Towards an Integrated Network-based Approach to Modeling the Dynamics of Invasive Plant Pests

Srinivasan Venkatramanan, Abhijin Adiga*, Achla Marathe, Stephen Eubank, Madhav Marathe
Biocomplexity Institute, Virginia Tech, Blacksburg
vsriniv, abhijin, achla, seubank, mmarathe@vbi.vt.edu

Rangaswamy Muniappan
IPM Innovation Lab, Virginia Tech, Blacksburg
rmuni@vt.edu

ABSTRACT

Globally, invasive alien species pose a growing threat to native ecosystems, health, food security and economic stability. The dynamics of biological invasions, particularly pests and pathogens associated with agricultural crops, are influenced not only by biological and climatic factors, but also human-centric activities such as trade, travel, supply chains, agricultural practices, etc. Modeling such a complex phenomenon requires integrating data, models and expertise from various disciplines. We survey the current modeling methods and identify their limitations, highlighting the computational and data challenges involved in modeling and simulating the spread of invasive species. We then envision a network-based modeling paradigm for building realistic models which represent and integrate various ecological and human-mediated pathways responsible for the rapid spread of pests. The framework will be presented in the context of a representative pest, the South American tomato leafminer which has been ravaging tomato crops across the globe. We believe that the methodology is generic enough to be easily adapted for other invasive alien species.

CCS Concepts

•Networks → Network structure; Network dynamics; •Computing methodologies → Modeling and simulation; •Applied computing → Agriculture; Forecasting;

Keywords

Alien invasive species; Integrated Pest management; Network synthesis; Modeling and Simulation

1. INTRODUCTION

One of the unintended fallouts of globalization has been the rapid increase in biological invasions causing unprecedented disruptions to native ecosystems [2], human, animal and plant health [13]. Such invasions compound the effects of climate change and threaten global food security and consequently economic stability [12]. In general, plant pathologists have acknowledged [3] the need to understand the dispersion dynamics of invasive pests that attack agricultural crops, causing significant yield reduction. In this case, the problem is further exacerbated if the invaded region is typically devoid of its natural predators, and trade & travel aid in its survival and rapid dispersal. This may lead to deployment of immediate myopic responses such as increased use of insecticides, trade restrictions, etc. While attempting to mitigate direct impacts due to crop loss, such measure may have indirect consequences on the environment, consumer-producer welfare, and at the extreme, political unrest. Thus there is an essential need to consider an integrated approach to pest management, spanning biological, ecological and socio-economic disciplines.

The entry, establishment and spread of invasive species is a complex process, with several factors affecting the dynamics. On one hand, we have ecological drivers such as climate, pest life-cycle, host availability, natural predators, etc. On the other we have anthropogenic drivers such as commodity flow, international trade, travel and migration, farming practices, and control measures ranging from pest management, trade & travel restrictions, phytosanitary measures, etc. A modeling framework should be capable of succinctly representing these components of the system at their appropriate scales as well as the interplay between them.

An ideal framework should provide reliable forecasts about the pest prevalence, with possibly causal explanations for the extent and dispersion. Further, in order to be practically useful, they should enable the policymaker to test various counter-factual scenarios, and evaluate the impact of different policies and interventions. In the context of aggressive pests, the model should help answer pertinent questions including: Which regions are highly vulnerable to invasion? Which locations need to be monitored for early detection and control? What are the primary and secondary effects of the prescribed control strategies? From the researcher's perspective, the model should guide data collection and assimilation, aid in hypothesis generation and explain the core dynamics. While these aspirations are well justified and widely acknowledged, the current repertoire of models do

*This author is an early career researcher interested in being considered for the travel fund.

DOI: 10.475/123_4

not completely satisfy the expectations, at least not in isolation.

In this paper, we will briefly survey the existing approaches to pest modeling and point out their limitations while building holistic models for pest invasion (Section 2); We then propose an interaction-based approach to modeling plant diseases (Section 3) with emphasis on the process of synthesizing network-based models for human-mediated pathways, while highlighting the research challenges of interest to the data science community. Finally, we discuss the potential analyses that can be carried out on such models and the impact they can have on policy design (Sections 4). Here, we will use the South American tomato leafminer (*Tuta absoluta*) as an illustrative example but the proposed methodology is generic enough to be extended to a wide range of biological invasions.

***Tuta Absoluta*:** *Tuta absoluta* or the South American tomato leafminer is a devastating tomato pest. It is native to South America and has spread throughout the Mediterranean, Europe and parts of Asia. It has been recently reported in sub-Saharan Africa, the Indian subcontinent and Central America (Panama and Costa Rica), and poses a serious threat to Mexico and the USA [1]. It can cause up to 100% loss in yield if no control measures are taken. It is known to overcome geographic and climatic barriers, surviving in greenhouses, and found in tomato crates/boxes. Insecticides have been the only control measure in deployment, though their effectiveness is limited due to pest's acquired resistance. No effective natural enemies have been identified for *T. absoluta*[4]. Studying the dispersion dynamics of such a rapidly spreading pest that severely affects a significant agricultural crop is both essential and timely.

2. RELATED WORK

Since invasive pest modeling is perched at the intersection of phyto-pathology, entomology and conservation ecology with palpable environmental and economic impact, we often find researchers addressing different aspects of the problem. In this section, we provide an overview of the various approaches. While we acknowledge the importance of these methods, we highlight their limitations especially when used in isolation.

Ecological models: These models are primarily aimed at predicting the geographic distribution and relative abundance of the species under various climate scenarios. The outputs of such models are typically in the form of a pest-risk map embedded into a geographic information system (GIS). There are two distinct approaches prevalent in the literature: niche- and process-based. Ecological Niche Models (ENM) (e.g. CLIMEX) are correlative models which map the known occurrence/absence records of the species to a set of climatic and edaphic variables thus characterizing the ecological niche of the species. On the other hand, process-based models (e.g., ILCYM) follow a deductive approach. Based on observations made in controlled field or laboratory studies, these models capture the physiological responses of a species to biotic and abiotic factors, providing the conditions in which the species can ideally persist and thrive.

Models for Anthropogenic dispersal: Given their flight range and life expectancy, the migratory capacity of *T. absoluta* is fairly limited. The very fact that it is rampant globally signifies the crucial role played by humans in its dispersion. Researchers have put forward disease models on complex networks as an approach for understanding long-distance dispersal pertaining to plant pathology [8]. Koch et al. [9] use empirical data on commercial trade and recreational transport to evaluate the potential forest pest invasion risk for more than 3000 urban areas in the U.S. Nopsa et al. [11] evaluate the structure of rail networks for grain transport to study the spread of post harvest diseases in wheat.

Intervention Models: A plethora of bioeconomic models that combine invasion ecology and economic analysis have been proposed [5] in the literature. By integrating human behavior models with the pest spread, one can make suitable recommendations for policymakers. Rebaudo et al. [14] study the success of IPM programs depending on the efficacy of information dissemination mechanisms in place, and the heterogeneity in farmer behavior. They use an ABM for modeling human mobility at the local level coupled with a cellular automaton model pest dynamics, to study the impact of farmer's mobility and knowledge of pest control on pest spread. Carrasco et al. [1] studied the effectiveness of control measures, when the compliance of farmers could depend on their learning/imitation behavior and the availability of information.

Limitations: There are several hurdles to implementing realistic integrated models for plant diseases. Cunniffe et al. [3] identify 13 challenges, broadly classifying them into those arising while modeling plant hosts, pathogens or control. Current modeling approaches either focus on a specific aspect of the problem or a small geographic area, thus resorting to a level of abstraction which does not capture the complexity of the problem. For instance, the ecological models, while capable of capturing climatic effects, cannot account for the proliferation of greenhouses where pests are provided with a year-round controlled environment free of climatic effects. On the other hand, there are no known models of anthropogenic dispersal that account for multiple modes (trade, travel, supply chain logistics) of dispersion. Finally, models for policy design rarely span across scales, from the farm level to the country level. They sidestep the fact that the socio-economic impact to individual farmers, tomato production firms and the national import/export need to be considered in conjunction.

3. AN INTEGRATED MODELING FRAMEWORK

Despite these challenges, one can be inspired by the development in computational epidemiology over the past couple of decades, where similar hurdles have been surmounted by the use of detailed computational models [10]. Aided by advances in surveillance, big data analytics and high performance computing, researchers in computational epidemiology have embraced this methodology with great success [7]. For instance, in computational epidemiology, modelers have traditionally used compartmental models implemented through coupled differential equations on completely mixed populations to understand the spread of diseases. Though they are quick to setup and analyze, they provide causal explanations or policy recommendations at a very

[1]https://gd.eppo.int/taxon/GNORAB/reporting

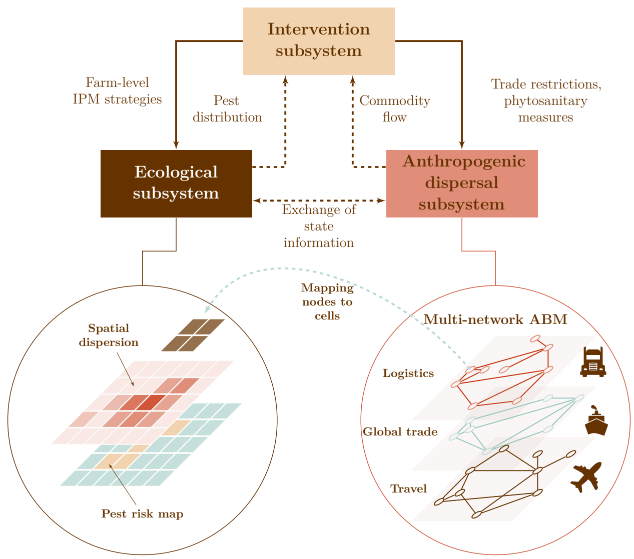

Figure 1: Integrated modeling framework for invasive plant pests

coarse scale, nor are they equipped to handle spatial and social heterogeneity. In recent years, data-driven network models have been very successful in this regard.

With origins in statistical physics, these complex network models have quickly gained traction across disciplines owing to the availability of (a) data that yields itself to a network abstraction, (b) a rich collection of mathematical tools from graph theory and dynamical systems, and (c) the computational power to run detailed processes on such networks. Further, network analysis is capable of revealing emergent patterns that may not be evident from raw statistical analysis of the available data. Viewing the data aggregated at different resolutions in conjunction with the interactions allows us to characterize the role played by different nodes and edges that lead to observed system behavior by means of their position in the network [11]. Studying networks with explicit hierarchy allows for policy design at various spatial and system scales.

In the study of invasive species, especially in the context of plant pathology, though there has been growing interest in using network models, [3, 8], efforts have been limited. Since human-induced pest spread has been widely acknowledged, mapping out these networks and studying their impact will be crucial to understand global spread of invasive pests. While complementing the ongoing research on food security and economic vulnerability via study of international food trade networks [6], they also allow positioning decades of entomological/ecological research in a wider context.

We envisage a multi-layered network-based modeling framework that has the potential to capture the various dynamics of biological invasion shown in Figure 1. The 'bottom' layer corresponds to the ecological subsystem which represents the influence of biotic and abiotic factors on the dynamics, provided in the form of a pest risk map. One can then embed a patch-based spatial model, say a cellular automata, which models the natural dispersion of pests across neighboring farms. The 'top' layer is used to model the effects of human activities of pest dispersal, conveniently modeled using a multi-layered hierarchial network capturing the flow of commodity and people at different spatial and temporal resolutions. This could capture farmer movement between fields, flows from farms to markets, flow of seedlings among greenhouses, flow of tomatoes along the supply chain to the consumer, all the way up to international trade and travel. The crucial aspect of integration is that nodes in this layer are embedded spatially, thus allowing the model fetch/feed state information from/to ecological subsystem. Though described as static, these networks can be extended temporally, by incorporating seasonal and long-term trends. Dispersion dynamics are then simulated using stochastic contagion processes on these networks. Finally, the intervention subsystem allows one to observe the state at any scale and implement policies in either of the subsystems.

3.1 Models for anthropogenic dispersal

We believe that implementation of the ecological subsystem can be bootstrapped using standard tools available in the community (e.g., CLIMEX, ILCYM). Our focus will be on the hierarchical network that models human influence on pest dispersion, which has received far less attention. In this section, we describe the stages in constructing the multi-layer network while bringing out the key challenges for the data/computational scientist.

(a) Network synthesis: Some domains, such as the study of online social networks, roadway networks, etc. have a natural advantage that data is gathered and presented in network form. For instance, in our context, FAO [2] provides a detailed trade matrix for tomatoes and its products, which has a natural network representation (weighted and directed), where countries represent the nodes, and export/import trade volume is assigned to the edges.

However, such an explicit network is usually unavailable at the regional level. As in computational epidemiology [7], when the network is not directly available, the onus is on the modeler to match the available data (such as activity patterns, household locations, population demographics) into a network paradigm. In the domain of food security, similar efforts have been carried out to map the livestock distribution in [15]. For instance, for the commodity flow within United States, sources such as National Agricultural Statistical Service (NASS) [3] provide county-level information such as harvest area, crop yield (intended for fresh market or processing), number of tomato farms. Other sources such as the Agricultural Marketing Service (AMS) [4] provide an incomplete view of the domestic flows, by tracking tomato crates between shipping points and a few chosen terminal markets. These are gathered for different purposes and thus need not work well together. Further some data may be unavailable due to commercial sensitivity. One may face several challenges while integrating these data sources including (i) mismatch in spatio-temporal scales (ii) inter-data discrepancies (iii) inherent data gaps. Further, for developing countries, the data is even sparser, requiring a more intensive effort in inferring the network. *Network inference by integrating such varied, incomplete and noisy sources is an interesting challenge in itself.*

(b) Structural and Dynamic analysis: Having obtained the network, one can use network measures to char-

[2]http://faostat.fao.org
[3]http://www.agcensus.usda.gov
[4]http://www.ams.usda.gov

acterize the importance of various nodes and interactions to the underlying process [11]. One can identify nodes that are highly vulnerable to, or critical for the dispersion process, by virtue of their location in the network [16]. *The key challenge is to identify the network measures that are appropriate and easy to interpret.*

In order to simulate, compare and forecast pest dispersion, one must define a stochastic contagion process on the multi-layer network. Though several standard disease models exist in epidemiology, none have yet been successfully adapted to the domain of plant pest dynamics. *The primary challenge is to appropriately choose the disease model, keeping in mind the available data to parametrize and validate the model.* Pest surveillance and interception are resource intensive, and in most cases, the timeline and extent of pest invasion is poorly documented. Adding to this, there are meteorological, market and political uncertainties, which make forecasting and policy evaluation extremely difficult. Having calibrated the model, *the uncertainties in data and the model itself must be reflected in the uncertainty quantification (UQ) of the forecasts.* Finally, complex agent-based models are *computationally intensive* and efforts need to be taken to efficiently simulate and analyze these models.

4. DISCUSSION

We believe that the proposed framework will lead to a more informed decision making and a better understanding of the underlying process. In addition to providing reliable forecasts, one can perform economic and environmental impact analysis on these networked models. One can study the impact of pest damage as well as of the proposed intervention policies. The explicit hierarchy in our framework allows to perform impact analysis across scales. Finally, we hope it would enable researchers across several disciplines to converge and collaborate on a common platform of practical and immediate need.

5. ACKNOWLEDGMENTS

This material is based upon work supported by the United States Agency for International Development under the Cooperative Agreement NO. AID-OAA-L-15-00001 Feed the Future Innovation Lab for Integrated Pest Management. The authors are also partially supported by DTRA CNIMS Contract HDTRA1-11-D-0016-0001 and NSF NetSE Grant CNS-1011769.

6. REFERENCES

[1] L. R. Carrasco, D. Cook, R. Baker, A. MacLeod, J. D. Knight, and J. D. Mumford. Towards the integration of spread and economic impacts of biological invasions in a landscape of learning and imitating agents. *Ecological Economics*, 76:95–103, 2012.

[2] T. A. Crowl, T. O. Crist, R. R. Parmenter, G. Belovsky, and A. E. Lugo. The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment*, 6(5):238–246, 2008.

[3] N. J. Cunniffe, B. Koskella, C. J. E. Metcalf, S. Parnell, T. R. Gottwald, and C. A. Gilligan. Thirteen challenges in modelling plant diseases. *Epidemics*, 10:6–10, 2015.

[4] N. Desneux, E. Wajnberg, K. A. Wyckhuys, G. Burgio, S. Arpaia, C. A. Narváez-Vasquez, J. González-Cabrera, D. C. Ruescas, E. Tabone, J. Frandon, et al. Biological invasion of european tomato crops by tuta absoluta: ecology, geographic expansion and prospects for biological control. *Journal of Pest Science*, 83(3):197–215, 2010.

[5] R. S. Epanchin-Niell and A. Hastings. Controlling established invaders: integrating economics and spread dynamics to determine optimal management. *Ecology letters*, 13(4):528–541, 2010.

[6] M. Ercsey-Ravasz, Z. Toroczkai, Z. Lakner, and J. Baranyi. Complexity of the international agro-food trade network and its impact on food safety. *PloS one*, 7(5):e37810, 2012.

[7] S. Eubank, H. Guclu, V. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 2004.

[8] M. J. Jeger, M. Pautasso, O. Holdenrieder, and M. W. Shaw. Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist*, 174(2):279–297, 2007.

[9] F. H. Koch, D. Yemshanov, R. D. Magarey, M. Colunga-Garcia, and W. D. Smith. Evaluating the forest pest invasion potential of trade-related and recreational transportation pathways. In *2011 International Conference on Ecology and Transportation (ICOET 2011)*, 2012.

[10] M. Marathe and A. K. S. Vullikanti. Computational epidemiology. *Communications of the ACM*, 56(7):88–96, 2013.

[11] J. F. H. Nopsa, G. J. Daglish, D. W. Hagstrum, J. F. Leslie, T. W. Phillips, C. Scoglio, S. Thomas-Sharma, G. H. Walter, and K. A. Garrett. Ecological networks in stored grain: Key postharvest nodes for emerging pests, pathogens, and mycotoxins. *BioScience*, 65(10):985–1002, 2015.

[12] D. Pimentel, S. McNair, J. Janecka, J. Wightman, C. Simmonds, C. OConnell, E. Wong, L. Russel, J. Zern, T. Aquino, et al. Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment*, 84(1):1–20, 2001.

[13] P. Pyšek and D. M. Richardson. Invasive species, environmental change and management, and health. *Annual Review of Environment and Resources*, 35:25–55, 2010.

[14] F. Rebaudo and O. Dangles. An agent-based modeling framework for integrated pest management dissemination programs. *Environmental modelling & software*, 45:141–149, 2013.

[15] T. P. Robinson, G. W. Wint, G. Conchedda, T. P. Van Boeckel, V. Ercoli, E. Palamara, G. Cinardi, L. D'Aietti, S. I. Hay, and M. Gilbert. Mapping the global distribution of livestock. *PloS one*, 9(5):e96084, 2014.

[16] S. Sutrave, C. Scoglio, S. A. Isard, J. S. Hutchinson, and K. A. Garrett. Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS One*, 7(6):e37793, 2012.