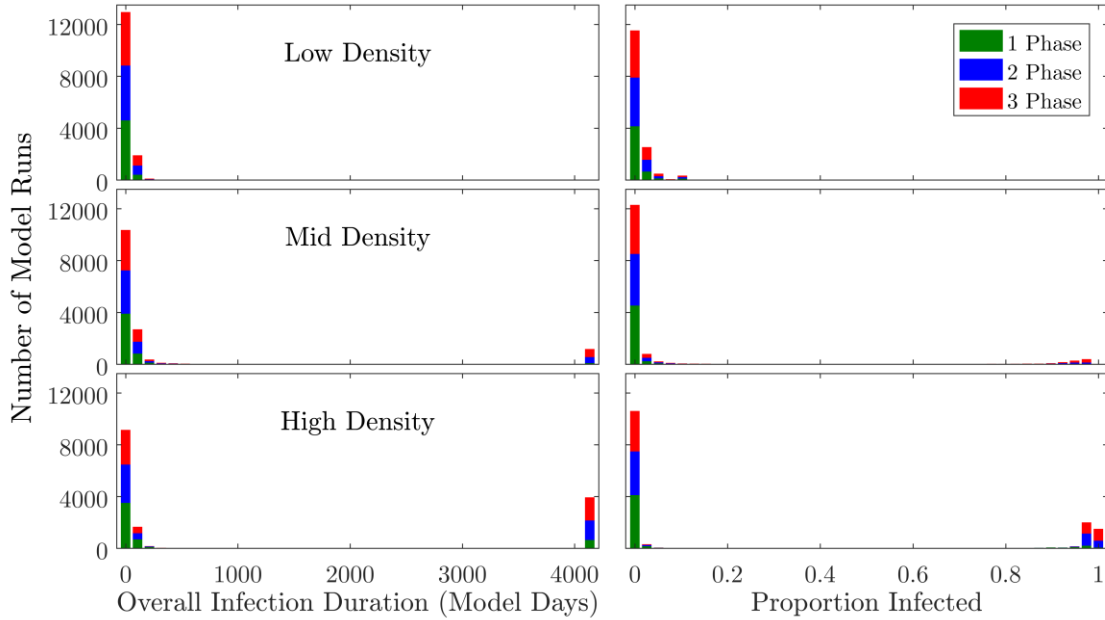
Figure 5: Histograms showing distribution of dependent variables. Infection duration data appear in the left column and proportion of infected agents in the right column, with color indicating producer specialization level. Low density runs were those with $0<N_p\leq500$, mid-density $500<N_p\leq1000$, and high-density $1000<N_p\leq1500$. Data were split into 40 bins.

Digging deeper into the behavior of the system within the subset of model runs that resulted in a long-duration "systemic" infection, we plot histograms including only runs in which the overall infection duration was ≥3000 model days (Figure 6). This analysis indicates that, at sufficiently-high N_p values, all three treatments sometimes

83

result in full-duration (4135 model day) epidemics. Once an infection reaches the systemic level, it is very unlikely to die off naturally prior to the end of the model run. However, even among the systemic outbreaks, the scale of spreading exhibits wider variability, with the high-specialization runs more likely to result in larger epidemics.
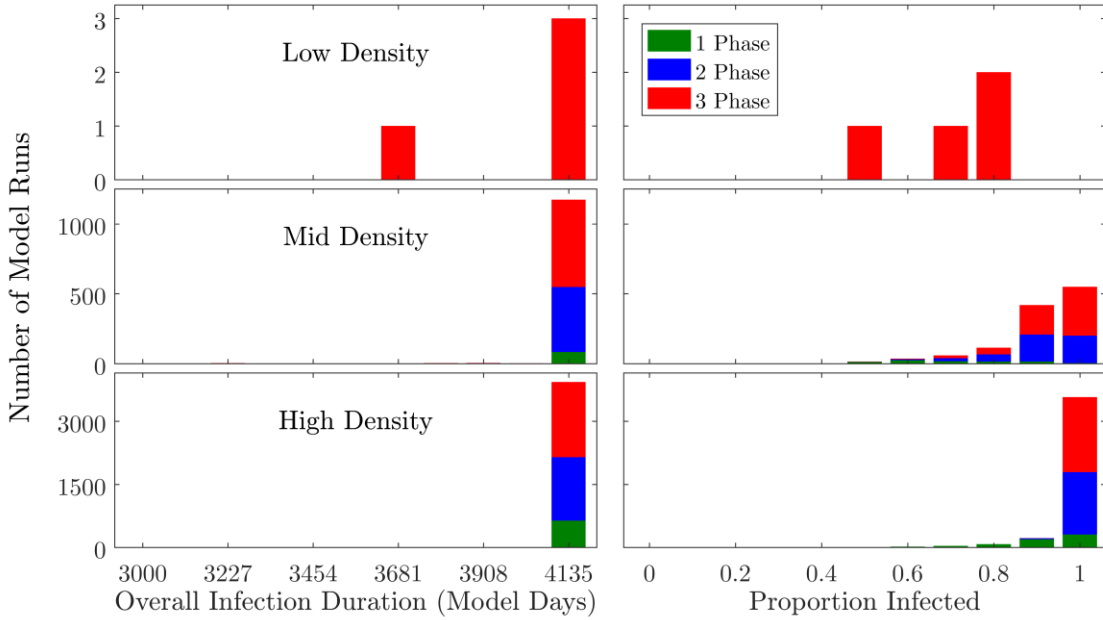
Figure 6: Right-censored histograms showing distribution of dependent variables. These plots are parallel to those in Figure 5, yet include only datapoints in which the infection duration was ≥3000 model days. Data were split into 11 bins.

Scatter plots of the raw data (Figure 7) reveal a nonlinearity, with N_p values below some value never igniting a globalized epidemic. This critical value appears to vary by specialization level. Kruskal-Wallis equality-of-populations rank tests (used

due to the non-normal data distributions) indicate that the data associated with each treatment differ significantly in terms of both infection duration and proportion infected (Table 3). Based on a cursory visual analysis, for the multi-phase systems, the critical region occurs at approximately $500 \leq N_p \leq 1000$, separating the unimodal phase (in which all infections are small and short) from the bimodal phase (in which infections are either small and short or large and long, but never in between). For the single-phase systems, the critical region would appear to occur around $600 \leq N_p \leq 1400$. Only within the critical regions do we ever observe mid-length or mid-scale outbreaks.
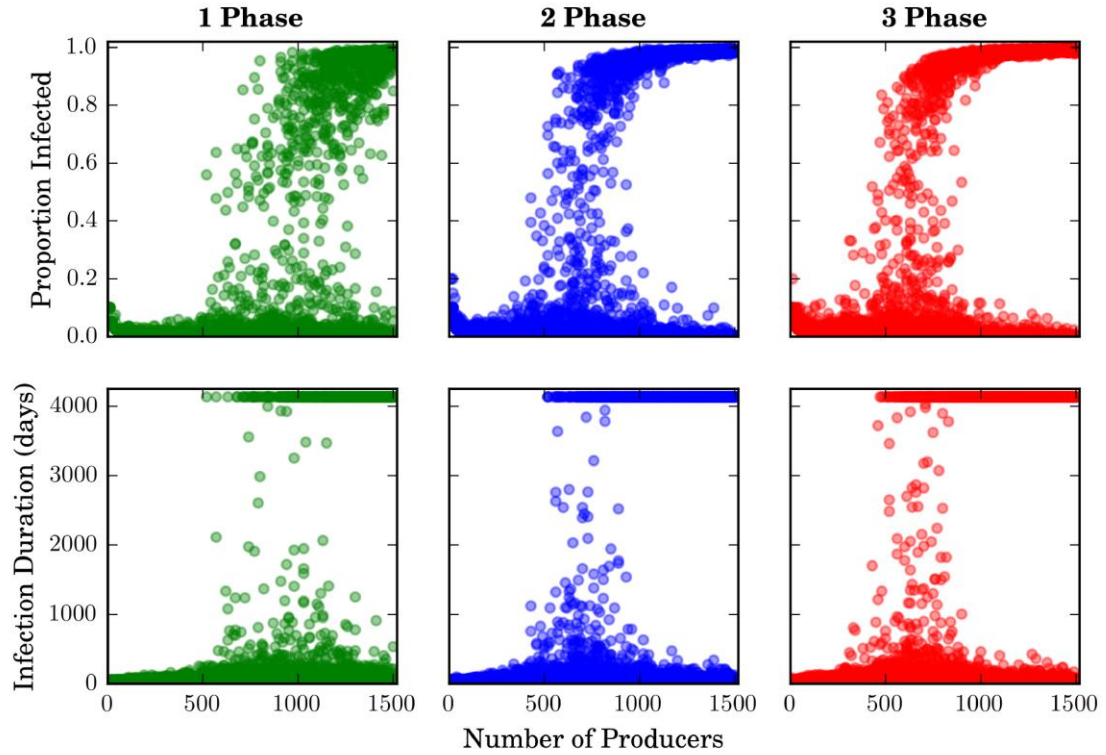
Figure 7: Scatter plots showing full model-output dataset for both dependent variables. Proportion of agents infected (cumulative) appears in the top row, and network-level infection duration in the bottom. Each point represents one of the 45,000 model runs (15,000 for each level of specialization).

Table 3: Kruskal-Wallis equality-of-populations rank test statistics.

Dependent Variable	Rank Sums			χ^2	d.f.	p
	1 Phase	2 Phase	3 Phase			
Proportion infected	2.99e+08	3.47e+08	3.67e+08	962.912	2	0.0001
Infection duration	3.09e+08	3.44e+08	3.59e+08	517.021	2	0.0001

To investigate the apparent percolation dynamics in the raw data, we plot the mean and 95% confidence interval for both dependent variables—as well as two alternative indicators of infection severity—over the full N_p range (Figure 8). These data, especially the metrics that focus on large-scale and long-term infection events (bottom row), provide further evidence of a percolation threshold. We also note that variability increases dramatically as N_p surpasses the critical region, with higher variability in the high-specialization data.
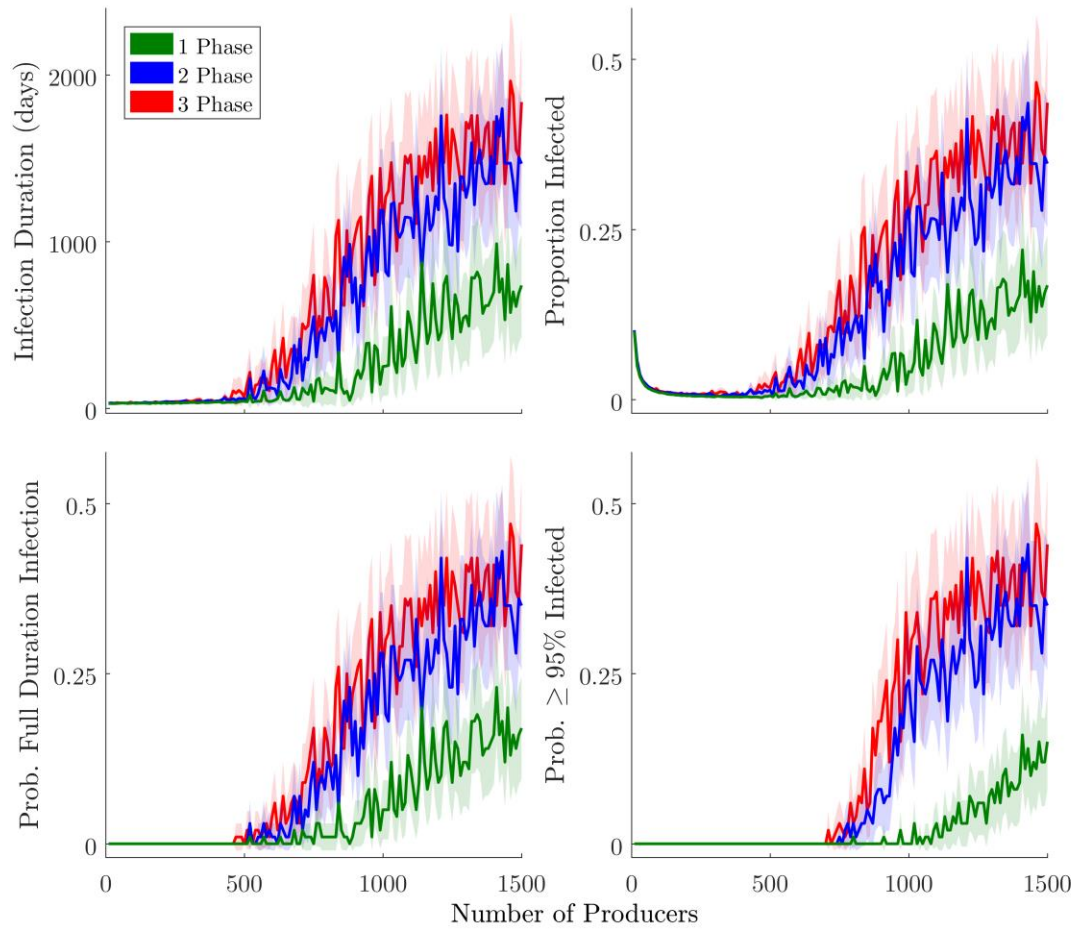
Figure 8: Percolation threshold visualizations. Lines plot average values for the 100 runs at each of 150 N_p levels, with corresponding color fields indicating 95% CI. Top left plot shows infection duration. Top right shows mean proportion infected (cumulative). Bottom left shows the fraction of runs resulting in a systemic network-level infection lasting the full duration of the model run (4135 model days). Bottom right shows the fraction of runs in which 95% or more of the agents became infected.

To further analyze the scaling behavior of the dependent variables, we apply

LOESS smoothing to the raw model output data. The N_p value at which the LOESS-

smoothed curve has the highest slope indicates the point at which outbreak severity

scales most abruptly with N_p, or the approximate percolation threshold. Figure 9

displays the results of this procedure, with the lower plots showing the slope of the

LOESS curves as a function of N_p. The N_p values corresponding to the maximum

slopes (indicated in the figures) would appear to correspond to the critical N_p ranges

observed visually in Figure 7 and Figure 8 for all treatments. On both epidemic severity

metrics, the three-phase treatment exhibits the lowest percolation threshold, as well as

the highest slope at this point, indicating that the high-specialization networks are

vulnerable to epidemic percolation at lower densities, and also exhibit a more-abrupt

phase-change. Although the difference between the two- and three-phase systems is

marginal, we can conclude that there is a marked differentiation between the behavior

of single- versus multi-phase systems at criticality.
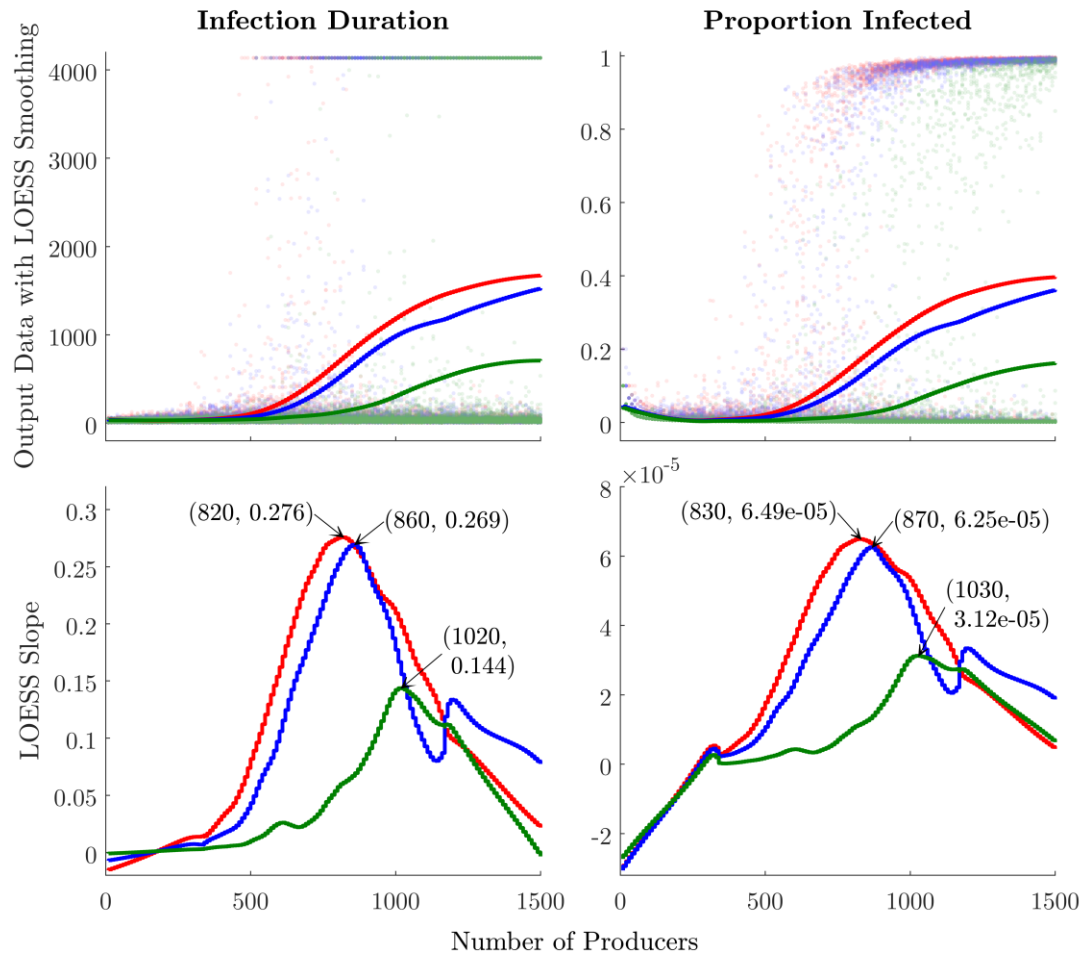
Figure 9: Finding percolation points numerically. Upper plots show raw model output data with LOESS-smoothed curves (span length = 0.45×*N*). Lower plots show the slope of each LOESS curve, with maximum-slope points annotated. Green represents 1-phase, blue 2-phase, and red 3-phase treatments.

3.4.2 Experiment 2

3.4.2.1 Contact network metrics

 Experiment one found that the greatest differences in percolation risk occur when stepping from single-phase to multi-phase systems. Here we investigate whether key network metrics may provide clues that explain this result from a network-theoretic perspective (Table 4 and Figure 10). The most striking feature in these data is that, in the single-phase networks, several of the network metrics simply do not scale with N_p as they do in the multi-phase scenarios. Network maps plotted from sample model runs (Figure 11) illustrate the fundamental structural difference underlying this result: in the single-phase systems, each producer is connected only to a single feed mill and a single slaughter plant. In light of this, it is clear why the average clustering coefficient $\langle C \rangle$—defined as the ratio of "closed triangles" to "total triangles"—will by definition be equal to zero for all single-phase runs. For the same reason, average degree $\langle k \rangle = 3$, k-core size $S_{kc} \approx N_p$, and k-core order $O_{kc} = 2$ also hold universally.

Table 4: Key contact network metrics for each producer specialization level, stratified across three producer density categories.

Network Metric		$0<N_p\leq500$			$500<N_p\leq1000$			$1000<N_p\leq1500$		
		Mean	95% CI		Mean	95% CI		Mean	95% CI	
Average Degree $\langle k \rangle$	1 Phase	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
	2 Phase	7.14	6.42	7.86	15.58	14.52	16.64	24.84	22.90	26.78
	3 Phase	6.70	5.90	7.50	16.62	15.67	17.57	26.88	24.98	28.78
Average Shortest Path Length $\langle l \rangle$	1 Phase	3.70	3.60	3.82	3.66	3.56	3.77	3.54	3.44	3.64
	2 Phase	3.82	3.72	3.92	3.58	3.52	3.64	3.57	3.51	3.62
	3 Phase	4.16	4.17	4.45	3.85	3.81	3.89	3.80	3.75	3.85
Clustering Coefficient C	1 Phase	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2 Phase	0.20	0.19	0.21	0.13	0.12	0.14	0.09	0.08	0.10
	3 Phase	0.19	0.18	0.20	0.13	0.12	0.13	0.09	0.08	0.09
k-core Order O_{kc}	1 Phase	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
	2 Phase	8.34	7.39	9.29	17.42	15.80	19.04	26.88	23.90	29.86
	3 Phase	8.58	7.47	9.69	19.88	18.09	21.67	29.78	26.71	32.85
k-core Size S_{kc}	1 Phase	308.40	266.63	350.17	812.76	772.16	853.36	1312.8	1272.2	1353.5
	2 Phase	66.12	52.13	80.11	242.42	200.33	284.51	547.80	481.56	614.04
	3 Phase	67.04	50.74	83.34	209.88	175.72	244.04	443.50	375.73	511.27
Number of k-cores N_{kc}	1 Phase	1.02	0.98	1.06	1.00	1.00	1.00	1.00	1.00	1.00
	2 Phase	1.12	1.03	1.21	1.04	0.98	1.10	1.00	1.00	1.00
	3 Phase	1.10	1.00	1.20	1.00	1.00	1.00	1.00	1.00	1.00
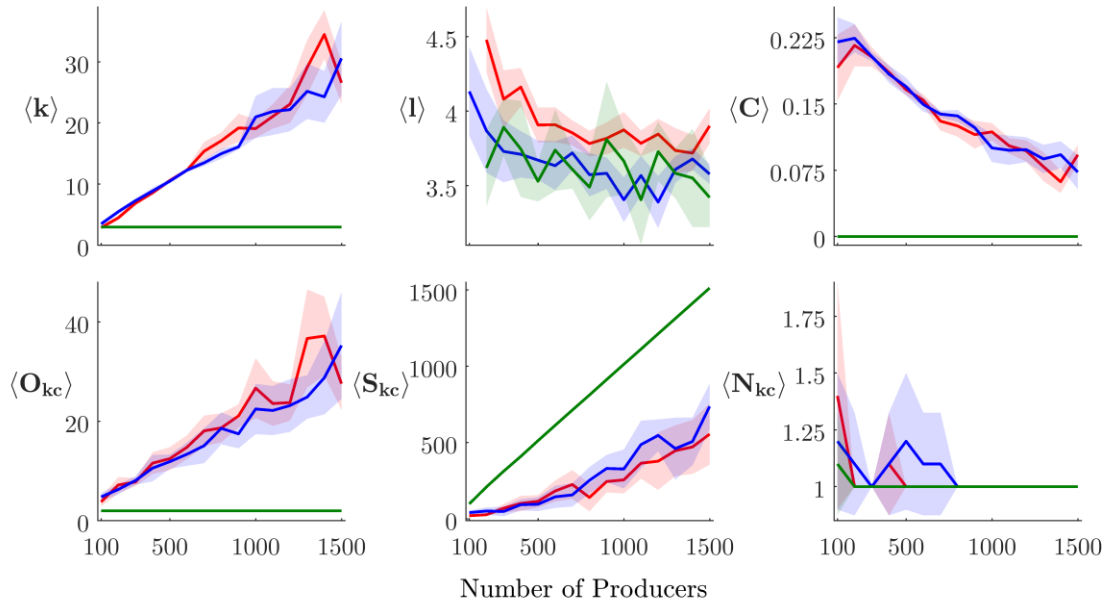
Figure 10: Correlating N$_p$ with contact network metrics. Key contact network metrics, calculated for each treatment. Lines plot averages for each N_p value; color fields show 95% CI. Green represents 1-phase, blue 2-phase, and red 3-phase treatments.

93
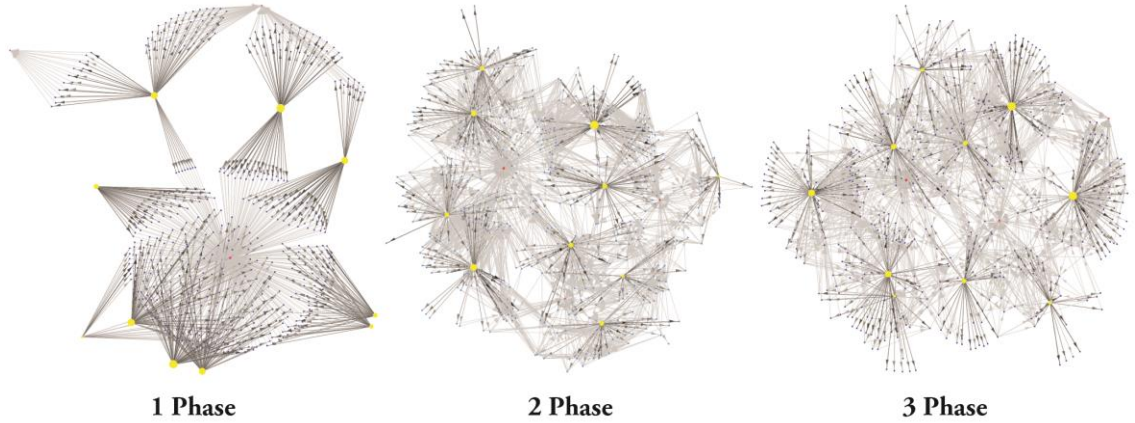
Figure 11: Visualizations of sample networks generated by the model under each level of producer specialization. N_p**=500 for each network. Nodes were positioned using a spring layout, and sized according to total number of contact events. Blue nodes are producers; yellow are feed mills, and red are slaughter plants.**

For the multi-phase networks, the explanation for the higher $\langle k \rangle$ is trivial: an entire interaction context is added, so there must be more edges. The more important realization is that the addition of the producer–producer interaction context can add bridging edges, resulting in elevated $\langle C \rangle$ values. Places where these bridges connect portions of the network that would otherwise have remained isolated represent clear risk points for disease outbreaks to become systemic.

But, can any of these metrics reliably predict epidemiological risk? For the multi-phase networks, many of the metrics seem to scale roughly linearly with N_p, with $\langle k \rangle$, $\langle O_{kc} \rangle$, and $\langle S_{kc} \rangle$ being positively correlated; and $\langle l \rangle$ and $\langle C \rangle$ being negatively correlated. Unfortunately, the lack of any significant nonlinearity suggests that the

94

percolation point cannot be reliably predicted *a priori* by tracking these network metrics as a network grows. Developing metrics that are effective predictors of disease spread risk in complex networks remains an area for future study.

3.4.2.2 Infected component network metrics

Examining metrics calculated on infected-component subgraphs (Table 5) provides insights into the network structures that underlie percolation, and how these structures differ between the single- and multi-phase networks. Overall, we note that only 226 of the 450 model runs conducted in experiment two resulted in an infection network, with the infection in the remaining 224 runs failing to spread beyond the initially-infected node. This mirrors the bimodal epidemic size distribution discussed above and shown in Figure 5 and Figure 6. As a result of this lower N, N_p values were divided into five bins for visualization (Figure 12).

Table 5: Key infected component network metrics for each producer specialization level, stratified across three producer density categories.

Network Metric		$0<N_p\leq500$			$500<N_p\leq1000$			$1000<N_p\leq1500$		
		Mean	95% CI		Mean	95% CI		Mean	95% CI	
Average Degree $\langle k^i \rangle$	1 Phase	1.05	0.95	1.14	1.26	0.99	1.53	1.76	1.33	2.19
	2 Phase	1.25	1.08	1.42	3.95	1.70	6.21	10.64	7.10	14.18
	3 Phase	1.25	0.96	1.54	5.33	3.09	7.58	13.41	9.10	17.72
Average Shortest Path Length $\langle l^i \rangle$	1 Phase	1.20	1.06	1.35	1.72	1.35	2.09	2.20	1.66	2.73
	2 Phase	1.63	1.34	1.91	2.35	1.81	2.88	2.70	2.24	3.15
	3 Phase	1.69	1.34	2.04	2.40	1.86	2.95	2.68	2.22	3.13
Clustering Coefficient C^i	1 Phase	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2 Phase	0.06	-0.02	0.13	0.09	0.03	0.15	0.07	0.04	0.09
	3 Phase	0.02	-0.01	0.05	0.07	0.04	0.11	0.05	0.03	0.08
k-core Order O_{kc}^i	1 Phase	1.18	0.96	1.40	1.48	1.09	1.87	1.67	1.25	2.08
	2 Phase	1.54	1.31	1.76	5.76	2.19	9.33	14.21	9.06	19.37
	3 Phase	1.54	1.11	1.97	7.37	3.81	10.93	14.44	9.07	19.80
k-core Size S_{kc}^i	1 Phase	2.64	2.21	3.06	39.61	-11.78	91.00	342.62	112.69	572.55
	2 Phase	4.64	3.03	6.26	47.67	11.24	84.09	105.96	44.94	166.99
	3 Phase	4.54	2.65	6.43	50.48	21.21	79.76	158.72	90.45	226.99
Number of k-cores N_{kc}^i	1 Phase	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	2 Phase	1.00	1.00	1.00	1.00	1.00	1.00	1.04	0.96	1.11
	3 Phase	1.00	1.00	1.00	1.00	1.00	1.00	1.03	0.97	1.09

Note: Reported values are averages for runs falling into each producer density category.
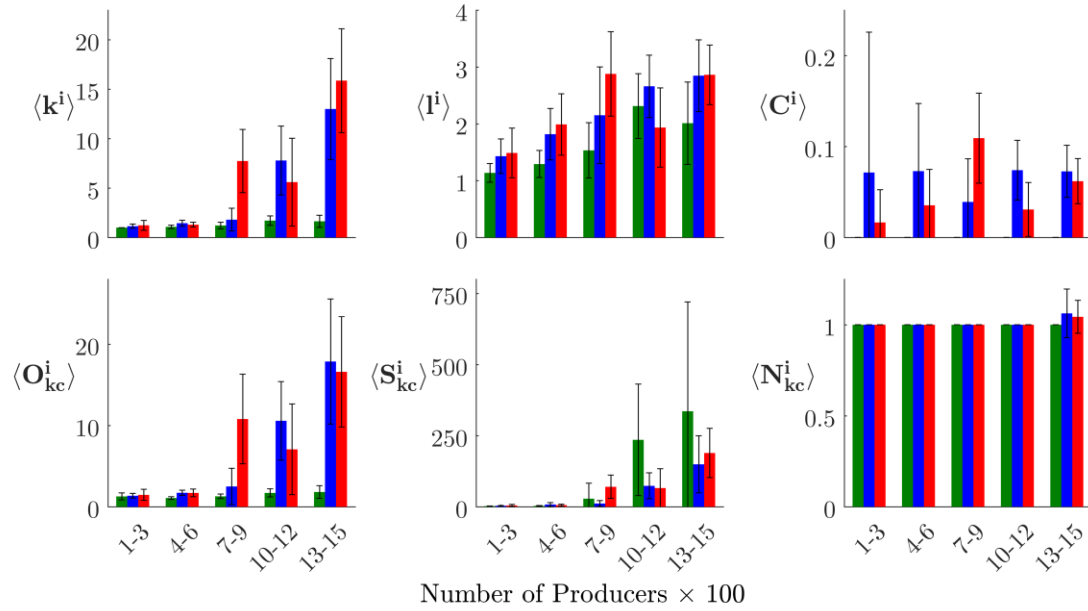
Figure 12: Correlating N$_{\mathbf{p}}$ with infected component network metrics. Key infected component network metrics, calculated for each treatment. Bars plot averages for five N_p ranges; whiskers show 95% CI. Green represents 1-phase, blue 2-phase, and red 3-phase treatments.

Data analyses reveal several differences between the overall contact networks and the infected component networks. Whereas $\langle C \rangle$ in the contact network drops as N_p rises, $\langle C^i \rangle$ remains relatively flat, suggesting that subgraphs with relatively-higher clustering than the rest of a network may be especially vulnerable to disease. Secondly, in the multi-phase systems, a nonlinearity would appear to exist in $\langle k^i \rangle$ with respect to N_p around the critical percolation regions. As a network grows larger, it becomes less likely that any two randomly-selected nodes will be linked, since only so many contacts

can occur in a given timeframe. Therefore, heavily-connected nodes will tend to be the ones whose edges happen to impinge upon an infected trading partner. $\langle O_{kc}^i \rangle$ and $\langle S_{kc}^i \rangle$ also appear to exhibit a similar non-linearity. Interestingly, for the multi-phase networks, $\langle O_{kc}^i \rangle$ would seem to more heavily reflect the percolation threshold, whereas for the single-phase networks, in which it was always the case that $1 < \langle O_{kc}^i \rangle < 2$, it is $\langle S_{kc}^i \rangle$ that balloons upon reaching the threshold.

3.5 Discussion

Our experimental results strongly suggest that, at least in the context of the model presented here, the risk of catastrophic infectious disease outbreaks may be inhibited by (a) sparser networks, and, perhaps more critically, (b) networks in which fewer contexts for interaction facilitate greater compartmentalization of inter-agent contact patterns, leading to both shorter and smaller outbreaks, as well as less uncertainty about whether a given outbreak will become systemic. These findings corroborate previous theoretical research into the network features that can promote large-scale epidemics [43, 44], as well as empirical studies that point to similar infection spread patterns having occurred in real-world outbreaks [2-4].

Despite the phase transitions in our data not being particularly "sharp," there is clear evidence of a nonlinearity in the scaling of epidemic severity with producer density in the hog production systems generated by our model. Quantifying the

98

producer densities at which adding additional producers to the system is most-strongly

correlated with an increased risk of catastrophic disease spread reveals a clear

differentiation between the epidemiological resilience of low- versus high-

specialization treatments.

As in many dynamical systems, we find that the critical region around the

percolation threshold acts as a border between a unimodal system state in which disease

outbreaks virtually always die out quickly, and a bimodal state in which large-scale,

systemic outbreaks are possible. It is only within the critical region that medium-

severity outbreaks are observed. This finding is important because it entails that, if the

size and/or duration of disease events in a growing livestock production network has

begun to show wider variability, this could be an indication that further increasing the

regional production density may not simply increase the risk linearly, but instead

accelerate the system toward a regime characterized by the possibility for catastrophic

epidemic events.

A limitation of the model concerns human behavioral adaptation in the face of

epidemiological threats. Whereas the model used here assumes that agents' behavioral

heuristics remain static, previous research has pointed to the potential for behavioral

adaptation—for example limiting contact as a disease becomes more prevalent—to

significantly affect the course of an outbreak [73]. The extent to which such adaptive

behavior may differentially-impact the disease resilience of livestock production

networks with varying levels of producer specialization or spatial density remains an area for future research.

A limitation of experiment two rests in our selection of network metrics, and our choice to binarize and symmetrize the networks for analysis. While binarization and symmetrization have been employed historically in network analysis—and while there is some evidence to suggest that this approach is valuable for the evaluation of spreading dynamics [65]—future studies will compare the efficacy these metrics to their weighted and/or directed counterparts as indicators of epidemiological vulnerability. It would also be useful to analyze vulnerability not only from the whole-graph perspective, but from the level of individual nodes.

Another area for future study is to investigate percolation dynamics within mixed systems of single- and multi-phase producers. This would lend further insight to optimize risk mitigation strategies in real-world networks, which generally contain multiple overlapping production systems. For example, it would be valuable to understand the extent to which the introduction of just a few multi-phase producers into a region dominated by farrow to finish farms may impact percolation risk.

3.6 Conclusion

Those concerned with preventing the spread of catastrophic diseases in the U.S. hog industry most-commonly promote the adoption of discrete biosecurity and biocontainment interventions at the premises level; strategies which may well be efficacious in many situations. However, epidemics are ultimately spread through

complex networks of interacting actors, and—as we have shown—the structure of a given network can have a dramatic impact on the epidemiological resilience of the system. As hog production grows denser and more spatially-consolidated, it will become increasingly vital to consider how operational decisions made at the farm level impinge upon the patterns of trade and contact that may become transmission vectors in the next outbreak.

While single-phase systems may be falling out of favor for reasons of production efficiency, our results suggest that industry practitioners, managers, and regulators would be wise to consider the biosecurity advantages associated with farrow to finish farms when developing best management practices to mitigate epidemiological risk. All else being equal, systems dominated by single-phase producers should theoretically be able to withstand significantly higher farm densities without a corresponding increase in the risk of large-scale disease percolation. This is because adding a producer–producer interaction context can form bridges between otherwise-isolated parts of a network, turning what could have been a short-term, localized outbreak into an ongoing, systemic one. In hog-dense regions such as Iowa, Illinois, and North Carolina—where disease is a constant threat—a turn back toward single-phase production may offer a means to increase system-wide disease resilience, even while maintaining high regional hog production capacity.

3.7 Supporting information

The RUSHPNBM v.0.8 ODD+D Protocol appears at the end of this dissertation as Appendix 1. The datasets are available upon request from the author.

3.7.1 S1 Protocol

RUSHPNBM v.0.8 ODD+D Protocol. Describes in detail the technical specifications of the model developed for this experiment, including parameter values, calibration details, and pseudocode representations of all functions.

3.7.2 S1 Dataset

Infection Data. Infection duration and size data for 45,000 model runs. Used in experiment one.

3.7.3 S2 Dataset

Network Data. Network edge list data for all agents, across 450 model runs. Used in experiment two.

3.8 References

[1]   AMASS SF, CLARK LK. BIOSECURITY CONSIDERATIONS FOR PORK PRODUCTION UNITS. JOURNAL OF SWINE HEALTH AND PRODUCTION. 1999;7(5):217–228.

[2]   FIRESTONE SM, WARD MP, CHRISTLEY RM, DHAND NK. THE IMPORTANCE OF LOCATION IN CONTACT NETWORKS: DESCRIBING EARLY EPIDEMIC SPREAD USING SPATIAL SOCIAL NETWORK ANALYSIS. PREVENTIVE VETERINARY MEDICINE. 2011;102(3):185–195.

[3]   FOURNIÉ G, GUITIAN J, DESVAUX S, CUONG VC, PFEIFFER DU, MANGTANI P, ET AL. INTERVENTIONS FOR AVIAN INFLUENZA A (H5N1) RISK MANAGEMENT IN LIVE BIRD MARKET NETWORKS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES. 2013;110(22):9177–9182.

[4]   O'DEA EB, SNELSON H, BANSAL S. STATE-LEVEL TRANSPORT FLOWS ARE PREDICTIVE OF THE DYNAMICS OF PORCINE EPIDEMIC DIARRHEA VIRUS. BIORXIV. 2015; P. 017178.

[5]   USDA APHIS. REFERENCE OF SWINE HEALTH AND HEALTH MANAGEMENT PRACTICES IN THE UNITED STATES. U.S. DEPARTMENT OF AGRICULTURE, ANIMAL AND PLANT HEALTH INSPECTION SERVICE; 2006.

[6]   GRANOVETTER M. THRESHOLD MODELS OF COLLECTIVE BEHAVIOR. AMERICAN JOURNAL OF SOCIOLOGY. 1978; P. 1420–1443.

[7]   AXELROD RM, AXELROD R, COHEN MD. HARNESSING COMPLEXITY: ORGANIZATIONAL IMPLICATIONS OF A SCIENTIFIC FRONTIER. BASIC BOOKS; 2000.

[8]   JANSSEN MA, OSTROM E. EMPIRICALLY BASED, AGENT-BASED MODELS. ECOLOGY AND SOCIETY. 2006;11(2):37.

[9]   LOWE J, GAUGER P, HARMON K, ZHANG J, CONNOR J, YESKE P, ET AL. ROLE OF TRANSPORTATION IN SPREAD OF PORCINE EPIDEMIC DIARRHEA VIRUS INFECTION, UNITED STATES. EMERGING INFECTIOUS DISEASES. 2014;20(5):872.

[10]   BOWMAN AS, KROGWOLD RA, PRICE T, DAVIS M, MOELLER SJ. INVESTIGATING THE INTRODUCTION OF PORCINE EPIDEMIC DIARRHEA VIRUS INTO AN OHIO SWINE OPERATION. BMC VETERINARY RESEARCH. 2015;11(1):38.

[11]   GRASSBERGER P. ON THE CRITICAL BEHAVIOR OF THE GENERAL EPIDEMIC PROCESS AND DYNAMICAL PERCOLATION. MATHEMATICAL BIOSCIENCES. 1983;63(2):157–172.

[12]   OHTSUKI T, KEYES T. KINETIC GROWTH PERCOLATION: EPIDEMIC PROCESSES WITH IMMUNIZATION. PHYSICAL REVIEW A. 1986;33(2):1223.

[13]   MOORE C, NEWMAN MEJ. EPIDEMICS AND PERCOLATION IN SMALL-WORLD NETWORKS. PHYSICAL REVIEW E. 2000;61(5):5678.

[14]   SANDER LM, WARREN CP, SOKOLOV IM, SIMON C, KOOPMAN J. PERCOLATION ON HETEROGENEOUS NETWORKS AS A MODEL FOR EPIDEMICS. MATHEMATICAL BIOSCIENCES. 2002;180(1):293–305.

[15]   SERRANO MN, BOGUNÁ M. PERCOLATION AND EPIDEMIC THRESHOLDS IN CLUSTERED NETWORKS. PHYSICAL REVIEW LETTERS. 2006;97(8):088701.

[16]   SORNETTE D. CRITICAL PHENOMENA IN NATURAL SCIENCES: CHAOS, FRACTALS, SELFORGANIZATION AND DISORDER: CONCEPTS AND TOOLS. SPRINGER; 2006. AVAILABLE FROM: HTTPS://BOOKS.GOOGLE.COM/BOOKS?ID=UYHQHV2HCPWC.

[17]   STAUFFER D, AHARONY A. INTRODUCTION TO PERCOLATION THEORY. CRC PRESS; 1994.

[18]   GRIMMETT G. WHAT IS PERCOLATION? SPRINGER; 1999.

[19]   ALBERT R, JEONG H, BARABÁSI AL. ERROR AND ATTACK TOLERANCE OF COMPLEX NETWORKS. NATURE. 2000;406(6794):378–382.

[20]   CARLSON JM, DOYLE J. HIGHLY OPTIMIZED TOLERANCE: ROBUSTNESS AND DESIGN IN COMPLEX SYSTEMS. PHYSICAL REVIEW LETTERS. 2000;84(11):2529.

[21]   CARLSON JM, DOYLE J. COMPLEXITY AND ROBUSTNESS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES. 2002;99(SUPPL 1):2538–2545.

[22]   NEWMAN MEJ, GIRVAN M, DOYNE FARMER J. OPTIMAL DESIGN, ROBUSTNESS, AND RISK AVERSION. PHYSICAL REVIEW LETTERS. 2002;89(2):028301.

[23]   KERMACK WO, MCKENDRICK AG. A CONTRIBUTION TO THE MATHEMATICAL THEORY OF EPIDEMICS. IN: PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON A: MATHEMATICAL, PHYSICAL AND ENGINEERING SCIENCES. VOL. 115. THE ROYAL SOCIETY; 1927. P. 700–721.

[24]   HEESTERBEEK JAP. MATHEMATICAL EPIDEMIOLOGY OF INFECTIOUS DISEASES: MODEL BUILDING, ANALYSIS AND INTERPRETATION. VOL. 5. JOHN WILEY & SONS; 2000.

[25]   HETHCOTE HW. THE MATHEMATICS OF INFECTIOUS DISEASES. SIAM REVIEW. 2000;42(4):599–653.

[26]   HARKO T, LOBO FSN, MAK MK. EXACT ANALYTICAL SOLUTIONS OF THE SUSCEPTIBLE-INFECTED-RECOVERED (SIR) EPIDEMIC MODEL AND OF THE SIR MODEL WITH EQUAL DEATH AND BIRTH RATES. APPLIED MATHEMATICS AND COMPUTATION. 2014;236:184–194.

[27]   ANDERSON RM, MAY RM, ANDERSON B. INFECTIOUS DISEASES OF HUMANS: DYNAMICS AND CONTROL. VOL. 28. WILEY ONLINE LIBRARY; 1992.

[28]   NEWMAN MEJ. SPREAD OF EPIDEMIC DISEASE ON NETWORKS. PHYSICAL REVIEW E. 2002;66(1).

[29]   AUGER P, MAGAL P, RUAN S. STRUCTURED POPULATION MODELS IN BIOLOGY AND EPIDEMIOLOGY. VOL. 1936. SPRINGER; 2008.

[30]   CLIFF AD. SPATIAL DIFFUSION: AN HISTORICAL GEOGRAPHY OF EPIDEMICS IN AN ISLAND COMMUNITY. 14. CUP ARCHIVE; 1981.

[31]   FUENTES MA, KUPERMAN MN. CELLULAR AUTOMATA AND EPIDEMIOLOGICAL MODELS WITH SPATIAL DEPENDENCE. PHYSICA A: STATISTICAL MECHANICS AND ITS APPLICATIONS. 1999;267(3):471–486.

[32]   KINZEL W. PHASE TRANSITIONS OF CELLULAR AUTOMATA. ZEITSCHRIFT FÜR PHYSIK B CONDENSED MATTER. 1985;58(3):229–244.

[33]   KEELING MJ. THE EFFECTS OF LOCAL SPATIAL STRUCTURE ON EPIDEMIOLOGICAL INVASIONS. PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON B: BIOLOGICAL SCIENCES. 1999;266(1421):859–867.

[34]   PFEIFER B, KUGLER K, TEJADA MM, BAUMGARTNER C, SEGER M, OSL M, ET AL. A CELLULAR AUTOMATON FRAMEWORK FOR INFECTIOUS DISEASE SPREAD SIMULATION. THE OPEN MEDICAL INFORMATICS JOURNAL. 2008;2:70.

[35]   KEELING MJ, WOOLHOUSE MEJ, MAY RM, DAVIES G, GRENFELL BT. MODELLING VACCINATION STRATEGIES AGAINST FOOT-AND-MOUTH DISEASE. NATURE. 2003;421(6919):136–142.

[36]   MIKLER AR, VENKATACHALAM S, ABBAS K. MODELING INFECTIOUS DISEASES USING GLOBAL STOCHASTIC CELLULAR AUTOMATA. JOURNAL OF BIOLOGICAL SYSTEMS. 2005;13(04):421–439.

[37]   WHITE SH, DEL REY AM, SÁNCHEZ GR. MODELING EPIDEMICS USING CELLULAR AUTOMATA. APPLIED MATHEMATICS AND COMPUTATION. 2007;186(1):193–202.

[38]   KEELING MJ, EAMES KTD. NETWORKS AND EPIDEMIC MODELS. JOURNAL OF THE ROYAL SOCIETY INTERFACE. 2005;2(4):295–307.

[39]   MORENO Y, PASTOR-SATORRAS R, VESPIGNANI A. EPIDEMIC OUTBREAKS IN COMPLEX HETEROGENEOUS NETWORKS. THE EUROPEAN PHYSICAL JOURNAL B - CONDENSED MATTER AND COMPLEX SYSTEMS. 2002;26(4):521–529. DOI:10.1140/EPJB/E20020122.

[40]   BARTHÉLEMY M, BARRAT A, PASTOR-SATORRAS R, VESPIGNANI A. VELOCITY AND HIERARCHICAL SPREAD OF EPIDEMIC OUTBREAKS IN SCALE-FREE NETWORKS. PHYSICAL REVIEW LETTERS. 2004;92(17):178701.

[41]   PASTOR-SATORRAS R, VESPIGNANI A. EPIDEMIC SPREADING IN SCALE-FREE NETWORKS. PHYSICAL REVIEW LETTERS. 2001;86(14):3200.

[42]   VESPIGNANI A. MODELLING DYNAMICAL PROCESSES IN COMPLEX SOCIO-TECHNICAL SYSTEMS. NATURE PHYSICS. 2012;8(1):32.

[43]   BAILEY NTJ. THE MATHEMATICAL THEORY OF INFECTIOUS DISEASES AND ITS APPLICATIONS. CHARLES GRIFFIN & COMPANY LTD, 5A CRENDON STREET, HIGH WYCOMBE, BUCKS HP13 6LE.; 1975.

[44]   WATTS DJ, MUHAMAD R, MEDINA DC, DODDS PS. MULTISCALE, RESURGENT EPIDEMICS IN A HIERARCHICAL METAPOPULATION MODEL. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. 2005;102(32):11157–11162.

[45]   BAJARDI P, POLETTO C, RAMASCO JJ, TIZZONI M, COLIZZA V, VESPIGNANI A. HUMAN MOBILITY NETWORKS, TRAVEL RESTRICTIONS, AND THE GLOBAL SPREAD OF 2009 H1N1 PANDEMIC. PLOS ONE. 2011;6(1):E16591.

[46]   COELHO FC, CRUZ OG, CODEÇO CT. EPIGRASS: A TOOL TO STUDY DISEASE SPREAD IN COMPLEX NETWORKS. SOURCE CODE FOR BIOLOGY AND MEDICINE. 2008;3(1):3. DOI:10.1186/1751-0473-3-3.

[47]   RAHMANDAD H, STERMAN J. HETEROGENEITY AND NETWORK STRUCTURE IN THE DYNAMICS OF DIFFUSION: COMPARING AGENT-BASED

AND DIFFERENTIAL EQUATION MODELS. MANAGEMENT SCIENCE. 2008;54(5):998–1014.

[48]   PEREZ L, DRAGICEVIC S. AN AGENT-BASED APPROACH FOR MODELING DYNAMICS OF CONTAGIOUS DISEASE SPREAD. INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS. 2009;8(1):50.

[49]   COLIZZA V, BARRAT A, BARTHÉLEMY M, VESPIGNANI A. THE ROLE OF THE AIRLINE TRANSPORTATION NETWORK IN THE PREDICTION AND PREDICTABILITY OF GLOBAL EPIDEMICS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. 2006;103(7):2015–2020.

[50]   FRIAS-MARTINEZ E, WILLIAMSON G, FRIAS-MARTINEZ V. AN AGENT-BASED MODEL OF EPIDEMIC SPREAD USING HUMAN MOBILITY AND SOCIAL NETWORK INFORMATION. IN: PRIVACY, SECURITY, RISK AND TRUST (PASSAT) AND 2011 IEEE THIRD INERNATIONAL CONFERENCE ON SOCIAL COMPUTING (SOCIALCOM), 2011 IEEE THIRD INTERNATIONAL CONFERENCE ON. IEEE; 2011. P. 57–64.

[51]   BAGNI R, BERCHI R, CARIELLO P. A COMPARISON OF SIMULATION MODELS APPLIED TO EPIDEMICS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION. 2002;5(3).

[52]   BRADHURST RA, ROCHE SE, GARNER MG, SAJEEV ASM, KWAN P. MODELLING THE SPREAD OF LIVESTOCK DISEASE ON A NATIONAL SCALE: THE CASE FOR A HYBRID APPROACH. IN: PROCEEDINGS OF THE 20TH INTERNATIONAL CONGRESS ON MODELLING AND SIMULATION (MODSIM2013), MODELLING AND SIMULATION SOCIETY OF AUSTRALIA AND NEW ZEALAND; 2013. P. 345–351.

[53]   GHANI AC, GARNETT GP. RISKS OF ACQUIRING AND TRANSMITTING SEXUALLY TRANSMITTED DISEASES IN SEXUAL PARTNER NETWORKS. SEXUALLY TRANSMITTED DISEASES. 2000;27(10):579–587.

[54]   EUBANK S, GUCLU H, KUMAR VSA, MARATHE MV. MODELLING DISEASE OUTBREAKS IN REALISTIC URBAN SOCIAL NETWORKS. NATURE. 2004;429(6988):180.

[55]   GOJOVIC MZ, SANDER B, FISMAN D, KRAHN MD, BAUCH CT. MODELLING MITIGATION STRATEGIES FOR PANDEMIC (H1N1) 2009. CANADIAN MEDICAL ASSOCIATION JOURNAL. 2009;181(10):673–680.

[56]   KEELING MJ, WOOLHOUSE MEJ, SHAW DJ, MATTHEWS L, CHASE-TOPPING M, HAYDON DT, ET AL. DYNAMICS OF THE 2001 UK FOOT AND MOUTH EPIDEMIC: STOCHASTIC DISPERSAL IN A HETEROGENEOUS LANDSCAPE. SCIENCE. 2001;294(5543):813–817. DOI:10.1126/SCIENCE.1065973.

[57]   NATALE F, GIOVANNINI A, SAVINI L, PALMA D, POSSENTI L, FIORE G, ET AL. NETWORK ANALYSIS OF ITALIAN CATTLE TRADE PATTERNS AND EVALUATION OF RISKS FOR POTENTIAL DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE. 2009;92(4):341–350.

[58]   SALATHÉ M, JONES JH. DYNAMICS AND CONTROL OF DISEASES IN NETWORKS WITH COMMUNITY STRUCTURE. PLoS COMPUTATIONAL BIOLOGY. 2010;6(4):E1000736.

[59]   CHRISTLEY R, ROBINSON S, LYSONS R, FRENCH N. NETWORK ANALYSIS OF CATTLE MOVEMENT IN GREAT BRITAIN. PROC SOC VET EPIDEMIOL PREV MED. 2005; P. 234–243.

[60]   BIGRAS-POULIN M, THOMPSON R, CHRIÉL M, MORTENSEN S, GREINER M. NETWORK ANALYSIS OF DANISH CATTLE INDUSTRY TRADE PATTERNS AS AN EVALUATION OF RISK POTENTIAL FOR DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE. 2006;76(1):11–39.

[61]   BIGRAS-POULIN M, BARFOD K, MORTENSEN S, GREINER M. RELATIONSHIP OF TRADE PATTERNS OF THE DANISH SWINE INDUSTRY ANIMAL MOVEMENTS NETWORK TO POTENTIAL DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE. 2007;80(2):143–165.

[62]   WATTS DJ, STROGATZ SH. COLLECTIVE DYNAMICS OF 'SMALL-WORLD'NETWORKS. NATURE. 1998;393(6684):440–442.

[63]   CHRISTLEY RM, PINCHBECK G, BOWERS R, CLANCY D, FRENCH N, BENNETT R, ET AL. INFECTION IN SOCIAL NETWORKS: USING NETWORK ANALYSIS TO IDENTIFY HIGH-RISK INDIVIDUALS. AMERICAN JOURNAL OF EPIDEMIOLOGY. 2005;162(10):1024–1031.

[64]   KITSAK M, GALLOS LK, HAVLIN S, LILJEROS F, MUCHNIK L, STANLEY HE, ET AL. IDENTIFICATION OF INFLUENTIAL SPREADERS IN COMPLEX NETWORKS. NATURE PHYSICS. 2010;6(11):888–893.

[65]   BELL DC, ATKINSON JS, CARLSON JW. CENTRALITY MEASURES FOR DISEASE TRANSMISSION NETWORKS. SOCIAL NETWORKS. 1999;21(1):1–21.

[66]   AASV. PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV). HTTPS://WWW.AASV.ORG/AASV AMERICAN ASSOCIATION OF SWINE VETERINARIANS; 2015.

[67]   MAY RM. USES AND ABUSES OF MATHEMATICS IN BIOLOGY. SCIENCE. 2004;303(5659):790–793.

[68]   JOHNSON D. THE TRIANGULAR DISTRIBUTION AS A PROXY FOR THE BETA DISTRIBUTION IN RISK ANALYSIS. JOURNAL OF THE ROYAL STATISTICAL SOCIETY: SERIES D (THE STATISTICIAN). 1997;46(3):387–398.

[69]   USDA NASS. HOGS AND PIGS: INVENTORY BY TYPE OF PRODUCER. U.S. DEPARTMENT OF AGRICULTURE, NATIONAL AGRICULTURAL STATISTICS SERVICE; 2012. AVAILABLE FROM: HTTP://AGCENSUS.USDA.GOV/PUBLICATIONS/2012/FULL_REPORT/VOLUME_1,_CHAPTER_1_US/ST99_1_020_023.PDF.

[70]   USDA FSIS. MEAT, POULTRY AND EGG PRODUCT INSPECTION DIRECTORY. U.S. DEPARTMENT OF AGRICULTURE, FOOD SAFETY AND INSPECTION SERVICE; 2015. AVAILABLE FROM: HTTP://WWW.FSIS.USDA.GOV/WPS/PORTAL/FSIS/TOPICS%20/INSPECTION/MPI-DIRECTORY.

[71]   JUNG K, SAIF LJ. PORCINE EPIDEMIC DIARRHEA VIRUS INFECTION: ETIOLOGY, EPIDEMIOLOGY, PATHOGENESIS AND IMMUNOPROPHYLAXIS. THE VETERINARY JOURNAL. 2015;204(2):134–143.

[72]   HAGBERG AA, SCHULT DA, SWART PJ. EXPLORING NETWORK STRUCTURE, DYNAMICS, AND FUNCTION USING NETWORKX. VAROQUAUX G, VAUGHT T, MILLMAN J, EDITORS. PASADENA, CA USA: PROCEEDINGS OF THE 7TH PYTHON IN SCIENCE CONFERENCE (SCIPY2008); 2008.

[73]   FENICHEL EP, CASTILLO-CHAVEZ C, CEDDIA MG, CHOWELL G, PARRA PAG, HICKLING GJ, ET AL. ADAPTIVE HUMAN BEHAVIOR IN

EPIDEMIOLOGICAL MODELS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES. 2011;108(15):6306–6311. DOI:10.1073/PNAS.1011250108.

CHAPTER 2: NETWORK META-METRICS: USING EVOLUTIONARY COMPUTATION TO IDENTIFY EFFECTIVE INDICATORS OF EPIDEMIOLOGICAL VULNERABILITY IN A LIVESTOCK PRODUCTION SYSTEM MODEL

4.1 Abstract

We developed an agent-based susceptible / infective model which simulates disease incursions in the hog production chain networks of three U.S. states. Agent parameters, contact network data, and epidemiological spread patterns are output after each model run. Key network metrics are then calculated, some of which pertain to overall network structure, and others to each node's positionality within the network. We run statistical tests to evaluate the extent to which each network metric predicts epidemiological vulnerability, finding significant correlations in some cases, but no individual metric that serves as a reliable risk indicator. To investigate the complex interactions between network structure and node positionality, we use a genetic programming (GP) algorithm to search for mathematical equations describing combinations of individual metrics—which we call "meta-metrics"—that may better predict vulnerability. We find that the GP solutions—the best of which combine both *global* and *node*-level metrics—are far better indicators of disease risk than any individual metric, with meta-metrics explaining up to 91% of the variability in agent vulnerability across all three study areas. We suggest that this methodology could be applied to aid livestock epidemiologists in the targeting of biosecurity interventions,

and also that the meta-metric approach may be useful to study a wide range of complex network phenomena.

4.2 Introduction

This paper reports on an experiment that leverages network analytics and evolutionary computation to identify indicators of network structure and node positionality that predict epidemiological vulnerability within simulated livestock production chains. We use an agent-based model (ABM) to generate network graphs and employ network analytical techniques, statistical analysis, and evolutionary computation to investigate the extent to which either single network metrics, or combinations of metrics, may serve as indicators of infection risk. Understanding these relationships will aid livestock production practitioners, managers, and epidemiologists in targeting interventions which may preempt the spread of socioeconomically-important diseases through livestock production networks.

The field of animal health has received considerable attention due to the significant economic impacts on livestock enterprises caused by recent epidemics, as well potential implications associated with maintaining a stable, secure food supply. Studies of disease dynamics—specifically, targeted interventions to prevent outbreaks—has therefore become an important area of study for both scientists and policymakers (Schoenbaum & Disney, 2003).

The Regional U.S. Hog Production Network Biosecurity Model (RUSH-PNBM) v.1.2 is an ABM developed to assess disease spread through regional livestock

production systems. The model uses a GIS-based spatial framework, with three important hog-producing U.S. states—North Carolina, Iowa, and Illinois—defining the study areas. Three types of agents—hog producers, feed mills, and slaughter plants—interact through the transfer of livestock and feed, based on parameters including industry role, size, and proximity. The model is calibrated to examine the spread of a fecal-oral disease such as Porcine Epidemic Diarrhea virus (PEDv), transmitted by either infected animals, contaminated feed, contaminated slaughter plant receiving areas, or contaminated transportation equipment (Schulz & Tonsor, 2015). Expert advisory panels, coupled with available statistical datasets, guided the development, calibration, and validation of the model's heuristics and parameters, with the goal being to capture critical complexities underpinning epidemiological spread dynamics observed in the real world.

Because epidemics are fundamentally phenomena that propagate through networks (social, business, transportation, etc.), the generation of suitably-realistic graph structures that position agents as nodes, and inter-agent contacts as edges, is a basic design principle of RUSH-PNBM. Weighted edge lists built up by tracking inter-agent contact and infection spread patterns are output after each model run. Multiple runs across the three study areas provide a dataset upon which network analytics are applied to evaluate epidemiological vulnerability.

To analyze the data, first, *global* metrics capturing overall structures associated with each study area network are statistically correlated with both average and agent-

level vulnerability. Second, *node*-level metrics indicating network positionality are correlated with individual agents' vulnerabilities. Finally, realizing that epidemiological risk is contingent on a complex interaction between both *global* and *node* metrics, we employ a procedure to identify "meta-metrics"—using a Genetic Programming (GP) evolutionary algorithm—that correlate with the vulnerability of nodes across three networks with differing structural typologies. Overall, we find that the meta-metrics identified by the GP serve as much better indicators of epidemiological risk than any individual metric. While the focus of this study is livestock epidemiology, we suggest that this novel methodology may be applied to evaluate a variety of outcomes in a diverse range of networked systems.

The following background section will first cover the contributions of computational science to explanatory epidemiological models at a general level. We then discuss the application of agent-based models to the study of disease spread. Next, we highlight relevant theoretical contributions from the field of network analytics. Finally, we discuss previous research applying network science and computer modeling to livestock epidemiology.

4.3 Background

4.3.1 Computational Epidemiology

While providing a firm foundation, it has become evident that observational epidemiological studies yield limited insights wherever there are large geographical areas involved, large numbers of infection cases, multiple potential sources and/or paths

of infection, inherent uncertainty, incomplete data collection, and/or other complications. Further, proposed interventions may act at different levels, targeting, for example, production practices, animal movements, or human-behavioral factors; as well as being either preventive or reactionary in nature (Garner et al., 2007; Garner & Hamilton, 2011).

In recent years, computational modeling has increasingly been employed to probe such complexities, revealing fundamental characteristics of disease propagation within complex systems (Garner & Hamilton, 2011; Parker & Epstein, 2011; Perez & Dragicevic, 2009; Perez et al., 2002). Using such approaches, researchers have identified key infection nodes and disease pathways, evaluated health-policy scenarios utilizing both preventive and reactionary interventions, and projected the economic impacts of several disease incursion threats (Robins et al., 2015; Alderton et al., 2016; Belkhiria et al., 2016; Bradhurst et al., 2016; Tracey et al., 2014; Arruda et al., 2016; Bagni et al., 2002; Schulz & Tonsor, 2015).

The strength of the modeling approach lies in its ability to distill real-world systems down to their core processes. However, the complex nature of biological systems has led some to question the extent to which models should be relied upon to forecast real-world disease incursions (Moss, 2008). For example, during the United Kingdom's 2001 Foot-and-Mouth Disease (FMD) epidemic, flawed predictive models were used to inform the culling policy (Kitching et al., 2006). To prevent such occurrences, it is incumbent upon modelers to recognize the limitations of their chosen

115

approaches when drawing conclusions (Bousquet et al., 1999; Barreteau et al., 2003; Garner et al., 2007; Klügl, 2008). We note up-front that RUSH-PNBM in its current form is not intended as a forecasting tool, but rather to understand and quantify the interactions between network structures and disease spread dynamics within parallel real-world systems more generally.

4.3.1.1 Epidemiological agent-based models

ABMs constitute a class of complex systems models in which the global dynamics of a system emerge as a result of many individuals' decision heuristics and resulting interaction patterns, rather than being defined from the top down. Since they often incorporate stochasticity, ABMs can be difficult to validate, and they can also incur high computational overhead due to the quantity of calculations required (Bradhurst et al., 2016). On the plus side, ABMs provide several advantages over other methodologies when the goal is to discover previously-unexplored patterns that emerge from heterogeneous behavioral patterns, environmental factors, and especially when agent behaviors are affected by the state of other agents and/or a changing environment (Auchincloss & Diez Roux, 2008; Parker & Epstein, 2011; An, 2012; Shi et al., 2014; Kaul & Ventikos, 2013). Used as "virtual laboratories," ABMs can unveil insights into the inter-agent interaction patterns that underpin macro-level results (Macal & North, 2010).

ABMs focused on disease spread typically include two main components: within-host progression and between-host transmission (Hunter et al., 2017). Within-

116

host progression describes the process of a pathogen infecting a given host, running its course, and eventually dying out. A common means of modeling this is the SIR framework, in which susceptible (S) agents may contract a disease and become infective; infective (I) agents can transmit the disease to others; and removed/recovered (R) agents are conceived of as either dead, or having acquired immunity to the disease (Anderson & May, 1979). SI, a common variant, allows for repeated reinfections. In general, parameters including infection probability and average infection duration mediate the dynamics of within-host progression.

Between-host transmission occurs probabilistically when a susceptible agent comes into contact with an infective one. Agents enter into contact either at a rate governed by a differential equation, or in the case of ABMs according to the scheduling of their individual decision heuristics. Transmission probabilities may also be heterogeneous, depending on factors such as inter-agent distance, differing agent parameters, or network positionality. Modeling the interactions between a network of agents permits the simulation of realistic between-host paths, as well as subsequent analysis of individual agent vulnerabilities (Barrett et al., 2008).

4.3.2 Network Analytics and Spreading Dynamics

Network analytic techniques have provided insights into spreading behavior in a wide range of contexts, including information spread, diffusion of innovations, disease transmission, and other phenomena. Researchers compare and contrast spreading on different typologies of algorithmically-generated networks (i.e. random graphs, small-

117

world, scale-free, etc.), as well as analyzing complex graphs constructed from real-world datasets and computer simulations (Wasserman & Faust, 1994).

The application of network science to epidemics has received considerable attention due to the ability of network models to simultaneously depict the global structure of a population as well as the personal interactions between individuals (Bell et al., 1999; Christley et al., 2005; El-Sayed et al., 2012). For example, the transmission pathways that mediate the spread of a sexually-transmitted disease will differ significantly from those of an airborne-transmitted disease, with the former characterized by interpersonal contact networks (Killworth et al., 1998), and the latter by global transportation patterns (Colizza et al., 2007).

Many algorithms have been developed to measure and mathematically-codify aspects of network structure and node positionality (Albert & Barabási, 2002). Owing largely to continual advancements in this network-analytical toolkit, researchers have increasingly worked to identify network metrics that correlate with spreading dynamics. Below we first present research examining connections between global network structures and spreading dynamics, and second those which focus on individual node positionality within a network. We then describe how these insights have been applied to the study of epidemiological spread through livestock production networks.

4.3.2.1 The role of network structure on disease spread

Several studies have examined connections between the overall structure of a network and the propensity with which that network promotes disease spread. Christley

118

et al. (2005) use an SIR model to compare spreading in undirected random networks versus small-world networks (which have more heterogeneous degree distributions), finding faster spreading but ultimately fewer infected individuals in the small-world networks. The small-world networks had greater clustering and significantly higher k-core densities. In both network typologies, nodes in the k-core were at a higher risk of infection. Comparing unweighted centrality measures, the authors find that degree centrality is about as good a predictor of a node's infection risk as random-walk, shortest-path, or farness centrality, while being simpler to compute.

Salathé & Jones (2010) investigate the role of community structure on spreading in networks. Community structure (a.k.a. modularity) in a graph exists when nodes are densely connected in "cohesive subgroups," with only a few bridging connections between groups. Using SIR simulations, the authors find that targeting high-degree nodes for immunization can be effective in low-modularity graphs, but with more community structure, targeting bridges between communities—identified by betweenness or random-walk centrality—becomes more effective.

Kitsak et al. (2010) run SIR and SI models on four large, complex real-world networks. The authors demonstrate that a node's "coreness" or "core number"—as determined by the k-shell decomposition procedure—can in many cases be a better predictor of spreading propensity than either degree or betweenness centrality. Further, in their SI model results, the boundaries of k-cores tended to determine where an infection would become systemic versus die out. This suggests that coreness interacts

119

with traditional indicators of risk in complex ways: for example, a peripheral yet high-degree / high-betweenness node may be less infective than a lower-degree or lower-betweenness node within the core. These findings point to the need for further research into the interplay between the overall structure of a graph and the position of a node within that graph when estimating the node's infectivity.

<u>4.3.2.2 The role of node positionality on disease spread</u>

Research has also focused on correlations between the positionality of individual nodes and their resultant disease risk. In an early effort, Rothenberg et al. (1995) use survey data to build a network tracking the spread of HIV within a small community. While their results are largely inconclusive—likely owing to the very low proportion of infected individuals—they identify several important methodological insights. First, they draw a distinction between egocentric network analysis (i.e. individual risk based on network positionality) vs. sociocentic analysis (i.e. how macro-level network structures impact group-level outcomes). Second, they consider whether weighted or unweighted network metrics are better indicators of risk. It is unclear whether, for example, a person who comes into infrequent contact with many other individuals would be more or less vulnerable than a person who comes into more frequent contact with fewer individuals. Since measures of centrality, assortativity, and other metrics can be based on either weighted or unweighted degree, the distinction is important (Wasserman & Faust, 1994).

Bell et al. (1999) calibrate an SIR model based on interview data to reflect the network structure of a small group of individuals, and use model output data from a series of runs to investigate correlations between several indicators of network positionality and two dependent variables: (a) vulnerability, or the number of times a node became infected during a run; and (b) infectivity, the number of times a node spread the infection during a run. As in Rothenberg et al. (1994), both weighted (a.k.a valued) and unweighted (a.k.a. dichotomized) versions of metrics were analyzed. The authors find that vulnerability is best predicted by unweighted versions of eigenvector centrality, information centrality, degree centrality, and in-degree prestige; whereas infectivity was most highly correlated with weighted metrics including out-degree centrality, influence centrality, degree centrality, eigenvector centrality, and power prestige. Similar studies using empirically-derived infection-spreading networks also find that multiple centrality measures correlate with infection risk (De et al., 2004).

Ghani & Garnett (2000) develop a stochastic SI simulation model of gonorrhea transmission through large (N=2000) networks of social partners. The authors use nested Poisson regression models to compare the relative influence of each node's unweighted degree (k), concurrency, k at distance $= 2$, k at distance $= 3$, closeness, betweenness, and information centrality on both vulnerability and infectivity. They find that k is highly significant in all cases, and concurrency improves the model fit for both vulnerability and infectivity. Increased "local" connectivity (k at distance $= 2$) improves the model fit only for vulnerability, whereas indicators of centrality within the full

121

network, including closeness and betweenness, are important in predicting infectivity. The authors conclude that vulnerability is primarily contingent upon the structure of local network neighborhoods, whereas infectivity depends more on the interactions betweeen individual behavior and global network positionality.

4.3.2.3 Application to livestock epidemiology

In the case of livestock disease outbreaks, the networks that underlie the movement of livestock, feed, supplies, workers, and visitors are often the ones that pathogens follow as an epidemic spreads. In addition to direct disease transmission resulting from animal movements, disease can be also transmitted indirectly via contaminated feed, fomites, and transportation vehicles (Fèvre et al., 2006). Network models of livestock disease spread encode the contact patterns between producers and other actors such as auction houses, animal fairs, slaughter plants, and feed mills; each of which can be analyzed as a node in the overall network.

The United States Department of Agriculture Animal and Plant Health Inspection Service currently utilizes two stochastic state-transition SIR models— InterSpread Plus (ISP) and the North American Animal Disease Spread Model / Animal Disease Spread Model (NAADSM/ADSM)—to simulate disease incursions and evaluate the efficacy of control strategies (Stevenson et al., 2013; Harvey et al., 2007). These models leverage empirical data to generate spatially-explicit networks of livestock production nodes, and can simulate disease transmission via either animal movements or airborne spread.

122

Network analytics can illuminate the flow of disease within a production system and provide insights into effective control strategies (Dubé et al., 2009). For example, livestock movement data reveal the most common pathways over which infection was transmitted during the 2001 Foot and Mouth Disease (FMD) outbreak in the UK (Mansley et al., 2003; Webb, 2005; Kao et al., 2006; Kiss et al., 2006; Ortiz-Pelaez et al., 2006; Dubé et al., 2009). Similarly, analysis of Danish animal movement networks has allowed researchers to forecast disease spread through that system (Bigras-Poulin et al., 2006, 2007).

Using an SI model-based approach, Natale et al. (2009) leverage animal traceback data to build networks of Italian cattle supply networks, and examine the influence of "seeding node" positionality upon the final extent of an epidemic, finding that the eigenvector centrality and closeness of the seeding node are strongly correlated with epidemic size. The authors find power law degree distributions for both animal shipment sizes and the number of shipments per node; characteristics of a "scale free" network with "small world" properties. These properties have also been observed in other animal movement networks (Bigras-Poulin et al., 2006, 2007).

Wiltshire (2018) uses an agent-based SI model to analyze the impact of increased network density and producer specialization on the size and scale of disease outbreaks in U.S. hog production systems, finding evidence of percolation-type phase changes, with more-specialized systems becoming vulnerable to catastrophic outbreaks at lower density levels.

123

4.3.3 Contributions of This Study

RUSH-PNBM makes several advances over previous SIR-type livestock epidemic models, producing a network with thousands of agents, multiple agent types with heterogeneous interaction heuristics, and empirically-derived agent locations and operational parameters. Existing models such as Harvey et al. (2007) and Stevenson et al. (2013) also use empirical data to generate spatially-explicit livestock supply chain networks, but include only livestock production unit actors, whereas RUSH-PNBM also incorporates feed mills and slaughter plants, which have been implicated in disease incursions historically.

Natale et al. (2009) include slaughter plants in their model, but only the movement of infected animals is considered as a transmission vector, whereas RUSH-PNBM simulates transmission resulting from contaminated feed, slaughter plant receiving areas, and returning livestock transportation equipment as well. Thus, while in Natale et al. (2009) slaughter plants serve only as sinks for animals moving through the production chain, in our model they also function as hubs that may facilitate spreading.

A further contribution of RUSH-PNBM is that it realistically accounts for temporal concurrency. Rather than inter-agent contact occurring with an equal probability at each time step, network edges in RUSH-PNBM come into existence only when the state of each agent (for example its current inventory), together with its individual decision heuristics, dictate that a transfer of animals or feed should proceed.

124

Although the current version of RUSH-PNBM was built to study PEDv transmission in three U.S. states, its general structure allows for re-parameterizations focused on increased context specificity and/or adaptation to other diseases, study areas, or livestock species. This flexibility suggests that the model could serve as a valuable tool for practitioners to assess epidemic risk in a variety of livestock production contexts.

Following other studies investigating the interplay between network structures and spreading dynamics (Bell et al., 1999; Ghani & Garnett, 2000; Colizza et al., 2007; Webb, 2005; Kao et al., 2006; Kiss et al., 2006; Ortiz-Pelaez et al., 2006; Dubé et al., 2009; Bigras-Poulin et al., 2006, 2007; Natale et al., 2009), we analyze network graphs (in this case output from RUSH-PNBM) to determine the extent to which key network metrics correlate with nodes' epidemiological vulnerability. Rothenberg et al. (1995) suggest that the fusion of egocentric and sociocentric analytics would be a valuable approach for future research. Whereas in much of the previous work in this area only a single network is analyzed, our study compares and contrasts networks from three distinct study areas, allowing us to investigate the role of both egocentric (*node*-level) and sociocentric (*global*) metrics upon a node's vulnerability.

Our approach to determining the impact of each metric upon vulnerability is also innovative. In much of the previous work in this area, bivariate statistical methods are utilized, with some authors using multivariate regressions to investigate the relative impact of a suite of metrics taken together (Ghani & Garnett, 2000). Rather than relying

125

on traditional statistical techniques, we employ evolutionary computation to search for more-complex relationships between network metrics. Using a genetic programming algorithm, we obtain a set of "meta-metrics" along a complexity / fitness Pareto front. We find that the GP solutions predict infection risk in our data far better than any metric taken singly, lending support for the efficacy of this methodology.

4.4 Model Description

This section provides an overview of the basic features of the RUSH-PNBM v.1.2, including initialization procedures, agent interaction heuristics, the epidemiological sub-model, parameterization and validation, and sensitivity analysis. The model was built using Anylogic v.8 multimethod modeling software. For additional details, reference the RUSH-PNBM v.1.2 ODD+D Protocol (Appendix 2). Table 6 gives the model's global parameters, the values explored in this experiment, and the datasets and/or information sources used in parameter calibration.

Table 6: Model parameters and values explored in the experiment. "EAP" indicates that the value was derived through expert advisory panel sessions. "FHPC" refers to the family-owned hog production chain system dataset.

Parameter	Value(s)	Source
Initial network parameters		
Study area	[NC, IA, IL]	-
Number of producers in study area	[2217, 6266, 2045]	Burdett et al. (2015)
Avg. Producer Capacity	[4015, 3265, 2264]	Burdett et al. (2015)
Proportion Farrow to Wean	[0.050, 0.026, 0.038]	Burdett et al. (2015)
Proportion Farrow to Feeder	[0.005, 0.010, 0.009]	Burdett et al. (2015)
Proportion Farrow to Finish	[0.554, 0.304, 0.635]	Burdett et al. (2015)
Proportion Wean to Feeder	[0.102, 0.064, 0.023]	Burdett et al. (2015)
Proportion Wean to Finish	[0.003, 0.077, 0.055]	Burdett et al. (2015)
Proportion Feeder to Finish	[0.286, 0.519, 0.241]	Burdett et al. (2015)
Number of slaughter plants in study area	[24, 18, 25]	USDA NASS (2014)
Number of feed mills in study area	[40, 114, 37]	Google search; EAP
Producer to slaughter plant λ	2	EAP
Producer to feed mill λ	1.5	EAP
Disease parameters		
Percent of producers initially infected	5%	-
Avg. producer infection duration (days)	40	Goede & Morrison (2016); EAP
Avg. slaughter plant infection duration (days)	7	EAP
Avg. feed mill infection duration (days)	25	EAP
Suckling pig mortality rate on infection	0.9	Goede & Morrison (2016); EAP
Nursery pig mortality rate on infection	0.4	Goede & Morrison (2016); EAP
Grow/finish hog mortality rate on infection	0.1	Goede & Morrison (2016); EAP
Producer disease spread probabilities		
Prob. producer will become infected if returning pig truck is contaminated	0.3	EAP
Prob. producer will become infected if feed truck is contaminated	0.8	EAP
Prob. feed truck will become contaminated if producer is infected	0.05	EAP

Prob. pig truck will become contaminated if producer is infected	0.2	EAP
Feed mill disease spread probabilities		
Prob. feed mill will become infected if returning feed truck is contaminated	0.1	EAP
Prob. feed truck will become contaminated if feed mill is infected	0.5	EAP
Slaughter plant disease spread probabilities		
Prob. slaughter plant receiving area will become infected if pig batch is infected	0.4	EAP
Prob. pig truck will become contaminated if receiving area is infected	0.2	EAP
Producer farrow, wean, and batch parameters		
Farrow to wean sow proportion (relative to total capacity)	0.6	EAP
Farrow to feeder sow proportion (relative to total capacity)	0.5	EAP
Farrow to finish sow proportion (relative to total capacity)	0.2	EAP
Annual number of piglets per sow	34	The Pig Site (2014)
Max. wean and batch frequency (days)	7	EAP
Min. batch size (as proportion of non-sow capacity)	0.05	EAP
Capacity under which producer has only one batch	20	EAP
Producer livestock transfer parameters		
Min. capacity similarity ratio	25	EAP
Max. producer to producer shipment distance (km)	150	EAP
Max. number of potential producer to producer trading partners	15	FHPC
Max. producer to producer shipment frequency (days)	5	FHPC
Feed mill parameters		
Avg. number of daily feed delivery trips per mill	10	FHPC
Number of producers visited per feed delivery trip λ	1	EAP

4.4.1 Agents and Initialization

Three types of hog production chain network agents, identified by industry experts as critical players in the transmission of fecal-oral diseases, are represented in the model, these being (a) hog producers, (b) feed mills, and (c) slaughter plants. Producer agents are assigned one of six industry roles based on the classification system used by the USDA and other industry analysts, these being (a) Farrow to Wean, (b) Farrow to Feeder, (c) Farrow to Finish, (d) Wean to Feeder, (e) Wean to Finish, and (f) Feeder to Finish.

The spatial extent of the model is defined by the boundaries of one of three U.S. states: (a) North Carolina, (b) Illinois, or (c) Iowa. These study areas were chosen because each is a major producer of hogs, yet key differences exist regarding the size of their networks and distribution of industry actors, allowing us to incorporate the impact of differences in global network structure into our analysis.

We use the Farm Location and Agricultural Production Simulator (FLAPS) tool—which draws upon USDA Census of Agriculture data along with aerial imaging to impute realistic distributions of livestock farms within a specified U.S. region (Burdett et al., 2015)—to set producer agent locations and key operational parameters including industry roles and capacities. Following FLAPS initialization, each producer generates a list of potential producer trading partners, constrained by maximum distance, size similarity, and maximum number parameters.

Livestock in the model are represented in batches (or metapopulations) of animals of the same age. Producer agents are initialized with one or several pig batches—depending on their capacity—within the correct age range for their industry roles. Farrowing producers periodically generate new piglets, which are subsequently batched as weaner pigs at a rate dependent on their sow inventory, according to industry standards.

Because FLAPS only covers animal production units, feed mill and slaughter plant locations are initialized by distributing them at random positions within each county in proportion to the number of producers in the county, with the overall numbers per state being derived from available datasets in conjunction with expert advisory panels. These non-producer agents are assigned service areas at initialization by having each producer agent connect to the nth-closest of each type, with n being drawn from Poisson distributions calibrated in consultation with industry experts. For example, in the case of feed mills, $\lambda=1.5$, indicating that a producer is most likely to purchase feed from the first- or second-closest mill, with fewer purchasing from the third-closest, fewer still from the fourth-closest, etc. Alternatively, for slaughter plants, $\lambda=2$, indicating that the most likely outcome is for a producer to ship hogs to the second-closest plant. A limitation of RUSH-PNBM is that livestock and feed transfers occur only between model agents; thus, since the model's spatial extent is a single U.S. state, interstate transportation and trade is not represented.

4.4.2 Behavioral Heuristics

Agent behavior rules are derived from a review of the primary literature as well as from industry expert advisory panels. A schematic showing the basic interaction flow appears in Figure 13. Livestock are transferred to appropriate agent(s) upon reaching the designated transfer age, and may be split and sent to multiple trading partners. Feed mill agents periodically generate delivery routes that visit a subset of producer agents within their service areas, with the number of deliveries per route being drawn from a Poisson distribution with $\lambda=1$.
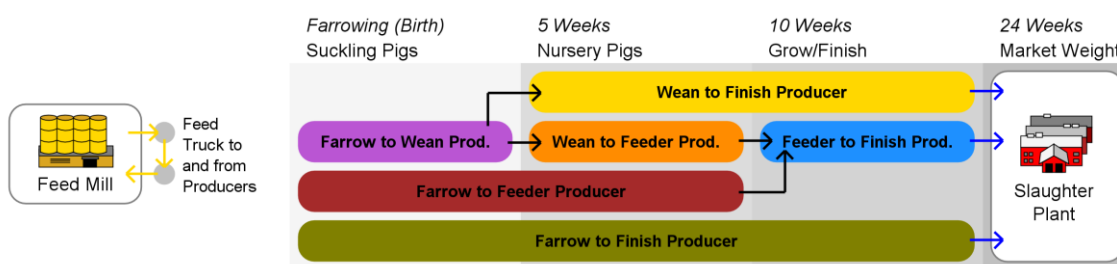
Figure 13: Agent connectivity key. Connections also illustrate potential disease transmission vectors.

4.4.3 Epidemiological Sub-Model

Agents in RUSH-PNBM exist in one of two states: susceptible (S) or infective (I). Infection spread becomes possible whenever a susceptible agent comes into contact with an infective agent. Several modes of inter-agent contact are represented in the

model (Figure 13), each corresponding to a potential disease transmission vector within the epidemiological sub-model.

Infected livestock transferred to a susceptible producer automatically infect the recipient. If a susceptible producer transfers livestock to an infective premises, there is a small probability that the infection may be brought back in the form of biological material which has contaminated the transportation equipment. Once a producer agent is infected, it is assumed that its entire premises becomes infected, due to the high reported virulence of PEDv once in a herd (Goede & Morrison, 2016). The infection event triggers a mortality calculation that decrements the infected agent's livestock inventory according to observed mortality rates from PEDv appropriate for the age of each pig batch. Producers remain infective for an average duration of 40 days before transitioning back to susceptible. If a producer ships out its entire livestock inventory, it is assumed the the premises is disinfected prior to receiving a new shipment.

Feed mills may become infected if a delivery truck that has previously been contaminated (by visiting an infective producer) returns to the mill. Mills remain infective for an average of 20 days. While a feed mill is infective, each delivery truck departing the mill may become contaminated. The truck may also become contaminated upon visiting an infective producer part-way through a route. If a contaminated truck visits a susceptible producer, that producer may be infected upon receiving a delivery of contaminated feed.

Slaughter plant receiving areas may become contaminated upon receiving a shipment of infected hogs. The plant's receiving area will remain infective for an average of five days, during which arriving transportation equipment may become contaminated. Transportation equipment that has been contaminated in this way may then spread the infection upon returning to the originating producer.

To skip the transient period and allow agent interactions to stabilize, the initial infection event, which randomly infects five percent of producer agents, occurs one model year post initialization.

4.4.4 Parameter Calibration and Validation

Due to the lack of precise animal movement data coupled with the inherent variability of epidemic events within complex networked systems, we are interested less in empirically-validating the model to be used as a forecasting tool, and more in developing sufficient structural- and face-validity to allow for a deeper understanding of the fundamental dynamics of the modeled systems (Klügl, 2008). Even given identical starting conditions, deviations in contact patterns over the course of a real-world disease incursion often render precise forecasts unfeasible (Moss, 2008). Our aim is rather to uncover and better understand the network features that lead to epidemiological vulnerability in livestock production systems more generally (Epstein, 2008).

Calibration procedures that leverage concrete historical data are often regarded as the best way to bring a model in line with empirical evidence. Unfortunately, there is

a marked lack of publicly-available data in the agricultural sector beyond aggregated county- or state-level statistics. To the extent that datasets containing explicit locations, operational parameters, livestock and feed movements, and disease histories exist; these data tend to be held by private enterprises, which view them as sensitive internal records. In light of this, following Windrum et al. (2007), we employ several alternative calibration procedures that have been widely-used in previous modeling endeavors in which fine-grained data are scarce.

The spatial locations and basic operational parameters of RUSH-PNBM agents associated with each study area are calibrated using the "indirect" approach, whereby stylized facts about the distribution of agents in the system are gleaned from statistical datasets (Windrum et al., 2007). While the FLAPS tool discussed above (Burdett et al., 2015) serves as our primary means to set agent locations and operational parameters, our team also gained access to internal records from a large family-owned hog production chain system, which are used to codify realistic contact rate and shipment size parameters (see Appendix 2).

Calibration of model elements that define how and when inter-agent contact occurs, as well as epidemiological sub-model parameters, were iteratively honed throughout the model development process following the "companion modeling" approach (Bousquet et al., 1999; Barreteau et al., 2003). During the scoping phase of model development, our research team convened two expert panels that included leading policymakers drawn from industry associations (e.g. the National Pork Board)

and the USDA, leading veterinary scientists, Agricultural Extension agents focused on livestock biosecurity from several states, veterinarians working with major livestock production chains, industry communications specialists, and agricultural economists. An additional two panels, involving many of these same experts, were convened during the model's development. Finally, a fifth expert panel was convened to discuss and interpret the results of the model runs shown here.

In these panel sessions, we used both targeted focus groups as well as quantitative survey questionnaires to elicit and hone parameter values. Through this process, model assumptions, data streams, and behavioral heuristics were shared, critiqued, and refined; and implications of model results were discussed. Detailed notes were taken during these expert panels, which were subsequently coded and analyzed. This information was used to bolster the face-validity of the distribution of epidemic patterns, scales, and durations produced by the model. Using this participatory methodology, the modeled system was brought in line with the collective understandings of stakeholders who are intimately familiar with the operational details of U.S. livestock production chains.

A common perception among our expert panelists concerned the complexity of disease transmission, the variation of disease spread across different states and production chains, and the relatively-meager understanding industry and USDA professionals possesses relative to identifying effective leverage points in production chains that are both most susceptible to disease, and most critical to its spread. Given

135

the complexity of the production chain—with its segmentation of livestock producer roles from farrow to finish—and the importance of feed mills and slaughter plants, industry experts called for the development of an ABM that could capture differences between regions and the configurations of their associated production chain networks.

4.4.5 Sensitivity Analysis

We conducted a sensitivity analysis focusing on four key model parameters. While the model includes many parameters (Table 6), these four were chosen because each focuses on a specific aspect of the model's network and/or epidemiological architecture. *Prob. producer to pig truck infection* highlights the impact of outgoing disease transmission from a producer agent to either another producer or a slaughter plant. *Avg. producer infection length* is an infection duration parameter, to which previous SIR-type models have been found to be sensitive. *Max producer connection distance* evaluates the effect of altering the global network structure from more localized to more spatially-diffused neighborhoods. Finally, *Prob. feed truck to feed mill infection* aims at elucidating the relative impact of infections stemming from the feed distribution network (versus the livestock transportation network) on system-wide disease resilience.

Each parameter is varied in steps between 50% and 150% of the baseline values given in Table 6, with ten replications per step. Table 7 shows the elasticity of the response variable—average vulnerability—across this range for each study area. Figure 14 visualizes the sensitivity analysis data and provides correlation statistics.

136

Table 7: Elasticity of response variable (Avg. Vulnerability) resulting from variation between -50% and 150% of baseline values for four key parameters.

Parameter	% Change in Avg. Vulnerability			
	North Carolina	Iowa	Illinois	All Study Areas
Prob. producer to pig truck infection	49.93	109.9	491.1	94.56
Avg. producer infection length (days)	2382	3164	6077	3010
Max. producer connection distance (km)	2.182	-20.32	108.6	7.976
Prob. feed truck to feed mill infection	36.21	114.8	119.4	50.82

Figure 14: Sensitivity analysis plots for four key parameters. Scatter points show average values at each step, colored regions show 95% CIs, and dashed lines show linear trends. Blue represents North Carolina, red Iowa, green Illinois, and black the combined dataset. Pearson correlation coefficients, *p*-values, and R^2 values of linear regressions appear in legends below each figure.

Results show that the model is moderately-sensitive to changes in *Prob. producer to pig truck infection* and *Prob. feed truck to feed mill infection*; with the

effect on average vulnerability being positive and significant in all study areas. It would appear that *Prob. producer to pig truck infection* has a larger impact on the response variable in Illinois, whereas *Prob. producer to pig truck infection* has the largest impact in North Carolina, reflecting differences in network structure between study areas. Despite state-to-state differences, the fact that increased infectivity in both hog transportation and feed distribution networks both led to similar increases in overall epidemiological vulnerability suggests that our epidemiological spread sub-model is relatively balanced between these two major modes of transmission. This ground-truths our model with respect to real-world findings implicating both contaminated feed and infected hogs / transportation equipment in the spread of PEDv and other diseases (Schulz & Tonsor, 2015).

We find that the model is not particularly sensitive to *Max producer connection distance*. The magnitude and direction of the effect varies between study areas, with only Illinois demonstrating a significant (positive, in this case) relationship to the response variable. While somewhat surprising, this result likely reflects the fact that, despite increasing the *Max producer connection distance*, the *Max. number of potential producer to producer trading partners* parameter remained constant (at 15) across all runs. This means that the average *k* of a given producer would only increase with *Max producer connection distance* in more spatially-diffused networks wherein a sizable proportion of producers did not have 15 potential producer trading partners within the baseline maximum connection distance of 150km. This is more likely to be the case in

Illinois, as there are relatively fewer producers within a relatively large area in that state.

By contrast, increasing *Avg. producer infection length* causes significant increases in average vulnerability across all study areas. In light of previous SIR / SI model studies, the observation that average infection duration heavily impacts average vulnerability is not a surprise. Further, the shape of the elasticity curves in the top right of Figure 14 suggest that percolation dynamics may exist, with the nonlinearity—or percolation threshold—being lowest for North Carolina and higher for the other two study areas, corroborating findings from Wiltshire (2018).

4.5 Experimental Design and Data Processing

The model was executed 50 times for each of the three study areas, maintaining all common parameters constant across repetitions. The total run count of 150 was chosen due to the large number of agents in each study area, along with the relative complexity of the model's livestock trade and epidemiological spread heuristics, rendering each run—along with its associated data analysis—quite computationally-intensive. The full experimental results encompass 539,300 individual agent-run combinations, with all network metrics and epidemiological statistics calculated for each, yielding a substantially large and rich dataset totaling approximately two million unique datapoints.

The numbers and spatial distributions of each type of agent were hard-coded and maintained across runs within each study area (Table 6). Further, immutable agent

140

properties set at model initialization—including the pool of potential trading partners—were also held across repetitions within each study area. After initialization, all stochastic events in a run utilize a random seed, meaning that—while the underlying network structure associated with each study area production chain remains constant—the transportation of livestock and feed, along with the resultant outbreak progression, is unique from one run to the next.

Throughout each run, a weighted, directed contact network was built up by tracking the number of times each agent contacted another agent—as a result of either delivering or receiving livestock or feed—with the final edge weights being equal to the number of contacts between connected agents. In a similar fashion, an infectivity network was constructed, with edge weights tracking the number of times the infection was passed between agents over the course of a run. An agent's vulnerability—the dependent variable used in the data analyses below—is defined as a node's in-degree within the infectivity network.

Networks generated during the experiment, along with key statistics associated with each agent, were output as tabular data at the conclusion of each run. Using the Python NetworkX 2.0 library (Varoquaux et al., 2008), the weighted edgelists describing each network were used to generate weighted, directed network graphs, allowing key *global* and *node*-level network metrics previously linked to epidemiological spreading to be evaluated (Ghani & Garnett, 2000; Colizza et al., 2007; Webb, 2005; Kao et al., 2006; Kiss et al., 2006; Ortiz-Pelaez et al., 2006; Dubé et al.,

2009; Bigras-Poulin et al., 2006, 2007; Natale et al., 2009). Where applicable, both weighted and unweighted versions of these metrics were calculated, in order to compare between the two (Rothenberg et al., 1995; Bell et al., 1999). The metrics utilized in this study are listed and briefly described inTable 8. For clarity, we distinguish between the set of *node*-level metrics pertaining to centrality, versus those capturing other aspects of node positionality. The full sets of *global* and *node* metrics used in the analysis were:

$$global=[\langle k\rangle,\langle k_w\rangle,k^{CV},k_w^{CV},\langle C\rangle,\langle C_w\rangle,r,r_w,p^{c_{max}}]$$

$$node=[k,k_w,k^{in},k_w^{in},k^{out},k_w^{out},C,C_w,c,C^B,C_w^B,C^{RW},C_w^{RW},C^E,C_w^E,C^D]$$

Table 8: Network metrics used in the GP analysis. Weighted (W) and unweighted (UW) statistics are indicated.

	Description	Ref.
Global Metrics		
UW Average Degree ($\langle k\rangle$)	Number of edges divided by number of nodes	(Albert & Barabási, 2002)
W Average Degree ($\langle k_w\rangle$)	Average number of contact events per node per run	(Albert & Barabási, 2002)
UW Degree CV ($\langle k^{CV}\rangle$)	Coefficient of variation for unweighted degree across all nodes	(Newman, 2003)
W Degree CV ($\langle k_w^{CV}\rangle$)	Coefficient of variation for weighted degree across all nodes	(Newman, 2003)
UW Average Clustering ($\langle C\rangle$)	Number of closed triplets divided by total number of triplets	(Luce & Perry, 1949)

W Average Clustering ($\langle C_w \rangle$)	Geometric average of the subgraph edge weights	(Saramäki et al., 2007)
UW Assortativity Coefficient ($\langle r \rangle$)	Level of similarity between the unweighted degrees of all nodes	(Newman, 2003)
W Assortativity Coefficient ($\langle r_w \rangle$)	Level of similarity between the weighted degrees of all nodes	(Newman, 2003)
k-core Fractional Size ($p^c max$)	Fraction of nodes that are within the main k-core	(Batagelj & Zaversnik, 2003)
Node Centrality Metrics		
UW Shortest Path Betweenness (C^B)	Fraction of shortest-paths passing through a node	(Freeman, 1977)
W Shortest Path Betweenness (C_w^B)	Same as above, but incorporating edge weights	(Freeman, 1977)
UW Random Walk Betweenness (C^{RW})	Fraction of random walks passing through a node	(Newman, 2005)
W Random Walk Betweenness (C_w^{RW})	Applies an electric current flow model across all nodes	(Brandes & Fleischer, 2005)
UW Eigenvector (C^E)	Nodes connected to other high-k nodes receive high scores	(Bonacich, 1987)
W Eigenvector (C_w^E)	Same as above, but incorporating edge weights	(Bonacich, 1987)
Degree (C^D)	Fraction of nodes to which a node is connected	(Albert & Barabási, 2002)
Node Positionality Metrics		
UW In-Degree (k^{in})	Number of incoming edges	(Albert & Barabási, 2002)
W In-Degree (k_w^{in})	Total weight of incoming edges	(Albert & Barabási, 2002)
UW Out-Degree (k^{out})	Number of outgoing edges	(Albert & Barabási, 2002)
W Out-Degree (k_w^{out})	Total weight of outgoing edges	(Albert & Barabási, 2002)
UW Clustering Coefficient (C)	Fraction of possible triangles through a node that exist	(Luce & Perry, 1949)
W Clustering Coefficient (C_w)	Geometric average of the subgraph edge weights	(Saramäki et al., 2007)
Coreness (c)	Largest value k of a k-core containing the node	(Batagelj & Zaversnik, 2003)

To generate meta-metrics—that is, formulae composed of multiple individual indicators—we utilized the Eureqa GP package (Schmidt & Lipson, 2009). Due to the large size of our dataset, we first sampled the data by randomly selecting 100 agents from each study area to include in the GP training and validation sets. Because of the lower numbers of feed mills and slaughter plants compared to producers, we ensured that all agent types were represented in the data by stipulating that a minimum of five agents of each type from each study area were included, yielding 16,050 total rows. The GP was trained on half of these data (selected randomly) and validated on the remaining half.

The GP algorithm was permitted to use constants, input variables, the four basic arithmetical operators, as well as exponential, logarithmic, and power functions; each of which was assigned a complexity value by which to weigh fitness vs. complexity tradeoffs between potential solutions (Table 9).

Table 9: Operators used in the GP analysis, with their assigned complexity values.

Operator	Complexity
Constant	1
Input Variable	1
Addition	1
Subtraction	1
Multiplication	1
Division	2
Exponential	3
Natural Logarithm	3
Power	3

The objective set for the GP was to maximize the R^2 coefficient between the function output and each node's vulnerability v—at varying levels of complexity—according to the function:

$$v=f(global,node)$$

The GP was allowed to run until 100% convergence was indicated. The algorithm executed 1.6×10^{10} function evaluations over the course of 112,378 generations. 14 total solutions were identified, ranging from *complexity* = 1 and $R^2 = 0.04$; to *complexity* = 24 and $R^2 = 0.96$.

4.6 Results

To analyze our experimental data, we begin by characterizing high-level properties of the production chain networks output by RUSH-PNBM. This allows for a generalized understanding of the similarities and differences between the hog production networks associated with our three study areas, as well as the basic connection patterns and epidemiological vulnerabilities associated with each agent type. Next, we examine the extent to which network metrics can predict individual nodes' infection vulnerabilities across the three states, examining both network-structural and node positionality factors. Finding that no individual metric adequately correlates with infection risk across study areas, we turn to the GP results to investigate

145

the extent to which formulae composed of multiple metrics—which we call "meta-metrics"—may predict epidemiological vulnerability.

4.6.1 Characterizing Study Area Production Networks

Figure 15 shows the degree distributions for each class of agent and study area. Weighted degree (k_w) represents the number of times an agent contacted another agent (i.e., sending or receiving either livestock or feed) over the course of a run, whereas unweighted degree (k) describes the number of unique agents with which an agent had contact.
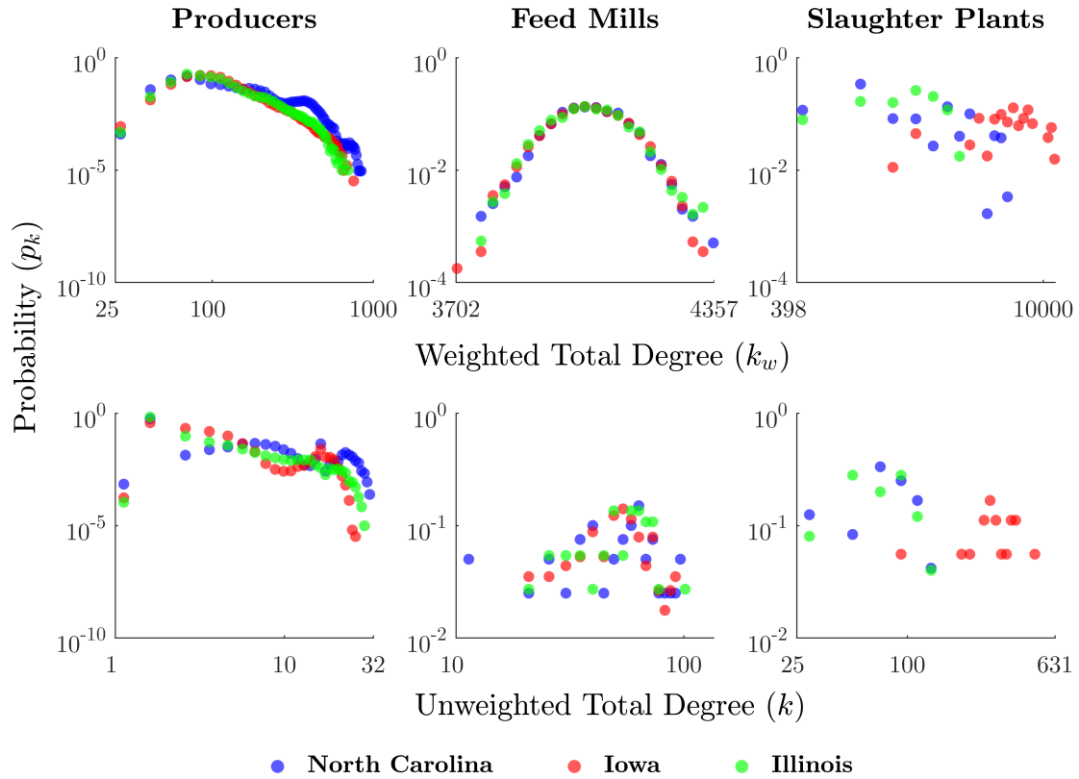
Figure 15: Weighted and unweighted degree distributions by agent type for each study area, across all runs. Plotted with log-scaled *x* and *y* axes. 60 bins were used for producers, and 25 for feed mills and slaughter plants.

Overall, producers tend to have the lowest k_w and k, followed in order by feed mills and slaughter plants. Feed mills' degrees are constrained to a smaller range, whereas the distributions for producers and slaughter plants are quite long-tailed. Slaughter plants in Iowa appear to have somewhat higher degrees on average than in the other study areas.

A linear degree distribution in log-log space demonstrates the power law relationship associated with scale-free networks, satisfying $P[X \geq x] \sim cx^{-\alpha}$ (Barabási & Albert, 1999). This appears to be roughly the case for producers, but it is less clear whether the other agent classes exhibit a power law degree distribution.

Additionally, we observe that the degree distribution for producers appears to have a different pattern in North Carolina compared with Iowa and Illinois, with North Carolina having more high-k producers. An examination of Figure 16 suggests that this may be related to the heavily-connected cluster surrounding Duplin County in the southeast of North Carolina.

North Carolina:
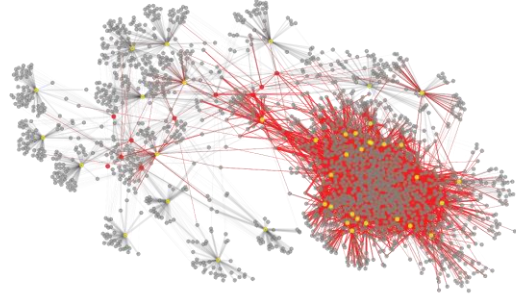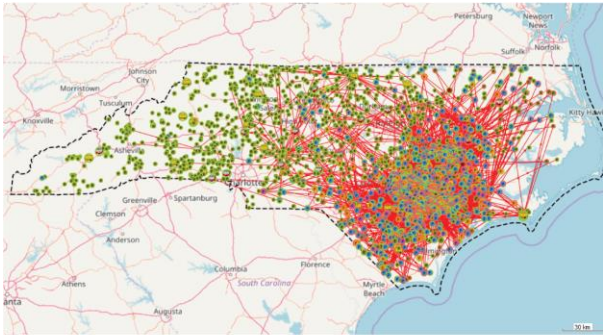
Iowa:
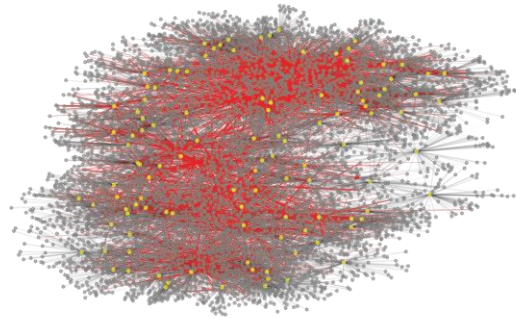
Illinois:
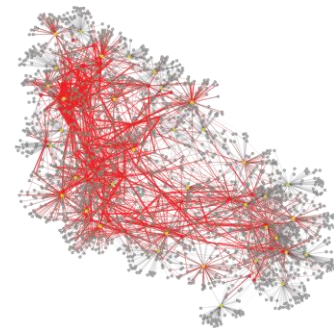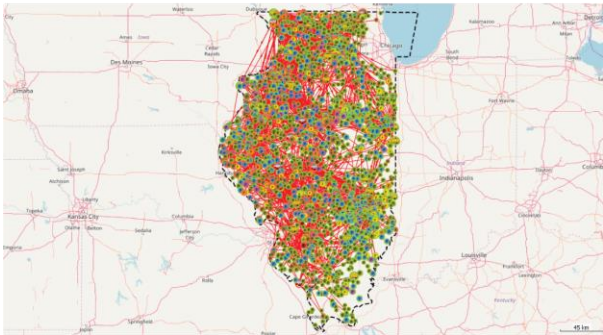
Figure 16: Sample model-generated production chain maps (left), and corresponding network graphs (right), for each study area. Maps show agents of each type (see Figure 13 for key) and infection-spreading edges in red. Graphs show contact (black) and infection-spreading (red) edges. Nodes are positioned using a spring layout. Gray nodes are producers, yellow are feed mills, and red are slaughter plants.

149

Table 10 gives epidemiological statistics and key network-structural indicators, averaged by study area. We find notable differences in infection dynamics between study areas, with North Carolina agents having the greatest chance of receiving the infection at least once during a model run, at $p^{inf} \approx 46\%$, versus 14% for Iowa, and 17% for Illinois. Agents in North Carolina also had the highest average vulnerability ($\langle v \rangle \approx 6$), as well as the longest average total infection duration ($t^{inf} \approx 101$ days).

Table 10: Mean values and 95% confidence intervals for epidemiological statistics and network metrics across agents in each study area.

	North Carolina			Iowa			Illinois		
	Mean	**95% CI**		**Mean**	**95% CI**		**Mean**	**95% CI**	
Epidemiological Statistics									
Vulnerability (v)	5.949	5.833	6.064	0.6504	0.6387	0.6620	0.9682	0.9480	0.9883
Infectivity (i)	5.949	5.858	6.039	0.6504	0.6310	0.6697	0.9682	0.9372	0.9992
Infection Probability (p^{inf})	0.4634	0.4606	0.4663	0.1423	0.1411	0.1435	0.1703	0.1680	0.1725
Spread Probability (p^{spread})	0.6253	0.6225	0.6281	0.3440	0.3424	0.3457	0.3403	0.3374	0.3432
Infection Duration (t^{inf})	101.2	100.5	101.9	14.97	14.87	15.07	19.94	19.69	20.19
Network Metrics									
Unweighted Degree (k)	7.910	7.838	7.982	6.226	6.161	6.292	5.298	5.228	5.367
Weighted Degree (k_w)	240.6	237.3	244.0	205.4	203.2	207.6	204.3	201.0	207.6
Unweighted In-Degree (k^{in})	3.955	3.904	4.006	3.113	3.053	3.173	2.649	2.596	2.702
Weighted In-Degree (k_w^{in})	120.3	118.8	121.8	102.7	101.4	104.0	102.1	100.9	103.4
Unweighted Out-Degree (k^{out})	3.955	3.903	4.006	3.113	3.085	3.142	2.649	2.601	2.696
Weighted Out-Degree (k_w^{out})	120.3	117.3	123.4	102.7	100.9	104.5	102.1	98.96	105.3
Unweighted Degree CV (k^{CV})	1.565	1.565	1.565	3.047	3.047	3.047	2.172	2.172	2.172
Weighted Degree CV (k_w^{CV})	2.385	2.385	2.385	3.110	3.110	3.110	2.703	2.703	2.703
Unweighted Clustering Coefficient (C)	0.005178	0.0050499	0.0055307	0.0078791	0.0077723	0.0080199	0.0088166	0.0085554	0.0090788
Weighted Clustering Coefficient (C_w)	0.00034997	0.00034810	0.00035385	0.001100	0.0010078	0.0011123	0.0012003	0.0011665	0.0012442
Unweighted Assortativity Coefficient (r)	-0.5009	-0.5009	-0.5009	-0.2490	-0.2490	-0.2490	-0.5859	-0.5859	-0.5859

Weighted Assortativity Coefficient (r_w)	-0.2989	-0.2989	-0.2989	-0.3610	-0.3610	-0.3610	-0.4931	-0.4931	-0.4931
k-Core Fractional Size ($p^c max$)	0.2101	0.2100	0.2101	0.2673	0.2673	0.2673	0.1380	0.1377	0.1383
Node Coreness (c)	4.481	4.464	4.499	3.328	3.323	3.332	2.787	2.780	2.795
Centrality Indicators									
Unweighted Shortest Path Betweenness (C^B)	0.001415	0.001384	0.001447	0.0005033	0.0004919	0.0005147	0.001591	0.001557	0.001625
Weighted Shortest Path Betweenness (C_w^B)	0.002875	0.002805	0.002945	0.0009989	0.0009845	0.0010103	0.0034330	0.0033346	0.0035133
Unweighted Random Walk Betweenness (C^{RW})	0.003213	0.003161	0.003266	0.001188	0.001172	0.001203	0.003534	0.003476	0.003592
Weighted Random Walk Betweenness (C_w^{RW})	0.003401	0.003341	0.003461	0.001319	0.001300	0.001338	0.003870	0.003800	0.003940
Unweighted Eigenvector (C^E)	0.01379	0.01370	0.01389	0.004814	0.004774	0.004854	0.01565	0.01556	0.01574
Weighted Eigenvector (C_w^E)	0.001365	0.001244	0.001486	0.001240	0.001197	0.001283	0.002102	0.001971	0.002233
Degree (C^D)	0.003469	0.003438	0.003501	0.0009733	0.0009630	0.0009836	0.002516	0.002483	0.002549

Examining the averaged network metrics provides some initial insights into these disease resilience discrepancies. While the Iowa networks have significantly more nodes, the North Carolina networks are the most densely-connected (higher $\langle k \rangle$ and $\langle k_w \rangle$). Differences are also apparent in the indicators associated with the networks' k-cores as well as average centrality values. In the sections below, we statistically evaluate the ways in which these patterns in the network data relate to epidemiological vulnerability.

4.6.2 Can Network Metrics Predict Infection Risk?

Since the probabilities of disease transmission and agents' behavioral heuristics do not change across runs, the disease risk discrepancies noted above must necessarily result from differences in contact networks. We use a three-pronged approach to analyze the relationships between network structure and epidemic risk factors in the RUSH-PNBM output data. First, we examine the relationships between *global* metrics that capture overarching features of each study area network, versus the average vulnerability of nodes in these networks, as well as individual agents' vulnerabilities. Second, we examine the relationships between *node*-level metrics that capture various aspects of positionality, versus each agent's vulnerability. Finally, we turn to the GP results to identify meta-metrics that may serve as better indicators of epidemiological risk.

4.6.2.1 Global network-structural factors

Figure 17 plots the *global* network metrics described in Table 8 against the average vulnerability of agents in each run. These findings may be analyzed in two ways: first, we can determine whether an overall trend exists across all three study areas; second, we can determine whether a trend exists across the 50 runs associated with each study area. Overall, we find that all global metrics with the exception of *k*-core fractional size do indeed correlate significantly with average vulnerability across all runs.

153
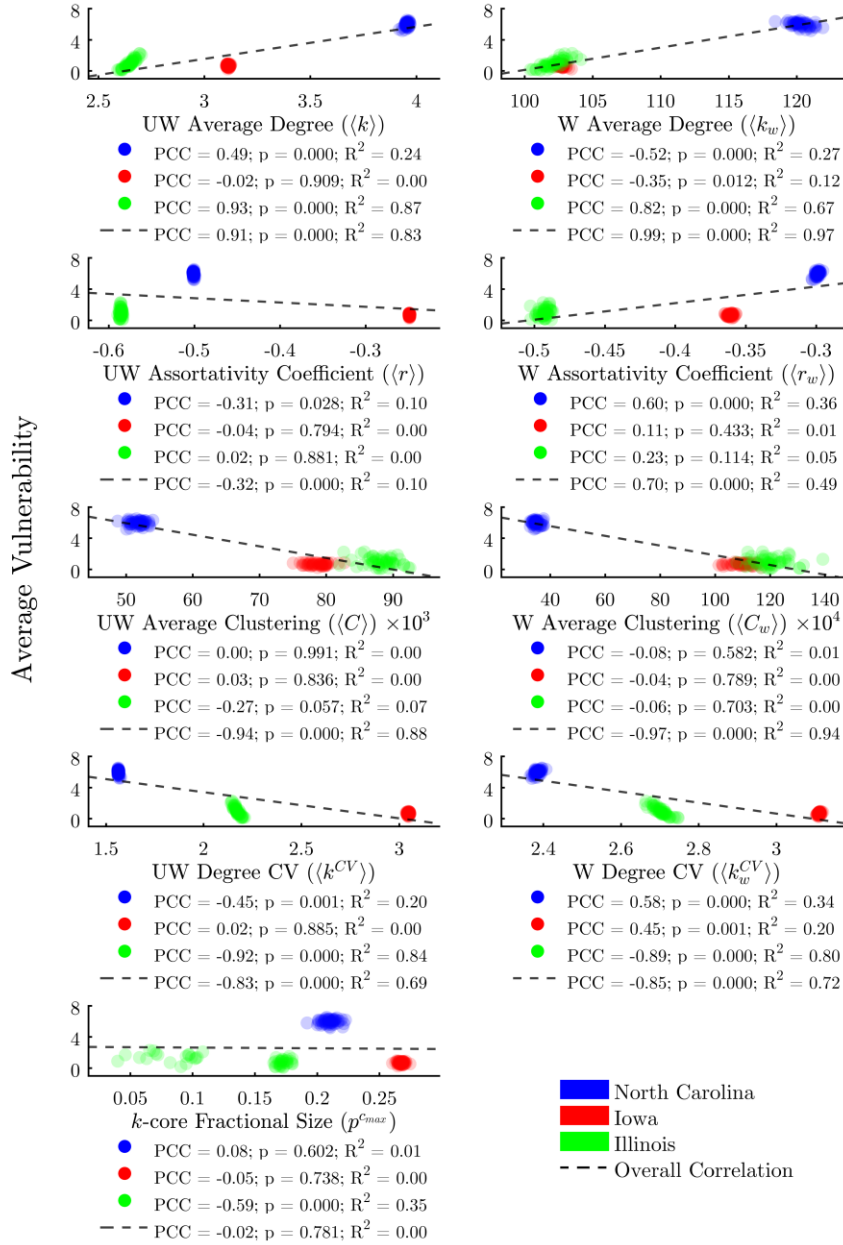
Figure 17: Correlations between key network-level metrics and the average vulnerability of nodes in each run. Weighted (W) and unweighted (UW) statistics are indicated. Scatter points are averages for each run, color coded by study area. Black lines show best fit, with Pearson correlation statistics and R^2 given in legends both by state and overall.

However, the situation is more complex when examining within-study-area trends. We find that weighted and unweighted average degree, unweighted assortativity, unweighted clustering, weighted and unweighted degree CV, and *k*-core fractional size all exhibit positive correlations in some states and negative in others. Further, the within-state correlations do not achieve significance for several of the metrics. No metric we assessed significantly correlates with average vulnerability both overall as well as within each study area.

While some *global* network metrics correlate strongly with the average vulnerability of agents in each network, we find that global metrics alone do little to predict the infection risk of individual nodes, owing to the wide range of vulnerabilities between agents. Figure 18 evaluates the same metrics as Figure 17, only this time we plot the individual vulnerability of each agent, rather than per-run averages. Here we find that, in fact, none of the *global* metrics taken alone is sufficient to explain more than 5% of the variability in the vulnerability of individual agents.
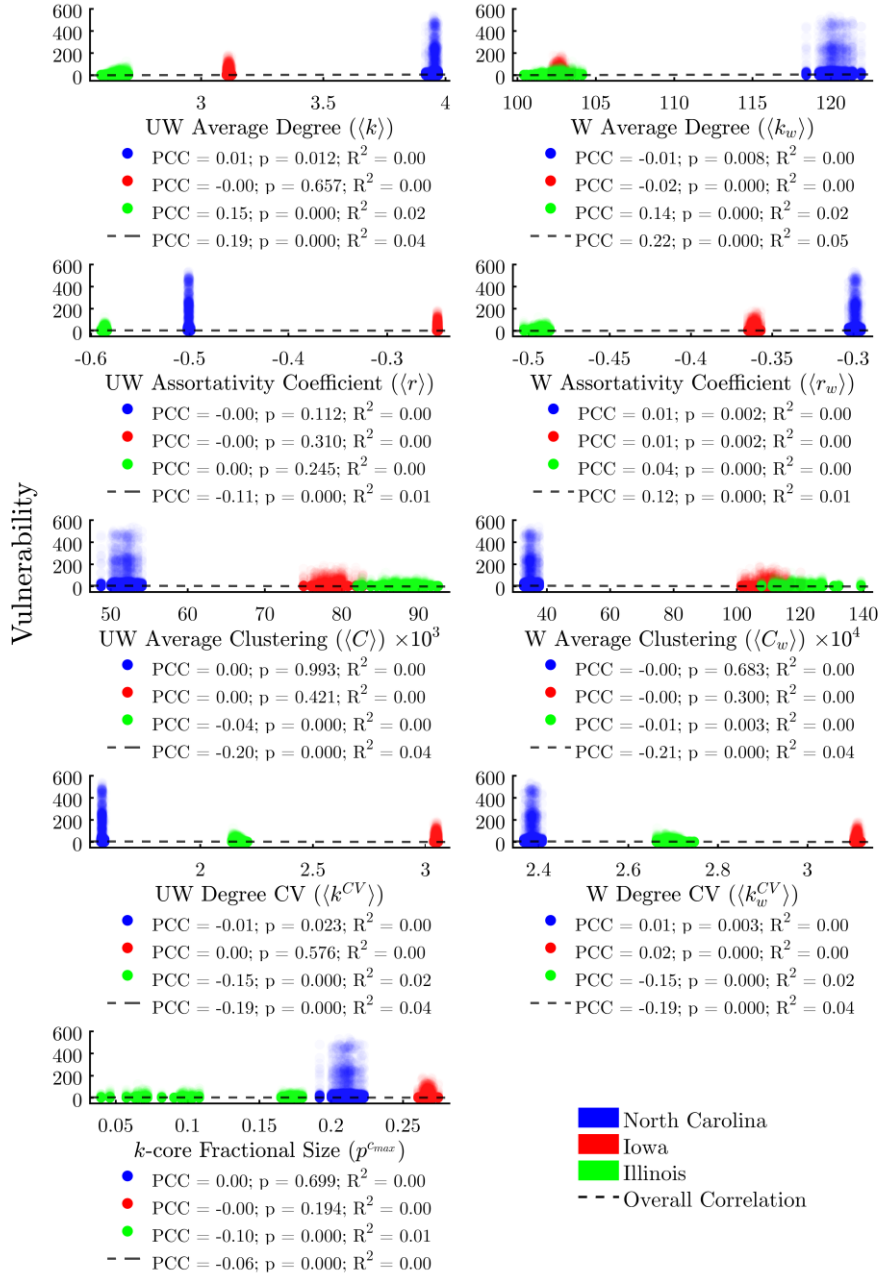
Figure 18: Correlations between key network-level metrics and node-level vulnerability. Weighted (W) and unweighted (UW) statistics are indicated. Scatter points represent each agent in each run, color coded by study area. Black lines show best fit, with Pearson correlation statistics and R^2 given in legends both by state and overall.

156

4.6.2.2 Node positionality factors

        For real-world livestock producers, production system managers, veterinarians, epidemiologists, and policymakers it would be useful to understand the relationships between the network positionality of a given actor and its risk of infection during an epidemic event. To investigate this, we plot each *node*-level metric from Table 8 against each agent's vulnerability. Figure 19 shows the various indicators of node centrality, while Figure 20 shows the remaining node positionality indicators.
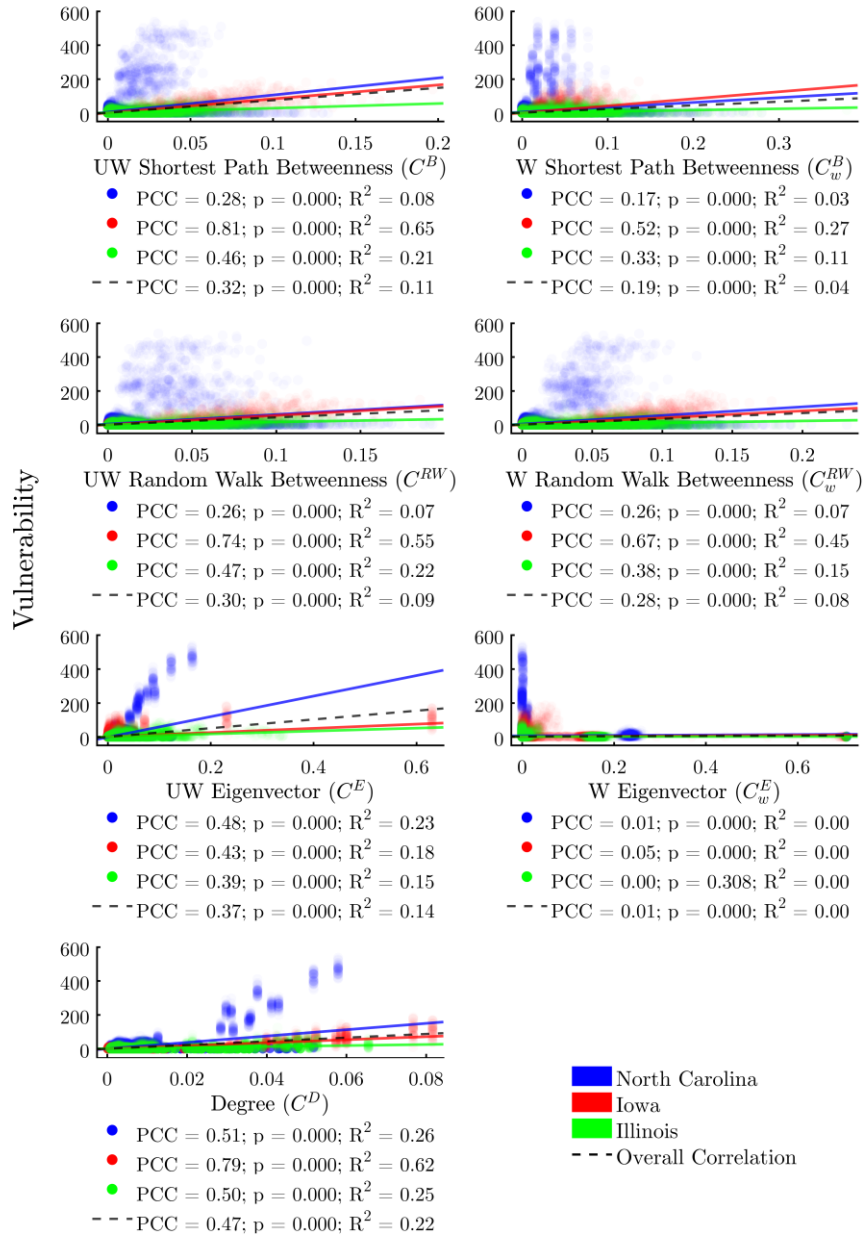
Figure 19: Correlations between centrality metrics and vulnerability. Weighted (W) and unweighted (UW) statistics are indicated. Scatter points represent each agent in each run, color coded by study area. Lines show best fit for each state as well as overall, with Pearson correlation statistics and R^2 given in legends both by state and overall.
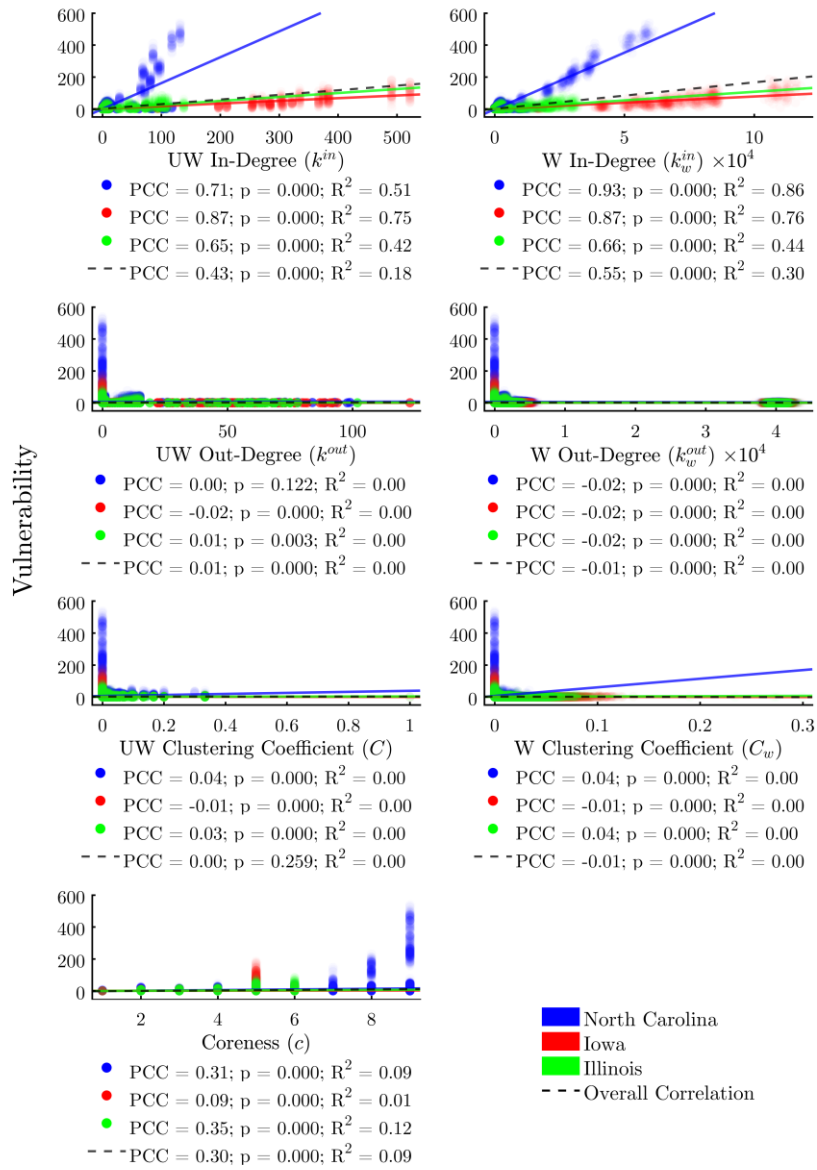
Figure 20: Correlations between key node-level metrics and vulnerability. Weighted (W) and unweighted (UW) statistics are indicated. Scatter points represent each agent in each run, color coded by study area. Lines show best fit for each state as well as overall, with Pearson correlation statistics and R^2 given in legends both by state and overall.

We find that all *node* centrality metrics significantly correlate with infection risk when using the full dataset composed of all three study areas (Figure 19). With the exception of weighted eigenvector, these trends hold within each study area as well. The centrality metric with the most consistent positive correlation to vulnerability across study areas would appear to be degree centrality.

However, linear regressions reveal that, despite achieving significant *p*-values, no centrality indicator explains more than 22% of the variability in the data across all study areas. Additionally, we find significant variability in the explanatory power of the correlations between study areas. This lends further evidence to the assertion that there is a complex interplay between the structure of a network overall and the efficacy of *node*-level metrics for determining risk.

Turning to the remaining *node* positionality indicators (Figure 20), we find—unsurprisingly—that both weighted and unweighted in-degree are relatively-strong predictors of vulnerability, with the weighted metric explaining 30% of the variability in the data. Simply stated, the more times a node is exposed to a potential threat by way of incoming livestock or feed, the more likely it is to be infected. We find that out-degree and clustering coefficient (both weighted and unweighted) do not explain a node's vulnerability. A node's coreness is a moderate (R^2=0.09) predictor of vulnerability across study areas, suggesting that subsets of more highly-interconnected nodes may be largely responsible for sustaining outbreaks.

160

4.6.2.3 Using Genetic Programming to develop better vulnerability indicators

      Our preliminary findings suggest that, while both *global* and *node*-level metrics correlate with epidemiological vulnerability in some contexts, there are likely complexities arising from the interplay between these factors, and no single indicator has a sufficiently high R^2 value to be of much use. The development of indicators that consistently identify risk—ideally accounting for both overall network structure and node positionality within a network—would be a boon for epidemiological analysts seeking to target biosecurity interventions in livestock production systems.

      To identify such meta-metrics, we employ a GP algorithm to search the solution space for mathematical combinations of individual indicators along a complexity / fitness Pareto front (Schmidt & Lipson, 2009). Of the 14 possible solutions generated by the GP, six were chosen for further analysis (Table 11). This subset of solutions aims to sample a range of complexity / fitness pairings, with solutions situated at "knees" of the Pareto front preferred where possible (Branke et al., 2004). To avoid overfitting the data, solutions beyond a complexity level of 14 are not considered, as there is very little improvement in fitness beyond that point.

Table 11: Six possible GP solutions along the complexity / fitness Pareto front. Fitness is defined as the R^2 value on the validation data. Red variables are from the *global* set; blue are from the *node* set—yet require a full network graph to calculate—and green are from the *node* set and require only *node*-level data. Constants are truncated to two significant figures. Equations are algebraically-simplified where applicable.

	Solution	Complexity	R^2	Max. Error
S1	$v = 0.0035 \cdot k_w^{in} \cdot c$	5	0.5496	240.8
S2	$v = 0.011 \cdot k_w^{in} \cdot c - k^{in}$	7	0.9082	102.4
S3	$v = 2.76 \times 10^{-6} \cdot k_w^{in} \cdot c / \langle C_w \rangle$	8	0.9331	120.9
S4	$v = 3.46 \times 10^{-7} \cdot k_w^{in} \cdot c^2 / \langle C_w \rangle$	10	0.9515	91.68
S5	$v = k_w^{in} \cdot \left(0.0010 \cdot c^2 - 14.3 \cdot \langle C_w \rangle \right)$	13	0.9531	83.64
S6	$v = k^{in} \cdot C_w^E + 3.42 \times 10^{-7} \cdot k_w^{in} \cdot c^2 / \langle C_w \rangle$	14	0.9562	75.79

Overall, it is notable that the better-performing of the GP solutions incorporate both *node*-level and *global* metrics. The most frequently-used *node*-level metrics are first and foremost weighted in-degree—which is used in all six possible solutions—followed by coreness, unweighted in-degree, unweighted total degree, and weighted eigenvector centrality. The most common (and indeed, only) *global* metric included in the GP formuae is weighted clustering.

Figure 21 plots the results of the GP solutions against the vulnerability of each node in the same manner as Figure 18, Figure 19, and Figure 20. We find that all of the GP solutions far outperform any individual metric, providing support for the value of the GP-derived meta-metric approach. The best of the GP solutions (S4 and S6) explain

162

91% of the variability in the response variable, while the best single metric (k_w^{in})

explains just 30%.
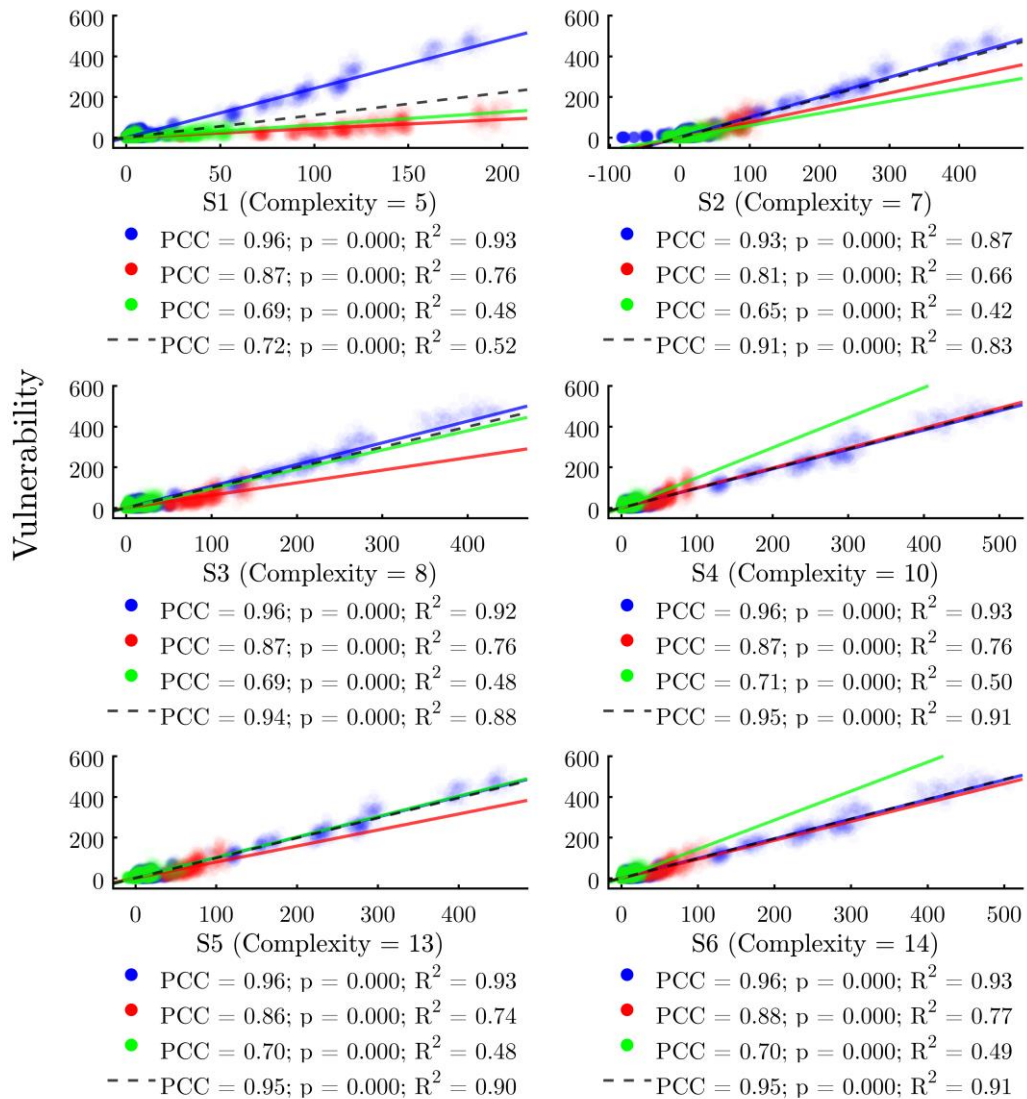
Figure 21: Correlations between six GP solutions and vulnerability. Scatter points represent each agent in each run, color coded by study area (blue = NC; red = IA; green = IL). Lines show best fit for each study area as well as overall, with Pearson correlation statistics and R^2 given in legends. Note that, unlike the R^2 values given in Table 11—which are calculated only on the GP validation set—the R^2 values here are calculated on the full dataset output from the model.

4.7 Discussion and Conclusions

In this study, we apply agent-based computer modeling, network analytics, and evolutionary computation to explore the impact of key network metrics on the epidemiological vulnerability of livestock production chain actors. We analyze the extent to which indicators describing the structure of a complex contact network can be combined with those pertaining to a node's positionality within that network to arrive at better risk indicators. To this end, we demonstrate the feasibility of using a GP algorithm to formulate such meta-metrics and, at least in this context, we find that the GP solutions outperform any single indicator of epidemiological vulnerability. Below, we discuss how the meta-metric method could be used by real-world practitioners, address methodological issues surrounding data availability and context specificity, and propose future research goals.

Previous research in this area has largely explored bivariate correlations between of a suite of metrics and epidemiological outcomes—at either the network or individual level—in both simulated and real-world networked systems (Rothenberg et al., 1995; Bell et al., 1999; De et al., 2004; Christley et al., 2005; Salathé & Jones, 2010; Kitsak et al., 2010). Multivariate statistical techniques have also been used to evaluate the relative effect of multiple *node*-level metrics on individual vulnerability (Ghani & Garnett, 2000). The GP methodology we have employed expands on this work, searching the solution space to identify mathematical relationships encompassing multiple metrics to predict a node's vulnerability across three networks with differing sizes, densities, and internal structures.

165

The ability of our GP approach to identify meta-metrics that capture complexities arising from the interactions between *global* and *node*-level network features suggests that this methodology may have widespread applicability. The particular networks evaluated in this experiment were state-level livestock production systems generated by a computer model, but we believe the procedures used to analyze the network data and identify meta-metrics could be applied to study phenomena on a wide range of graphs, both empirically- and computationally-derived. While we have applied the meta-metric approach to study epidemiological dynamics, meta-metrics could equally be formulated to predict other important graph properties, such as the probability that a node will add or prune edges over time, the frequency with which a node will enter some state, or virtually any other outcome variable of interest.

4.7.1 Global vs. Node Metrics as Indicators of Vulnerability

We find that the *global* metrics, while serving as reasonably-accurate predictors of average vulnerability within each study area, do little to predict the risk of an individual agent within each run, as a result of the high variability in vulnerability across agents of different classes, and with different network positionality characteristics (Figure 18).

Several of the *node*-level metrics, such as degree centrality and weighted in-degree, perform reasonably well as predictors of an agent's vulnerability (R^2=0.22 and 0.30 respectively). However, many of the *node* metrics demonstrate differential

correlation coefficients across each study area, indicating that risk cannot be accurately predicted by *node*-level indicators alone (Figure 19 and Figure 20).

4.7.2 Evaluating GP-Derived Meta-Metrics

Turning to the GP solutions, we find that the best-performing meta-metrics (S3–S6) incorporate both *global* and *node*-level indicators. Further, among the *node* metrics, we note that some of the calculations—for example weighted in-degree—require only local information from each node for calculation. However, other *node* metrics—for example coreness—require that a full graph of the network is first codified. Notably, all six of the GP solutions we analyzed do require at least one instance of a *node* metric that falls into this latter category (primarily coreness in this case), emphasizing the interplay between *global* network structures and *node* positionality in the prediction of infection risk.

Weighing the six meta-metric solutions against one another, S4 and S6 are the best performing of the set, each achieving a high overall R^2 value of 0.91 on the full dataset. Table 11 reveals that S6, while more complex, cuts the maximum error from 91.68 for S4 to 75.79 for S6, suggesting that the additional complexity pays off in terms of predictive accuracy in this context. Overall, we conclude that the GP approach to determining effective predictors of epidemiological risk by combining indicators of network structure with node positionality represents an effective and under-explored mechanism to evaluate the infection vulnerability of individual nodes within networks.

4.7.3 Application to Real-World Decision-Making

Aside from demonstrating the applicability of our methodology, this work is also intended to inform practitioners in their efforts to strategically-allocate resources which improve epidemiological resilience in existing livestock production systems. While we do not wish to imply that the results presented here can reliably predict the course of any real-world disease incursion, our findings suggest several novel ways for livestock epidemiologists to approach risk analysis.

Perhaps the most important takeaway from our study is that risk is a function of both network structure and individual decision-making. Biosecurity measures—such as shower-in-shower-out facilities, lines of separation, or all-in-all-out protocols—have proven effective in limiting the scope of livestock disease outbreaks. While a producer's primary motivation to adopt such measures is rooted in individual risk reduction, from a systems perspective, our research suggests that biosecurity protocols would be vastly more effective if implemented at especially-vulnerable network loci. The methods presented in this study allow for the identification of nodes in a system most likely to play a role in disease spread, which can then inform targeted biosecurity interventions to reduce systemic risk.

The lack of available fine-grained U.S. animal movement data currently represents a major hurdle for network analysts and modelers interested in livestock disease risk mitigation. Animal traceback protocols, which track the movement of livestock between premises over their full lifespans, remain controversial in the U.S.,

likely due to attitudes around privacy. While consumer protection (recalls, etc.) is the most-cited reason for the implementation of traceback protocols, in light of our findings, an additional consideration is that traceback data could be used to codify accurate network representations of livestock supply chains, as has been done in other countries (Caporale et al., 2001; Natale et al., 2009).

Given sufficient input data, RUSH-PNBM could be parameterized to guide on-the-ground policy and managerial decisions. By encoding the interaction patterns and animal movements between actors in real-world production systems into network graphs, and then incorporating data tracking the spread of a historical outbreak through the system, it becomes possible to apply the methods we have presented to actual disease incursion threat scenarios.

Model use cases could focus on different scales. For example, national agencies such as the USDA could use RUSH-PNBM to target nationwide interventions—such as the placement of truck-wash sites—or to propose health and safety regulations focused on potential "hubs" like slaughter plants and feed mills. Alternatively, owners and managers of private livestock production chain systems could employ specifically-parameterized versions of the model to assess disease risks within their own systems. This would facilitate both insights into the individual network loci at which biosecurity protocols could do the most good, as well as analyses of how reorganization of network structures could improve system-wide disease resilience.

4.7.4 Expert Panelist Feedback

The variation of individual node vulnerabilities found in our results did not surprise our panelists: the relationship between an individual node's vulnerability and its position and/or function in the network matters. The potential of our model to identify those nodes with the highest disease risk casts light on an important informational need raised in our panel sessions. Coupled with finer-grained and more-detailed data (such as the widespread use of animal traceback protocols), panelists agreed that RUSH-PNBM—together with the meta-metric methodology—could represent a valuable tool which may be leveraged to identify where best to target biosecurity measures.

Interestingly, many of the expert panelists were less interested in some of the broader structural implications that may be drawn from this study. For example, despite its demonstrated biosecurity implications, minimizing the bridging links within livestock production networks by reducing producer over-specialization was not universally deemed desirable. However, more recent engagement between our research team and owners of mid-sized, leading-edge production chain systems has revealed an openness to rethinking the role of network connectivity structures upon biosecurity within the context of their own operational networks.

4.7.5 Limitations of Our Approach

Macro-scale socio-economic and political factors can directly and indirectly influence the economic incentives, disease patterns, and network-structural parameters

embedded in RUSH-PNBM. While this study demonstrates the limited generalizability of our model across three states in the U.S., application of the model to other socio-economic and political contexts (e.g. the EU) will require careful re-calibration and validation of model parameters, including securing relevant datasets and convening additional expert panels. Extensibility of RUSH-PNBM to generate medium- to long-term (i.e. 30- to 50-year) scenarios would also require refinement of simulation processes and parameters to account for projected changes in economic and political conditions within which the simulated networks operate.

As noted above, practical implementation of this new meta-metric approach (i.e. to guide policymaking) will require better contextualization of the network structures and underlying theoretical assumptions governing network dynamics. Due to complex government-industry relationships and competition within the livestock production industry, empirical data concerning underlying network structures, animal movements, epidemic spread data, and other important factors are not typically easily available. Fortunately, ABMs—acting as virtual laboratories (Macal & North, 2010)—enable testing of alternate assumptions about network structures and their underlying dynamics. As described above, widespread adoption of animal traceback protocols would go a long way toward solving the data availability issue.

4.7.6 Directions for Future Research

Previous studies investigating the impact of network positionality on epidemiological risk differentiate between vulnerability—which we have used as our

dependent variable—and infectivity, i.e. the frequency with which an actor spreads the disease to others (Rothenberg et al., 1995; Ghani & Garnett, 2000). Understanding both factors is important to guide the implementation of interventions which curb disease spread, as some protocols are designed to prevent incoming disease threats, whereas others are aimed at limiting outgoing infection spread. An area for future research is to develop GP-derived meta-metrics that correlate with a node's infectivity rather than its vulnerability. For example, it would be interesting to determine whether the finding of Ghani & Garnett (2000)—that infectivity is more contingent on global network structures—is corroborated. This would offer a more complete picture of the overall risk environment.

In addition, future versions of the RUSH-PNBM model itself will be increasingly honed. For example, our team will conduct survey research to evaluate the economic and human dimensions underlying choices to adopt methods and technologies aimed at bolstering producer-level biosecurity. Another future research goal is to use experimental gaming data to determine the circumstances under which individuals choose to adopt biosecurity protocols in risky and/or uncertain decision environments (Merrill et al., 2018). We can then program the agents in RUSH-PNBM to adapt their behavior under parallel conditions in the model. Coupling such insights with network-based risk analyses—especially those calibrated around real-world production chain systems—will allow for a deeper understanding of the human-behavioral solutions that may promote the adoption of risk-mitigating protocols, as well

172

as helping to focus biosecurity protocol adoption upon the nodes at which it will have the greatest effect on system-wide disease resilience.

4.8 Acknowledgements

4.9 References

ALBERT, R. & BARABÁSI, A.-L. (2002). STATISTICAL MECHANICS OF COMPLEX NETWORKS. REVIEWS OF MODERN PHYSICS, 74(1), 47

ALDERTON, S., MACLEOD, E. T., ANDERSON, N. E., SCHATEN, K., KULESZO, J., SIMUUNZA, M., WELBURN, S. C. & ATKINSON, P. M. (2016). A MULTI-HOST AGENT-BASED MODEL FOR A ZOONOTIC, VECTOR-BORNE DISEASE: A CASE STUDY ON TRYPANOSOMIASIS IN EASTERN PROVINCE, ZAMBIA. PLOS NEGLECTED TROPICAL DISEASES, 10(12), E0005252

AN, L. (2012). MODELING HUMAN DECISIONS IN COUPLED HUMAN AND NATURAL SYSTEMS: REVIEW OF AGENT-BASED MODELS. ECOLOGICAL MODELLING, 229, 25–36

ANDERSON, R. M. & MAY, R. M. (1979). POPULATION BIOLOGY OF INFECTIOUS DISEASES: PART I. NATURE, 280(5721), 361

ARRUDA, A. G., FRIENDSHIP, R., CARPENTER, J., GREER, A. & POLJAK, Z. (2016). EVALUATION OF CONTROL STRATEGIES FOR PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME (PRRS) IN SWINE

BREEDING HERDS USING A DISCRETE EVENT AGENT-BASED MODEL. PLOS ONE, 11(11), E0166596

AUCHINCLOSS, A. H. & DIEZ ROUX, A. V. (2008). A NEW TOOL FOR EPIDEMIOLOGY: THE USEFULNESS OF DYNAMIC-AGENT MODELS IN UNDERSTANDING PLACE EFFECTS ON HEALTH. AMERICAN JOURNAL OF EPIDEMIOLOGY, 168(1), 1–8

BAGNI, R., BERCHI, R. & CARIELLO, P. (2002). A COMPARISON OF SIMULATION MODELS APPLIED TO EPIDEMICS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 5(3)

BARABÁSI, A.-L. & ALBERT, R. (1999). EMERGENCE OF SCALING IN RANDOM NETWORKS. SCIENCE, 286(5439), 509–512

BARRETEAU, O., ANTONA, M., D'AQUINO, P., AUBERT, S., BOISSAU, S., BOUSQUET, F., DARÉ, W., ETIENNE, M., LE PAGE, C., MATHEVET, R. ET AL. (2003). OUR COMPANION MODELLING APPROACH. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 6(1)

BARRETT, C. L., BISSET, K. R., EUBANK, S. G., FENG, X. & MARATHE, M. V. (2008). EPISIMDEMICS: AN EFFICIENT ALGORITHM FOR SIMULATING THE SPREAD OF INFECTIOUS DISEASE OVER LARGE REALISTIC SOCIAL NETWORKS. IN HIGH PERFORMANCE COMPUTING, NETWORKING, STORAGE AND ANALYSIS, (PP. 1–12). IEEE

BATAGELJ, V. & ZAVERSNIK, M. (2003). AN O (M) ALGORITHM FOR CORES DECOMPOSITION OF NETWORKS. ARXIV PREPRINT CS/0310049

BELKHIRIA, J., ALKHAMIS, M. A. & MARTÍNEZ-LÓPEZ, B. (2016). APPLICATION OF SPECIES DISTRIBUTION MODELING FOR AVIAN INFLUENZA SURVEILLANCE IN THE UNITED STATES CONSIDERING THE NORTH AMERICA MIGRATORY FLYWAYS. SCIENTIFIC REPORTS, 6, 33161

BELL, D. C., ATKINSON, J. S. & CARLSON, J. W. (1999). CENTRALITY MEASURES FOR DISEASE TRANSMISSION NETWORKS. SOCIAL NETWORKS, 21(1), 1–21

BIGRAS-POULIN, M., BARFOD, K., MORTENSEN, S. & GREINER, M. (2007). RELATIONSHIP OF TRADE PATTERNS OF THE DANISH SWINE INDUSTRY ANIMAL MOVEMENTS NETWORK TO POTENTIAL DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE, 80(2), 143–165

BIGRAS-POULIN, M., THOMPSON, R., CHRIÉL, M., MORTENSEN, S. & GREINER, M. (2006). NETWORK ANALYSIS OF DANISH CATTLE INDUSTRY TRADE PATTERNS AS AN EVALUATION OF RISK POTENTIAL FOR DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE, 76(1), 11–39

BONACICH, P. (1987). POWER AND CENTRALITY: A FAMILY OF MEASURES. AMERICAN JOURNAL OF SOCIOLOGY, 92(5), 1170–1182

BOUSQUET, F., BARRETEAU, O., LE PAGE, C., MULLON, C. & WEBER, J. (1999). AN ENVIRONMENTAL MODELLING APPROACH: THE USE OF MULTI-AGENT SIMULATIONS. ADVANCES IN ENVIRONMENTAL AND ECOLOGICAL MODELLING, 113, 122

BRADHURST, R., ROCHE, S. E., EAST, I., KWAN, P. & GARNER, M. (2016). IMPROVING THE COMPUTATIONAL EFFICIENCY OF AN AGENT-BASED SPATIOTEMPORAL MODEL OF LIVESTOCK DISEASE SPREAD AND CONTROL. ENVIRONMENTAL MODELLING & SOFTWARE, 77, 1–12

BRANDES, U. & FLEISCHER, D. (2005). CENTRALITY MEASURES BASED ON CURRENT FLOW. IN ANNUAL SYMPOSIUM ON THEORETICAL ASPECTS OF COMPUTER SCIENCE, (PP. 533–544). SPRINGER

BRANKE, J., DEB, K., DIEROLF, H. & OSSWALD, M. (2004). FINDING KNEES IN MULTI-OBJECTIVE OPTIMIZATION. IN INTERNATIONAL CONFERENCE ON PARALLEL PROBLEM SOLVING FROM NATURE, (PP. 722–731). SPRINGER

BURDETT, C. L., KRAUS, B. R., GARZA, S. J., MILLER, R. S. & BJORK, K. E. (2015). SIMULATING THE DISTRIBUTION OF INDIVIDUAL LIVESTOCK FARMS AND THEIR POPULATIONS IN THE UNITED STATES: AN EXAMPLE USING DOMESTIC SWINE (SUS SCROFA DOMESTICUS) FARMS. PLOS ONE, 10(11), E0140338

CAPORALE, V., GIOVANNINI, A., DI FRANCESCO, C. & CALISTRI, P. (2001). IMPORTANCE OF THE TRACEABILITY OF ANIMALS AND ANIMAL PRODUCTS IN EPIDEMIOLOGY. REVUE SCIENTIFIQUE ET TECHNIQUE–OIE, 20(2), 372–378

CHRISTLEY, R. M., PINCHBECK, G., BOWERS, R., CLANCY, D., FRENCH, N., BENNETT, R. & TURNER, J. (2005). INFECTION IN SOCIAL NETWORKS:

USING NETWORK ANALYSIS TO IDENTIFY HIGH-RISK INDIVIDUALS. AMERICAN JOURNAL OF EPIDEMIOLOGY, 162(10), 1024–1031

COLIZZA, V., BARRAT, A., BARTHÉLEMY, M. & VESPIGNANI, A. (2007). PREDICTABILITY AND EPIDEMIC PATHWAYS IN GLOBAL OUTBREAKS OF INFECTIOUS DISEASES: THE SARS CASE STUDY. BMC MEDICINE, 5(1), 34

DE, P., SINGH, A. E., WONG, T., YACOUB, W. & JOLLY, A. (2004). SEXUAL NETWORK ANALYSIS OF A GONORRHOEA OUTBREAK. SEXUALLY TRANSMITTED INFECTIONS, 80(4), 280–285

DUBÉ, C., RIBBLE, C., KELTON, D. & MCNAB, B. (2009). A REVIEW OF NETWORK ANALYSIS TERMINOLOGY AND ITS APPLICATION TO FOOT-AND-MOUTH DISEASE MODELLING AND POLICY DEVELOPMENT. TRANSBOUNDARY AND EMERGING DISEASES, 56(3), 73–85

EL-SAYED, A. M., SCARBOROUGH, P., SEEMANN, L. & GALEA, S. (2012). SOCIAL NETWORK ANALYSIS AND AGENT-BASED MODELING IN SOCIAL EPIDEMIOLOGY. EPIDEMIOLOGIC PERSPECTIVES & INNOVATIONS, 9(1), 1

EPSTEIN, J. M. (2008). WHY MODEL? JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 11(4), 12

FÈVRE, E. M., BRONSVOORT, B. M. D. C., HAMILTON, K. A. & CLEAVELAND, S. (2006). ANIMAL MOVEMENTS AND THE SPREAD OF INFECTIOUS DISEASES. TRENDS IN MICROBIOLOGY, 14(3), 125–131

FREEMAN, L. C. (1977). A SET OF MEASURES OF CENTRALITY BASED ON BETWEENNESS. SOCIOMETRY, (PP. 35–41)

GARNER, M. & HAMILTON, S. (2011). PRINCIPLES OF EPIDEMIOLOGICAL MODELLING. REVUE SCIENTIFIQUE ET TECHNIQUE–OIE, 30(2), 407

GARNER, M. G., DUBÉ, C., STEVENSON, M. A., SANSON, R. L., ESTRADA, C. & GRIFFIN, J. (2007). EVALUATING ALTERNATIVE APPROACHES TO MANAGING ANIMAL DISEASE OUTBREAKS–THE ROLE OF MODELLING IN POLICY FORMULATION. VET. ITAL, 43(2), 285–298

GHANI, A. C. & GARNETT, G. P. (2000). RISKS OF ACQUIRING AND TRANSMITTING SEXUALLY TRANSMITTED DISEASES IN SEXUAL PARTNER NETWORKS. SEXUALLY TRANSMITTED DISEASES, 27(10), 579–587

GOEDE, D. & MORRISON, R. B. (2016). PRODUCTION IMPACT & TIME TO STABILITY IN SOW HERDS INFECTED WITH PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV). PREVENTIVE VETERINARY MEDICINE, 123, 202 – 207. HTTPS://DOI.ORG/10.1016/J.PREVETMED.2015.11.010
URL
HTTP://DX.DOI.ORG/HTTPS://DOI.ORG/10.1016/J.PREVETMED.2015.11.010

HARVEY, N., REEVES, A., SCHOENBAUM, M. A., ZAGMUTT-VERGARA, F. J., DUBÉ, C., HILL, A. E., CORSO, B. A., MCNAB, W. B., CARTWRIGHT, C. I. & SALMAN, M. D. (2007). THE NORTH AMERICAN ANIMAL DISEASE SPREAD MODEL: A SIMULATION MODEL TO ASSIST DECISION MAKING IN EVALUATING ANIMAL DISEASE INCURSIONS. PREVENTIVE VETERINARY MEDICINE, 82(3-4), 176–197

HUNTER, E., MAC NAMEE, B. & KELLEHER, J. D. (2017). A TAXONOMY FOR AGENT-BASED MODELS IN HUMAN INFECTIOUS DISEASE EPIDEMIOLOGY. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 20(3), 2. 10.18564/JASSS.3414
URL HTTP://DX.DOI.ORG/10.18564/JASSS.3414

KAO, R., DANON, L., GREEN, D. & KISS, I. (2006). DEMOGRAPHIC STRUCTURE AND PATHOGEN DYNAMICS ON THE NETWORK OF LIVESTOCK MOVEMENTS IN GREAT BRITAIN. PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON B: BIOLOGICAL SCIENCES, 273(1597), 1999–2007

KAUL, H. & VENTIKOS, Y. (2013). INVESTIGATING BIOCOMPLEXITY THROUGH THE AGENT-BASED PARADIGM. BRIEFINGS IN BIOINFORMATICS, 16(1), 137–152

KILLWORTH, P. D., MCCARTY, C., BERNARD, H. R., SHELLEY, G. A. & JOHNSEN, E. C. (1998). ESTIMATION OF SEROPREVALENCE, RAPE, AND HOMELESSNESS IN THE UNITED STATES USING A SOCIAL NETWORK APPROACH. EVALUATION REVIEW, 22(2), 289–308

KISS, I. Z., GREEN, D. M. & KAO, R. R. (2006). THE NETWORK OF SHEEP MOVEMENTS WITHIN GREAT BRITAIN: NETWORK PROPERTIES AND THEIR IMPLICATIONS FOR INFECTIOUS DISEASE SPREAD. JOURNAL OF THE ROYAL SOCIETY INTERFACE, 3(10), 669–677

KITCHING, R. P., THRUSFIELD, M. V. & TAYLOR, N. M. (2006). USE AND ABUSE OF MATHEMATICAL MODELS: AN ILLUSTRATION FROM THE 2001 FOOT AND MOUTH DISEASE EPIDEMIC IN THE UNITED KINGDOM. REVUE SCIENTIFIQUE ET TECHNIQUE–OIE, 25(1), 293

KITSAK, M., GALLOS, L. K., HAVLIN, S., LILJEROS, F., MUCHNIK, L., STANLEY, H. E. & MAKSE, H. A. (2010). IDENTIFICATION OF INFLUENTIAL SPREADERS IN COMPLEX NETWORKS. NATURE PHYSICS, 6(11), 888–893

KLÜGL, F. (2008). A VALIDATION METHODOLOGY FOR AGENT-BASED SIMULATIONS. IN PROCEEDINGS OF THE 2008 ACM SYMPOSIUM ON APPLIED COMPUTING, (PP. 39–43). ACM

LUCE, R. D. & PERRY, A. D. (1949). A METHOD OF MATRIX ANALYSIS OF GROUP STRUCTURE. PSYCHOMETRIKA, 14(2), 95–116

MACAL, C. & NORTH, M. (2010). TUTORIAL ON AGENT-BASED MODELLING AND SIMULATION. JOURNAL OF SIMULATION, 4(3), 151–162

MANSLEY, L., DUNLOP, P., WHITESIDE, S. & SMITH, R. (2003). EARLY DISSEMINATION OF FOOT-AND-MOUTH. THE VETERINARY RECORD, 153, 43–50

MERRILL, S. C., KOLIBA, C. J., ZIA, A., PARKER, J., MOEGENBURG, S., SELLNOW, T., WILTSHIRE, S. W., BUCINI, G., DANEHY, C. & SMITH, J. M. (2018). ADOPTION OF LIVESTOCK BIOSECURITY PRACTICES AMIDST ENVIRONMENTAL AND SOCIAL UNCERTAINTY: EVIDENCE FROM A RANDOMIZED EXPERIMENTAL GAME. PLOS ONE. UNDER REVIEW

MOSS, S. (2008). ALTERNATIVE APPROACHES TO THE EMPIRICAL VALIDATION OF AGENT-BASED MODELS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 11(1), 5 URL HTTP://JASSS.SOC.SURREY.AC.UK/11/1/5.HTML

NATALE, F., GIOVANNINI, A., SAVINI, L., PALMA, D., POSSENTI, L., FIORE, G. & CALISTRI, P. (2009). NETWORK ANALYSIS OF ITALIAN CATTLE TRADE PATTERNS AND EVALUATION OF RISKS FOR POTENTIAL DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE, 92(4), 341–350

NEWMAN, M. E. (2003). MIXING PATTERNS IN NETWORKS. PHYSICAL REVIEW E, 67(2), 026126

NEWMAN, M. E. (2005). A MEASURE OF BETWEENNESS CENTRALITY BASED ON RANDOM WALKS. SOCIAL NETWORKS, 27(1), 39–54

ORTIZ-PELAEZ, A., PFEIFFER, D., SOARES-MAGALHAES, R. & GUITIAN, F. (2006). USE OF SOCIAL NETWORK ANALYSIS TO CHARACTERIZE THE PATTERN OF ANIMAL MOVEMENTS IN THE INITIAL PHASES OF THE 2001 FOOT AND MOUTH DISEASE (FMD) EPIDEMIC IN THE UK. PREVENTIVE VETERINARY MEDICINE, 76(1-2), 40–55

PARKER, J. & EPSTEIN, J. M. (2011). A DISTRIBUTED PLATFORM FOR GLOBAL-SCALE AGENT-BASED MODELS OF DISEASE TRANSMISSION. ACM TRANSACTIONS ON MODELING AND COMPUTER SIMULATION (TOMACS), 22(1), 2

PEREZ, A. M., WARD, M. P., CHARMANDARIÁN, A. & RITACCO, V. (2002). SIMULATION MODEL OF WITHIN-HERD TRANSMISSION OF BOVINE TUBERCULOSIS IN ARGENTINE DAIRY HERDS. PREVENTIVE VETERINARY MEDICINE, 54(4), 361–372

PEREZ, L. & DRAGICEVIC, S. (2009). AN AGENT-BASED APPROACH FOR MODELING DYNAMICS OF CONTAGIOUS DISEASE SPREAD. INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS, 8(1), 50

ROBINS, J., BOGEN, S., FRANCIS, A., WESTHOEK, A., KANAREK, A., LENHART, S. & EDA, S. (2015). AGENT-BASED MODEL FOR JOHNE'S DISEASE DYNAMICS IN A DAIRY HERD. VETERINARY RESEARCH, 46(1), 68

ROTHENBERG, R. B., POTTERAT, J. J., WOODHOUSE, D. E., DARROW, W. W., MUTH, S. Q. & KLOVDAHL, A. S. (1995). CHOOSING A CENTRALITY MEASURE: EPIDEMIOLOGIC CORRELATES IN THE COLORADO SPRINGS STUDY OF SOCIAL NETWORKS. SOCIAL NETWORKS, 17(3-4), 273–297

SALATHÉ, M. & JONES, J. H. (2010). DYNAMICS AND CONTROL OF DISEASES IN NETWORKS WITH COMMUNITY STRUCTURE. PLOS COMPUTATIONAL BIOLOGY, 6(4), E1000736

SARAMÄKI, J., KIVELÄ, M., ONNELA, J.-P., KASKI, K. & KERTESZ, J. (2007). GENERALIZATIONS OF THE CLUSTERING COEFFICIENT TO WEIGHTED COMPLEX NETWORKS. PHYSICAL REVIEW E, 75(2), 027105

SCHMIDT, M. & LIPSON, H. (2009). DISTILLING FREE-FORM NATURAL LAWS FROM EXPERIMENTAL DATA. SCIENCE, 324(5923), 81–85

SCHOENBAUM, M. A. & DISNEY, W. T. (2003). MODELING ALTERNATIVE MITIGATION STRATEGIES FOR A HYPOTHETICAL OUTBREAK OF FOOT-AND-MOUTH DISEASE IN THE UNITED STATES. PREVENTIVE VETERINARY MEDICINE, 58(1-2), 25–52

SCHULZ, L. L. & TONSOR, G. T. (2015). ASSESSMENT OF THE ECONOMIC IMPACTS OF PORCINE EPIDEMIC DIARRHEA VIRUS IN THE UNITED STATES. JOURNAL OF ANIMAL SCIENCE, 93(11), 5111–5118

SHI, Z. Z., WU, C.-H. & BEN-ARIEH, D. (2014). AGENT-BASED MODEL: A SURGING TOOL TO SIMULATE INFECTIOUS DISEASES IN THE IMMUNE SYSTEM. OPEN JOURNAL OF MODELLING AND SIMULATION, 2(01), 12

STEVENSON, M., SANSON, R., STERN, M., O'LEARY, B., SUJAU, M., MOLES-BENFELL, N. & MORRIS, R. (2013). INTERSPREAD PLUS: A SPATIAL AND STOCHASTIC SIMULATION MODEL OF DISEASE IN ANIMAL POPULATIONS. PREVENTIVE VETERINARY MEDICINE, 109(1-2), 10–24

THE PIG SITE (2014). PIG FARM TARGETS: STOCKMANSHIP STANDARDS. ARCHIVED BY WEBCITE AT HTTP://WWW.WEBCITATION.ORG/6YPMDXS0E URL HTTP://WWW.THEPIGSITE.COM/STOCKSTDS/3/PIG-FARM-TARGETS/

TRACEY, J. A., BEVINS, S. N., VANDEWOUDE, S. & CROOKS, K. R. (2014). AN AGENT-BASED MOVEMENT MODEL TO ASSESS THE IMPACT OF LANDSCAPE FRAGMENTATION ON DISEASE TRANSMISSION. ECOSPHERE, 5(9), 1–24

USDA NASS (2014). LIVESTOCK SLAUGHTER, 2013 SUMMARY: FEDERALLY INSPECTED PLANTS AND HEAD SLAUGHTERED BY SPECIES, STATES AND UNITED STATES, 2013. ARCHIVED BY WEBCITE AT HTTP://WWW.WEBCITATION.ORG/6YPM8S3HF URL HTTP://USDA.MANNLIB.CORNELL.EDU/USDA/NASS/LIVESLAUSU/2010S/2014/LIVESLAUSU-04-21-2014.PDF

VAROQUAUX, G., VAUGHT, T. & MILLMAN, J. (EDS.) (2008). EXPLORING NETWORK STRUCTURE, DYNAMICS, AND FUNCTION USING NETWORKX.

PASADENA, CA USA: PROCEEDINGS OF THE 7TH PYTHON IN SCIENCE CONFERENCE (SCIPY2008)

WASSERMAN, S. & FAUST, K. (1994). SOCIAL NETWORK ANALYSIS: METHODS AND APPLICATIONS, VOL. 8. CAMBRIDGE UNIVERSITY PRESS

WEBB, C. R. (2005). FARM ANIMAL NETWORKS: UNRAVELING THE CONTACT STRUCTURE OF THE BRITISH SHEEP POPULATION. PREVENTIVE VETERINARY MEDICINE, 68(1), 3–17

WILTSHIRE, S. W. (2018). USING AN AGENT-BASED MODEL TO EVALUATE THE EFFECT OF PRODUCER SPECIALIZATION ON THE EPIDEMIOLOGICAL RESILIENCE OF LIVESTOCK PRODUCTION NETWORKS. PLOS ONE, 13(3), E0194013

WINDRUM, P., FAGIOLO, G. & MONETA, A. (2007). EMPIRICAL VALIDATION OF AGENT-BASED MODELS: ALTERNATIVES AND PROSPECTS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 10(2), 8 URL HTTP://JASSS.SOC.SURREY.AC.UK/10/2/8.HTML

CHAPTER 3: AN EVOLUTIONARY-COMPUTATIONAL METHOD TO OPTIMIZE DYNAMIC, MULTI-MODAL INTERVENTION-DRIVEN CHANGE TOWARD GLOBAL FOOD SECURITY

5.1 Abstract

A major priority of the UN Sustainable Development Goals (SDGs) is to address the looming threat of increased hunger and food insecurity in the coming decades. While major progress has been made over that past half-century, achieving this goal going forward—especially in light of the multiple changing dynamics affecting food availability such as a growing population, dietary shifts, and climate change—has become increasingly challenging, requiring assessment of multiple potential options including both ways to increase food supply and minimize food demand. Food security is in part a component of the difference between food supply and demand over a given time interval. Unfortunately, there is not likely to be a single solution that ensures global food availability in the coming decades. However, it has been proposed that public investments into multiple interventions, or "wedges," (e.g. diet changes, climate change adaptation, and increasing yields) may work together to increase future food supply and reduce future demand. To optimize such public policy investments requires knowledge about how the change resulting from each intervention is likely to affect supply and demand, especially considering that investments in some interventions may have complex interaction effects. In addition to the ultimate change in food availability resulting from an investment, the time required for that change to

unfold, the expected baseline change (in the absence of any intervention), and the maximum possible level of change from a given wedge are all important parameters when planning an investment schedule. Here we examine multiple interventions to meet food availability needs by 2050 using a computational method to optimize public investments with the twin goals of minimizing both the supply / demand gap and the overall cost. Solutions take the form of investment schedules indicating the optimal allocation of public funds devoted to each wedge over several decades. While future research is necessary to precisely hone parameter values, solutions from the proof of concept model presented here indicate that investment schedules prioritizing near-term interventions across the board, including demand-reducing wedges such as family planning and supply-increasing wedges such as yield gaps and climate adaptation, together with a bump in funding for key supply-increasing interventions starting around 2035, represent the most financially-efficient means of closing the supply / demand gap.

5.2 Introduction

5.2.1 Problem Statement

Global food demand in the year 2050 is estimated to increase by between 116% and 219% from 2010 demand (Keating and Carberry 2010, Alexandratos et al. 2012, Valin et al. 2014, Pardey et al. 2014, Keating et al. 2014, Cole et al. 2018), driven largely by projected increases in the global population and shifting diets.

Global population is expected to increase from 7.2 to almost 10 billion people by 2050 (Brown and Cameron 2000, Heikkinen 2014, UN 2015). However, simultaneously, diets are shifting toward both increased overall per-capita caloric demand, as well as higher proportions of animal-derived protein consistent with a "nutrition transition" (Popkin 2001, Popkin et al. 2012). Both of these dietary shifts are being largely driven by rising per-capita wealth, with an individual in a high-income country currently demanding 256% more overall calories and 430% more protein than an individual in a lower-income country (Brown and Cameron 2000, Amine et al. 2003, Herrero et al. 2013). With average global wealth (in terms of per-capita GDP) growing at around 2.5% per annum, both of these factors will account for a large proportion of projected demand looking ahead to 2050 (Tilman et al. 2011).

These increasing demands for food correspond with existing and projected future challenges to food security. In 2017, more than 820 million people were undernourished in the world, the third year in a row that food insecurity rates climbed despite decades of previous decreases. Such increases have been attributed to instability resulting from violence and war in certain regions and increasing extreme climate events in others (FAO 2018). Solutions to these inequalities have been debated for decades, and are embodied within the Sustainable Development Goals, particularly Goal Two to attain a safe and adequate food supply. Achieving this goal will thus require a number of strategies that can aim to address these inequalities and consider

food system investments that can both potentially increase food supply while also addressing food demands.

While food security involves multiple components including food availability, accessibility, utilization, and stability (Napoli et al. 2011), food supply and demand both affect all dimensions of food security. Several solutions have been proposed to curb increasing food demand. For example, Ranganathan et al. (2016) assess the ways in which changes in dietary patterns could theoretically achieve favorable nutritional outcomes, while lowering agricultural production demand, and also decreasing environmental externalities. Garnett (2016) notes that food policy interventions often aim to influence social and cultural values rather than focusing on agricultural technology—the former being more complex but also equally if not more important in addressing global food security—highlighting the tradeoffs at play in the strategic allocation of funds.

On the supply side of the coin, technological improvements in agriculture, such as those associated with the "green revolution," have kept approximate pace with demand (distribution issues notwithstanding) for the past several decades, however serious questions exist as to whether supply-increasing innovations can keep up with demand going forward, especially considering that many of our current intensive agricultural practices have negative long-term effects on the ability of the land to maintain current yields (Eswaran et al. 2001, Nelson et al. 2009, Van Ittersum et al. 2013, Sadras et al. 2015, Cole et al. 2018). While many propose simply

increasing the land area devoted to agriculture, sustainable intensification theorists focus on increasing agricultural efficiency, curbing land degradation, reducing waste, and adapting to climate change as means of increasing supply (Borlaug 2008, Tilman et al. 2002, Foley et al. 2011, Garnett et al. 2013).

In the absence of political interventions, the gap between food supply and demand is likely to widen significantly, with a maximum deficit predicted somewhere around 2050, after which the global population is expected to level off or possibly even decrease as industrialization reduces birth rates (Foley et al. 2011, Keating et al. 2014). Ensuring that the supply curve meets the demand curve over the next several decades is critical for policymakers for a number of reasons. Human wellbeing requires access to safe, healthy, and nutritionally-sufficient levels of culturally-appropriate foods; and ensuring an adequate global food supply is an important component of achieving this goal (Sen 1999). Further, from a geopolitical perspective, regional food insecurity has often led to political instability and conflict (Lagi et al. 2011, FAO 2018).

5.2.2 Sustainable Development Goals

In 2012, the United Nations Rio+20 Summit laid out a set of sustainable development goals (SDGs) to fight global poverty and secure human wellbeing (Griggs et al. 2013). As of September 2015, governments have committed to achieving these SDGs by 2030 (Schmidt-Traub and Shah 2015). While the scope of the SDGs is wide, SDG Two is specifically concerned with ensuring a safe, adequate food supply to feed

186

our growing population, while also minimizing the environmental impact of food production. Achieving SDG Two requires reducing global hunger to less than 5% of the population, 20% improvements in agricultural nutrient use efficiency, and several environmental goals aimed at curbing soil and water degradation and reducing the climatological footprint of agricultural production (especially from animal products). While these broad targets have been codified and agreed upon—and indicators of success identified—significant uncertainty remains as to the best means of meeting SDG Two, especially considering the complex interactions between interventions (Robert et al. 2005).

Working toward the establishment of a practical plan of action, Schmidt-Traub and Shah (2015) propose an analytical framework to weigh investment priorities into the SDGs, identifying eight key investment areas, one of which is food security and sustainable agriculture. The authors then estimate the share of public and private financing that should be devoted to each investment area, as well as the economic and social tradeoffs that must be considered, concluding that overall about 1.5-2.5% of global GDP needs to be allocated to achieve the SDGs. Within the food security and sustainable agriculture investment area, the authors identify an array of necessary investments, including "…responses to emergencies; improving nutrition; meeting the special needs of smallholder farmers or artisanal fishermen; maintaining and restoring productive soils; rural infrastructure; increasing the productivity and sustainability of commercial agriculture; reducing greenhouse gas emissions from agriculture; …

increasing the resilience of agriculture to climate change; [and] research and development for agriculture and food security" (Schmidt-Traub and Shah 2015, p.58).

Schmidt-Traub and Shah (2015) go on to evaluate four existing publications which investigate investment needs to meet SDG Two, arriving at a preliminary estimate to meet the goal of around $148 billion per year globally between 2016 and 2030. While some estimates of the annual cost required to reduce global hunger to acceptable levels—such as Laborde et al.'s (2016) calculation based on the computable general equilibrium MIRAGRODEP economic model—project lower investment requirements ($11 billion annually in this case), these models do not consider the complex interactions between production efficiency, consumption patterns, changing diets, nutritional requirements, environmental impacts, and unavoidable climate change.

To account for these important factors, Schmidt-Traub and Shah's (2015) report highlights several important knowledge gaps that require further research to bolster the robustness of projected investment requirements. They call for the need for integrated assessment models that account for the "multiple and complex" relationships between investments in each sub-area, and for more research focused on dietary change, climate-resilient production practices, nutrition, and other factors.

Other researchers have also identified the need to grapple with the complex interactions between interventions designed to address the SDGs, asserting that the international negotiations that gave rise to the SDGs were inhibited by the siloed

188

knowledge-base underlying each sub-goal, thus neglecting to establish a framework

through which investment tradeoffs may be analyzed, and funding prioritized over time

(Pongiglione 2015, Nillson et al. 2016). Nillson et al. (2016) introduce a seven-point

scale describing how goals may range from "indivisible" (one cannot be accomplished

without the other) to "cancelling" (the goals are mutually exclusive). For example, they

describe how climate-change mitigation interacts with food security: on one hand—due

to agriculture's GHG footprint—climate mitigation efforts may constrain food

production to some extent; while at the same time, a stable climate is required for a

resilient food supply. Similarly, land taken out of cultivation for reasons of ecosystem

protection will constrain food production.

These studies suggest that there is a clear need for additional research on the

strategies necessary for achieving SDG Two, particularly the need to account for the

complex interactions and tradeoffs between individual goals. Developing an effective

plan of action will require a decision support system capable of weighing such tradeoffs

and prioritizing intervention investments over time.

5.2.3 Previous Research into the Complexity of Multiple Food Security Interventions

Keating et al. (2014) analyze the complex nature of these food system

interactions in a "wedge model" for achieving global food security. The authors build

upon the seminal work of Pacala and Socolow (2004), which attempts to quantify

how—using a variety of interventions, each aimed at tackling a smaller sub-issue—the

overall problem of global climate change could be effectively solved. Both efforts

acknowledge that climate change and future food security, like other complex issues

facing our global society, are something of a "wicked problem" (Churchman 1967);

that is, a conundrum that rests upon probabilistic analyses, and cannot be solved by any

single intervention.

To assess these complexities, Keating et al. propose three "mega wedges"—

namely: reducing demand; filling the production gap; and avoiding losses—that, taken

together, should be able to overcome the 127×10^{15} kcal deficit the world is projected to

face between 2010 and 2050 (Figure 22). Under this food wedge model, each "mega

wedge" is in turn composed of several individual wedges that each contribute to filling

some amount of the projected gap. Through an expert informant advisory process, the

authors attempt to identify the percentage of resources that should be allocated to each
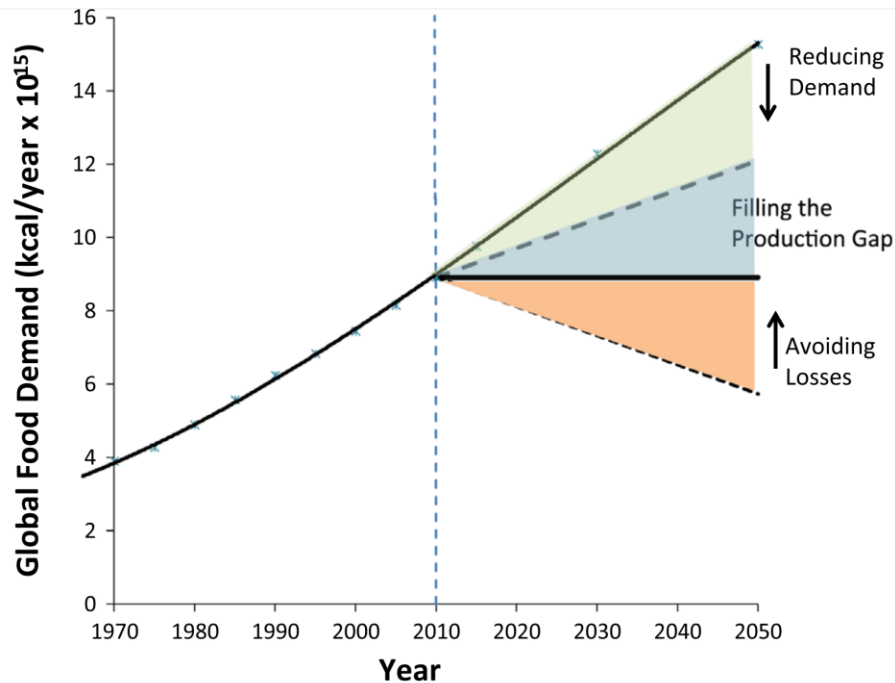
wedge.

Figure 22: Theoretical food security wedge model from Keating et al. (2014). The top line shows the historical and projected global food demand, and the bottom line shows food supply. The colored regions illustrate how the wedge model could reduce demand and increase supply to fill the projected supply / demand gap.

While Keating et al.'s (2014) framework represents a valuable theoretical contribution, it is hampered by a number of critical oversimplifications, some of which are inherently identified by food security scholars as being necessary for achieving SDG Two (Robert et al. 2005, Pongiglione 2015, Schmidt-Traub and Shah 2015, Nillson et al. 2016). Firstly, it is assumed that all interventions occur simultaneously, in 2010. Second, it is assumed that the effect of each intervention will proceed linearly, beginning at the intervention date and extending *ad infinitum*, with the effect on supply

191

or demand unfolding at a constant rate. Finally, the authors suggest no mechanism to govern the limitations of each wedge given ongoing investments.

From a pragmatic perspective, it is clear that the food wedge model requires re-thinking if it is to be useful as a practicable decision-support tool. Similar to the criticisms of SDG Two implementation discussed above (Robert et al. 2005, Pongiglione 2015, Schmidt-Traub and Shah 2015, Nillson et al. 2016), the food wedge framework needs to bridge between siloed knowledge bases, accounting for factors including interactions between interventions and the time required for an intervention to come to fruition to determine an effective prioritization schedule for policy actions.

Several efforts have been made to pick up where Keating et al.'s (2014) model left off. Kummu et al. (2017) review recent literature to examine the potential for interventions directed at diet change, food loss reduction, and yield gap closure (in terms of both nutrient and water utilization) to increase food supplies to the projected 2050 food demand, which scholars estimate may be as much as a 100% increase over current demand levels. Following Foley et al. (2011) and Keating et al. (2014), among others, they argue that there is a need to de-emphasize individual solutions—especially those that derive from increased resource use—and instead focus on the interplay between various opportunities targeting both supply and demand.

Kummu et al. (2017) calculate the projected potential level of change for both production-side measures and consumption-side measures based on country-specific data and existing production models. In a departure from the food wedge framework—

192

which considers intervention effects in terms of kcals—the authors instead encode each

measure as a percentage increase or decrease over current supply or demand.

Multiplying together the effects of all production-side measures with current production

levels in kcal by country, and repeating the process for consumption-side measures, the

authors calculate that the potential does indeed exist to close the projected supply /

demand gap, although they identify several limitations, such as the fact that climate

change is not considered in their model. Going forward, the authors call for several

methodological developments to obtain a more accurate, integrated understanding.

They identify three broad categories that need further research, (a) linking

developments, such as case studies and localized field research, (b) supportive

developments, such as scenario analyses and feasibility studies, and (c) core

developments, the most critical of which is the inclusion of dynamic feedback between

interventions.

5.2.4 Contributions of Our Study

In light of the identified need to incorporate complex relationships between

intervention and prioritization schedules into the implementation plan for SDG Two

(Robert et al. 2005, Pongiglione 2015, Schmidt-Traub and Shah 2015, Nillson et al.

2016)—i.e. to address interventions "in combination, not just in parallel" (Kummu et

al. 2017)—we aim to develop the wedge framework into an integrated decision support

tool that can be used to guide and prioritize policymaking. Following previous work in

this area (Foley et al. 2011, Keating et al. 2014, Kummu et al. 2017), we divide wedges

into two basic categories: supply-side interventions (i.e., increasing production) or demand-side interventions (i.e., reducing consumption). An additional insight which governs how wedges interact is that some wedges (e.g. increasing land use) will manifest as increases or reductions in supply or demand, whereas others (e.g. changing diets) will instead serve to multiply or divide the existing level of supply or demand. We therefore introduce the sub-classifications of additive / subtractive versus multiplicative wedges.

We can expect each wedge to contribute a certain level of baseline change (change in supply or demand that would happen in the absence of any intervention), and this change will unfold over some period of time. Also, given a certain quantity of global public investment into a given wedge, a corresponding quantity of change may be expected, and, once again, this will occur over some characteristic timeframe. For example, expanding land resources used for agriculture could theoretically be a relatively quick-acting intervention, whereas significantly influencing human dietary choices can be expected to take years, or perhaps even generations. Finally, each wedge is associated with a maximum possible level of change. For example, there is only so much additional land that could reasonably be converted to agricultural use, such that at a certain point additional investments will exhibit diminishing returns.

Based on the above, this paper presents the computational method we have developed to formalize the above parameters using a series of three mathematical functions to characterize each wedge. We then demonstrate how an evolutionary

algorithm can be employed to optimize wedge investments over time, report on an experiment we conducted to optimally parameterize the algorithm, and finally discuss preliminary results.

5.3 Methods

5.3.1 Optimization of Intervention-Driven Change

Based on each wedge's parameters, our algorithm stacks the curve associated with each wedge's baseline together with the curves resulting from public policy allocations into that wedge over time. The optimization procedure we have developed generates an intervention schedule represented by a vector whose values indicate the fraction of total annual funding available for food security that should be allocated to each wedge in each year. The effects of a given intervention depend on previous interventions, the magnitude of investment, and the shape of the wedge's characteristic intervention curve.

The goal is to optimize the investment schedule such that supply meets demand throughout the entire time period of interest, for the least possible amount of public investment. The problem is therefore a multi-objective optimization, encompassing both fitness in terms of the supply-demand difference, as well as cost. Each optimal solution will rest on the Pareto front defined by these twin objectives.

A given solution takes the form of a schedule of global public policy interventions into each wedge—which could be implemented at either the national (e.g. a country's agricultural oversight agency) or international (e.g. the U.N. FAO) level—

over the time period of interest (e.g. from present day to 2050). Because of the interdependency between the effects of the wedges upon one another—for example the fact that the effect of multiplying wedges upon the final supply or demand curve depends on previously-implemented reducing/increasing interventions, and vice-versa—the optimization problem becomes quite complex. An evolutionary computational approach capable of identifying near-optimal solutions within a solution space governed by a system of interdependent equations would seem to be the only feasible means to tackle such a problem. Here we approach the optimization by leveraging a multi-objective differential evolution (MODE) algorithm (Reynoso-Meza et al. 2014).

5.3.2 Wedges Used in the Experiment

While we base our work on Keating et al. (2014), the number of wedges included in our experiment has been cut down from the 14 originally identified to seven key wedges, described below. While future versions of our wedge optimization algorithm will incorporate at least 14 wedges, for this proof of concept—in order to limit computational overhead—we chose to focus on the six wedges identified by Keating et al. (2014) as especially important, in addition to a seventh, family planning, which has been identified as a critical factor toward addressing SDG Two (Starbird et al. 2016). We retain the wedge numbering scheme from Keating et al. (2014), but wedge categorizations have been updated to clarify their use in our algorithm. While we cannot claim that the present analysis is sufficiently-parameterized to forecast

196

precise results in real-world systems, we believe this project is valuable as it identifies

"stylized facts" about the system, which can help stakeholders to understand and

effectively grapple with key questions surrounding complex tradeoff landscapes, as

well as generating new insights to ground future research efforts (Epstein 2008).

Wedge parameters required for our model were derived from existing empirical

data and primary literature sources (Table 12, and described below). It is important to

note that many of these values, though based on an analysis of the scientific literature

from each domain, largely represent the best-estimates of our team. The experiment

presented here is meant mainly as a proof of concept to demonstrate our optimization

procedure. Future work will involve efforts to more-realistically parameterize the

wedges based on continued scientific research and expert input.

Table 12: Wedge Parameter Values Used in the Experiment

Num.	Description	Class	Δ_{base}	$t(\Delta_{base}/2)$	n_{base}	Δ_i	$t(\Delta_i/2)$	n_i	Δ_{max}	a
5	Land expansion	↑ Supply	1.05×10^{15}	50	1	4×10^{14}	3	2	1.6×10^{16}	2
12	Soil/water degradation	↑ Supply	0	–	–	7×10^{14}	5	2	1.1×10^{16}	2
8	Yield gaps	× Supply	1.05	20	1.25	1.03	8	2	1.1	10
14	Climate adaptation	× Supply	0.8	25	2	1.1	10	2	1	2
15	Birth control	↓ Demand	1.28×10^{16}	50	1.25	-1.2×10^{15}	15	3	-2.2×10^{16}	2
2	Over-consumption	× Demand	1.05	25	1.25	0.98	5	2	0.75	10
3	Diet	× Demand	1.05	35	1.25	0.98	14	2	0.6	10

The most critical parameters for each wedge are as follows: Δ_{base} represents the expected level of baseline change in the absence of any intervention. $t(\Delta_{base}/2)$ indicates the number of years into the future at which half of the baseline change is expected to have unfolded. Similarly, Δ_i is the expected level of change resulting from expending a full year's global food security policy budget solely on the wedge, and $t(\Delta_i/2)$ is the expected number of years into the future at which half of that intervention-driven change is expected to have unfolded. Finally, Δ_{max} represents the theoretical upper limit for change associated with each wedge. The other parameters are less critical, tuning more subtle aspects of the equations, such as the transition slope. The specific role of each parameter in the mathematical functions underlying our model are detailed in "5.3.4.1 Wedge functions and parameter encoding."

5.3.2.1 Supply increasing wedges

- *Expanding agricultural land resources (Wedge 5)*

The USDA reports that global land use for agricultural production increased from 1.65 billion acres in 1961 to around 2.32 billion acres as of 2016, an increase of 40.6% (USDA PSD 2017). Theoretically, there may be as much as 6 billion acres of additional land that could be allocated for agriculture—informing our Δ_{max} value of 1.6×10^{16} kcal—although this figure includes reserved and protected land which, if utilized, would have significant impacts on environmental resources (FAO 2011, FDI 2011).

198

Land quality has a large impact on per-acre kcal production, with 80% of potential yield presently met with prime land, 40-60% with good land, and < 40% with marginal land. In many cases, it is marginal land that has not yet been cultivated, which will limit the productivity potential of newly-cultivated land (Barbier 2004). As such, the impetus to cultivate new land—while projected to continue gradually—has already tapered off and is expected to plateau in the coming decades (FAO 2011).

Although more research is needed, given the factors above, we estimate that baseline land expansion may increase by around 12% between 2015 and 2050. With current global food production around 9×10^{15} kcal (Keating et al. 2014), this yields a Δ_{base} of 1.05×10^{15} kcal. However, due to the marginal nature of much of the potential new acreage, we anticipate a relatively long $t(\Delta_{base}/2)$ of 50 years. In contrast, we anticipate that intervention-driven change may unfold quickly for this wedge, since adding new land for agriculture could be incentivized with simple zoning / regulatory changes, and as such we estimate a Δ_i of 4×10^{14} kcal, and a $t(\Delta_i/2)$ of three years.

- *Avoiding soil and water degradation (Wedge 12)*

Approximately 25% of cultivated land is currently considered highly degraded, with another 44% being moderately degraded (FAO 2011), causing, on average, 20-40% reductions in yield (Eswaran et al. 2001, Morgan 2009). Some of this land is inherently marginal (Barbier 2004), but it has been suggested that many commonly-used intensive agricultural practices, together with alternative land uses such as

199

urbanization, have led to the degradation of previously-prime land and water resources (Bindraban 2012). However, better management practices in recent years, such as limited tillage, have begun to curb the problem, and the UN FAO currently does not consider it a major threat (Gomiero 2016). Given the uncertainties involved, we set the Δ_{base} for this wedge to zero.

Multiple factors influence the stability of agricultural landscapes, presenting many routes for improvements to zoning, soil management, runoff control, and water quality that may begin to reverse soil and water degradation (Gomiero 2016). Historically, the American dust bowl disaster shows the potential for degraded soil to come back into production, given sufficient policy action. While more research is required to determine the projected effects of interventions in this area, we have used an estimated Δ_i of 7×10^{14} kcal and a $t(\Delta_i/2)$ of five years in the current model.

5.3.2.2 Supply multiplying wedges

- *Closing yield gaps in existing crop and/or livestock production systems (Wedge 8)*

Current estimates for potential yield improvements vary by crop type. For example, while average rainfed wheat yields could improve by 25%, irrigated corn already achieves 90% of its potential yield (Van Ittersum et al. 2013, Sadras et al. 2015). Investing in yield gaps is multiplicative, since it increases per-acre production efficiency on existing farmland.

The global trend over the past century has been toward increased per-acre yields—largely due to technological improvements resulting from e.g. the "green

revolution"—yet in recent years yields of key staple crops have stagnated and are not expected to keep pace with demand going forward (Ray et al. 2013, Pradhan et al. 2015). While potential gains are possible—especially in developing regions including much of Africa—in the absence of a fundamental paradigm shift, researchers project only minor improvements in average global yields over the coming decades (Van Ittersum et al. 2013). In light of these findings, we use a Δ_{base} multiplier of 1.05 (i.e., a 5% gain), and a $t(\Delta_{base}/2)$ of 20 years in the model.

Despite the fact that the effectiveness of the 20th century paradigm to address yield gaps is diminishing, there are still many interventions that have the potential to increase production efficiency over the longer term (Rodriguez et al. 2018). The current state of the art involves a focus on sustainable intensification, including precision agriculture, new management practices, localized cropping decisions, and other factors (Foley et al. 2011, Mueller et al. 2012, Cassidy et al. 2013, Steffen et al. 2015, Rockström et al. 2017). While a full analysis of the potential effects of sustainable intensification remains outside the scope of the current project, it is clear that improvements are possible in this area, and as such we set Δ_i to 1.028 (i.e., a 2.8% gain) and $t(\Delta_i/2)$ to 8 years, accounting for the fact that yield improvement innovations have historically spread rather slowly across the globe (Feder and Umali 1993).

- *Adaptation to climate change that can't be avoided (Wedge 14)*

Climate change is expected to have a major negative impact on agriculture, with estimated annual kcal production losses of 20% over current levels projected by 2050 if

systems do not adapt (Nelson et al. 2009). Thus, we set our Δ_{base} multiplier at 0.8,

and our $t(\Delta_{base}/2)$ at 25 years for this wedge.

While the baseline multiplier indicates a loss in productivity in the absence of

any intervention, investment in the adaptation of agricultural systems to maintain

productivity in the face of a changing climate may have the potential to mitigate or in

some cases even reverse these projected losses, for example where a warming climate

increases the length of the growing season, or where increased CO_2 accelerates crop

growth. By breeding new varietals of drought-tolerant crops and educating farmers to

select cultivars that are more amenable to their changing local climates, the negative

baseline trend can be mitigated to some extent (Bisbis et al. 2018). In light of these

findings, we estimate our Δ_i at 1.1 (i.e., a 10% gain) and $t(\Delta_i/2)$ to 10 years, owing to

the expected lag between the development of climate-smart agricultural practices and

their eventual adoption by farmers on the ground (Burke and Emerick 2016).

5.3.2.3 Demand reducing wedges

- *Curbing population growth by expanding access to contraceptives (Wedge 15)*

While this wedge was not included in Keating et al. (2014), evidence suggests

that it may in fact be a crucial leverage point, since more than anything else, rising

demand is driven by population. The number of humans is expected to increase from

7.55 billion to 9.77 billion by 2050 and 11.2 billion by 2100 (UN 2015). These

numbers are staggering, as each person consumes an average of 1073 kcal annually

(Brown and Cameron 2000). Of course, the situation is complex, since for example

individuals in lower-income countries—where contraception tends to be less available, and fertility rates are higher—also consume less kcal per capita than their richer neighbors. In light of these projections, we set the Δ_{base} to 1.28×10^{16} kcal for this wedge, albeit with a very long $t(\Delta_{base}/2)$ of 50 years, indicating that population growth will be a steady demand-increasing baseline factor over the next century.

Despite the baseline increase, providing widespread access to contraception and family planning education could drastically curb human population growth, with a theoretical possible reduction—assuming one child per woman—of three billion individuals, equating to 3.219×10^{12} kcal annually, by 2045 (Heikkinen 2014). In light of its huge impact on the overall food security problem, SDG Two identifies family planning as one of the most powerful levers toward reducing hunger in the coming decades (Starbird et al. 2016). We set our Δ_i at -1.2×10^{15} kcal, representing a significant demand reduction. However, achieving this will require a long-term commitment, since reproduction choices are wrapped up in cultural norms, likely requiring generational shifts in public opinion (Heikkinen 2014). As such, we estimate $t(\Delta_i/2)$ for this wedge to be 15 years.

5.3.2.4 Demand multiplying wedges

- *Reducing over-consumption by humans (Wedge 2)*

Over-consumption of food—where caloric intake is higher than caloric expenditure in an individual person—results in increased demand that is not necessary

for healthy human livelihoods. This wedge is multiplicative, since it is contingent upon the number of humans on the planet. Globally, adults consume 2940 cal. on average daily (Brown and Cameron 2000), whereas the recommended healthy caloric intake is around 2000 cal. (Amine et al. 2003), though this is highly dependent on individual lifestyles and characteristics. This means that, in theory, this wedge could reduce demand by around 30%, although we choose a more realistic figure of 25% reduction to inform our Δ_{max} value of 0.75.

Global per-capita calorie consumption has increased 31% in the last half-decade, largely driven by ongoing industrialization and development in formerly-underdeveloped regions (Roser and Ritchie 2019). This consumption trend is projected to continue in underdeveloped areas, yet it has already tapered off in the world's most-developed countries. As such, we roughly estimate a Δ_{base} value of 1.05 (a modest 5% increase) for this wedge, with a $t(\Delta_{base}/2)$ of 25 years. The impact of intervention-driven change designed to reduce overconsumption is also fairly uncertain, although the literature does indicate that interventions to shift dietary choices can be effective to a certain extent (Racey et al. 2016). Here we estimate a Δ_i of 0.98 (a 2% reduction) and $t(\Delta_i/2)$ of 5 years. Given the uncertainty in the literature, more-precisely forecasting trends and projecting intervention impacts toward reducing overconsumption would be a valuable avenue for future research.

- *Rebalancing the livestock component in human diets (Wedge 3)*

Global average meat consumption has increased from 23 kg per person per year in 1961 to 42 kg per person per year in 2011, an 82% change (Sans and Combris 2015). The energetic requirements to raise livestock have in some cases been estimated to be as high as 100 times more costly than if rowcrops were used to feed humans directly (Herrero et al. 2013), although the practices, geographical location, and other factors associated with animal production complicate this figure. Nevertheless, increased consumption of animal-derived products will on average require higher overall levels of agricultural production, in terms of kcal potential for human nutrition (Jalava et al. 2016).

It has been widely reported that rising incomes tend to inflate per-capita consumption of animal-derived food products, however there is also an income threshold—around $36,000 per capita—beyond which meat consumption has been found to decline (Vranken et al. 2014). With global incomes generally rising on average, under baseline conditions meat consumption in wealthy areas is projected to plateau and eventually decrease, whereas in areas experiencing rapid economic development—notably China—demand for animal products will continue to rise. While the rapid increase in meat consumption observed over the last half-century is already tapering off, there is still an overall upward trend globally. In light of these countervailing factors, we estimate a Δ_{base} value of 1.05 (a 5% increase in kcal demand) for this wedge, with a $t(\Delta_{base}/2)$ of 35 years.

Despite the complexities associated with the impact of animal agriculture on food production, climate, and human nutrition, a modest reduction in the average proportion of animal-based foods in human diets represents a critical multiplicative lever to reduce pressure on global food supplies. Like overconsumption, the efficacy of policies designed to effect dietary change are difficult to predict, but literature shows that eating habits can in fact be nudged, albeit gradually (Racey et al. 2016). We therefore estimate a Δ_i of 0.98 (a 2% reduction in kcal demand) and $t(\Delta_i/2)$ of 14 years for this wedge, with a Δ_{max} of 0.6.

5.3.3 Genome Representation

Evolutionary algorithms represent the variable values associated with each potential solution as a "genome." Here, the genome used to encode each solution was implemented as a vector of floating-point values between zero and one. The genome can be divided into a series of repeating blocks, one for each intervention year. Each gene value within these blocks encodes the percentage of global funding available for public investments into food security allocated to each wedge in that year. Using the notation $g_{w,y}$, where g is the gene value for wedge w in intervention year y, and with n_w and n_y being the number of wedges and the number of intervention years, respectively, the final genome takes the general form:

$$
\begin{bmatrix}
g_{1,1} & g_{2,1} & \cdots & g_{n_w,1} \\
g_{1,2} & g_{2,2} & \cdots & g_{n_w,2} \\
& & \vdots & \\
g_{1,n_y} & g_{2,n_y} & \cdots & g_{n_w,n_y}
\end{bmatrix}
$$

5.3.4 Objective Function

The development of an objective function that meets the requirements of a "curve-stacking" optimization problem like the one we have outlined represents a key contribution of our work. Such a function must handle interventions that fall into each of the four wedge classifications described above. The function must appropriately sum or multiply the supply or demand effects resulting from multiple individual interventions for each wedge, with each intervention occurring in a specific year and unfolding over time according to several compounding parameters associated with the wedge, as well as the expenditure magnitude represented by the genome value. Finally, it must compute the fitness associated with the overall intervention schedule described by each genome according to two distinct objectives: the area between the final supply and demand curves, and the total summed implementation cost.

The objective function takes as arguments vectors specifying the wedges from each category to be used in the analysis, the start and end year for the optimization (i.e. 2020 through 2050), the intervention frequency (i.e. re-evaluating investment priorities every five years), the initial supply and demand figures (in kcal), as well as several parameters associated with deriving the final fitness values (discussed below).

207

5.3.4.1 Wedge functions and parameter encoding

Each wedge has nine associated parameters—according to the values given in Table 12 above—stored in a structure array. First is the wedge name, and second is its class, which can be one of "supply increasing," "supply multiplying," "demand reducing," or "demand multiplying." The units for increasing / decreasing classes are represented in kcal, while the multiplicative wedges result in multiplier values that are applied to either the global supply or demand curve.

The remaining numerical parameters are employed within three functions, each of which is called where applicable by the objective function to calculate the overall multiplier or kcal Δ vector resulting from the series of interventions associated with each wedge. These wedge functions are: (a) the baseline function, (b) the intervention-driven change function, and (c) the diminishing returns function.

(a) The baseline function for each wedge is a Hill function of the form:

$$f(x) = \frac{\Delta_{base} x^{n_{base}}}{(t(\Delta_{base}/2))^{n_{base}} + x^{n_{base}}}$$

Where x is the number of elapsed years since the start year, Δ_{base} is the ultimate level of baseline increase or decrease (or, in the case of multiplicative wedges, the maximum or minimum baseline multiplier value), n_{base} is the slope of the transition (as $n_{base} \to \infty$ the Hill function becomes a step function; we found values between one and three to be realistic), and $t(\Delta_{base}/2)$ is the number of years before half of the Δ_{base} value will have been reached.

(b) The intervention-driven change function for each wedge was another Hill function of the same form as the baseline function, but calibrated to output a vector the length of the number of years between the intervention year and the end year. The intervention parameters are used in the intervention-driven change function in the same way as the baseline parameters in the baseline function, substituting Δ_{base} for Δ_i, $t(\Delta_{base}/2)$ for $t(\Delta_i/2)$, and n_{base} for n_i. The difference is that, in the case of intervention-driven change, the effect on the supply or demand curve only begins the year in which an intervention is scheduled. Δ_i represents the total change (in terms of either multiplier or kcal Δ) associated with allocating 100% of a given intervention year's food policy resources to that wedge. $t(\Delta_i/2)$ represents the time in years from the intervention year until the year in which the change would equal $\Delta_i/2$. n_i encodes the transition slope associated with the intervention. The function and its associated parameters are visualized in Figure 23.
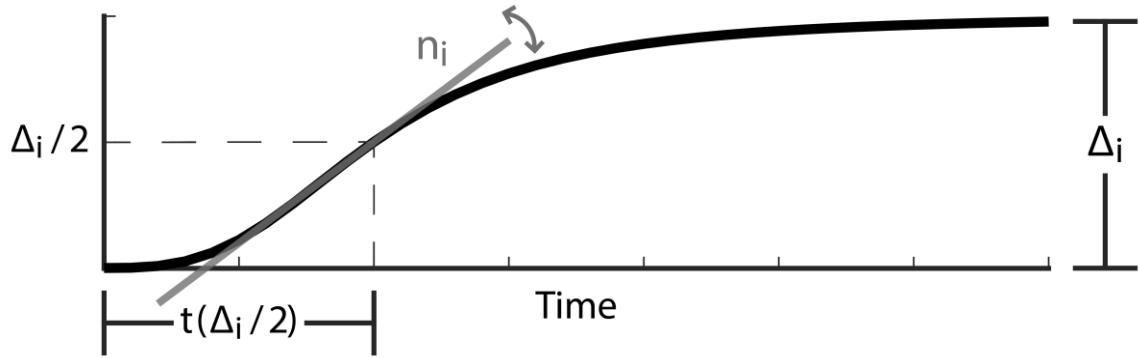
Figure 23: Hill function of the form used for both the baseline and intervention-driven change associated with each wedge, showing the effect of the three key parameters on the shape of the resultant curve.

(c) Finally, the diminishing returns parameters describe how additional investments into a wedge will begin to become less and less effective. For example, there is only so much additional agricultural efficiency that we can reasonably hope to achieve, and at some point additional investments into that wedge will begin to pay smaller and smaller dividends. We found that a modified logistic growth/decay function suited the diminishing returns situation well, since, when properly parameterized, it has a $y \approx x$ trajectory initially, followed by a gradual transition to a specified limit. The diminishing returns function took the following general form:

$$f(k) = \frac{a\Delta_{max}}{1 + e^{-ak/\Delta_{max}}} - \Delta_{max}$$

Where Δ_{max} describes the ultimate maximum level of change that may be gained through investments into each wedge (see Figure 24), and a encodes the transition

210

abruptness as regards the diminishment of intervention effectiveness as more

investments are made. The function should initially yield $y \approx x$ (i.e. the outcome from

the first intervention should equal Δ_i), while further equal investments in the same

wedge should yield outcomes $<\Delta_i$. We found that $a = 2$ produced such an outcome for

the additive / subtractive wedges; while $a = 10$ produced the same pattern for the

multiplicative wedges.
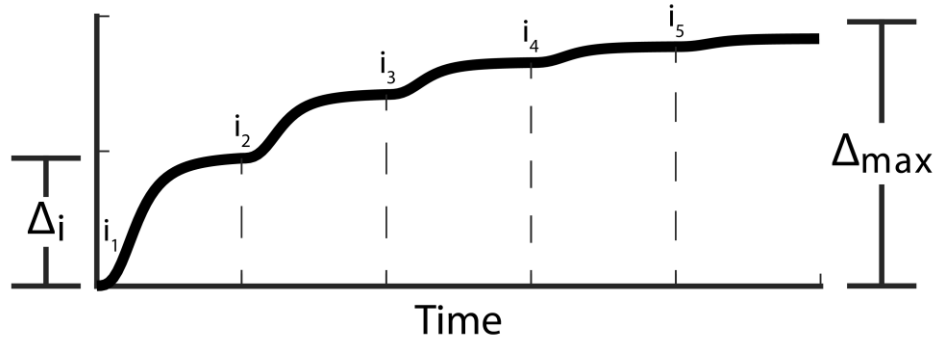
Figure 24: Effect of the diminishing returns function on five equal intervention investments $[i_1 \cdots i_5]$

into the same wedge. Note that Δ_{max} is a limit to which a series of interventions may approach but

never reach.

We have chosen to use Hill functions and logistic growth / decay functions here

because they were deemed to best suit the problem at hand. Were a similar

optimization method to be used for other applications, the structure of the algorithm

makes it trivial to use whatever mathematical functions are most appropriate to

parameterize each wedge. Note that we have demonstrated only the supply-increasing ($y \propto x$) versions in the equations and figures above. For demand-reducing ($y \propto -x$) wedges, the equations were modified such that they were mirrored across the x axis, and then translated up where necessary.

5.3.4.2 Curve stacking procedure

As noted above, food security wedges may act either to increase supply, or to decrease demand. For brevity, we will discuss only the supply side curve stacking procedure here, with the demand side calculations simply being a repeat of the same process. Within each broad category, wedges may be broken down into two sub-classes: (a) multiplying wedges, and (b) increasing/reducing wedges.

To track the supply multiplying wedge effects in aggregate, an overall supply multiplier vector of ones the length of the years vector is initialized. For each multiplying wedge, a separate tracking vector of the same length is initialized by calling the baseline change function using that wedge's baseline parameters as arguments. A loop then traverses each genome index associated with the wedge. If the genome value is non-zero, the intervention-driven change function—again using the wedge's individual parameters as arguments—is executed for the intervention year through the end year. The magnitude of the resulting intervention multiplier vector is then multiplied by the expenditure fraction represented by the genome value, which is always ≤1. Next, the corresponding segment of the wedge multiplier vector is multiplied by the intervention multiplier vector. The individual intervention vectors are

212

also stored separately for later plotting. Once all intervention multipliers associated with a given wedge are combined into a single vector, that vector is run through the wedge's diminishing returns function. Finally, each supply multiplying wedge's associated vector is multiplied together to generate the overall supply multiplier vector.

The stacking procedure is similar for the supply increasing wedges, except that each wedge's effect vector is initialized with values of zero, and instead of multiplying vectors together, they are summed. Once again, the procedure is to first establish a baseline vector for a wedge, then modify this vector by adding in each intervention-driven change vector for the wedge, and finally run the resulting summed vector through the diminishing returns function. Once all supply increasing wedge vectors are calculated in this way, they are summed into a single overall supply increase vector.

To generate the final supply curve, a vector the length of the years vector is initialized with each value being equal to the initial supply level in kcal. This is then summed with the overall supply increase vector. Finally, the resulting vector is multiplied by the overall supply multiplier vector to yield the final supply curve. As noted above, this full procedure is repeated—with appropriate modifications—for the demand-side wedges to calculate the final demand curve.

5.3.4.3 Fitness calculation

The optimization problem described here encompasses two distinct objectives, namely: (O1) minimize the total deficit between food supply and demand in each year, and (O2) minimize the average annual cost of global food security interventions.

213

We call the fitness value associated with (O1) "supply / demand (or SD) fitness." This is calculated by traversing the data in one-year segments, and using trapezoidal numerical integration on each segment to find the area between the supply and demand curves. If supply is lower than demand, the area value is added to the total area difference. In initial testing, we found that penalizing overproduction allowed the algorithm to hone in on more efficient solutions. In the real world, this would correspond to limiting food waste. To capture this, we stipulate that in years in which supply exceeds demand, fitness is penalized by multiplying the area associated with the overproduction by the "overproduction penalty factor" parameter (=1.1) before adding it to the total area difference. Finally, the total area difference is normalized by dividing it by the area between the pre-intervention baseline supply and demand curves, giving us the SD fitness figure.

We call the fitness value associated with (O2) "cost fitness." Cost fitness is calculated by dividing the sum of the genome by Δ_t, the number of years between the end year and start year. This encodes the average annual percentage of possible global food security policy resources required to obtain the solution.

5.3.5 Differential Evolution Model

In selecting a process to optimize the objective function described above, we considered factors such as how the genome was implemented, what the solution space may look like, problem dimensionality, and computational requirements.

214

The differential evolution model, pioneered by Storn and Price (1997), like all evolutionary algorithms, is based on multi-generational mutation, recombination, and selection. In each generation, the algorithm selects three population members at random. Mutation proceeds by adding the weighted difference between the first two vectors to the third. The resulting "donor vector" is then recombined with a "target vector" to form a new vector called the "trial vector." Finally, the fitness of the target and trial vectors are compared, and the better of the two is selected for admission to the next generation.

Since we have two potentially-competing objectives, an algorithm that outputs a set of possible solutions along an approximate Pareto front (i.e. a multi-objective algorithm) was necessary. Several multi-objective extensions of the basic DE algorithm have been implemented (Xue et ql. 2003, Reynoso-Meza et al. 2014). Our procedure utilizes a Matlab MODE implementation called spMODEx, which incorporates a novel method of spherical pruning into the basic MODE algorithm (Storn and Price 1997). We modified the code such that it calls a custom population generation function appropriate for the problem at hand, retrieves and passes the appropriate wedge parameter and experimental design arguments through to the objective function, and utilizes a p-factor parameter—which is multiplied by problem dimensionality to calculate population size, as is common in other DE implementations—rather than a fixed population size. We also wrapped the algorithm

code in a driver function which initializes and executes the experiment described below, and developed extensive scripts to appropriately analyze and visualize results.

Aside from being comparatively-straightforward to implement, DE algorithms have performed very well on a range of benchmarks (Storn and Price 1997, Das and Suganthan 2011). Since, at least to the authors' knowledge, this study represents the first attempt to optimize a curve-stacking problem of this type using an evolutionary algorithm, several unknowns exist about the fitness landscape and solution space. For example, whether or not the problem is linearly separable or unimodal is somewhat unknown. The fact that DE has proved very versatile, and has relatively-few tunable parameters, whose actions are well-understood, suggests it to be a good fit for a novel class of optimization problems (Das and Suganthan 2011).

One thing we do know is that our class of optimizations are, or at least could be, very high-dimensionality problems. Dimensionality can be calculated here based on three problem parameters, as follows:

$$D = \frac{n_w \Delta_t}{f_i}$$

Where n_w is the number of wedges, Δ_t is the difference between start year and end year, and f_i is the intervention frequency. For example, in the experiment below, we use a 30-year time window, a five-year intervention frequency, and 7 wedges, meaning the problem takes on a dimensionality of $D=42$.

Another factor in our optimization is its relatively-expensive objective function evaluation procedure, which requires several loops and some amount of pre-processing overhead. Because of its low space-complexity, DE has proven especially adept at handling both high-dimensionality problems and those with relatively-expensive objective function evaluations (Das and Suganthan 2011), once again suggesting that it is well suited to our optimization problem.

5.3.6 Experimental Design

An iterative process of parameter tuning was undertaken to hone in on baseline parameter values which consistently produced high-quality solutions. To evaluate and optimize the MODE algorithm for the curve-stacking problem described above, we conduct a sensitivity analysis experiment which systematically varies several key model parameters, and analyzes the impact of these changes on solution fitness. We also examine the effect of each of these reparameterizations on average computation time. Finally, we delve deeper into the best set of solutions and comment on the patterns and tradeoffs therein.

<u>5.3.6.1 Parameters of interest</u>

First we explore several categorical parameters (Table 13). Recombination method indicates the algorithm used to merge the genomes of the target vector and donor vector to form the trial vector, with most DE implementations historically employing binomial recombination (Xue et al. 2003, Reynoso-Meza et al. 2014). Selection strategy affects how the target vector is compared to the trial vector to

determine which will be admitted to the next generation. Here we compare traditional

dominance-based ("Push") selection to Reynoso-Meza et al.'s (2014) novel spherical

pruning ("SphP") approach. Normalization algorithm indicates the specific algorithm

to be used in the spherical pruning process.

Table 13: Categorical Parameters and Values Explored

Parameter	Baseline	Values Explored
Recomb. Method	'Binomial'	['Binomial,' 'Lineal']
Selection Strat.	'SphP'	['SphP,' 'Push']
Normalize. Algo.	'Euclidean'	['Euclidean,' 'Manhattan,' 'Infinite,' 'Physical']

Next, we evaluate how changing the values of key numerical parameters affects

results, as well as the models' degree of sensitivity to change when increasing and

decreasing the baseline values by a known percentage. In this case, we evaluate the

baseline value against $0.75 \times$ baseline and $1.25 \times$ baseline for each parameter of interest

(Table 14). Number of initial interventions corresponds to the number of non-zero

genome entries assigned randomly to each initial population member. *p*-factor

indicates the size of the initial population, with population size equaling *p*-factor $\times$

problem dimensionality (*D*=42 in this case). Scaling factor—variously known as

"weighting" or "mutation" factor—affects the extent to which the difference between

the first two vectors changes, or mutates, the third vector when forming the donor

vector, and should be between zero and two. Crossover probability is the chance that

each element of the donor vector will enter the trial vector (Storn and Price 1997).

Finally, we tune the maximum number of generations to investigate the point at which the increased run-time associated with more generations no longer yields sufficiently-better solutions to continue increasing the generation limit.

Table 14: Numerical Parameters and Values Explored

Parameter	0.75 × Baseline	Baseline	1.25 × Baseline
N. Init. Intervs.	15	20	25
P Factor	3.75	5	6.25
Scaling Factor	0.16875	0.225	0.28125
Crossover Prob.	0.225	0.3	0.375
Max. Generations	96	128	160

5.3.6.2 Experiment setup and analysis

Each of the parameterizations detailed above was run 15 times, which was generally sufficient to evaluate the performance of each treatment against the others. For parameters where results were inconclusive after 15 runs (Normalize. Algo., N. Init. Intervs., and Max. Generations), an additional 15 repetitions were completed for each parameterization. A total of 495 runs were conducted for this experiment. The solutions along the Pareto front from each run were saved in a data structure, along with their accompanying fitness values, and the time elapsed for each run. For each parameterization, a combined Pareto front was generated, composed of the fittest individuals across all repetitions. To evaluate the effect of changing parameter values on model performance, we also generate a combined Pareto front composed of the best individuals across all values evaluated for each parameter.

The share of individuals in the combined Pareto front which were drawn from each parameterization serves as a reliable indicator of solution quality. Realizing that, for the problem at hand, the solutions we're most interested in are those with very favorable supply-demand fitness (i.e. we want to ensure that the world is fed first and foremost), we define a solution acceptability window, with a maximum SD fitness of 0.0175, and a maximum cost fitness of 0.25. Within this window, all solutions come very close to fully satisfying the SD objective, while also rejecting the few solutions with extremely-high cost (which may also be overfitting the data). We then calculate the Pareto front share for each parameterization, constrained within this window. This value—the Pareto front share within the acceptability window—serves as the primary indicator of solution quality used to parameterize the model in this experiment.

Finally, we compute averages and standard deviations for the runtimes to evaluate how each parameter change affects the time required for the algorithm to complete. Our goal is to weigh the efficacy of the solutions generated by each parameterization against any runtime penalty to arrive at a suitable baseline parameterization that consistently produces good results, without requiring excessive computation time.

Baseline values were honed using an iterative series of experimental reparameterizations. The results below reflect the best baseline parameter values we were able to achieve and show that deviating from these baselines negatively affects the performance and/or computation time of the algorithm.

220

5.4 Results

We begin by examining the experimental data to evaluate how the baseline spMODEx parameterization we have settled on for this use case weighs tradeoffs between performance and execution speed. Next, we analyze the full set of baseline-parameterized runs to explore similarities and differences among solutions within the acceptability window.

5.4.1 Model Parameterization Experimental Results

Figure 25 shows experimental results for the categorical parameters. We find that the binomial recombination method, which is typically preferred in most DE implementations, is superior both in terms of performance and speed (Storn and Price 1997, Xue et al. 2003). However, the spherical pruning selection strategy developed by Reynoso-Meza et al. (2014) outperforms the dominance-based strategy canonically used in DE algorithms, confirming that, at least for this class of problems, spherical pruning is a worthwhile amendment to the MODE algorithm.
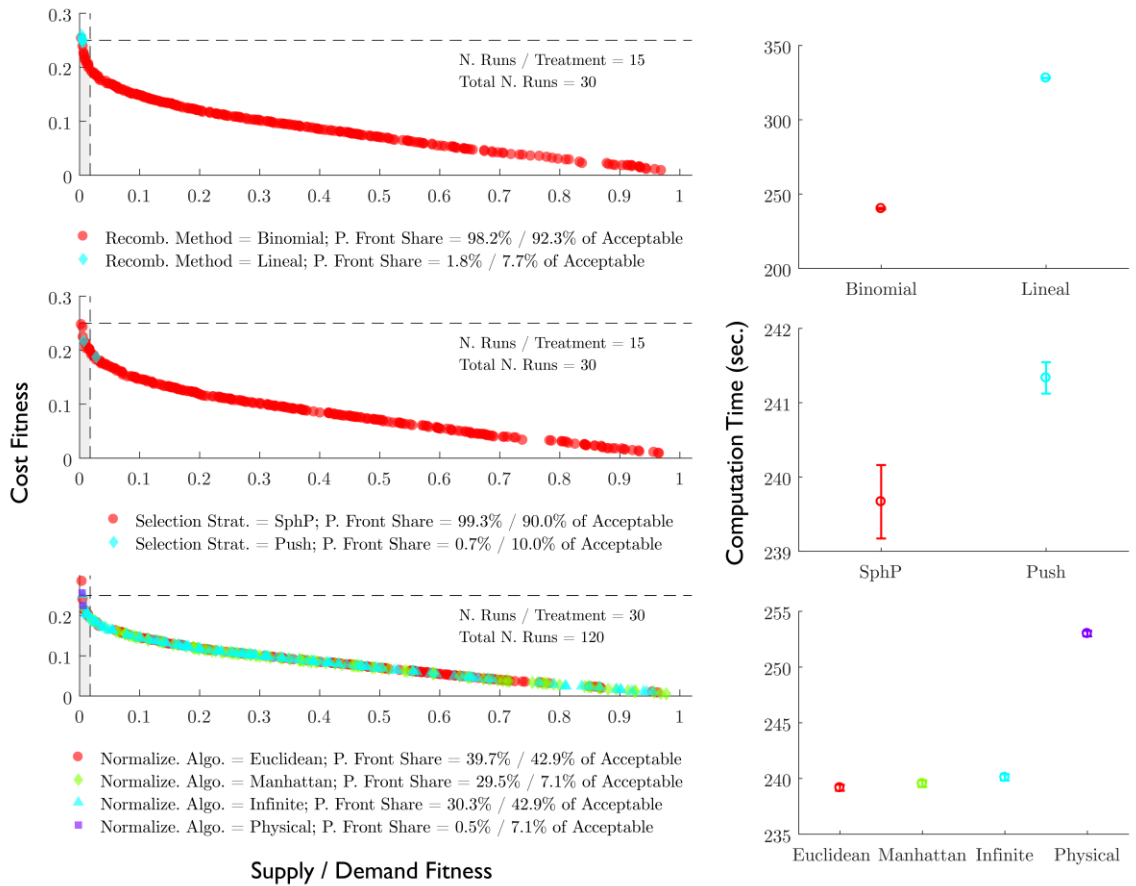
Figure 25: Left plots show SD Fitness / Cost Fitness Pareto fronts for categorical parameters. Gray areas bounded by dashed lines indicate solution acceptability window. The share of the Pareto front generated by each parameterization appears in the legend, both over the full range, as well as strictly within the acceptability window. Right plots show mean and std. errors of computation time per run for each parameterization.

While Euclidean, Manhattan, and Infinite normalization algorithms show little difference overall in terms of either performance or execution time, we find that for this use case the Euclidean algorithm is preferable, capturing more of the Pareto front share

overall, and an equal amount as the Infinite algorithm within the acceptability window; while slightly edging out the other algorithms on speed.

Next, we turn to the numerical parameters, each of which was evaluated at 0.75 × baseline, baseline, and 1.25 × baseline levels (Figure 26). Experimental results show that the model is quite sensitive to the number of initial interventions in each genome, with interesting tradeoffs between performance overall, performance within the acceptability window, and speed. Whereas 15 initial interventions yielded the highest overall performance and the fastest execution time, it captured only 22.2% of acceptable solutions. On the other hand, 25 initial interventions captured the highest proportion of acceptable solutions—despite generating only 0.8% of overall solutions—but this came at the cost of ≈20 second longer run times. We conclude that in this case 20 initial interventions offers a happy medium, generating many acceptable solutions without sacrificing a great deal of execution time. We also noticed that increasing the number of initial interventions tended to produce intervention schedules with higher numbers of smaller interventions. While potentially more precise, these more complex solutions may come at the cost of being more difficult to for policymakers to parse and effectively utilize. For a global problem like adequate food production, the "simpler" solutions—i.e. those that rely on fewer individual intervention events—may be easier to implement and communicate to stakeholders across the globe than intervention schedules requiring very precise and timely coordinated actions.

223
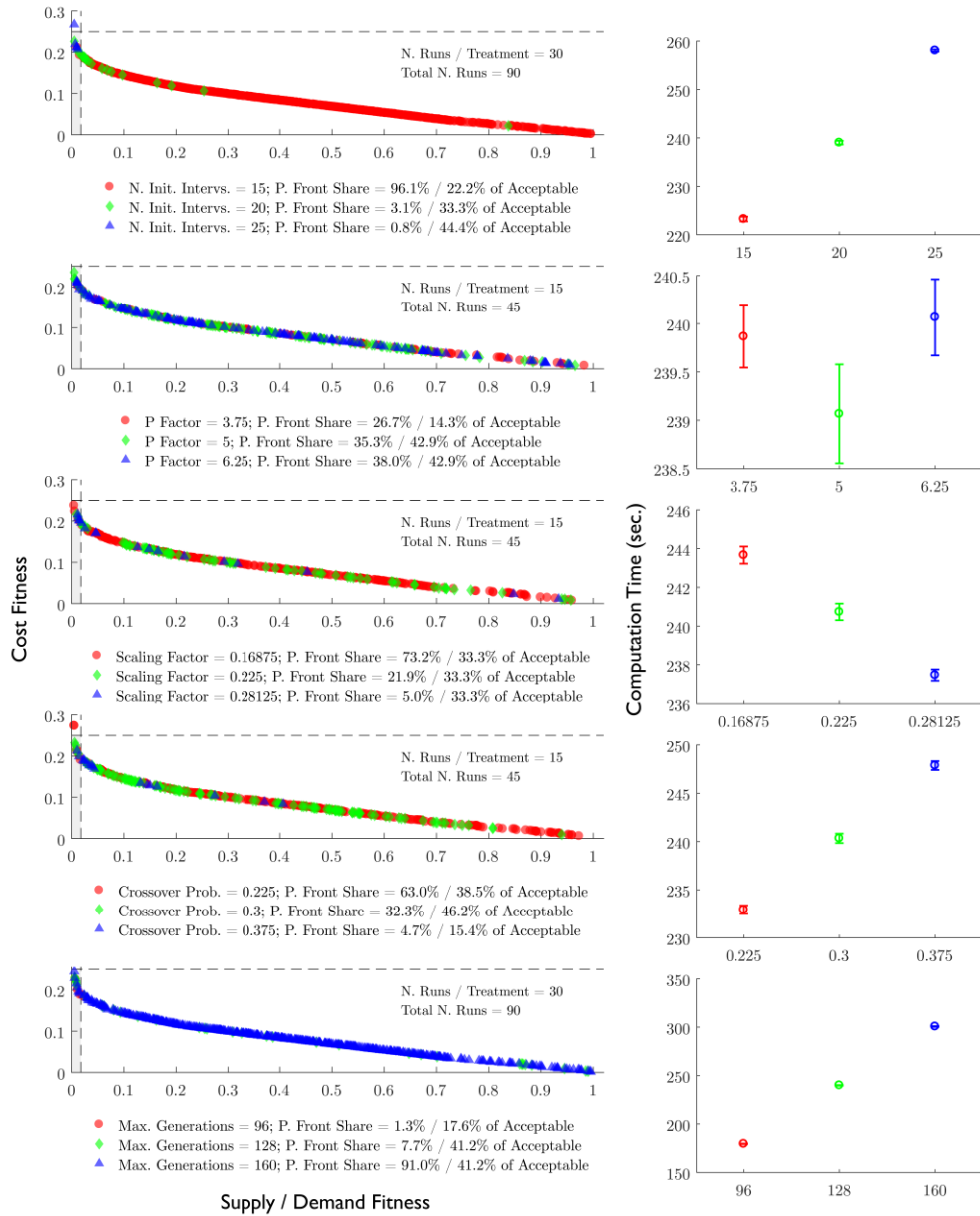
Figure 26: Left plots show SD Fitness / Cost Fitness Pareto fronts for numerical parameters. Gray areas bounded by dashed lines indicate solution acceptability window. The share of the Pareto front generated by each parameterization appears in the legend, both over the full range, as well as strictly within the acceptability window. Right plots show mean and std. errors of computation time per run for each parameterization.

224

Overall, the model was minimally sensitive to changes in p-factor, which dictates the size of the initial population. This finding echoes previous studies which find that, due to its design, the efficacy of DE is less reliant on population numbers than many other evolutionary algorithms (Storn and Price 1997, Xue et al. 2003, Das and Suganthan 2011). We found that a p-factor of five yielded an equal proportion of acceptable solutions to any other value we assessed, while also executing slightly faster on average.

Turning to scaling factor, the model would appear to be sensitive with regard to speed, but less so with regard to performance, especially within the acceptability window. While each value we assessed produced an equal share of acceptable solutions, lower values tended to generate more Pareto front solutions overall but were also more computationally costly. Thus, we arrived at a baseline value of 0.225, as this offers a compromise between the three factors.

For crossover probability, higher values require more computational time, while also decreasing performance on the overall Pareto front. We found that a value of 0.3 yielded the highest proportion of acceptable solutions, while maintaining moderate execution times.

Finally, as should be expected, increasing the maximum number of generations had a large impact on execution time. It is desirable to choose a value that achieves the desired results without unduly burdening the model user with overly long computation

225

times. In this case, we found that a limit of 128 generations yielded the same
percentage of acceptable solutions as higher generation limits, while also preserving
reasonable execution times.

5.4.2 Best Acceptable Solutions

To determine the best overall solution(s) to the optimization problem at hand,
we concatenate results from all baseline-parameterized runs to calculate a final Pareto
front, and focus in solely on the subset of solutions within the acceptability window.
This acceptable Pareto front contains 19 solutions (Figure 27), three of which—located
at the "knees" of the front—are selected for detailed analysis.
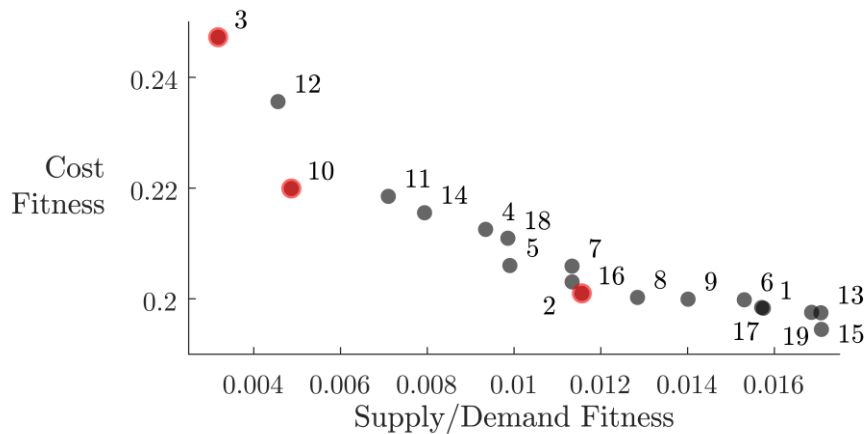
Figure 27: SD Fitness / Cost Fitness Pareto front for the 19 baseline-parameterized solutions within the acceptability window. Red solutions (3, 10, & 16) are detailed in Figure 28.

Figure 28 plots the intervention schedules associated with each of the three
optimal solutions (3, 10, and 16 from Figure 27). The top plots of Figure 28 show the

effect of each intervention on either global food supply (cool colors) or demand (warm colors). Note that it is possible for multiple interventions to occur in a single intervention year. Essentially, the colored regions expanding from the intervention year and extending out toward 2050 can be thought of as the "wedges" of Keating et al. (2014). The cost fitness and SD fitness associated with each solution are also given (with lower values being better in this case). The bottom plots of Figure 28 indicate the level of resource allocation (in terms of percentage of global public policy budget available for food security) devoted to each wedge in each intervention year.
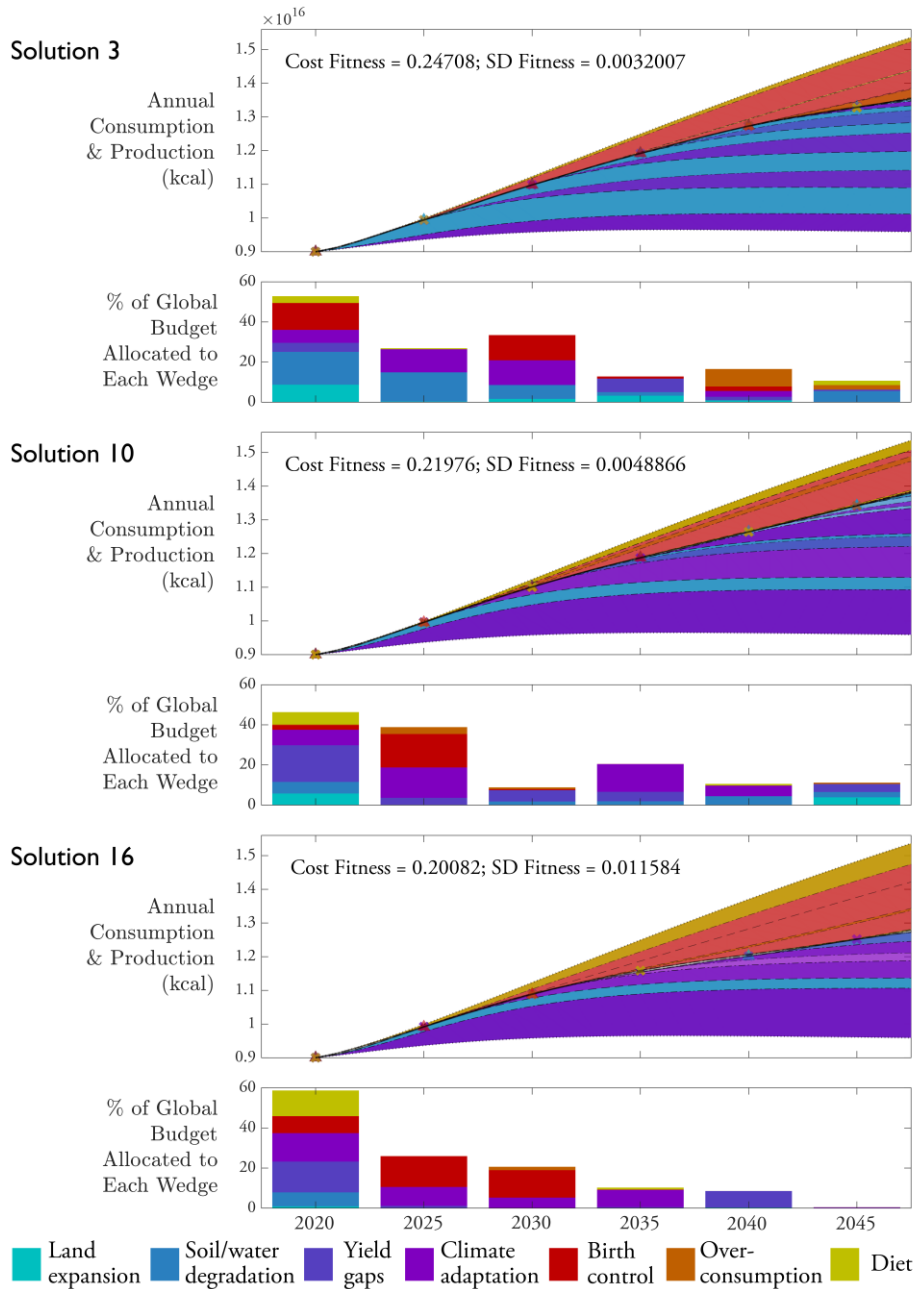
Figure 28: Example solutions—located at the "knees" of the Pareto front—within the acceptability window. Upper plots for each solution show the baseline supply and demand curves, along with the effect of each wedge intervention called for by the solution on either supply (cool colors) or demand (warm colors). Lower plots show the intervention investment schedule indicated by each solution.

228

5.4.3 Prioritizing Wedge Interventions Based on DE Solutions

Taken together, the three solutions identified above share several similarities as well as demonstrating key differences (Table 15). Clearly, there is more than one way to achieve an effective food security intervention schedule, emphasizing the complex tradeoff landscape at play. Solutions 3 and 10 prioritize supply-side interventions, with 3 emphasizing curbing soil and water degradation, and 10 focusing more on closing yield gaps (e.g. by sustainably-intensifying agricultural production). Both of these wedges offer substantial opportunities to bolster agricultural output over the coming decades, and policymakers must carefully weigh the tradeoffs between the two. Solution 16 differs somewhat in that it relies more heavily on curbing demand through birth control and reducing the average quantity of meat in human diets. Notably, what all three solutions have in common, regardless of the overall makeup of supply versus demand side wedges, is that the majority of investments invest heavily into climate adaptation (on the supply side) and/or birth control (on the demand side).

Table 15: Overall Average Allocation to Each Wedge for the Three Best Acceptable Solutions

	Allocation Percentage		
Wedge	**S3**	**S10**	**S16**
Land expansion	9.72%	7.09%	0.99%
Soil/water degradation	29.45%	12.17%	5.59%
Yield gaps	8.99%	26.57%	20.27%
Climate adaptation	21.69%	30.82%	30.49%
Birth control	19.06%	14.31%	30.13%
Over-consumption	7.14%	3.93%	1.57%
Diet	3.95%	5.11%	10.97%

Among the three solutions, 16 has the lowest cost fitness, indicating that an acceptable global food security outcome could be met most cost-effectively by following this intervention schedule. On the other hand, 3 has the lowest SD Fitness, meaning that this intervention schedule should yield the highest level of global food availability throughout the study period.

While drilling down into individual solutions illuminates the potential tradeoffs at play within the solution landscape, another analysis strategy is to investigate overall patterns across the set of acceptable solutions. Figure 29 shows the average level of allocation to each wedge when considering all 19 acceptable solutions. Here we observe several key takeaways. First off, it is notable that the most efficient solutions prioritize heavy up-front investments across the board. By acting earlier rather than later, less resources will ultimately be required to solve the problem. Second, examining Table 15 together with the overall average allocation—over all intervention years—across acceptable solutions (Table 16), we find that early interventions (up

through 2030) into birth control—a long-term yet potent demand-side intervention—appear consistently across many solutions and make up the majority (≈23%) of total supply side allocations. Meanwhile, supply-side interventions—primarily yield gaps and climate adaptation, at ≈28% and ≈23%, respectively—should receive steady investment over the short term, and perhaps even a bump in funding around 2035.
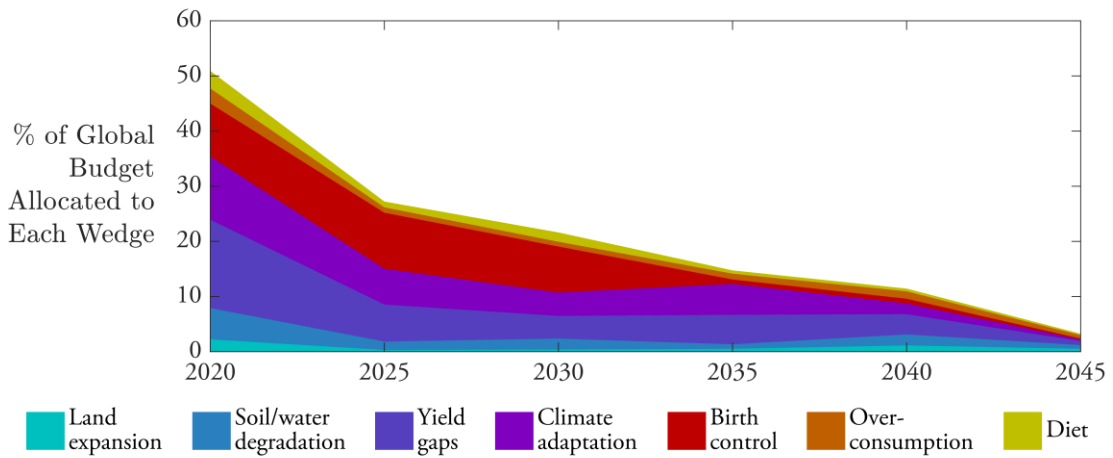
Figure 29: Average percent of global budget allocated to each wedge in each intervention year across all 19 solutions within the acceptability window.

Table 16: Overall Average Allocation to Each Wedge across All Intervention Years for 19

Acceptable Solutions, with Peak Investment Years

Wedge	Allocation Percentage	Top Two Investment Years
Yield gaps	28.3411%	2020, 2025
Climate adaptation	23.2418%	2020, 2025
Birth control	23.1549%	2025, 2020
Soil/water degradation	9.8146%	2020, 2040
Over-consumption	5.8547%	2020, 2040
Diet	5.6264%	2020, 2030
Land expansion	3.9665%	2020, 2040

While the approach of averaging acceptable solutions illuminates overall

patterns and trends associated with our proof of concept results, it is important to note

that this method of analysis has limitations, and should not necessarily be relied upon to

determine an optimal course of action once the model is more-accurately

parameterized. For the conceptual model presented here, averaging solutions is

informative, but in the future we will also consider how solution patterns change across

the Pareto front in order to better highlight the tradeoffs at play across the fitness

landscape. A related approach would be to test the extent to which the fitness scores of

averaged solutions fare against the set of individual solutions. While outside the scope

of the present study, this remains an area for future research.

5.5 Discussion

UN Sustainable Development Goal Two commits world governments to the

twin objectives of drastically reducing worldwide hunger as well as mitigating the

worst of the environmental impacts associated with agricultural production by the year

2030 (Griggs et al. 2013, Schmidt-Traub and Shah 2015). Achieving this goal will not

be simple, since the effects of multiple interventions—including family planning,

human dietary change, agricultural production practices, waste reduction, climate

adaptation, and more—need to be carefully weighed (Robert et al. 2005). The primary

policymaking challenges revolve around the facts that the set of potential interventions

interact with one another in complex ways, take different amounts of time to unfold,

and act on diverse structures within our complex food system. These factors make

investment prioritization toward achieving SDG Two very complex (Pongiglione 2015,

Nillson et al. 2016). For this reason, previous scholarship in this area has highlighted

the need to develop integrated assessment models that can effectively weigh tradeoffs

and optimize intervention investment schedules, accounting for complex interactions

between the set of proposed solutions (Schmidt-Traub and Shah 2015). In this paper,

we have demonstrated the efficacy of a multi-objective differential evolution algorithm

to generate solutions to the type of complex, dynamic, intervention-driven problems

that characterize SDG Two.

Our study was aimed at developing and experimentally-analyzing a novel

computational method to parse and optimize food policy investment schedules,

following up on the basic wedge framework of Pacala and Socolow (2004) and later

Keating et al. (2014). The wedge framework provides a valuable lens through which

systems-level problems may be broken down into manageable pieces. Keating et al.

(2014) use an expert informant advisory process to identify the theoretical potential—in

233

terms overall contribution percentage—of each wedge toward ensuring adequate future

global food supplies. Our work picks up on this theoretical grounding, providing

insights into not only the level of public policy focus each wedge should receive in a

general sense, but also the best way to schedule interventions into each wedge over

time. For example, while Keating et al. (2014) simply conclude that 14.8% of

resources should be devoted toward yield gaps and 10.4% toward soil and water

degradation, our model provides a structure for not only the overall quantity of

resources allocated to each wedge, but also the ways in which investments in each

wedge should be prioritized over time, accounting for factors including time lags and

inter-wedge interaction effects. We hope that our computational methods—together

with domain-level expert collaboration—may eventually provide an actionable tool

which can be utilized to guide real-world investment prioritization decisions.

The objective function we have developed to pair with the MODE

algorithm (Storn and Price 1997, Reynoso-Meza et al. 2014) accounts for all inter-

wedge interactions as it "stacks" both the baseline and intervention-driven change

curves across multiple wedges, intervention years, and levels of investment. To

account for the complex interactions between multiple interventions over time, we

build on the existing wedge model, differentiating between additive / subtractive

wedges and multiplicative wedges, as well as between those which increase food

supply versus decrease demand.

We go on to experimentally investigate the sensitivity of the algorithm to changes in eight key MODE parameters. Through this parameterization process, we arrive at suitable baseline values which account for both solution fitness and execution time. While we find that parameter changes do meaningfully affect the algorithm's performance, the model is generally robust, as minor fluctuations in most parameter values do not cause marked effects on either the cost or SD fitness of the final solutions. This can be seen in the left columns of Figure 25 and Figure 26, which show that—despite differences in Pareto front share—solutions from all parameterizations we explored generally fall within a consistently-shaped curve, suggesting that most solutions we obtained are reasonably-close to the true Pareto optima. Thus, while optimizing the parameterization for fitness and computational speed is an important goal, gains are mostly made at the margins.

Using the baseline parameterization, the algorithm produces a variety of different solutions within the acceptability window of cost and SD fitness. The extent to which the end user wishes to prioritize each objective can thus be effectively weighed. The study presented here illuminates several insights about intervention investment priorities in the area of future global food availability, demonstrating the complex tradeoffs involved in such an optimization problem. For example, prioritizing key interventions sooner rather than later can drastically decrease the total resources required to achieve long-term policy goals. Further, complex choices must be made between e.g. interventions which have a high ultimate potential for change (large Δ_{max})

235

but take many years to come to fruition (large $t(\Delta_i/2)$), versus those with a lower Δ_{max}

and a faster-acting $t(\Delta_i/2)$. Other important factors that may be incorporated into future

model versions include ease of implementation, uncertainty when predicting

intervention outcomes, and scale of action. Such considerations illustrate the value—

and in fact, necessity—of computational methods to aid policymakers and practitioners

in such complex prioritization tasks. Simply put, no unaided human or group of

humans is capable of intuiting solutions to this class of problems.

　　While each solution within the acceptable Pareto front differs to some extent,

averaging solutions over multiple restarts reveals that the algorithm generally identifies

similar macro-level investment prioritization patterns across all acceptable solutions

(Figure 29). Since, for this type of optimization problem, there may very well be

multiple intervention schedules which lead to similar cost and SD fitness values, we

suggest that a multiple-restart approach represents a valuable tool for policymakers or

others who wish to apply an algorithm like the one we have presented here to a real-

world domain. However, as discussed above, a systematic evaluation of trends among

acceptable solutions along the Pareto front should also be conducted for a better

understanding of the solution landscape.

　　While we focus on global food security policy prioritization, we contend that

this methodology could be applied to a wide variety of domains in which investment

prioritization amidst a complex tradeoff landscape is called for. For example, in the

business context it may be desirable to achieve a certain minimal level of net profit for

236

each quarter, while investing any remaining income back into the business in various ways. This type of problem has many similarities to the one we've presented here, and a similar schedule of investments could be optimized to allow the business owner to maximize the effectiveness of her investments over time.

5.5.1 Limitations

We wish to reiterate that the model as it currently exists is parameterized based on the work of Keating et al. (2014), together with a review of key scientific literature pertinent to each wedge. This represents a limitation of the study, as there is certainly more domain-level expertise available than the scope of this study has allowed to be incorporated into the model as it currently stands. To transform this proof of concept into a practicable decision support tool, we intend to convene panels of experts who have devoted their careers to studying each wedge, and work collaboratively with these stakeholders to hone wedge parameters such that they represent the best current scientific understanding. A fully-calibrated model would also include, at minimum, all 14 food security wedges originally identified by Keating et al. (2014). An expert advisory process could help to solidify the final set of wedges to employ in the analysis.

For a variety of reasons, discussed above, we have chosen to employ the spMODEx algorithm here (Reynoso-Meza et al. 2014). Despite the demonstrated success of this approach, another avenue for future research is to investigate how other evolutionary algorithms perform against our current MODE implementation. For example, the multi-objective CMA-ES algorithm has also proven very capable on high-

237

dimensionality problems, and outperforms DE on some non-separable problems with many local minima (Rahnamayan and Dieras 2008). Comparing the two algorithms on the class of problems described here was outside the scope of this experiment but may be a valuable avenue for future research.

Additionally, while our model accounts for inter-wedge interactions in the context of sufficient food production, we don't consider how investments into these wedges may have other implications and effects. For example, land expansion is likely to be plagued by negative environmental externalities. Meanwhile, dietary shifts away from meat may have unintended public health consequences. These important considerations remain active topics of academic study and could potentially be incorporated into future versions of our model.

A final limitation, identified by Kummu et al. (2017), concerns the loci of control associated with global policy implementation. Even if an intervention investment schedule to address global hunger is optimized at the global level, this does not guarantee food security in all regional contexts. Further research is required to better understand the inequalities that underlie these heterogeneous effects and develop targeted policy solutions that can be effectively carried out within especially-vulnerable locales.

5.6 Conclusion

The looming threat of global food insecurity is likely one of the most critical public policy challenges we currently face. We believe our technical approach,

although leaving ample room for future research, represents a valuable potential tool to optimize investments in global food policy over the coming decades. Aside from other "wicked problems" facing society as a whole—such as climate change, which has also been described through a wedge model lens—several other uses for our methodology are feasible. Overall, we hope that the method we have developed can to help tackle both global food security, as well as a variety of other dynamic intervention-driven change problems where the goal is to continually achieve some end (i.e., make two curves match at every interval) at a minimal cost.

5.7 Acknowledgments

5.8 References

ALEXANDRATOS, N., BRUINSMA, J., & OTHERS. (2012). WORLD AGRICULTURE TOWARDS 2030/2050: THE 2012 REVISION. RETRIEVED FROM HTTP://LARGE.STANFORD.EDU/COURSES/2014/PH240/YUAN2/DOCS/AP106E.PDF

AMINE, E., BABA, N., BELHADJ, M., DEURENBERY-YAP, M., DJAZAYERY, A., FORRESTER, T., . . . MBUYAMBA, J. (2002). DIET, NUTRITION AND THE PREVENTION OF CHRONIC DISEASES: REPORT OF A JOINT WHO/FAO EXPERT CONSULTATION: WORLD HEALTH ORGANIZATION.

BARBIER, E. B. (2004). EXPLAINING AGRICULTURAL LAND EXPANSION AND DEFORESTATION IN DEVELOPING COUNTRIES. AMERICAN JOURNAL OF AGRICULTURAL ECONOMICS, 86(5), 1347 - 1353.

BINDRABAN, P. S., VAN DER VELDE, M., YE, L., VAN DEN BERG, M., MATERECHERA, S., KIBA, D. I., . . . HOOGMOED, M. (2012). ASSESSING THE IMPACT OF SOIL DEGRADATION ON FOOD PRODUCTION. CURRENT OPINION IN ENVIRONMENTAL SUSTAINABILITY, 4(5), 478-488.

BISBIS, M. B., GRUDA, N., & BLANKE, M. (2018). POTENTIAL IMPACTS OF CLIMATE CHANGE ON VEGETABLE PRODUCTION AND PRODUCT QUALITY–A REVIEW. JOURNAL OF CLEANER PRODUCTION, 170, 1602-1620.

BORLAUG, N. E. (2008). FEEDING A WORLD OF 10 BILLION PEOPLE: OUR 21ST CENTURY CHALLENGE PERSPECTIVES IN WORLD FOOD AND AGRICULTURE 2004 (PP. 31-56). AMES, IOWA, USA: IOWA STATE PRESS.

BROWN, P., & CAMERON, L. (2000). WHAT CAN BE DONE TO REDUCE OVERCONSUMPTION? ECOLOGICAL ECONOMICS, 32(1), 27 - 41.

BURKE, M., & EMERICK, K. (2016). ADAPTATION TO CLIMATE CHANGE: EVIDENCE FROM US AGRICULTURE. AMERICAN ECONOMIC JOURNAL: ECONOMIC POLICY, 8(3), 106-140.

CASSIDY, E. S., WEST, P. C., GERBER, J. S., & FOLEY, J. A. (2013). REDEFINING AGRICULTURAL YIELDS: FROM TONNES TO PEOPLE NOURISHED PER HECTARE. ENVIRONMENTAL RESEARCH LETTERS, 8(3), 034015.

CHURCHMAN, C. W. (1967). WICKED PROBLEMS [GUEST EDITORIAL]. MANAGEMENT SCIENCE, 14(4), B141-B142.

COLE, M. B., AUGUSTIN, M. A., ROBERTSON, M. J., & MANNERS, J. M. (2018). THE SCIENCE OF FOOD SECURITY. NPJ SCIENCE OF FOOD, 2(1), 14.

DAS, S., & SUGANTHAN, P. (2011). DIFFERENTIAL EVOLUTION: A SURVEY OF THE STATE-OF-THE-ART. IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, 15(1), 4-31.

DELUCIA, E. H., GOMEZ-CASANOVAS, N., GREENBERG, J. A., HUDIBURG, T. W., KANTOLA, I. B., LONG, S. P., . . . PARTON, W. J. (2014). THE THEORETICAL LIMIT TO PLANT PRODUCTIVITY. ENVIRONMENTAL SCIENCE AND TECHNOLOGY, 48(16), 9471-9477.

EPSTEIN, J. M. (2008). WHY MODEL? JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 11(4), 12.

ESWARAN, H., LAL, R., & REICH, P. (2001). LAND DEGRADATION: AN OVERVIEW. RESPONSE TO LAND DEGRADATION, 20-35.

FAO. (2004). THE ETHICS OF SUSTAINABLE AGRICULTURAL INTENSIFICATION.

FAO. (2009). GLOBAL AGRICULTURE TOWARDS 2050. HIGH LEVEL EXPERT FORUM-HOW TO FEED THE WORLD 2050, 1-4.

FAO. (2010). INVESTING IN FOOD SECURITY. FOOD SECURITY, 59(NOVEMBER), 2-20.

FAO. (2011). THE STATE OF THE WORLD'S LAND AND WATER RESOURCES FOR FOOD AND AGRICULTURE (SOLAW)—MANAGING SYSTEMS AT RISK. RETRIEVED FROM HTTP://WWW.FAO.ORG/DOCREP/017/I1688E/I1688E.PDF

FAO. (2011). THE STATE OF THE WORLD'S LAND AND WATER RESOURCES FOR FOOD AND AGRICULTURE. MANAGING SYSTEMS AT RISK (VOL. 2).

FAO. (2018). THE STATE OF FOOD SECURITY AND NUTRITION IN THE WORLD 2018. RETRIEVED FROM HTTP://WWW.FAO.ORG/STATE-OF-FOOD-SECURITY-NUTRITION/EN/.

FDI. (2011). THE FUTURE PROSPECTS FOR GLOBAL ARABLE LAND. RETRIEVED FROM HTTP://WWW.FUTUREDIRECTIONS.ORG.AU/PUBLICATION/THE-FUTURE-PROSPECTS-FOR-GLOBAL-ARABLE-LAND/

FEDER, G., & UMALI, D. L. (1993). THE ADOPTION OF AGRICULTURAL INNOVATIONS: A REVIEW. TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE, 43(3-4), 215-239.

FEDOROFF, N. V., BATTISTI, D. S., BEACHY, R. N., COOPER, P. J., FISCHHOFF, D. A., HODGES, C. N., . . . REYNOLDS, M. P. (2010). RADICALLY RETHINKING AGRICULTURE FOR THE 21ST CENTURY. SCIENCE (NEW YORK, N.Y.), 327(5967), 833-834.

FISCHER, G., HIZSNYIK, E., & WIBERG, D. (2011). SCARCITY AND ABUNDANCE OF LAND RESOURCES: COMPETING USES AND THE

SHRINKING LAND RESOURCE BASE. SOLAW BACKGROUND THEMATIC REPORT - TR02, 58.

FOLEY, J. A., RAMANKUTTY, N., BRAUMAN, K. A., CASSIDY, E. S., GERBER, J. S., JOHNSTON, M., . . . WEST, P. C. (2011). SOLUTIONS FOR A CULTIVATED PLANET. NATURE, 478(7369), 337.

GARNETT, T. (2016). PLATING UP SOLUTIONS. SCIENCE, 353(6305).

GARNETT, T., APPLEBY, M. C., BALMFORD, A., BATEMAN, I. J., BENTON, T. G., BLOOMER, P., . . . HERRERO, M. (2013). SUSTAINABLE INTENSIFICATION IN AGRICULTURE: PREMISES AND POLICIES. SCIENCE, 341(6141).

GOMIERO, T. (2016). SOIL DEGRADATION, LAND SCARCITY AND FOOD SECURITY: REVIEWING A COMPLEX CHALLENGE. SUSTAINABILITY, 8(3), 281.

GRIGGS, D., STAFFORD-SMITH, M., GAFFNEY, O., ROCKSTRÖM, J., ÖHMAN, M. C., SHYAMSUNDAR, P., . . . NOBLE, I. (2013). POLICY: SUSTAINABLE DEVELOPMENT GOALS FOR PEOPLE AND PLANET. NATURE, 495(7441), 305.

HANRAHAN, C. (2015). GLOBAL FOOD SECURITY BY THE NUMBERS. (SEPTEMBER), 1-20.

HEIKKINEN, N. (2014). BIRTH CONTROL COULD HELP THE ENVIRONMENT, BUT NOT QUICKLY.

HERRERO, M., & OTHERS. (2013). BIOMASS USE, PRODUCTION, FEED EFFICIENCIES, AND GREENHOUSE GAS EMISSIONS FROM GLOBAL LIVESTOCK SYSTEMS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 110(52), 20888-20893.

IRAC. (2011). PREVENTION AND MANAGEMENT OF INSECTICIDE RESISTANCE IN VECTORS OF PUBLIC HEALTH IMPORTANCE.

ITTERSUM, V., & OTHERS. (2013). YIELD GAP ANALYSIS WITH LOCAL TO GLOBAL RELEVANCE-A REVIEW. FIELD CROPS RESEARCH, 143, 4-17.

JALAVA, M., GUILLAUME, J. H., KUMMU, M., PORKKA, M., SIEBERT, S., & VARIS, O. (2016). DIET CHANGE AND FOOD LOSS REDUCTION: WHAT IS

THEIR COMBINED IMPACT ON GLOBAL WATER USE AND SCARCITY? EARTH'S FUTURE, 4(3), 62-78.

KEATING, B., & CARBERRY, P. (2010). SUSTAINABLE PRODUCTION, FOOD SECURITY AND SUPPLY CHAIN IMPLICATIONS.

KEATING, B., & OTHERS. (2014). FOOD WEDGES: FRAMING THE GLOBAL FOOD DEMAND AND SUPPLY CHALLENGE TOWARDS 2050. GLOBAL FOOD SECURITY, 3(3 – 4), 125 - 132.

KUMMU, M., FADER, M., GERTEN, D., GUILLAUME, J. H., JALAVA, M., JÄGERMEYR, J., . . . VARIS, O. (2017). BRINGING IT ALL TOGETHER: LINKING MEASURES TO SECURE NATIONS' FOOD SUPPLY. CURRENT OPINION IN ENVIRONMENTAL SUSTAINABILITY, 29, 98-117.

LABORDE, D., BIZIKOVA, L., LALLEMANT, T., & SMALLER, C. (2016). ENDING HUNGER: WHAT WOULD IT COST? : INTERNATIONAL INSTITUTE FOR SUSTAINABLE DEVELOPMENT.

LAGI, M., BERTRAND, K. Z., & BAR-YAM, Y. (2011). THE FOOD CRISES AND POLITICAL INSTABILITY IN NORTH AFRICA AND THE MIDDLE EAST. SSRN ELECTRONIC JOURNAL, 1-15.

MORGAN, R. P. C. (2009). SOIL EROSION AND CONSERVATION: JOHN WILEY & SONS.

MUELLER, N. D., GERBER, J. S., JOHNSTON, M., RAY, D. K., RAMANKUTTY, N., & FOLEY, J. A. (2012). CLOSING YIELD GAPS THROUGH NUTRIENT AND WATER MANAGEMENT. NATURE, 490(254).

NACHTERGAELE, F., PETRI, M., & BIANCALANI, R. (2011). LAND DEGRADATION. LAL, R. {\&} STEWART, BA, NIGER-DELTA DEVELOPMENT COMMISSION (1999): A BRIEF HISTORY.

NAPOLI, M., DE MURO, P., & MAZZIOTTA, M. (2011). TOWARDS A FOOD INSECURITY MULTIDIMENSIONAL INDEX (FIMI). MASTER IN HUMAN DEVELOPMENT AND FOOD SECURITY.

NELSON, G., & OTHERS. (2009). CLIMATE CHANGE: IMPACT ON AGRICULTURE AND COSTS OF ADAPTATION. 19.

NILSSON, M., GRIGGS, D., & VISBECK, M. (2016). POLICY: MAP THE INTERACTIONS BETWEEN SUSTAINABLE DEVELOPMENT GOALS. NATURE NEWS, 534(7607), 320.

OLSON, R. (2015). DIETARY GUIDELINES FOR AMERICANS. RETRIEVED FROM

PARDEY, P. G., BEDDOW, J. M., HURLEY, T. M., BEATTY, T. K. M., & EIDMAN, V. R. (2014). A BOUNDS ANALYSIS OF WORLD FOOD FUTURES: GLOBAL AGRICULTURE THROUGH TO 2050. AUSTRALIAN JOURNAL OF AGRICULTURAL AND RESOURCE ECONOMICS, 58(4), 571-589.

PONGIGLIONE, F. (2015). THE NEED FOR A PRIORITY STRUCTURE FOR THE SUSTAINABLE DEVELOPMENT GOALS. JOURNAL OF GLOBAL ETHICS, 11(1), 37-42.

POPKIN, B. M. (2001). THE NUTRITION TRANSITION AND OBESITY IN THE DEVELOPING WORLD. THE JOURNAL OF NUTRITION, 131(3), 871S-873S.

POPKIN, B. M., ADAIR, L. S., & NG, S. W. (2012). GLOBAL NUTRITION TRANSITION AND THE PANDEMIC OF OBESITY IN DEVELOPING COUNTRIES. NUTRITION REVIEWS, 70(1), 3-21.

PRADHAN, P., FISCHER, G., VAN VELTHUIZEN, H., REUSSER, D. E., & KROPP, J. P. (2015). CLOSING YIELD GAPS: HOW SUSTAINABLE CAN WE BE? PLOS ONE, 10(6), E0129487.

RACEY, M., O'BRIEN, C., DOUGLAS, S., MARQUEZ, O., HENDRIE, G., & NEWTON, G. (2016). SYSTEMATIC REVIEW OF SCHOOL-BASED INTERVENTIONS TO MODIFY DIETARY BEHAVIOR: DOES INTERVENTION INTENSITY IMPACT EFFECTIVENESS? JOURNAL OF SCHOOL HEALTH, 86(6), 452-463.

RAHNAMAYAN, S., & DIERAS, P. (2008, MAY). EFFICIENCY COMPETITION ON N-QUEEN PROBLEM: DE VS. CMA-ES. PAPER PRESENTED AT THE 2008 CANADIAN CONFERENCE ON ELECTRICAL AND COMPUTER ENGINEERING.

RANGANATHAN, J., VENNARD, D., WAITE, R., DUMAS, P., LIPINSKI, B., SEARCHINGER, T., & GLOBAGRI-WRR. (2016). SHIFTING DIETS FOR A SUSTAINABLE FOOD FUTURE. (APRIL), 90.

RAY, D. K., MUELLER, N. D., WEST, P. C., & FOLEY, J. A. (2013). YIELD TRENDS ARE INSUFFICIENT TO DOUBLE GLOBAL CROP PRODUCTION BY 2050. PLOS ONE, 8(6), 1-8.

REYNOSO-MEZA, G., SANCHIS, J., BLASCO, X., & GARCÍA-NIETO, S. (2014). PHYSICAL PROGRAMMING FOR PREFERENCE DRIVEN EVOLUTIONARY MULTI-OBJECTIVE OPTIMIZATION. APPLIED SOFT COMPUTING, 24, 341-362.

ROBERT, K. W., PARRIS, T. M., & LEISEROWITZ, A. A. (2005). WHAT IS SUSTAINABLE DEVELOPMENT? GOALS, INDICATORS, VALUES, AND PRACTICE. ENVIRONMENT: SCIENCE AND POLICY FOR SUSTAINABLE DEVELOPMENT, 47(3), 8-21.

ROCKSTRÖM, J., WILLIAMS, J., DAILY, G., NOBLE, A., MATTHEWS, N., GORDON, L., . . . STEDUTO, P. (2017). SUSTAINABLE INTENSIFICATION OF AGRICULTURE FOR HUMAN PROSPERITY AND GLOBAL SUSTAINABILITY. AMBIO, 46(1), 4-17.

RODRIGUEZ, D., DE VOIL, P., HUDSON, D., BROWN, J., HAYMAN, P., MARROU, H., & MEINKE, H. (2018). PREDICTING OPTIMUM CROP DESIGNS USING CROP MODELS AND SEASONAL CLIMATE FORECASTS. SCIENTIFIC REPORTS, 8(1), 2231.

ROSER, M., & RITCHIE, H. (2019). OUR WORLD IN DATA: FOOD PER PERSON. RETRIEVED FROM HTTPS://OURWORLDINDATA.ORG/FOOD-PER-PERSON

SADRAS, V. O., & OTHERS. (2015). YIELD GAP ANALYSIS OF FIELD CROPS, METHODS AND CASE STUDIES.

SANS, P., & COMBRIS, P. (2015). WORLD MEAT CONSUMPTION PATTERNS: AN OVERVIEW OF THE LAST FIFTY YEARS (1961–2011). MEAT SCIENCE, 109, 106-111.

SCHMIDT-TRAUB, G., & SHAH, A. (2015). INVESTMENT NEEDS TO ACHIEVE THE SUSTAINABLE DEVELOPMENT GOALS. PARIS AND NEW YORK: SUSTAINABLE DEVELOPMENT SOLUTIONS NETWORK.

SEN, A. K. (1999). DEVELOPMENT AS FREEDOM. 1-17.

STARBIRD, E., NORTON, M., & MARCUS, R. (2016). INVESTING IN FAMILY PLANNING: KEY TO ACHIEVING THE SUSTAINABLE DEVELOPMENT GOALS. GLOBAL HEALTH: SCIENCE AND PRACTICE, 4(2), 191-210.

STEFFEN, W., RICHARDSON, K., ROCKSTRÖM, J., CORNELL, S. E., FETZER, I., BENNETT, E. M., . . . DE WIT, C. A. (2015). PLANETARY BOUNDARIES: GUIDING HUMAN DEVELOPMENT ON A CHANGING PLANET. SCIENCE, 347(6223), 1259855.

STORN, R., & PRICE, K. (1997). DIFFERENTIAL EVOLUTION--A SIMPLE AND EFFICIENT HEURISTIC FOR GLOBAL OPTIMIZATION OVER CONTINUOUS SPACES. JOURNAL OF GLOBAL OPTIMIZATION, 11(4), 341-359.

TILMAN, D., BALZER, C., HILL, J., & BEFORT, B. L. (2011). GLOBAL FOOD DEMAND AND THE SUSTAINABLE INTENSIFICATION OF AGRICULTURE. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 108(50), 20260-20264.

TILMAN, D., CASSMAN, K. G., MATSON, P. A., NAYLOR, R., & POLASKY, S. (2002). AGRICULTURAL SUSTAINABILITY AND INTENSIVE PRODUCTION PRACTICES. NATURE, 418(6898), 671-677.

UN. (2015). WORLD POPULATION PROSPECTS KEY FINDINGS AND ADVANCE TABLES (9). RETRIEVED FROM

USDA PSD. (2017). TOTAL ACRES HARVESTED (WORLD), 13 PRIMARY CROPS. 1960- 2016 MARKETING YEARS. RETRIEVED FROM HTTPS://COMMUNITY.ENDNOTE.COM/T5/ENDNOTE-STYLES-FILTERS-AND/INSTITUTIONAL-AUTHOR-FORMATTING/TD-P/5962

VALIN, H., SANDS, R. D., VAN DER MENSBRUGGHE, D., NELSON, G. C., AHAMMAD, H., BLANC, E., . . . HEYHOE, E. (2014). THE FUTURE OF FOOD DEMAND: UNDERSTANDING DIFFERENCES IN GLOBAL ECONOMIC MODELS. AGRICULTURAL ECONOMICS, 45(1), 51-67.

VAN ITTERSUM, M. K., CASSMAN, K. G., GRASSINI, P., WOLF, J., TITTONELL, P., & HOCHMAN, Z. (2013). YIELD GAP ANALYSIS WITH LOCAL TO GLOBAL RELEVANCE—A REVIEW. FIELD CROPS RESEARCH, 143, 4-17.

VRANKEN, L., AVERMAETE, T., PETALIOS, D., & MATHIJS, E. (2014). CURBING GLOBAL MEAT CONSUMPTION: EMERGING EVIDENCE OF A SECOND NUTRITION TRANSITION. ENVIRONMENTAL SCIENCE & POLICY, 39, 95-106.

WALKER, P. (1983). CROP LOSSES: THE NEED TO QUANTIFY THE EFFECTS OF PESTS, DISEASES AND WEEDS ON AGRICULTURAL PRODUCTION. AGRICULTURE, ECOSYSTEMS AND ENVIRONMENT, 9(2), 119-158.

WICHELNS, D., & OTHERS. (2007). LOOKING AHEAD TO 2050: SCENARIOS OF ALTERNATIVE INVESTMENT APPROACHES. WATER FOR FOOD, WATER FOR LIFE: A COMPREHENSIVE ASSESSMENT OF WATER MANAGEMENT, 91-145.

XUE, F., SANDERSON, A. C., & GRAVES, R. J. (2003). PARETO-BASED MULTI-OBJECTIVE DIFFERENTIAL EVOLUTION. PAPER PRESENTED AT THE THE 2003 CONGRESS ON EVOLUTIONARY COMPUTATION, 2003. CEC'03.

CONCLUSIONS, FUTURE RESEARCH PLANS, AND POLICY IMPLICATIONS

6.1 Overarching Conclusions

It has long been recognized that the food system is a complex web of interacting actors operating at different scales, in different capacities, and with different goals. As such, systems theorists highlight the need to consider potential nonlinearities, adaptations, and unanticipated consequences as researchers and policymakers consider interventions designed to improve some aspect of the food system (Meadows 2008). While efforts have been made to develop frameworks describing connectivity patterns between different food systems actors (e.g. Sobal et al. 1998, Ericksen 2008), integrating data and academic findings across multiple connected sub-systems to project real-world outcomes can be difficult if not impossible without the use of computational tools. In short, modern problems require modern solutions.

This dissertation has presented three applications of complex systems modeling which inform the discussion surrounding critical food systems problem areas. Each chapter describes the development of a computational model toward the ultimate goal of creating a useful Digital Decision Support System (DDSS). Further research efforts will be required to reach that objective (discussed below), but the work that I have presented here serves as a foundation upon which myself and other researchers continue to build.

While the development of a practicable final product is an important goal, this dissertation also argues that the modeling process itself can be a valuable transdisciplinary research tool. Using a companion modeling approach, we show how key stakeholders and experts can be involved across all phases of model development, including scoping, parameterization and calibration, experimental design, interpretation of findings, and dissemination of results (Bousquet et al. 1999, Barreteau et al. 2003). In the case of RUSH-PNBM, our team convened periodic expert stakeholder panel sessions to guide and ground-truth our technical work. In the case of the food wedge investment optimization model—which is presently at a proof of concept stage—we are in the process of developing partnerships with stakeholder organizations and domain-level experts, with whom we will work as model development proceeds (discussed in the Directions for Future Research section below).

Overall, I believe there are clear synergies between the ongoing developments in complex systems modeling and the need for tools to identify novel solutions to and new ways of thinking about key food systems problems. The continuing integration of food systems and complex systems scholarship represents a valuable path forward, especially in light of our ever-increasing supply of data, and our expanding recognition that food systems research—like virtually all human pursuits—cannot be adequately studied without seriously considering the web of interconnections between diverse actors and processes. By thinking beyond traditional academic silos—and employing the latest transdisciplinary methods, tools, and technologies—food systems scholarship

249

will continue to produce the cutting edge, practically-motivated, participatory research that is so necessary to address society's need to find new "outside the box" solutions to the myriad wicked problems we will face in the decades to come.

In following sections, I discuss how the work I have presented here fits into a larger food systems context. First, I outline how RUSH-PNBM fits into other ongoing research efforts to better understand the leverage points to prevent catastrophic livestock disease outbreaks. I then discuss how researchers have continued to use and modify the model I originally developed to study a range of issues surrounding livestock biosecurity. I also acknowledge key limitations in both RUSH-PNBM and the food wedge optimization model, and describe future research plans which address current shortcomings. Finally, as the journal article format of the chapters only allowed for limited discussion of on-the-ground implications, I go into additional detail about how the work I've presented could inform the policy debate to bring about positive real-world change.

6.2 Taking a Systems Approach to Livestock Biosecurity

The Social Ecological Gaming and Simulation (SEGS) lab at UVM, with which I have worked closely in the development of RUSH-PNBM, is interested in applying systems thinking to complex issues that incorporate both a social-science and natural-science component. As part of the ADB-CAP project, our goal is to develop a practical DDSS that can be leveraged at multiple scales of action to reduce disease vulnerability in the livestock production industry.

250

We work across and incorporate three distinct levels of analysis: operational, tactical, and strategic (Figure 30). The operational level explores decision-making at the farm employee level, and largely concerns employees' willingness to put prescribed biosecurity protocols into action. The tactical level focuses on biosecurity investments and implementation of single-site biosecurity measures at the level of an individual farm or private production system manager. Finally, the strategic level focuses on macro-scale outcomes stemming from the complex interaction patterns between agents across the wider production chain network. While Chapters 1 and 2 of this dissertation focus primarily on the strategic level, my own research is part of a bigger project which takes a systems perspective, aiming to integrate all three levels of analysis.
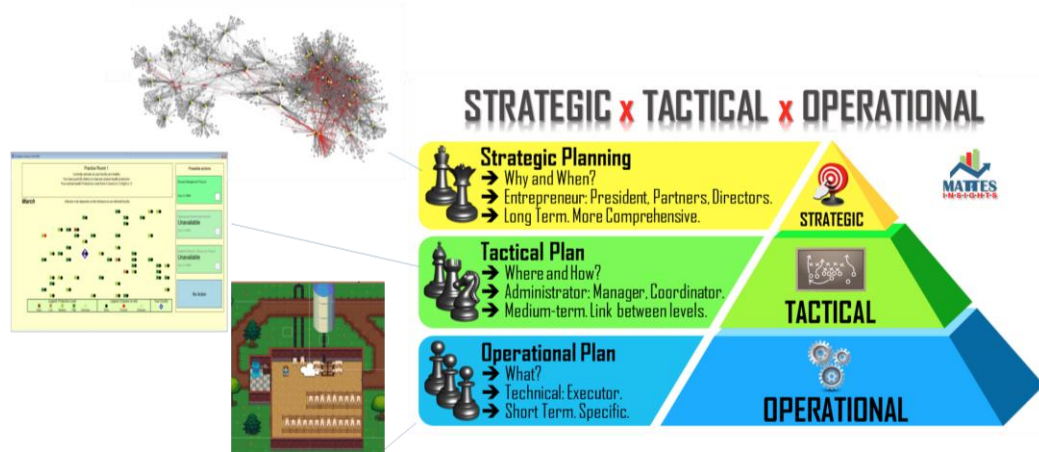
Mattes, J. (2015). Strategic Planning. Mattes Insights. Retrieved: 10/26/16
http://jonatasmattes.blogspot.com/2015/05/strategic-planning.html

Figure 30: Levels of risk management analysis and their application within the ADB-CAP project.

251

At the operational level, SEGS has developed an experimental biosecurity compliance game to determine how different methods of risk communication may impact a farm employee's choice to follow established protocols. We find that farm-level decisions matter, and that educating and incentivizing farm workers to follow biosecurity protocols is a major key to mitigating disease spread. Our initial results suggest that visual threat gauges are most effective at communicating the uncertainties associated with the disease threat environment, as opposed to traditional means such as verbal presentations or textual pamphlets (Merrill et al. 2019b).

At the tactical level, we have implemented another experimental game to understand the factors underlying a farm manager's decision to invest money toward biosecurity protocol adoption. Here we find that the availability of information about the prevalence, location, and peer-behavior related to a disease in the system can either stimulate or inhibit biosecurity adoption in complex ways. For example, a disease which is known to be creeping spatially closer to a decision-maker's premises tends to encourage adoption behavior, whereas awareness of peers' biosecurity adoption decisions is not correlated with increased adoption (Merrill et al. 2019a). Based on these game data, we identify three major clusters, differing based on risk preference. Findings from this experimental game have subsequently been used to inform agent decision heuristics in RUSH-PNBM (Bucini et al. 2019).

Finally, the strategic level is where RUSH-PNBM comes into play. The ABM methodology allows for the integration of insights from the operational and tactical

252

levels of analysis to inform a systems-level view. Using RUSH-PNBM, we have

illuminated the negative impact on disease resilience of the ongoing trend toward

producer specialization (Chapter 1). We have also explored the effects of network

structure and node positionality on epidemiological risk and proposed new methods to

measure and quantify an agent's vulnerability (Chapter 2). In the following sections, I

outline the ongoing work that is being conducted based upon the RUSH-PNBM

framework I originally developed, including incorporation of experimental game results

to inform agents' decision heuristics, limitations of our model, and directions for future

development.

6.3 The Continuing Legacy of RUSH-PNBM

My SEGS lab colleagues and I, working under the ADB-CAP project, have

leveraged my RUSH-PNBM model framework, adapting it to study additional

phenomena of interest related to biosecurity within livestock production systems.

Whereas RUSH-PNBM v.0.8 and v.1.2 focused primarily on the impact of network

structure on the disease resilience of livestock production systems, the model has been

expanded in several ways, according to the specific requirements of our research

questions, additional expert panelist feedback, deeper dives into the available statistical

data, and results from behavioral economics games.

Together with a cohort of agricultural economists, the development team has

added an economic sub-model to RUSH-PNBM. By tracking the sales of livestock—

along with standard prices for suckling pigs, weaners, and finishing hogs—we can

calculate each producer agent's cash-on-hand budget at any given time during a model run. The going rate for a pig at each life cycle stage is modified according to supply and demand in the system, since data show that diminishing supplies (e.g. during acute periods of disease outbreaks) does not generally not inhibit pork demand and will therefore put upward pressure on the price of livestock (Paarlberg 2014, Schulz and Tonsor 2015). The finances of each agent will be used to develop cost-loss functions to guide decision-making surrounding biosecurity protocol implementation, which may be as much as $2 in additional costs per pig (Dritz 2018).

In addition, the team has worked to add several realism-enhancing modifications shown empirically to impact disease spread. For example, farm visitation events—which have been identified as a weak point in producer biosecurity—open an additional possibility for an infection to be introduced (Lachappelle et al. 2017). An environmental avenue of infection is also modeled, representing the potential spread of PEDv through survival of the virus in the on-farm environment—e.g. in manure storage facilities (Goyal 2014, Tun et al. 2016, Murai et al. 2018)—or through exogenous mechanisms such as airborne transmission (Alonso et al. 2014).

Epidemiological data show that the 2013 – present U.S. PEDv epidemic is characterized by periodic peaks (in the winter) and valleys (in the summer) of newly-infected premises (AASV 2018). The frequency of these oscillations correlates with the impact of seasonal change on virus infectivity. These annual virulence cycles have

254

also been implemented in the current model version using a sinusoidal sweep of infection risk parameters throughout each model year.

The epidemiological data show that the peak level of infected premises topped out in 2013 and has continually decreased with each subsequent annual virulence cycle (AASV 2018). Since the disease's morbidity and mortality rates did not change over this period, these data suggest that farmers must have made behavioral changes which increased their biosecurity as the epidemic continued. To capture this dynamic, we have developed a suite of functions in RUSH-PNBM that modulate how agents' decision heuristics adapt to their surrounding risk environment.

Essentially, agents in the latest model version periodically choose whether or not to implement biosecurity protocols which decrease their chances of contracting the infection. Experimental gaming data collected by the SEGS lab through our protocol adoption game (Merrill et al. 2019a) show that, as infections on nearby farms increase, managers will become more likely to invest in biosecurity, even if it means taking a financial hit up front. However, the decision-making patterns of our experimental participants are complex. Factors such as risk attitude, availability of information, and psychological distancing were all found to play important roles.

Of primary importance in the current phase of RUSH-PNBM development is the incorporation of this nuanced game data into the decision heuristics of model agents. Our team has employed statistical clustering analysis to identify behavioral trends within the game data, finding three distinct clusters which differ according to

risk attitude. Risk-averse individuals tend to implement biosecurity protocols whenever even a few infections are present in the network. "Opportunists" don't invest in biosecurity when perceived risk is low but begin to invest more heavily when risk is high. Finally, risk-seekers gamble with biosecurity, generally investing very little, and only when the perceived epidemic environment is extremely dangerous.

To model information availability—a primary research question in our protocol adoption game—we implement a new agent type: veterinarians. Each vet has a service area and periodically surveys the farms in that area, recording the infection status of each. The proportion of farms that are currently infected is then reported to all farms in a vet's service area, serving as a proxy for the information availability criterion in the protocol adoption game. Each agent's probability of investing in additional biosecurity is governed by a logistic function based on the proportion of infected farms within its vet network, together with its risk attitude.

Psychological distancing is a phenomenon describing why—among other things—the longer in the past an event occurred, the less likely it is to be a salient decision factor (Yi et al. 2006). Our expert panelists believe this effect is significant when it comes to livestock biosecurity. Specifically, it has been observed that compliance with biosecurity protocols on the part of farm employees tends to fall off the longer it has been since the farm experienced an infection event. The latest version of RUSH-PNBM models this "potential biosecurity vs. actual biosecurity" by imposing a linear drop-off in compliance rate over time—which corresponds to increases in

256

infection probabilities—for as long as a producer has not experienced an infection event.

Using this new version of RUSH-PNBM, an experiment was conducted which systematically swept across treatments varying the relative proportion of risk-averse versus risk-taking producers (Bucini et al. 2019). Results show that the system's epidemiological resilience is sensitive to fairly-minor changes to this parameter. We find that there is a significant difference in median PEDv incidence between the "baseline" scenario runs and the "27.5% averse" scenario runs, which include just 10% more risk-averse producers than the baseline. We also find that increasing the number of risk-seeking producers in the system leads to more variability in epidemic severity between runs, supporting the percolation dynamics reported in Wiltshire (2018), and suggesting that the addition of just a few risk-seekers to a production system could hamper the efforts of policymakers who rely on disease forecasting to guide reactive decision-making in the face of an epidemic threat.

6.4 Limitations and Directions for Future Research

6.4.1 Current RUSH-PNBM Development Initiatives

RUSH-PNBM is still under active development, and as such there are a number of important features we intend to implement as the ADB-CAP project progresses to address limitations we have identified. Firstly, we will work to link the budgetary / economic sub-model with agents' biosecurity implementation decision-making, since,

in the real world, a hog producer could not implement a new biosecurity protocol without sufficient capital on hand, or at least sufficient liquidity to justify a bank loan.

Another limitation of the current model is that it does not consider hog or feed transfers extending beyond the boundaries of the U.S. state being modeled. With a great deal of livestock movement across state borders recorded every year, this likely represents a critical factor governing disease spread. We intend to implement a new "out of state" agent class, which will serve as both a source and sink for some subset of producers' livestock shipments. Incoming out-of-state shipments of infected animals or contaminated transportation equipment may then be evaluated as possible disease vectors.

In the longer term, the SEGS team will continue to hone the model to further enhance its context-specificity and realism. In the present version, the agents are adaptive, reacting to perceived disease risk within their local regions by potentially implementing new biosecurity precautions. However, they are still not "intelligent," i.e., they don't learn from the past and adjust their behavior accordingly. Instead, the risk preference of each agent is static throughout each model run. This represents a limitation of the model as it currently stands. A long-term goal of RUSH-PNBM development is to model adaptive behavior over time by imbuing each agent with a "memory" whereby previous events and their outcomes may impact their risk preference. Another means of developing agent intelligence we are pursuing is to use

experimental gaming data to train an AI algorithm, such as an artificial neural network, and use those results to inform model agents' decision heuristics.

Finally, we are also interested in risk factors stemming from the intentional release of an infective agent into the system. Bioterrorism of this type poses a serious threat to our critical food production infrastructure, and as such has received considerable attention in recent years (Madden et al. 2002, Murch 2003, Buehler et al. 2003, Parnell et al. 2010). We propose to evaluate attacker / defender dynamics using an experimental game. The data gleaned from this experiment can then be incorporated into a new version of RUSH-PNBM to evaluate the potential impact of attacks and policy responses on systemwide disease resilience.

6.4.2 Future RUSH-PNBM Research Goals

Being a food systems project, practical application of our model toward solving real-world problems remains a fundamental goal. We intend to develop RUSH-PNBM to the point that it can serve as an effective DDSS. To this end, we plan to build the model into a standalone application that can be easily utilized by industry stakeholders. This will involve development of a user interface which allows a general audience to interact with and extract value from the model in an intuitive manner.

We also plan to leverage RUSH-PNBM to project the systems-level impacts of managerial and employee decision-making across multiple levels of risk management, i.e. strategic, tactical, and operational. Of primary concern going forward is to develop a better understanding of the connections between human behavior and the resilience of

259

livestock production systems in the face of disease incursions. RUSH-PNBM will allow us to test multiple biosecurity scenarios, ranging from low adoption and compliance to full adoption and compliance. We will then be able to evaluate the resilience of our modeled livestock production systems—as indicated by the extent to which small perturbations (i.e. just a few more risk-seeking producers) may impact systems-level outcomes—as well as to suggest policy interventions that could improve disease resilience. Key questions include: how can we effectively nudge managers and workers toward more risk-averse behaviors? How does the means of information transmission (i.e. visual vs. verbal vs. textual) impinge upon biosecurity protocol compliance rates? How will changes in biosecurity protocol adoption and/or compliance affect key outcomes such as food availability or farm viability?

Another potential value proposition of RUSH-PNBM is its flexibility with regard to scale. While we have so far focused on entire U.S. states, we are currently exploring potential partnerships with several mid-scale private hog production chain enterprises to put our work to use "where the rubber meets the road." Despite encountering significant hurdles concerning data availability and quality, we are still committed to pursuing an industry partnership and developing a specifically-calibrated version of the model based on that partner's own network of operations, shipment patterns, existing biosecurity protocols, etc. Working together, we could then leverage RUSH-PNBM to identify risk points and develop effective biosecurity strategies which enhance the resilience of our partner's production chain to disease incursions.

A final research goal is to engage directly with stakeholders to gain feedback and put our results into action, bridging research outputs with practice and policy. We aim to hold a series of symposia with the theme of improving educational and risk communication messages and related policies designed to enhance adoption of and compliance with biosecurity practices in food animal production. Invited stakeholders will have an interest in promoting biosecurity to protect food animals from pest and disease threats. In these symposia, we will share our research and receive feedback on what we have accomplished, focusing on development of briefs and white papers to translate our results into more accessible formats that can be effectively utilized by those engaged in setting current policy and practice. We especially aim to attract attention for potential future collaboration, promote the development of systems-level tools, and secure funding to continue transdisciplinary work that integrates social and agricultural sciences.

6.4.3 Food Wedge Model Research Plan

Our food wedge investment optimization model represents significant progress toward the ultimate goal of building the wedge framework (Pacala and Socolow 2004, Keating et al. 2014) into a practicable DDSS. The model as currently parameterized proves out the technical methodology we have developed, representing the wedges as a system of equations, developing a fitness function to integrate the effects of multiple wedge interventions, and leveraging a multi-objective evolutionary-computational algorithm to optimize an investment schedule. However, significant work is still

261

required if the model is to become a practicable DDSS. Specifically, we need to transition to a stakeholder-driven, transdisciplinary research process by gathering a cohort of qualified individuals to weigh in on the project, including domain-level experts focused on each individual wedge, as well as the practitioners and policymakers who will ultimately use to model to inform real-world decision-making.

To this end, we are currently in the process of identifying and reaching out to stakeholders using a snowball sampling approach, with a food-policy advocacy organization—such as CGIAR's Research Program on Climate Change, Agriculture and Food Security (CCAFS)—acting as the central bridging organization. Our aim would then be to reformulate the wedge model around our partner organization's existing policy goals. This could involve shifting the scale of action: the model may be most useful at a smaller geographic scale, perhaps at a regional or country level, where specific investment decisions are being made, rather than a global level, where current investment efforts are often hindered by geopolitical challenges. We will also undertake an iterative process of wedge identification and parameterization which dovetails with existing policy proposals our partner organization already has on the table.

To accomplish the reparameterization, we have developed a survey instrument—which is presently undergoing institutional review board approval—to be issued to the expert domain-level stakeholders identified collaboratively alongside our partner organization. This survey asks participants to weigh in on the parameters listed

262

in Table 12, honing values according to their own areas of expertise. In this sense, we seek to obtain a larger range of potential impacts from a given wedge than currently utilized, as most of the wedge parameterization has been developed through single studies or estimates, often large scale and with high levels of uncertainty. We will then use a Delphi method to further modify parameter values through an iterative process asking participants to re-evaluate their original estimates based on feedback from their peers. Armed with a more realistic range of potential parameters for each wedge, future model iterations may provide a more reliable estimate of investment priorities. We believe this expert stakeholder engagement protocol to be the best way to ground the model in cross-disciplinary knowledge. With food systems being an explicitly transdisciplinary research approach devoted to real-world change, this next phase of the project is critical to its success, and we look forward to the opportunity to transition the project from theory to praxis.

6.5 Policy Implications by Chapter

6.5.1 Chapter One

The primary goal of Chapter One was to explore the effect of network structure on the disease resilience of livestock production chains. The impetus for the project stemmed from the realization that the current biosecurity intervention recommendations within livestock production systems are rooted in a paradigm that considers actions at the level of individual actors to be the basic unit of analysis. While many of these interventions—including lines of separation, truck wash infrastructure, shower-in-

shower-out facilities, all-in-all-out protocols, limited visitor access, etc.—have proven

reasonably effective in the fight against disease spread, we wondered whether a systems

perspective on this issue may cast light on different aspects of the problem, potentially

illuminating new, outside-the-box solutions. Specifically, we were interested in the

livestock industry trend toward utilizing multiple production sites, each specialized in a

particular life cycle stage of the animals.

The key policy takeaway from this work is informed by the multi-scale risk

management framework shown in Figure 30. Current thinking about disease

prevention in the industry tends to be focused on the operational and tactical levels of

analysis; i.e., what can a farm manager do to prevent disease incursions on his or her

farm, either through adoption of new biosecurity protocols or facilitation of employee

compliance with existing ones? While these measures can certainly be effective,

focusing solely on operational and tactical interventions ignores the broader strategic

picture, which stems from the complex interactions between production chain actors,

and thus demands some degree of collective action. While biosecurity measures at the

farm level may well diminish the disease transmission probability associated with a

given hog or feed shipment, simply avoiding that shipment in the first place—by

reconfiguring the network structure—likely represents a much more powerful leverage

point to bolster system resilience.

Of course, collective action problems are notoriously difficult, as achieving

consensus surrounding mutually-beneficial behavior changes can often result in

discrepancies between what is good for the individual versus what is good for the community. However, there are likely existing loci of control in livestock production systems which could be leveraged to effect the type of structural change that is called for. At the national level, USDA regulations may play a role, but achieving widespread consensus in our current antagonistic political environment may unfortunately render that leverage point relatively ineffective.

The level of individual private production chain management may be the lower-hanging fruit. It is in no one's best interest for a catastrophic disease to spread through a livestock production system, so the incentive to increase systemwide resilience is shared by capitalistic owners, government regulators, and animal product consumers alike. Hog production in the U.S. is currently something of an oligopoly, with just a few corporations dictating production practices across a majority of the industry. If managers of these major production chains were to realize the potential biosecurity advantages associated with a commitment to single-phase, farrow-to-finish production, it could have a huge impact on the disease resilience not just of their own enterprises, but more widely through regional and even national trade networks. While additional economic research is required, it is quite possible that reorganizing production network structures in this way—rather than investing huge sums into farm-level biosecurity measures—may well be both more efficacious and more cost-effective in the long term, representing a win-win across the board.

6.5.2 Chapter Two

Chapter Two focuses on the interaction between node positionality and network structure on the disease risks a given livestock production system actor may encounter during an epidemic event. We wish to reiterate that the specific meta-metrics we have identified to forecast vulnerability are not necessarily broadly applicable, being based on model simulations rather than any actual real-world production system. However, we hope that the methodology we have developed—given sufficient parameterization data and stakeholder input—may one day reach beyond the realm of computational science to inform on-the-ground policy.

Despite their limitations, several general policy takeaways may be drawn from our results. First and foremost, we find that epidemiological vulnerability is a function of both network structure and individual decision-making. Echoing implications from Chapter One, our findings suggest that biosecurity measures focused on individual actors—such as shower-in-shower-out facilities, lines of separation, or all-in-all-out protocols—are not the only way to approach livestock biosecurity.

From a tactical perspective, a producer may implement a new biosecurity measure to reduce his / her own vulnerability in the face of a perceived threat. From an operational perspective, employees may comply with biosecurity protocols if they receive effective messaging and are appropriately incentivized. These factors are currently being explored through ongoing research as part of the ADB-CAP project (Merrill et al. 2019a, 2019b, Bucini et al. 2019). However, from a strategic, systems-

266

level perspective, the identification of especially-vulnerable network loci would allow for a more intelligent and effective allocation of resources to curb epidemiological risk throughout the system. This latter factor is the goal of the meta-metric approach presented in Chapter Two.

Examining the meta-metric solutions in detail (Table 11), we find that the individual metrics included in the solutions have diverse implications. The suite of interventions that may be implemented at each level of risk analysis (i.e. strategic, tactical, and operational) each have their own data-gathering requirements. In light of this observation, we break the network metrics used in the analysis into three categories to interpret the GP solutions.

First is the set of *global* metrics—shown in red in Table 11—out of which weighted clustering coefficient was included in the all of the best-performing solutions. A higher clustering coefficient indicates that nodes tend to cohere into tightly-knit subgroups characterized by a relatively-high density of local ties, and a relative paucity of ties between clusters. In our solutions, clustering coefficient was included as either a denominator or a subtractive element in each function, suggesting that local clustering puts downward pressure on systemic vulnerability. This result is consistent with the observation from Chapter One (Figure 11) that networks with low specialization tend to include tight-knit communities of producers, feed mills, and slaughter plants; with few bridging links that may spread the infection to other areas of the network.

267

Next we examine the *node*-level metrics, which may themselves be broken down into two sub-classifications. First are the metrics that can only be known once a full graph representation is obtained, shown in blue in Table 11. While these metrics indicate a node's positionality within the overall network, a full codification of the network is required for calculation. This fact has implications for decision-makers intending to employ the meta-metric approach to curb vulnerability in a real-world system. Both *global* metrics and those *node*-level metrics that require a full graph codification necessitate a strategic level of risk analysis and action. Such analyses require a full dataset tracking animal movements through the system (Caporale et al., 2001), and as such, would likely require top-down action. In the meta-metrics we identify, the most common individual indicator from this category is coreness. Coreness is based on the *k*-core decomposition technique, which generates a series of maximal subgraphs in which each node has at least degree *k*. The coreness of each node equals *k* if it is a member of the *k*-core but not the (*k*+1)-core. For example, in plain terms, the two-core is the set of all nodes that have at least two connections to other nodes and are also connected to at least one of the other two-core subgraph members. Thus, a higher coreness indicates that a node is both heavily-connected, and also that it is connected to other equally-heavily-connected nodes. Coreness is a factor in the numerator of all six meta-metric solutions identified in Table 11, indicating that high-coreness nodes may be especially vulnerable to disease incursions.

The other metric in this category—which appears only in the numerator of solution six—is weighted Eigenvector centrality. Eigenvector centrality is based on the sum of the centrality values of all neighbors of a given node. Like coreness, it indicates inclusion within a subgroup of relatively-heavily-connected nodes, although for weighted Eigenvector centrality the connectivity criterion is the total number of livestock or feed transfers during each model run, whereas for coreness it is the number of trading partners. In either case, it would appear that the nodes that are members of a central, heavily-connected subgraph—as measured by either coreness or Eigenvector centrality—are more vulnerable to disease, all else being equal.

Finally, we examine the *node*-level metrics that can be calculated based only on information available to each individual agent (shown in green in Table 11). These metrics represent a tactical perspective, as they only require the internal records of an individual farm operator, and as such may be valuable in the prediction of disease vulnerability where a full network graph cannot be generated, which is unfortunately quite likely at present due to the aforementioned lack of transparency as regards livestock and feed movements (Caporale et al., 2001). Here we find that weighted and/or unweighted in-degree are included in the numerator of all six meta-metrics. This result is fairly trivial on the surface, since clearly more incoming deliveries—as well as more inbound trading partners—both open additional pathways through which disease may be spread to a given premises. However, despite its relative simplicity, the importance of these metrics cannot be ignored, with weighted in-degree explaining

269

30% of the variability in agent-level vulnerability across the board. Once again, this echoes findings from Chapter One, suggesting that increasing farms' connectivity can do nothing but decrease their epidemiological resilience, and that therefore every effort should be made to limit both the number of incoming shipments to each producer, and the number of trading partners with which it interacts.

6.5.3 Chapter Three

Chapter Three details a novel method for optimizing investment schedules into tough, dynamic, systems-level problems. Here we focus on global food supply and security, although we suggest that our methodology may find wide applications. To date, we have taken significant strides toward developing the food wedge framework of Keating (2014) into a DDSS that can be effectively employed to help stakeholders grapple with investment priorities into multiple, competing, interacting interventions.

As discussed in the Limitations and Directions for Future Research section above, the model currently stands as a proof of concept. It is not sufficiently-parameterized to conclude that the intervention investment schedules we have obtained should be employed to guide real-world decisions. However, we believe that the stakeholder-driven, transdisciplinary research protocol we have proposed can bring the model to a point where it may be effectively utilized by a partnering policy-driven organization or government actor. Despite its current limitations, we believe the methodology we have developed may have important policy implications, potentially serving as a valuable tool to help decision-makers parse the complexity of interventions

270

in complex problem domains. This is particularly true as there is increasing focus on achieving the Sustainable Development Goals (including Goal Two, which aims to curb global hunger and is most relevant to this work). However key barriers to success in SDG implementation still remain, concerning both the scale of implementation and the investments and funding mechanisms required (Nillson 2016).

While the ultimate goal is to develop a reliable DDSS to guide food wedge investments, the preliminary findings presented in Chapter Three cast light on the potential value proposition we can expect from a fully-parameterized model. For example, we find that the algorithm effectively weighs tradeoffs between investments in demand reduction, investments in supply increases, and the overall cost associated with an intervention schedule. Further, our results demonstrate the model's ability to deal with the temporal tradeoffs inherent in the problem, for example the decision to invest more up-front to decrease necessary expenditures in future years, or considerations which weigh interventions that can be expected to take a long time to come to fruition against those that yield quick results. We look forward to continuing this line of research, building upon the technical foundation we have created, and convening a cohort of expert stakeholders with whom we will work to help identify solutions to the looming global food security crisis.

6.6 References

AASV. (2018). PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV): WHAT'S NEW THIS WEEK? RETRIEVED FROM HTTPS://WWW.AASV.ORG/RESOURCES/PEDV/PEDVWHATSNEW.PHP

271

ALONSO, C., GOEDE, D. P., MORRISON, R. B., DAVIES, P. R., ROVIRA, A., MARTHALER, D. G., & TORREMORELL, M. (2014). EVIDENCE OF INFECTIVITY OF AIRBORNE PORCINE EPIDEMIC DIARRHEA VIRUS AND DETECTION OF AIRBORNE VIRAL RNA AT LONG DISTANCES FROM INFECTED HERDS. VETERINARY RESEARCH, 45(1), 73.

BARRETEAU, O., ANTONA, M., D'AQUINO, P., AUBERT, S., BOISSAU, S., BOUSQUET, F., . . . MATHEVET, R. (2003). OUR COMPANION MODELLING APPROACH. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 6(1).

BOUSQUET, F., BARRETEAU, O., LE PAGE, C., MULLON, C., & WEBER, J. (1999). AN ENVIRONMENTAL MODELLING APPROACH: THE USE OF MULTI-AGENT SIMULATIONS. ADVANCES IN ENVIRONMENTAL AND ECOLOGICAL MODELLING, 113, 122.

BUCINI, G., MERRILL, S. C., CLARK, E., MOEGENBURG, S. M., ZIA, A., KOLIBA, C. J., . . . SMITH, J. M. (2019). RISK ATTITUDES AFFECT LIVESTOCK BIOSECURITY DECISIONS WITH RAMIFICATIONS FOR DISEASE CONTROL IN A SIMULATED PRODUCTION SYSTEM. FRONTIERS IN VETERINARY SCIENCE (ACCEPTED PENDING REVISIONS).

BUEHLER, J. W., BERKELMAN, R. L., HARTLEY, D. M., & PETERS, C. J. (2003). SYNDROMIC SURVEILLANCE AND BIOTERRORISM-RELATED EPIDEMICS. EMERGING INFECTIOUS DISEASES, 9(10), 1197.

CAPORALE, V., GIOVANNINI, A., DI FRANCESCO, C., & CALISTRI, P. (2001). IMPORTANCE OF THE TRACEABILITY OF ANIMALS AND ANIMAL PRODUCTS IN EPIDEMIOLOGY. REVUE SCIENTIFIQUE ET TECHNIQUE-OFFICE INTERNATIONAL DES EPIZOOTIES, 20(2), 372-378.

DRITZ, S. (2018). [FINANCIAL IMPLICATIONS OF BIOSECURITY DECISION-MAKING IN THE U.S. HOG INDUSTRY (PERSONAL COMMUNICATION)].

ERICKSEN, P. J. (2008). CONCEPTUALIZING FOOD SYSTEMS FOR GLOBAL ENVIRONMENTAL CHANGE RESEARCH. GLOBAL ENVIRONMENTAL CHANGE, 18(1), 234-245.

GOYAL, S. (2014). PEDV RESEARCH UPDATES: ENVIRONMENTAL STABILITY OF PED (PORCINE EPIDEMIC DIARRHEA VIRUS). RETRIEVED FROM HTTPS://WWW.AASV.ORG/PEDV/RESEARCH/13_215.PDF

KEATING, B. A., HERRERO, M., CARBERRY, P. S., GARDNER, J., & COLE, M. B. (2014). FOOD WEDGES: FRAMING THE GLOBAL FOOD DEMAND AND SUPPLY CHALLENGE TOWARDS 2050. GLOBAL FOOD SECURITY, 3(3), 125-132.

LACHAPELLE, V., LETELLIER, A., FRAVALO, P., BRASSARD, J., & L'HOMME, Y. (2017). DYNAMICS OF VIRUS DISTRIBUTION IN A DEFINED SWINE PRODUCTION NETWORK USING ENTERIC VIRUSES AS MOLECULAR MARKERS. APPL. ENVIRON. MICROBIOL., 83(4), E03187-03116.

MEADOWS, D. H. (2008). THINKING IN SYSTEMS: A PRIMER: CHELSEA GREEN PUBLISHING.

MERRILL, S. C., KOLIBA, C. J., MOEGENBURG, S. M., ZIA, A., PARKER, J., SELLNOW, T., . . . SMITH, J. M. (2019A). DECISION-MAKING IN LIVESTOCK BIOSECURITY PRACTICES AMIDST ENVIRONMENTAL AND SOCIAL UNCERTAINTY: EVIDENCE FROM AN EXPERIMENTAL GAME. PLOS ONE (ARTICLE IN PRESS).

MERRILL, S. C., MOEGENBURG, S., KOLIBA, C. J., ZIA, A., TRINITY, L., CLARK, E., . . . SMITH, J. M. (2019B). WILLINGNESS TO COMPLY WITH BIOSECURITY IN LIVESTOCK FACILITIES: EVIDENCE FROM EXPERIMENTAL SIMULATIONS. FRONTIERS IN VETERINARY SCIENCE (ACCEPTED PENDING REVISIONS).

MURAI, K., MORIGUCHI, S., HAYAMA, Y., KOBAYASHI, S., MIYAZAKI, A., TSUTSUI, T., & YAMAMOTO, T. (2018). MATHEMATICAL MODELING OF PORCINE EPIDEMIC DIARRHEA VIRUS DYNAMICS WITHIN A FARROW-TO-FINISH SWINE FARM TO INVESTIGATE THE EFFECTS OF CONTROL MEASURES. PREVENTIVE VETERINARY MEDICINE, 149, 115-124.

MURCH, R. S. (2003). MICROBIAL FORENSICS: BUILDING A NATIONAL CAPACITY TO INVESTIGATE BIOTERRORISM. BIOSECURITY AND BIOTERRORISM: BIODEFENSE STRATEGY, PRACTICE, AND SCIENCE, 1(2), 117-122.

NILSSON, M., GRIGGS, D., & VISBECK, M. (2016). POLICY: MAP THE INTERACTIONS BETWEEN SUSTAINABLE DEVELOPMENT GOALS. NATURE NEWS, 534(7607), 320.

PAARLBERG, P. (2014). UPDATED ESTIMATED ECONOMIC WELFARE IMPACTS OF PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV). RETRIEVED FROM

PACALA, S., & SOCOLOW, R. (2004). STABILIZATION WEDGES: SOLVING THE CLIMATE PROBLEM FOR THE NEXT 50 YEARS WITH CURRENT TECHNOLOGIES. SCIENCE, 305(5686), 968-972.

PARNELL, G. S., SMITH, C. M., & MOXLEY, F. I. (2010). INTELLIGENT ADVERSARY RISK ANALYSIS: A BIOTERRORISM RISK MANAGEMENT MODEL. RISK ANALYSIS: AN INTERNATIONAL JOURNAL, 30(1), 32-48.

SCHULZ, L. L., & TONSOR, G. T. (2015). ASSESSMENT OF THE ECONOMIC IMPACTS OF PORCINE EPIDEMIC DIARRHEA VIRUS IN THE UNITED STATES. JOURNAL OF ANIMAL SCIENCE, 93(11), 5111-5118.

SOBAL, J., KHAN, L. K., & BISOGNI, C. (1998). A CONCEPTUAL MODEL OF THE FOOD AND NUTRITION SYSTEM. SOCIAL SCIENCE & MEDICINE, 47(7), 853-863.

TUN, H. M., CAI, Z., & KHAFIPOUR, E. (2016). MONITORING SURVIVABILITY AND INFECTIVITY OF PORCINE EPIDEMIC DIARRHEA VIRUS (PEDv) IN THE INFECTED ON-FARM EARTHEN MANURE STORAGES (EMS). FRONTIERS IN MICROBIOLOGY, 7, 265.

WHEELIS, M., CASAGRANDE, R., & MADDEN, L. V. (2002). BIOLOGICAL ATTACK ON AGRICULTURE: LOW-TECH, HIGH-IMPACT BIOTERRORISM: BECAUSE BIOTERRORIST ATTACK REQUIRES RELATIVELY LITTLE SPECIALIZED EXPERTISE AND TECHNOLOGY, IT IS A SERIOUS THREAT TO US AGRICULTURE AND CAN HAVE VERY LARGE ECONOMIC REPERCUSSIONS. BIOSCIENCE, 52(7), 569-576.

WILTSHIRE, S. W. (2018). USING AN AGENT-BASED MODEL TO EVALUATE THE EFFECT OF PRODUCER SPECIALIZATION ON THE EPIDEMIOLOGICAL RESILIENCE OF LIVESTOCK PRODUCTION NETWORKS. PLOS ONE, 13(3), E0194013.

COMPREHENSIVE LIST OF REFERENCES

AASV. (2015). PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV). RETRIEVED FROM HTTPS://WWW.AASV.ORG/AASV%20WEBSITE/RESOURCES/DISEASES/PORCINEEPIDEMICDIARRHEA.PHP:

AASV. (2018). PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV): WHAT'S NEW THIS WEEK? RETRIEVED FROM HTTPS://WWW.AASV.ORG/RESOURCES/PEDV/PEDVWHATSNEW.PHP

ADGER, W. N., EAKIN, H., & WINKELS, A. (2009). NESTED AND TELECONNECTED VULNERABILITIES TO ENVIRONMENTAL CHANGE. FRONTIERS IN ECOLOGY AND THE ENVIRONMENT, 7(3), 150-157. DOI:10.1890/070148

AJELLI, M., GONÇALVES, B., BALCAN, D., COLIZZA, V., HU, H., RAMASCO, J. J., . . . VESPIGNANI, A. (2010). COMPARING LARGE-SCALE COMPUTATIONAL APPROACHES TO EPIDEMIC MODELING: AGENT-BASED VERSUS STRUCTURED METAPOPULATION MODELS. BMC INFECTIOUS DISEASES, 10(1), 190.

ALAVI, M., & LEIDNER, D. E. (2001). KNOWLEDGE MANAGEMENT AND KNOWLEDGE MANAGEMENT SYSTEMS: CONCEPTUAL FOUNDATIONS AND RESEARCH ISSUES. MIS QUARTERLY, 107-136.

ALBERT, R., & BARABÁSI., A.-L. (2002). STATISTICAL MECHANICS OF COMPLEX NETWORKS. REVIEWS OF MODERN PHYSICS, 74(1), 47.

ALBERT, R., JEONG, H., & BARABÁSI, A.-L. (2000). ERROR AND ATTACK TOLERANCE OF COMPLEX NETWORKS. NATURE, 406(6794), 378-382.

ALDERTON, S., MACLEOD, E. T., ANDERSON, N. E., SCHATEN, K., KULESZO, J., SIMUUNZA, M., . . . ATKINSON, P. M. (2016). A MULTI-HOST AGENT-BASED MODEL FOR A ZOONOTIC, VECTOR-BORNE DISEASE: A CASE STUDY ON TRYPANOSOMIASIS IN EASTERN PROVINCE, ZAMBIA. PLOS NEGLECTED TROPICAL DISEASES, 10(12), E0005252.

ALEXANDRATOS, N., BRUINSMA, J., & OTHERS. (2012). WORLD AGRICULTURE TOWARDS 2030/2050: THE 2012 REVISION. RETRIEVED FROM

HTTP://LARGE.STANFORD.EDU/COURSES/2014/PH240/YUAN2/DOCS/AP10
6E.PDF

ALONSO, C., GOEDE, D. P., MORRISON, R. B., DAVIES, P. R., ROVIRA, A.,
MARTHALER, D. G., & TORREMORELL, M. (2014). EVIDENCE OF
INFECTIVITY OF AIRBORNE PORCINE EPIDEMIC DIARRHEA VIRUS AND
DETECTION OF AIRBORNE VIRAL RNA AT LONG DISTANCES FROM
INFECTED HERDS. VETERINARY RESEARCH, 45(1), 73.

AMASS, S. F., & CLARK, L. K. (1999). BIOSECURITY CONSIDERATIONS FOR
PORK PRODUCTION UNITS. JOURNAL OF SWINE HEALTH AND
PRODUCTION, 7(5), 217-228.

AMINE, E., BABA, N., BELHADJ, M., DEURENBERY-YAP, M., DJAZAYERY, A.,
FORRESTER, T., . . . MBUYAMBA, J. (2002). DIET, NUTRITION AND THE
PREVENTION OF CHRONIC DISEASES: REPORT OF A JOINT WHO/FAO
EXPERT CONSULTATION: WORLD HEALTH ORGANIZATION.

AN, L. (2012). MODELING HUMAN DECISIONS IN COUPLED HUMAN AND
NATURAL SYSTEMS: REVIEW OF AGENT-BASED MODELS. ECOLOGICAL
MODELLING, 229, 25-36.

ANDERSON, R. M., & MAY, R. M. (1979). POPULATION BIOLOGY OF
INFECTIOUS DISEASES: PART I. NATURE, 280(5721), 361.

ANDERSON, R. M., MAY, R. M., & ANDERSON, B. (1992). INFECTIOUS
DISEASES OF HUMANS: DYNAMICS AND CONTROL (VOL. 28): WILEY
ONLINE LIBRARY.

ARRUDA, A. G., FRIENDSHIP, R., CARPENTER, J., GREER, A., & POLJAK, Z.
(2016). EVALUATION OF CONTROL STRATEGIES FOR PORCINE
REPRODUCTIVE AND RESPIRATORY SYNDROME (PRRS) IN SWINE
BREEDING HERDS USING A DISCRETE EVENT AGENT-BASED MODEL. PLOS
ONE, 11(11), E0166596.

AUCHINCLOSS, A. H., & DIEZ ROUX, A. V. (2008). A NEW TOOL FOR
EPIDEMIOLOGY: THE USEFULNESS OF DYNAMIC-AGENT MODELS IN
UNDERSTANDING PLACE EFFECTS ON HEALTH. AMERICAN JOURNAL OF
EPIDEMIOLOGY, 168(1), 1-8.

AUGER, P., MAGAL, P., & RUAN, S. (2008). STRUCTURED POPULATION
MODELS IN BIOLOGY AND EPIDEMIOLOGY (VOL. 1936): SPRINGER.

AXELROD, R. M., AXELROD, R., & COHEN, M. D. (2000). HARNESSING COMPLEXITY: ORGANIZATIONAL IMPLICATIONS OF A SCIENTIFIC FRONTIER: BASIC BOOKS.

AZZARRI, C., BACOU, M., COX, C. M., GUO, Z., & KOO, J. (2016). SUBNATIONAL SOCIO-ECONOMIC DATASET AVAILABILITY. NATURE CLIM. CHANGE, 6(2), 115-116. DOI:10.1038/NCLIMATE2842

HTTP://WWW.NATURE.COM/NCLIMATE/JOURNAL/V6/N2/ABS/NCLIMATE2842.HTML#SUPPLEMENTARY-INFORMATION

BAERENKLAU, K. A. (2005). TOWARD AN UNDERSTANDING OF TECHNOLOGY ADOPTION: RISK, LEARNING, AND NEIGHBORHOOD EFFECTS. LAND ECONOMICS, 81(1), 1-19.

BAGNI, R., BERCHI, R., & CARIELLO, P. (2002). A COMPARISON OF SIMULATION MODELS APPLIED TO EPIDEMICS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 5(3).

BAILEY, N. T. J. (1975). THE MATHEMATICAL THEORY OF INFECTIOUS DISEASES AND ITS APPLICATIONS: CHARLES GRIFFIN \& COMPANY LTD, 5A CRENDON STREET, HIGH WYCOMBE, BUCKS HP13 6LE.

BAJARDI, P., POLETTO, C., RAMASCO, J. J., TIZZONI, M., COLIZZA, V., & VESPIGNANI, A. (2011). HUMAN MOBILITY NETWORKS, TRAVEL RESTRICTIONS, AND THE GLOBAL SPREAD OF 2009 H1N1 PANDEMIC. PLOS ONE, 6(1), E16591.

BAKKER, K. (2012). WATER SECURITY: RESEARCH CHALLENGES AND OPPORTUNITIES. SCIENCE, 337(6097), 914-915. DOI:10.1126/SCIENCE.1226337

BALL, P. (2004). CRITICAL MASS: HOW ONE THING LEADS TO ANOTHER: MACMILLAN.

BARABÁSI, A.-L., & ALBERT, R. (1999). EMERGENCE OF SCALING IN RANDOM NETWORKS. SCIENCE, 286(5439), 509-512.

BARBIER, E. B. (2004). EXPLAINING AGRICULTURAL LAND EXPANSION AND DEFORESTATION IN DEVELOPING COUNTRIES. AMERICAN JOURNAL OF AGRICULTURAL ECONOMICS, 86(5), 1347 - 1353.

BARRETEAU, O., ANTONA, M., D'AQUINO, P., AUBERT, S., BOISSAU, S., BOUSQUET, F., . . . MATHEVET, R. (2003). OUR COMPANION MODELLING APPROACH. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 6(1).

BARRETT, C. B. (2013). FOOD SECURITY AND SOCIOPOLITICAL STABILITY: OUP OXFORD.

BARRETT, C. L., BISSET, K. R., EUBANK, S. G., FENG, X., & MARATHE, M. V. (2008). EPISIMDEMICS: AN EFFICIENT ALGORITHM FOR SIMULATING THE SPREAD OF INFECTIOUS DISEASE OVER LARGE REALISTIC SOCIAL NETWORKS. PAPER PRESENTED AT THE HIGH PERFORMANCE COMPUTING, NETWORKING, STORAGE AND ANALYSIS.

BARTHÉLEMY, M., BARRAT, A., PASTOR-SATORRAS, R., & VESPIGNANI, A. (2004). VELOCITY AND HIERARCHICAL SPREAD OF EPIDEMIC OUTBREAKS IN SCALE-FREE NETWORKS. PHYSICAL REVIEW LETTERS, 92(17), 178701.

BATAGELJ, V., & ZAVERSNIK, M. (2003). AN O (M) ALGORITHM FOR CORES DECOMPOSITION OF NETWORKS. ARXIV PREPRINT CS/0310049.

BECK, U. (2010). REMAPPING SOCIAL INEQUALITIES IN AN AGE OF CLIMATE CHANGE: FOR A COSMOPOLITAN RENEWAL OF SOCIOLOGY*. GLOBAL NETWORKS, 10(2), 165-181. DOI:10.1111/J.1471-0374.2010.00281.X

BECK, U., & GRANDE, E. (2010). VARIETIES OF SECOND MODERNITY: THE COSMOPOLITAN TURN IN SOCIAL AND POLITICAL THEORY AND RESEARCH. THE BRITISH JOURNAL OF SOCIOLOGY, 61(3), 409-443. DOI:10.1111/J.1468-4446.2010.01320.X

BECKER, A., INOUE, S., FISCHER, M., & SCHWEGLER, B. (2012). CLIMATE CHANGE IMPACTS ON INTERNATIONAL SEAPORTS: KNOWLEDGE, PERCEPTIONS, AND PLANNING EFFORTS AMONG PORT ADMINISTRATORS. CLIMATIC CHANGE, 110(1), 5-29.

BELKHIRIA, J., ALKHAMIS, M. A., & MARTÍNEZ-LÓPEZ, B. (2016). APPLICATION OF SPECIES DISTRIBUTION MODELING FOR AVIAN INFLUENZA SURVEILLANCE IN THE UNITED STATES CONSIDERING THE NORTH AMERICA MIGRATORY FLYWAYS. SCIENTIFIC REPORTS, 6, 33161.

BELL, D. C., ATKINSON, J. S., & CARLSON, J. W. (1999). CENTRALITY MEASURES FOR DISEASE TRANSMISSION NETWORKS. SOCIAL NETWORKS, 21(1), 1-21.

BIGRAS-POULIN, M., BARFOD, K., MORTENSEN, S., & GREINER, M. (2007). RELATIONSHIP OF TRADE PATTERNS OF THE DANISH SWINE INDUSTRY ANIMAL MOVEMENTS NETWORK TO POTENTIAL DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE, 80(2), 143-165.

BIGRAS-POULIN, M., THOMPSON, R. A., CHRIÉL, M., MORTENSEN, S., & GREINER, M. (2006). NETWORK ANALYSIS OF DANISH CATTLE INDUSTRY TRADE PATTERNS AS AN EVALUATION OF RISK POTENTIAL FOR DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE, 76(1), 11-39.

BINDRABAN, P. S., VAN DER VELDE, M., YE, L., VAN DEN BERG, M., MATERECHERA, S., KIBA, D. I., . . . HOOGMOED, M. (2012). ASSESSING THE IMPACT OF SOIL DEGRADATION ON FOOD PRODUCTION. CURRENT OPINION IN ENVIRONMENTAL SUSTAINABILITY, 4(5), 478-488.

BISBIS, M. B., GRUDA, N., & BLANKE, M. (2018). POTENTIAL IMPACTS OF CLIMATE CHANGE ON VEGETABLE PRODUCTION AND PRODUCT QUALITY–A REVIEW. JOURNAL OF CLEANER PRODUCTION, 170, 1602-1620.

BOCCARA, N. (2010). MODELING COMPLEX SYSTEMS: SPRINGER SCIENCE & BUSINESS MEDIA.

BONACICH, P. (1987). POWER AND CENTRALITY: A FAMILY OF MEASURES. AMERICAN JOURNAL OF SOCIOLOGY, 92(5), 1170-1182.

BORLAUG, N. E. (2008). FEEDING A WORLD OF 10 BILLION PEOPLE: OUR 21ST CENTURY CHALLENGE PERSPECTIVES IN WORLD FOOD AND AGRICULTURE 2004 (PP. 31-56). AMES, IOWA, USA: IOWA STATE PRESS.

BORN, B., & PURCELL, M. (2006). AVOIDING THE LOCAL TRAP: SCALE AND FOOD SYSTEMS IN PLANNING RESEARCH. JOURNAL OF PLANNING EDUCATION AND RESEARCH, 26(2), 195-207.

BOUSQUET, F., BARRETEAU, O., LE PAGE, C., MULLON, C., & WEBER, J. (1999). AN ENVIRONMENTAL MODELLING APPROACH: THE USE OF

MULTI-AGENT SIMULATIONS. ADVANCES IN ENVIRONMENTAL AND ECOLOGICAL MODELLING, 113, 122.

BOWMAN, A. S., KROGWOLD, R. A., PRICE, T., DAVIS, M., & MOELLER, S. J. (2015). INVESTIGATING THE INTRODUCTION OF PORCINE EPIDEMIC DIARRHEA VIRUS INTO AN OHIO SWINE OPERATION. BMC VETERINARY RESEARCH, 11(1), 38.

BRADHURST, R. A., ROCHE, S. E., EAST, I. J., KWAN, P., & GARNER, M. G. (2016). IMPROVING THE COMPUTATIONAL EFFICIENCY OF AN AGENT-BASED SPATIOTEMPORAL MODEL OF LIVESTOCK DISEASE SPREAD AND CONTROL. ENVIRONMENTAL MODELLING \& SOFTWARE, 77, 1-12.

BRADHURST, R. A., ROCHE, S. E., GARNER, M. G., SAJEEV, A. S. M., & KWAN, P. (2013). MODELLING THE SPREAD OF LIVESTOCK DISEASE ON A NATIONAL SCALE: THE CASE FOR A HYBRID APPROACH. PAPER PRESENTED AT THE PROCEEDINGS OF THE 20TH INTERNATIONAL CONGRESS ON MODELLING AND SIMULATION (MODSIM2013), MODELLING AND SIMULATION SOCIETY OF AUSTRALIA AND NEW ZEALAND.

BRANDES, U., & FLEISCHER, D. (2005). CENTRALITY MEASURES BASED ON CURRENT FLOW. PAPER PRESENTED AT THE ANNUAL SYMPOSIUM ON THEORETICAL ASPECTS OF COMPUTER SCIENCE.

BRANKE, J., DEB, K., DIEROLF, H., & OSSWALD, M. (2004). FINDING KNEES IN MULTI-OBJECTIVE OPTIMIZATION. PAPER PRESENTED AT THE INTERNATIONAL CONFERENCE ON PARALLEL PROBLEM SOLVING FROM NATURE.

BROWN, P. M., & CAMERON, L. D. (2000). WHAT CAN BE DONE TO REDUCE OVERCONSUMPTION? ECOLOGICAL ECONOMICS, 32(1), 27-41.

BRYCESON, K. P., & SMITH, C. S. (2008). ABSTRACTION AND MODELLING OF AGRI-FOOD CHAINS AS COMPLEX DECISION MAKING SYSTEMS. RETRIEVED FROM

BUCINI, G., MERRILL, S. C., CLARK, E., MOEGENBURG, S. M., ZIA, A., KOLIBA, C. J., . . . SMITH, J. M. (2019). RISK ATTITUDES AFFECT LIVESTOCK BIOSECURITY DECISIONS WITH RAMIFICATIONS FOR DISEASE

CONTROL IN A SIMULATED PRODUCTION SYSTEM. FRONTIERS IN VETERINARY SCIENCE (ACCEPTED PENDING REVISIONS).

BUEHLER, J. W., BERKELMAN, R. L., HARTLEY, D. M., & PETERS, C. J. (2003). SYNDROMIC SURVEILLANCE AND BIOTERRORISM-RELATED EPIDEMICS. EMERGING INFECTIOUS DISEASES, 9(10), 1197.

BURDETT, C. L., KRAUS, B. R., GARZA, S. J., MILLER, R. S., & BJORK, K. E. (2015). SIMULATING THE DISTRIBUTION OF INDIVIDUAL LIVESTOCK FARMS AND THEIR POPULATIONS IN THE UNITED STATES: AN EXAMPLE USING DOMESTIC SWINE (SUS SCROFA DOMESTICUS) FARMS. PLOS ONE, 10(11), E0140338.

BURKE, M., & EMERICK, K. (2016). ADAPTATION TO CLIMATE CHANGE: EVIDENCE FROM US AGRICULTURE. AMERICAN ECONOMIC JOURNAL: ECONOMIC POLICY, 8(3), 106-140.

CAPORALE, V., GIOVANNINI, A., DI FRANCESCO, C., & CALISTRI, P. (2001). IMPORTANCE OF THE TRACEABILITY OF ANIMALS AND ANIMAL PRODUCTS IN EPIDEMIOLOGY. REVUE SCIENTIFIQUE ET TECHNIQUE-OFFICE INTERNATIONAL DES EPIZOOTIES, 20(2), 372-378.

CARLSON, J. M., & DOYLE, J. (2000). HIGHLY OPTIMIZED TOLERANCE: ROBUSTNESS AND DESIGN IN COMPLEX SYSTEMS. PHYSICAL REVIEW LETTERS, 84(11), 2529.

CARLSON, J. M., & DOYLE, J. (2002). COMPLEXITY AND ROBUSTNESS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 99(SUPPL 1), 2538-2545.

CASSIDY, E. S., WEST, P. C., GERBER, J. S., & FOLEY, J. A. (2013). REDEFINING AGRICULTURAL YIELDS: FROM TONNES TO PEOPLE NOURISHED PER HECTARE. ENVIRONMENTAL RESEARCH LETTERS, 8(3), 034015.

CGIAR. (2014). CLIMATE CHANGE, AGRICULTURE AND FOOD SECURITY (CCAFS) DATASETS. RETRIEVED FROM: HTTPS://DATAVERSE.HARVARD.EDU/DATAVERSE/CCAFSBASELINE

CHALLINOR, A. J., ADGER, W. N., & BENTON, T. G. (2017). CLIMATE RISKS ACROSS BORDERS AND SCALES. NATURE CLIM. CHANGE, 7(9), 621-623. DOI:10.1038/NCLIMATE3380

CHALLINOR, A. J., KOEHLER, A. K., RAMIREZ-VILLEGAS, J., WHITFIELD, S., & DAS, B. (2016). CURRENT WARMING WILL REDUCE YIELDS UNLESS MAIZE BREEDING AND SEED SYSTEMS ADAPT IMMEDIATELY. NATURE CLIM. CHANGE, 6(10), 954-958. DOI:10.1038/NCLIMATE3061

HTTP://WWW.NATURE.COM/NCLIMATE/JOURNAL/V6/N10/ABS/NCLIMATE3061.HTML#SUPPLEMENTARY-INFORMATION

CHICHAKLY, K. J., & EPPSTEIN, M. J. (2013). DISCOVERING DESIGN PRINCIPLES FROM DOMINATED SOLUTIONS. IEEE ACCESS, 1, 275-289.

CHRISTLEY, R. M., PINCHBECK, G. L., BOWERS, R. G., CLANCY, D., FRENCH, N. P., BENNETT, R., & TURNER, J. (2005). INFECTION IN SOCIAL NETWORKS: USING NETWORK ANALYSIS TO IDENTIFY HIGH-RISK INDIVIDUALS. AMERICAN JOURNAL OF EPIDEMIOLOGY, 162(10), 1024-1031.

CHRISTLEY, R. M., ROBINSON, S. E., LYSONS, R., & FRENCH, N. P. (2005). NETWORK ANALYSIS OF CATTLE MOVEMENT IN GREAT BRITAIN. PROC. SOC. VET. EPIDEMIOL. PREV. MED, 234-243.

CHURCHMAN, C. W. (1967). WICKED PROBLEMS [GUEST EDITORIAL]. MANAGEMENT SCIENCE, 14(4), B141-B142.

CLIFF, A. D. (1981). SPATIAL DIFFUSION: AN HISTORICAL GEOGRAPHY OF EPIDEMICS IN AN ISLAND COMMUNITY: CUP ARCHIVE.

COBURN, B. J., WAGNER, B. G., & BLOWER, S. (2009). MODELING INFLUENZA EPIDEMICS AND PANDEMICS: INSIGHTS INTO THE FUTURE OF SWINE FLU (H1N1). BMC MEDICINE, 7(1), 30.

COELHO, F. C., CRUZ, O. G., & CODEÇO, C. T. (2008). EPIGRASS: A TOOL TO STUDY DISEASE SPREAD IN COMPLEX NETWORKS. SOURCE CODE FOR BIOLOGY AND MEDICINE, 3(1), 3. DOI:10.1186/1751-0473-3-3

COHEN, R., EREZ, K., BEN-AVRAHAM, D., & HAVLIN, S. (2000). RESILIENCE OF THE INTERNET TO RANDOM BREAKDOWNS. PHYSICAL REVIEW LETTERS, 85(21), 4626.

COLE, M. B., AUGUSTIN, M. A., ROBERTSON, M. J., & MANNERS, J. M. (2018). THE SCIENCE OF FOOD SECURITY. NPJ SCIENCE OF FOOD, 2(1), 14.

COLIZZA, V., BARRAT, A., BARTHÉLEMY, M., & VESPIGNANI, A. (2006). THE ROLE OF THE AIRLINE TRANSPORTATION NETWORK IN THE PREDICTION AND PREDICTABILITY OF GLOBAL EPIDEMICS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 103(7), 2015-2020.

COLIZZA, V., BARRAT, A., BARTHÉLEMY, M., & VESPIGNANI, A. (2007). PREDICTABILITY AND EPIDEMIC PATHWAYS IN GLOBAL OUTBREAKS OF INFECTIOUS DISEASES: THE SARS CASE STUDY. BMC MEDICINE, 5(1), 34.

CORTEZ, M. H., & WEITZ, J. S. (2013). DISTINGUISHING BETWEEN INDIRECT AND DIRECT MODES OF TRANSMISSION USING EPIDEMIOLOGICAL TIME SERIES. THE AMERICAN NATURALIST, 181(2), E43-E52.

DAS, S., & SUGANTHAN, P. (2011). DIFFERENTIAL EVOLUTION: A SURVEY OF THE STATE-OF-THE-ART. IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, 15(1), 4-31.

DE, P., SINGH, A. E., WONG, T., YACOUB, W., & JOLLY, A. M. (2004). SEXUAL NETWORK ANALYSIS OF A GONORRHOEA OUTBREAK. SEXUALLY TRANSMITTED INFECTIONS, 80(4), 280-285.

DEB, K., PRATAP, A., AGARWAL, S., & MEYARIVAN, T. (2002). A FAST AND ELITIST MULTIOBJECTIVE GENETIC ALGORITHM: NSGA-II. IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, 6(2), 182-197.

DELUCIA, E. H., GOMEZ-CASANOVAS, N., GREENBERG, J. A., HUDIBURG, T. W., KANTOLA, I. B., LONG, S. P., . . . PARTON, W. J. (2014). THE THEORETICAL LIMIT TO PLANT PRODUCTIVITY. ENVIRONMENTAL SCIENCE AND TECHNOLOGY, 48(16), 9471-9477.

DEPPISCH, S., & HASIBOVIC, S. (2013). SOCIAL-ECOLOGICAL RESILIENCE THINKING AS A BRIDGING CONCEPT IN TRANSDISCIPLINARY RESEARCH ON CLIMATE-CHANGE ADAPTATION. NATURAL HAZARDS, 67(1), 117-127.

DRITZ, S. (2018). [FINANCIAL IMPLICATIONS OF BIOSECURITY DECISION-MAKING IN THE U.S. HOG INDUSTRY (PERSONAL COMMUNICATION)].

DUBÉ, C., RIBBLE, C., KELTON, D., & MCNAB, B. (2009). A REVIEW OF NETWORK ANALYSIS TERMINOLOGY AND ITS APPLICATION TO FOOT-

AND-MOUTH DISEASE MODELLING AND POLICY DEVELOPMENT. TRANSBOUNDARY AND EMERGING DISEASES, 56(3), 73-85.

DUBOIS, O. (2011). THE STATE OF THE WORLD'S LAND AND WATER RESOURCES FOR FOOD AND AGRICULTURE: MANAGING SYSTEMS AT RISK: EARTHSCAN.

ELLISON, G., & FUDENBERG, D. (1993). RULES OF THUMB FOR SOCIAL LEARNING. POLITICAL ECONOMY, 101(4), 612-643.

EL-SAYED, A. M., SCARBOROUGH, P., SEEMANN, L., & GALEA, S. (2012). SOCIAL NETWORK ANALYSIS AND AGENT-BASED MODELING IN SOCIAL EPIDEMIOLOGY. EPIDEMIOLOGIC PERSPECTIVES \& INNOVATIONS, 9(1), 1.

EPSTEIN, J. M. (2008). WHY MODEL? JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 11(4), 12.

ERICKSEN, P. J. (2008). CONCEPTUALIZING FOOD SYSTEMS FOR GLOBAL ENVIRONMENTAL CHANGE RESEARCH. GLOBAL ENVIRONMENTAL CHANGE, 18(1), 234-245.

ESWARAN, H., LAL, R., & REICH, P. (2001). LAND DEGRADATION: AN OVERVIEW. RESPONSE TO LAND DEGRADATION, 20-35.

EUBANK, S., GUCLU, H., KUMAR, V. S. A., & MARATHE, M. V. (2004). MODELLING DISEASE OUTBREAKS IN REALISTIC URBAN SOCIAL NETWORKS. NATURE, 429(6988), 180.

FAO. (2004). THE ETHICS OF SUSTAINABLE AGRICULTURAL INTENSIFICATION.

FAO. (2009). GLOBAL AGRICULTURE TOWARDS 2050. HIGH LEVEL EXPERT FORUM-HOW TO FEED THE WORLD 2050, 1-4.

FAO. (2010). INVESTING IN FOOD SECURITY. FOOD SECURITY, 59(NOVEMBER), 2-20.

FAO. (2011). THE STATE OF THE WORLD'S LAND AND WATER RESOURCES FOR FOOD AND AGRICULTURE (SOLAW)—MANAGING SYSTEMS AT RISK. RETRIEVED FROM HTTP://WWW.FAO.ORG/DOCREP/017/I1688E/I1688E.PDF

FAO. (2011). THE STATE OF THE WORLD'S LAND AND WATER RESOURCES FOR FOOD AND AGRICULTURE. MANAGING SYSTEMS AT RISK (VOL. 2).

FAO. (2018). THE STATE OF FOOD SECURITY AND NUTRITION IN THE WORLD 2018. RETRIEVED FROM HTTP://WWW.FAO.ORG/STATE-OF-FOOD-SECURITY-NUTRITION/EN/.

FDI. (2011). THE FUTURE PROSPECTS FOR GLOBAL ARABLE LAND. RETRIEVED FROM HTTP://WWW.FUTUREDIRECTIONS.ORG.AU/PUBLICATION/THE-FUTURE-PROSPECTS-FOR-GLOBAL-ARABLE-LAND/

FEDER, G., & UMALI, D. L. (1993). THE ADOPTION OF AGRICULTURAL INNOVATIONS: A REVIEW. TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE, 43(3-4), 215-239.

FEDOROFF, N. V., BATTISTI, D. S., BEACHY, R. N., COOPER, P. J., FISCHHOFF, D. A., HODGES, C. N., . . . REYNOLDS, M. P. (2010). RADICALLY RETHINKING AGRICULTURE FOR THE 21ST CENTURY. SCIENCE (NEW YORK, N.Y.), 327(5967), 833-834.

FENICHEL, E. P., CASTILLO-CHAVEZ, C., CEDDIA, M. G., CHOWELL, G., PARRA, P. A. G., HICKLING, G. J., . . . VILLALOBOS, C. (2011). ADAPTIVE HUMAN BEHAVIOR IN EPIDEMIOLOGICAL MODELS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 108(15), 6306-6311.

FÈVRE, E. M., BRONSVOORT, B. M. D. C., HAMILTON, K. A., & CLEAVELAND, S. (2006). ANIMAL MOVEMENTS AND THE SPREAD OF INFECTIOUS DISEASES. TRENDS IN MICROBIOLOGY, 14(3), 125-131.

FIRESTONE, S. M., WARD, M. P., CHRISTLEY, R. M., & DHAND, N. K. (2011). THE IMPORTANCE OF LOCATION IN CONTACT NETWORKS: DESCRIBING EARLY EPIDEMIC SPREAD USING SPATIAL SOCIAL NETWORK ANALYSIS. PREVENTIVE VETERINARY MEDICINE, 102(3), 185-195.

FISCHER, A., ARNOLD, A., & GIBBS, M. (1996). INFORMATION AND THE SPEED OF INNOVATION ADOPTION. AMERICAN JOURNAL OF AGRICULTURAL ECONOMICS, 78(4), 1073-1081.

FISCHER, G., HIZSNYIK, E., & WIBERG, D. (2011). SCARCITY AND ABUNDANCE OF LAND RESOURCES: COMPETING USES AND THE SHRINKING LAND RESOURCE BASE. SOLAW BACKGROUND THEMATIC REPORT - TR02, 58.

FOLEY, J. A., RAMANKUTTY, N., BRAUMAN, K. A., CASSIDY, E. S., GERBER, J. S., JOHNSTON, M., . . . WEST, P. C. (2011). SOLUTIONS FOR A CULTIVATED PLANET. NATURE, 478(7369), 337.

FORAN, T. (2015). NODE AND REGIME: INTERDISCIPLINARY ANALYSIS OF WATER-ENERGY-FOOD NEXUS IN THE MEKONG REGION. WATER ALTERNATIVES, 8(1).

FORAN, T., BUTLER, J. R., WILLIAMS, L. J., WANJURA, W. J., HALL, A., CARTER, L., & CARBERRY, P. S. (2014). TAKING COMPLEXITY IN FOOD SYSTEMS SERIOUSLY: AN INTERDISCIPLINARY ANALYSIS. WORLD DEVELOPMENT, 61, 85-101.

FOURNIÉ, G., GUITIAN, J., DESVAUX, S., CUONG, V. C., PFEIFFER, D. U., MANGTANI, P., & GHANI, A. C. (2013). INTERVENTIONS FOR AVIAN INFLUENZA A (H5N1) RISK MANAGEMENT IN LIVE BIRD MARKET NETWORKS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 110(22), 9177-9182.

FRAITURE, C. D., WICHELNS, D., ROCKSTRÖM, J., KEMP-BENEDICT, E., ERIYAGAMA, N., GORDON, L. J., . . . KARLBERG, L. (2007). LOOKING AHEAD TO 2050: SCENARIOS OF ALTERNATIVE INVESTMENT APPROACHES.

FRANCIS, C., LIEBLEIN, G., BRELAND, T., SALOMONSSON, L., GEBER, U., SRISKANDARAJAH, N., & LANGER, V. (2008). TRANSDISCIPLINARY RESEARCH FOR A SUSTAINABLE AGRICULTURE AND FOOD SECTOR. AGRONOMY JOURNAL, 100(3), 771-776.

FREEMAN, L. C. (1977). A SET OF MEASURES OF CENTRALITY BASED ON BETWEENNESS. SOCIOMETRY, 35-41.

FRIAS-MARTINEZ, E., WILLIAMSON, G., & FRIAS-MARTINEZ, V. (2011). AN AGENT-BASED MODEL OF EPIDEMIC SPREAD USING HUMAN MOBILITY AND SOCIAL NETWORK INFORMATION. PAPER PRESENTED AT THE PRIVACY, SECURITY, RISK AND TRUST (PASSAT) AND 2011 IEEE THIRD INERNATIONAL CONFERENCE ON SOCIAL COMPUTING (SOCIALCOM), 2011 IEEE THIRD INTERNATIONAL CONFERENCE ON.

FUENTES, M. A., & KUPERMAN, M. N. (1999). CELLULAR AUTOMATA AND EPIDEMIOLOGICAL MODELS WITH SPATIAL DEPENDENCE. PHYSICA A: STATISTICAL MECHANICS AND ITS APPLICATIONS, 267(3), 471-486.

GARNER, M. G., DUBÉ, CAROLINE, STEVENSON, M. A., SANSON, R. L., ESTRADA, C., & GRIFFIN, J. (2007). EVALUATING ALTERNATIVE APPROACHES TO MANAGING ANIMAL DISEASE OUTBREAKS--THE ROLE OF MODELLING IN POLICY FORMULATION. VET. ITAL, 43(2), 285-298.

GARNER, M. G., & HAMILTON, S. A. (2011). PRINCIPLES OF EPIDEMIOLOGICAL MODELLING. REVUE SCIENTIFIQUE ET TECHNIQUE--OIE, 30(2), 407.

GARNETT, T. (2016). PLATING UP SOLUTIONS. SCIENCE, 353(6305).

GARNETT, T., APPLEBY, M. C., BALMFORD, A., BATEMAN, I. J., BENTON, T. G., BLOOMER, P., . . . HERRERO, M. (2013). SUSTAINABLE INTENSIFICATION IN AGRICULTURE: PREMISES AND POLICIES. SCIENCE, 341(6141).

GASSERT, F., M. LANDIS, M. LUCK, P. REIG, AND T. SHIAO. (2014). AQUEDUCT GLOBAL MAPS 2.1. RETRIEVED FROM: HTTP://WWW.WRI.ORG/PUBLICATION/AQUEDUCT-METADATA-GLOBAL

GHANI, A. C., & GARNETT, G. P. (2000). RISKS OF ACQUIRING AND TRANSMITTING SEXUALLY TRANSMITTED DISEASES IN SEXUAL PARTNER NETWORKS. SEXUALLY TRANSMITTED DISEASES, 27(10), 579-587.

GOEDE, D., & MORRISON, R. B. (2016). PRODUCTION IMPACT & TIME TO STABILITY IN SOW HERDS INFECTED WITH PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV). PREVENTIVE VETERINARY MEDICINE, 123, 202 - 207.

GOJOVIC, M. Z., SANDER, B., FISMAN, D., KRAHN, M. D., & BAUCH, C. T. (2009). MODELLING MITIGATION STRATEGIES FOR PANDEMIC (H1N1) 2009. CANADIAN MEDICAL ASSOCIATION JOURNAL, 181(10), 673-680.

GOMIERO, T. (2016). SOIL DEGRADATION, LAND SCARCITY AND FOOD SECURITY: REVIEWING A COMPLEX CHALLENGE. SUSTAINABILITY, 8(3), 281.

GOYAL, S. (2014). PEDV RESEARCH UPDATES: ENVIRONMENTAL STABILITY OF PED (PORCINE EPIDEMIC DIARRHEA VIRUS). RETRIEVED FROM HTTPS://WWW.AASV.ORG/PEDV/RESEARCH/13_215.PDF

GRANOVETTER, M. (1978). THRESHOLD MODELS OF COLLECTIVE BEHAVIOR. AMERICAN JOURNAL OF SOCIOLOGY, 1420-1443.

GRASSBERGER, P. (1983). ON THE CRITICAL BEHAVIOR OF THE GENERAL EPIDEMIC PROCESS AND DYNAMICAL PERCOLATION. MATHEMATICAL BIOSCIENCES, 63(2), 157-172.

GREEN, L. E., & MEDLEY, G. F. (2002). MATHEMATICAL MODELLING OF THE FOOT AND MOUTH DISEASE EPIDEMIC OF 2001: STRENGTHS AND WEAKNESSES. RESEARCH IN VETERINARY SCIENCE, 73(3), 201-205.

GRIGGS, D., STAFFORD-SMITH, M., GAFFNEY, O., ROCKSTRÖM, J., ÖHMAN, M. C., SHYAMSUNDAR, P., . . . NOBLE, I. (2013). POLICY: SUSTAINABLE DEVELOPMENT GOALS FOR PEOPLE AND PLANET. NATURE, 495(7441), 305.

GRIMMETT, G. (1999). WHAT IS PERCOLATION? : SPRINGER.

HAGBERG, A. A., SCHULT, D. A., & SWART, P. J. (2008, AUGUST). EXPLORING NETWORK STRUCTURE, DYNAMICS, AND FUNCTION USING NETWORKX, PASADENA, CA USA.

HAMMOND, R. A., & DUBÉ, L. (2012). A SYSTEMS SCIENCE PERSPECTIVE AND TRANSDISCIPLINARY MODELS FOR FOOD AND NUTRITION SECURITY. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 109(31), 12356-12363.

HANRAHAN, C. (2015). GLOBAL FOOD SECURITY BY THE NUMBERS. (SEPTEMBER), 1-20.

HANSEN, N., & OSTERMEIER, A. (1996). ADAPTING ARBITRARY NORMAL MUTATION DISTRIBUTIONS IN EVOLUTION STRATEGIES: THE COVARIANCE MATRIX ADAPTATION. PAPER PRESENTED AT THE EVOLUTIONARY COMPUTATION, 1996., PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON.

HANSON, S., NICHOLLS, R., RANGER, N., HALLEGATTE, S., CORFEE-MORLOT, J., HERWEIJER, C., & CHATEAU, J. (2011). A GLOBAL RANKING OF PORT CITIES WITH HIGH EXPOSURE TO CLIMATE EXTREMES. CLIMATIC CHANGE, 104(1), 89-111.

HARKO, T., LOBO, F. S. N., & MAK, M. K. (2014). EXACT ANALYTICAL SOLUTIONS OF THE SUSCEPTIBLE-INFECTED-RECOVERED (SIR)

EPIDEMIC MODEL AND OF THE SIR MODEL WITH EQUAL DEATH AND BIRTH RATES. APPLIED MATHEMATICS AND COMPUTATION, 236, 184-194.

HARVEY, N., REEVES, A., SCHOENBAUM, M. A., ZAGMUTT-VERGARA, F. J., DUB\'E, CAROLINE, HILL, A. E., . . . SALMAN, M. D. (2007). THE NORTH AMERICAN ANIMAL DISEASE SPREAD MODEL: A SIMULATION MODEL TO ASSIST DECISION MAKING IN EVALUATING ANIMAL DISEASE INCURSIONS. PREVENTIVE VETERINARY MEDICINE, 82(3-4), 176-197.

HEESTERBEEK, J. A. P. (2000). MATHEMATICAL EPIDEMIOLOGY OF INFECTIOUS DISEASES: MODEL BUILDING, ANALYSIS AND INTERPRETATION (VOL. 5): JOHN WILEY \& SONS.

HEIKKINEN, N. (2014). BIRTH CONTROL COULD HELP THE ENVIRONMENT, BUT NOT QUICKLY. SCIENTIFIC AMERICAN.

HERRERO, M., HAVLÍK, P., VALIN, H., NOTENBAERT, A., RUFINO, M. C., THORNTON, P. K., . . . OBERSTEINER, M. (2013). BIOMASS USE, PRODUCTION, FEED EFFICIENCIES, AND GREENHOUSE GAS EMISSIONS FROM GLOBAL LIVESTOCK SYSTEMS. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 110(52), 20888-20893.

HETHCOTE, H. W. (2000). THE MATHEMATICS OF INFECTIOUS DISEASES. SIAM REVIEW, 42(4), 599-653.

HUNTER, E., MAC NAMEE, B., & KELLEHER, J. D. (2017). A TAXONOMY FOR AGENT-BASED MODELS IN HUMAN INFECTIOUS DISEASE EPIDEMIOLOGY. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 20(3), 2.

IGEL, C., HANSEN, N., & ROTH, S. (2007). COVARIANCE MATRIX ADAPTATION FOR MULTI-OBJECTIVE OPTIMIZATION. EVOLUTIONARY COMPUTATION, 15(1), 1-28.

INGRAM, J. (2011). A FOOD SYSTEMS APPROACH TO RESEARCHING FOOD SECURITY AND ITS INTERACTIONS WITH GLOBAL ENVIRONMENTAL CHANGE. FOOD SECURITY, 3(4), 417-431.

IPCC. (2014). CLIMATE CHANGE 2014: IMPACTS, ADAPTATION, AND VULNERABILITY. PART A: GLOBAL AND SECTORAL ASPECTS. CONTRIBUTION OF WORKING GROUP II TO THE FIFTH ASSESSMENT REPORT OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE.

(V. R. B. C.B. FIELD, D.J. DOKKEN, K.J. MACH, M.D. MASTRANDREA, T.E. BILIR, M. CHATTERJEE, K.L. EBI, Y.O. ESTRADA, R.C. GENOVA, B. GIRMA, E.S. KISSEL, A.N. LEVY, S. MACCRACKEN, P.R. MASTRANDREA, L.L. WHITE ED.). CAMBRIDGE, UNITED KINGDOM AND NEW YORK, NY, USA: CAMBRIDGE UNIVERSITY PRESS.

IRAC. (2011). PREVENTION AND MANAGEMENT OF INSECTICIDE RESISTANCE IN VECTORS OF PUBLIC HEALTH IMPORTANCE.

ISTFETC ACTION. (2006). LIVING ROADMAP FOR COMPLEX SYSTEMS SCIENCE.

JALAVA, M., GUILLAUME, J. H., KUMMU, M., PORKKA, M., SIEBERT, S., & VARIS, O. (2016). DIET CHANGE AND FOOD LOSS REDUCTION: WHAT IS THEIR COMBINED IMPACT ON GLOBAL WATER USE AND SCARCITY? EARTH'S FUTURE, 4(3), 62-78.

JANSSEN, M. A., & OSTROM, E. (2006). EMPIRICALLY BASED, AGENT-BASED MODELS. ECOLOGY AND SOCIETY, 11(2), 37.

JOHNSON, D. (1997). THE TRIANGULAR DISTRIBUTION AS A PROXY FOR THE BETA DISTRIBUTION IN RISK ANALYSIS. JOURNAL OF THE ROYAL STATISTICAL SOCIETY: SERIES D (THE STATISTICIAN), 46(3), 387-398.

JUNG, K., & SAIF, L. J. (2015). PORCINE EPIDEMIC DIARRHEA VIRUS INFECTION: ETIOLOGY, EPIDEMIOLOGY, PATHOGENESIS AND IMMUNOPROPHYLAXIS. THE VETERINARY JOURNAL, 204(2), 134-143.

KAO, R. R., DANON, L., GREEN, D. M., & KISS, I. Z. (2006). DEMOGRAPHIC STRUCTURE AND PATHOGEN DYNAMICS ON THE NETWORK OF LIVESTOCK MOVEMENTS IN GREAT BRITAIN. PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON B: BIOLOGICAL SCIENCES, 273(1597), 1999-2007.

KAUL, H., & VENTIKOS, Y. (2013). INVESTIGATING BIOCOMPLEXITY THROUGH THE AGENT-BASED PARADIGM. BRIEFINGS IN BIOINFORMATICS, 16(1), 137-152.

KEATING, B., & CARBERRY, P. (2010). SUSTAINABLE PRODUCTION, FOOD SECURITY AND SUPPLY CHAIN IMPLICATIONS.

KEATING, B. A., HERRERO, M., CARBERRY, P. S., GARDNER, J., & COLE, M. B. (2014). FOOD WEDGES: FRAMING THE GLOBAL FOOD DEMAND AND

SUPPLY CHALLENGE TOWARDS 2050. GLOBAL FOOD SECURITY, 3(3), 125-132.

KEELING, M., WOOLHOUSE, M., MAY, R., DAVIES, G., & GRENFELL, B. (2003). MODELLING VACCINATION STRATEGIES AGAINST FOOT-AND-MOUTH DISEASE. NATURE, 421(6919), 136-142.

KEELING, M. J. (1999). THE EFFECTS OF LOCAL SPATIAL STRUCTURE ON EPIDEMIOLOGICAL INVASIONS. PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON B: BIOLOGICAL SCIENCES, 266(1421), 859-867.

KEELING, M. J., & EAMES, K. T. D. (2005). NETWORKS AND EPIDEMIC MODELS. JOURNAL OF THE ROYAL SOCIETY INTERFACE, 2(4), 295-307.

KEELING, M. J., WOOLHOUSE, M. E. J., MAY, R. M., DAVIES, G., & GRENFELL, B. T. (2003). MODELLING VACCINATION STRATEGIES AGAINST FOOT-AND-MOUTH DISEASE. NATURE, 421(6919), 136-142.

KEELING, M. J., WOOLHOUSE, M. E. J., SHAW, D. J., MATTHEWS, L., CHASE-TOPPING, M., HAYDON, D. T., . . . GRENFELL, B. T. (2001). DYNAMICS OF THE 2001 UK FOOT AND MOUTH EPIDEMIC: STOCHASTIC DISPERSAL IN A HETEROGENEOUS LANDSCAPE. SCIENCE, 294(5543), 813-817.

KERMACK, W. O., & MCKENDRICK, A. G. (1927). A CONTRIBUTION TO THE MATHEMATICAL THEORY OF EPIDEMICS. PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON. SERIES A, CONTAINING PAPERS OF A MATHEMATICAL AND PHYSICAL CHARACTER, 115(772), 700-721.

KILLWORTH, P. D., MCCARTY, C., BERNARD, H. R., SHELLEY, G. A., & JOHNSEN, E. C. (1998). ESTIMATION OF SEROPREVALENCE, RAPE, AND HOMELESSNESS IN THE UNITED STATES USING A SOCIAL NETWORK APPROACH. EVALUATION REVIEW, 22(2), 289-308.

KINZEL, W. (1985). PHASE TRANSITIONS OF CELLULAR AUTOMATA. ZEITSCHRIFT FÜR PHYSIK B CONDENSED MATTER, 58(3), 229-244.

KISS, I. Z., GREEN, D. M., & KAO, R. R. (2006). THE NETWORK OF SHEEP MOVEMENTS WITHIN GREAT BRITAIN: NETWORK PROPERTIES AND THEIR IMPLICATIONS FOR INFECTIOUS DISEASE SPREAD. JOURNAL OF THE ROYAL SOCIETY INTERFACE, 3(10), 669-677.

KITCHING, R. P., THRUSFIELD, M. V., & TAYLOR, N. M. (2006). USE AND ABUSE OF MATHEMATICAL MODELS: AN ILLUSTRATION FROM THE 2001

FOOT AND MOUTH DISEASE EPIDEMIC IN THE UNITED KINGDOM. REVUE SCIENTIFIQUE ET TECHNIQUE--OIE, 25(1), 293.

KITSAK, M., GALLOS, L. K., HAVLIN, S., LILJEROS, F., MUCHNIK, L., STANLEY, H. E., & MAKSE, H. A. (2010). IDENTIFICATION OF INFLUENTIAL SPREADERS IN COMPLEX NETWORKS. NATURE PHYSICS, 6(11), 888-893.

KLÜGL, F. (2008). A VALIDATION METHODOLOGY FOR AGENT-BASED SIMULATIONS. PAPER PRESENTED AT THE PROCEEDINGS OF THE 2008 ACM SYMPOSIUM ON APPLIED COMPUTING.

KOETSE, M. J., & RIETVELD, P. (2009). THE IMPACT OF CLIMATE CHANGE AND WEATHER ON TRANSPORT: AN OVERVIEW OF EMPIRICAL FINDINGS. TRANSPORTATION RESEARCH PART D: TRANSPORT AND ENVIRONMENT, 14(3), 205-221.

KONAR, M., DALIN, C., SUWEIS, S., HANASAKI, N., RINALDO, A., & RODRIGUEZ-ITURBE, I. (2011). WATER FOR FOOD: THE GLOBAL VIRTUAL WATER TRADE NETWORK. WATER RESOURCES RESEARCH, 47(5), N/A-N/A. DOI:10.1029/2010WR010307

KOOPMAN, J. S., JACQUEZ, J. A., WELCH, G. W., SIMON, C. P., FOXMAN, B., POLLOCK, S. M., . . . LANGE, K. (1997). THE ROLE OF EARLY HIV INFECTION IN THE SPREAD OF HIV THROUGH POPULATIONS. JOURNAL OF ACQUIRED IMMUNE DEFICIENCY SYNDROMES, 14(3), 249-258.

KUMMU, M., FADER, M., GERTEN, D., GUILLAUME, J. H., JALAVA, M., JÄGERMEYR, J., . . . VARIS, O. (2017). BRINGING IT ALL TOGETHER: LINKING MEASURES TO SECURE NATIONS' FOOD SUPPLY. CURRENT OPINION IN ENVIRONMENTAL SUSTAINABILITY, 29, 98-117.

LABORDE, D., BIZIKOVA, L., LALLEMANT, T., & SMALLER, C. (2016). ENDING HUNGER: WHAT WOULD IT COST? : INTERNATIONAL INSTITUTE FOR SUSTAINABLE DEVELOPMENT.

LACHAPELLE, V., LETELLIER, A., FRAVALO, P., BRASSARD, J., & L'HOMME, Y. (2017). DYNAMICS OF VIRUS DISTRIBUTION IN A DEFINED SWINE PRODUCTION NETWORK USING ENTERIC VIRUSES AS MOLECULAR MARKERS. APPL. ENVIRON. MICROBIOL., 83(4), E03187-03116.

LAGI, M., BERTRAND, K. Z., & BAR-YAM, Y. (2011). THE FOOD CRISES AND POLITICAL INSTABILITY IN NORTH AFRICA AND THE MIDDLE EAST. SSRN ELECTRONIC JOURNAL, 1-15.

LEMMEN, D. S., & WARREN, F. J. (2004). CLIMATE CHANGE IMPACTS AND ADAPTATION: A CANADIAN PERSPECTIVE.

LENTON, T. M., HELD, H., KRIEGLER, E., HALL, J. W., LUCHT, W., RAHMSTORF, S., & SCHELLNHUBER, H. J. (2008). TIPPING ELEMENTS IN THE EARTH'S CLIMATE SYSTEM. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 105(6), 1786-1793.

LIU, J., MOONEY, H., HULL, V., DAVIS, S. J., GASKELL, J., HERTEL, T., . . . LI, S. (2015). SYSTEMS INTEGRATION FOR GLOBAL SUSTAINABILITY. SCIENCE, 347(6225). DOI:10.1126/SCIENCE.1258832

LOBELL, D. B., SCHLENKER, W., & COSTA-ROBERTS, J. (2011). CLIMATE TRENDS AND GLOBAL CROP PRODUCTION SINCE 1980. SCIENCE, 333(6042), 616-620. DOI:10.1126/SCIENCE.1204531

LOTKA, A. J. (1925). ELEMENTS OF PHYSICAL BIOLOGY. BALTIMORE: WILLIAMS AND WILKINS: NEW YORK: DOVER PUBLICATIONS.

LOWE, J., GAUGER, P., HARMON, K., ZHANG, J., CONNOR, J., YESKE, P., . . . MAIN, R. (2014). ROLE OF TRANSPORTATION IN SPREAD OF PORCINE EPIDEMIC DIARRHEA VIRUS INFECTION, UNITED STATES. EMERGING INFECTIOUS DISEASES, 20(5), 872.

LUCE, R. D., & PERRY, A. D. (1949). A METHOD OF MATRIX ANALYSIS OF GROUP STRUCTURE. PSYCHOMETRIKA, 14(2), 95-116.

MACAL, C. M., & NORTH, M. J. (2010). TUTORIAL ON AGENT-BASED MODELLING AND SIMULATION. JOURNAL OF SIMULATION, 4(3), 151-162.

MANSLEY, L. M., DUNLOP, P. J., WHITESIDE, S. M., & SMITH, R. G. H. (2003). EARLY DISSEMINATION OF FOOT-AND-MOUTH. THE VETERINARY RECORD, 153, 43-50.

MAY, R. M. (2004). USES AND ABUSES OF MATHEMATICS IN BIOLOGY. SCIENCE, 303(5659), 790-793.

MEADOWS, D. (2001). DANCING WITH SYSTEMS. WHOLE EARTH, 106, 58-63.

MEADOWS, D. H. (2008). THINKING IN SYSTEMS: A PRIMER: CHELSEA GREEN PUBLISHING.

MELILLO, J. M., RICHMOND, T., & YOHE, G. (2014). CLIMATE CHANGE IMPACTS IN THE UNITED STATES. THIRD NATIONAL CLIMATE ASSESSMENT.

MÉNDEZ, V. E., BACON, C. M., & COHEN, R. (2013). AGROECOLOGY AS A TRANSDISCIPLINARY, PARTICIPATORY, AND ACTION-ORIENTED APPROACH. AGROECOLOGY AND SUSTAINABLE FOOD SYSTEMS, 37(1), 3-18.

MERRILL, S. C., KOLIBA, C. J., MOEGENBURG, S. M., ZIA, A., PARKER, J., SELLNOW, T., . . . SMITH, J. M. (2019). DECISION-MAKING IN LIVESTOCK BIOSECURITY PRACTICES AMIDST ENVIRONMENTAL AND SOCIAL UNCERTAINTY: EVIDENCE FROM AN EXPERIMENTAL GAME. PLOS ONE (ARTICLE IN PRESS).

MERRILL, S. C., MOEGENBURG, S., KOLIBA, C. J., ZIA, A., TRINITY, L., CLARK, E., . . . SMITH, J. M. (2019). WILLINGNESS TO COMPLY WITH BIOSECURITY IN LIVESTOCK FACILITIES: EVIDENCE FROM EXPERIMENTAL SIMULATIONS. FRONTIERS IN VETERINARY SCIENCE (ACCEPTED PENDING REVISIONS).

MIKLER, A. R., VENKATACHALAM, S., & ABBAS, K. (2005). MODELING INFECTIOUS DISEASES USING GLOBAL STOCHASTIC CELLULAR AUTOMATA. JOURNAL OF BIOLOGICAL SYSTEMS, 13(04), 421-439.

MILLER, J. H., & PAGE, S. E. (2009). COMPLEX ADAPTIVE SYSTEMS: AN INTRODUCTION TO COMPUTATIONAL MODELS OF SOCIAL LIFE (VOL. 17): PRINCETON UNIVERSITY PRESS.

MONTGOMERY, M. R. (2008). THE URBAN TRANSFORMATION OF THE DEVELOPING WORLD. SCIENCE, 319(5864), 761-764. DOI:10.1126/SCIENCE.1153012

MOORE, C., & NEWMAN, M. E. J. (2000). EPIDEMICS AND PERCOLATION IN SMALL-WORLD NETWORKS. PHYSICAL REVIEW E, 61(5), 5678.

MORENO, Y., PASTOR-SATORRAS, R., & VESPIGNANI, A. (2002). EPIDEMIC OUTBREAKS IN COMPLEX HETEROGENEOUS NETWORKS. THE EUROPEAN

PHYSICAL JOURNAL B - CONDENSED MATTER AND COMPLEX SYSTEMS, 26(4), 521-529.

MORGAN, R. P. C. (2009). SOIL EROSION AND CONSERVATION: JOHN WILEY & SONS.

MOSS, S. (2008). ALTERNATIVE APPROACHES TO THE EMPIRICAL VALIDATION OF AGENT-BASED MODELS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 11(1), 5.

MUELLER, N. D., GERBER, J. S., JOHNSTON, M., RAY, D. K., RAMANKUTTY, N., & FOLEY, J. A. (2012). CLOSING YIELD GAPS THROUGH NUTRIENT AND WATER MANAGEMENT. NATURE, 490(254).

MURAI, K., MORIGUCHI, S., HAYAMA, Y., KOBAYASHI, S., MIYAZAKI, A., TSUTSUI, T., & YAMAMOTO, T. (2018). MATHEMATICAL MODELING OF PORCINE EPIDEMIC DIARRHEA VIRUS DYNAMICS WITHIN A FARROW-TO-FINISH SWINE FARM TO INVESTIGATE THE EFFECTS OF CONTROL MEASURES. PREVENTIVE VETERINARY MEDICINE, 149, 115-124.

MURCH, R. S. (2003). MICROBIAL FORENSICS: BUILDING A NATIONAL CAPACITY TO INVESTIGATE BIOTERRORISM. BIOSECURITY AND BIOTERRORISM: BIODEFENSE STRATEGY, PRACTICE, AND SCIENCE, 1(2), 117-122.

NACHTERGAELE, F., PETRI, M., BIANCALANI, R., VAN LYNDEN, G., VAN VELTHUIZEN, H., & BLOISE, M. (2011). AN INFORMATION DATABASE FOR LAND DEGRADATION ASSESSMENT AT GLOBAL LEVEL. GLOBAL LAND DEGRADATION INFORMATION SYSTEM (GLADIS).

NAPOLI, M., DE MURO, P., & MAZZIOTTA, M. (2011). TOWARDS A FOOD INSECURITY MULTIDIMENSIONAL INDEX (FIMI). MASTER IN HUMAN DEVELOPMENT AND FOOD SECURITY.

NASA. (2017). NASA EARTH EXCHANGE GLOBAL DAILY DOWNSCALED PROJECTIONS (NEX-GDDP). RETRIEVED FROM: HTTPS://CDS.NCCS.NASA.GOV/NEX-GDDP/

NASS, U. (2012). HOGS AND PIGS: INVENTORY BY TYPE OF PRODUCER. RETRIEVED FROM HTTP://AGCENSUS.USDA.GOV/PUBLICATIONS/2012/FULL_REPORT/VOLUME_1,_CHAPTER_1_US/ST99_1_020_023.PDF

NASS, U. (2014). LIVESTOCK SLAUGHTER, 2013 SUMMARY: FEDERALLY INSPECTED PLANTS AND HEAD SLAUGHTERED BY SPECIES, STATES AND UNITED STATES, 2013.

NATALE, F., GIOVANNINI, A., SAVINI, L., PALMA, D., POSSENTI, L., FIORE, G., & CALISTRI, P. (2009). NETWORK ANALYSIS OF ITALIAN CATTLE TRADE PATTERNS AND EVALUATION OF RISKS FOR POTENTIAL DISEASE SPREAD. PREVENTIVE VETERINARY MEDICINE, 92(4), 341-350.

NELSON, G. C., ROSEGRANT, M. W., KOO, J., ROBERTSON, R., SULSER, T., ZHU, T., . . . BATKA, M. (2009). CLIMATE CHANGE: IMPACT ON AGRICULTURE AND COSTS OF ADAPTATION (VOL. 21): INTL FOOD POLICY RES INST.

NEWMAN, M. E., GIRVAN, M., & FARMER, J. D. (2002). OPTIMAL DESIGN, ROBUSTNESS, AND RISK AVERSION. PHYSICAL REVIEW LETTERS, 89(2), 028301.

NEWMAN, M. E. J. (2002). SPREAD OF EPIDEMIC DISEASE ON NETWORKS. PHYSICAL REVIEW E, 66(1).

NEWMAN, M. E. J. (2003). MIXING PATTERNS IN NETWORKS. PHYSICAL REVIEW E, 67(2), 026126.

NEWMAN, M. E. J. (2005). A MEASURE OF BETWEENNESS CENTRALITY BASED ON RANDOM WALKS. SOCIAL NETWORKS, 27(1), 39-54.

NEWMAN, M. E. J., GIRVAN, M., & DOYNE FARMER, J. (2002). OPTIMAL DESIGN, ROBUSTNESS, AND RISK AVERSION. PHYSICAL REVIEW LETTERS, 89(2), 028301.

NILSSON, M., GRIGGS, D., & VISBECK, M. (2016). POLICY: MAP THE INTERACTIONS BETWEEN SUSTAINABLE DEVELOPMENT GOALS. NATURE NEWS, 534(7607), 320.

O'DEA, E. B., SNELSON, H., & BANSAL, S. (2015). STATE-LEVEL TRANSPORT FLOWS ARE PREDICTIVE OF THE DYNAMICS OF PORCINE EPIDEMIC DIARRHEA VIRUS. BIORXIV, 017178.

OHTSUKI, T., & KEYES, T. (1986). KINETIC GROWTH PERCOLATION: EPIDEMIC PROCESSES WITH IMMUNIZATION. PHYSICAL REVIEW A, 33(2), 1223.

OLSON, R. (2015). DIETARY GUIDELINES FOR AMERICANS. RETRIEVED FROM

Ortiz-Pelaez, A., Pfeiffer, D. U., Soares-Magalhaes, R. J., & Guitian, F. J. (2006). Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. Preventive veterinary medicine, 76(1-2), 40-55.

Ostrom, E. (2010). Polycentric systems for coping with collective action and global environmental change. Global Environmental Change, 20(4), 550-557. doi:https://doi.org/10.1016/j.gloenvcha.2010.07.004

Otto, I. M., Biewald, A., Coumou, D., Feulner, G., Kohler, C., Nocke, T., . . . Beck, U. (2015). Socio-economic data for global environmental change research. Nature Clim. Change, 5(6), 503-506. doi:10.1038/nclimate2593

Paarlberg, P. (2014). Updated estimated economic welfare impacts of porcine epidemic diarrhea virus (PEDV). Retrieved from

Pacala, S., & Socolow, R. (2004). Stabilization wedges: solving the climate problem for the next 50 years with current technologies. Science, 305(5686), 968-972.

Pardey, P. G., Beddow, J. M., Hurley, T. M., Beatty, T. K. M., & Eidman, V. R. (2014). A bounds analysis of world food futures: Global agriculture through to 2050. Australian Journal of Agricultural and Resource Economics, 58(4), 571-589.

Parker, J., & Epstein, J. M. (2011). A distributed platform for global-scale agent-based models of disease transmission. ACM transactions on modeling and computer simulation (TOMACS), 22(1), 2.

Parnell, G. S., Smith, C. M., & Moxley, F. I. (2010). Intelligent adversary risk analysis: A bioterrorism risk management model. Risk Analysis: An International Journal, 30(1), 32-48.

Parry, M., Rosenzweig, C., Iglesias, A., Fischer, G., & Livermore, M. (1999). Climate change and world food security: a new assessment. Global Environmental Change, 9(Supplement 1), S51-S67. doi:https://doi.org/10.1016/S0959-3780(99)00018-7

PARRY, M., ROSENZWEIG, C., & LIVERMORE, M. (2005). CLIMATE CHANGE, GLOBAL FOOD SUPPLY AND RISK OF HUNGER. PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B: BIOLOGICAL SCIENCES, 360(1463), 2125-2138. DOI:10.1098/RSTB.2005.1751

PASTOR-SATORRAS, R., & VESPIGNANI, A. (2001). EPIDEMIC SPREADING IN SCALE-FREE NETWORKS. PHYSICAL REVIEW LETTERS, 86(14), 3200.

PEREZ, A. M., WARD, M. P., CHARMANDARIÁN, A., & RITACCO, V. (2002). SIMULATION MODEL OF WITHIN-HERD TRANSMISSION OF BOVINE TUBERCULOSIS IN ARGENTINE DAIRY HERDS. PREVENTIVE VETERINARY MEDICINE, 54(4), 361-372.

PEREZ, L., & DRAGICEVIC, S. (2009). AN AGENT-BASED APPROACH FOR MODELING DYNAMICS OF CONTAGIOUS DISEASE SPREAD. INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS, 8(1), 50.

PERROT, N., TRELEA, I.-C., BAUDRIT, C., TRYSTRAM, G., & BOURGINE, P. (2011). MODELLING AND ANALYSIS OF COMPLEX FOOD SYSTEMS: STATE OF THE ART AND NEW TRENDS. TRENDS IN FOOD SCIENCE & TECHNOLOGY, 22(6), 304-314.

PFEIFER, B., KUGLER, K., TEJADA, M. M., BAUMGARTNER, C., SEGER, M., OSL, M., . . . WURZ, M. (2008). A CELLULAR AUTOMATON FRAMEWORK FOR INFECTIOUS DISEASE SPREAD SIMULATION. THE OPEN MEDICAL INFORMATICS JOURNAL, 2, 70.

PONGIGLIONE, F. (2015). THE NEED FOR A PRIORITY STRUCTURE FOR THE SUSTAINABLE DEVELOPMENT GOALS. JOURNAL OF GLOBAL ETHICS, 11(1), 37-42.

POPKIN, B. M. (2001). THE NUTRITION TRANSITION AND OBESITY IN THE DEVELOPING WORLD. THE JOURNAL OF NUTRITION, 131(3), 871S-873S.

POPKIN, B. M., ADAIR, L. S., & NG, S. W. (2012). GLOBAL NUTRITION TRANSITION AND THE PANDEMIC OF OBESITY IN DEVELOPING COUNTRIES. NUTRITION REVIEWS, 70(1), 3-21.

PRADHAN, P., FISCHER, G., VAN VELTHUIZEN, H., REUSSER, D. E., & KROPP, J. P. (2015). CLOSING YIELD GAPS: HOW SUSTAINABLE CAN WE BE? PLOS ONE, 10(6), E0129487.

PUGA, D., & VENABLES, A. J. (1996). THE SPREAD OF INDUSTRY: SPATIAL AGGLOMERATION IN ECONOMIC DEVELOPMENT. JOURNAL OF THE JAPANESE AND INTERNATIONAL ECONOMIES, 10(4), 440-464.

RACEY, M., O'BRIEN, C., DOUGLAS, S., MARQUEZ, O., HENDRIE, G., & NEWTON, G. (2016). SYSTEMATIC REVIEW OF SCHOOL-BASED INTERVENTIONS TO MODIFY DIETARY BEHAVIOR: DOES INTERVENTION INTENSITY IMPACT EFFECTIVENESS? JOURNAL OF SCHOOL HEALTH, 86(6), 452-463.

RAHMANDAD, H., & STERMAN, J. (2008). HETEROGENEITY AND NETWORK STRUCTURE IN THE DYNAMICS OF DIFFUSION: COMPARING AGENT-BASED AND DIFFERENTIAL EQUATION MODELS. MANAGEMENT SCIENCE, 54(5), 998-1014.

RAHNAMAYAN, S., & DIERAS, P. (2008, MAY). EFFICIENCY COMPETITION ON N-QUEEN PROBLEM: DE VS. CMA-ES. PAPER PRESENTED AT THE 2008 CANADIAN CONFERENCE ON ELECTRICAL AND COMPUTER ENGINEERING.

RANGANATHAN, J., VENNARD, D., WAITE, R., DUMAS, P., LIPINSKI, B., SEARCHINGER, T., & GLOBAGRI-WRR. (2016). SHIFTING DIETS FOR A SUSTAINABLE FOOD FUTURE. (APRIL), 90.

RAY, D. K., MUELLER, N. D., WEST, P. C., & FOLEY, J. A. (2013). YIELD TRENDS ARE INSUFFICIENT TO DOUBLE GLOBAL CROP PRODUCTION BY 2050. PLOS ONE, 8(6), 1-8.

RESCHER, N. (1998). PREDICTING THE FUTURE: AN INTRODUCTION TO THE THEORY OF FORECASTING: SUNY PRESS.

REYNOSO-MEZA, G., SANCHIS, J., BLASCO, X., & GARCÍA-NIETO, S. (2014). PHYSICAL PROGRAMMING FOR PREFERENCE DRIVEN EVOLUTIONARY MULTI-OBJECTIVE OPTIMIZATION. APPLIED SOFT COMPUTING, 24, 341-362.

ROBERT, K. W., PARRIS, T. M., & LEISEROWITZ, A. A. (2005). WHAT IS SUSTAINABLE DEVELOPMENT? GOALS, INDICATORS, VALUES, AND PRACTICE. ENVIRONMENT: SCIENCE AND POLICY FOR SUSTAINABLE DEVELOPMENT, 47(3), 8-21.

ROBINS, J., BOGEN, S., FRANCIS, A., WESTHOEK, A., KANAREK, A., LENHART, S., & EDA, S. (2015). AGENT-BASED MODEL FOR JOHNE'S DISEASE DYNAMICS IN A DAIRY HERD. VETERINARY RESEARCH, 46(1), 68.

ROCKSTRÖM, J., WILLIAMS, J., DAILY, G., NOBLE, A., MATTHEWS, N., GORDON, L., . . . STEDUTO, P. (2017). SUSTAINABLE INTENSIFICATION OF AGRICULTURE FOR HUMAN PROSPERITY AND GLOBAL SUSTAINABILITY. AMBIO, 46(1), 4-17.

RODRIGUEZ, D., DE VOIL, P., HUDSON, D., BROWN, J., HAYMAN, P., MARROU, H., & MEINKE, H. (2018). PREDICTING OPTIMUM CROP DESIGNS USING CROP MODELS AND SEASONAL CLIMATE FORECASTS. SCIENTIFIC REPORTS, 8(1), 2231.

ROGERS, E. M. (2010). DIFFUSION OF INNOVATIONS: SIMON AND SCHUSTER.

ROSENZWEIG, C., & PARRY, M. L. (1994). POTENTIAL IMPACT OF CLIMATE CHANGE ON WORLD FOOD SUPPLY. NATURE, 367(6459), 133-138.

ROSER, M., & RITCHIE, H. (2019). OUR WORLD IN DATA: FOOD PER PERSON. RETRIEVED FROM HTTPS://OURWORLDINDATA.ORG/FOOD-PER-PERSON

ROTHENBERG, R. B., POTTERAT, J. J., WOODHOUSE, D. E., DARROW, W. W., MUTH, S. Q., & KLOVDAHL, A. S. (1995). CHOOSING A CENTRALITY MEASURE: EPIDEMIOLOGIC CORRELATES IN THE COLORADO SPRINGS STUDY OF SOCIAL NETWORKS. SOCIAL NETWORKS, 17(3-4), 273-297.

RUBEN, R., VERHAGEN, J., & PLAISIER, C. (2019). THE CHALLENGE OF FOOD SYSTEMS RESEARCH: WHAT DIFFERENCE DOES IT MAKE? SUSTAINABILITY, 11(1), 171.

RYAN, B., & GROSS, N. C. (1943). THE DIFFUSION OF HYBRID SEED CORN IN TWO IOWA COMMUNITIES. RURAL SOCIOLOGY, 8(1), 15.

SADRAS, V. (2015). YIELD GAP ANALYSIS OF FIELD CROPS: METHODS AND CASE STUDIES.

SALATHÉ, MARCEL, & JONES, J. H. (2010). DYNAMICS AND CONTROL OF DISEASES IN NETWORKS WITH COMMUNITY STRUCTURE. PLOS COMPUTATIONAL BIOLOGY, 6(4), E1000736.

SANDER, L. M., WARREN, C. P., SOKOLOV, I. M., SIMON, C., & KOOPMAN, J. (2002). PERCOLATION ON HETEROGENEOUS NETWORKS AS A MODEL FOR EPIDEMICS. MATHEMATICAL BIOSCIENCES, 180(1), 293-305.

SANS, P., & COMBRIS, P. (2015). WORLD MEAT CONSUMPTION PATTERNS: AN OVERVIEW OF THE LAST FIFTY YEARS (1961–2011). MEAT SCIENCE, 109, 106-111.

SARAMÄKI, J., KIVELÄ, M., ONNELA, J.-P., KASKI, K., & KERTESZ, J. (2007). GENERALIZATIONS OF THE CLUSTERING COEFFICIENT TO WEIGHTED COMPLEX NETWORKS. PHYSICAL REVIEW E, 75(2), 027105.

SCHELLING, T. C. (2006). MICROMOTIVES AND MACROBEHAVIOR: WW NORTON & COMPANY.

SCHMIDT, M., & LIPSON, H. (2009). DISTILLING FREE-FORM NATURAL LAWS FROM EXPERIMENTAL DATA. SCIENCE, 324(5923), 81-85.

SCHMIDT, M., & LIPSON, H. (2013). EUREQA (VERSION 1.24)[SOFTWARE]. NUTONIAN, SOMERVILLE, MASS, USA.

SCHMIDT-TRAUB, G., & SHAH, A. (2015). INVESTMENT NEEDS TO ACHIEVE THE SUSTAINABLE DEVELOPMENT GOALS. PARIS AND NEW YORK: SUSTAINABLE DEVELOPMENT SOLUTIONS NETWORK.

SCHOENBAUM, M. A., & DISNEY, W. T. (2003). MODELING ALTERNATIVE MITIGATION STRATEGIES FOR A HYPOTHETICAL OUTBREAK OF FOOT-AND-MOUTH DISEASE IN THE UNITED STATES. PREVENTIVE VETERINARY MEDICINE, 58(1-2), 25-52.

SCHULZ, L. L., & TONSOR, G. T. (2015). ASSESSMENT OF THE ECONOMIC IMPACTS OF PORCINE EPIDEMIC DIARRHEA VIRUS IN THE UNITED STATES. JOURNAL OF ANIMAL SCIENCE, 93(11), 5111-5118.

SCOTT, D. M., NOVAK, D. C., AULTMAN-HALL, L., & GUO, F. (2006). NETWORK ROBUSTNESS INDEX: A NEW METHOD FOR IDENTIFYING CRITICAL LINKS AND EVALUATING THE PERFORMANCE OF TRANSPORTATION NETWORKS. JOURNAL OF TRANSPORT GEOGRAPHY, 14(3), 215-227.

SEN, A. K. (1999). DEVELOPMENT AS FREEDOM. 1-17.

SERRANO, M. Á., & BOGUNÁ, M. (2006). PERCOLATION AND EPIDEMIC THRESHOLDS IN CLUSTERED NETWORKS. PHYSICAL REVIEW LETTERS, 97(8), 088701.

SHI, Z. Z., WU, C.-H., & BEN-ARIEH, D. (2014). AGENT-BASED MODEL: A SURGING TOOL TO SIMULATE INFECTIOUS DISEASES IN THE IMMUNE SYSTEM. OPEN JOURNAL OF MODELLING AND SIMULATION, 2(01), 12.

SOBAL, J., KHAN, L. K., & BISOGNI, C. (1998). A CONCEPTUAL MODEL OF THE FOOD AND NUTRITION SYSTEM. SOCIAL SCIENCE & MEDICINE, 47(7), 853-863.

SORNETTE, D. (2006). CRITICAL PHENOMENA IN NATURAL SCIENCES: CHAOS, FRACTALS, SELFORGANIZATION AND DISORDER: CONCEPTS AND TOOLS: SPRINGER SCIENCE & BUSINESS MEDIA.

SQUAZZONI, F., & BOERO, R. (2010). COMPLEXITY-FRIENDLY POLICY MODELLING. INNOVATION IN COMPLEX SYSTEMS. LONDON: ROUTLEDGE.

STARBIRD, E., NORTON, M., & MARCUS, R. (2016). INVESTING IN FAMILY PLANNING: KEY TO ACHIEVING THE SUSTAINABLE DEVELOPMENT GOALS. GLOBAL HEALTH: SCIENCE AND PRACTICE, 4(2), 191-210.

STAUFFER, D., & AHARONY, A. (1994). INTRODUCTION TO PERCOLATION THEORY: CRC PRESS.

STEFFEN, W., RICHARDSON, K., ROCKSTRÖM, J., CORNELL, S. E., FETZER, I., BENNETT, E. M., . . . DE WIT, C. A. (2015). PLANETARY BOUNDARIES: GUIDING HUMAN DEVELOPMENT ON A CHANGING PLANET. SCIENCE, 347(6223), 1259855.

STEVENSON, M. A., SANSON, R. L., STERN, M. W., O'LEARY, B. D., SUJAU, M., MOLES-BENFELL, N., & MORRIS, R. S. (2013). INTERSPREAD PLUS: A SPATIAL AND STOCHASTIC SIMULATION MODEL OF DISEASE IN ANIMAL POPULATIONS. PREVENTIVE VETERINARY MEDICINE, 109(1-2), 10-24.

STORN, R., & PRICE, K. (1997). DIFFERENTIAL EVOLUTION--A SIMPLE AND EFFICIENT HEURISTIC FOR GLOBAL OPTIMIZATION OVER CONTINUOUS SPACES. JOURNAL OF GLOBAL OPTIMIZATION, 11(4), 341-359.

SULLIVAN, J., AULTMAN-HALL, L., & NOVAK, D. (2009). A REVIEW OF CURRENT PRACTICE IN NETWORK DISRUPTION ANALYSIS AND AN

ASSESSMENT OF THE ABILITY TO ACCOUNT FOR ISOLATING LINKS IN TRANSPORTATION NETWORKS. TRANSPORTATION LETTERS, 1(4), 271-280.

THE PIG SITE. (2014). PIG FARM TARGETS: STOCKMANSHIP STANDARDS.

THOMPSON, J., & SCOONES, I. (2009). ADDRESSING THE DYNAMICS OF AGRI-FOOD SYSTEMS: AN EMERGING AGENDA FOR SOCIAL SCIENCE RESEARCH. ENVIRONMENTAL SCIENCE & POLICY, 12(4), 386-397.

THORNTON, P. K., WHITBREAD, A., BAEDEKER, T., CAIRNS, J., CLAESSENS, L., BAETHGEN, W., . . . HERRERO, M. (2018). A FRAMEWORK FOR PRIORITY-SETTING IN CLIMATE SMART AGRICULTURE RESEARCH. AGRICULTURAL SYSTEMS, 167, 161-175.

TILMAN, D., BALZER, C., HILL, J., & BEFORT, B. L. (2011). GLOBAL FOOD DEMAND AND THE SUSTAINABLE INTENSIFICATION OF AGRICULTURE. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 108(50), 20260-20264.

TILMAN, D., CASSMAN, K. G., MATSON, P. A., NAYLOR, R., & POLASKY, S. (2002). AGRICULTURAL SUSTAINABILITY AND INTENSIVE PRODUCTION PRACTICES. NATURE, 418(6898), 671-677.

TRACEY, J. A., BEVINS, S. N., VANDEWOUDE, S., & CROOKS, K. R. (2014). AN AGENT-BASED MOVEMENT MODEL TO ASSESS THE IMPACT OF LANDSCAPE FRAGMENTATION ON DISEASE TRANSMISSION. ECOSPHERE, 5(9), 1-24.

TSAI, Y., ZIA, A., KOLIBA, C., BUCINI, G., GUILBERT, J., & BECKAGE, B. (2015). AN INTERACTIVE LAND USE TRANSITION AGENT-BASED MODEL (ILUTABM): ENDOGENIZING HUMAN-ENVIRONMENT INTERACTIONS IN THE WESTERN MISSISQUOI WATERSHED. LAND USE POLICY, 49, 161-176.

TUN, H. M., CAI, Z., & KHAFIPOUR, E. (2016). MONITORING SURVIVABILITY AND INFECTIVITY OF PORCINE EPIDEMIC DIARRHEA VIRUS (PEDV) IN THE INFECTED ON-FARM EARTHEN MANURE STORAGES (EMS). FRONTIERS IN MICROBIOLOGY, 7, 265.

UK COMMITTEE ON CLIMATE CHANGE. (2017). UK CLIMATE CHANGE RISK ASSESSMENT 2017. RETRIEVED FROM

HTTPS://WWW.THECCC.ORG.UK/TACKLING-CLIMATE-CHANGE/PREPARING-FOR-CLIMATE-CHANGE/UK-CLIMATE-CHANGE-RISK-ASSESSMENT-2017/.

UN. (2015). WORLD POPULATION PROSPECTS KEY FINDINGS AND ADVANCE TABLES (9). RETRIEVED FROM

UNFAO. (2017). AGRO-MAPS. RETRIEVED FROM: HTTP://KIDS.FAO.ORG/AGROMAPS/

UNFAO. (2017). FOOD BALANCE SHEETS DATA. RETRIEVED FROM: HTTP://WWW.FAO.ORG/FAOSTAT/EN/#DATA/FBS

UNFAO. (2017). GLOBAL INFORMATION AND EARLY WARNING SYSTEM, FOOD PRICE MONITORING AND ANALYSIS TOOL. RETRIEVED FROM: HTTP://WWW.FAO.ORG/GIEWS/FOOD-PRICES/TOOL/PUBLIC/#/HOME

UNFAO LADA. (2017). GLOBAL LAND DEGRADATION INFORMATION SYSTEM (GLADIS) DATABASE. RETRIEVED FROM: HTTP://WWW.FAO.ORG/NR/LADA/GLADIS/GLADIS_DB/

UNHCR. (2012). THE STATE OF THE WORLD'S REFUGEES: IN SEARCH OF SOLIDARITY. OXFORD UNIVERSITY PRESS RETRIEVED FROM HTTP://WWW.UNHCR.ORG/EN-US/PUBLICATIONS/SOWR/4FC5CECA9/STATE-WORLDS-REFUGEES-2012-SEARCH-SOLIDARITY.HTML.

US CENSUS BUREAU. (2012). COMMODITY FLOW SURVEY, 2012 DATA. RETRIEVED FROM: HTTPS://WWW.CENSUS.GOV/ECON/CFS/

USAID. (2016). DEMOGRAPHIC AND HEALTH SURVEYS (DHS) DATA. RETRIEVED FROM: HTTP://DHSPROGRAM.COM/DATA/AVAILABLE-DATASETS.CFM

USDA APHIS. (2006). REFERENCE OF SWINE HEALTH AND HEALTH MANAGEMENT PRACTICES IN THE UNITED STATES. RETRIEVED FROM

USDA FSIS. (2015). MEAT, POULTRY AND EGG PRODUCT INSPECTION DIRECTORY. RETRIEVED FROM HTTP://WWW.FSIS.USDA.GOV/WPS/PORTAL/FSIS/TOPICS%20/INSPECTION/MPI-DIRECTORY

USDA PSD. (2017). Total Acres Harvested (World), 13 Primary Crops. 1960- 2016 Marketing Years. Retrieved from https://community.endnote.com/t5/EndNote-Styles-Filters-and/Institutional-author-formatting/td-p/5962

Valin, H., Sands, R. D., Van der Mensbrugghe, D., Nelson, G. C., Ahammad, H., Blanc, E., . . . Heyhoe, E. (2014). The future of food demand: understanding differences in global economic models. Agricultural Economics, 45(1), 51-67.

Van der Vorst, J. G., Beulens, A. J., & van Beek, P. (2000). Modelling and simulating multi-echelon food systems. European journal of operational research, 122(2), 354-366.

Van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2013). Yield gap analysis with local to global relevance—a review. Field Crops Research, 143, 4-17.

Vermeulen, S. J., Campbell, B. M., & Ingram, J. S. (2012). Climate change and food systems. Annual review of environment and resources, 37, 195-222.

Vespignani, A. (2012). Modelling dynamical processes in complex socio-technical systems. Nature physics, 8(1), 32.

Volterra, V. (1927). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi: C. Ferrari.

Vranken, L., Avermaete, T., Petalios, D., & Mathijs, E. (2014). Curbing global meat consumption: Emerging evidence of a second nutrition transition. Environmental science & policy, 39, 95-106.

Walker, P. (1983). Crop losses: The need to quantify the effects of pests, diseases and weeds on agricultural production. Agriculture, Ecosystems and Environment, 9(2), 119-158.

Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications (Vol. 8): Cambridge university press.

Watts, D. J., Muhamad, R., Medina, D. C., & Dodds, P. S. (2005). Multiscale, resurgent epidemics in a hierarchical metapopulation model. Proceedings of the National Academy

OF SCIENCES OF THE UNITED STATES OF AMERICA, 102(32), 11157-11162.

WATTS, D. J., & STROGATZ, S. H. (1998). COLLECTIVE DYNAMICS OF 'SMALL-WORLD' NETWORKS. NATURE, 393(6684), 440-442.

WEBB, C. R. (2005). FARM ANIMAL NETWORKS: UNRAVELING THE CONTACT STRUCTURE OF THE BRITISH SHEEP POPULATION. PREVENTIVE VETERINARY MEDICINE, 68(1), 3-17.

WEST, P. C., GERBER, J. S., ENGSTROM, P. M., MUELLER, N. D., BRAUMAN, K. A., CARLSON, K. M., . . . SIEBERT, S. (2014). LEVERAGE POINTS FOR IMPROVING GLOBAL FOOD SECURITY AND THE ENVIRONMENT. SCIENCE, 345(6194), 325-328. DOI:10.1126/SCIENCE.1246067

WHEELER, T., & VON BRAUN, J. (2013). CLIMATE CHANGE IMPACTS ON GLOBAL FOOD SECURITY. SCIENCE, 341(6145), 508-513. DOI:10.1126/SCIENCE.1239402

WHEELIS, M., CASAGRANDE, R., & MADDEN, L. V. (2002). BIOLOGICAL ATTACK ON AGRICULTURE: LOW-TECH, HIGH-IMPACT BIOTERRORISM: BECAUSE BIOTERRORIST ATTACK REQUIRES RELATIVELY LITTLE SPECIALIZED EXPERTISE AND TECHNOLOGY, IT IS A SERIOUS THREAT TO US AGRICULTURE AND CAN HAVE VERY LARGE ECONOMIC REPERCUSSIONS. BIOSCIENCE, 52(7), 569-576.

WHITE, S. H., DEL REY, A. M., & SÁNCHEZ, G. R. (2007). MODELING EPIDEMICS USING CELLULAR AUTOMATA. APPLIED MATHEMATICS AND COMPUTATION, 186(1), 193-202.

WICHELNS, D., & OTHERS. (2007). LOOKING AHEAD TO 2050: SCENARIOS OF ALTERNATIVE INVESTMENT APPROACHES. WATER FOR FOOD, WATER FOR LIFE: A COMPREHENSIVE ASSESSMENT OF WATER MANAGEMENT, 91-145.

WILSON, K. J., FALKINGHAM, J., MELLING, H., & DE ABREU, R. (2004). SHIPPING IN THE CANADIAN ARCTIC: OTHER POSSIBLE CLIMATE CHANGE SCENARIOS. PAPER PRESENTED AT THE GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 2004. IGARSS'04. PROCEEDINGS. 2004 IEEE INTERNATIONAL.

WILTSHIRE, S. W. (2018). USING AN AGENT-BASED MODEL TO EVALUATE THE EFFECT OF PRODUCER SPECIALIZATION ON THE EPIDEMIOLOGICAL RESILIENCE OF LIVESTOCK PRODUCTION NETWORKS. PLOS ONE, 13(3), E0194013.

WINDRUM, P., FAGIOLO, G., & MONETA, A. (2007). EMPIRICAL VALIDATION OF AGENT-BASED MODELS: ALTERNATIVES AND PROSPECTS. JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION, 10(2), 8.

WORLD BANK. (2016). SUBNATIONAL POPULATION DATABASE. RETRIEVED FROM: HTTPS://DATA.WORLDBANK.ORG/DATA-CATALOG/SUBNATIONAL-POPULATION

WU, T., BLACKHURST, J., & O'GRADY, P. (2007). METHODOLOGY FOR SUPPLY CHAIN DISRUPTION ANALYSIS. INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH, 45(7), 1665-1682.

XUE, F., SANDERSON, A. C., & GRAVES, R. J. (2003). PARETO-BASED MULTI-OBJECTIVE DIFFERENTIAL EVOLUTION. PAPER PRESENTED AT THE THE 2003 CONGRESS ON EVOLUTIONARY COMPUTATION, 2003. CEC'03.

YI, R., GATCHALIAN, K. M., & BICKEL, W. K. (2006). DISCOUNTING OF PAST OUTCOMES. EXPERIMENTAL AND CLINICAL PSYCHOPHARMACOLOGY, 14(3), 311.

YOUNG, R., ONSTAD, C., BOSCH, D., & ANDERSON, W. (1989). AGNPS: A NONPOINT-SOURCE POLLUTION MODEL FOR EVALUATING AGRICULTURAL WATERSHEDS. JOURNAL OF SOIL AND WATER CONSERVATION, 44(2), 168-173.

YU, Q., WU, W., YANG, P., LI, Z., XIONG, W., & TANG, H. (2012). PROPOSING AN INTERDISCIPLINARY AND CROSS-SCALE FRAMEWORK FOR GLOBAL CHANGE AND FOOD SECURITY RESEARCHES. AGRICULTURE, ECOSYSTEMS & ENVIRONMENT, 156, 57-71.

ZIA, A., BOMBLIES, A., SCHROTH, A. W., KOLIBA, C., ISLES, P. D., TSAI, Y., . . . TURNBULL, S. (2016). COUPLED IMPACTS OF CLIMATE AND LAND USE CHANGE ACROSS A RIVER–LAKE CONTINUUM: INSIGHTS FROM AN INTEGRATED ASSESSMENT MODEL OF LAKE CHAMPLAIN'S MISSISQUOI BASIN, 2000–2040. ENVIRONMENTAL RESEARCH LETTERS, 11(11), 114026.

APPENDICES

7.1 Appendix 1 – ODD+D Protocol: Regional U.S. Hog Production Network Biosecurity Model v.0.8

Serge Wiltshire, Food Systems Department, University of Vermont

9/28/2017

The following model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm et al. 2006[1], 2010[2]), with ODD+D amendments as proposed by Müller et al. (2012[3]).

I) Overview

Purpose

The Regional U.S. Hog Production Network Biosecurity Model (RUSHPNBM), version 0.8, is an agent-based model developed to assess both supply chain network level and human-behavioral factors relevant to the spread of socioeconomically-important diseases through the U.S. hog production chain. RUSHPNBM has been developed using AnyLogic 7 software, which uses the

---

[1] Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, Goss-Custard J, Grand T, Heinz SK, Huse G, Huth A, Jepsen JU, Jørgensen C, Mooij WM, Müller B, Pe'er G, Piou C, Railsback SF, Robbins AM, Robbins MM, Rossmanith E, Rüger N, Strand E, Souissi S, Stillman RA, Vabø R, Visser U, DeAngelis DL. (2006). A standard protocol for describing individual-based and agent-based models. Ecological Modelling 198:115-126.

[2] Grimm V, Berger U, DeAngelis DL, Polhill G, Giske J, Railsback SF. (2010). The ODD protocol: a review and first update. Ecological Modelling 221: 2760-2768.

[3] Müller, B., Angermueller, F., Drees, R., Dressler, G., Groeneveld, J., Klassert, C., ... & Schwarz, N. (2012). Describing Human Decisions in Agent-Based Social-Ecological Models-ODD+ D an Extension of the ODD Protocol. Available at SSRN 2044736.

Java programming language. Model calibration was undertaken via available statistical datasets coupled with an iterative expert informant advisory process. The model generates realistic production chain networks of producers, feed mills, and slaughter plants at the spatial scale of 431,200 square kilometers. The epidemiological spread model is of the Susceptible Infective (SI) type, with infections transmitted between agents probabilistically based on patterns of trade and contact. Disease spread probabilities associated with the different types of inter-agent contact have been calibrated by reference to epidemiological data concerning disease spread dynamics associated with previous real-world epidemic events in the hog industry, as well as input from livestock veterinarians.

The model was designed for use by university researchers, industry practitioners, veterinary specialists, and government agencies wishing to analyze the dynamics and consequences of disease spread in the U.S. hog production chain under varying assumptions concerning disease characteristics, production chain network structures, and implementation of biosecurity measures and agent behaviors that may prevent or curb catastrophic outbreaks.

Entities, State Variables, and Scales

Three classifications of hog production chain network agents, identified by industry experts as critical players in the transmission of disease, are represented in the model. These are (a) producers, (b) feed mills, and (c) slaughter plants. Producer agents are assigned one of five industry roles based on the USDA's classification system for hog producers, these being (a) Farrow to Wean, (b) Wean to Feeder (a.k.a. Nursery), (c) Feeder to Finish (a.k.a. Finish Only), (d) Farrow to Feeder, and (e) Farrow to Finish. Figure 1 below shows each agent type, its graphical representation in the model, and an outline of the heuristics that govern inter-agent contact patterns. Tables 1-2 below show the agent-level attributes and baseline parameters used in the simulations.
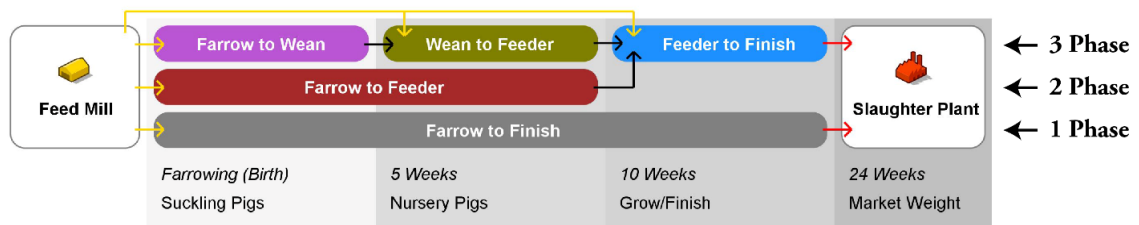
Figure 1: Structure of connections between agents, including hoofstock age transfer conditions where applicable.

309

Table 1.1: Parameters and variables common to all agents

Attribute	Description
Static parameters (set at initialization)	
My name	String representation encoding agent class and index (for tracking network connections)
State variables (may change throughout simulation run)	
Infectivity state	Either "clean" or "infected".
Has been infected	Flag indicating whether the agent was ever infected during a run.
Contact network out degree & in degree	List of the other agents with whom each agent had contact throughout a run, as well as the number of times contact between the two agents occurred. Out degree is incremented whenever an agent sends animals or feed; in-degree whenever animals or feed are received.
Infection network out degree	List of the other agents to whom an agent spread the disease, along with the number of times infection spread occurred between the two agents throughout a run.

Table 1.2: Parameters and variables for producer agents

Attribute	Description
Static parameters (set at initialization)	
Farm category	Encodes which of the 5 producer industry roles this agent falls into.
Total capacity	Total animal capacity.
My transferee producers	List of potential producer trading partner agent objects.
My slaughter plant	Link to slaughter plant agent object (finishing producers only).
My feed mill	Link to feed mill agent object.
State variables (may change throughout simulation run)	
Current pig inventory	Number of pigs currently on premises.
Pig batch tracker	Pigs are represented by a data structure encoding the size of each batch of stock currently making up the

310

	agent's hoofstock inventory, along with a timestamp representing that pig batch's "birthday", which is used to calculate the age of the batch.
Pig shipments in & out	Lists of sizes (number of animals) of each shipment incoming from and outgoing to the producer (used for calibration).
Feed deliveries in	Number of feed deliveries incoming to the producer (used for calibration).

Table 1.3: Parameters and variables for feed mill agents

Attribute	Description
Static parameters (set at initialization)	
My transferee producers	List of producer agent objects within service area.
State variables (may change throughout simulation run)	
Truck infected	Flag indicating whether the feed mill's delivery truck is currently infected.
Feed deliveries out	Number of outgoing feed deliveries to producers (used for calibration).

Table 1.4: Parameters and variables for slaughter plant agents

Attribute	Description
Static parameters (set at initialization)	
My transferee producers	List of producer agent objects within service area.
State variables (may change throughout simulation run)	
Pig shipments in	List of sizes (number of animals) of each shipment incoming from producers (used for calibration).

311

Table 2: Parameters remaining fixed throughout each model run

Parameter	Baseline Value(s)	Source
Network Makeup		
Area of network region (km2)	431,200	
Number of producer agents in the model	750	[4]
Number of livestock per producer [normal distribution, rounded to integer]	$\mu = 1{,}000; \sigma = 300; x \geq 50$	XE, ID, [5]
Number of producer production phases	[1, 2, 3]	
Number of slaughter plant agents in the model	3	[6]
Number of feed mill agents in the model	10	XE, ID
Epidemiological Characteristics		[7]
Suckling pig mortality rate	0.95	
Nursery pig mortality rate	0.6	
Grow/finish hog mortality rate	0.1	
Length of producer infection (days) [triangular distribution]	$\mu = 30; 0 \leq x \leq 60$	
Length of slaughter plant contamination (days) [triangular distribution]	$\mu = 5; 0 \leq x \leq 10$	
Farrowing		
Frequency of farrowing (days)	30	[8]
Minimum farrowing quantity as a proportion of producer capacity	0.25	XE
Producer–Producer Contact		
Maximum producer to producer connection distance (km)	100	XE, ID
Maximum frequency of pig shipments (/week)	1	XE, ID
Minimum transfer quantity as a proportion of transferee capacity	0.25	XE
Prob. of infection via trailer returning from infected transferee	0.15	XE
Feed Mill–Producer Contact		

[4] FLAPS datasets, generated using tool available at: http://flaps.biology.colostate.edu/
[5] Ibid.
[6] http://usda.mannlib.cornell.edu/usda/nass/LiveSlauSu/2010s/2014/LiveSlauSu-04-21-2014.pdf
[7] https://www.aasv.org/aasv%20website/Resources/Diseases/PorcineEpidemicDiarrhea.php
[8] http://www.thepigsite.com/stockstds/3/pig-farm-targets/

Frequency of feed distribution trips (days)	1	XE, ID
Percent of producers in feed mill service area visited per trip	15	XE, ID
Prob. that truck will be contaminated upon visiting an infected producer	0.15	XE
Prob. that contaminated truck will infect subsequent producers on route	0.15	XE
Producer–Slaughter Plant Contact		
Prob. that infected hogs will contaminate slaughter plant receiving area	0.75	XE
Prob. of infection via truck returning from infected slaughter plant	0.15	XE

Note: "XE" indicates that the value was derived using the best estimates of industry expert informants. "ID" refers to a proprietary industry dataset utilized for calibration.

The model's parameters and functions controlling pig movement and feed deliveries were further specified with the help of data provided by a Family Farm Company from the U.S. (per confidentiality, the company's name is not disclosed here). The database contains two-year records of each pig movement and each feed delivery involving producers in the Family system. The Family Farm Company consists of a network of 161 producer partners that raise pigs from birth to market. The pig movement records were used to derive realistic estimates of transfer frequencies and number of animals per transfer, as well as reinforcing USDA farm size and operational statistics (Table 2; Figure 2). The feed delivery records were used to estimate delivery frequencies (Figure 3).

313

Table 3: Average annual number of pig shipments per agent (producer/supplier/packer/costumer) in the Family Farm Company.

FROM	TO	N shipments FROM	N shipments TO	Avg. pigs/shipment
Wean to Finish	Wean to Finish	1.9	1.6	259
Wean to Finish	Finishing	2.5	1.7	608
Finishing	Finishing	0.5	0.5	176
Nursery	Nursery	1.2	0.8	161
Nursery	Wean to Finish	1.5	1.5	652
Nursery	Finishing	45.7	8.6	512
Supplier	Nursery	156.6	41.2	582
Supplier	Wean to Finish	60.9	10.8	406
Nursery	Customer	0.8	4	3
Finishing	Packer Customer	27.9	276.6	153
Wean to Finish	Packer Customer	22.4	179.5	153
Nursery	Packer Customer	0.5	1	2

Note: The columns list the following information. "FROM": the type of shipping agent (producer/supplier). "TO": receiving agent (producer/packer/costumer). "N shipments FROM": average annual shipments per shipping agent. "N shipments TO": average annual shipments per receiving agent. "Avg. pigs/shipment": average number of pigs per shipment.
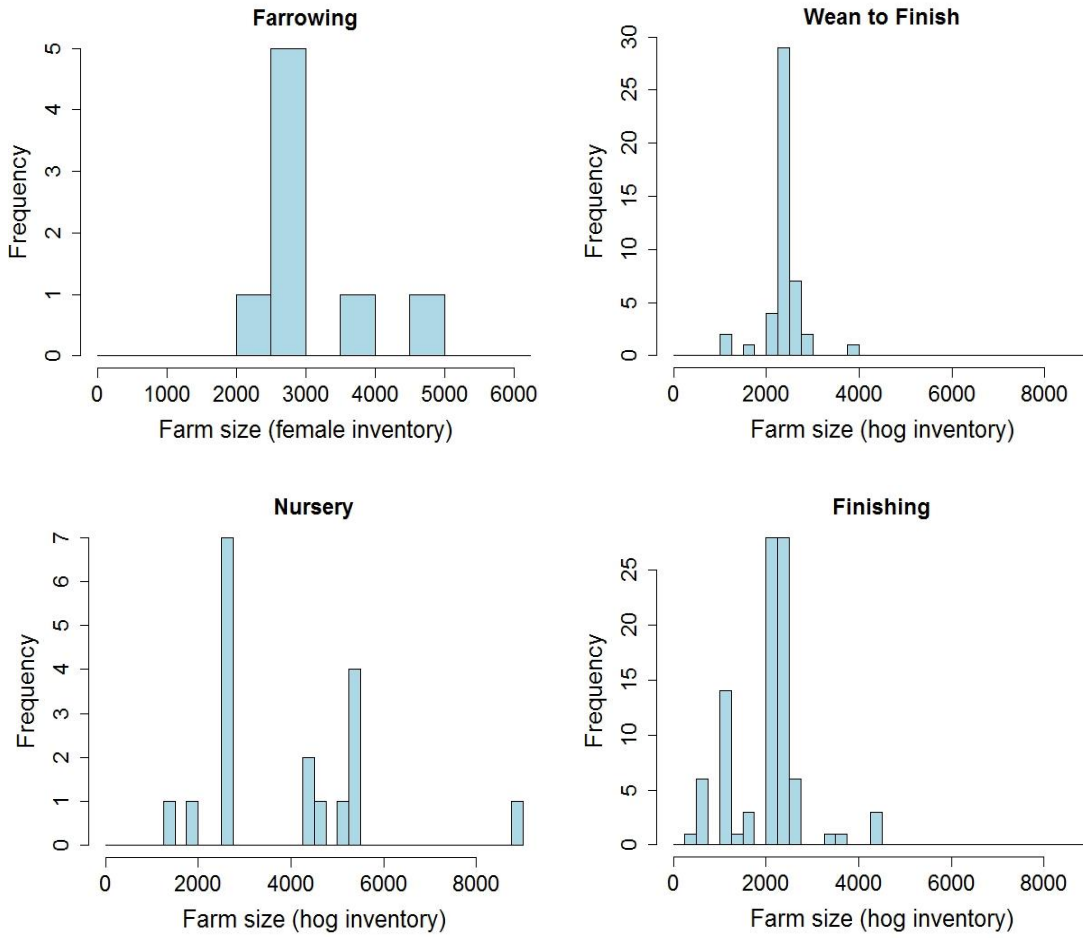
Figure 2: Farm size distribution by farm type (farrowing, wean-to-finish, nursery, finishing) in the Family Farm Company.
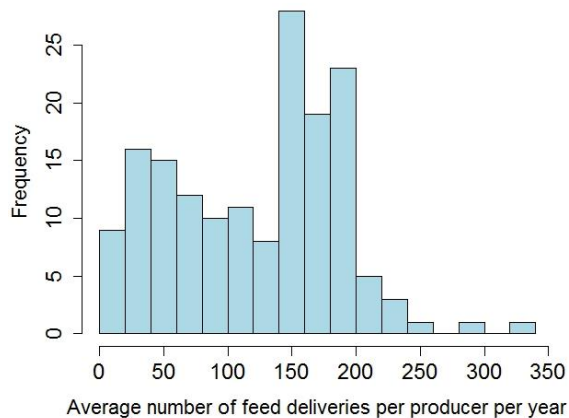
Figure 3: Average number of feed deliveries per producer per year in the system of Family Farm Company.

Finally, a team of experts in veterinary medicine and in agent-based modeling has followed the development of the model and collaborated in parametrizing, calibrating and ground truthing it:

- Julie Smith (DVM, PhD, https://asci.uvm.edu/?Page=faculty/smith/homepage.html)
- Steve Dritz (DVM, PhD, Swine Specialist, https://www.vet.k-state.edu/education/dmp/faculty-staff/faculty/dritz/)
- Asim Zia (PhD, http://www.uvm.edu/~azia/)
- Christopher Koliba (PhD, http://www.uvm.edu/~ckoliba/index.htm)

How is space included in the model?

The model is spatially situated in a continuous, two-dimensional environment representing 880 x 490 km. Distances between agents are calculated "as the crow flies." In some cases, distance is a factor in determining inter-agent contact patterns (detailed below).

How is time represented in the model?

The model's time scale is based on real-world days, with the initial model date set to January 1st, 2012. 2012 was chosen because FLAPS initialization data are drawn from the 2012 USDA Census of Agriculture. The model's stop date can be set as desired depending on the experimental phenomena the user is interested in studying, with a default setting of January 1st, 2022.

316

<u>Process Overview and Scheduling</u>

All event scheduling in the model follows a Last-In-First-Out (LIFO) protocol. Four classes of functions define the operation of the model, in order of the point(s) in the simulation that they occur. First are the initialization functions, which define how the agents will be physically situated in the space, set each agent's individual parameters, and identify lists of potential trading partners based on the classification and industry role of the agent, as well as spatial proximity to other agents. Second are the cyclically-executing functions, which make up the agents' decision rules, determining how and when contact between agents will occur (through the transfer of livestock and the distribution of feed), and thereby opening potentials for infection to spread. These functions also determine and implement the consequences of an infection upon the agent. Third is the initial infection function, which occurs after the initial transient period. Finally, fourth are the set of functions facilitating the output of model data for further analysis, including post-experiment scripts to parse model outputs and analyze results across multiple runs.

II) Design Concepts

<u>Theoretical and Empirical Background</u>

Because real-world epidemics are fundamentally phenomena which propagate through networks (social, business, transportation, etc.), the formulation of a suitably-realistic network structure within which agents operate is a fundamental basic principle of the model. A corollary to this basic principle concerns the model's balance between context specificity and analytic transparency. The model's network generation algorithm strives to maintain sufficient context specificity to capture the critical complexities underpinning observed epidemiological spread phenomena, while bracketing superfluous elements of real-world production chain networks that have not been implicated in previous epidemiological events. For example, the model contains only feed mill, producer, and slaughter plant agent typologies, because these were identified by industry experts as the critical players underpinning disease spread. Whereas in real-world hog production chain networks there may be a multitude of other actor typologies (i.e. auction houses, equipment suppliers, construction contractors, insurance agents, and many more), these were intentionally excluded from the model's design to simplify analysis.

Another guiding principle is the spatial framework of the model. In many epidemiological studies, agent density has been shown to impinge directly upon spread characteristics. With high enough density, complex phenomena such as percolation thresholds may emerge. To study this, the model was designed with the ability to flexibly change the density of agents in space.

On what assumptions is/are the agents' decision model(s) based?

The primary assumptions driving agent behavior relate to trade patterns associated with the industry role each agent plays. For example, it is assumed that, as soon as their livestock batches reach the transfer age appropriate for their industry role, producer agents will search the agent space for appropriate trading partners.

Another assumption concerns the characteristic distance over which agents may interact. This can be adjusted by tuning the maximum connection distance parameter. Since specific spatial location data were not available in either the USDA statistics or the Family Farm System dataset, baseline values were estimated in consultation with industry experts.

Learning

The agents' decision rules remain non-adaptive at this stage of model development. The agents' action heuristics are based on their industry roles, and are designed to realistically replicate throughput in the production chain system as a whole. Thus, an agent will transfer hoofstock to an appropriate trading partner as soon as possible, farrowing will proceed regularly wherever a producer has sufficient excess capacity, and feed deliveries will take place at a set frequency. The agents' behavior does not change as a result of model conditions, for example the presence of a disease within the network, however each agent will necessarily adapt to conditions resulting from the factors such as the number of other agents in the space, or the currently-available spare capacity of its trading partners. In future model versions, adaptive agent behavior will be implemented to reflect the decision-making heuristics of real-world hog producers, as identified through data gathering efforts presently underway.

Collectives

Livestock in the model may be considered as collectives, as they are encoded in batches of animals of the same age, and with the same infectivity status. If a producer is infected, it is assumed that all livestock on the premises become infected. This simplifying assumption follows from veterinary reports on the virility of PEDv, which tends to quickly sweep through entire herds. From a standpoint of practicality, encoding livestock in this manner was also desirable because it significantly reduces the computational time required for each run.

In addition, while not defined explicitly as such, groups of agents in the model exhibit emergent collective characteristics due to their distribution within the model's spatial framework. For example, in densely-packed areas, groups of agents tend to interact heavily within connected clusters, leading to localized disease outbreaks. This type of collective behavior is not directly

318

imposed, but rather emerges from interactions between the model parameterization and agents' fixed behavioral rules.

Heterogeneity

As described in the *Entities, state variables, and scales* section above, agents fall into three main categories: (a) producers, (b) slaughter plants, and (c) feed mills. Producer agents are assigned one of five industry roles based on the USDA classification system for hog producers. An agent's industry role determines the initial age of its hoofstock, its hoofstock age transfer condition, as well as its set of potential trading partners. These relationships are visualized in Figure 1. Agents' decision-making heuristics also vary according to their class. For example, a farrow-to-wean producer will only send pigs to wean-to-feeder producers.

Stochasticity

RUSHPNBM uses stochasticity for initialization of agent locations and parameters, as well as for controlling infection spread. A random seed is used, such that all runs are different. These stochastic features ensure that the contact patterns that unfold in each model run are never repeated.

Draws from distribution functions (normal, uniform, and triangular) are utilized in some cases. For example, the age of the pig groups initially associated with each producer are drawn from a uniform distribution bounded according to the producer's industry role. Triangular distributions underlie the time agents will remain infected before transitioning back to the susceptible state.

Stochasticity is also used in all disease-spread events. Uniform probability distributions returning "true" if a randomly-drawn value between zero and one is less than *p* are used to determine if the infection will spread whenever contact between a susceptible and an infected agent occurs. Different probability values are used for each mode of transmission.

Observation

The model tracks in real-time the current hoofstock inventory of all producers in the model, the number of currently infected hoofstock, the number of currently infected producers, and the cumulative number of infected producers, which can be output as time-series data to examine infection-spread dynamics. The total infection duration is also recorded.

In addition, we track the flow of feed and livestock between different types of agents in order to calibrate model parameters to reflect real-world data, for example the distribution of hog shipment sizes and delivery frequencies characteristic of real hog supply chain networks.

319

Finally, a contact network adjacency matrix with link weights encoding the number of times each agent interacted throughout the model run is exported as tabular data after each run, and later parsed using a series of Python functions. An infection-spreading network is similarly tracked, output, and parsed. Key statistics on trade and infectivity patterns across a series of model runs—both at the individual agent as well as the whole-network level—may then be analyzed.

<u>Emergence</u>

Emergent phenomena in the present model occur as a result of the interaction between agents' behavioral heuristics and structural elements of the model, for example the network configuration and disease spread characteristics specified by the user. This could take the form of differential spread characteristics—such as in the observation of percolation thresholds—resulting from user-input parameters concerning network makeup, probabilities, or duration parameters.

III) Details

<u>Implementation Details</u>

The model was implemented using AnyLogic version 7 software, which relies upon the Java programming language for all scripts and functions. The sections below use pseudocode to describe in detail the algorithmic structures underlying each model function.

Notes on pseudocode used in this document:

- The characters "//" will be used to designate a comment (i.e., the line of text following the "//" is not part of the actual function logic).
- Parameters referenced in all functions refer to those associated with the agent object from which a function has been called. In some cases, to disambiguate, the terms "self" or "my" may be used to refer to the function-calling agent object or its associated parameters.
- "`ADD OR INCREMENT [sender] in [receiver]'s [network edge list]`" is defined here as:
  ```
  IF [sender] is not in [receiver]'s [network edge list] ADD [sender] to
          [receiver]'s [network edge list] with contact counter set to 1
  ELSE INCREMENT contact counter associated with [sender] in [receiver]'s
          [network edge list] by 1
  ```
- "`RANDOM DRAW using [probability]`" is defined here as the Boolean value resulting from:

```
(DRAW random number from uniform distribution between 0 and 1) <
      [probability]
```

Is the model accessible, and if so where?

While the raw source code for the model is not accessible, the pseudocode below may be used as a guide to understand the model's structure and logic with a high level of detail.

<u>Initialization</u>

The model is initialized by progressing through a series of initialization functions. Agent parameters such as location are set as each agent object is generated by the model. Next, producer agents initialize their operational parameters and define their networks of potential trading partners to be referenced throughout the model run.

Initialize agent locations:
All agents are placed at a random location in the continuous 2D space.

```
FOR EACH agent object
    SET x coordinate to RANDOM INTEGER between 0 and 880
    SET y coordinate to RANDOM INTEGER between 0 and 490
```

Producer agent initialize category function:
The initialize category function sets the industry role of producers according to the specialization level to be evaluated.

```
IF (specialization level = "low")
    SET farm category to "farrow to finish"
ELSE IF (specialization level = "medium")
    SET farm category to RANDOM DRAW from ["farrow to feeder", "feeder
          to finish"]
ELSE IF (specialization level = "high")
    SET farm category to RANDOM DRAW from ["farrow to wean", "wean to
          feeder", "feeder to finish"]
```

Producer agent initialize farm function:
The initialize farm function sets the producer's maximum capacity and adds one pig batch equal to this capacity to the pig batch tracker, with an age corresponding to the producer's industry role.

```
SET total capacity to MAX of 50 and (ROUND to INTEGER (DRAW from normal
      distribution with μ = 1,000 and σ = 300))
```

```
ADD pig batch to pig batch tracker with size equal to total capacity AND
    birthday equal to to a random integer between the maximum and
    minimum age of a pig for the agent's industry role
SET current inventory to total capacity
```

Producer agent initialize network function:

Once the agents' locations and operational parameters have been initialized, a network initialization function generates a set of potential trading partners. All producer agents are assigned to the nearest feed mill, and finishing producers are also assigned to the nearest slaughter plant. A pool of potential transferee producers is also generated for each non-finishing producer according to their industry role. These relationships are shown in Figure 1. The potential transferee producers in this pool are limited by the maximum producer-to-producer connection distance parameter.

```
SET potential farms list to (SORT by distance (FILTER other producer
    agents s.t. (industry role of other producer is the next step in
    the production chain) AND (distance to the other producer <= max
    producer-producer connection distance global parameter)))
IF industry role is a finishing type
    SET my slaughter plant to closest slaughter plant
SET my feed mill to closest feed mill
ADD self to my feed mill's "links to farms" list
```

Is the initialization always the same, or is it allowed to vary among simulations?

The initialization of the spatial location, operational characteristics, and potential trading partners for each agent, and initial livestock ages differ between runs. However, the distributions of from which these values are drawn, as well as the basic heuristics controlling the behavior of each type of agent, do not change.

Are the initial values chosen arbitrarily or based on data?

Initialization parameters rely upon several datasets, including the University of Colorado / USDA FLAPS system, USDA NASS data, USDA APHIS data, Google Maps queries, proprietary industry datasets, and expert input. For details, see the Entities, State Variables, and Scales section above.

<u>Submodels</u>

Producer agent cyclically-executing functions:
Farrow, wean, and batch piglets function:

If a farm that farrows piglets (Farrow to Wean, Farrow to Feeder, or Farrow to Finish types) is left with excess capacity after a livestock transfer, a periodic farrowing function fills that capacity with a new batch of piglets, whose birthday is set to the current model day. Once again, to eliminate unrealistically-small pig groups, a minimum farrowing size as a proportion of the farrowing farm's total capacity is required for the farrowing function to proceed. Thus, a farm which is already almost at maximum capacity will not farrow a new batch of piglets until another batch has been shipped to an appropriate trading partner.

```
**Recurrence time is the frequency of farrowing global parameter**
IF industry role is a farrowing type
    Number to wean and batch = remaining pig capacity
    IF (number to wean and batch >= (minimum farrowing batch proportion
            * my capacity))
        ADD number to wean and batch and birthday = current day to pig
                batch tracker
        INCREMENT pig inventory by batch size
```

Ship to transferee farms function:

Non-finishing producers transfer hoofstock to a transferee farm when the hoofstock reach the age corresponding to the transfer condition associated with the industry role of the producer. This function periodically evaluates whether the transfer age requirement of a pig batch has been met. If so, the producers in the transferring producer's pool of possible trading partners are sequentially evaluated to determine whether they are able to receive the shipment. To eliminate the transfer of unrealistically-small groups of livestock, transfers will only proceed if the pig batch size exceeds the minimum transfer quantity, as a proportion of the transferee's total capacity.

If the transferring producer is infected but the transferee is not, the transferred hoofstock will automatically spread the infection to the transferee producer. If the transferee producer is infected but the transferring producer is not, the "delivery trailer" returning from the infected transferee producer may infect the transferring producer according to a probability set at model initialization.

The birthday parameter associated with the batch of transferred stock is maintained as it is passed to the transferee, such that the pig batch will once again be appropriately transferred to the next production phase at the correct transfer age. In the rare case that a pig batch exceeds

the slaughtering age before a suitable transferee producer could be located, it is culled to make room for a new batch of pigs.

```
    **Recurrence time is the maximum frequency of pig shipments global
        parameter**
IF industry role is NOT a finishing type
    FOR EACH pig batch meeting age transfer requirement
        FOR EACH transferee in my transferee producers
            IF (batch size <= transferee's spare capacity) AND (batch size
                    >= transferee's minimum batch size)
                IF (transferee's infectivity state is "infected") AND
                        (infectivity state is "clean")
                    DECREMENT batch size according to mortality rate global
                            parameter associated with pigs' age
                REMOVE pig batch from pig batch tracker
                DECREMENT pig inventory by batch size
                ADD pig batch and birthday to transferee's pig batch
                        tracker
                INCREMENT transferee's pig inventory by batch size

                // update contact network trackers
                ADD OR INCREMENT transferee in contact network out-degree
                        list
                ADD OR INCREMENT self in transferee's contact network in-
                        degree list

                // update pig shipment trackers
                ADD batch size to pig shipments out list
                ADD batch size to transferee's pig shipments in list

                // infection brought to transferee via infected pigs
                IF infectivity state is "infected"
                    SET transferee's infectivity state to "infected"
                    ADD OR INCREMENT transferee in infection-spreading
                            network out degree list

                // infection brought home via trailer from transferee farm
                IF (transferee's infectivity state is "infected") AND
                        (RANDOM DRAW using Prob. pig truck will become
                        contaminated if producer is infected) AND (RANDOM
                        DRAW using Prob. producer will become infected if
                        returning pig truck is contaminated)
```

324

```
                    SET infectivity state to "infected"
                    ADD OR INCREMENT self in transferee's infection-
                            spreading network out degree list


            // cull pigs that are too old and were never able to be transferred
            FOR EACH pig batch over 168 days old REMOVE batch from pig batch
                    tracker
```

Ship to slaughter plant function:

Finishing producers (Feeder to Finish and Farrow to Finish types) ship hoofstock to either an auction house (as described above), or directly to their slaughter plant, as soon as the hoofstock reach the designated slaughtering age. If the transferring producer is infected, the receiving area of the slaughter plant may become contaminated according to a probability set at model initialization. If the receiving area of the slaughter plant is already contaminated, the "delivery trailer" returning to the transferring producer may carry the infection back to that producer according to another probability set at model initialization.

```
        IF industry role is a finishing type
            FOR EACH pig batch meeting age transfer requirement
                REMOVE pig batch from pig batch tracker
                DECREMENT pig inventory by batch size

                // update contact network trackers
                ADD OR INCREMENT slaughter plant in contact network out-degree
                        list
                ADD OR INCREMENT self in slaughter plant's contact network in-
                        degree list

                // update pig shipment trackers
                ADD batch size to pig shipments out list
                ADD batch size to slaughter plant's pig shipments in list

                // infection brought to slaughter plant via infected pigs
                IF (infectivity state is "infected") AND (RANDOM DRAW using Prob.
                        slaughter plant receiving area will become infected if pig
                        batch is infected)
                    SET slaughter plant's infectivity state to "infected"
                    ADD OR INCREMENT slaughter plant in infection-spreading
                            network out degree list
```

325

```
              // infection brought home via trailer from slaughter plant
              IF (slaughter plant's infectivity state is "infected") AND
                    (RANDOM DRAW using prob. pig truck will become
                    contaminated if receiving area is infected) AND (RANDOM
                    DRAW using prob. producer will become infected if
                    returning pig truck is contaminated)
                SET infectivity state to "infected"
                ADD OR INCREMENT self in slaughter plant's infection-spreading
                    network out degree list
```

Feed mill agent cyclically-executing functions

Feed mills periodically generate delivery routes encompassing a subset of producers within their latent feed-mill-to-producer link set. Each route encompasses a subset of the producers in the feed mill's service area, with the number of stops in each trip resulting from a draw from a Poisson distribution. While there is no actual "feed truck" object in the model, the logic of the following function is based on the way such a truck would move between agents and possibly spread disease.

Beginning from the mill, this conceptual feed truck will visit the previously-drawn number of randomly-selected producers within the feed mill's service area before finally returning to the feed mill. If the feed mill is infected, the truck may be contaminated initially. Should the truck encounter an infected producer on its route, it may become contaminated at that point. Once a truck is contaminated, the infection may be spread to subsequent producers on the route. If a contaminated truck returns to the feed mill, the mill itself may become infected.

Distribute feed function:

```
        **Recurrence time is the frequency of feed deliveries global parameter**
        // generate delivery route
        FOR number of producers in service area * percent of producers in feed
              mill service area visited per trip
          ADD random producer in service area (that is not already in delivery
                route list) to delivery route list

        // parse infectivity consequences of delivery route
        FOR EACH producer in delivery route list
          // update contact network trackers
          ADD OR INCREMENT producer in contact network out-degree list
          ADD OR INCREMENT self in producer's contact network in-degree list

          //infected truck infects farm it's delivering to
```

```
IF (truck infected is true) AND (RANDOM DRAW using prob. producer
      will become infected if feed truck is contaminated)
    SET producer's infectivity state to "infected"
    ADD OR INCREMENT producer in infection-spreading network out
        degree list


//truck becomes infected from delivery to infected farm
IF (producer's infectivity state is "infected") AND (RANDOM DRAW
      using prob. feed truck will become contaminated if producer is
      infected)
    SET truck infected to "true"
```

Initial infection function

The system is initialized with all agents free of infection. After one model year has passed, an infection is introduced to a random producer agent. The reason for the one-year lag is to skip the transient period and allow the model to stabilize before analyzing the effect of an introduced disease. This lag is necessary because, as in a real production chain, a certain amount of slack, or a difference between the theoretical production capacity and actual production, is characteristic in the modeled production chains. In the model, this economic slack is due to the producers sometimes temporarily operating at less than maximum hoofstock capacity until an appropriate shipment of livestock becomes available. In general, after about 9 months, the level of slack in the model has stabilized.

```
**Function is called only once, after one model year**
SET one randomly chosen producer agent's infectivity state to "infected"
```

Infection control functions

Susceptible/infective state charts:

Each agent has an embedded state chart which encodes its infectivity status, with clean and infected states corresponding to the classical susceptible/infective framework. Should an agent become infected, a function is called which calculates the number of its stock that will die of the disease. The proportion of livestock that succumb to the disease is based on their age, with uniform mortality rates set at model initialization for suckling pigs, nursery pigs, and grow/finish hogs. After die-off is calculated for pig batches of each life stage within an infected producer's inventory, the producer's inventory data are updated accordingly. An agent will remain infected for a duration drawn from a triangular distribution whose mean length in days is controlled by parameters specific to each agent type.
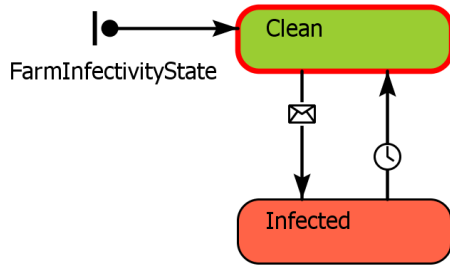
Figure X: Infectivity State Chart

Calculate mortality function:

```
FOR EACH pig batch in pig batch tracker
    DECREMENT pig batch size by global parameter encoding mortality
            proportion appropriate for age of pigs
```

Data output

At the conclusion of each run, the model determines whether a parameter variation experiment is being conducted, and, if so, a line is added to an output csv file containing the number of producers in that run, the total infection duration, the proportion of agents that had been infected, and the specialization level.

If desired, a contact network adjacency matrix with link weights encoding the number of times each agent interacted throughout the model run, as well as the specialization level and repetition number, is generated. This is accomplished by looping through all agents and adding a line to a csv file for each entry in the agent's contact tracker. Subsequent network analysis may then be performed using external scripts.

Finally, for calibration purposes, it is possible to output the flow of feed and livestock between all agents throughout the run. A csv is generated that encodes the size and date of each transfer, as well as the IDs of the sending and receiving agents. Agent parameters such as classification and operational details are exported to another csv. The resulting data may then be compared to real-world data, and parameters tuned so that agents' actions in the model more closely mirror the distribution of hog shipment sizes and delivery frequencies observed in real hog supply chain networks.

7.2 Appendix 2 – ODD+D Protocol: Regional U.S. Hog Production Network Biosecurity Model v.1.2

Serge Wiltshire, Food Systems Department, University of Vermont

The following model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm et al. 2006, 2010), with ODD+D amendments as proposed by Müller et al. (2013).

I) Overview

I.i Purpose

I.i.a What is the purpose of the study?

The Regional U.S. Hog Production Network Biosecurity Model (RUSHPNBM), version 1.2, is an agent-based susceptible / infective model developed to assess both supply chain network level and human-behavioral factors relevant to the spread of socioeconomically-important diseases through regional U.S. hog production chain networks. RUSHPNBM has been developed using AnyLogic v.8 software. The model is calibrated to represent hog production within the U.S. states of North Carolina, Iowa, and Illinois. These three states serve as case studies, since they are all major hog producers, while also having interesting supply chain network features that differentiate them from one another.

Model calibration was undertaken using available datasets coupled with an iterative expert advisory panel process. The model uses agricultural statistics and model calibration tools to generate realistic production chain networks of producers, feed mills, and slaughter plants within the spatial bounds of each study area state.

The epidemiological spread submodel is of the susceptible / infective (SI) type, with infections transmitted between agents probabilistically based on patterns of trade and contact honed through industry expert advisory panels and a review of the primary literature. Disease spread

probabilities associated with the different types of inter-agent contact have been calibrated by reference to epidemiological data concerning disease spread dynamics associated with previous real-world epidemic events in the hog industry, as well as input from livestock veterinary professionals within the expert panel sessions.

I.i.b For whom is the model designed?

The model was designed for use by university researchers, industry practitioners, veterinary specialists, and government agencies wishing to analyze the dynamics and consequences of disease spread in U.S. hog production systems under varying assumptions concerning disease characteristics, production chain network structures, and implementation of biosecurity measures and agent behaviors that may prevent or curb catastrophic outbreaks.

I.ii Entities, State Variables, and Scales

I.ii.a What kinds of entities are in the model?

Three classifications of hog production chain network agents, identified by industry experts as critical players in the transmission of disease, are represented in the model. These are (a) producers, (b) feed mills, and (c) slaughter plants. Producer agents are assigned one of six industry roles. Five of these encompass the USDA's classification system for hog producers, these being (a) Farrow to Wean, (b) Wean to Feeder (a.k.a. Nursery), (c) Feeder to Finish (a.k.a. Finish Only), (d) Farrow to Feeder, and (e) Farrow to Finish. Upon the advice of industry experts, an additional producer classification, (f) Wean to Finish—which has recently become more popular in the industry—is also included in the model. Figure 1 below shows each agent type, its graphical representation, and an outline of the heuristics that govern inter-agent contact patterns.
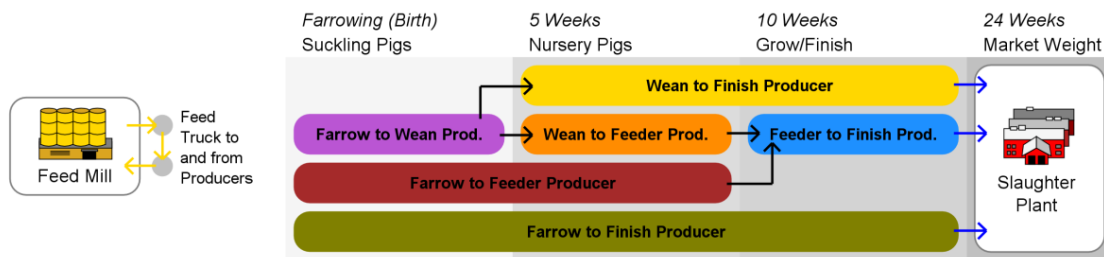
Figure 1: Structure of connections between agents, including livestock age transfer conditions where applicable

I.ii.b By what attributes (i.e. state variables and parameters) are these entities characterized?

Each agent class has a specific set of state variables and parameters relevant to its industry role. These are given in Tables 1–5, below.

Table 1: Parameters and variables common to all agents

Attribute	Description
Static parameters (set at initialization)	
My name	String representation encoding agent class and index (for tracking network connections)
State variables (may change throughout simulation run)	
Infectivity state	Either "clean" or "infected".
Has been infected	Flag indicating whether the agent was ever infected during a run.
Total infection duration	Counter indicating the total number of days an agent was infected during a run.
Contact network out degree & in degree	List of the other agents with whom each agent had contact throughout a run, as well as the number of times contact between the two agents occurred. Out degree is incremented whenever an agent sends animals or feed; in-degree whenever animals or feed are received.
Infection network out degree	List of the other agents to whom an agent spread the disease, along with the number of times infection spread occurred between the two agents throughout a run.

Table 2: Parameters and variables for producer agents

Attribute	Description
Static parameters (set at initialization)	
Farm category	Encodes which of the 6 producer industry roles this agent falls into.
Total capacity	Total animal capacity.
Number of sows	Number of sows (calculated using parameters for the ratio of sows to pigs associated with each producer classification).
Non-sow capacity	Total capacity minus number of sows.
My transferee producers	List of potential producer trading partner agent objects.
My slaughter plant	Link to slaughter plant agent object (finishing producers only).

331

My feed mill	Link to feed mill agent object.

State variables (may change throughout simulation run)

Current piglet inventory	Number of piglets currently available to wean and batch.
Last weaning day	Timecode of the last date piglets were batched and encoded in the pig batch tracker.
Current non-sow pig inventory	Number of pigs (that is, neither piglets or sows) currently on premises.
Pig batch tracker	Non-sow pigs are represented by a data structure encoding the size of each batch of stock currently making up the agent's hoofstock inventory, along with a timestamp representing that pig batch's "birthday", which is used to calculate the age of the batch.
Pig shipments in & out	Lists of sizes (number of animals) of each shipment incoming from and outgoing to the producer (used for calibration).
Feed deliveries in	Number of feed deliveries incoming to the producer (used for calibration).

Table 3: Parameters and variables for feed mill agents

Attribute	Description
State variables (may change throughout simulation run)	
Truck infected	Flag indicating whether the feed mill's delivery truck is currently infected.
Feed deliveries out	Number of outgoing feed deliveries to producers (used for calibration).

Table 4: Parameters and variables for slaughter plant agents

Attribute	Description
State variables (may change throughout simulation run)	
Pig shipments in	List of sizes (number of animals) of each shipment incoming from producers (used for calibration).

I.ii.c What are the exogenous factors/drivers of the model?

Exogenous factors include the spatial and operational distributions of agents of each class within each study area (Table 5), as well as parameters held constant across model runs (Table 6). These exogenous parameters can be broken down into general disease parameters, disease

spread probabilities specific to each modality of inter-agent contact, and parameters pertaining to each agent class.

Table 5: Study area network parameters

Parameter	Baseline Value			Data Source(s)
	North Carolina	*Iowa*	*Illinois*	
Num. producers	2217	6266	2045	Burdett et al. (2015)
Avg. producer capacity	4015	3265	2264	Burdett et al. (2015)
Proportion farrow to wean	0.050	0.026	0.038	Burdett et al. (2015)
Proportion farrow to feeder	0.005	0.010	0.009	Burdett et al. (2015)
Proportion farrow to finish	0.554	0.304	0.635	Burdett et al. (2015)
Proportion wean to feeder	0.102	0.064	0.023	Burdett et al. (2015)
Proportion wean to finish	0.003	0.077	0.055	Burdett et al. (2015)
Proportion feeder to finish	0.286	0.519	0.241	Burdett et al. (2015)
Num. slaughter plants	24	18	25	USDA NASS (2014)
Num. feed mills	40	114	37	Google search; EAP

Table 6: Parameters common to all study areas, remaining fixed throughout each model run

Parameter	Description / Units	Baseline Value	Data Source(s)
Disease parameters			
Percent to infect	Percent of producers initially infected	5%	–
Avg. producer infection length	Avg. duration of producer infection (days)	40	Goede & Morrison (2016); EAP
Avg. slaughter plant infection length	Avg. duration of slaughter plant infection (days)	7	EAP
Avg. feed mill infection length	Avg. duration of feed mill infection (days)	25	EAP
Suckling mortality rate	Proportion of suckling pigs dying if infected	0.9	Goede & Morrison (2016); EAP
Nursery mortality rate	Proportion of nursery pigs dying if infected	0.4	Goede & Morrison (2016); EAP
Grow/finish mortality rate	Proportion of grow/finish hogs dying if infected	0.1	Goede & Morrison (2016); EAP
Producer disease spread probabilities			
Prob. producer will become infected if returning pig truck is contaminated		0.3	EAP
Prob. producer will become infected if delivered feed is contaminated		0.8	EAP
Prob. feed truck will become contaminated if producer is infected		0.05	EAP
Prob. pig truck will become contaminated if producer is infected		0.2	EAP
Feed mill disease spread probabilities			
Prob. feed mill will become infected if returning feed truck is contaminated		0.1	EAP
Prob. feed truck will become contaminated if feed mill is infected		0.5	EAP
Slaughter plant disease spread probabilities			
Prob. slaughter plant receiving area will become infected if pig batch is infected		0.4	EAP
Prob. pig truck will become contaminated if receiving area is infected		0.2	EAP
Producer farrow, wean, and batch parameters			
Farrow to wean sow proportion	Relative to total capacity	.6	EAP
Farrow to feeder sow proportion	Relative to total capacity	.5	EAP
Farrow to finish sow proportion	Relative to total capacity	.2	EAP
Annual piglets per sow	Number of piglets	34	The Pig Site (2014)
Max. frequency of weaning	"farrow, wean and batch" function freq. (days)	7	EAP
Min. batch size	As proportion of total capacity	0.05	EAP
Capacity under which one batch	Small producers have only one pig batch	20	EAP
Producer to producer transfer parameters			
Min. capacity similarity ratio	Trading producers cannot be greatly	25	EAP

	different sizes		
Max. producer connection distance	Max. distance between trading producers (km)	150	EAP
Max. potential transferees	Max. number of producer trading partners	15	FHPC
Max. shipment frequency	"evaluate pig shipments" function freq. (days)	5	FHPC
Feed mill parameters			
Producer to feed mill proximity λ	Producers connect to λth closest feed mill (Poisson)	1.5	EAP
Avg. daily trips	Avg. num. daily feed deliveries per mill	10	FHPC
Num. producers visited λ	Expected to visit λ per delivery (Poisson)	1	EAP
Slaughter plant parameters			
Producer to slaughter plant proximity λ	Producers connect to λth closest slaughter plant (Poisson)	2	EAP

Note: "EAP" indicates that the value was derived through expert advisory panel sessions. "FHPC" refers to the family-owned hog production chain system dataset.

I.ii.d If applicable, how is space included in the model?

The model is spatially situated in a continuous, two-dimensional GIS environment. Distances between agents are calculated "as the crow flies" and measured in kilometers. In some cases, distance is a factor in determining inter-agent contact patterns.

I.ii.e What are the temporal and spatial resolutions and extents of the model?

The model's time scale is based on real-world days, with the initial model date set to January 1st, 2012. 2012 was chosen because FLAPS initialization data are drawn from the 2012 USDA Census of Agriculture (Burdett et al. 2015). The model's stop date can be set as desired depending on the experimental phenomena the user is interested in studying. Model time is continuous (i.e., events may occur part-way through a day). The model's spatial extents correspond to the extents of the selected U.S. state study area. Space is also continuous in the model, utilizing a two-dimensional GIS framework.

I.iii Process Overview and Scheduling

I.iii.a What entity does what, and in what order?

Four classes of functions define the operation of the model, presented in order of the point(s) in the simulation that they occur (see the *Implementation Details* section). First are the initialization functions, which define how the agents will be physically situated in the space, set

335

each agent's individual operational parameters, and identify lists of potential trading partners based on the classification and industry role of the agent, as well as spatial proximity to other agents. Second are the cyclically-executing functions, which make up the agents' decision rules, determining how and when contact between agents will occur (through the transfer of livestock and the distribution of feed), and thereby opening potentials for infection to spread. These functions also determine and implement the consequences of an infection upon the agent. Third is the initial infection function, which is called after the initial transient period in each run. Finally, fourth are the set of functions facilitating the output of model data for further analysis, including post-experiment scripts to parse model outputs and analyze results across multiple runs. All event scheduling in the model follows a Last-In-First-Out (LIFO) protocol.

II) Design Concepts

II.i Theoretical and Empirical Background

II.i.a Which general concepts, theories or hypotheses are underlying the model's design at the system level or at the level(s) of the submodel(s) (apart from the decision model)? What is the link to complexity and the purpose of the model?

Because real-world epidemics are fundamentally phenomena which propagate through networks (social, business, transportation, etc.), the formulation of a suitably-realistic network structure within which agents operate is a fundamental basic principle of the model. A corollary to this basic principle concerns the model's balance between context specificity and analytic transparency. The model's network generation algorithm strives to maintain sufficient context specificity to capture the critical complexities underpinning observed epidemiological spread phenomena, while bracketing superfluous elements of real-world production chain networks which have not been implicated in previous epidemiological events. For example, the model contains only feed mill, producer, and slaughter plant agent classes, because these were identified by industry experts in our Delphi panels as the critical players underpinning the spread of fecal-oral livestock diseases. Whereas in real-world hog production chain networks there may be a multitude of other actor typologies (e.g. equipment suppliers, construction contractors, insurance agents, auction houses, and many more), these were intentionally excluded from the model's design to simplify analysis.

Another guiding principle is the geospatial situation of the model within real U.S. states. The states of North Carolina, Iowa, and Illinois were chosen because they produce a large number of the nation's hogs, as well as being amongst the most hog-dense states. In many epidemiological studies, agent density has been shown to impinge directly upon spread characteristics. With high enough density, complex phenomena such as percolation thresholds may emerge (Wiltshire, 2018).

336

II.i.b On what assumptions is/are the agents' decision model(s) based?

The primary set of assumptions driving agent behavior relate to trade patterns associated with the industry role each agent plays, with agents in the model operating in accordance with general industry norms. For example, it is assumed that, as soon as their livestock batches reach the transfer age appropriate for their industry role, producer agents will search the agent space for trading partners until an appropriate partner is found, at which point the pig batch will be immediately transferred.

Another assumption concerns the spatial locations of non-producer agents. Since fine-grained spatial data were not available—and the FLAPS tool only covers livestock production units (Burdett et al. 2015)—locations of non-producer agents are initialized by distributing them at random positions within each county, in proportion to the number of producers in the county.

Several assumptions also come into play concerning the distance, similarity, and number of other industry actors with which each agent may interact. These assumptions were parameterized using the maximum distance, minimum capacity similarity, and maximum connection number global parameters, as well as the characteristic connection distance "lambda" values that underlie how feed mill and slaughter plant service areas are generated. For example, in the case of feed mills, $\lambda = 1.5$, indicating that most producers purchase feed from the nearest or second-nearest mill, with fewer purchasing from the third-nearest, fewer still from the fourth-nearest, etc. Alternatively, for slaughter plants, $\lambda = 2$, indicating that the most likely outcome is for a producer to ship hogs to the second-closest plant. We note that a limitation of the model is that, since all agents exist within the bounds of a single U.S. state, inter-state trade is not accounted for.

A further assumption is our representation of livestock in batches (or metapopulations) of animals of the same age. This was primarily done to reduce computational overhead (vs. storing each animal's parameters individually). In addition to having the same theoretical birthday, it is assumed that if a batch is infected, all of its members are infected.

Finally, it is assumed that if a producer agent becomes infected, all of its livestock batches become infected. While a simplifying assumption to be sure, this is reasonably realistic, owing to the high observed virulence of the PED virus, which tends to sweep quickly through entire herds.

II.i.c Why is/are certain decision model(s) chosen?

Agents' decision heuristics primarily relate to their day-to-day operations, and rely on parameters including industry role, size, and spatial location. Based on industry standards along with individual parameters, agents make decisions pertaining to when and with whom inter-agent contact will occur. These contact patterns go on to impact the susceptible / infective

state variable of each agent, according to probabilities associated with each modality of inter-agent contact

II.i.d If the model/submodel (e.g. the decision model) is based on empirical data, where do the data come from?

We use the Farm Location and Agricultural Production Simulator (FLAPS) tool—which draws upon USDA Census of Agriculture data along with aerial imaging to impute realistic distributions of livestock farms within a specified U.S. region—to set producer agent locations and key operational parameters including industry roles and capacities (Burdett et al. 2015). While the FLAPS tool serves as our primary means to set production unit locations and operational parameters, our team also gained access to internal records from a large family-owned hog production chain system—identified as "FHPC" in Table 6—which was used to impute realistic contact rate and shipment size parameters. Several other sources of empirical data were also used to parameterize the model, also indicated in Table 6.

II.i.e At which level of aggregation were the data available?

The FLAPS system uses several core datasets to impute producer agent locations and operational characteristics. The system primarily relies upon the 2012 USDA Census of Agriculture, which is aggregated at the county level. However, FLAPS also leverages aerial imaging land use data to position agents more precisely within each county. Slaughter plant data from USDA NASS (2014) were aggregated at the level of the U.S. state.

II.ii Individual Decision-Making

II.ii.a What are the subjects and objects of the decision-making? On which level of aggregation is decision-making modelled? Are multiple levels of decision making included?

Decision-making is modeled at the level of the agent, be it a producer, feed mill, or slaughter plant. When transferring livestock, for example, a producer agent will wait until a pig batch matches the appropriate age corresponding to its industry role. This subject will then search among its potential trading partners (objects, in this case), which were pre-selected at model initialization to be of the appropriate industry role, until a suitable agent is found that has sufficient capacity to accept the shipment.

II.ii.b What is the basic rationality behind agent decision-making in the model? Do agents pursue an explicit objective or have other success criteria?

Agents in the model act according to accepted industry operational standards. They do not pursue a specific "objective," per se, other than to efficiently take in new livestock and ship them out at the appropriate life cycle stage. Future versions of the model will incorporate adaptive agent decision-making, for example allowing for decisions which will reduce an agent's vulnerability if a disease is present in the network.

II.ii.c How do agents make their decisions?

A series of cyclically-executing functions (outlined in the *Implementation Details* section) govern how and when agents make decisions. These are based on the industry role of each agent.

II.ii.d Do the agents adapt their behavior to changing endogenous and exogenous state variables? And if yes, how?

Producer agents desiring to transfer livestock to the next production phase adapt their behavior based on the operational variables associated with their potential trading partners. This primarily comes down to finding a partner within a certain distance, of the appropriate industry role, of sufficiently-similar size, and with sufficient excess capacity to accept the shipment. The number of pigs housed at each production unit is constantly updated as the model runs, so the agents have to perform this search with each new outgoing shipment. However, agents in the current model do not adapt their decisions based on the prevalence of disease in the system, or other global factors.

II.ii.e Do social norms or cultural values play a role in the decision-making process?

Agents follow basic industry standards when making their decisions, which could be considered a kind of social norm. Values do not play into their choices, however.

II.ii.f Do spatial aspects play a role in the decision process?

Producer agents looking to transfer animals to another producer are constrained by a maximum distance parameter. The service areas of slaughter plant and feed mill agents are governed by Poisson distributions, with producers most likely to connect to the λth-closest of each.

II.ii.g Do temporal aspects play a role in the decision process?

Producer agents will only transfer livestock to their producer trading partners if the partner has sufficient excess capacity. Since the inventory of each agent is constantly in flux, the time when the transfer function is executed will determine the trading partner that is chosen. Farrowing producers will also wait to batch weaner pigs until the quantity of piglets is greater than or equal to the minimum batch size, as a proportion of their capacity.

II.ii.h To which extent and how is uncertainty included in the agents' decision rules?

Agent behavior does not account for uncertainty in the current model. However, we are pursuing behavioral research which will be used to parameterize agents in future model versions to react to uncertainty as regards disease prevalence in the networks.

II.iii Learning

II.iii.a Is individual learning included in the decision process? How do individuals change their decision rules over time as consequence of their experience?

The agents' decision rules remain non-adaptive in the current model. Decision heuristics are based on industry roles, and are designed to realistically replicate throughput in the production chain system as a whole. Thus, an agent will transfer animals to an appropriate trading partner as soon as possible, farrowing will proceed regularly wherever a producer has sufficient excess capacity, and feed deliveries take place at a set frequency. The agents' behavior does not change as a result of model conditions, for example the presence of a disease within the network, however each agent will necessarily adapt to market conditions resulting from the available spare capacity of its trading partners.

II.iii.b Is collective learning implemented in the model?

No.

II.iv Individual Sensing

II.iv.a What endogenous and exogenous state variables are individuals assumed to sense and consider in their decisions? Is the sensing process erroneous?

Producer agents "sense" the operational variables associated with potential trading partners when making decisions concerning livestock transfers. An agent's perception of these factors is

not erroneous, as it is based on a direct query of the potential recipients' operational variables at the time the transfer is desired.

II.iv.b What state variables of which other individuals can an individual perceive? Is the sensing process erroneous?

The primary state variables agents perceive when making livestock transfer decisions are the potential trading partner's industry role, capacity, inventory, and proximity.

II.iv.c What is the spatial scale of sensing?

At model initialization, producer agents generate a list of potential producer trading partners only within a given distance (150 km). It could be said that their "sensing" of other producers does not extend beyond this distance. Note that (based on the Poisson distributions discussed above), producers may very well interact with slaughter plants and feed mills beyond the 150-km limit imposed upon producer to producer transfers.

II.iv.d Are the mechanisms by which agents obtain information modelled explicitly, or are individuals simply assumed to know these variables?

Agents are simply assumed to know the relevant operational variables necessary to make the choice of which trading partner to choose.

II.iv.e Are the costs for cognition and the costs for gathering information explicitly included in the model?

No.

II.v Individual Prediction

II.v.a Which data do the agents use to predict future conditions?

Agents do not engage in predictive behavior concerning the likely outcomes of their actions. They simply operate according to industry standards.

II.v.b What internal models are agents assumed to use to estimate future conditions or consequences of their decisions?

None, although future versions of the model may incorporate the Theory of Planned Behavior with regard to agent decisions to increase biosecurity measures or limit livestock movements in response to a perceived disease threat in the system.

II.v.c Might agents be erroneous in the prediction process, and how is it implemented?

Since prediction is not currently a decision-making factor, no.

II.vi Interaction

II.vi.a Are interactions among agents and entities assumed as direct or indirect?

Interactions among agents take the form of transfers of livestock and feed. These transfers are assumed to be mediated by transportation equipment, and so can be conceived as indirect. For example, if a truck delivers livestock to an infected premises, a probability parameter governs whether the truck will become contaminated. The truck is then assumed to return to its origin, at which point (if it is contaminated), another probability parameter determines whether the truck contamination will result in infection of the original agent. Similarly, a feed truck originating from a mill may become contaminated upon reaching an infected producer premises, and then may pass the infection along to another producer on the same route, or to the mill itself upon its return.

II.vi.b On what do the interactions depend?

Interactions depend upon the operational variables associated with each agent. Producer agents will only initiate a transfer when a pig batch reaches the designated transfer age. Spatial proximity and underlying network structure—encoded at model initialization—also mediate interactions. For example, feed mills periodically generate delivery routes which encompass a subset of producer agents within their service areas.

II.vi.c If the interactions involve communication, how are such communications represented?

The only communication that is represented during livestock transfers is the size and age of pig batches, along with the spare capacity of the potential trading partner. This may be conceived as a phone call or email to potential trading partners inquiring whether they are able to accept

342

the batch; or alternatively as predetermined business arrangements, common in private production system networks.

II.vi.d If a coordination network exists, how does it affect the agent behavior? Is the structure of the network imposed or emergent?

Agents coordinate only in the sense that, during livestock transfers, the receiving agent must concurrently have the spare capacity to accept the shipment. Thus, to the extent that coordination occurs in the model, it is emergent, not imposed.

II.vii Collectives

II.vii.a Do the individuals form or belong to aggregations that affect and are affected by the individuals? Are these aggregations imposed by the modeler or do they emerge during the simulation?

Livestock in the model may be considered as collectives (or metapopulations), as they are encoded in groups of animals with the same theoretical age. This is imposed by the model structure.

In addition, while not defined explicitly as such, groups of agents in the model exhibit emergent collective characteristics due to their differential spatial distribution across the model's GIS space. For example, in densely-packed areas, groups of agents tend to interact heavily within connected clusters, potentially leading to localized disease outbreaks. This type of emergent collective behavior is not directly imposed by the modeler, although the fixed spatial location and network structure that is imposed at model initialization impacts how and where such phenomena arise.

II.vii.b How are collectives represented?

Metapopulations of livestock are tracked as passive objects within a concurrent hash map data structure associated with each agent object. This data structure encodes the size and age of each pig batch.

Emergent "collectives" of heavily-connected agent clusters are captured by tracking agent contact patterns over the course of each run, and outputting these data as weighted edge lists.

II.viii Heterogeneity

II.viii.a Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?

As described in the *Entities, state variables, and scales* section above, agents fall into three main classes: (a) producers, (b) slaughter plants, and (c) feed mills. Producer agents are assigned one of six industry roles, based on the USDA classification system for hog producers, along with expert advisement. A producer agent's industry role determines the initial age of its livestock, its livestock age transfer condition, as well as the appropriate trading partners which make up its set of potential trading partners. These relationships are visualized in Figure 1.

II.viii.b Are the agents heterogeneous in their decision-making? If yes, which decision models or decision objects differ between the agents?

Agents' decision-making will differ depending upon their industry roles. For example, a farrow-to-wean producer will only send pigs to wean-to-feeder or wean-to-finish producers; slaughter plant agents will only receive hogs from finishing producers within their service areas; and feed mills will only distribute to producers within their service areas.

II.ix Stochasticity

II.ix.a What processes (including initialization) are modelled by assuming they are random or partly random?

RUSHPNBM uses both fixed-seed stochasticity (for initialization) as well as random-seed stochasticity (to mediate infection spread). All initialization procedures that are not drawn directly from the FLAPS data— for example the list of potential transferees for each producer— utilize a fixed seed in all draws from stochastic functions. This is important because we are interested in analyzing the dynamics associated with disease risk, and changing the basic supply chain network structure across runs would confound results. We also use fixed-seed draws from custom distributions—i.e., those based on the number of observations in a series of categories—to establish the spatial location of slaughter plants and feed mills to correspond with producer density by county.

Poisson distributions are utilized in some cases in model initialization, for example to determine which slaughter plant or feel mill to which each producer will connect. Poisson distributions are used in the model because they require only a λ parameter corresponding to the expected value of the distribution; because they are discrete, returning a whole-number; and because they cannot return a value less than zero. For initialization procedures, the fixed seed is utilized in

these Poisson draws, maintaining the same initial "latent" network structure for each study area across runs.

The remainder of the stochasticity in the model uses a random seed, yielding a dataset representing a distribution of contact and infection patterns across model runs. Random-seeded stochasticity is first used to populate each producer's initial pig batches. Both the size and age of the pig groups associated with each producer are drawn from a uniform distribution bounded according to the producer's industry role. Since each producer will always start with a slightly different animal inventory, this ensures that the trade patterns that unfold throughout each model run are not repeated exactly. Random-seeded Poisson distributions are used to determine the number of producers to visit for each feed distribution trip.

Random-seed stochasticity is also used for all disease-spread calculations. Uniform probability distributions returning "true" if a randomly-drawn value between zero and one is less than p are used to determine if the infection will spread. Thus, whereas the disease spread probabilities stay constant across runs, the result of any given random draw using these probabilities may differ.

Finally, a random-seeded triangular distribution centered at the average infection duration for each agent type, and limited to the range between 50% and 150% of this value, is used to determine the length of time an agent will remain in the infected state. Triangular distributions were used here because they are a good stand-in for the normal distribution, while offering an intuitive means to establish upper and lower limits.

II.x Observation

II.x.a What data are collected from the ABM for testing, understanding and analyzing it and how and when are they collected?

The model tracks in real-time the current livestock inventory of all producers in the model, the number of currently infected animals, the number of currently infected agents, and the cumulative number of infected agents, which can be output as time-series data to examine infection-spread dynamics. Tabular data including each agent's class, operational parameters, and other information is also output at the conclusion of each run.

In addition, a contact network adjacency matrix with link weights encoding the number of times each agent interacted throughout the model run is exported as tabular data after each run, and later parsed using a series of Python functions. An infection-spreading network is similarly tracked, output, and parsed. Key statistics on trade and infectivity patterns across a series of model runs—both at the individual agent as well as the whole-network level—may then be analyzed.

Finally, for calibration purposes, we include an option to track the flow of feed and livestock between different types of agents, for example the distribution of hog shipment sizes and delivery frequencies, which can then be compared with available real-world data.

II.x.b What key results, outputs or characteristics of the model are emerging from the individuals? (Emergence)

Emergent phenomena in the model occur as a result of the contact patterns mediated by agents' decision heuristics. For example, this could take the form of differential disease spread dynamics resulting from the structures of contact networks that emerge throughout each run.

III) Details

<u>III.i Implementation Details</u>

III.i.a How has the model been implemented?

The model was implemented using AnyLogic v.8 software, which relies upon the Java programming language for all scripts and functions. The sections below use pseudocode to describe in detail the algorithmic structures underlying each model function.

Notes on pseudocode used in this document:

- The characters "//" will be used to designate a descriptive comment (i.e., the line of text following the "//" is not part of the actual function logic).
- Parameters referenced in all functions refer to those associated with the agent object from which a function has been called. In some cases, to disambiguate, the terms "self" or "my" may be used to refer to the function-calling agent object or its associated parameters.
- "ADD OR INCREMENT [sender] in [receiver]'s [network edge list]" is defined here as:
  ```
  IF [sender] is not in [receiver]'s [network edge list] ADD [sender] to
          [receiver]'s [network edge list] with contact counter set to 1
  ELSE INCREMENT contact counter associated with [sender] in [receiver]'s
          [network edge list] by 1
  ```
- "RANDOM DRAW using [probability]" is defined here as the Boolean value resulting from:
  ```
  (DRAW random number from uniform distribution between 0 and 1) <
          [probability]
  ```

346

III.i.b Is the model accessible, and if so where?

While the source code for the model is not accessible due to limitations of the AnyLogic software, the pseudocode below explicates the code at a high level of detail. In principle, these pseudocode functions may be used to implement the model using any desired programming language.

III.ii Initialization

III.ii.a What is the initial state of the model world, i.e. at time t = 0 of a simulation run?

The model is initialized by progressing through a series of functions. Several agent parameters are set as each agent object is generated by the model. Next, further agent parameters are set by reference to the model database. Finally, upon completion of the preceding, producer agents initialize their networks of potential trading partners to be referenced throughout the model run.

▪   Initialization functions called from main object (in order of function calls):
Initialize map view function:

```
SET GIS map boundaries and zoom on U.I. dashboard to correspond to study
        area
```

Initialize agents function:

Numbers, locations, and typological distributions of producer agents within the model are generated heuristically using the Farm Location and Agricultural Production Simulator (FLAPS) system developed through a collaboration between Colorado State University and the United States Department of Agriculture (Burdett et al. 2015). FLAPS parses USDA National Agricultural Statistics Service (USDA NASS) databases along with land use data to impute spatially-explicit datasets depicting the distribution of livestock production units throughout the desired study area. Thus, while not representing actual farm locations, the producer agents in the model are distributed geographically and with characteristics including production volume and classification category in such a way as to be consistent with real-world distributions.

An implicit assumption we have made is that the distribution of slaughter plants and feed mills mirrors the distribution of producers by county. Once a non-producer agent's county has been assigned by drawing from this distribution, the agent is placed at a random set of coordinates within the selected county. Thus, counties with higher producer density will tend to have higher numbers of non-producer industry actors as well.

```
FOR EACH producer agent
    // read in data for each agent from database and set appropriately
    READ latitude and longitude from database table corresponding to study
          area
    SET agent's spatial location
    READ total capacity from database table corresponding to study area
    SET agent's total capacity
    READ producer industry role from database table corresponding to study
          area
    // in consultation with industry experts, producers classified as
          "other" in the USDA NASS data are assumed to be "wean to feeder"
    IF (database query result = "other")
       SET agent's industry role to "wean to finish"
    ELSE
       SET agent's industry role to correspond with database query result
    SET agent's icon color to match producer type
    IF (agent is a farrowing type)
       SET agent's sow inventory to MAX of 1 and (total capacity * sow
             ratio parameter appropriate for agent's type)
    ELSE
       SET agent's sow inventory to zero
    SET agent's non-sow capacity to (total capacity – sow inventory)


    // iteratively generate initial pig batches
    SET minimum batch size to (agent's total capacity * minimum batch size
          as proportion of capacity parameter)
    WHILE (minimum batch size < remaining capacity)
       IF (agent's total capacity <= capacity under which a producer is
             assumed to have only one batch parameter)
          SET batch size to agent's non-sow capacity
       ELSE
          SET batch size to a random integer between minimum batch size and
                non-sow capacity
          IF ((batch size + current pig inventory) > non-sow capacity)
             SET batch size to minimum batch size
       SET batch birthday to a random integer between the maximum and
             minimum age of a pig for the agent's industry role
       ADD batch size and birthday to pig tracker
       INCREMENT agent's current inventory by batch size
```

348

```
// distribute other agents by county to correspond with producer density
FOR EACH county in study area
    READ number of producers in county from database
    ADD county name to county distribution array as many times as there are
            producers in that county
FOR EACH slaughter plant agent
    DRAW RANDOM county name from county distribution array
    SET agent's spatial location to a random point inside the county drawn
REPEAT above FOR loop for feed mill agents


// now that locations and characteristics are set, run each producer's
        network initialization function in turn
FOR EACH producer agent
    CALL agent's "initialize network" function
```

▪ Producer agent initialize network function:

Once the agents' locations and industry roles have been initialized, a network initialization function generates a set of potential trading partners for each agent. All producer agents are assigned one feed mill, and finishing producers are also assigned one slaughter plant, both connections being to the nth-closest of that agent type, with n being drawn from a Poisson distribution using the appropriate λ parameter. A pool of potential transferee producers is also generated for each non-finishing producer according to their industry role. These relationships are shown in Figure 1. The potential transferee producers in this potential transferee pool are filtered according to (a) the maximum producer-to-producer connection distance parameter, (b) the minimum capacity similarity ratio parameter, as well as (c) the maximum number of transferee producers parameter.

```
IF industry role is NOT a finishing type
    Potential farms list = FILTER other producer agents s.t. (industry role
            of other producer is the next step in the production chain) AND
            (distance to the other producer <= max producer-producer
            connection distance global parameter) AND NOT ((total capacity /
            other producer's total capacity >= minimum capacity similarity
            ratio parameter) OR (other producer's total capacity / total
            capacity >= minimum capacity similarity ratio parameter)))
    WHILE (there are still potential farms AND (my transferee producers <=
            max number of transferee producers global parameter))
        DRAW RANDOM from potential farms list and ADD to my transferee
            producers
    IF my transferee producers is empty
        ADD nearest farm of appropriate industry role
IF industry role is a finishing type
```

```
            SET my slaughter plant to Nth closest slaughter plant where N is drawn
                 from a Poisson distribution where lambda = the global proximity
                 lambda parameter for slaughter plant connections
       SET my feed mill to Nth closest feed mill where N is drawn from a Poisson
              distribution where lambda = the global proximity lambda parameter
              for feed mill connections
```

III.ii.b Is the initialization always the same, or is it allowed to vary among simulations?

The initialization of the spatial location, operational characteristics, and potential trading partners for each agent remains consistent across runs within each of our three study area states. Thus, there are in essence three distinct initial states with regard to the above parameters, defined by the study areas. However, the initial livestock population housed at each producer premises differs between runs, as do the real-time trading choices and infection spread patterns experienced by each agent throughout the run. For more detail, see the *Stochasticity* section.

III.ii.c Are the initial values chosen arbitrarily or based on data?

Initialization parameters rely upon several datasets, including the University of Colorado / USDA FLAPS system (Burdett et al. 2015), USDA NASS data (USDA NASS 2014), Google Maps queries, and livestock industry internal records. For more details, see the initialization function descriptions and pseudocode above, section *III.iv.c*, as well as Table 6.

III.iii Input Data

III.iii.a Does the model use input from external sources such as data files or other models to represent processes that change over time?

The model relies upon an external database to store many of the initialization parameters. Once set, these values remain static throughout each model run.

III.iv Submodels

III.iv.a What, in detail, are the submodels that represent the processes listed in 'Process Overview and Scheduling'?

▪ Producer agent cyclically-executing functions:
Farrow, wean, and batch piglets function:

If a farm which farrows piglets (Farrow to Wean, Farrow to Feeder, or Farrow to Finish types) is left with excess capacity after a livestock transfer, a farrowing function fills that capacity with a new batch of piglets, whose birthday is set to the current model day. Once again, to eliminate unrealistically-small pig groups, a minimum farrowing size as a proportion of the farrowing farm's total capacity is required for the farrowing function to proceed. Thus, a farm which is already almost at maximum capacity will not farrow a new batch of piglets until another batch has been shipped to an appropriate trading partner.

```
**Recurrence time is the frequency of weaning global parameter**
IF industry role is a farrowing type
    // calculate number of farrowed piglets ready to wean and batch
    Current piglet inventory = MIN of remaining pig capacity and (days since
            last weaning day * number of sows * (global parameter for piglets
            weaned annually / 365))


    IF infectivity state is "infected"
        DECREMENT current piglet inventory according to suckling mortality
                rate global parameter
    // wean and batch piglets
    Number to wean and batch = MIN of current piglet inventory and remaining
            pig capacity
    IF (number to wean and batch >= my minimum batch size)
        ADD number to wean and batch and birthday (current day – 35) to pig
                batch tracker
        INCREMENT non-sow pig inventory by batch size
        DECREMENT piglet inventory by batch size
        SET last weaning day to current day
```

Evaluate pig shipments function:

Non-finishing producers transfer hoofstock to a transferee farm as soon as the hoofstock reach the age corresponding to the transfer condition associated with the industry role of the producer. If it is determined that the transfer age requirement of a pig batch has been met, the transferee producers in the transferring producer's pool of possible producer trading partners are sequentially evaluated to determine whether they are able to receive the shipment. To eliminate the transfer of unrealistically-small groups of livestock, transfers will only proceed if the pig batch size exceeds the minimum transfer quantity, as a proportion of the transferee's total capacity. If the excess capacity of a potential transferee producer is less than the size of

351

the pig batch, the pig batch will be split such that the transferree producer's capacity will be filled, and the remaining animals will stay with the transferring producer. The transferring producer will then continue to assess producers until all remaining pigs in the pig batch have been transferred to appropriate trading partners.

If the transferring producer is infected but the transferee is not, the transferred hoofstock will automatically spread the infection to the transferee producer. If the transferee producer is infected but the transferring producer is not, the "delivery trailer" returning from the infected transferee producer may infect the transferring producer according to a probability set at model initialization.

The birthday parameter associated with the batch of transferred stock is maintained as it is passed to the transferee(s), such that the pig batch will once again be appropriately transferred to the next production phase at the correct transfer age. In the rare case that a pig batch grows too old before a suitable transferee producer can be located, the pig group is culled, making room for a new batch of pigs.

Finishing producers (Feeder to Finish and Farrow to Finish types) ship hoofstock their slaughter plant as soon as the hoofstock reach the designated slaughtering age. If the transferring producer is infected, the receiving area of the slaughter plant may become contaminated according to a probability set at model initialization. If the receiving area of the slaughter plant is already contaminated, the "delivery trailer" returning to the transferring producer may carry the infection back to that producer according to another probability set at model initialization.

```
**Recurrence time is the maximum frequency of pig shipments global
        parameter**
// to eliminate continually changing inventory levels during execution
SUSPEND farrow, wean, and batch piglets function countdown


// determine where pigs will be shipped
IF industry role is NOT a finishing type
    FOR EACH pig batch meeting age transfer requirement
        FOR EACH transferee in my transferee producers
            IF (batch size <= transferee's spare non-sow capacity) AND (batch
                    size >= transferee's minimum batch size)
                IF (transferee's infectivity state is "infected") AND
                        (infectivity state is "clean")
                    DECREMENT batch size according to mortality rate global
                            parameter associated with pigs' age
                REMOVE pig batch from pig batch tracker
                DECREMENT non-sow pig inventory by batch size
                ADD pig batch and birthday to transferee's pig batch tracker
                INCREMENT transferee's non-sow pig inventory by batch size
```

```
                    // update contact network trackers
                    ADD OR INCREMENT transferee in contact network out-degree list
                    ADD OR INCREMENT self in transferee's contact network in-
                            degree list

                    // update pig shipment trackers
                    ADD batch size to pig shipments out list
                    ADD batch size to transferee's pig shipments in list

                    // infection brought to transferee via infected pigs
                    IF infectivity state is "infected"
                        SET transferee's infectivity state to "infected"
                        ADD OR INCREMENT transferee in infection-spreading network
                                out degree list

                    // infection brought home via trailer from transferee farm
                    IF (transferee's infectivity state is "infected") AND (RANDOM
                            DRAW using Prob. pig truck will become contaminated if
                            producer is infected) AND (RANDOM DRAW using Prob.
                            producer will become infected if returning pig truck is
                            contaminated)
                        SET infectivity state to "infected"
                        ADD OR INCREMENT self in transferee's infection-spreading
                                network out degree list

            // cull pigs that are too old and were never able to be transferred
            FOR EACH pig batch over 168 days old
                REMOVE batch from pig batch tracker
                DECREMENT non-sow pig inventory by batch size

    ELSE IF industry role is a finishing type
        REMOVE pig batch from pig batch tracker
        DECREMENT non-sow pig inventory by batch size

        // update contact network trackers
        ADD OR INCREMENT slaughter plant in contact network out-degree list
        ADD OR INCREMENT self in slaughter plant's contact network in-degree
                list

        // update pig shipment trackers
        ADD batch size to pig shipments out list
```

```
        ADD batch size to slaughter plant's pig shipments in list

        // infection brought to slaughter plant via infected pigs
        IF (infectivity state is "infected") AND (RANDOM DRAW using Prob.
              slaughter plant receiving area will become infected if pig batch
              is infected)
            SET slaughter plant's infectivity state to "infected"
            ADD OR INCREMENT slaughter plant in infection-spreading network out
                degree list

        // infection brought home via trailer from slaughter plant
        IF (slaughter plant's infectivity state is "infected") AND (RANDOM DRAW
              using Prob. pig truck will become contaminated if receiving area
              is infected) AND (RANDOM DRAW using Prob. producer will become
              infected if returning pig truck is contaminated)
            SET infectivity state to "infected"
            ADD OR INCREMENT self in slaughter plant's infection-spreading
                network out degree list

    IF (non-sow pig inventory is zero)
        SET infectivity state to "clean"
    RESUME farrow, wean, and batch piglets function countdown
```

- Feed mill agent cyclically-executing functions

Feed mills periodically generate delivery routes encompassing a subset of producers within their latent feed-mill-to-producer link set. Each route encompasses a subset of the producers in the feed mill's service area, with the number of stops in each trip resulting from a draw from a Poisson distribution. While there is no actual "feed truck" object in the model, the logic of the following function is based on the way such a truck would move between agents and possibly spread disease.

Beginning from the mill, this conceptual feed truck will visit the previously-drawn number of randomly-selected producers within the feed mill's service area before finally returning to the feed mill. If the feed mill is infected, the truck may be contaminated initially. Should the truck encounter an infected producer on its route, it may become contaminated at that point. Once a truck is contaminated, the infection may be spread to subsequent producers on the route. If a contaminated truck returns to the feed mill, the mill itself may become infected.

Distribute feed function:

```
**Recurrence time is the frequency of feed deliveries global parameter**
// determine whether truck is initially infected
IF (infectivity state is "infected") AND (RANDOM DRAW using Prob. feed
        truck will become contaminated if feed mill is infected)
    SET truck infected to "true"
ELSE
    SET truck infected to "false"


// determine number of farms on delivery route
SET number to visit to MIN of (number of farms in service area) and (DRAW
        from Poisson distribution with lambda equal to global parameter
        encoding average number of producers visited per route)


// generate delivery route
FOR number to visit
    ADD random producer in service area (that is not already in delivery
            route list) to delivery route list


// parse infectivity consequences of delivery route
FOR EACH producer in delivery route list
    // update contact network trackers
    ADD OR INCREMENT producer in contact network out-degree list
    ADD OR INCREMENT self in producer's contact network in-degree list

    //infected truck infects farm it's delivering to
    IF (truck infected is true) AND (RANDOM DRAW using Prob. producer will
            become infected if feed truck is contaminated)
        SET producer's infectivity state to "infected"
        ADD OR INCREMENT producer in infection-spreading network out degree
                list

    //truck becomes infected from delivery to infected farm
    IF (producer's infectivity state is "infected") AND (RANDOM DRAW using
            Prob. feed truck will become contaminated if producer is
            infected)
        SET truck infected to "true"
        SET "truck-infecting producer" to current producer

// infected truck infects feed mill
```

```
IF (truck infected is "true") AND (RANDOM DRAW using Prob. feed mill will
        become infected if returning feed truck is contaminated)
    SET infectivity state to "infected"
    ADD OR INCREMENT self in truck-infecting producer's infection-spreading
            network out degree list
```

- Initial infection function

The system is initialized with all agents free of infection. After one model year has passed, an infection is introduced to a random subset of producer agents. The proportion of agents which are infected by the initial infection function may be set at model initialization.

The reason for the one-year lag is to skip the transient period and allow the model to stabilize before analyzing the effect of an introduced disease. This lag is necessary because, as in a real production chain, a certain amount of slack, or a difference between the theoretical production capacity and actual production, is characteristic in the modeled production chains. In the model, this economic slack is due to the producers sometimes temporarily operating at less than maximum hoofstock capacity until an appropriate shipment of livestock becomes available. In general, after about 9 months, the level of slack in the model has stabilized.

Initial infection function:

```
    **Function is called only once, after one model year**
    FOR number to infect global parameter
        SET randomly chosen producer agent's infectivity state to "infected"
```

- Infection control functions

Susceptible/infective state charts:

Each agent has an embedded state chart which encodes its infectivity status. Should an agent become infected, a function is called which calculates the number of its stock which are to die of the disease. The proportion of livestock which succumb to the disease is based on the age of the pig groups, with uniform mortality rates set at model initialization for suckling pigs, nursery pigs, and grow/finish hogs. After die-off is calculated for pig groups of each life stage within an infected producer's inventory, the producer's inventory data are updated accordingly. An agent will remain infected for a duration whose mean length in days is controlled by parameters specific to each agent type by drawing from a triangular distribution limited to between 50% and 150% of the mean value. Upon transition back to "clean," the total infection duration is updated for later analysis. In some cases, such as when a producer has an inventory of zero after transferring pigs, the transition back to a "clean" state may also be triggered manually.
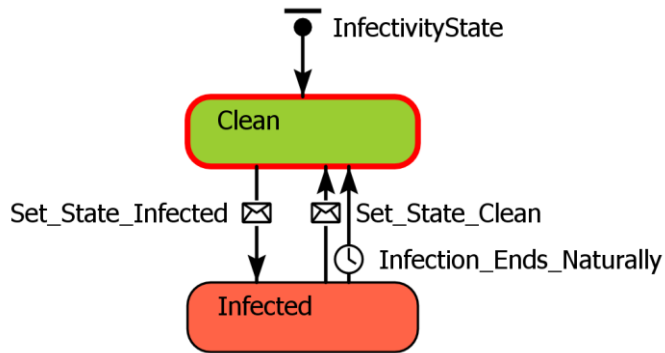
356

Figure 2: Infectivity State Chart

Livestock mortality calculation:

Calculate mortality function:

```
DECREMENT piglet inventory by global parameter encoding piglet mortality
        proportion
FOR EACH pig batch in pig batch tracker
    DECREMENT pig batch size by global parameter encoding mortality
            proportion appropriate for age of pigs
```

III.iv.b What are the model parameters, their dimensions and reference values?

See Table 6.

III.iv.c How were the submodels designed or chosen, and how were they parameterized and then tested?

Due to the inherent variability of epidemic events within complex networked systems, we are interested less in empirically-validating the model to be used as a forecasting tool, and more in developing sufficient structural and face validity to allow for a deeper understanding of the dynamics of the modeled systems. Even given identical starting conditions, deviations in contact patterns over the course of a real-world disease incursion render precise forecasts unfeasible. For example, while reducing outcome volatility in RUSHPNBM could easily be accomplished by eliminating stochasticities associated with disease transmission, calibrating the model such that outcomes correspond precisely to a single observed epidemic event misses the point. Our aim is rather to uncover and better understand the fundamental network features that lead to epidemiological vulnerability in livestock production systems more generally.

357

Calibration and validation procedures that leverage concrete historical data are often regarded as the best way to bring a model in line with empirical evidence. Unfortunately, there is a marked lack of publicly-available data in the agricultural sector beyond aggregated county- or state-level statistics. To the extent that datasets containing explicit locations, operational parameters, livestock and feed movements, and disease histories exist; these data tend to be held by private enterprises, which view them as sensitive internal records. In light of this, following Windrum et al. (2007), we employ several alternative calibration procedures that have been widely-used in previous modeling endeavors in which fine-grained data are scarce.

The spatial locations and basic operational parameters of RUSHPNBM agents associated with each study area are calibrated using the ``indirect'' approach, whereby stylized facts about the distribution of agents in the system are gleaned from statistical datasets. Statistical datasets used in this process include the FLAPS output data, USDA data, and livestock and feed movement records we obtained from a large U.S. family-owned hog production chain system (discussed below).

To calibration additional model elements that define how and when inter-agent contact occurs, as well as epidemiological submodel parameters, we leveraged an iterative companion modeling approach (Barreteau et al. 2003). 2015 to 2017, we convened several Delphi panels consisting of livestock industry and veterinary experts at national research team meetings and livestock veterinary conferences. In these meetings, we used both qualitative focus groups as well as questionnaires to elicit and hone parameter values. Using this participatory methodology, the modeled system was brought in line with the collective understandings of stakeholders who are intimately familiar with the operational details of U.S. livestock production systems. As model development progressed, these same experts also provided input to ensure the face validity of the distribution of epidemic patterns, scales, and durations produced by the model.

The model's parameters and functions controlling pig movement and feed deliveries were further validated with the help of data provided by a large U.S. family-owned hog production chain system (as a result of our confidentiality agreement, the company's name is not disclosed here). The database contains two-year records of each pig movement and each feed delivery involving producers in the system, although spatial data on premises locations were not provided. The family-owned hog production chain system consists of a network of 161 producer partners that raise pigs from birth to market. This production chain system has two characteristics that allowed for the parameterization and validation of RUSHPNBM. The first characteristic is that the farm sizes vary from small (300 pigs) to large (8800 pigs, Figure 3). The second characteristic is that pigs are grown at specialized sites including farrowing, wean-to-finish, nursery, and finishing premises; and are moved across the production network according to their growing stage. The pig movement records were used to derive realistic estimates of hog transfer frequencies and number of hogs per transfer relative to farm size, including both producer to producer transfers, and producer to slaughter plant transfers (Table 7). The feed delivery records were used to estimate delivery frequencies (Figure 4).
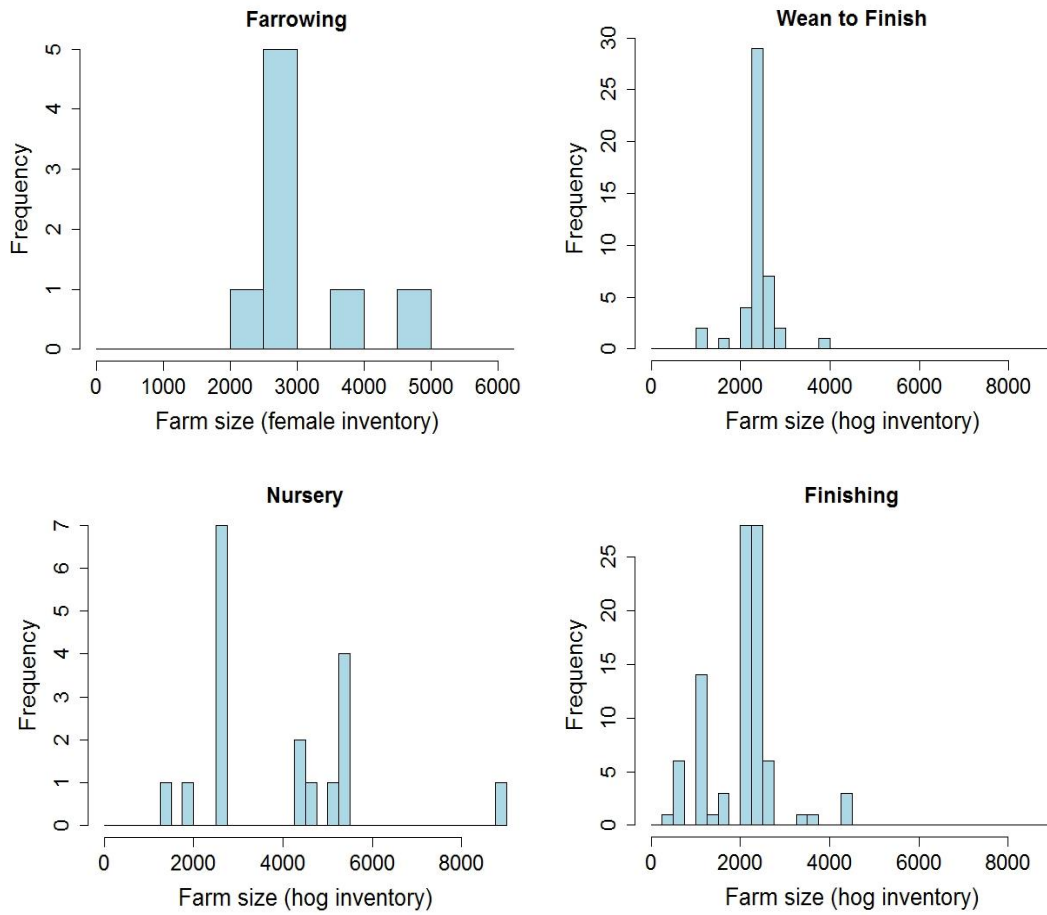
358

Figure 3: Farm size distribution by farm type (farrowing, wean-to-finish, nursery, finishing)

Table 7: Average annual number of pig shipments within the study period by operational classification

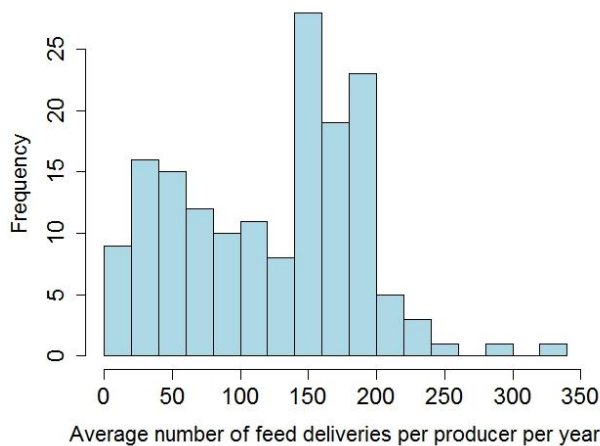| FROM	TO	N shipments FROM	N shipments TO	Avg. shipment size
Wean to Finish	Wean to Finish	1.9	1.6	259
Wean to Finish	Finishing	2.5	1.7	608
Finishing	Finishing	0.5	0.5	176
Nursery	Nursery	1.2	0.8	161
Nursery	Wean to Finish	1.5	1.5	652
Nursery	Finishing	45.7	8.6	512
Supplier	Nursery	156.6	41.2	582
Supplier	Wean to Finish	60.9	10.8	406
Nursery	Customer	0.8	4	3
Finishing	Packer / Customer	27.9	276.6	153
Wean to Finish	Packer / Customer	22.4	179.5	153
Nursery	Packer / Customer	0.5	1	2

Figure 4: Average number of feed deliveries per producer per year

Finally, we conducted a sensitivity analysis focusing on four key parameters, each representing a specific aspect of the model's architecture. Each parameter is varied in steps between 50% and 150% of the baseline values given in Table 6, with ten replications per step. Table 8 shows the elasticity of the response variable---average vulnerability---across this range. Figure 5 visualizes the sensitivity analysis data along with linear correlations.

360

Table 8: Elasticity of response variable (Avg. Vulnerability) resulting from variation between -50% and 50% of baseline values for four key parameters

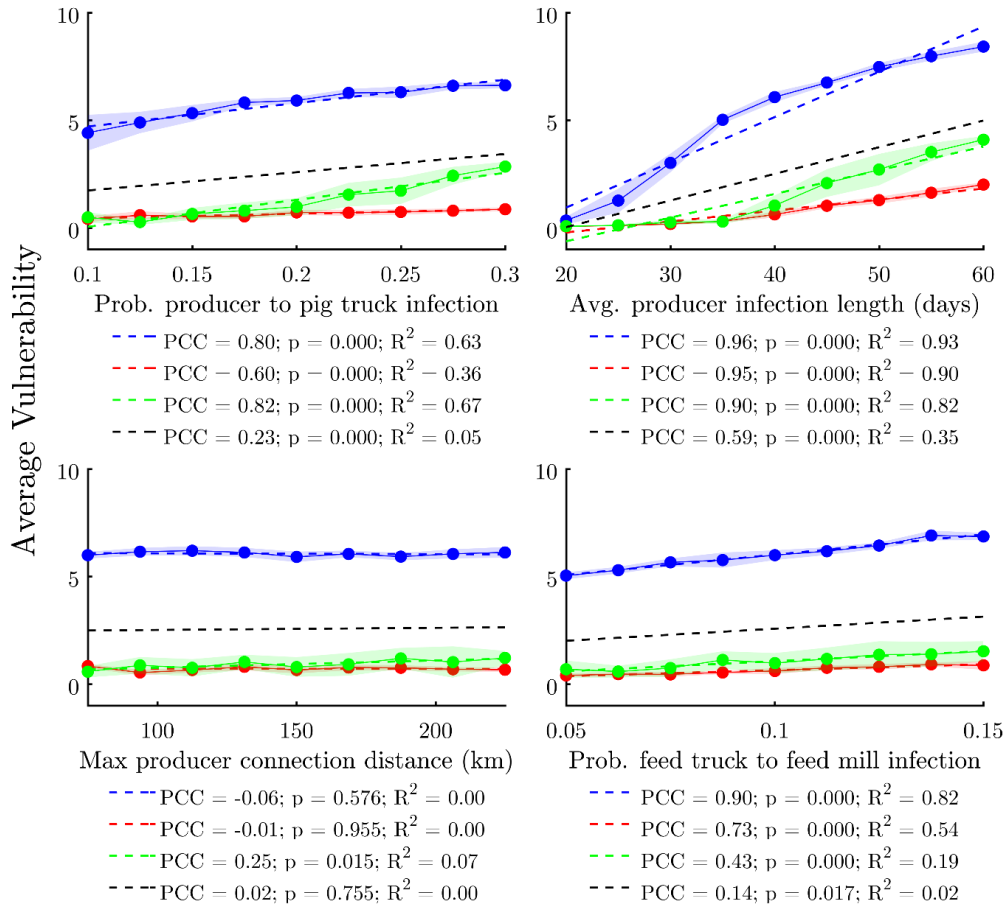| Parameter | % Change in Avg. Vulnerability | | | |
	North Carolina	Iowa	Illinois	All Study Areas
Prob. producer to pig truck infection	49.93	109.9	491.1	94.56
Avg. producer infection length (days)	2382	3164	6077	3010
Max producer connection distance (km)	2.182	-20.32	108.6	7.976
Prob. feed truck to feed mill infection	36.21	114.8	119.4	50.82

Figure 5: Sensitivity analysis plots for four key parameters. Scatter points show average values at each step, colored regions show 95% CIs, and dashed lines show linear trends. Blue represents North Carolina, red Iowa, green Illinois, and black the combined dataset. Pearson correlation coefficients, p-values, and R^2 values of linear regressions appear in legends below each figure.

Results of the sensitivity analysis show that the model is moderately-sensitive to changes in the probability of producer to pig truck infections and the probability of feed truck to feed mill infections; with the effect on average vulnerability being positive in all study areas, and generally highest in Illinois. For maximum producer connection distance, the magnitude and direction of the effect varies considerably between study areas, with Iowa demonstrating a negative relationship with the response variable. The R^2 values suggest that, overall, the model is not particularly sensitive to maximum producer connection distance. By contrast, increasing the average producer infection length causes significant increases in average vulnerability across all study areas. In light of previous SIR / SI model studies, the observation that average infection duration heavily impacts average vulnerability is not a surprise. The shape of the elasticity curves suggest that percolation dynamics may exist, with the nonlinearity---or percolation threshold---being lowest for North Carolina and higher for the other two study areas, corroborating findings from Wiltshire (2018).

References:

Barreteau, O., Antona, M., D'Aquino, P., Aubert, S., Boissau, S., Bousquet, F., . . . Mathevet, R. (2003). Our companion modelling approach. *Journal of Artificial Societies and Social Simulation, 6*(1).

Burdett, C. L., Kraus, B. R., Garza, S. J., Miller, R. S., & Bjork, K. E. (2015). Simulating the distribution of individual livestock farms and their populations in the United States: An example using domestic swine (Sus scrofa domesticus) farms. *PloS one, 10*(11), e0140338.

Goede, D., & Morrison, R. B. (2016). Production impact & time to stability in sow herds infected with porcine epidemic diarrhea virus (PEDV). *Preventive Veterinary Medicine, 123*, 202 - 207.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., . . . others. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological modelling, 198*(1-2), 115-126.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological modelling, 221*(23), 2760-2768.

Muller, B., Bohn, F., Dressler, G., Groeneveld, J., Klassert, C., Martin, R., . . . Schwarz, N. (2013). Describing human decisions in agent-based models: ODD+ D, an extension of the ODD protocol. *Environmental Modelling & Software, 48*, 37-48.

The Pig Site. (2014). Pig Farm Targets: Stockmanship Standards. Retrieved from http://www.thepigsite.com/stockstds/3/pig-farm-targets/

USDA NASS. (2014). Livestock Slaughter, 2013 Summary: Federally Inspected Plants and Head Slaughtered by Species, States and United States, 2013. http://usda.mannlib.cornell.edu/usda/nass/LiveSlauSu/2010s/2014/LiveSlauSu-04-21-2014.pdf

Wiltshire, S. W. (2018). Using an agent-based model to evaluate the effect of producer specialization on the epidemiological resilience of livestock production networks. *PloS one, 13*(3), e0194013.

Windrum, P., Fagiolo, G., & Moneta, A. (2007). Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation, 10*(2), 8.