ABSTRACT

MONTGOMERY, KELLYN PAIGE. Analytical Approaches for Plant Pest Management Across the Biosecurity Continuum. (Under the direction of Dr. Helena Mitasova).

Non-native plant pests and pathogens threaten biodiversity, ecosystem function, food security, and economic livelihoods. Preventing, minimizing, and mitigating damage from invasive non-native plant pests involve dynamic, multiscale strategies that are agile enough to respond to rapidly changing conditions while balancing economic and environmental trade-offs. Geospatial analytical approaches are uniquely suited to address this challenge by leveraging spatio-temporal data to understand patterns of outbreaks and create predictive models that support management decision-making. In this dissertation, I develop and apply geospatial tools and analytics for plant pest management at global, national, and local scales. In Chapter 2, I present a species-agnostic, spatio-temporal stochastic network model called PoPS (Pest or Pathogen Spread) Global that couples international trade networks with core drivers of biological invasions to forecast the spread of non-native plant pests. The modular design of the framework supports early, proactive responses for emerging pests even when limited data are available and enables forecasts at different spatial and temporal resolutions. I demonstrate the framework using a case study of the invasive planthopper spotted lanternfly, *Lycorma delicatula* (White) (Hemiptera: Fulgoridae), and test management intervention scenarios. Simulating spotted lanternfly spread through trade networks, the model predicted that Japan, S. Korea, the U.S., and Germany had the highest probabilities of bridgehead population establishment. In Chapter 3, I develop an open-source software package called PoPS (Pest or Pathogen Spread) Border for measuring inspection outcomes for consignments with variations in size, cargo configuration, contamination rates, and contaminant arrangements for designing effective risk-based sampling protocols. I use the tool to estimate contamination rates from historical interception data, quantify trade-offs in effectiveness and workload for inspection strategies, and identify vulnerabilities in sampling protocols as changes in cargo configurations and contamination occur. This work represents first steps toward a decision support tool for creating dynamic inspection protocols that respond to changes in available resources, workload, and commerce trends. In Chapter 4, I investigate the use of 3D crop canopy mapping with unmanned aerial systems (UAS) for field-scale crop scouting and stress detection. Using nutrient stress in flue-cured tobacco as a case study, I computed structural metrics and a visible band spectral index

from high resolution canopy surface models and orthoimages to assess the relationship between plant nutrient status and UAS-derived metrics. Combining information about canopy structure and spectral reflectance increased model fit for all measured nutrients compared to spectral alone. These results demonstrate that an important relationship exists between relative canopy shape and crop health that can be leveraged to improve the usefulness of low-cost UAS for remote crop stress monitoring. The research presented in this dissertation advances multiscale, interdisciplinary approaches that contribute to reducing global migration of non-native species, improving biosecurity efforts at ports of entry, and detecting field-scale crop stress.

Analytical Approaches for Plant Pest Management Across the Biosecurity Continuum.

by
Kellyn Paige Montgomery

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Geospatial Analytics

Raleigh, North Carolina
2021

APPROVED BY:

_____	_____
Helena Mitasova Committee Chair	Ross Meentemeyer
_____	_____
Jean Ristaino	Anders Huseth
_____	
Yu Takeuchi	

BIOGRAPHY

Kellyn was born in Norfolk, Virginia and raised in Wytheville, Virginia. She completed her undergraduate degree in Geospatial and Environmental Analysis at Virginia Tech in Blacksburg, Virginia. She also earned a Master of Science in Geography at Virginia Tech where she studied the spatial and gender dimensions of integrated pest management adoption by small-scale farmers in rural Uganda. She worked for seven years in project and program management across multiple sectors, including industry, local government, and non-profit, with a focus on sustainable agriculture and community development. Through that experience, she decided to return to graduate school to develop her technical skills in geospatial data science for agriculture and natural resource management. She came to North Carolina State University to work under the direction of Dr. Helena Mitasova in 2017 as a graduate student in the Marine, Earth and Atmospheric Sciences department before joining the Center for Geospatial Analytics as a student in the first cohort of the Geospatial Analytics doctoral program in 2018.

ACKNOWLEDGMENTS

First, I thank my advisor, Helena Mitasova, for warmly welcoming me to her lab group and for her constant encouragement and guidance from the first day we met. I also want to thank Ross Meentemeyer for exposing me to exciting research in biological invasions through his lab group and the many opportunities he gave me throughout my research assistantship with the USDA APHIS. The combined support from Dr. Mitasova and Dr. Meentemeyer ensured my success as a doctoral student and researcher. I am grateful for the many collaborators I've worked with at the Center for Integrated Pest Management and USDA APHIS, including Yu Takeuchi, Cathy Sue Katsar, Sunil Kumar, Manuel Colunga, Ernie Hain, Steve Hong, Allison Neeley, and many others. In particular, thank you to Dr. Takeuchi whose support and advocacy throughout my research assistantship has been invaluable. Thank you to Anders Huseth and Jean Ristaino for the candid, insightful conversations that helped me put my ideas into context as I developed my research. I also want to thank the professors, Josh Gray, Laura Tateosian, and others, who taught me the foundations of geospatial data science to be successful in my research and beyond. To the Center for Geospatial Analytics community in general, including faculty, staff, and students, I am so grateful for the positive, supportive environment you created that has resulted in one of the most enriching experiences of my life. I also want to acknowledge my cohort of fellow inaugural doctoral students who have made the Ph.D. experience so much fun. Lastly, I am endlessly grateful to my mother, Pamela, and father, Paul, for their unconditional love and support and my dog, Hank, and cat, Phoebe, for helping me stay active and sane throughout it all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1 The impact of invasive plant pests and diseases on global food security

Food security and agricultural sustainability are complex, global challenges that require balancing the nutritional demands of a growing human population with the imperative to conserve biodiversity and natural resources. In 2020, nearly one in three people globally did not have adequate access to food, an issue exacerbated by the COVID-19 pandemic (FAO, 2021). Feeding the human population in 2050 is projected to require a 70% increase in food production (FAO, 2009). Conventional agricultural production practices, however, have detrimental impacts on the environment through habitat loss, freshwater nutrient loading, groundwater depletion, soil degradation, and carbon emissions (Harrison et al., 2012). To support the need for increased productivity while minimizing environmental degradation, conservation efforts must increase at the same time. Sustainable agricultural intensification is a way to approach these challenges that aims to maximize crop yields without expanding into non-cultivated lands or causing environmental harm (Pretty & Bharucha, 2018). A major component of sustainable agricultural intensification is minimizing crop loss from pests and pathogens, which have been estimated to cause 8 to 41% yield loss in major grain crops worldwide (Savary et al., 2019). Furthermore, crop losses to agricultural pests and pathogens are predicted to increase with rising global temperatures caused by climate change (Derocles et al., 2018; Deutsch et al., 2018; Elad & Pertot, 2014). Rates of non-native plant pests and pathogens invasions are also increasing, caused by globalization and unintentional transport of pest species to countries where they were previously absent (Bebber et al., 2019; Hulme, 2009; Ristaino et al., 2021). Alien species invasions can cause widespread damage to natural and managed ecosystems, resulting in species extinctions and habitat alteration. Agriculture is particularly at risk due to its dependency on healthy ecosystems with natural predators to keep pest populations under control (Carter et al., 2004).

Preventing, minimizing, and mitigating damage from invasive non-native plant pests and diseases involves strategic efforts at multiple scales. Internationally, regulatory policies on phytosanitary standards for traded goods help minimize the transport of pests through trade networks. Pre-border phytosanitary programs are also used to increase the capacity of trading partner countries to manage and prevent pests and to perform a screening of commodities prior to transit. At national borders, incoming consignments, packing material, and containers are

inspected for contaminants, which may include soil, arthropods, pathogens, or any other foreign material. Contaminated consignments are treated to eliminate the contaminants, if possible, or are prevented from crossing borders. Post-border in the United States, coordinated surveillance programs monitor local areas for newly established non-native pests to detect invasions as early as possible. Once a pest is introduced, quarantine and eradication efforts are used to slow population growth and spread, especially to sensitive ecosystems, crop producing regions, or other areas where invasion impacts would be significant. This broad spectrum of intervention strategies for non-native pest management is collectively known as the biosecurity continuum (Magarey et al., 2009). Efforts across the biosecurity continuum are supported by data science approaches and decision analytics to aid managers in targeting limited resources and policymaking. In this chapter, I provide a review of the analytical approaches most relevant to this dissertation that have been used to inform plant biosecurity measures from international phytosanitary regulations to field-level pest monitoring. I conclude with an introduction to the analytical approaches contributed by this dissertation and future directions of biological invasions data science.

2 Analytical approaches used across biosecurity continuum

Understanding the drivers of biological invasions is critical to improving our ability to anticipate sources and timing of global pest spread. Historical data have been used to establish a clear relationship between international trade networks and non-native species invasions (Chapman et al., 2017; Turbelin et al., 2017). These studies document the dominant role of economic development and globalization in human-mediated pest spread and provide evidence of the predictive power of trade to support biosecurity analytics. Other studies have found that certain taxa are particularly invasive and are associated with trade pathways by comparing species distributions and assemblages across regions (Liebhold et al., 2016). Invasion pathways have been categorized based on their spatio-temporal patterns to facilitate prioritization of biosecurity efforts (Bertelsmeier et al., 2017; Faulkner et al., 2020). Seebens et al. (2017) quantified rates of biological invasions and showed that they have dramatically increased over the past two centuries with no signs of slowing. However, Sikes et al. (2018) presented encouraging evidence that biosecurity efforts have reduced rates of invasions in focal

commodities, including crops and livestock, further highlighting the importance of data-driven techniques to improve targeted management.

Researchers have developed predictive models to forecast invasions by leveraging their close relationship with trade and travel patterns. Many approaches use statistical or machine learning models with climate information and historical trade or air travel volumes to predict the potential for future invasions and identify areas most at risk (Levine & D'Antonio, 2003; Paini et al., 2016; Tatem, 2009; Tatem & Hay, 2007). Faulkner et al. (2017) presented an approach for evaluating the vulnerability of specific ports to new invasions by analyzing the route characteristics of ocean-going vessels, such as the duration and frequency of trips and how often they link environmentally similar ports within the distribution of known invasive species. Other approaches have focused on trade or transport network topology to identify hubs for spread to focus biosecurity on those nodes in the network (Banks et al., 2015; Floerl et al., 2009; Paini & Yemshanov, 2012). Many studies have demonstrated the bridgehead effect, in which successfully invasive populations become additional sources for invasions and accelerate global spread. Studies are revealing that invasion pathways are more complex than previously thought, with repeat introductions from multiple population sources and genetic admixture that can increase invasive population fitness and establishment success (Bertelsmeier et al., 2018; Garnas et al., 2016). Increasing availability of genetic information has also made it possible to simulate the likely invasion routes using approximate Bayesian computation, which has provided strong evidence for the role played by bridgehead populations and transportation hubs (Blumenfeld et al., 2021; Boheemen et al., 2017; Lombaert et al., 2010). Population genetics and historical border interception data have also been useful for demonstrating global spread pathways through bridgehead populations for multiple invasive insects (Bertelsmeier et al., 2018; Blumenfeld & Vargo, 2020; Correa et al., 2019) and pathogens (Fontaine et al., 2020). Approaches for modeling complex patterns of global species movement and establishment through trade and bridgehead populations show promise for helping phytosanitary agencies anticipate and prioritize future invasions.

Phytosanitary inspections of imported goods are a crucial first line of defense to minimize the movement of non-native species across national borders. However, given the volume of goods imported daily and resource limitations, inspectors employ a wide range of sampling strategies that prioritize detection over consistency, resulting in nonstatistical sampling

and limitations to using the data collected for statistical inference (International Plant Protection Convention, 2016; Saccaggi et al., 2016). Extensive work has been done on ways to shift towards targeted risk-based sampling to match inspection effort to risk levels. For example, shipment risk has been quantified using statistical models that relate pest interceptions to commodity, country of origin, and shipment size in cut flowers (Areal et al., 2008), propagative material (Kim et al., 2019; Surkov, van der Werf, et al., 2008), fruits and vegetables (Lichtenberg & Olson, 2018), and goods coming through specific ports of entry (Caton et al., 2006). Similar models have also been developed to quantify risk of pest and pathogen transport in passenger baggage (Gottwald et al., 2019; Szyniszewska et al., 2016)**.** More complex inspection optimization models also incorporate inspection workload and exporter response to phytosanitary policies to dynamically allocate inspection resources (O. Jones et al., 2017; Rossiter & Hester, 2017; Springborn et al., 2016; Surkov, Oude Lansink, et al., 2008). While these models have potential to be useful for designing efficient inspection programs, ultimately, they are only effective if accurate estimates of risk are available (Powell, 2015). Several studies have proposed sampling approaches to simultaneously prioritize minimizing pest slippage (i.e., proportion of pests or contaminants that fail to be intercepted by border inspections) and maximizing information gathering to continually monitor contamination rates (Chen et al., 2018; Robinson et al., 2011, 2015; Trouvé & Robinson, 2020)**.** Inspections at national borders play an important role in the biosecurity continuum by minimizing entry of non-native pest species while also providing useful information on what pests are not being detected.

The number of non-native organisms missed by border inspections that then successfully enter new areas is referred to as propagule pressure. Propagule pressure is an important predictor of invasion success and can be used to inform invasive species management and policy (Lockwood et al., 2005; Reaser et al., 2008). While direct measurement of propagule pressure is not feasible, many proxy measures have been used in the literature, including border interceptions and trade volumes (Lockwood et al., 2009). Several studies have provided summaries of border interceptions to characterize rates of pest arrival through different trade pathways (Jenkins et al., 2014; Liebhold et al., 2006; McCullough et al., 2006; Work et al., 2005). Others have attempted to explain invasion rates using historical interception data with statistical models, although many found interceptions to be a poor predictor of establishment (Brockerhoff et al., 2014; Caley et al., 2015; Eschen et al., 2015). Bacon et al. (2014) showed

that interception data was predictive of establishment only when combined with other key environmental suitability measures, like climate and host availability. An inverse approach for measuring relative propagule pressure was used in Bacon et al. (2012) by comparing an inspection effectiveness metric (trade volume per interception) between countries to identify areas that are likely failing to intercept incoming pests. Turner et al. (2020) used interception data in a stochastic process-based model of arrival, interception, and establishment to shed light on species that are likely to establish but go undetected at ports of entry. While interception data can provide valuable insights, their usefulness for predicting invasions are limited by lack of statistical sampling and the complex relationship between propagule pressure and invasion success (Lockwood et al., 2009).

Once an alien species is transported to new areas, it is important for pest managers to understand where it could establish to assess potential impacts. There are many approaches for estimating the potential distribution of a species, depending on data availability and the geographic extent to be mapped (Elith, 2017; Venette et al., 2010). Species distribution models that use a correlative approach, such as MaxEnt (Phillips et al., 2006), model the relationship between species occurrence data and environmental covariates to predict areas suitable for establishment. This approach is useful when mapping risk over a large geographic extent when little is known about species biology. However, these models come with a risk of overfitting to species presence data (Venette et al., 2010). Also, Václavík & Meentemeyer (2012) demonstrated that the potential range tends to be underpredicted at the earliest stages of an invasion when few occurrence points are available in the invaded range. Bebber et al. (2019) presented a statistical approach for handling global underreporting of invasive species occurrences by predicting species presences and psuedo-absences using regional socioeconomic data and environmental suitability. Alternatively, mechanistic models of species distribution, such as CLIMEX (Kriticos et al., 2015), are based on an understanding of the processes that limit species distribution, such as phenology (i.e., developmental temperature thresholds) and climatic tolerances (e.g., cold stress, heat stress, dry stress, and moisture stress thresholds). Species distribution models generally do not take dispersal into account and instead provide information about the potential range if dispersal limitations are removed.

Modeling human-mediated and natural dispersal mechanisms to capture propagule pressure in combination with environmental suitability can provide useful information to target

post-border management. Many modeling approaches are available that capture invasive species establishment and spread, ranging from static network models to dynamic spatio-temporal grid-based diffusion models to multiscale models incorporating local and long-range dispersal (Buchadas et al., 2017; Frost et al., 2019; Meentemeyer et al., 2012; Yemshanov et al., 2017). Dynamic simulations are especially useful for supporting decision making by providing information about risk uncertainty and the range of possible invasion outcomes (Buchadas et al., 2017; Yemshanov et al., 2017). Dynamic network models have been used for biological invasions to simulate human-mediated dispersal through transportation networks (Frost et al., 2019). For example, marine biological invasions have been simulated in multiple studies by coupling maritime shipping networks with environmental conditions at ports to forecast future invasions (Keller et al., 2011; Sardain et al., 2019; Seebens et al., 2013, 2016). International airline travel has been used with demographic information to predict areas where exotic infectious pathogens are likely to be transmitted to human populations (Gottwald et al., 2019). Network models have also been developed for regional scale spread of forest insect pests and pathogens (Ferrari et al., 2014; Koch et al., 2011, 2014) and highly localized, spatially explicit spread of a plant fungus (Brooks et al., 2008).

Dynamic grid-based spread models use host distributions with local or long-range dispersal kernels to simulate stochastic spread of pests or pathogens across landscapes over time (Filipe et al., 2012; C. Jones et al., 2021; Meentemeyer et al., 2011). Petrasova et al. (2020) presented an approach for allowing users of a stochastic, spatially explicit grid-based spread simulation to interactively test management options. Multiscale models simulate multiple dispersal mechanisms and often use a combination of modeling approaches. As an example from human epidemiology, Balcan et al. and Colizza et al. (2010; 2006) used a network model to capture global and regional movement of infectious individuals and a SIR epidemiological model to capture local disease dynamics. Seebens et al. (2019) also used a network model of global marine invasive species transport with local population growth models to simulate bridgehead population formation. Although computational models have been extensively used to study biological invasions, most models of terrestrial plant pest invasions have focused on either local or regional scales and do not attempt to model the multiscale processes of bridgehead effects.

The invasion models described above can be used to help target on-the-ground surveillance of priority pests. For example, mechanistic species distribution models are being

developed by researchers at the Center for Integrated Pest Management for use by the Cooperative Agricultural Pest Survey (CAPS) to plan and prioritize surveillance for harmful plant pests in the United States (NC State University, 2021; Purdue University, 2021). More complex approaches for prioritizing surveys have also been developed, including modeling domestic commodity pathways to identify the highest risk areas (Colunga-Garcia et al., 2010, 2013), using models of disease dynamics to prioritize areas with highest infection likelihood (Parnell et al., 2017), and prioritizing data collection in areas of highest uncertainty to improve model performance (Bourhis et al., 2020). Approaches used for plant pest surveillance range from on-the-ground surveys with lures and traps to web scraping and text mining social media. Emerging technologies for plant pathogen diagnostics, such as biosensors and hyperspectral imaging, show promise for improving early detection to prevent disease outbreaks (Hu et al., 2020; Ristaino et al., 2021). Crowdsourced and citizen science data has also become an important resource for information about where pest species are being discovered globally, reducing latency between data collection and availability (Brown et al., 2020; Jarić et al., 2021). The success of surveillance for preventing invasions has been linked to communication between stakeholders, coordination, and data sharing across the globe, particularly for low-income countries (Carvajal-Yepes et al., 2019; Milne et al., 2020).

Remote sensing via near-range devices, manned or unmanned aerial vehicles, and satellites is also an important tool for pest and disease surveillance. This broad category of tools includes spectral, thermal, fluorescence, synthetic aperture radar (SAR), and light detection and ranging (lidar) sensors (Zhang et al., 2019). Spectral image analysis has had the most applications for pest and disease detection by leveraging distinct spectral reflectance signatures of plant tissue experiencing different types of stress. Hyperspectral sensors, which measure spectral reflectance with very fine resolution, in contrast to broad resolution multispectral sensors, are a promising frontier in remote sensing for crop stress detection. However, hyperspectral remote sensing is currently limited by sensor costs, data volumes, difficulty of data interpretation, and limited information on stress spectral signatures (Hu et al., 2020; Mahlein, 2016; Mahlein et al., 2018). Lidar has been used to measure airborne pest densities (Dwivedi et al., 2020) and create detailed point clouds of crop geometry (Madec et al., 2017; Paulus, 2019; Zhang et al., 2019)**.** The photogrammetric technique Structure from Motion (SfM) can also be used to compute 3D point clouds and highly detailed surface models that provide valuable

information about vegetation height and heterogeneity (Carrivick et al., 2016). For example, Hunt & Rondon (2017) used very high-resolution canopy surface models to identify foliar pest damage in potatoes. The biggest challenges for using remote sensing for pest and disease surveillance are early detection and distinguishing the source of the stress, which may be caused by multiple sources (Zhang et al., 2019).

3 New approaches for pest management

Many analytical approaches have been developed to improve our understanding and ability to predict biological invasions of plant pests and pathogens. While some multiscale models exist for simulating human disease pandemics and marine biological invasions (Balcan et al., 2010; Seebens et al., 2019), no such frameworks have been developed for forecasting global spread of terrestrial plant pests and pathogens through bridgehead populations over space and time. In Chapter 2, I present a species-agnostic, spatio-temporal stochastic network model that couples international trade networks with core drivers of biological invasions—climate suitability, host availability, and propagule pressure—quantified through open, globally available databases to forecast the spread of non-native plant pests. The modular design of the framework supports early, proactive responses for emerging pests even when limited data are available and enables forecasts at different spatial and temporal resolutions. I demonstrate the framework using a case study of the invasive planthopper spotted lanternfly, *Lycorma delicatula* (White) (Hemiptera: Fulgoridae), and test management intervention scenarios. Simulating spotted lanternfly spread through trade networks, the model predicted that Japan, S. Korea, the U.S., and Germany had the highest probabilities of bridgehead population establishment. This global view of phytosanitary pandemics provides crucial information for anticipating biological invasions, quantifying transport pathways risk levels, and allocating resources to safeguard native plant health, agricultural crops, and ecosystem function.

Phytosanitary inspections at ports of entry play an important role in preventing entry of non-native species. Records of the pests found during these inspections provide valuable information on what species are being successfully intercepted. However, nonstatistical sampling approaches limit the ability to make inferences about propagule pressure and actual risk levels of trade pathways. Phytosanitary agencies need better tools for quantifying pests going undetected and designing risk-based inspection strategies appropriate for changing operational conditions. In

Chapter 3, I present an open-source consignment inspection simulator for measuring inspection outcomes under various cargo contamination scenarios to support recommendations for inspection protocols and estimating pest slippage rates. I use the tool to estimate contamination rates from interception data collected using statistical sampling, quantify trade-offs in effectiveness and workload for inspection strategies, and identify vulnerabilities in sampling protocols as changes in cargo configurations and contamination occur. These use cases demonstrate how this simulation approach permits testing inspection strategies and measuring quantities that would otherwise be impossible in a field-based setting. This work represents the first steps toward a decision support tool for creating dynamic inspection protocols that respond to changes in available resources, workload, and commerce trends.

Surveillance and monitoring for priority, invasive plant pests and pathogens requires a range of approaches including improved sensors for in-field detection, predictive modeling and analytics, formal and informal networks for data collection, and remote sensing. Small unmanned aerial systems (UAS) have emerged in recent years as a versatile remote sensing tool that can provide precisely timed, fine-grained data for crop monitoring that more accurately accounts for spatial variations in crop stress across a field. UAS sensors with high spectral resolution used to compute informative vegetation indices, however, are practically limited by high cost and data dimensionality. In Chapter 4, I present an approach that extends spectral analysis for remote crop monitoring to investigate the relationship between crop health and 3D canopy structure using low-cost UAS equipped with consumer-grade RGB cameras. Using nutrient stress in flue-cured tobacco as a case study, my results demonstrate that an important relationship exists between relative canopy shape and crop health that can be leveraged to improve the usefulness of low-cost UAS for remote crop stress monitoring.

Managing agricultural pests and pathogens is a complex, multiscale task that requires dynamic response to abiotic and biotic drivers while balancing economic and environmental trade-offs. Geospatial analytical approaches are uniquely suited to address this challenge by leveraging diverse spatio-temporal data to understand patterns of outbreaks and create predictive models that support management decision-making. This dissertation applies geospatial tools and data analytics for agricultural pest management at global, national, regional, and local scales to advance interdisciplinary approaches that contribute to reducing global migration of non-native species, improving biosecurity efforts at ports of entry, and detecting field-scale crop stress.

REFERENCES

Areal, F. J., Touza, J., MacLeod, A., Dehnen-Schmutz, K., Perrings, C., Palmieri, M. G., & Spence, N. J. (2008). Integrating drivers influencing the detection of plant pests carried in the international cut flower trade. *Journal of Environmental Management*, *89*(4), 300–307. https://doi.org/10.1016/j.jenvman.2007.06.017

Bacon, S. J., Aebi, A., Calanca, P., & Bacher, S. (2014). Quarantine arthropod invasions in Europe: The role of climate, hosts and propagule pressure. *Diversity and Distributions*, *20*(1), 84–94. https://doi.org/10.1111/ddi.12149

Bacon, S. J., Bacher, S., & Aebi, A. (2012). Gaps in border controls are related to quarantine alien insect invasions in Europe. *PLoS ONE 7*(10): e47689. https://doi.org/10.1371/journal.pone.0047689

Balcan, D., Gonçalves, B., Hu, H., Ramasco, J. J., Colizza, V., & Vespignani, A. (2010). Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science*, *1*(3), 132–145. https://doi.org/10.1016/j.jocs.2010.07.002

Banks, N. C., Paini, D. R., Bayliss, K. L., & Hodda, M. (2015). The role of global trade and transport network topology in the human-mediated dispersal of alien species. *Ecology Letters*, *18*(2), 188–199. https://doi.org/10.1111/ele.12397

Bebber, D. P., Field, E., Gui, H., Mortimer, P., Holmes, T., & Gurr, S. J. (2019). Many unreported crop pests and pathogens are probably already present. *Global Change Biology*, *25*(8), 2703–2713. https://doi.org/10.1111/gcb.14698

Bertelsmeier, C., Ollier, S., Liebhold, A., & Keller, L. (2017). Recent human history governs global ant invasion dynamics. *Nature Ecology & Evolution*, *1*(7), 1–8. https://doi.org/10.1038/s41559-017-0184

Bertelsmeier, C., Ollier, S., Liebhold, A. M., Brockerhoff, E. G., Ward, D., & Keller, L. (2018). Recurrent bridgehead effects accelerate global alien ant spread. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(21), 5486–5491. https://doi.org/10.1073/pnas.1801990115

Blumenfeld, A. J., Eyer, P.-A., Husseneder, C., Mo, J., Johnson, L. N. L., Wang, C., Kenneth Grace, J., Chouvenc, T., Wang, S., & Vargo, E. L. (2021). Bridgehead effect and multiple introductions shape the global invasion history of a termite. *Communications Biology*, *4*(1), 1–12. https://doi.org/10.1038/s42003-021-01725-x

Blumenfeld, A. J., & Vargo, E. L. (2020). Geography, opportunity and bridgeheads facilitate termite invasions to the United States. *Biological Invasions*, *22*(11), 3269–3282. https://doi.org/10.1007/s10530-020-02322-5

Boheemen, L. A. van, Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion

characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Molecular Ecology*, *26*(20), 5421–5434. https://doi.org/10.1111/mec.14293

Bourhis, Y., Bell, J. R., Bosch, F. van den, & Milne, A. E. (2020). Deep learning for monitoring network optimisation—A practical example using a national insect survey. *BioRxiv*, 2020.05.14.095539. https://doi.org/10.1101/2020.05.14.095539

Brockerhoff, E. G., Kimberley, M., Liebhold, A. M., Haack, R. A., & Cavey, J. F. (2014). Predicting how altering propagule pressure changes establishment rates of biological invaders across species pools. *Ecology*, *95*(3), 594–601. https://doi.org/10.1890/13-0465.1

Brooks, C. P., Antonovics, J., & Keitt, T. H. (2008). Spatial and temporal heterogeneity explain disease dynamics in a spatially explicit network model. *The American Naturalist*, *172*(2), 149–159. https://doi.org/10.1086/589451

Brown, N., Pérez-Sierra, A., Crow, P., & Parnell, S. (2020). The role of passive surveillance and citizen science in plant health. *CABI Agriculture and Bioscience*, *1*(1), 17. https://doi.org/10.1186/s43170-020-00016-5

Buchadas, A., Vaz, A. S., Honrado, J. P., Alagador, D., Bastos, R., Cabral, J. A., Santos, M., & Vicente, J. R. (2017). Dynamic models in research and management of biological invasions. *Journal of Environmental Management*, *196*, 594–606. https://doi.org/10.1016/j.jenvman.2017.03.060

Caley, P., Ingram, R., & De Barro, P. (2015). Entry of exotic insects into Australia: Does border interception count match incursion risk? *Biological Invasions*, *17*(4), 1087–1094. https://doi.org/10.1007/s10530-014-0777-z

Carrivick, J., Smith, M., & Quincey, D. (2016). *Structure from Motion in the Geosciences* (1st ed.). Wiley-Blackwell.

Carter, C. A., Chalfant, J. A., & Goodhue, R. E. (2004). Invasive species in agriculture: A rising concern. *Western Economics Forum*, *Journal Article*. http://purl.umn.edu/27979

Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K. A., Giovani, B., Saunders, D. G. O., Kamoun, S., Legg, J. P., Verdier, V., Lessel, J., Neher, R. A., Day, R., Pardey, P., Gullino, M. L., Records, A. R., Bextine, B., Leach, J. E., Staiger, S., & Tohme, J. (2019). A global surveillance system for crop diseases. *Science*, *364*(6447), 1237–1239. https://doi.org/10.1126/science.aaw1572

Caton, B. P., Dobbs, T. T., & Brodel, C. F. (2006). Arrivals of hitchhiking insect pests on international cargo aircraft at Miami International Airport. *Biological Invasions*, *8*(4), 765–785. https://doi.org/10.1007/s10530-005-3736-x

Chapman, D., Purse, B., Roy, H., & Bullock, J. (2017). Global trade networks determine the distribution of invasive non-native species. *Global Ecology and Biogeography*, *26*, 907–917.

Chen, C., Epanchin-Niell, R. S., & Haight, R. G. (2018). Optimal inspection of imports to prevent invasive pest introduction. *Risk Analysis*, *38*(3), 603–619. https://doi.org/10.1111/risa.12880

Colizza, V., Barrat, A., Barthélemy, M., & Vespignani, A. (2006). The modeling of global epidemics: Stochastic dynamics and predictability. *Bulletin of Mathematical Biology*, *68*(8), 1893–1921. https://doi.org/10.1007/s11538-006-9077-9

Colunga-Garcia, M., Haack, R., Magarey, R., & Borchert, D. (2013). Understanding trade pathways to target biosecurity surveillance. *NeoBiota*, *18*, 103–118. http://dx.doi.org/10.3897/neobiota.18.4019

Colunga-Garcia, M., Magarey, R. A., Haack, R. A., Gage, S. H., & Qi, J. (2010). Enhancing early detection of exotic pests in agricultural and forest ecosystems using an urban-gradient framework. *Ecological Applications*, *20*(2), 303–310. https://doi.org/10.1890/09-0193.1

Correa, M. C. G., Palero, F., Malausa, T., Crochard, D., Zaviezo, T., & Lombaert, E. (2019). European bridgehead effect in the worldwide invasion of the obscure mealybug. *Biological Invasions*, *21*(1), 123–136. https://doi.org/10.1007/s10530-018-1809-x

Derocles, S. A. P., Lunt, D. H., Berthe, S. C. F., Nichols, P. C., Moss, E. D., & Evans, D. M. (2018). Climate warming alters the structure of farmland tritrophic ecological networks and reduces crop yield. *Molecular Ecology*, *27*(23), 4931–4946. https://doi.org/10.1111/mec.14903

Deutsch, C. A., Tewksbury, J. J., Tigchelaar, M., Battisti, D. S., Merrill, S. C., Huey, R. B., & Naylor, R. L. (2018). Increase in crop losses to insect pests in a warming climate. *Science (New York, N.Y.)*, *361*(6405), 916–919. https://doi.org/10.1126/science.aat3466

Dwivedi, M., Shadab, M. H., & Santosh, V. R. (2020). Insect pest detection, migration and monitoring using radar and LiDAR systems. In A. K. Chakravarthy (Ed.), *Innovative pest management approaches for the 21st century: Harnessing automated unmanned technologies* (pp. 61–76). Springer. https://doi.org/10.1007/978-981-15-0794-6_4

Elad, Y., & Pertot, I. (2014). Climate change impacts on plant pathogens and plant diseases. *Journal of Crop Improvement*, *28*(1), 99–139. https://doi.org/10.1080/15427528.2014.865412

Elith, J. (2017). Predicting distributions of invasive species. In A. P. Robinson, M. A. Burgman, M. Nunn, & T. Walshe (Eds.), *Invasive species: Risk assessment and management* (pp. 93–129). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781139019606.006

Eschen, R., Rigaux, L., Sukovata, L., Vettraino, A. M., Marzano, M., & Grégoire, J.-C. (2015). Phytosanitary inspection of woody plants for planting at European Union entry points: A practical enquiry. *Biological Invasions*, *17*(8), 2403–2413. https://doi.org/10.1007/s10530-015-0883-6

FAO. (2009). *Global agriculture towards 2050* (High Level Expert Forum - How to Feed the World in 2050). FAO.

FAO, I. (2021). *The state of food security and nutrition in the world 2021: Transforming food systems for food security, improved nutrition and affordable healthy diets for all*. FAO, IFAD, UNICEF, WFP and WHO. https://doi.org/10.4060/cb4474en

Faulkner, K. T., Hulme, P. E., Pagad, S., Wilson, J. R. U., & Robertson, M. P. (2020). Classifying the introduction pathways of alien species: Are we moving in the right direction? *NeoBiota*, *62*, 143–159. https://doi.org/10.3897/neobiota.62.53543

Faulkner, K. T., Robertson, M. P., Rouget, M., & Wilson, J. R. U. (2017). Prioritising surveillance for alien organisms transported as stowaways on ships travelling to South Africa. *PLoS ONE*, *12*(4). https://doi.org/10.1371/journal.pone.0173340

Ferrari, J. R., Preisser, E. L., & Fitzpatrick, M. C. (2014). Modeling the spread of invasive species using dynamic network models. *Biological Invasions*, *16*(4), 949–960. https://doi.org/10.1007/s10530-013-0552-6

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

Floerl, O., Inglis, G. J., Dey, K., & Smith, A. (2009). The importance of transport hubs in stepping-stone invasions. *Journal of Applied Ecology*, *46*(1), 37–45. https://doi.org/10.1111/j.1365-2664.2008.01540.x

Fontaine, M. C., Labbé, F., Dussert, Y., Delière, L., Richart-Cervera, S., Giraud, T., & Delmotte, F. (2021). Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*. Current Biology, 31(10), 2155-2166.e4. https://doi.org/10.1016/j.cub.2021.03.009

Frost, C. M., Allen, W. J., Courchamp, F., Jeschke, J. M., Saul, W.-C., & Wardle, D. A. (2019). Using network theory to understand and predict biological invasions. *Trends in Ecology & Evolution*, *34*(9), 831–843. https://doi.org/10.1016/j.tree.2019.04.012

Garnas, J. R., Auger-Rozenberg, M.-A., Roques, A., Bertelsmeier, C., Wingfield, M. J., Saccaggi, D. L., Roy, H. E., & Slippers, B. (2016). Complex patterns of global spread in invasive insects: Eco-evolutionary and management consequences. *Biological Invasions*, *18*(4), 935–952. https://doi.org/10.1007/s10530-016-1082-9

Gottwald, T., Luo, W., Posny, D., Riley, T., & Louws, F. (2019). A probabilistic census-travel model to predict introduction sites of exotic plant, animal and human pathogens. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1776), 20180260. https://doi.org/10.1098/rstb.2018.0260

Harrison, R. M., Chadwick, D., Hester, R. E., Kibblewhite, M. G., Sakrabani, R., Deeks, L., Cloy, J. M., Rees, B., Smith, P., & Smith, K. (2012). *Environmental impacts of modern agriculture*. Royal Society of Chemistry. http://ebookcentral.proquest.com/lib/ncsu/detail.action?docID=1185495

Hu, Y., Wilson, S., Schwessinger, B., & Rathjen, J. P. (2020). Blurred lines: Integrating emerging technologies to advance plant biosecurity. *Current Opinion in Plant Biology*, *56*, 127–134. https://doi.org/10.1016/j.pbi.2020.04.011

Hulme, P. E. (2009). Trade, transport and trouble: Managing invasive species pathways in an era of globalization. *The Journal of Applied Ecology*, *46*(1), 10–18. https://doi.org/10.1111/j.1365-2664.2008.01600.x

Hunt, E. R., & Rondon, S. I. (2017). Detection of potato beetle damage using remote sensing from small unmanned aircraft systems. *Journal of Applied Remote Sensing*, *11*(2), 026013. https://doi.org/10.1117/1.JRS.11.026013

International Plant Protection Convention. (2016). *Methodologies for sampling of consignments* (International Standards for Phytosanitary Measures No. 31). https://www.ippc.int/en/publications/588/

Jarić, I., Bellard, C., Correia, R. A., Courchamp, F., Douda, K., Essl, F., Jeschke, J. M., Kalinkat, G., Kalous, L., Lennox, R. J., Novoa, A., Proulx, R., Pyšek, P., Soriano–Redondo, A., Souza, A. T., Vardi, R., Veríssimo, D., & Roll, U. (2021). Invasion culturomics and iEcology. *Conservation Biology*, *35*(2), 447–451. https://doi.org/10.1111/cobi.13707

Jenkins, D. A., Mizell, R. F., Van Bloem, S., Whitmire, S., Wiscovitch, L., Zaleski, C., & Goenaga, R. (2014). An analysis of arthropod interceptions by APHIS-PPQ and Customs and Border Protection in Puerto Rico. *American Entomologist*, *60*(1), 44–57. https://doi.org/10.1093/ae/60.1.44

Jones, C., Jones, S., Petrasova, A., Petras, V., Gaydos, D., Skrip, M., Takeuchi, Y., Bigsby, K., & Meentemeyer, R. (2021). Iteratively forecasting invasions with PoPS and a little help from our friends. *Frontiers In Ecology And The Environment*. https://doi.org/doi:10.1002/fee.2357

Jones, O., Robinson, A. P., Shield, M., & Sibley, J. (2017). The allocation of inspection resources. In A. P. Robinson, M. A. Burgman, M. Nunn, & T. Walshe (Eds.), *Invasive species: Risk assessment and management* (pp. 1–16). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781139019606.002

Keller, R. P., Drake, J. M., Drew, M. B., & Lodge, D. M. (2011). Linking environmental conditions and ship movements to estimate invasive species transport across the global shipping network. *Diversity and Distributions*, *17*(1), 93–102. https://doi.org/10.1111/j.1472-4642.2010.00696.x

Kim, B., Hong, S. C., Egger, D., Katsar, C. S., & Griffin, R. L. (2019). Predictive modeling and categorizing likelihoods of quarantine pest introduction of imported propagative

commodities from different countries. *Risk Analysis*, *39*(6), 1382–1396. https://doi.org/10.1111/risa.13252

Koch, F. H., Yemshanov, D., Colunga-Garcia, M., Magarey, R. D., & Smith, W. D. (2011). Potential establishment of alien-invasive forest insect species in the United States: Where and how many? *Biological Invasions*, *13*, 969–985.

Koch, F. H., Yemshanov, D., Haack, R. A., & Magarey, R. D. (2014). Using a network model to assess risk of forest pest spread via recreational travel. *PLOS ONE*, *9*(7), 10.

Kriticos, D., Maywald, G., Yonow, T., Zurcher, E., Herrmann, N., & Sutherst, R. (2015). *CLIMEX. Version 4. Exploring the effects of climate on plants, animals and diseases*.

Levine, J. M., & D'Antonio, C. M. (2003). Forecasting biological invasions with increasing international trade. *Conservation Biology*, *17*(1), 322–326. https://doi.org/10.1046/j.1523-1739.2003.02038.x

Lichtenberg, E., & Olson, L. J. (2018). The fruit and vegetable import pathway for potential invasive pest arrivals. *PLoS One*, *13*(2), e0192280. https://doi.org/10.1371/journal.pone.0192280

Liebhold, A. M., Work, T. T., McCullough, D. G., & Cavey, J. F. (2006). Airline baggage as a pathway for alien insect species invading the United States. *American Entomologist*, *52*(1), 48–54. https://doi.org/10.1093/ae/52.1.48

Liebhold, A. M., Yamanaka, T., Roques, A., Augustin, S., Chown, S. L., Brockerhoff, E. G., & Pyšek, P. (2016). Global compositional variation among native and non-native regional insect assemblages emphasizes the importance of pathways. *Biological Invasions*, *18*(4), 893–905. https://doi.org/10.1007/s10530-016-1079-4

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, *20*(5), 223–228. https://doi.org/10.1016/j.tree.2005.02.004

Lockwood, J. L., Cassey, P., & Blackburn, T. M. (2009). The more you introduce the more you get: The role of colonization pressure and propagule pressure in invasion ecology. *Diversity and Distributions*, *15*(5), 904–910. https://doi.org/10.1111/j.1472-4642.2009.00594.x

Lombaert, E., Guillemaud, T., Cornuet, J.-M., Malausa, T., Facon, B., & Estoup, A. (2010). Bridgehead effect in the worldwide invasion of the biocontrol Harlequin Ladybird. *PLoS One; San Francisco*, *5*(3), e9743. https://doi.org/10.1371/journal.pone.0009743

Madec, S., Baret, F., de Solan, B., Thomas, S., Dutartre, D., Jezequel, S., Hemmerlé, M., Colombeau, G., & Comar, A. (2017). High-throughput phenotyping of plant height: Comparing unmanned aerial vehicles and ground LiDAR estimates. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.02002

Magarey, R. D., Colunga-Garcia, M., & Fieselmann, D. A. (2009). Plant biosecurity in the United States: Roles, responsibilities, and information needs. *BioScience*, *59*(10), 875–884. https://doi.org/10.1525/bio.2009.59.10.9

Mahlein, A.-K. (2016). Plant disease detection by imaging sensors – Parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*, *100*(2), 241–251. https://doi.org/10.1094/PDIS-03-15-0340-FE

Mahlein, A.-K., Kuska, M. T., Behmann, J., Polder, G., & Walter, A. (2018). Hyperspectral sensors and imaging technologies in phytopathology: State of the art. *Annual Review of Phytopathology*, *56*(1), 535–558. https://doi.org/10.1146/annurev-phyto-080417-050100

McCullough, D. G., Work, T. T., Cavey, J. F., Liebhold, A. M., & Marshall, D. (2006). Interceptions of nonindigenous plant pests at US ports of entry and border crossings over a 17-year period. *Biological Invasions*, *8*(4), 611. https://doi.org/10.1007/s10530-005-1798-4

Meentemeyer, R. K., Cunniffe, N. J., Cook, A. R., Filipe, J. A. N., Hunter, R. D., Rizzo, D. M., & Gilligan, C. A. (2011). Epidemiological modeling of invasion in heterogeneous landscapes: Spread of sudden oak death in California (1990-2030). *Ecosphere*, *2*(2), art17-24. https://doi.org/10.1890/ES10-00192.1

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

Milne, A. E., Gottwald, T., Parnell, S. R., Chavez, V. A., & Bosch, F. van den. (2020). What makes or breaks a campaign to stop an invading plant pathogen? *PLOS Computational Biology*, *16*(2), e1007570. https://doi.org/10.1371/journal.pcbi.1007570

NC State University. (2021). *SAFARIS CAPS*. SAFARIS. https://safaris.cipm.info/safarispestmodel/

Paini, D. R., Sheppard, A. W., Cook, D. C., De Barro, P. J., Worner, S. P., & Thomas, M. B. (2016). Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7575–7579. https://doi.org/10.1073/pnas.1602205113

Paini, D. R., & Yemshanov, D. (2012). Modelling the arrival of invasive organisms via the international marine shipping network: A khapra beetle study. *PloS One*, *7*(9), e44589. https://doi.org/10.1371/journal.pone.0044589

Parnell, S., van den Bosch, F., Gottwald, T., & Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: An epidemiological perspective. *Annual Review of Phytopathology*, *55*(1), 591–610. https://doi.org/10.1146/annurev-phyto-080516-035334

Paulus, S. (2019). Measuring crops in 3D: Using geometry for plant phenotyping. *Plant Methods*, *15*. https://doi.org/10.1186/s13007-019-0490-0

Petrasova, A., Gaydos, D. A., Petras, V., Jones, C. M., Mitasova, H., & Meentemeyer, R. K. (2020). Geospatial simulation steering for adaptive management. *Environmental Modelling & Software*, *133*, 104801. https://doi.org/10.1016/j.envsoft.2020.104801

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Powell, M. R. (2015). Risk-based sampling: I don't want to weight in vain. *Risk Analysis*, *35*(12), 2172–2182. https://doi.org/10.1111/risa.12415

Pretty, J., & Bharucha, Z. P. (2018). *Sustainable intensification of agriculture: Greening the world's food economy*. Routledge. https://doi.org/10.4324/9781138638044

Purdue University. (2021). *CAPS program resource and collaboration site*. CAPS. http://caps.ceris.purdue.edu/

Reaser, J. K., Meyerson, L. A., & Von Holle, B. (2008). Saving camels from straws: How propagule pressure-based prevention policies can reduce the risk of biological invasion. *Biological Invasions*, *10*(7), 1085–1098. https://doi.org/10.1007/s10530-007-9186-x

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

Robinson, A. P., Burgman, M. A., & Cannon, R. (2011). Allocating surveillance resources to reduce ecological invasions: Maximizing detections and information about the threat. *Ecological Applications*, *21*(4), 1410–1417. https://doi.org/10.1890/10-0195.1

Robinson, A. P., Chisholm, M., Mudford, R., & Maillardet, R. (2015). Ad hoc solutions to estimating pathway non-compliance rates using imperfect and incomplete information. In F. Jarrad, S. Low-Choy, & K. Mengersen (Eds.), *Biosecurity surveillance: Quantitative approaches* (pp. 167–180). CABI. https://doi.org/10.1079/9781780643595.0167

Rossiter, A., & Hester, S. M. (2017). Designing biosecurity inspection regimes to account for stakeholder incentives: An inspection game approach. *Economic Record*, *93*(301), 277–301. https://doi.org/10.1111/1475-4932.12315

Saccaggi, D. L., Karsten, M., Robertson, M. P., Kumschick, S., Somers, M. J., Wilson, J. R. U., & Terblanche, J. S. (2016). Methods and approaches for the management of arthropod border incursions. *Biological Invasions*, *18*(4), 1057–1075. https://doi.org/10.1007/s10530-016-1085-6

Sardain, A., Sardain, E., & Leung, B. (2019). Global forecasts of shipping traffic and biological invasions to 2050. *Nature Sustainability*, *2*(4), 274–282. https://doi.org/10.1038/s41893-019-0245-y

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, *3*(3), 430–439. https://doi.org/10.1038/s41559-018-0793-y

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., Blasius, B., Brundu, G., Capinha, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Fuentes, N., Jäger, H., … Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*(1), 14435. https://doi.org/10.1038/ncomms14435

Seebens, H., Briski, E., Ghabooli, S., Shiganova, T., MacIsaac, H. J., & Blasius, B. (2019). Non-native species spread in a complex network: The interaction of global transport and local population dynamics determines invasion success. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1901), 20190036. https://doi.org/10.1098/rspb.2019.0036

Seebens, H., Gastner, M. T., Blasius, B., & Franck Courchamp. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters*, *16*(6), 782–790. https://doi.org/10.1111/ele.12111

Seebens, H., Schwartz, N., Schupp, P. J., & Blasius, B. (2016). Predicting the spread of marine species introduced by global shipping. *From the Cover*, *113*(20), 5646–5651. https://doi.org/10.1073/pnas.1524427113

Sikes, B. A., Bufford, J. L., Hulme, P. E., Cooper, J. A., Johnston, P. R., & Duncan, R. P. (2018). Import volumes and biosecurity interventions shape the arrival rate of fungal pathogens. *PLOS Biology*, *16*(5), e2006025. https://doi.org/10.1371/journal.pbio.2006025

Springborn, M. R., Lindsay, A. R., & Epanchin-Niell, R. S. (2016). Harnessing enforcement leverage at the border to minimize biological risk from international live species trade. *Journal of Economic Behavior & Organization*, *132*, 98–112. https://doi.org/10.1016/j.jebo.2016.03.011

Surkov, I. V., Oude Lansink, A. G. J. M., van Kooten, O., & van der Werf, W. (2008). A model of optimal import phytosanitary inspection under capacity constraint. *Agricultural Economics*, *38*(3), 363–373. https://doi.org/10.1111/j.1574-0862.2008.00306.x

Surkov, I. V., van der Werf, W., van Kooten, O., & Oude Lansink, A. G. J. M. (2008). Modeling the rejection probability in plant imports. *Phytopathology*, *98*(6), 728–735. https://doi.org/10.1094/PHYTO-98-6-0728

Szyniszewska, A. M., Leppla, N. C., Huang, Z., & Tatem, A. J. (2016). Analysis of seasonal risk for importation of the Mediterranean Fruit Fly, *Ceratitis capitata* (Diptera: Tephritidae), via air passenger traffic arriving in Florida and California. *Journal of Economic Entomology*, *109*(6), 2317–2328. https://doi.org/10.1093/jee/tow196

Tatem, A. J. (2009). The worldwide airline network and the dispersal of exotic species: 2007–2010. *Ecography*, *32*(1), 94–102. https://doi.org/10.1111/j.1600-0587.2008.05588.x

Tatem, A. J., & Hay, S. I. (2007). Climatic similarity and biological exchange in the worldwide airline transportation network. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1617), 1489–1496. https://doi.org/10.1098/rspb.2007.0148

Trouvé, R., & Robinson, A. P. (2020). Estimating consignment-level infestation rates from the proportion of consignment that failed border inspection: Possibilities and limitations in the presence of overdispersed data. *Risk Analysis*, *n/a*(n/a). https://doi.org/10.1111/risa.13592

Turbelin, A. J., Malamud, B. D., Francis, R. A., & Sykes, M. (2017). Mapping the global state of invasive alien species: Patterns of invasion and policy responses. *Global Ecology and Biogeography*, *26*(1), 78–92. https://doi.org/10.1111/geb.12517

Turner, R. M., Plank, M. J., Brockerhoff, E. G., Pawson, S., Liebhold, A., & James, A. (2020). Considering unseen arrivals in predictions of establishment risk based on border biosecurity interceptions. *Ecological Applications*, *30*(8), e02194. https://doi.org/10.1002/eap.2194

Václavík, T., & Meentemeyer, R. K. (2012). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, *18*(1), 73–83. https://doi.org/10.1111/j.1472-4642.2011.00854.x

Venette, R. C., Kriticos, D. J., Magarey, R. D., Koch, F. H., Baker, R. H. A., Worner, S. P., Raboteaux, N. N. G., McKenney, D. W., Dobesberger, E. J., Yemshanov, D., Barro, P. J. D., Hutchison, W. D., Fowler, G., Kalaris, T. M., & Pedlar, J. (2010). Pest risk maps for invasive alien species: A roadmap for improvement. *BioScience*, *60*(5), 349–362. https://doi.org/10.1525/bio.2010.60.5.5

Work, T. T., McCullough, D. G., Cavey, J. F., & Komsa, R. (2005). Arrival rate of nonindigenous insect species into the United States through foreign trade. *Biological Invasions*, *7*(2), 323. https://doi.org/10.1007/s10530-004-1663-x

Yemshanov, D., Koch, F. H., Coulston, J. W., & Smith, W. D. (2017). Mapping risks and impacts of invasive alien species with dynamic simulation models. In A. P. Robinson, M. A. Burgman, M. Nunn, & T. Walshe (Eds.), *Invasive species: Risk assessment and management* (pp. 130–151). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781139019606.007

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

CHAPTER 2: FORECASTING GLOBAL SPREAD OF INVASIVE PESTS AND PATHOGENS THROUGH INTERNATIONAL TRADE

Under peer review at Global Change Biology

Authors

Kellyn Montgomery[1], Chelsey Walden-Schreiner[1], Chris Jones[1], Ben Seliger[1], Thom Worm[1], Ariel Saffer[1], Laura Tateosian[1], Makiko Shukunobe[1], Sunil Kumar[2], Ross K. Meentemeyer[1]

Affiliations

[1]Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, USA

[2]United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Plant Protection and Quarantine (PPQ), Raleigh, North Carolina, USA

Author Contributions

Kellyn M. wrote the manuscript; significantly contributed to the model and framework design and calibration; and wrote code for the data pipeline, equation development, testing, and visualizations. Chelsey W.-S. contributed to the manuscript writing and performed code implementation and model calibration. Chris J. wrote initial code for model implementation. Computational and interface components were performed by Thom W., Ariel S., and Ben S. The host map was derived by Ben S. and web scraping and text mining was performed by Laura T. and Makiko S. Sunil K. and Ross M. provided extensive feedback and guidance on model development and case study. All authors contributed to the framework conceptual design and provided editorial review of the manuscript.

Data Availability

Software and tools developed through this research are open-source and freely available under GNU General Public License. The PoPS Global code, along with version history, is compiled and available via Zenodo (Walden-Schreiner et al., 2021). An early version of the PoPS Global visualization dashboard to view the case study results can be accessed via https://pops-global.ngrok.io/. The national phytosanitary capacity data were obtained from Dr.

Regan Early (Early et al., 2016). All other data used are open and available at the references cited within the manuscript text.

Abstract

Non-native plant pests and pathogens threaten biodiversity, ecosystem function, food security, and economic livelihoods. As new invasive populations establish, often as an unintended consequence of international trade, they can become additional sources of introductions, accelerating global spread through bridgehead effects. While the study of non-native pest spread has leveraged computational models to provide critical insights into drivers and dynamics of biological invasions and inform management, efforts have focused on local or regional scales and are challenged by complex transmission networks arising from bridgehead population establishment. This paper presents a species-agnostic, spatio-temporal stochastic network model called PoPS (Pest or Pathogen Spread) Global that couples international trade networks with core drivers of biological invasions—climate suitability, host availability, and propagule pressure—quantified through open, globally available databases to forecast the spread of non-native plant pests. The modular design of the framework supports early, proactive responses for emerging pests even when limited data are available and enables forecasts at different spatial and temporal resolutions. We demonstrate the framework using a case study of the invasive planthopper spotted lanternfly (*Lycorma delicatula*). The model correctly hindcasted historical, known spotted lanternfly introductions and identified potential bridgehead populations contributing to global spread. The case study also demonstrated how the framework can be used to test management intervention scenarios, such as limiting imports along high-risk pathways. This global view of phytosanitary pandemics provides crucial information for anticipating biological invasions, quantifying transport pathways risk levels, and allocating resources to safeguard plant health, agriculture, and natural resources.

Keywords: biological invasions, bridgehead effects, international trade, iterative forecasting, network model, plant pests and pathogens

1 Introduction

Worldwide, plant pests and pathogens (hereafter "pests" for brevity) are responsible for destroying up to 40% of crop yields and disrupting forest ecosystems through mortality of keystone tree species (Fones et al., 2020; Paini et al., 2016; Pyšek & Richardson, 2010; Savary et al., 2019). Invasions by non-native pests have increased in recent decades and show no signs of slowing (Seebens et al., 2017), with corresponding damage and management costs roughly doubling every six years since 1970 (Diagne et al., 2021). This acceleration is largely driven by globalization and international trade, which facilitates long-distance movement beyond natural geographic barriers (Brasier, 2008; Epanchin-Niell et al., 2021; Chapman et al., 2017). The establishment of successful invasive populations, termed bridgehead populations, then become additional pest sources and accelerate global spread (Bertelsmeier et al., 2018; Bertelsmeier & Keller, 2018; Lombaert et al., 2010). Furthermore, propagule transport from these bridgehead populations results in complex patterns of repeated introductions and population admixture that can increase invasive population fitness and establishment success (Garnas et al., 2016; Rius & Darling, 2014). Accounting for global transport and bridgehead effects in invasive plant pest risk assessments can help inform prevention and management efforts taking place at the range of spatial scales known as the biosecurity continuum (Magarey et al., 2009) and improve global spread forecasts.

Research in biological invasions has increasingly turned to mathematical models and computational simulations to understand drivers and dynamics of invasive species spread (Douma et al., 2016; Lewis et al., 2016; Parnell et al., 2017). Population models have been used to predict species survival and growth following an introduction, incorporating complicated interspecific interactions to model biological control outcomes (Lewis et al., 2016; S. Petrovskii et al., 2005; S. V. Petrovskii et al., 2005). Species distribution models, meanwhile, provide insights for invasive pest management by identifying areas where a pest could establish based on statistical relationships between current distributions and relevant environmental factors (Elith & Leathwick, 2009; Václavík & Meentemeyer, 2009; Wakie et al., 2020). Spatially explicit epidemiological simulations that capture local and long-distance dispersal mechanisms are especially useful for predicting patterns and rates of invasive plant pest spread (DeAngelis & Yurek, 2017; Jones et al., 2021; Meentemeyer et al., 2011). Network models are well suited for modeling complex human-mediated dispersal and can be coupled with grid-based diffusion

models (Banks et al., 2015; Harwood et al., 2009; Strickland et al., 2015). Although computational models have been extensively used to study biological invasions and provide valuable insights, most models focus on local or regional scales of establishment and spread and do not attempt to model the multiscale processes of bridgehead effects. The immensely challenging task of invasive plant pest management, however, must be supported by forecasts that take into account local, regional, and global dynamics to understand complex patterns of movement and invasion success (Chapman et al., 2017; Garnas et al., 2016; Meentemeyer et al., 2012).

Multiscale network modeling frameworks that couple global models of organism movement with local or regional models of population dynamics have been used to predict rates and patterns of pandemic spread among humans and, more recently, invasive species. The GLEAMvis framework for human infectious disease modeling (Balcan et al., 2010) uses global airline and regional commuter data to simulate movement of infected individuals between geographically dispersed metapopulations and compartmental models to simulate spread dynamics within each metapopulation. The flexible Spatiotemporal Epidemiological Modeller (STEM; Ford et al., 2006) also couples network models with local compartmental or agent-based disease models. For invasive marine species, Seebens et al. (2013; 2016) developed a probabilistic network model that couples ballast water transport through international shipping networks with port biogeographic conditions to identify invasion routes and predict establishment. These multiscale modeling approaches represent a promising paradigm for global biological invasion forecasting; however, their applicability to phytosanitary risk analysis is limited by the unique challenges posed by plant pests. Plant pests often exhibit long detection lag times (e.g., years vs. days), complex host preferences, temporal and spatial irregularity of host availability, and diverse dispersal mechanisms (Crooks, 2005; Cunniffe et al., 2015). Furthermore, traits of emerging plant pests are often not well understood, and occurrence data for model calibration and validation are limited. Phytosanitary agencies tasked with preventing and managing emerging invasive plant pests need tools that can quickly be deployed for a broad range of pest-host systems when data are limited and simulate global transport and bridgehead population establishment.

This paper introduces PoPS (Pest or Pathogen Spread) Global—a species-agnostic, stochastic simulation and forecasting framework adapted for plant pests and pathogens from the

approach developed by Seebens et al. (2013; 2016). PoPS Global is part of a suite of tools developed to support phytosanitary management decisions along the biosecurity continuum, including PoPS, a generalized spatially explicit, discrete-time model used for simulating landscape spread of plant pests and pathogens (Jones et al., 2021) and PoPS Border, a consignment inspection simulator used for designing dynamic, risk-based sampling protocols at ports of entry (Montgomery et al., 2021). PoPS Global is broadly applicable for forecasting plant pest invasions accelerated by international trade and bridgehead populations. It can be adapted for a wide range of pest-host systems, targeting pest species associated with specific commodities. The modular design enables quick implementation to support management of poorly understood emerging pests using general, open data, while also providing options to integrate more specialized information when available. PoPS Global is multiscale, addressing calls to incorporate temporal dynamics of local spread and global transmission concurrently; multiscale dynamics (e.g., latency between invasion and population build up) are critical to understanding the timing of pandemics (Balcan et al., 2009; Chapman et al., 2017). The framework also allows for scenario testing by adjusting inputs to explore the impact of management (e.g., phytosanitary efforts, trade policy) or pest ecology (e.g., polyphagy, host availability). This paper describes the modeling framework, highlighting the modular design, core drivers, and data requirements, and presents a case study of an emerging plant pest, spotted lanternfly (*Lycorma delicatula* (White) (Hemiptera: Fulgoridae)), to demonstrate how the framework can be used to forecast global species movement and simulate two potential management scenarios. We conclude by highlighting opportunities for expansion of this open-source framework and priorities for data and future research.

2 Methods

The PoPS Global forecasting framework integrates an automated open data pipeline that supports model calibration and validation, scenario testing and analytics to inform management decisions, and iterative updating as new data become available (Figure 1). The open data pipeline acquires and formats pest observations, core driver data, and optional modules for input into the model. A calibration loop computes model parameters, validates results, and enables sensitivity analyses (Figure 1.2). Users can visualize forecast and scenario results (Figure 1.1) to inform biosecurity management decisions, including the deployment of novel surveillance approaches

like web scraping and text mining to automatically monitor online information sources. The model is then confronted with new pest occurrence data as they become available to update and improve the forecast (Figure 1.3). This framework provides a practical approach for leveraging currently existing data to quickly provide quantitative forecasts of plant pest invasions and to iteratively improve the forecast over time.

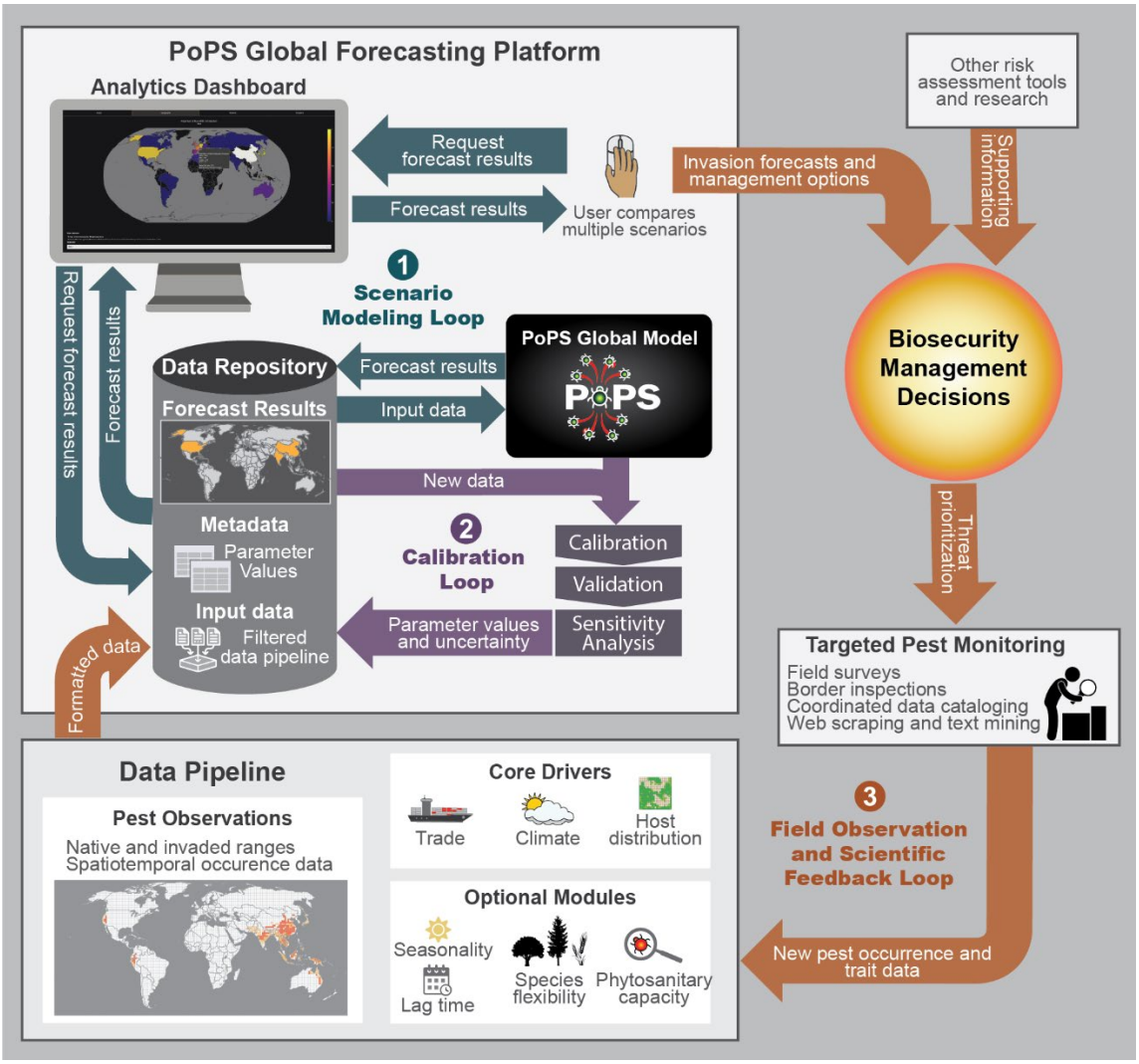


Figure 1. The PoPS Global forecasting framework includes an automated data pipeline using open data sources that acquires, aggregates, and formats model input data to support three iterative components: 1) a scenario modeling loop, 2) a calibration loop, and 3) a field observation and scientific feedback loop. Results from these loops are visualized on an analytics dashboard to support user interaction and decision making, incorporating new data as they become available to iteratively update and improve forecast accuracy.

2.1 Modeling Framework

PoPS Global is a spatio-temporal stochastic network modeling approach wherein network nodes represent geographical areas (e.g., countries, regions, ports) and bidirectional network edges (i.e., connections between nodes) represent the movement of goods via trade pathways. Potential plant pest import and export is modeled along these pathways by integrating global trade data, pest occurrence, host species distribution, and climate conditions. The model incorporates changing environmental and host conditions (e.g., annual agricultural crops or perennial forest species) and the impact of biosecurity and control measures on spread (Figure 2a–b), which can be updated as new data become available (Dietze et al., 2018; Jones et al., 2021; Parnell et al., 2017). The model predicts the probability of introduction (i.e., successful entry and establishment) for every node in the network. Nodes with successful introductions then become bridgehead populations with the potential for transmitting the pest in the subsequent time step, or after an optional latency period, described in more detail in Subsection 2.2 (Figure 2c–d).

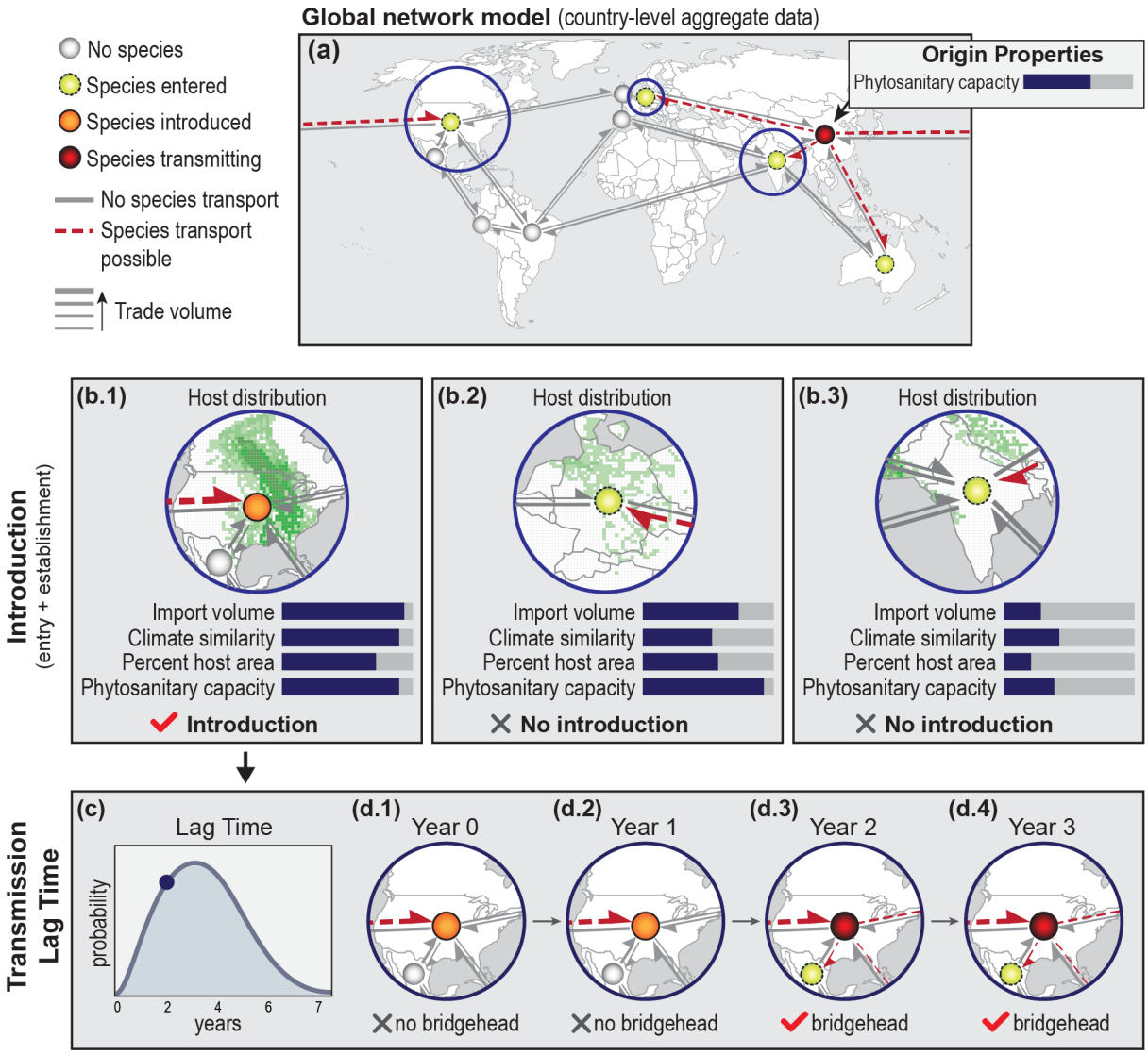
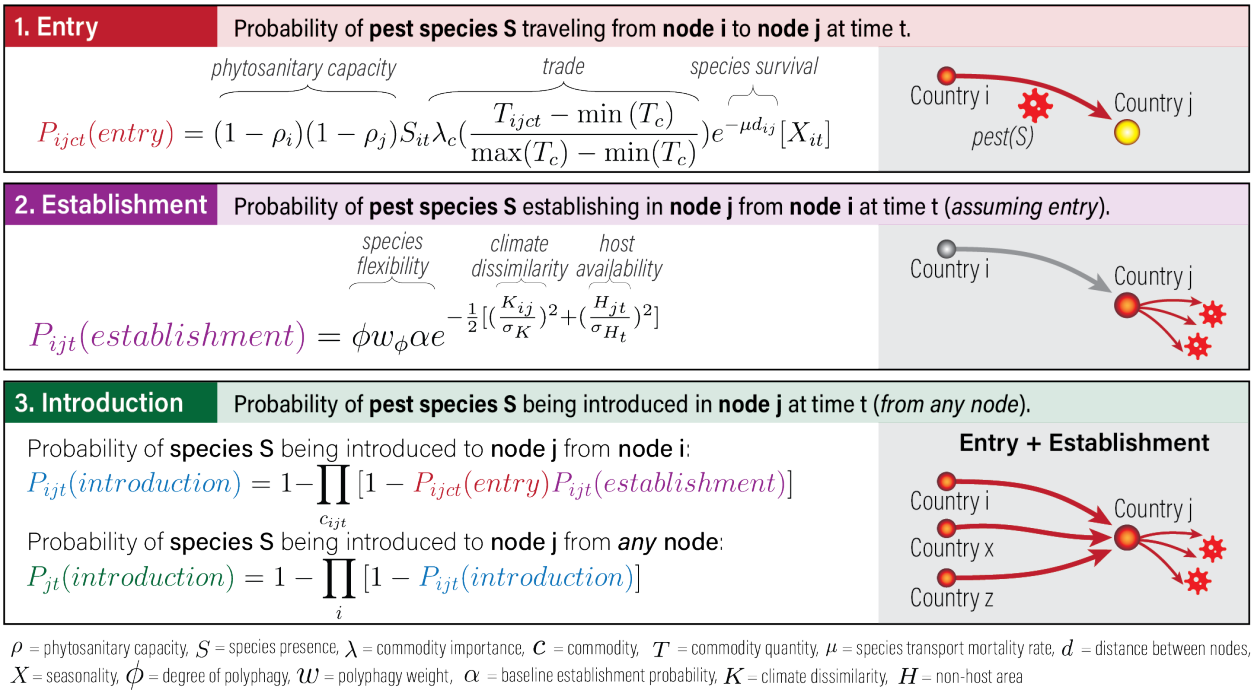


Figure 2. The PoPS Global spatio-temporal network model (a) simulates spread from node to node via edges weighted by trade volume. (b) The environmental conditions, phytosanitary capacity, and import volume (proxy for propagule pressure) in the destination nodes determine the probability of pest introduction. Following a predicted introduction, the destination can become a bridgehead population in subsequent time steps (c–d), either immediately or after a lag drawn from a distribution (c).

The model consists of three equations (Figure 3) calculating separate but related probabilities: 1) entry, 2) establishment, and 3) introduction. These terms align with definitions used by the United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHIS) and correspond, respectively, to transport, introduction, and establishment as

defined by Blackburn et al. (2011). Probability of entry (Figure 3.1) captures processes controlling movement between globally distributed nodes. It is a function of the amount of traded goods capable of transporting the pest, the likelihood of a pest surviving the journey, and, optionally, the phytosanitary capacity of importing and exporting countries. Probability of establishment (Figure 3.2) captures conditions and ecological processes within a node area. Establishment probability increases with environmental suitability, which is modeled as a Gaussian function of the climate dissimilarity between the two trading nodes, and percent area without host species in the destination node. Optionally, the probability can be adjusted by the pest's ability to survive on multiple hosts (e.g., number of host taxonomic families weighted by phylogenetic diversity of hosts). Probability of introduction (Figure 3.3) is a function of the probability of entry (inter-node processes) and the probability of establishment (intra-node processes), and is used in a binomial distribution to determine if a successful introduction occurs.



ρ = phytosanitary capacity, S = species presence, λ = commodity importance, c = commodity, T = commodity quantity, μ = species transport mortality rate, d = distance between nodes, X = seasonality, ϕ = degree of polyphagy, w = polyphagy weight, α = baseline establishment probability, K = climate dissimilarity, H = non-host area

Figure 3. The forecasting framework uses a graph structure with nodes (geographical areas) and edges (transport of pests) driven by three equations for calculating event probabilities within the network: 1) probability of entry represents inter-node processes; 2) probability of establishment represents intra-node processes; 3) probability of introduction is a combined outcome of inter- and intra-node processes.

2.2 Drivers and Data

PoPS Global applies broad theoretical concepts of biological invasion drivers at a global scale to provide an estimate of the likelihood and timing of pest arrival (Bellard et al., 2016; Hulme, 2009; Lenzner et al., 2019). Three core drivers are required to run the network model: quantity (i.e., value or weight) of goods traded between nodes over time, climate dissimilarity between nodes, and the area within nodes without hosts. The spatial scale represented by the nodes determines the degree to which these drivers are aggregated and should be chosen for each use case based on data availability and the desired network resolution. The amount of traded goods is obtained from the United Nations Comtrade database, which provides historical annual or monthly records of traded goods using the Harmonized Commodity Description and Coding System (HS) for countries since the 1960s (DESA/UNSD, 2020). The U.S. Bureau of Labor Statistics Consumer Price Index (CPI) is used to adjust trade values for inflation. Trade values can be limited to include only goods known to transport the species being modeled to provide a better proxy for propagule pressure (Bacon et al., 2012; Chapman et al., 2017).

Climate dissimilarity and non-host area are computed only for areas that are likely to be commodity destinations and, therefore, where the pest will have the opportunity to establish. We use the Global Human Influence Index to identify areas within each node with direct human influence (e.g., cities, working lands) and accessibility (e.g., highways), retaining areas above a threshold index value of 16 (Sanderson et al., 2002; WCS & CIESIN-CU, 2005). This threshold was approximated by identifying the lowest index values assigned to developed areas and major transportation corridors to ensure those areas were retained in the analysis. To determine climate dissimilarity between source and destination nodes, we use the Köppen-Geiger Climate Classification (Beck et al., 2018) and compute the percent area of the destination node that contains the same climate class(es) as the origin node, then subtract from 100%. Non-host area is computed as the percent node area that does not contain, or is not suitable for, at least one host. For agricultural hosts, harvested area per country is obtained from the Food and Agriculture Organization Statistics (FAOSTAT) database (FAO, 2020). Maps of other hosts can be created with species distribution models, such as Maximum Entropy (MAXENT; Phillips et al., 2006), using input data from the Global Biodiversity Information Facility (GBIF) species distribution database (GBIF, 2021) or the Centre for Agriculture and Biosciences International (CABI) Crop Protection Compendium (CABI, 2021).

Beyond the core required inputs, several additional optional components can be included: phytosanitary capacity of each node, degree of pest polyphagy, seasonality in the likelihood of a pest being transported, and transmission lag. Phytosanitary capacity represents measures taken by origin and destination countries to prevent or mitigate cargo contamination. We use an index developed by Early et al. (2016) based on a review of Convention on Biological Diversity national reports to determine a country's capacity to proactively address invasive species threats. The degree of pest polyphagy, specified as a count of host families, can be used to increase the likelihood of establishment for more generalist species. The seasonality component is used to constrain transmission of the pest to certain months depending on species biology. If seasonality is not incorporated, the pest can be transported between nodes during any month. The transmission lag specifies the time between when a pest is introduced to a node and when that node can become a source to other nodes in the network, representing local population growth and spread. The transmission lag can be set to a static interval (e.g., 3 years) or can be dynamic and stochastic (e.g., drawn from a gamma distribution informed by literature or available data). With a dynamic transmission lag and repeated introductions, a new lag time is drawn for each reintroduction and the one resulting in the earliest date that the node could become a source is chosen, simulating the effect of increasing propagule pressure. If lag time is not used in the model, a node can become a source of the pest in the time step immediately following introduction. See Supporting Information S1.1 for additional description of model and data assumptions.

2.3 Decision Analytics

The PoPS Global modeling framework was collaboratively developed with USDA APHIS practitioners tasked with rapid response to threats from emerging plant pests. Visualization of global spread over time allows users to analyze predicted locations and timing of introductions, with associated levels of uncertainty, and compare outcomes of different intervention strategies. We are working with analysts and managers to create a web-based decision analytics dashboard that will use the open data pipeline to run the model and test scenarios and provide multiple, interactive views of the model output. The first dashboard iteration (available at https://pops-global.ngrok.io/) allows users to view the model results as a world map with country-level summary statistics of introduction probability, timing, and sources

(Figure 4a). Model results can also be viewed as a network graph and filtered to nodes of interest to examine spread pathways from native and bridgehead pest populations (Figure 4b). Statistics summarizing transmission network characteristics, e.g., the most susceptible and transmissive nodes, are also provided. Graphs of forecasted introductions over time show predicted rates of pandemic spread. Users can toggle between model runs to compare outcomes of management scenarios. Collectively, these visualizations and statistics provide insight into the spread network to support decision making.

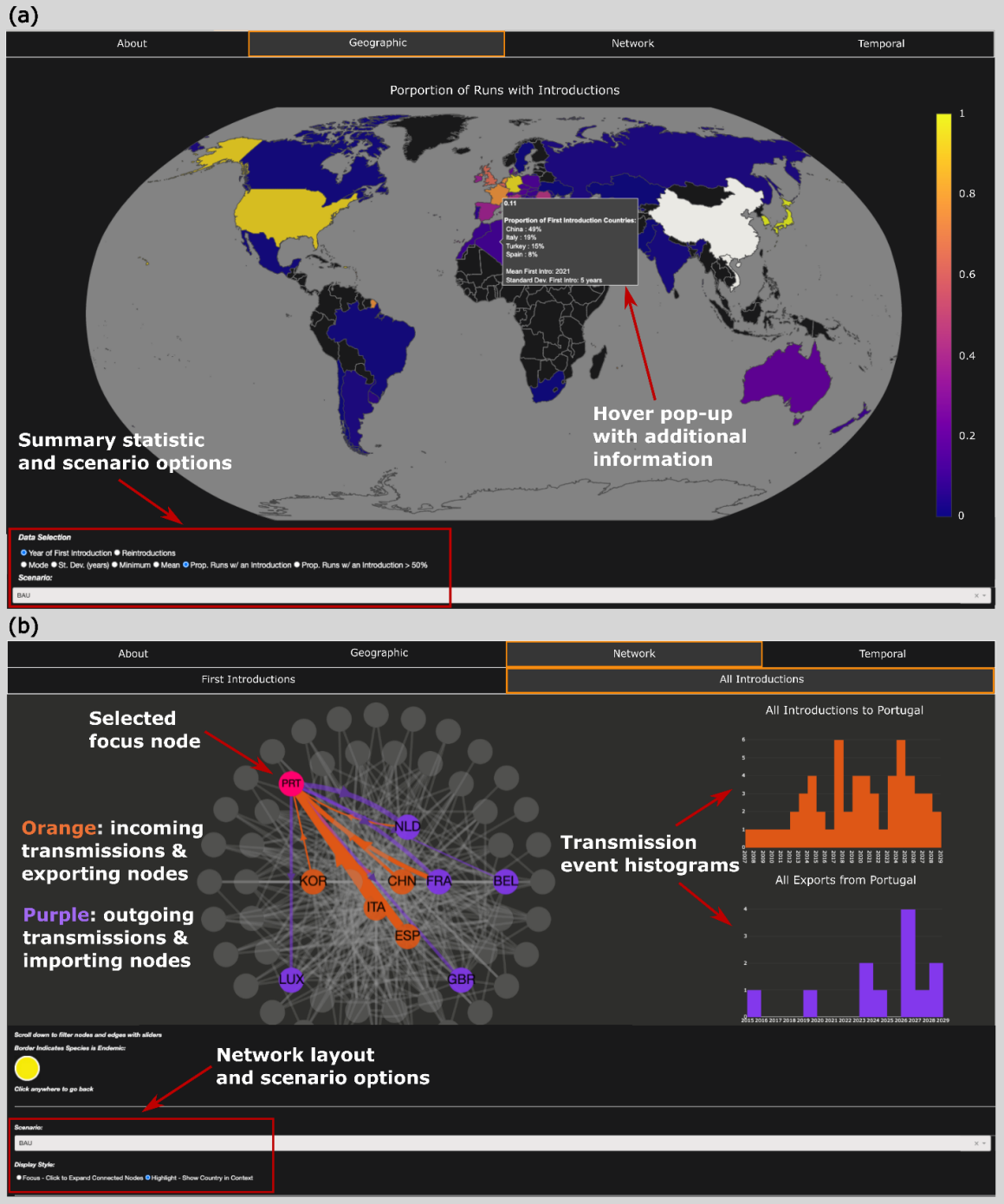


Figure 4. The first iteration of the PoPS Global analytics dashboard (https://pops-global.ngrok.io/) displays (a) a geographic view of forecast results with country-level summary statistics of pest introduction probability, timing, and sources and (b) an interactive pest transmission network, which allows users to filter the network and explore the timing and magnitude of predicted spread pathways. See Supporting Information S4 for additional screenshots and description of the analytics dashboard.

2.4 Applying the Framework: A Case Study of Spotted Lanternfly

We demonstrate the framework by simulating the global spread of spotted lanternfly, a phytophagous planthopper and emerging pest of many commercially important fruit plants and timber. Spotted lanternfly is native to China and possibly other subtropical regions of Southeast Asia (Dara et al., 2015; H. Kim et al., 2013; J. G. Kim et al., 2011). We used China and Vietnam as origin countries (Lee et al., 2019; Wakie et al., 2020) and ran the simulation from 2006 to 2029. We modeled transmission lag as a stochastic process by drawing a lag time at each simulated introduction from a gamma probability distribution ($\alpha = 4$, $\beta = 1$), with a mean lag time of 4 years. The lag time distribution was parameterized based on a review of thirty-five invasive insects in Japan that found the mean detection lag time was 3.8 years when insects with extremely long latency periods were excluded (Kiritani & Yamamura, 2003).

Spotted lanternfly feeds on more than 70 plant species but has a strong preference for tree of heaven (*Ailanthus altissima*; Murman et al., 2020). We created a tree of heaven host map by using observations from the GBIF database to create a MAXENT species suitability model (Supporting Information S1.2). Spotted lanternfly can hitchhike long distances by depositing egg masses on surfaces like trees, vehicles, or stone. Experts suspect that egg masses arrived in the United States on imported stone, although plant material and wood packing material are also suspected carriers (Dara et al., 2015; USDA-APHIS-PPQ-CPHST, 2018). In our simulation, we used monthly imported values (2019 U.S.$) of stone commodities (HS 6801, 6802, 6803, 6804) obtained from the UN Comtrade database and included only stone commodities that provide a solid substrate for carrying eggs, given that adults are very unlikely to survive long-distance transport. We created a simple trade forecast to 2029 by randomly drawing a value for each month per country from their respective 2015–2019 values. We did not forecast a change in trade patterns for this case study; however, trade increase or decrease scenarios can easily be included in PoPS Global by applying a percent change to the drawn forecast data. We incorporated seasonality by constraining the potential for transporting eggs to only those months suitable for egg-laying in each origin country (Dara et al., 2015; USDA-APHIS-PPQ-CPHST, 2018; Wakie et al., 2020). GBIF records for spotted lanternfly include first observations for three established populations outside the native range: the Republic of Korea in 2009, Japan in 2010, and the U.S. in 2014. These three observations were used to calibrate the network model (Supporting Information S2).

Each input module used in the framework can be manipulated to simulate scenarios and understand the impact of changes. For example, available host area could be decreased to simulate efforts to eradicate invasive tree of heaven. From a national phytosanitary perspective, an agency might explore how to disrupt the transmission network, such as by increasing resources for phytosanitary activities or reducing imports of goods known to carry pests from areas with established populations. To demonstrate how the framework can be used for testing scenarios, we compared two potential responses to the discovery of spotted lanternfly in Pennsylvania in 2014. In one scenario, we disrupted the transmission network by stopping U.S. imports of stone commodities from the native countries, China and Vietnam, beginning in 2014. In the other, we used the PoPS Global forecasting framework to design a management plan for limiting stone imports by identifying where and when countries with potential bridgehead populations were likely to export spotted lanternfly to the U.S.

3 Results

We used PoPS Global to simulate the international spread of spotted lanternfly through trade between 2006 and 2029, identifying potential global patterns of invasion pressure over space and time (Figure 5). Introductions in the three countries with known invasive populations (the Republic of Korea, Japan, and the U.S.) were predicted in nearly all stochastic realizations. Introductions were also consistently predicted for several European countries including Germany, France, the Netherlands, Belgium, and the United Kingdom. First introduction year is relatively consistent for the Republic of Korea and Japan and more variable for the U.S. and Germany (Figure 5). The framework's network approach illuminates likely indirect pathways of pest movement via bridgehead populations over time. We identified several potential spotted lanternfly bridgehead populations and quantified their potential for exporting the pest to other countries with environmental conditions suitable for establishment (Figure 6).

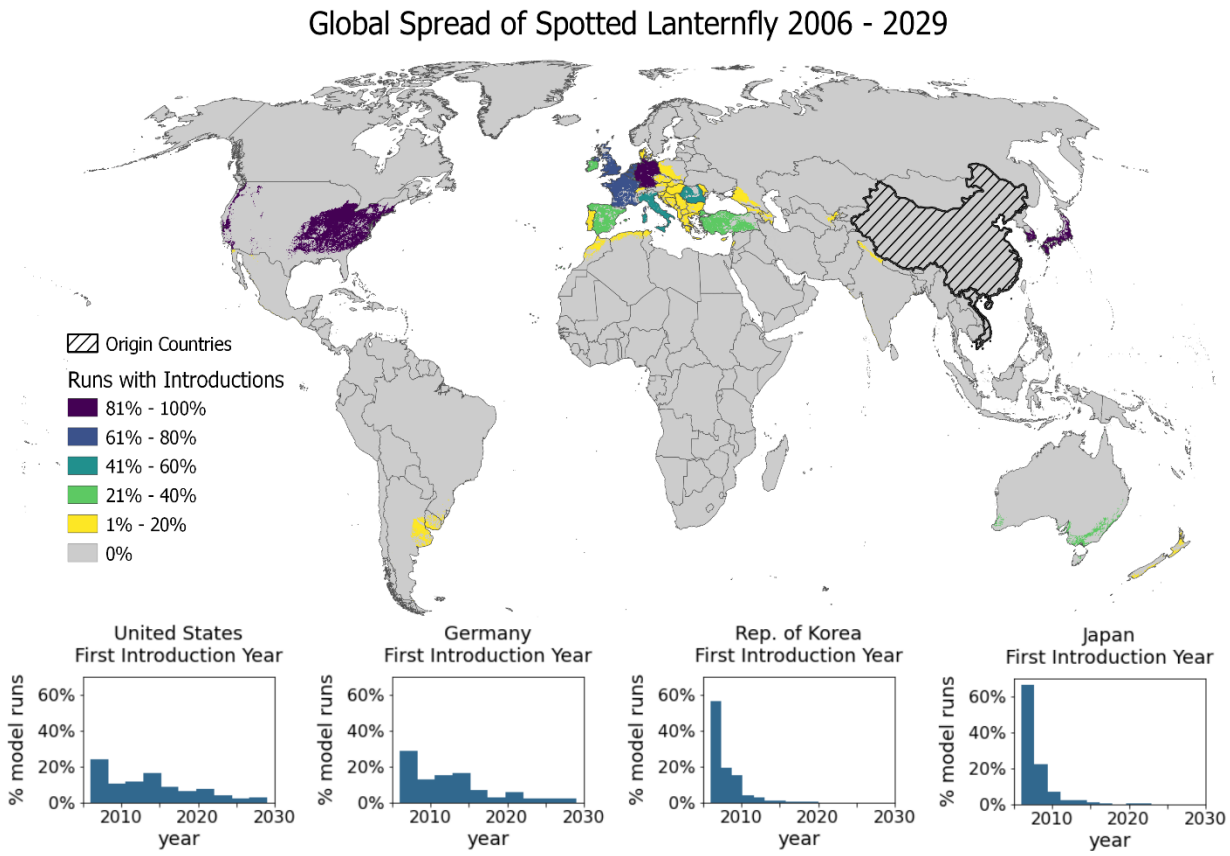


Figure 5. Predicted global distribution of spotted lanternfly by 2029. Colors indicate the proportion of the 1,000 stochastic model realizations that predicted introductions for each country. Histograms of predicted first introduction years are shown for the countries most consistently predicted to be invaded.

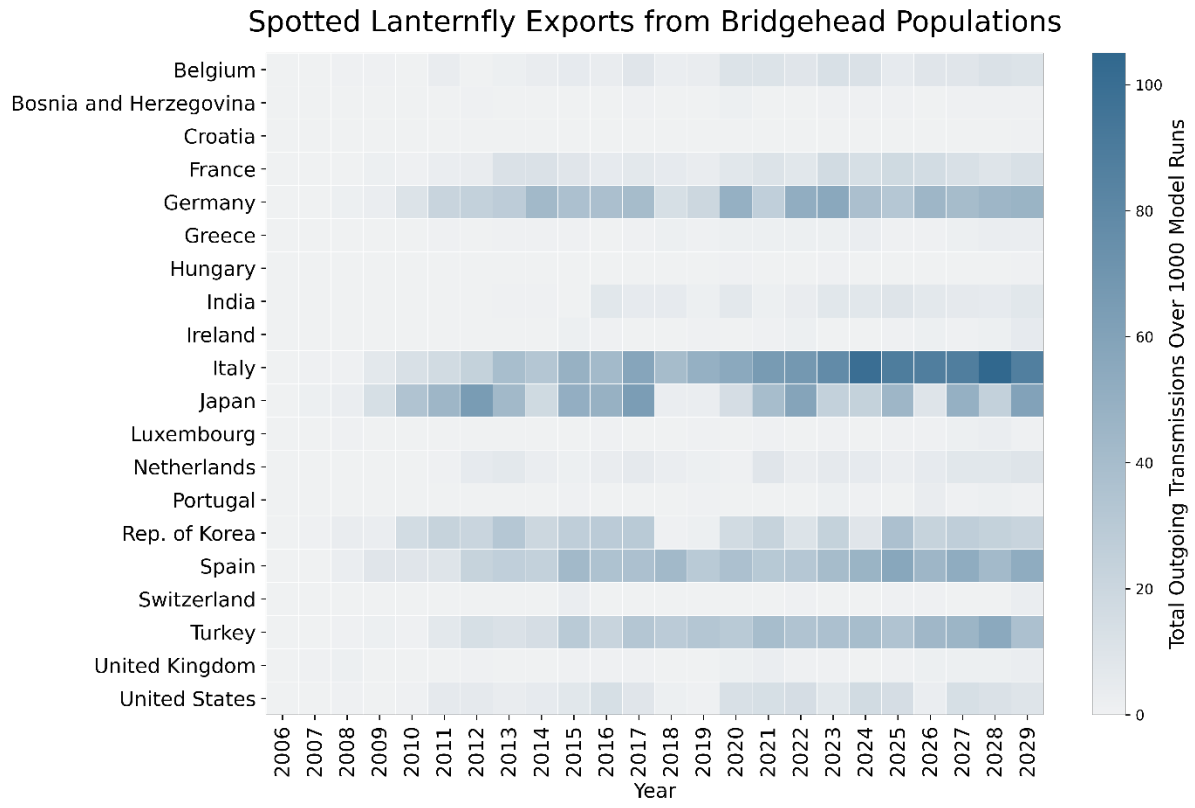


Figure 6. Frequency of exports predicted from bridgehead populations, summed over 1,000 stochastic model realizations.

We identified five spotted lanternfly transport pathways to the U.S. that could be disrupted by reducing imports of stone goods (Figure 7); one pathway is directly from a native country (China) and four are from bridgehead populations (the Republic of Korea, Japan, Italy, and Turkey). In the simulation-informed management scenario, we reduced imports of stone goods from China and Vietnam and each of the bridgehead populations by 100% in 2014 or the simulated 10th percentile first export year (Figure 7), whichever was latest. Disrupting direct pathways of transport from only native countries (Figure 8) reduced the number of annual introductions predicted after 2014, but some introductions continued to occur. Disrupting the transmission pathways from the native countries as well as the bridgehead populations predicted by the simulation resulted in nearly zero additional introduction after 2014 (Figure 8).

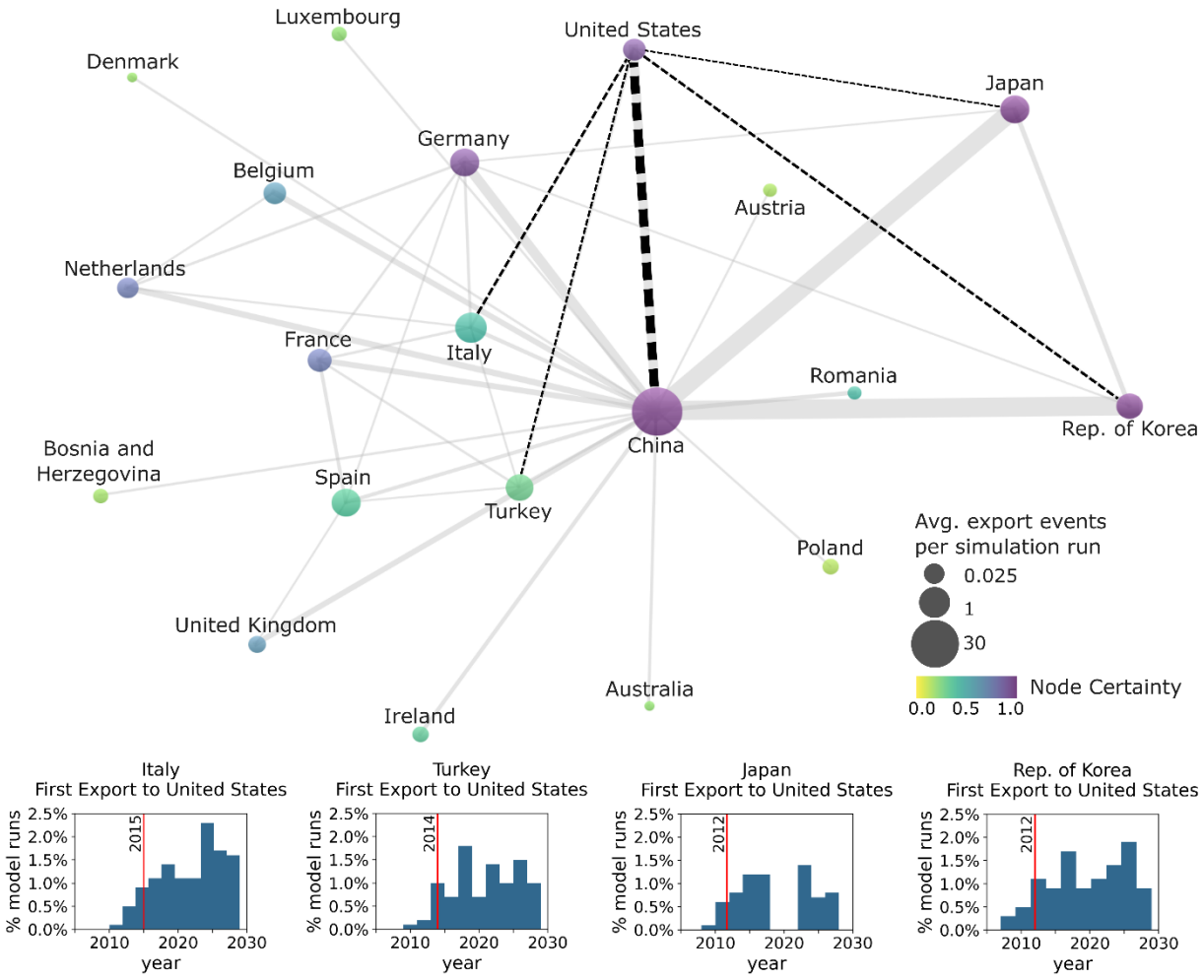


Figure 7. Simulated spotted lanternfly transmission network. Node size indicates the number of times the country exported the pest between 2006–2029. Node color represents certainty, i.e., the proportion of 1,000 stochastic realizations predicting transmission to or from the node. Edge weight represents the number of transmissions between countries. Edges directly connected to the U.S. are dashed black lines. The network is filtered to include nodes with at least 150 incoming or outgoing transport events and edges with at least 75 transport events over 1,000 stochastic model realizations. Histograms of the simulated first export event to the U.S. are shown for each bridgehead node; the vertical red line marks the 10th percentile year, which was used as a conservative estimate of the earliest export event.

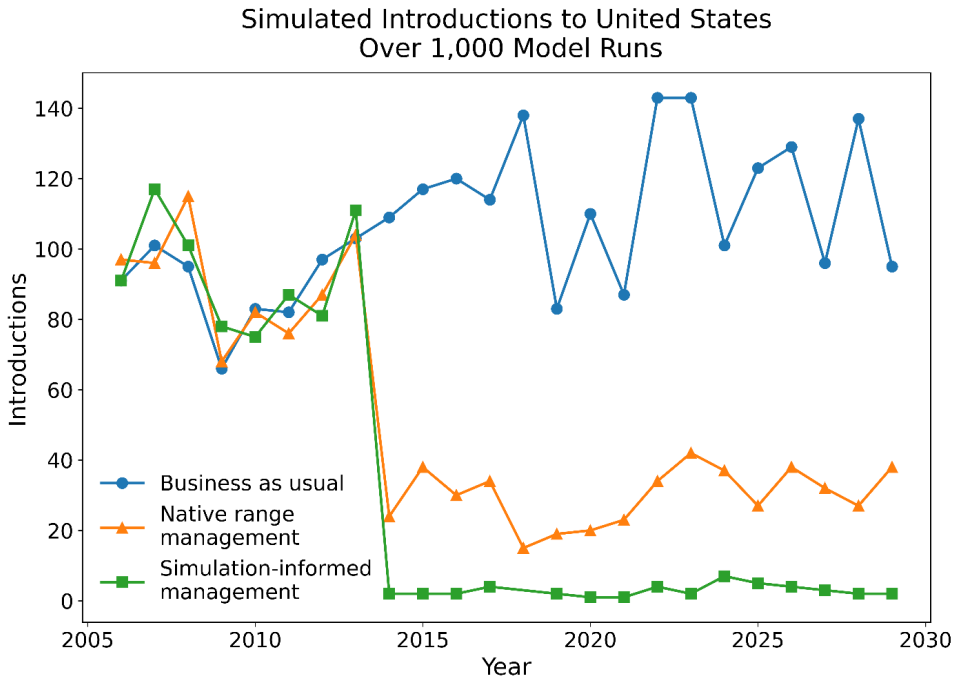


Figure 8. Cumulative number of annual introductions to the U.S. with no change in stone commodity trading (business as usual, blue line) vs. two management scenarios: native range management that reduces U.S. imports of stone commodities from China and Vietnam starting in 2014 (orange line), and simulation-informed management that uses a PoPS Global forecast to determine when and where to reduce U.S. imports of stone commodities from native and bridgehead populations (green line).

4 Discussion

Our case study of spotted lanternfly demonstrates the utility of the PoPS Global framework to hindcast and forecast the international spread of emerging pests and identify potential bridgehead populations. Identifying these bridgeheads is important for directing pre-border biosecurity efforts in infested areas and for tailoring inspection of goods from bridgehead countries at ports of entry (Epanchin-Niell et al., 2021). PoPS Global successfully predicted spotted lanternfly spread to the three known invaded countries to date (Japan, the Republic of Korea, and the U.S.) as well as identified potential bridgehead populations in Europe and quantified pathway risk.

The PoPS Global framework addresses critical needs for forecasting plant pest invasion pathways when very little is known about past movement. It leverages globally available open

data to implicitly model the effect of propagule pressure from international trade, and each simulated introduction encompasses the entire establishment process, from overcoming movement barriers to developing a self-sustaining population (Blackburn et al., 2011). The transmission lag time provides a simplified, stochastic conceptualization of the time required for a newly introduced population to successfully reproduce and be subsequently exported. This is a necessary simplification as population spread is highly variable and difficult to generalize at a global scale (Garnas et al., 2016; Vose, 2008). Lag time will depend on myriad factors such as intra- and interspecific interactions (e.g., competition, predation, Allee effect), dispersal mechanisms, detection rates, and other complex processes (Crooks, 2005; Morimoto et al., 2019). However, if subsequent introductions are simulated, the impact of increasing propagule pressure is incorporated by drawing another transmission lag from a probability distribution and using the shortest time drawn. While the effect of reintroductions and population growth are not explicitly modeled, the likelihood of short transmission lag times is higher for countries with simulated repeat introductions.

Fine resolution temporal data on the timing of species introductions are often not readily available and will be very challenging to obtain with high accuracy due to inherent lags in detection (Crooks, 2005). Temporal imprecision is likely to persist in this type of forecast. Initial model conditions, including species distribution and simulation start year, are a source of forecast uncertainty for PoPS Global. Species distributions, in both native and invaded ranges, are often poorly documented, limiting the ability to estimate the potential for propagule transport between countries (Bebber et al., 2019; McGeoch et al., 2010; Rouget et al., 2016). The simulation is also sensitive to the start year used. We conceptualized "start year" as the year that an unobserved event results in species range expansion and propagule transport. The optimal simulation start year can be calibrated alongside the other unknown model parameters, alpha and lambda, but we suggest that future application of the framework incorporate start year as a random variable along with other model parameters and estimate parameter distributions through approximate Bayesian computation (Minter & Retkute, 2019).

Despite limitations, the PoPS Global simulation provides insights into broad patterns of introduction timing and sequence (e.g., native country A transmits to bridgehead country B, which then transmits to country C after a latency period). There are opportunities to increase the spatial resolution by reducing the area represented by nodes to regions around ports of entry,

similar to subpopulation areas surrounding airports used by Balcan et al. (2010). Modeling transport between smaller regions could be especially useful for large countries with concentrated areas of susceptible hosts, such as fruit-producing regions of the U.S. West Coast. Higher spatial resolution also provides opportunities for explicitly modeling natural dispersal mechanisms and regional transmission between adjacent nodes. Modeling multiple dispersal mechanisms within the network could be achieved by designating edge types based on Euclidean distance between nodes, with nearby nodes connected by edges describing natural dispersal and more distant nodes connected by trade-driven edges (Brooks et al., 2008). When adequate global data are available, PoPS Global can also be coupled with spatially explicit models of landscape-scale spread (Jones et al., 2021; Meentemeyer et al., 2011). Further research is needed to understand how the large uncertainties in temporal and spatial data used for model calibration may limit the usefulness of modeling species movement at higher resolutions.

Creating actionable insights when very little information is available for model calibration and validation is challenging. However, the threats posed by emerging plant pests and pathogens necessitate quick, data-driven decisions. A near-term forecasting approach creates opportunities for frequent, iterative updating with new observations to improve understanding of global biological invasions, build predictive capacity, and reduce uncertainty (Dietze et al., 2018). The PoPS Global framework can be used to identify when and where data are needed to create dynamic, adaptive monitoring efforts in response to forecast uncertainties (Figure 1.3). Novel techniques, such as web scraping and text mining, represent an opportunity to leverage free, publicly available information from online news articles, social media, technical forums, and research journals to collect up-to-date information about where pest species are being discovered globally, reducing latency between data collection and availability (Jarić et al., 2021). Critical to improving our ability to forecast invasions is coordinated global surveillance to catalog pest occurrence data from native and introduced ranges, species traits, and the potential impacts of spread (Carvajal-Yepes et al., 2019). In addition to these data, biosecurity efforts need to be framed within a global context by treating the spread of invasive plant pests as pandemics and learning from advances made in global public health (Hulme, 2021; Ristaino et al., 2021). The PoPS Global framework provides a modular approach for risk assessment that enables phytosanitary agencies to quickly evaluate and visualize globally connected invasion processes at multiple scales.

Acknowledgements

This research was supported by funding from the United States Department of Agriculture's Animal and Plant Health Inspection Service (APHIS) grant awarded to CJ and RKM (grant number AP19PPQS&T00C236). This paper does not necessarily reflect the views of APHIS. Thank you to Dr. Ron Sequeira and Dr. Yu Takeuchi for inspiration and guidance during conceptualization of the modeling framework. We would also like to thank Dr. Megan Skrip for her helpful review and editorial comments, Dr. Shannon Jones for creating Figures 1-3, and Dr. Regan Early for sharing the national phytosanitary capacity data.

REFERENCES

Bacon, S. J., Bacher, S., & Aebi, A. (2012). Gaps in border controls are related to quarantine alien insect invasions in Europe. *PLoS One*, *7*(10). Agricultural & Environmental Science Collection; ProQuest Central; SciTech Premium Collection. https://doi.org/10.1371/journal.pone.0047689

Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J., & Vespignani, A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, *106*(51), 21484–21489. https://doi.org/10.1073/pnas.0906910106

Balcan, D., Gonçalves, B., Hu, H., Ramasco, J. J., Colizza, V., & Vespignani, A. (2010). Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science*, *1*(3), 132–145. https://doi.org/10.1016/j.jocs.2010.07.002

Banks, N. C., Paini, D. R., Bayliss, K. L., & Hodda, M. (2015). The role of global trade and transport network topology in the human-mediated dispersal of alien species. *Ecology Letters*, *18*(2), 188–199. https://doi.org/10.1111/ele.12397

Bebber, D. P., Field, E., Gui, H., Mortimer, P., Holmes, T., & Gurr, S. J. (2019). Many unreported crop pests and pathogens are probably already present. *Global Change Biology*, *25*(8), 2703–2713. https://doi.org/10.1111/gcb.14698

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 180214. https://doi.org/10.1038/sdata.2018.214

Bellard, C., Leroy, B., Thuiller, W., Rysman, J.-F., & Courchamp, F. (2016). Major drivers of invasion risks throughout the world. *Ecosphere*, *7*(3), e01241. https://doi.org/10.1002/ecs2.1241

Bertelsmeier, C., & Keller, L. (2018). Bridgehead Effects and Role of Adaptive Evolution in Invasive Populations. *Trends in Ecology & Evolution*, *33*(7), 527–534. https://doi.org/10.1016/j.tree.2018.04.014

Bertelsmeier, C., Ollier, S., Liebhold, A. M., Brockerhoff, E. G., Ward, D., & Keller, L. (2018). Recurrent bridgehead effects accelerate global alien ant spread. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(21), 5486–5491. https://doi.org/10.1073/pnas.1801990115

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., Wilson, J. R. U., & Richardson, D. M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, *26*(7), 333–339. https://doi.org/10.1016/j.tree.2011.03.023

Brasier, C. M. (2008). The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathology*, *57*(5), 792–808.

Brooks, C. P., Antonovics, J., & Keitt, T. H. (2008). Spatial and temporal heterogeneity explain disease dynamics in a spatially explicit network model. *The American Naturalist*, *172*(2), 149–159. https://doi.org/10.1086/589451

CABI. (2021). *Invasive Species Compendium*. www.cabi.org/isc

Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K. A., Giovani, B., Saunders, D. G. O., Kamoun, S., Legg, J. P., Verdier, V., Lessel, J., Neher, R. A., Day, R., Pardey, P., Gullino, M. L., Records, A. R., Bextine, B., Leach, J. E., Staiger, S., & Tohme, J. (2019). A global surveillance system for crop diseases. *Science (New York, N.Y.)*, *364*(6447), 1237–1239. https://doi.org/10.1126/science.aaw1572

Chapman, D., Purse, B., Roy, H., & Bullock, J. (2017). Global trade networks determine the distribution of invasive non-native species. *Global Ecology and Biogeography*, *26*, 907–917.

Crooks, J. A. (2005). Lag times and exotic species: The ecology and management of biological invasions in slow-motion. *Ecoscience*, *12*(3), 316–329. https://doi.org/10.2980/i1195-6860-12-3-316.1

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

Dara, S. K., Barringer, L., & Arthurs, S. P. (2015). *Lycorma delicatula* (Hemiptera: Fulgoridae): A New Invasive Pest in the United States. *Journal of Integrated Pest Management*, *6*(20). https://doi.org/10.1093/jipm/pmv021

DeAngelis, D. L., & Yurek, S. (2017). Spatially explicit modeling in ecology: A review. *Ecosystems*, *20*(2), 284–300. https://doi.org/10.1007/s10021-016-0066-z

Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I., Salles, J.-M., Bradshaw, C. J. A., & Courchamp, F. (2021). High and rising economic costs of biological invasions worldwide. *Nature*, *592*(7855), 571–576. https://doi.org/10.1038/s41586-021-03405-6

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lunch, C. K., Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C., & White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*, *115*(7), 1424–1432. https://doi.org/10.1073/pnas.1710231115

Douma, J. C., Pautasso, M., Venette, R. C., Robinet, C., Hemerik, L., Mourits, M. C. M., Schans, J., & van der Werf, W. (2016). Pathway models for analysing and managing the

introduction of alien plant pests – an overview and categorization. *Ecological Modelling*, *339*, 58–67. https://doi.org/10.1016/j.ecolmodel.2016.08.009

Early, R., Bradley, B. A., Dukes, J. S., Lawler, J. J., Olden, J. D., Blumenthal, D. M., Gonzalez, P., Grosholz, E. D., Ibañez, I., Miller, L. P., Sorte, C. J. B., & Tatem, A. J. (2016). Global threats from invasive alien species in the twenty-first century and national response capacities. *Nature Communications*, *7*(1), 12485. https://doi.org/10.1038/ncomms12485

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Epanchin-Niell, R., McAusland, C., Liebhold, A., Mwebaze, P., & Springborn, M. R. (2021). Biological invasions and international trade: Managing a moving target. *Review of Environmental Economics and Policy*, 000–000. https://doi.org/10.1086/713025

Fones, H. N., Bebber, D. P., Chaloner, T. M., Kay, W. T., Steinberg, G., & Gurr, S. J. (2020). Threats to global food security from emerging fungal and oomycete crop pathogens. *Nature Food*, *1*(6), 332–342. https://doi.org/10.1038/s43016-020-0075-0

Food and Agriculture Organization of the United Nations. (2020). *FAOSTAT Database*. FAO. http://www.fao.org/faostat

Ford, D. A., Kaufman, J. H., & Eiron, I. (2006). An extensible spatial and temporal epidemiological modelling system. *International Journal of Health Geographics*, *5*(1), 4. https://doi.org/10.1186/1476-072X-5-4

Garnas, J. R., Auger-Rozenberg, M.-A., Roques, A., Bertelsmeier, C., Wingfield, M. J., Saccaggi, D. L., Roy, H. E., & Slippers, B. (2016). Complex patterns of global spread in invasive insects: Eco-evolutionary and management consequences. *Biological Invasions*, *18*(4), 935–952. https://doi.org/10.1007/s10530-016-1082-9

GBIF. (2021). *What is GBIF?* https://www.gbif.org/what-is-gbif

Harwood, T. D., Xu, X., Pautasso, M., Jeger, M. J., & Shaw, M. W. (2009). Epidemiological risk assessment using linked network and grid based modelling: *Phytophthora ramorum* and *Phytophthora kernoviae* in the UK. *Ecological Modelling*, *220*(23), 3353–3361. https://doi.org/10.1016/j.ecolmodel.2009.08.014

Hulme, P. E. (2009). Trade, transport and trouble: Managing invasive species pathways in an era of globalization. *The Journal of Applied Ecology*, *46*(1), 10–18. https://doi.org/10.1111/j.1365-2664.2008.01600.x

Hulme, P. E. (2021). Advancing one biosecurity to address the pandemic risks of biological invasions. *BioScience*, *biab019*. https://doi.org/10.1093/biosci/biab019

Jarić, I., Bellard, C., Correia, R. A., Courchamp, F., Douda, K., Essl, F., Jeschke, J. M., Kalinkat, G., Kalous, L., Lennox, R. J., Novoa, A., Proulx, R., Pyšek, P., Soriano–Redondo, A.,

Souza, A. T., Vardi, R., Veríssimo, D., & Roll, U. (2021). Invasion culturomics and iEcology. *Conservation Biology*, *35*(2), 447-451. https://doi.org/10.1111/cobi.13707

Jones, C., Jones, S., Petrasova, A., Petras, V., Gaydos, D., Skrip, M., Takeuchi, Y., Bigsby, K., & Meentemeyer, R. (2021). Iteratively forecasting invasions with PoPS and a little help from our friends. *Frontiers In Ecology And The Environment*. https://doi.org/doi:10.1002/fee.2357

Kim, H., Kim, M., Kwon, D. H., Park, S., Lee, Y., Huang, J., Kai, S., Lee, H.-S., Hong, K.-J., Jang, Y., & Lee, S. (2013). Molecular comparison of *Lycorma delicatula* (Hemiptera: Fulgoridae) isolates in Korea, China, and Japan. *Journal of Asia-Pacific Entomology*, *16*(4), 503–506. https://doi.org/10.1016/j.aspen.2013.07.003

Kim, J. G., Lee, E.-H., Seo, Y.-M., & Kim, N.-Y. (2011). Cyclic Behavior of *Lycorma delicatula* (Insecta: Hemiptera: Fulgoridae) on Host Plants. *Journal of Insect Behavior*, *24*(6), 423. https://doi.org/10.1007/s10905-011-9266-8

Kiritani, K., & Yamamura, K. (2003). Exotic insects and their pathways for invasion. In G. M. Ruiz & J. T. Carlton (Eds.), *Invasive species: Vectors and management strategies* (pp. 44–67). Island Press.

Lee, D.-H., Park, Y.-L., & Leskey, T. C. (2019). A review of biology and management of *Lycorma delicatula* (Hemiptera: Fulgoridae), an emerging global invasive species. *Journal of Asia-Pacific Entomology*, *22*(2), 589–596. https://doi.org/10.1016/j.aspen.2019.03.004

Lenzner, B., Leclère, D., Franklin, O., Seebens, H., Roura-Pascual, N., Obersteiner, M., Dullinger, S., & Essl, F. (2019). A framework for global twenty-first century scenarios and models of biological invasions. *BioScience*, *69*(9), 697–710. https://doi.org/10.1093/biosci/biz070

Lewis, M. A., Petrovskii, S. V., & Potts, J. R. (2016a). Invasion in a multispecies system. In M. A. Lewis, S. V. Petrovskii, & J. R. Potts (Eds.), *The Mathematics Behind Biological Invasions* (pp. 107–154). Springer International Publishing. https://doi.org/10.1007/978-3-319-32043-4_4

Lewis, M. A., Petrovskii, S. V., & Potts, J. R. (2016b). *The mathematics behind biological invasions* (Vol. 44). https://doi.org/10.1007/978-3-319-32043-4

Lombaert, E., Guillemaud, T., Cornuet, J.-M., Malausa, T., Facon, B., & Estoup, A. (2010). Bridgehead effect in the worldwide invasion of the biocontrol Harlequin Ladybird. *PLoS One; San Francisco*, *5*(3), e9743. http://dx.doi.org.prox.lib.ncsu.edu/10.1371/journal.pone.0009743

Magarey, R. D., Colunga-Garcia, M., & Fieselmann, D. A. (2009). Plant biosecurity in the United States: Roles, responsibilities, and information needs. *BioScience*, *59*(10), 875–884. https://doi.org/10.1525/bio.2009.59.10.9

McGeoch, M. A., Butchart, S. H. M., Spear, D., Marais, E., Kleynhans, E. J., Symes, A., Chanson, J., & Hoffmann, M. (2010). Global indicators of biological invasion: Species numbers, biodiversity impact and policy responses. *Diversity and Distributions*, *16*(1), 95–108. https://doi.org/10.1111/j.1472-4642.2009.00633.x

Meentemeyer, R. K., Cunniffe, N. J., Cook, A. R., Filipe, J. A. N., Hunter, R. D., Rizzo, D. M., Gilligan, C. A., Filipe, J. A. N., Hunter, R. D., Rizzo, D. M., & Gilligan, C. A. (2011). Epidemiological modeling of invasion in heterogeneous landscapes: Spread of sudden oak death in California (1990-2030). *Ecosphere*, *2*(2), art17. https://doi.org/10.1890/es10-00192.1

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

Minter, A., & Retkute, R. (2019). Approximate Bayesian Computation for infectious disease modelling. *Epidemics*, *29*, 100368. https://doi.org/10.1016/j.epidem.2019.100368

Montgomery, K., Petras, V., Katsar, C. S., & Takeuchi, Y. (2021). Contaminated consignment simulation to support risk-based inspection design [Manuscript submitted for publication].

Morimoto, N., Kiritani, K., Yamamura, K., & Yamanaka, T. (2019). Finding indications of lag time, saturation and trading inflow in the emergence record of exotic agricultural insect pests in Japan. *Applied Entomology and Zoology*, *54*(4), 437–450. https://doi.org/10.1007/s13355-019-00640-2

Murman, K., Setliff, G. P., Pugh, C. V., Toolan, M. J., Canlas, I., Cannon, S., Abreu, L., Fetchen, M., Zhang, L., Warden, M. L., Wallace, M., Wickham, J., Spichiger, S.-E., Swackhamer, E., Carrillo, D., Cornell, A., Derstine, N. T., Barringer, L., & Cooperband, M. F. (2020). Distribution, survival, and development of spotted lanternfly on host plants found in North America. *Environmental Entomology*, *49*(6), 1270–1281. https://doi.org/10.1093/ee/nvaa126

Paini, D. R., Sheppard, A. W., Cook, D. C., De Barro, P. J., Worner, S. P., & Thomas, M. B. (2016). Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences*, *113*(27), 7575–7579. https://doi.org/10.1073/pnas.1602205113

Parnell, S., van den Bosch, F., Gottwald, T., & Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: An epidemiological perspective. *Annual Review of Phytopathology*, *55*(1), 591–610. https://doi.org/10.1146/annurev-phyto-080516-035334

Petrovskii, S., Morozov, A., & Li, B.-L. (2005). Regimes of biological invasion in a predator-prey system with the Allee effect. *Bulletin of Mathematical Biology*, *67*(3), 637. https://doi.org/10.1016/j.bulm.2004.09.003

Petrovskii, S. V., Malchow, H., Hilker, F. M., & Venturino, E. (2005). Patterns of patchy spread in deterministic and stochastic models of biological invasion and biological control. *Biological Invasions*, *7*(5), 771–793. https://doi.org/10.1007/s10530-005-5217-7

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pyšek, P., & Richardson, D. M. (2010). Invasive species, environmental change and management, and health. *Annual Review of Environment and Resources*, *35*(1), 25–55. https://doi.org/10.1146/annurev-environ-033009-095548

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

Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution*, *29*(4), 233–242. https://doi.org/10.1016/j.tree.2014.02.003

Rouget, M., Robertson, M. P., Wilson, J. R. U., Hui, C., Essl, F., Renteria, J. L., & Richardson, D. M. (2016). Invasion debt – quantifying future biological invasions. *Diversity and Distributions*, *22*(4), 445–456. https://doi.org/10.1111/ddi.12408

Sanderson, E. W., Jaiteh, M., Levy, M. A., Redford, K. H., Wannebo, A. V., & Woolmer, G. (2002). The Human Footprint and the Last of the Wild: The human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. *BioScience*, *52*(10), 891–904. https://doi.org/10.1641/0006-3568(2002)052[0891:THFATL]2.0.CO;2

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, *3*(3), 430–439. https://doi.org/10.1038/s41559-018-0793-y

Seebens, H., Gastner, M. T., & Blasius, B. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters*, *16*(6), 782–790. https://doi.org/10.1111/ele.12111

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., Blasius, B., Brundu, G., Capinha, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Fuentes, N., Jäger, H., … Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*(1), 14435. https://doi.org/10.1038/ncomms14435

Seebens, H., Schwartz, N., Schupp, P. J., & Blasius, B. (2016). Predicting the spread of marine species introduced by global shipping. *Proceedings of the National Academy of Sciences*, *113*(20), 5646–5651. https://doi.org/10.1073/pnas.1524427113

Strickland, C., Dangelmayr, G., Shipman, P. D., Kumar, S., & Stohlgren, T. J. (2015). Network spread of invasive species and infectious diseases. *Ecological Modelling*, *309–310*(Journal Article), 1–9. https://doi.org/10.1016/j.ecolmodel.2015.04.010

USDA-APHIS-PPQ-CPHST. (2018). *Technical working group summary report spotted lanternfly, Lycorma delicatula (White, 1845)*.

Václavík, T., & Meentemeyer, R. K. (2009). Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling*, *220*(23), 3248–3258. https://doi.org/10.1016/j.ecolmodel.2009.08.013

Vose, D. (2008). *Risk analysis: A quantitative guide* (3rd ed.). Chichester, England ; Hoboken, NJ : Wiley, [2008]. https://catalog.lib.ncsu.edu/catalog/NCSU2163700

Wakie, T. T., Neven, L. G., Yee, W. L., & Lu, Z. (2020). The establishment risk of *Lycorma delicatula* (Hemiptera: Fulgoridae) in the United States and globally. *Journal of Economic Entomology*, *113*(1), 306–314. https://doi.org/10.1093/jee/toz259

Walden-Schreiner, C., Jones, C., & Montgomery, K. (2021). *ncsu-landscape-dynamics/PoPS-Global—Beta release* (Version v0.2-beta) [Computer software]. Zenodo. http://doi.org/10.5281/zenodo.5026883

Wildlife Conservation Society - WCS & Center for International Earth Science Information Network - CIESIN - Columbia University. (2005). *Last of the Wild Project, Version 2, 2005 (LWP-2): Global Human Influence Index (HII) Dataset (Geographic)*. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4BP00QC

CHAPTER 3: CONTAMINATED CONSIGNMENT SIMULATION TO SUPPORT RISK-BASED INSPECTION DESIGN

Under peer review at Risk Analysis

Authors

Kellyn Montgomery[1], Vaclav Petras[1], Yu Takeuichi[2], and Catherine S. Katsar[3]

Affiliations

[1]Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, USA

[2]NSF Center for Integrated Pest Management, North Carolina State University, Raleigh, NC, USA

[3]United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Plant Protection and Quarantine (PPQ), Raleigh, North Carolina, USA

Author Contributions

Kellyn M. led the design of the simulation; wrote code for the contamination and inspection modules, the simulation output, and the output visualizations; performed the simulation validation; designed and conducted the use cases; and wrote the manuscript. Vaclav P. made significant contributions to the ideation and design of the simulation; wrote code to create the Python package, automated code testing, and the prototype version of the simulation; and provided editorial review of the manuscript. Cathy Sue K. and Yu T. provided guidance on simulation design and use cases and manuscript editorial reviews.

Data Availability

Software and tools developed through this research are open-source and freely available under GNU General Public License. The source code for PoPS Border is at https://github.com/ncsu-landscape-dynamics/popsborder. Jupyter notebooks containing code used for the validation tests and the use cases can be found in the Github repository as well. The imported cut flower consignment inspection data collected by the USDA APHIS Agriculture Quarantine Inspection Monitoring (AQIM) program from January 2011 - October 2020 are not publicly available.

Abstract

Invasive non-native plant pests can cause extensive environmental and economic damage and are very difficult to eradicate once established. Phytosanitary inspections that aim to prevent biological invasions by limiting movement of non-native plant pests across borders are a critical component of the biosecurity continuum. Inspections can also provide valuable information about when and where plant pests are crossing national boundaries. However, only a limited portion of the massive volume of goods imported daily can be inspected, necessitating a highly targeted, risk-based strategy. Furthermore, since inspections must prioritize detection and efficiency, their outcomes generally cannot be used to make inferences about risk for cargo pathways as a whole. Phytosanitary agencies need better tools for quantifying pests going undetected and designing risk-based inspection strategies appropriate for changing operational conditions. In this research, we present PoPS (Pest or Pathogen Spread) Border, an open-source consignment inspection simulator for measuring inspection outcomes under various cargo contamination scenarios to support recommendations for inspection protocols and estimate pest slippage rates. We used the tool to estimate contamination rates of historical interception data, quantify trade-offs in effectiveness and workload for inspection strategies, and identify vulnerabilities in sampling protocols as changes in cargo configurations and contamination occur. These use cases demonstrate how this simulation approach permits testing inspection strategies and measuring quantities that would otherwise be impossible in a field-based setting. This work represents the first steps toward a decision support tool for creating dynamic inspection protocols that respond to changes in available resources, workload, and commerce trends.

Keywords: biological invasions, phytosanitary inspections, simulation

1 Introduction

Phytosanitary inspections of imported goods serve as an important biosecurity measure to prevent the transport of non-native species that can become invasive and cause extensive economic and environmental harm. However, given the volumes of goods imported into the United States, totaling 2.67 trillion dollars in 2019, it is infeasible for the limited number of border inspection personnel to fully inspect all incoming consignments (U.S. Customs and Border Protection, 2021). Efficiently targeting inspection resources through programs designed to expedite commodities with low risk, such as risk-based sampling and commodity release programs, are important strategies for minimizing the entry of non-native species under constrained resources (Kim et al., 2019; Robinson et al., 2017). Optimization of sampling strategies to detect low frequency contaminants with limited resources has been relatively well studied, with many recommendations for maximizing inspection efficiency (Chen et al., 2018; International Plant Protection Convention, 2016; Robinson et al., 2011; Robinson et al., 2015; Rossiter & Hester, 2017; Saccaggi et al., 2016; Surkov et al., 2008). In practice, however, inspections face diverse challenges and constraints often unaccounted for in optimization models related to limited personnel or other resources, commerce dynamics driving changes in import/export volumes and shifts in packaging trends, or prioritization of other objectives, like screening for trade of illicit items. Tools are needed to quantify inspection outcomes under various scenarios to choose the best strategy for specific conditions.

Another critical aspect of border biosecurity efforts is quantifying where and when pests are going undetected and entering domestic areas. Understanding pest propagule pressure can help prioritize post-border surveillance and management efforts (Lockwood et al., 2009). Many national phytosanitary agencies store decades of port interception records. However, most historical inspection data cannot be used to quantify patterns of pest slippage (i.e., proportion of pests or contaminants that fail to be intercepted by border inspections), because these data are generally collected during border security inspections designed to prioritize detection over statistical inference (Caley et al., 2015; Saccaggi et al., 2016). Sampling methods are often not recorded and vary widely by inspection objectives, resource availability, workload, and officer training (Saccaggi et al., 2016). This results in nonstatistical samples that cannot be used to draw conclusions about consignments or pathways as a whole. Many phytosanitary agencies conduct initiatives specifically designed to gather statistically valid interception data, such as the United

States Department of Agriculture (USDA) Animal and Plant Health Inspection Service (APHIS) Agriculture Quarantine Inspection Monitoring (AQIM) program (USDA APHIS PPQ, 2011). However, these programs are limited and do not collect data representing all incoming trade pathways. Directly measuring pest slippage rates is not possible in operational, field-based settings; therefore, alternative approaches are needed for leveraging limited data to estimate what is passing through international borders undetected.

Commerce trends, operational resources, and the mandate to streamline international trade make the optimal pest inspection strategy a moving target (Epanchin-Niell et al., 2021). There is a need for more flexible tools to help inspection agencies prevent entry of non-native species and identify vulnerabilities within inspection protocols to quantify pest slippage, while also responding to commerce and operational dynamics. In this research, we present PoPS (Pest or Pathogen Spread) Border -- an open-source software package designed to measure inspection outcomes under various cargo contamination scenarios to support recommendations for inspection protocols and estimate pest slippage rates. PoPS Border is part of a suite of tools developed to support phytosanitary management decisions along the biosecurity continuum, including PoPS, a generalized spatially explicit, discrete-time model used for simulating landscape spread of plant pests and pathogens (Jones et al., 2021) and PoPS Global, a species-agnostic, stochastic simulation and framework for forecasting global plant pest invasions (Montgomery et al., 2021). PoPS Border can be used to create synthetic data representing consignments with variations in sizes, cargo configurations, contamination rates, and contaminant distributions to test approaches for calculating sample sizes and selecting units for inspection (Figure 1). The effectiveness of the approaches for detecting contaminants can be quantified and compared to trade-offs in the amount of work and resources required. We demonstrate possible applications of the simulation via three use cases. First, we use PoPS Border to estimate contamination rates of cut flower consignments using high-quality inspection records from AQIM. Second, we quantify the trade-offs in inspection effectiveness and work required for various sampling strategies. Finally, we compare inspection outcomes for shipments with variations in cargo configuration and contamination. This work represents the first steps towards an agile decision support tool for designing dynamic inspection protocols that effectively respond to changes in available resources, workload, and commerce trends.

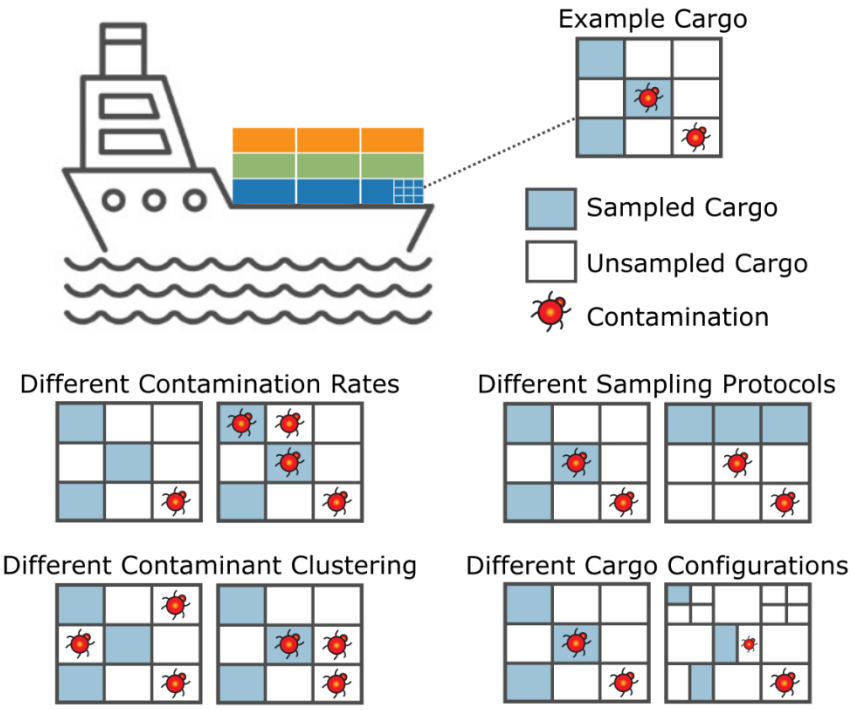


Figure 1. The consignment inspection simulation can be used to evaluate the effect of different contaminant quantities, contaminant clustering, sampling protocols, and cargo configurations on inspection outcomes. Blue squares represent sampled units.

2 Inspection Simulation Framework

PoPS Border is an open-source Python package for simulating inspections of contaminated consignments. The tool generates numerical representations of consignments, contaminates items within the consignments, samples units for inspection, and records the inspection outcomes. PoPS Border uses hypothetical scenarios for cargo configurations and contamination to create fully synthetic consignment data. Alternatively, it can be configured to produce realistic scenarios by providing additional information, such as phytosanitary inspections records to recreate real consignments or modeled estimates of noncompliance to simulate variations in pathway risk. We divided the simulation process into three steps - consignment generation, contamination, and inspection - with configurable parameters for each (Figure 2).

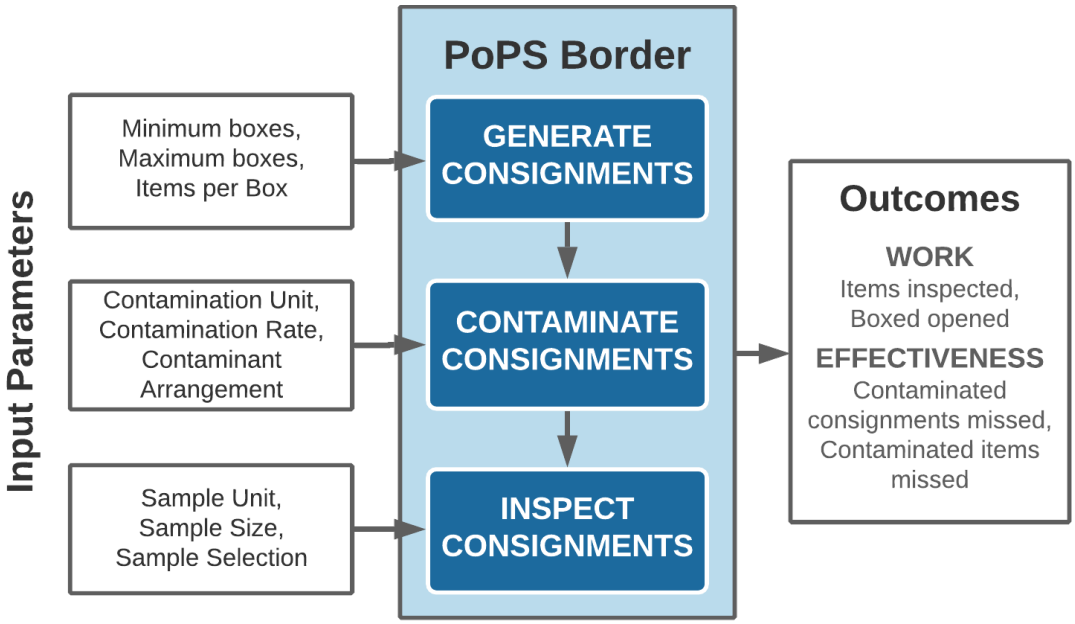


Figure 2. The PoPS Border framework consists of three components – consignment generation, consignment contamination, and consignment inspection. The user provides configuration parameters for each component. The simulation outputs include outcomes measuring inspection effort and effectiveness.

2.1 Consignment Generation

Consignments are generated as an array of items stratified into boxes and are given attributes describing the consignment origin, commodity, arrival date, port of entry, and mode of transport. If generating parameter-based synthetic consignments, the number of boxes in each consignment is chosen using a discrete uniform distribution to vary the consignment sizes between user-defined minimum and maximum number of total boxes. Each box is then filled with a specified number of items. Additional consignment attributes are assigned from lists of possible values provided by the user. Alternatively, if historical consignment inspection records are provided as a comma separated value (CSV) file, consignments will be generated to match each recorded shipment's size and attributes. If desired, the consignment attributes can be used to configure other simulation parameters. For example, the number of items per box can vary by commodity or mode of transport if criteria are provided.

2.2 Consignment Contamination

Approaches for contaminating consignments in PoPS Border are based on the contamination unit, rate, and arrangement. Contamination unit indicates whether individual items or entire boxes in the consignment should be contaminated. Contamination rate (i.e., the proportion of units in a consignment that are contaminated) can be fixed for all consignments or treated as a random variable and stochastically selected for each consignment from a beta probability distribution with user-defined shape parameters. The number of units to contaminate in each consignment is computed using the selected contamination rate and contamination unit as:

$$\mathrm{n_{contaminate}} = \mathrm{N} \times \mathrm{r}$$

where $n_{contaminate}$ is the number of units that will be contaminated in the consignment, N is the total number of contamination units in the consignment (boxes or sub-box items), and r is the contamination rate.

The user can also specify how contaminants are arranged within the consignments. The units to contaminate may be selected randomly or in a clustered arrangement. Examples of random and clustered contamination arrangements are shown in Figure 3 for the item contamination unit and Figure 4 for the box contamination unit. When using a clustered arrangement, the number of units to contaminate is divided by the user-defined cluster size to determine the total number of clusters needed. To ensure the clusters do not overlap, the array of units is divided into sections that are the size of one contaminant cluster, and the sections are randomly selected for each cluster. When using the item contamination unit with a clustered arrangement, the items to contaminate within each selected section are chosen either continuously or randomly. If done continuously, the contaminated units are placed sequentially, side by side. If done randomly, two parameters must be specified – total cluster width and number of contaminated items per cluster. The cluster width is the range over which items may be randomly selected for contamination. These two parameters determine the density of contaminated units within the clusters.

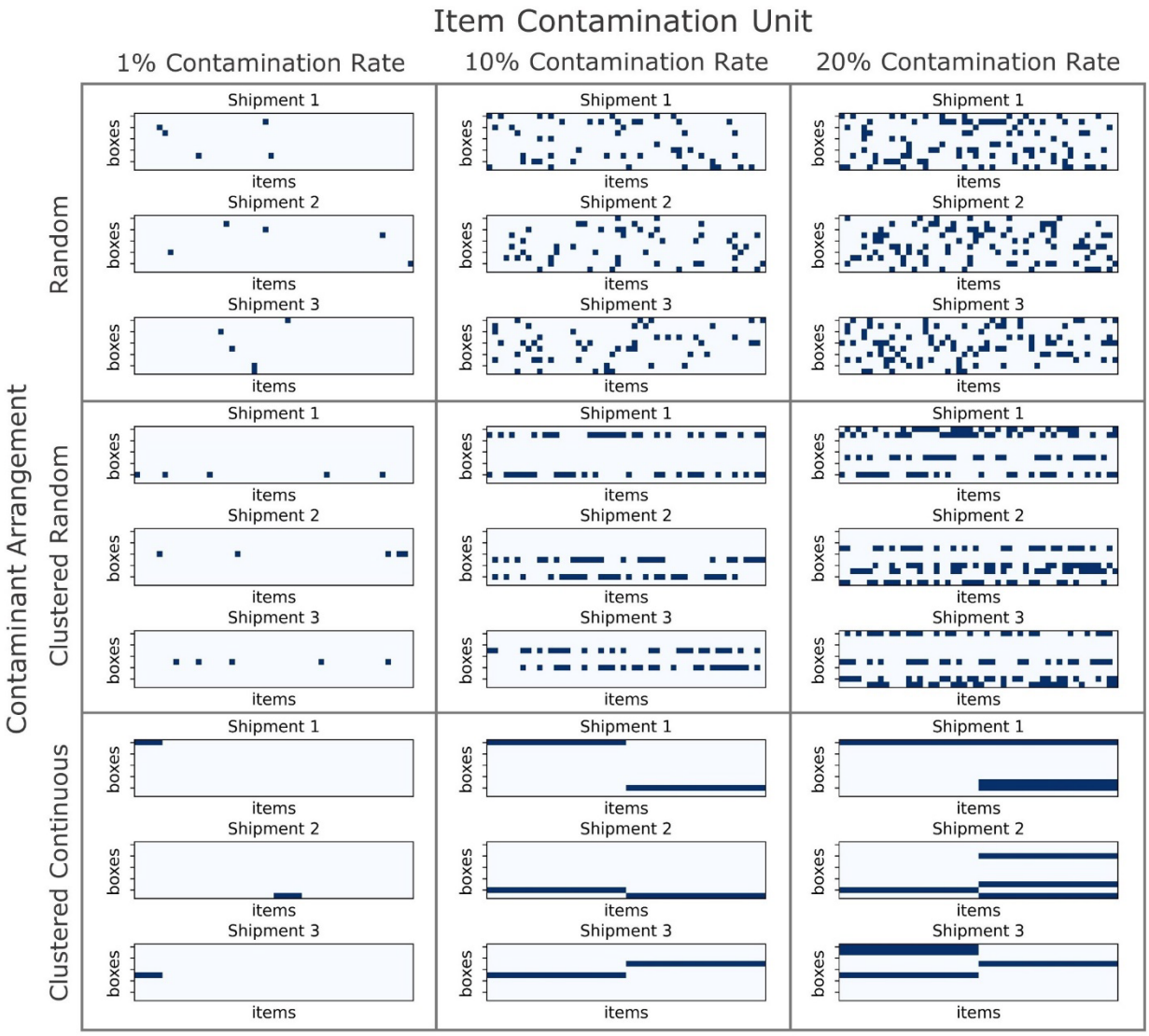


Figure 3. Examples of contaminated consignments with various contaminant arrangements at 1%, 10%, and 20% contamination rates using the item contamination unit. Each example includes three simulated consignments containing 10 boxes each (rows) with 50 items per box (columns). The dark blue grid cells represent contaminated items. The clustered arrangements use 25 contaminated items per cluster and a cluster width of 50 items.

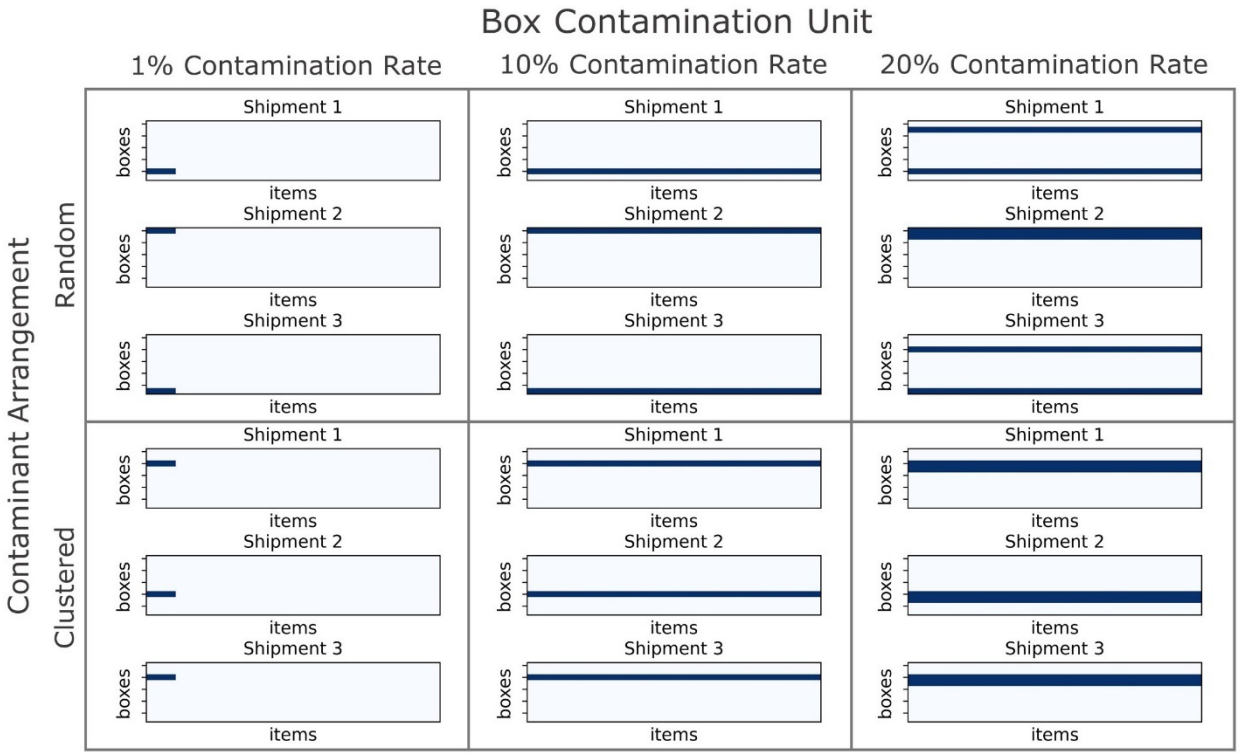


Figure 4. Examples of consignments with various contaminant arrangements at 1%, 10%, and 20% contamination rates using the box contamination unit. Each example includes three simulated consignments containing 10 boxes each (rows) with 50 items per box (columns). The dark blue grid cells represent contaminated items. The clustered contaminant arrangement is indistinguishable from the random arrangement when using the box contamination unit, unless the contamination rate is high enough to contaminate more than one box.

2.3 Consignment Inspection

Inspection strategies used in PoPS Border are a combination of inspection unit, sample size calculation method, and sample selection method. The inspection unit is used to define the total number of inspectable units in the consignment from which a sample is taken. The simulation is designed to accept two possible inspection units - boxes or items - although further packaging hierarchy could be implemented in the future. Options for sample size calculation methods include fixed number of units, fixed proportion of units, and using the hypergeometric distribution with user-defined detection and confidence levels (Fosgate, 2009; International Plant Protection Convention, 2016). The sample size when using the hypergeometric method is calculated as:

$$n_{sample} = \left(1 - (1 - \beta)^{\frac{1}{D \times N}}\right)\left(N - \frac{D \times N - 1}{2}\right)$$

where n_{sample} is the sample size, β is the confidence level, D is the detection level, and N is the number of inspection units in the consignment. If sample units are selected randomly, the hypergeometric sample size should detect contamination rates above or equal to the specified detection level with the level of confidence specified. The user can also specify a contamination tolerance level so that the simulation tracks the number of missed consignments with contamination rates below a phytosanitary threshold. This can be useful if a low level of contamination is considered tolerable or unavoidable and the user would like to adjust inspection slippage rates to only include missed consignments with contamination rates above the phytosanitary threshold.

For all sample selection methods, the selected units are inspected assuming 100% efficacy, so that if the unit is contaminated, it will be detected. The sample selection method options include choosing units uniform randomly, convenience style (i.e., select first N inspection units until sample size is met), systematically at a user-specified interval, or in clusters. Figures 5 and 6 show examples of simulated inspections using various sample sizes and selection methods. Cluster selection works by selecting groups of items (e.g., boxes) to make up the required total sample size. The cluster selection method is only available when using items as the inspection unit, as boxes represent the cluster unit. The user must specify the method for selecting the boxes to sample from, which can be systematic at a specified interval or uniformly random, and the minimum proportion of each box to inspect. If the sample size cannot be reached with the cluster specifications, the proportion of the boxes to inspect is automatically increased. For example, if interval clusters are used and every nth box is sampled from, the proportion of each box inspected must be high enough to reach the sample size. This is illustrated in the larger sample size example in Figure 5, where the cluster sizes were increased from 25% of each box sampled used in the random clusters, to approximately 55% of each box sampled in the interval clusters. The total number of clusters needed to reach the overall sample size is computed as:

$$n_{clusters} = \frac{n_{sample}}{i \times p}$$

where $n_{clusters}$ is the number of sample clusters required, n_{sample} is the total sample size, i is the number of items per box, and p is the proportion of each box that will be inspected.

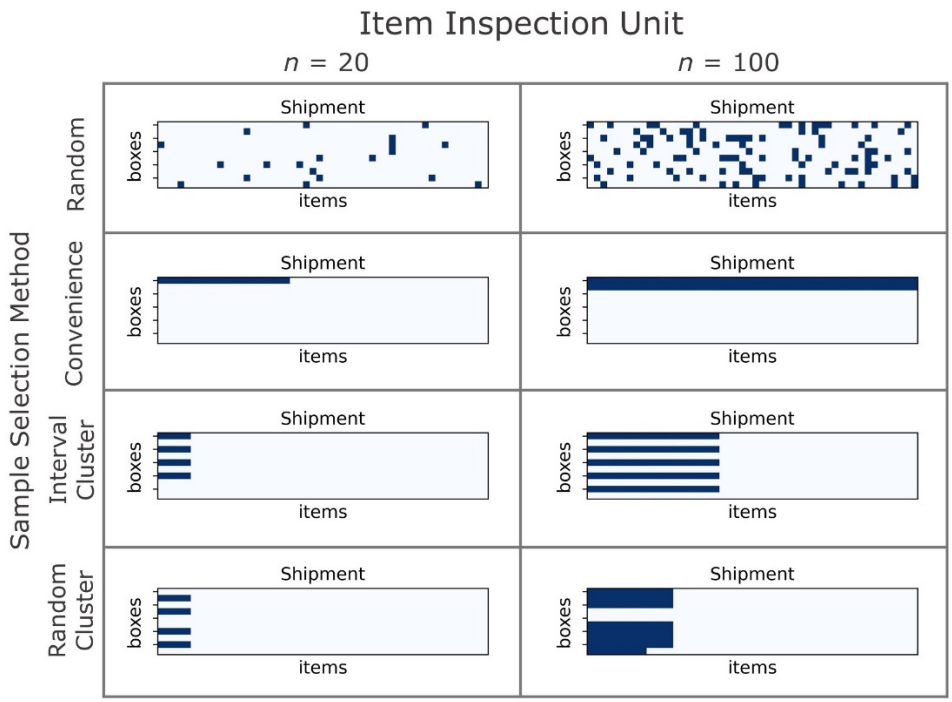


Figure 5. Examples of consignments sampled with various sample sizes and selection methods using item inspection unit. Each example includes one simulated consignment containing twenty boxes each (rows) with 100 items per box (columns). The dark blue grid cells represent inspected items.

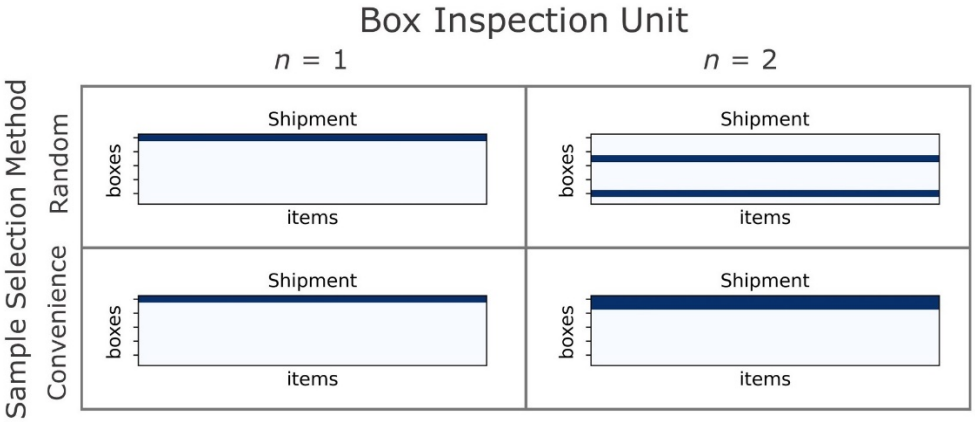


Figure 6. Examples of consignments sampled with various sample sizes and selection methods using box inspection unit. Each example includes one simulated consignment containing twenty boxes each (rows) with 100 items per box (columns). The dark blue grid cells represent inspected items. Box inspection unit sample sizes are rounded to the nearest integer.

2.4 Inspection Outcomes

For each inspected consignment, several metrics are recorded to quantify the inspection effort and effectiveness. To estimate the overall amount of work done, the number of items inspected and the number of boxes opened is recorded. The information recorded for measuring inspection effectiveness includes the successful detection of an existing contamination, the total number of contaminated units in the consignment (regardless of detection), and the number of contaminated units in the inspected sample. The outcomes are measured based on two scenarios - the entire sample is fully inspected, or the inspection stops after the first contamination is detected. Recording metrics for these two scenarios for each inspection provides additional information about the trade-offs between inspection effort and gathering statistically robust data on contamination rates. Inspections are simulated for a user-defined number of consignments, and the inspection results are aggregated. The simulation can be repeated multiple times to get a set of average outcomes across many stochastic runs. See Table 1 for the full list of inspection outcomes.

Table 1. Summary of inspection simulation outputs.

Effectiveness	Number of missed contaminated consignments Number of missed contaminated consignments within tolerance threshold Average missed contamination rate Number of intercepted contaminated consignments Average intercepted contamination rate Number of missed contaminants Number of intercepted contaminants
Work	Number of boxes opened if sample ends at detection Number of boxes opened if sample is fully inspected Number of items inspected if sample ends at detection Number of items inspected if sample is fully inspected Total number of contaminants found in samples if sample ends at detection Total number of contaminants found in samples if sample is fully inspected

3. Inspection Simulation Une Cases

To demonstrate possible applications of PoPS Border, we created three example use cases that incorporated data from imported cut flower consignment inspections conducted by the USDA APHIS Agriculture Quarantine Inspection Monitoring (AQIM) program from January

2011 - October 2020. Since these inspections used statistically valid sampling approaches, we can use the inspection outcomes to calibrate PoPS Border and do experiments that would otherwise not be possible. We used the AQIM inspection data with the simulation to estimate the consignment contamination rates and quantify inspection outcomes when deviations from inspection protocols or variations in consignment characteristics occur.

3.1 Estimate Contamination Rates from High Quality Inspection Data

Estimating contamination rates from inspection data is often impossible due to inconsistent, nonstatistical sampling. However, when inspection data obtained with known, statistically valid inspection methods are available, PoPS Border can be used to estimate consignment contamination rates by recreating the inspections and calibrating the contamination configuration to achieve similar inspection outcomes. The AQIM inspection data (8,051 records) included information about the inspected consignment size, inspection unit, sample size, sample selection method, and the inspection outcomes (pest detected or not). The data were filtered to include inspections that used a box inspection unit and sample sizes specified by the hypergeometric distribution, leaving 3,313 records. The remaining records included consignments with sizes ranging from 8 to 21,364 boxes, with an average size of 304 boxes. However, the number of stems contained in each box was not recorded and we assumed 200 stems per box for the use cases.

Using PoPS Border, we estimated a contamination rate probability distribution for the inspected cut flower consignments. First, we estimated the mean contamination rate of the AQIM inspected cut flower consignments with the maximum likelihood estimation method described by Chen et al. (2018) and Trouvé and Robinson (2020). A generalized linear model with a Bernoulli error term and a complementary $\log - \log$ link function was fitted to the vector of binary consignment inspection data (0: passed inspection, 1: failed inspection due to presence of pest). An offset of $\log$(sample size) was used to account for the number of inspected units per inspection. The mean contamination rate was calculated using the model intercept in the inverse complementary \(log\)-\(log\) function as:

$$r = 1 - e^{-e^{\lambda}}$$

where r is the mean contamination rate and \(\lambda\) is the model intercept.

We then computed shape parameters of a beta probability distribution so that the distribution had a mean equal to the estimated mean contamination rate with an arbitrary standard deviation. We generated 3,313 consignments matching the AQIM shipment sizes and contaminated them using the parameterized contamination rate distribution, the item contamination unit, and a clustered random arrangement with 40 items per cluster. We then iteratively adjusted the standard deviation of the contamination rate distribution and reran the simulation until the simulated inspections resulted in approximately the same proportion of failed consignments as the AQIM inspections, which had a consignment failure rate of 0.049 (Figure 7). See Table 2 for the definition of consignment failure rate and other important related terms used in this study. The calibrated contamination rate distribution provides an estimation of the contamination levels of the AQIM inspected consignments.

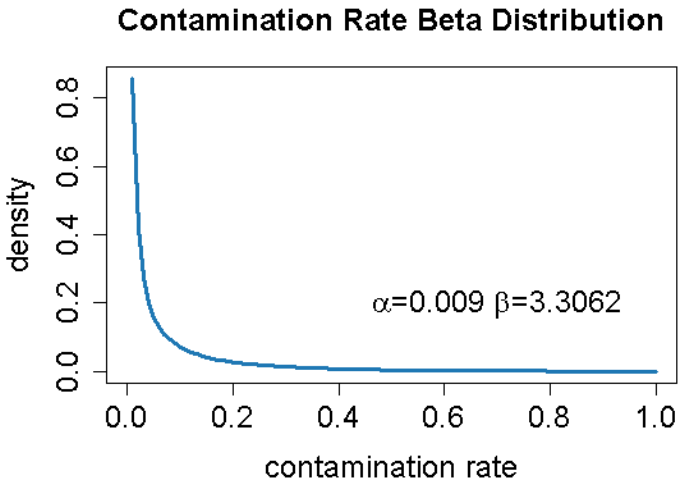


Figure 7. Beta probability distribution of contamination rates estimated to match the 0.049 consignment failure rate of the AQIM inspection data. The mean contamination rate is 0.0027 and standard deviation is 0.0282.

Table 2. Glossary of important terms.

consignment failure rate	Proportion of all consignments intercepted by inspections. This quantity can be determined in an operational setting without knowledge of the overall number of contaminated consignments.
interception rate	Proportion of contaminated consignments intercepted by inspections. This quantity can only be determined in an experimental or simulated setting with knowledge of the overall number of contaminated consignments.
noncompliance rate	Proportion of all consignments with contamination. This quantity can only be determined in an experimental or simulated setting with knowledge of the overall number of contaminated consignments.
missed contaminants	Count of total contaminated items in consignments missed by the inspections. This quantity is an estimate of the number of released contaminants.
average missed contamination rate	Average contamination rate of contaminated consignments not detected by inspections. This quantity gives an estimate of the average quality of goods being released.

3.2 Measure the Effect of Deviations in Inspection Protocols

PoPS Border is also useful for measuring trade-offs in effectiveness and work required for different inspection approaches. We ran inspection sampling scenarios with a fixed contamination configuration to observe the relative impacts of changes in inspection protocols. We generated 3,313 consignments to match the shipment sizes in the AQIM records with the contamination configuration estimated in Section 3.1 (beta contamination rate distribution, item contamination unit, clustered random contaminant arrangement, 40 items per cluster). We applied twenty-two inspection approaches to the consignments and the outcomes were averaged over one hundred simulation runs (Table 3). The box inspection unit with hypergeometric sample sizes achieved the highest interception rates, but also required the most items to be inspected. The item unit inspections with a 2% sample size also performed well but required many boxes to be opened and items inspected. Overall, the item unit inspections with random selection performed well; however, ensuring random selection of individual items packaged within boxes or bags would likely be operationally impractical. The cluster selection method may provide a feasible compromise by allowing inspectors to randomly or systematically select boxes from which to sample a cluster of items. There was no difference in interception rates between convenience and random selection when using the box inspection unit. Since contaminant clusters are placed throughout the consignments randomly, the likelihood of finding the contaminants is not dependent on where a sample is selected within a consignment. In general,

there was an inverse relationship between interception rate and average missed contamination rate. Assessing metrics of inspection effectiveness and efficiency, as shown in the four plots in Figure 8, provides a way to evaluate the relative trade-offs of different inspection approaches.

Table 3. Inspection scenario configurations and inspection outcomes.

Sample Size Method	Selection Method	Interception Rate	Avg. Missed Contamination Rate	Missed Contaminants	Boxes Opened	Items Inspected
box, hypergeometric, 0.05 detection	random	0.81	0.0008	4,927 (0.9%)	28	5,731
	convenience	0.82	0.0008	5,611 (1.0%)	28	5,731
box, hypergeometric, 0.1 detection	random	0.71	0.0015	13,760 (2.6%)	18	3,702
	convenience	0.72	0.0015	12,341 (2.4%)	18	3,702
box, proportion 0.02	random	0.13	0.0261	32,648 (6.2%)	5	1,198
	convenience	0.13	0.0259	32,268 (6.1%)	5	1,198
item, hypergeometric, 0.05 detection	random	0.38	0.0033	29,903 (5.7%)	28	58
	convenience	0.07	0.0252	336,590 (64.0%)	1	58
	interval cluster	0.11	0.0216	257,325 (49.0%)	2	58
	random cluster	0.11	0.0212	303,364 (57.7%)	2	58
item, hypergeometric, 0.1 detection	random	0.30	0.0056	60,167 (11.4%)	18	28
	convenience	0.04	0.0302	412,760 (78.5%)	1	28
	interval cluster	0.08	0.0250	334,327 (63.6%)	1	28
	random cluster	0.07	0.0256	365,245 (69.4%)	1	28
item, proportion, 0.02	random	0.52	0.0026	1,343 (0.3%)	298	1,215
	convenience	0.16	0.0200	27,808 (5.3%)	6	1,215
	interval cluster	0.22	0.0155	17,738 (3.4%)	61	1,215
	random cluster	0.22	0.0157	17,807 (3.4%)	61	1,215

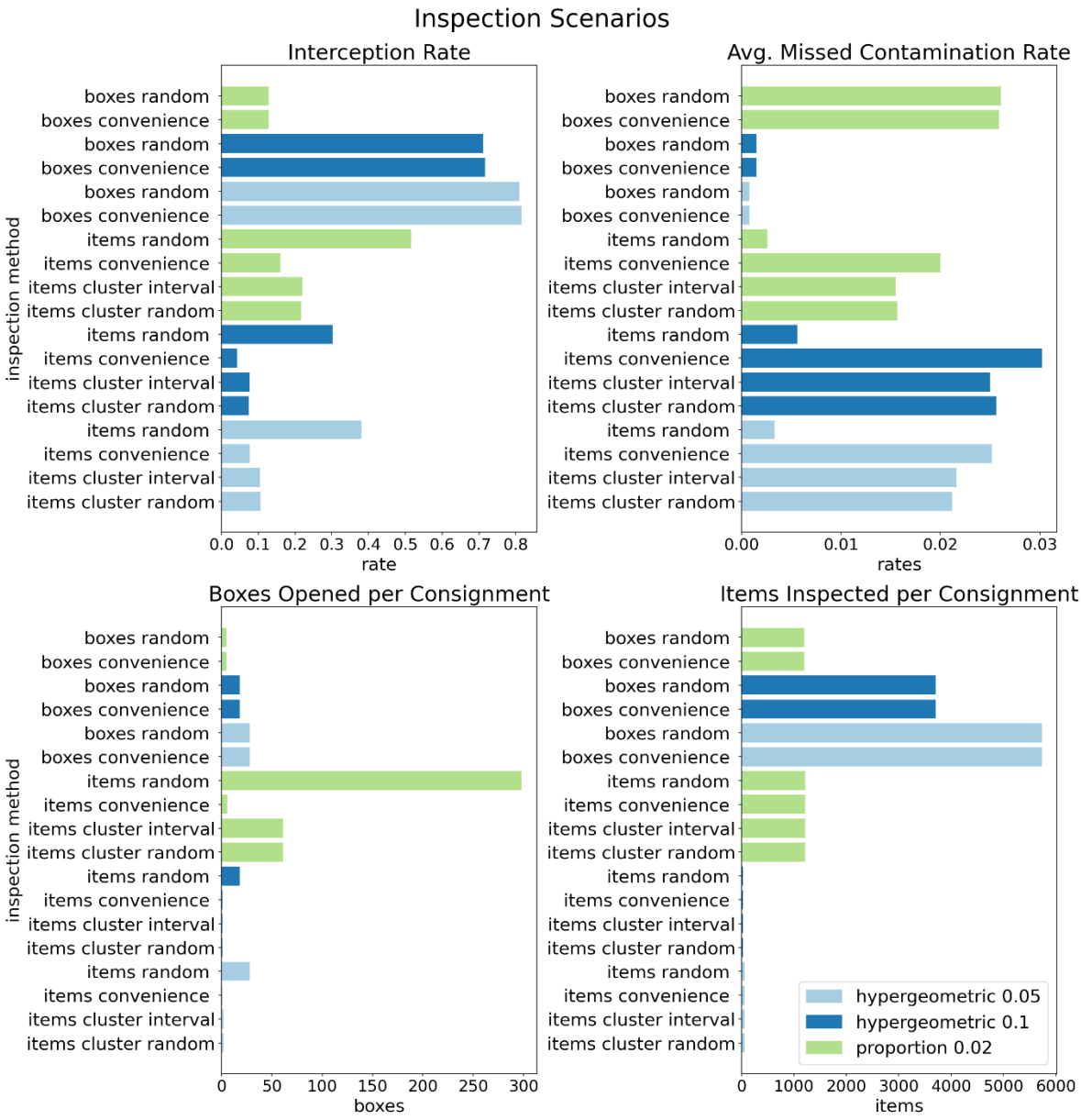


Figure 8. Outcomes of eighteen inspection protocols applied to 3,313 consignments generated to match the AQIM consignments. The outcomes are averaged over one hundred simulation runs. The scenarios used a contamination rate distribution with mean rate of 0.0027 and standard deviation of 0.0282, which resulted in a noncompliance rate of 0.07. The upper plots show the inspection effectiveness, and the lower plots show the work required.

3.3 Measure the Effect of Changes in Consignment Characteristics

PoPS Border is also useful for understanding how inspection outcomes differ with variations in consignments. We applied a fixed inspection protocol (box inspection unit,

hypergeometric sample with 0.1 detection level and 0.95 confidence level, random selection) to various consignments to represent changes in packaging, contaminant arrangements, and contamination rate variability.

3.3.1 Cargo Packaging Scenarios

Cargo using different modes of transport or going to different types of customers are often associated with different packaging standards. PoPS Border can be used to measure the workload required for inspecting shipments packaged in different ways and test alternative strategies to balance interception rates with available resources. We ran PoPS Border with three scenarios representing approximately 10,000,000 items shipped via maritime pathways with large consignments and boxes, air pathways with mid-sized consignments and boxes, and direct-to-consumer with very small consignments and boxes (Table 4). We applied the fixed inspection protocol for each scenario and averaged the outcomes over one hundred stochastic runs. Using the box inspection unit and hypergeometric sample for direct-to-consumer cargo containing very few boxes resulted in about 60% of all items and boxes being inspected and, therefore, very high inspection effectiveness (Figure 9). The maritime cargo required inspections of approximately 18,000 items per consignment on average, but since there were fewer consignments overall, only 20% of the total items were inspected compared to nearly 40% of items in the air cargo. The increase in inspection effort for air cargo resulted in a 9% increase in interception rates, but the actual percentage of missed contaminants was higher than in the marine cargo. The inspection protocol only missed consignments with very low contamination rates and was, therefore, very effective for marine cargo. PoPS Border can be used in this way to evaluate the suitability of inspection protocols for various cargo packaging scenarios.

Table 4. Cargo packaging scenario configurations and inspection outcomes.

Cargo Scenarios	Consignments Simulated	Boxes per Consignment (min - max)	Items per Box	Interception Rate	Avg. Missed Contamination Rate	Missed Contaminants
Maritime	110	100 - 160	700	0.57	0.0000	148 (0.6%)
Air	833	20 - 100	200	0.66	0.0020	420 (1.6%)
Direct-to-consumer	4,000	1 - 50	100	0.82	0.0040	563 (2.1%)

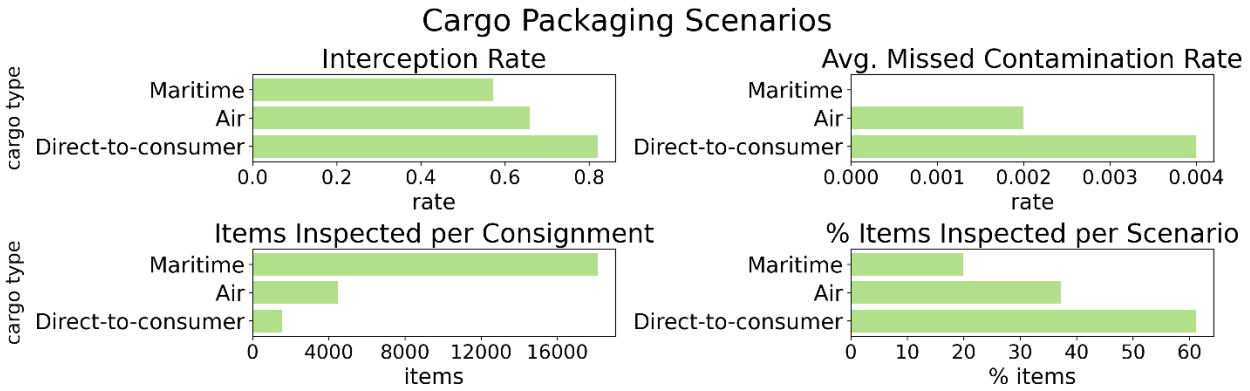


Figure 9. Inspection outcomes for 10,000,000 items packaged using three cargo scenarios. The maritime scenario represents large consignments (100 - 160 boxes) with 700 items per box. The air scenario uses mid-sized consignments (20 - 100 boxes) with 200 items per box. The direct-to-consumer scenario represents very small consignments (1 - 50 boxes) with 100 items per box.

3.3.2 Contaminant Arrangement Scenarios

The spatial distribution of contaminants within consignments influences how easily they are detected by inspections. However, information about contaminant arrangement is very difficult to obtain. We can use PoPS Border to understand the implications of potential contaminant arrangement scenarios for various inspection approaches. We generated 3,313 consignments to match the shipment sizes in the AQIM records and applied four scenarios of contaminant arrangements (Table 5). We ran the simulation with the fixed inspection protocol and averaged the outcomes over one hundred stochastic runs. Overall, interception rates decreased and average missed contamination rates increased as contaminants became more clustered (Figure 10). Larger cluster sizes resulted in much higher average missed contamination rates and missed contaminants than the small clusters, despite the overall interception rates being only slightly lower. It is important to note that when the box contamination unit is used in the simulation, all items within boxes are contaminated until the contamination rate is achieved. This means that contaminants will always be highly clustered with the box contamination unit even when boxes are selected randomly for contamination. The lower interception rates with the box contamination unit reflect this high contaminant clustering.

Table 5. Contaminant arrangement scenario configurations and inspection outcomes.

Contamination Unit	Contaminant Arrangement Scenario	Units per Cluster	Interception Rate	Avg. Missed Contamination Rate	Missed Contaminants
box	random		0.64	0.0056	52,643 (9.6%)
item	random		0.89	0.0001	279 (0.1%)
	small clusters (clustered random)	25	0.74	0.0010	7,947 (1.4%)
	large clusters (clustered random)	400	0.71	0.0048	49,996 (10.8%)

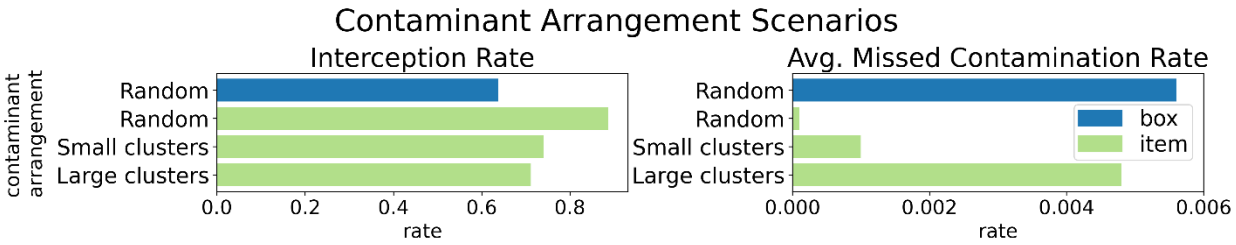


Figure 10. Inspection outcomes for scenarios comparing contaminant arrangements. The scenarios use combinations of item or box contamination unit and random or clustered arrangements.

3.3.3. Contamination Rate Variability Scenarios

As shown in Section 3.1, it may be possible to estimate the mean contamination rate from statistically robust inspection data, but the rate variance may be unknown. We used PoPS Border to compare inspection outcomes of three potential contamination rate variability scenarios. We generated 3,313 consignments matching the AQIM records and applied contamination rate distributions with a mean of 0.0027 and three different standard deviations (Table 6). We ran the simulation with the fixed inspection protocol and averaged the outcomes over one hundred stochastic runs. The interception rates were highest for the high variability scenarios due to higher likelihood of larger contamination rates, which are easier to detect by inspection (Figure 11). However, the average missed contamination rate is the highest for the high variability box unit scenario due to increased chances of missing consignments with high contamination rates and high contaminant clustering. The lower interception rates for the box contamination unit overall also reflect higher contaminant clustering. This use case

demonstrates that when the mean contamination rate is very low, higher variability in rates will result in better inspection outcomes. It also highlights that the inspection simulation results are very sensitive to the contamination unit used.

Table 6. Contamination rate variability scenario configurations and inspection outcomes.

Contamination Unit	Contamination Variability Scenario	Contamination Rate Mean (Std. Dev.)	Interception Rate	Avg. Missed Contamination Rate	Missed Contaminants
box	High	0.0027 (0.05)	0.76	0.0063	14,808 (6.3%)
	Mid	0.0027 (0.03)	0.73	0.0061	20,343 (6.3%)
	Low	0.0027 (0.01)	0.67	0.0043	124,928 (43.6%)
item	High	0.0027 (0.05)	0.82	0.0015	1,466 (0.5%)
	Mid	0.0027 (0.03)	0.79	0.0016	4,780 (1.6%)
	Low	0.0027 (0.01)	0.73	0.0016	36,167 (12.8%)

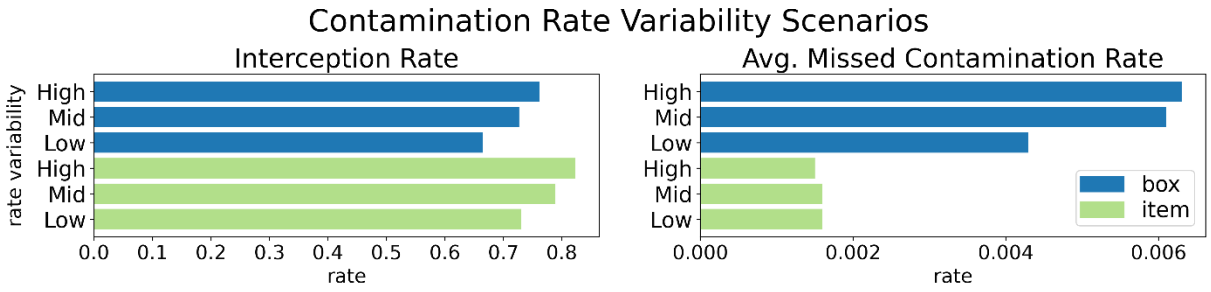


Figure 11. Comparison of inspection outcomes for contamination rate variability scenarios. The scenarios use combinations of item or box contamination unit and contamination rate distributions with low, mid, and high standard deviations.

4 Discussion

We have presented PoPS Border, an open-source tool for evaluating inspection strategies for detecting rare occurrences, and three example use cases for the tool. The three use cases presented here are for cut flower pest inspections; however, PoPS Border can be used for many types of inspections, including industrial quality control sampling, law enforcement, or border control inspections. PoPS Border can be applied with minimal inputs or can be further enhanced with additional data to increase sophistication of the consignments and inspections simulated. For example, results from pathway risk analysis models could be used to inform contamination rates that vary by origin, season, mode of transport, and commodity. Cargo configurations, such as number of items per box, could be made more realistic by setting the related parameters to vary by commodity type or mode of transport. The amount of work required for inspections could be improved by incorporating accurate estimates of the man hours required to inspect various types of commodities. This work represents the first steps towards building an inspection protocol evaluation tool that can be iteratively improved to reflect the realities of border inspection operations.

PoPS Border can be used to evaluate risk-based sampling methods or consignment release programs under various conditions. For example, the example use cases show simulated outcomes for hypergeometric sampling approaches, which are common in many risk-based sampling programs. As shown in Figure 8, the outcomes are highly sensitive to the inspection unit and selection approach used, with clear trade-offs in high interception rate and the amount of work required. PoPS Border can also enable development of consignment release programs, which release low risk consignments without inspection according to an inspection schedule. Although the simulation does not currently include a realistic release program, the basic mechanics have been tested with a simple prototype based on the National Cut Flower Release Program (NCFRP). In the prototype release program, consignments that fit specified origin and commodity criteria are released without inspection unless the simulated arrival date falls on a user-defined inspection calendar (Figure 12). Further development of the tool's release program features could provide valuable information on program vulnerabilities and a way to test proposed program updates.

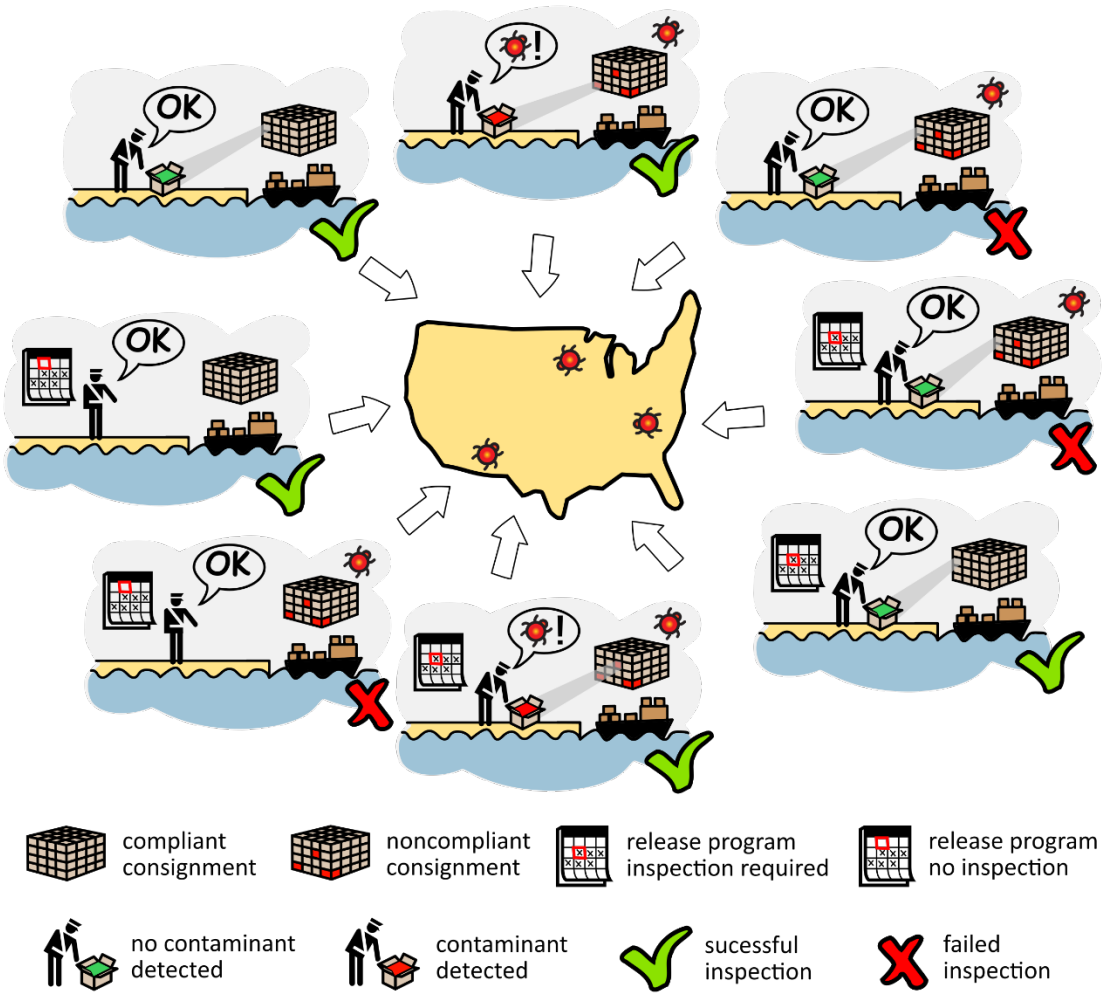


Figure 12. Conceptual diagram of the prototype consignment release program simulation. Each scenario illustrates a possible inspection outcome under the release program.

As shown in the use cases, PoPS Border can be used to extend statistically robust, high quality inspection data to simulate incoming consignment contamination levels and estimate slippage rates under specific inspection protocols. If applied using actual shipment volumes coming through ports of entry, PoPS Border could provide an estimate of regional propagule pressure and inform pest surveillance efforts. Future enhancements of the tool should include an automated method for calibrating contamination rate distributions from inspection data. However, since inspection outcomes are influenced by how contaminants are clustered within consignments, better information on contaminant arrangement is needed to precisely estimate contamination rates. There may be opportunities to gain insights on contaminant arrangement

by using PoPS Border with contamination rate data. As better information on contaminant arrangement becomes available, the methods for generating clusters in the simulation can be improved. For example, future implementations should use dynamic cluster sizes based on contamination rates.

The inspection approaches currently implemented in PoPS Border were chosen in collaboration with USDA APHIS analysts and include primary sampling methods used by inspectors. However, additional approaches could be implemented in future versions of the tool. In the current tool, if a contaminated item is selected for inspection, the contaminant will always be detected. Future implementations should include a user-defined efficacy rate so that some inspections do not detect contaminated items. This variable inspection efficacy could be used to simulate different inspection techniques, technologies, or conditions. As the sophistication of the tool is increased, more complex scenarios with commingled commodities and multiple types of contaminants can be simulated. For example, contaminants may be characterized as quarantine significant or not, so that some detected contaminants are considered not actionable and released without treatment. Further enhancements of the simulated inspections could provide opportunities to evaluate inspection protocols for more diverse types of contamination with variations in risk and actions taken.

5 Conclusion

Preventing non-native pests and pathogens from crossing borders is very challenging due to the massive volume of goods passing through ports of entry daily. Phytosanitary agencies need better tools to understand how trade-offs in inspection efficiency and effectiveness may change with shifts in operations and commerce. By simulating consignments and inspections, PoPS Border provides a way to measure contaminant slippage, quantify workload, and perform computational experiments to refine inspection protocols before deploying them in the field. Further development of a web-based analytics dashboard for running PoPS Border and visualizing results would provide a valuable decision support tool to help agencies create more agile, risk-based inspection strategies.

Acknowledgements

This study was conducted as part of graduate study and made possible, in part, by

cooperative agreements from U.S. Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Plant Protection and Quarantine (PPQ), Science and Technology. This paper does not necessarily reflect the views of APHIS. We would like to thank E. Hain, M. Colunga, S. Hong, and A. Neeley for guidance on the study, R. Hallberg for providing editorial suggestions, and D. Gaydos, S. Jones, and A. Petrasova for contributing to the design and creation of Figures 1 and 12.

REFERENCES

Caley, P., Ingram, R., & De Barro, P. (2015). Entry of exotic insects into Australia: Does border interception count match incursion risk? *Biological Invasions*, *17*(4), 1087–1094. https://doi.org/10.1007/s10530-014-0777-z

Chen, C., Epanchin-Niell, R. S., & Haight, R. G. (2018). Optimal inspection of imports to prevent invasive pest introduction. *Risk Analysis*, *38*(3), 603–619. https://doi.org/10.1111/risa.12880

Epanchin-Niell, R., McAusland, C., Liebhold, A., Mwebaze, P., & Springborn, M. R. (2021). Biological invasions and international trade: Managing a moving target. *Review of Environmental Economics and Policy*, 000–000. https://doi.org/10.1086/713025

Fosgate, G. T. (2009). Practical Sample Size Calculations for Surveillance and Diagnostic Investigations. *Journal of Veterinary Diagnostic Investigation*, *21*(1), 3–14. https://doi.org/10.1177/104063870902100102

International Plant Protection Convention. (2016). *Methodologies for sampling of consignments* (International Standards for Phytosanitary Measures No. 31). https://www.ippc.int/en/publications/588/

Jones, C., Jones, S., Petrasova, A., Petras, V., Gaydos, D., Skrip, M., Takeuchi, Y., Bigsby, K., & Meentemeyer, R. (2021). Iteratively forecasting invasions with PoPS and a little help from our friends. *Frontiers In Ecology And The Environment*. https://doi.org/doi:10.1002/fee.2357

Kim, B., Hong, S. C., Egger, D., Katsar, C. S., & Griffin, R. L. (2019). Predictive modeling and categorizing likelihoods of quarantine pest introduction of imported propagative commodities from different countries. *Risk Analysis*, *39*(6), 1382–1396. https://doi.org/10.1111/risa.13252

Lockwood, J. L., Cassey, P., & Blackburn, T. M. (2009). The more you introduce the more you get: The role of colonization pressure and propagule pressure in invasion ecology. *Diversity and Distributions*, *15*(5), 904–910. https://doi.org/10.1111/j.1472-4642.2009.00594.x

Montgomery, K., Walden-Schreiner, C., Chris Jones, Ben Seliger, Thom Worm, Ariel Saffer, Laura Tateosian, Makiko Shukunobe, Sunil Kumar, & Ross Meentemeyer. (2021). *Forecasting global spread of invasive pests and pathogens through international trade [Manuscript submitted for publication]*.

Robinson, A., Burgman, M. A., & Cannon, R. (2011). Allocating surveillance resources to reduce ecological invasions: Maximizing detections and information about the threat. *Ecological Applications*, *21*(4), 1410–1417. https://doi.org/10.1890/10-0195.1

Robinson, A. P., Chisholm, M., Mudford, R., & Maillardet, R. (2015). Ad hoc solutions to estimating pathway non-compliance rates using imperfect and incomplete information. In

F. Jarrad, S. Low-Choy, & K. Mengersen (Eds.), *Biosecurity surveillance: Quantitative approaches* (pp. 167–180). CABI. https://doi.org/10.1079/9781780643595.0167

Robinson, A. P., Walshe, T., Burgman, M. A., & Nunn, M. (2017). *Invasive Species: Risk Assessment and Management*. Cambridge University Press.

Rossiter, A., & Hester, S. M. (2017). Designing biosecurity inspection regimes to account for stakeholder incentives: An inspection game approach. *Economic Record*, *93*(301), 277–301. https://doi.org/10.1111/1475-4932.12315

Saccaggi, D. L., Karsten, M., Robertson, M. P., Kumschick, S., Somers, M. J., Wilson, J. R. U., & Terblanche, J. S. (2016). Methods and approaches for the management of arthropod border incursions. *Biological Invasions*, *18*(4), 1057–1075. https://doi.org/10.1007/s10530-016-1085-6

Surkov, I. V., Oude Lansink, A. G. J. M., van Kooten, O., & van der Werf, W. (2008). A model of optimal import phytosanitary inspection under capacity constraint. *Agricultural Economics*, *38*(3), 363–373. https://doi.org/10.1111/j.1574-0862.2008.00306.x

Trouvé, R., & Robinson, A. P. (2020). Estimating consignment-level infestation rates from the proportion of consignment that failed border inspection: Possibilities and limitations in the presence of overdispersed data. *Risk Analysis*. https://doi.org/10.1111/risa.13592

U.S. Customs and Border Protection. (2021). *Trade Statistics*. https://www.cbp.gov/newsroom/stats/trade

USDA APHIS PPQ. (2011). *Agricultural Quarantine Inspection Monitoring (AQIM) Handbook*. USDA APHIS PPQ.

CHAPTER 4: MEASURES OF CANOPY STRUCTURE FROM LOW-COST UAS FOR MONITORING CROP NUTRIENT STATUS

Reprint

Montgomery, K., Henry, J. B., Vann, M. C., Whipker, B. E., Huseth, A. S., & Mitasova, H. (2020). Measures of Canopy Structure from Low-Cost UAS for Monitoring Crop Nutrient Status. *Drones*, *4*(3), 36. https://doi.org/10.3390/drones4030036

Author Contributions (using standard formatting from journal)

Conceptualization, K.M. and H.M.; Methodology, K.M., H.M., J.B.H., B.W., and M.C.V.; Validation, K.M. and H.M.; Formal analysis, K.M.; Investigation, K.M. and J.B.H.; Resources, K.M., J.B.H., B.W., and M.C.V.; Writing—original draft preparation, K.M.; Writing—review and editing, K.M., H.M., J.B.H., M.C.V., and A.S.H.; visualization, K.M.; funding acquisition, K.M, J.B.H., B.W., and M.C.V. All authors have read and agreed to the published version of the manuscript.

Data Availability

The canopy surface models, orthoimages, and water depth model are available upon request. The code used for the canopy surface model analysis, spectral and water flow index computations, and statistical modeling are available at https://github.com/kellynm/canopy_structure.

Article

Measures of Canopy Structure from Low-Cost UAS for Monitoring Crop Nutrient Status

Kellyn Montgomery [1,*], Josh B. Henry [2,3,†], Matthew C. Vann [2], Brian E. Whipker [3], Anders S. Huseth [1,4] and Helena Mitasova [1,5]

[1] Center for Geospatial Analytics, North Carolina State University, Raleigh, NC 27695, USA; ashuseth@ncsu.edu (A.S.H.); hmitaso@ncsu.edu (H.M.)
[2] Crop and Soil Sciences, North Carolina State University, Raleigh, NC 27695, USA; jbhenry2@ncsu.edu (J.B.H.); matthew_vann@ncsu.edu (M.C.V.)
[3] Horticultural Science, North Carolina State University, Raleigh, NC 27695, USA; bwhipker@ncsu.edu
[4] Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA
[5] Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695, USA
* Correspondence: kpmontgo@ncsu.edu
† Current address: The Scotts Miracle-Gro Company, Marysville, OH 43040, USA.

Received: 6 June 2020; Accepted: 17 July 2020; Published: 22 July 2020

Abstract: Deriving crop information from remotely sensed data is an important strategy for precision agriculture. Small unmanned aerial systems (UAS) have emerged in recent years as a versatile remote sensing tool that can provide precisely-timed, fine-grained data for informing management responses to intra-field crop variability (e.g., nutrient status and pest damage). UAS sensors with high spectral resolution used to compute informative vegetation indices, however, are practically limited by high cost and data dimensionality. This research extends spectral analysis for remote crop monitoring to investigate the relationship between crop health and 3D canopy structure using low-cost UAS equipped with consumer-grade RGB cameras. We used flue-cured tobacco as a case study due to its known sensitivity to fertility variation and nutrient-specific symptomology. Fertilizer treatments were applied to induce plant health variability in a 0.5 ha field of flue-cured tobacco. Multi-view stereo images from three UAS surveys collected during crop development were processed into orthoimages used to compute a visible band spectral index and photogrammetric point clouds using Structure from Motion (SfM). Plant structural metrics were then computed from detailed high resolution canopy surface models (0.05 m resolution) interpolated from the photogrammetric point clouds. The UAS surveys were complimented by nutrient status measurements obtained from plant tissues. The relationships between foliar nitrogen (N), phosphorus (P), potassium (K), and boron (B) concentrations and the UAS-derived metrics were assessed using multiple linear regression. Symptoms of N and K deficiencies were well captured and differentiated by the structural metrics. The strongest relationship observed was between canopy shape and N foliar concentration (adj. $r^2 = 0.59$, increasing to adj. $r^2 = 0.81$ when combined with the spectral index). B foliar concentration was consistently better predicted by canopy structure with a maximum adj. $r^2 = 0.41$ observed at the latest growth stage surveyed. Overall, combining information about canopy structure and spectral reflectance increased model fit for all measured nutrients compared to spectral alone. These results suggest that an important relationship exists between relative canopy shape and crop health that can be leveraged to improve the usefulness of low cost UAS for precision agriculture.

Keywords: unmanned aerial systems; precision agriculture; canopy structure; crop surface model; visible band index; flue-cured tobacco

1. Introduction

One of the primary goals of precision agriculture is identifying intra-field variability to enable a management response to plant stress that will maximize yield. Currently, farmers monitor crop health in high-value crops through labor-intensive scouting, destructive field sampling, and costly laboratory assays. Remote sensing, however, is increasingly being used as an alternative tool for crop monitoring that more accurately accounts for spatial variations in crop stress across a field [1]. Small unmanned aerial systems (UAS) have emerged in recent years as a versatile remote sensing tool used by scientists and agricultural producers for collecting data at very high spatial and temporal resolutions [2,3]. UAS can provide precisely-timed, fine-grained data for informing spatially variable management responses for maximizing crop productivity while minimizing natural resource degradation.

One of the most promising application of remote sensing in agriculture is identifying symptoms from biotic and abiotic plant stressors, such as nutrient deficiencies, pest pressure, and harsh environmental conditions. In particular, monitoring plant nutrition with remote sensing has been successfully demonstrated by numerous studies [4–7]. Symptoms of deficiencies of essential macronutrients, like nitrogen (N), phosphorus (P), and potassium (K), can often be differentiated visually, making remote identification possible. These nutrients perform crucial roles in plant development and are needed in relatively high concentrations at specific plant growth stages [8]. Nutrient toxicity is also important to monitor for many crops, such as boron (B) in flue-cured tobacco, which has negative effects on leaf quality due to increased nicotine and other alkaloids [9–11]. Remote monitoring of crop nutrient status with UAS can provide on-demand, fine-grained spatial information to improve crop productivity and minimize over-application of fertilizers if integrated into existing precision agriculture technology used on-board commercial tractors and implements.

Nutrient deficiencies often cause visible changes in foliar coloration, appearing yellow (chlorotic), white (bleached), brown (necrotic), red, or black [12]. These well-documented deficiencies can be measured by sensors as changes in spectral reflectance and enable noninvasive diagnostics of plant nutrient disorders. Vegetation indices, like Normalized Difference Vegetation Index (NDVI), measured by broadband multispectral sensors have been shown to strongly correlate with crop health and are widely used in agriculture and throughout the literature [13–16]. Hyperspectral sensors with very fine spectral resolution also have the potential to measure unique spectral signatures for nutrient disorders [7,17,18]. However, practical application of UAS hyperspectral imagery to diagnose nutrient disorders in production agriculture remains limited by high costs, high data dimensionality, and the need for advanced data analysis skills. UAS with optical cameras, in contrast, are inexpensive and often supported by cloud-based data storage and processing, making them easier to use by farmers without remote sensing data processing skills [2]. Optical cameras provide visible band imagery that can be used to calculate vegetation indices, like Visible Atmospheric Resistant Index (VARI) and Triangular Greenness Index (TGI) [19]. These RGB indices have been shown to correlate well with leaf area index and chlorophyll content [19,20]. Chlorophyll, which reflects highly in the green spectrum, is a good indicator of many plant stressors with symptoms of foliar chlorosis or yellowing [21]. Unfortunately, the small spectral range of optical cameras make RGB vegetation indices limited in usefulness.

In addition to natural color images, 3D point clouds and highly detailed surface models that provide valuable information about vegetation height and heterogeneity can be computed from RGB images using the photogrammetric technique Structure from Motion (SfM) [22]. The high-resolution representation of crop canopy surfaces presents an opportunity to explore the relationship between canopy structure and important crop characteristics, like stress and productivity. For example, Hunt and Rondon [23] used very high resolution canopy surface models to identify foliar pest damage in potatoes. Other studies have shown success in correlating yield and biomass with structural metrics generally used to characterize geomorphological surfaces [24–26]. Parker and Russ [24] measured spatial covariance of LIDAR-derived outer forest canopy through rugosity (standard deviation of canopy height), hypsographs, fractal dimension, integral scale, and elevation relief ratio to characterize conifer forest canopy structural complexity. A similar technique was used by Li et al. [26] and Li, Niu,

Chen and Li [25] with photogrammetric point clouds to estimate biomass and leaf area index (LAI) in maize. Studies have also used simple crop height and relative canopy cover to estimate barley and cotton yields [27,28]. Research on methods for characterizing crop canopy structure and the potential for deriving valuable crop information from SfM canopy surface models has been limited, however, and should be further explored.

A common goal within precision agriculture is improving crop yield and farm profitability through the use of technology and data analytics. The current costs and barriers to practical use associated with high spectral resolution imagery and active laser scanning increase the importance of improving the predictive power of inexpensive UAS-acquired data. In this study, we used a small UAS platform equipped with a consumer-grade RGB camera to investigate the use of canopy structural metrics combined with spectral reflectance to monitor the onset of nutrient disorder symptomology in a case study of flue-cured tobacco. This approach goes beyond traditional spectral analysis for remote crop monitoring to explore the relationship between crop health and patterns in canopy height spatial variance measured from high resolution, photogrammetric canopy surface models. Our results suggest that an important relationship exists between relative canopy shape and crop nutrition that can be leveraged to improve the usefulness of low cost UAS for precision agriculture.

2. Materials and Methods

2.1. Study Site

A field experiment took place in Wilson, NC (35°47′42″ N lat, 77°56′52″ W long) during the 2019 summer growing season to investigate the use of UAS imagery for detecting nutrient disorders in flue-cured tobacco. The study site was located on a commercial farm and was planted and maintained by the farm managers. The site had a mean elevation of 47.5 m and relatively uniform, low gradient topography with mean slope of 1.46 degrees tilted eastward. Soils across the study site were relatively homogeneous sandy loam. Six soil samples evenly distributed across the field were collected and analyzed for nutrient content, pH, humic matter, and cation exchange capacity (CEC) prior to the experiment (NCDA & CS Agronomic Services Division; Raleigh, NC). The soil properties were similar across the samples with an average of 5.68 pH, 3.12 $cmol_c$/kg CEC, and 0.32% humic matter. Weather conditions during the experiment were typical of the area's humid subtropical climate, with an average daily air temperature of 25.6 °C and average daily precipitation of 2.7 mm. During the two weeks prior to the first data collection, the study site received 97 mm of rain. Total rainfall during the remaining four weeks of the experiment was 30 mm.

Flue-cured tobacco seedlings were transplanted on 16 April 2019 into a 0.5 ha field in thirty-six 10 m × 16 m blocks. Each block consisted of four rows with 1.16 m row spacing and 0.58 m in-row spacing between plants. To induce variability and extremes in nutrient availability within the field, 6N-6P-18K fertilizer (Tobacco Super Rainbow; Agrium U.S., Inc., Denver, CO, USA) was applied to the blocks in early May (3–4 true leaves) at four different rates using block randomization. Nine blocks each received 0, 50, 100, or 200% of the recommended fertilizer rate for flue-cured tobacco production as published by Fisher [29]. To induce nutrient disorders, three small strips of twelve plants (approximately 1.7 m × 4.7 m) within each block were given a fertilizer treatment deficient of N, P, or K or a toxic amount of B. Within the blocks fertilized at 0 or 50% rates, the three strips were fertilized with a custom fertilizer made to replicate the 6N-6P-18K fertilizer devoid of either N, P, or K. Within the blocks fertilized at 100 or 200%, the three strips were given an excess application of B at a rate of 4.50, 9.00, or 18.1 kg/ha in addition to the 6N-6P-18K fertilizer. Figure 1 shows the study site with the block and strip fertilizer treatments. Flowering and topping began on June 15 with all plants topped by 3 July. Chemical sucker control was used to inhibit auxiliary bud growth after topping.

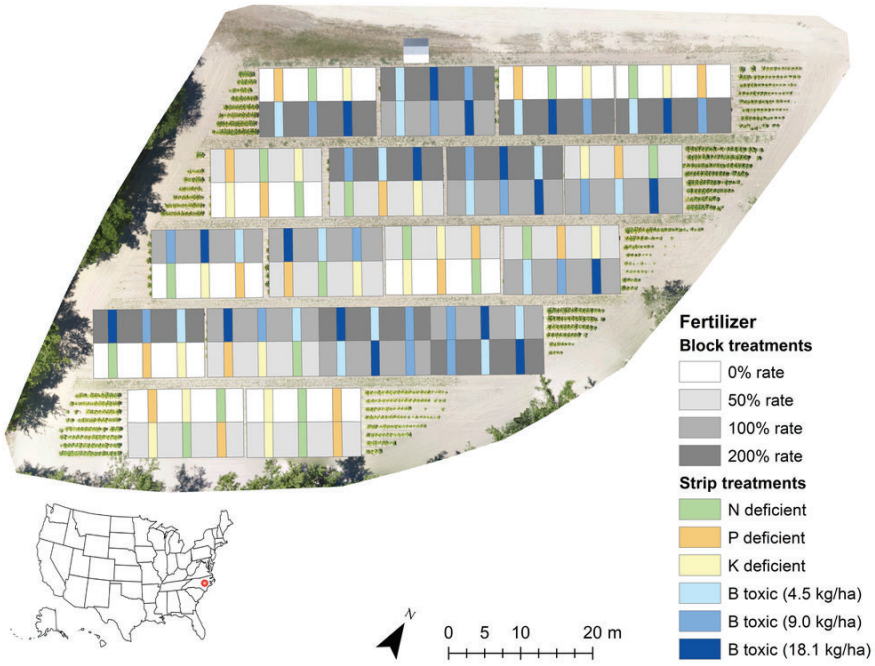


Figure 1. Field experiment site in Wilson, NC. Flue-cured tobacco was planted in thirty-six blocks each containing three strips fertilized with treatments to induce variability in nutrient availability throughout the 0.5 ha field.

2.2. Data Collection

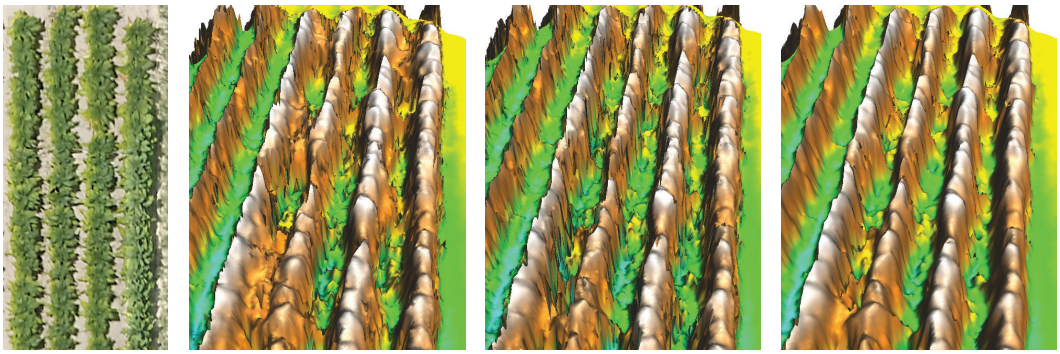
UAS surveys to collect aerial imagery of the study site were conducted on 19 June, 3 July, and 17 July starting approximately four weeks after the fertilizer treatments were applied. The UAS data were collected by PrecisionHawk (Raleigh, NC, USA) using a DJI Phantom Pro 4 equipped with a 20 mp RGB camera. At least twenty-two images were collected during each survey using a grid flight pattern at approximately 67 m altitude with a nadir camera angle. To ensure matched vegetation features across multiple images, 80% side and forward image overlap was used. Ground control points were not used as the images were initially collected to be used for visual inspection. Errors in X,Y position and elevation were corrected using orthoimagery and LIDAR data as described in Section 2.3.

Foliar tissue samples were collected from each treatment replicate at approximately the same time as the UAS flights by Dr. Josh Henry and field research assistants from the Crop and Soil Sciences and Horticulture Science Departments at NC State. Leaf tissue samples were pooled resulting in three samples per fertilizer treatment to reduce the total number of samples and lab analysis costs. The leaf tissues were dried, ground, and analyzed for elemental nutrient concentrations of N, P, K, and B by AgSource Laboratories (Lincoln, NE, USA).

2.3. Data Processing

Photogrammetry software (Agisoft Metashape, version 1.5.4.8885) was used to process the UAS images into 3D point clouds and orthoimages for the three survey dates. The multi-view stereo images were aligned by matching features in multiple images to create a sparse point cloud. The sparse point cloud was then filtered to include only points with reconstruction uncertainty less than 50 and reprojection error less than 0.35 pixels. A dense point cloud was created using images downscaled by a factor of four (high reconstruction quality setting). Point filtering was disabled to preserve gaps between plants and other heterogeneous features in the point geometry representing the plant canopy (Figure 2). The density of the resulting point clouds was approximately 900 points/m^2. The average

X,Y and Z camera location error was 2.29 m and 0.46 m, respectively. The average point reprojection error was 0.56 pixels.



(**a**) (**b**) (**c**) (**d**)

Figure 2. Comparison of surface models resulting from variations in point cloud filtering in Agisoft Metashape and interpolation smoothing parameters. (**a**) Orthoimage of tobacco plants represented in surface models shown at right. (**b**) Noisy surface with poor representation of gaps between plant rows created using mild depth filtering and no surface smoothing. The point filtering algorithm removed important bare ground points between plant rows. (**c**) Noisy surface with better representation of gaps between plant rows created using no depth filtering and no surface smoothing. (**d**) Less noisy surface with good representation of gaps between plants created using no depth filtering and smoothing splines with tension.

The point clouds were then interpolated to create smoothed digital surface models (DSM) using a regularized spline with tension method implemented in GRASS GIS 7.8 [30,31]. The tension parameter tunes the surface stiffness and the smoothing parameter controls the resulting surface's deviation from the point cloud (see the relevant equations and explanation of the parameters in [30,31]). Tension and smoothing parameter settings were iteratively adjusted to reduce surface noise and retain realistic canopy features (tension: 20, smoothing: 5) (Figure 2). A spatial resolution of 0.05 m was used for the interpolated surface model based on point density and plant feature dimensions. To correct errors in X,Y position, the resulting DSMs and UAS-derived orthoimages were georeferenced to 2017 state-wide orthoimagery (0.15 m resolution) [32].

To correct errors in elevation, the DSMs were compared to a 0.05 m resolution digital elevation model (DEM) interpolated from LIDAR bare earth points (9.25 cm vertical linear root mean square error (RMSEz)) collected in 2015 for the NC Floodplain Mapping Program [33]. Orthoimages created from the UAS data were used to identify approximately 30 bare earth points well distributed across the study area that should have elevations values that match the DEM. The DSMs were sampled at each of these bare earth points and the sample point values were interpolated using splines with tension to create smoothed surfaces representing error in the z-values for each DSM (tension: 40, smoothing: 5). The error surfaces were then subtracted from the DSMs to correct the elevation values. The corrected DSMs had bare earth elevation values within $\pm$0.05 m of the DEM. Crop surface models (CSM) with z-values normalized for topography were created by subtracting the bare earth DEM from each DSM. Note that bare earth points from UAS data could also be used to create a DEM when lidar data is not available [34].

A mask was created to exclude non-tobacco features, such as soil, rocks, and other vegetation, from computation of the spectral vegetation index and the structural metrics that were not intended to characterize canopy gaps (described in Sections 2.4 and 2.5). Features less than 0.2 m in height were considered non-tobacco. This height threshold was effective in removing features that were not of interest while retaining canopy features from the most nutrient deficient plants, which were at least 0.6 m in height at the first survey.

2.4. Canopy Surface Analysis

Using ArcMap 10.7, plot boundaries were drawn around each treatment strip with approximately 0.5 m between each polygon to reduce edge effects, resulting in 4.7 m × 1.2 m areas for each strip. Four plot polygons of equal size were drawn in each treatment block as well (Figure 3). The plot polygons were used to extract pixel values representing crop height in each treatment area and compute metrics quantifying structural characteristics of the canopy. Vertical and horizontal patterns in canopy structure were of interest for their potential to predict crop health. Metrics were chosen to quantify central tendency of crop height, canopy structural complexity, and canopy structural uniformity. Table 1 summarizes the metrics used.

Figure 3. Plot polygons corresponding to variable nutrient treatments across the study area. The plot polygons were used to extract canopy height and geometry from the canopy surface models.

Table 1. Summary of canopy metrics computed from unmanned aerial systems (UAS) data.

Metric	Canopy Characteristic Measured	Notes
Crop height mean (m)	central tendency of plant height	
Crop height std. dev. (m)	variability of plant height	
Histogram skewness	uniformity of plant height	
Histogram kurtosis	uniformity of plant height	Excess kurtosis
Rumple index (RI)	canopy surface roughness, complexity	[35]
Canopy relief ratio (CRR)		[36]
median	central tendency of local relative canopy shape	
interquartile range	variability of local relative canopy shape	
Moran's I	local similarity of plant height	[37]
Water flow index	spatial distribution of overland water flow	[38]
Triangular Greenness Index (TGI):		[19]
mean	central tendency of per pixel chlorophyll content	
standard deviation	variability of per pixel chlorophyll content	

Nutrient disorders have been shown to correlate with flue-cured tobacco size and shape in multiple prior studies [39–42]. To quantify plant size and variability in canopy height, mean and standard deviation of height were computed. Skewness and excess kurtosis of the histogram of canopy heights within each plot were used to quantify canopy uniformity and closure [43]. Highly skewed canopy heights may measure uniformity and relative plant shape within a plot (i.e., ratio of height and width of plants). High histogram kurtosis is associated with a very high peak and heavy tails compared to a normal distribution, which may indicate canopy closure, or similar heights across a plot.

Low kurtosis indicates a more uniform distribution of heights, which may be a good indicator of the presence of gaps between plants. Rumple index (RI) was used to quantify three-dimensional canopy complexity in both the horizontal and vertical directions [35,44]. RI is a measure of surface roughness, calculated as the ratio between surface area and the two-dimensional projected area on a plane below.

The metrics described thus far do not account for the spatial pattern of canopy structure and cannot, for example, distinguish between clusters of pixels with low heights and dispersed pixels with low heights. To better account for local spatial patterns in canopy structure, canopy relief ratio (CRR) was calculated for a moving window of approximately 0.25 m $\times$ 0.25 m (5 $\times$ 5 pixels) within each plot [36]. CRR measures how close the mean height is to the minimum and maximum heights within the chosen neighborhood. The neighborhood size was chosen to correspond to one quarter of a single plant to measure variability in relative shape within a plant. Several neighborhood sizes were tested to identify which resulted in metrics that explained the most variability in nutrient concentrations. CRR was calculated as

$$CRR = \frac{mean(height) - min(height)}{max(height) - min(height)} \quad (1)$$

This metric was intended to measure relative shape of the canopy. Generally, canopy peaks will have CRR values close to 1, valleys will be close 0, and slopes or flat areas will have values centered around 0.5. To summarize the results of the moving window CRR calculation, the median and interquartile range (IQR) for each plot were used.

In addition to CRR, spatial patterns in canopy height were measured using Moran's I with an approximately 0.55 m $\times$ 0.55 m (11 $\times$ 11 pixels) Queen's case weights matrix [37,45,46]. The weights matrix defines the area and direction in which plant heights are expected to be spatially correlated and was chosen to correspond to the in-row plant spacing, or the approximate area of one mature tobacco plant. These metrics were intended to quantify the similarity of local canopy structure across each plot. Plots with repeating patterns of canopy geometry due to uniform plant growth were expected to have high spatial autocorrelation, whereas plots with highly heterogeneous plant sizes and shapes were expected to have low spatial autocorrelation.

2.5. Visible Band Spectral Metrics

Triangular greenness index (TGI) was used to approximate chlorophyll content and vegetation vigor from broadband visible wavelength reflectance. TGI is based on the spectral features of chlorophyll and can be used for crop nitrogen management [19]. TGI was calculated as

$$TGI = \frac{(\lambda_{Red} - \lambda_{Blue})(\rho_{Red} - \rho_{Green}) - (\lambda_{Red} - \lambda_{Green})(\rho_{Red} - \rho_{Blue})}{2} \quad (2)$$

where λ is the center wavelength and ρ is the pixel value of the respective bands. The band center wavelengths used were 670 nm, 550 nm, and 480 nm for red, green, and blue, respectively, as proposed for broadband digital cameras by Hunt et al. [19]. For each UAS survey, TGI was computed for all pixels containing tobacco vegetation using the red, green, and blue bands and summarized per plot using mean and standard deviation.

2.6. Water Flow Index

To account for spatial variability in soil moisture and possible movement of nutrients within the soil caused by water flow, an overland water flow index was computed. The index was derived from an overland flow simulation in GRASS GIS 7.8.0 [38,47]. We simulated a steady state water flow over a 0.05 m resolution LIDAR DEM of the study area resulting in per pixel water accumulation depth [33]. The water depth values were normalized to create an index representing the spatial distribution of water flow for the site. The sum of the per pixel index values was used to summarize water flow for each plot.

2.7. Statistics

The computed canopy metrics were aggregated using the mean value for each combination of three plots used during the foliar tissue sample collection resulting in a single set of canopy metrics per lab sample. The metrics were then z-score = standardized to facilitate comparison of statistical results. The relationship between the canopy metrics and foliar concentrations of N, P, K, and B was evaluated using ordinary least squares regression [48]. To quantify multicollinearity between the regression predictors, the variance inflation factor (VIF), an index measuring the increase in regression coefficient variance caused by collinearity, was computed for each predictor [48]. Any predictor variable with a VIF greater than 4 was omitted from subsequent analysis due to the redundancy of the information provided and to reduce model overfitting [49]. Multicollinear variables removed include height standard deviation and height histogram skewness.

The canopy metrics were then used in ordinary least squares regression to predict not only nutrient concentrations measured at the time of the UAS surveys, but also all prior measurements of nutrient concentrations resulting in six sets of regression models. This was done to investigate the timing of canopy response to the nutrient application. First, the structural and spectral metrics were used in the models separately to predict nutrient concentrations. Then, the combined metrics were used to determine if model fit could be improved by using a combination of structural and spectral information. For each combination of variables, Akaike's information criteria with small sample bias adjustment (AICc) was used to select the highest quality model to reduce the risk of overfitting [50]. The adjusted coefficient of determination (adj. r^2) was used to evaluate the models.

3. Results

3.1. Foliar Nutrient Concentrations

A summary of foliar tissue nutrient concentrations and correlation coefficients for the three sample dates are shown in Tables 2 and 3. The fertilizer treatments applied across the study plots were effective in creating plot-level variability in N, K, and B response. Symptoms of N deficiency were visible early in plant development and were well-captured by each UAS survey. Plots with K deficiency were not observed until after the first UAS survey, with symptoms becoming visible approximately six weeks after the fertilizer treatments were applied. High variability in foliar B concentration was achieved throughout the study site with very high concentrations observed at each sample date. P concentrations, however, were present in the study area soils prior to fertilizer treatments at levels 2–5 times greater than optimal making it impossible to induce a deficit in the plants. P is relatively immobile in soils, which has resulted in a build up of stores in agricultural soils over time caused by fertilizer over-application. As very little variation in foliar P concentrations were observed, the regression models for this nutrient were not informative and therefore not reported here.

Table 2. Foliar nutrient concentrations.

Sample Date	Growth Stage	Nutrient	Mean	St. Dev.	Min	Max	Sufficiency Range [51]
17 June	early growth, 9 weeks post-transplant	N (%)	4.439	0.798	2.226	5.308	4.0–5.0
		P (%)	0.382	0.057	0.257	0.508	0.2–0.5
		K (%)	4.186	0.798	2.490	5.880	2.5–3.5
		B (ppm)	73.584	57.798	23.730	250.410	18–75
1 July	flowering, 11 weeks post-transplant	N (%)	3.552	0.804	1.382	4.396	3.5–4.5
		P (%)	0.287	0.036	0.208	0.357	0.2–0.5
		K (%)	4.303	0.942	1.980	6.080	2.5–3.5
		B (ppm)	206.823	225.862	28.260	856.170	18–75
18 July	maturity, 13 weeks post-transplant	N (%)	2.964	0.666	1.450	4.171	2.25–3.0
		P (%)	0.247	0.036	0.165	0.313	0.17–0.5
		K (%)	3.015	0.614	1.360	4.140	1.6–3.0
		B (ppm)	152.947	180.258	33.870	746.360	18–75

Table 3. Correlation matrices of foliar nutrient concentrations.

	19 June				3 July				17 July			
	N	**P**	**K**	**B**	**N**	**P**	**K**	**B**	**N**	**P**	**K**	**B**
N	1				1				1			
P	0.639 ***	1			0.037	1			0.466 ***	1		
K	0.572 ***	0.547 ***	1		0.523 ***	−0.007	1		0.476 ***	−0.067	1	
B	0.219	0.043	0.220	1	0.112	−0.011	0.093	1	0.405 ***	0.166	0.245	1

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

3.2. UAS-Derived Canopy Metrics

Table 4 shows a summary of the metrics computed to quantify canopy structure and spectral reflectance. Two structural metrics, crop height standard deviation and histogram skewness, were removed from the analysis due to collinearity. The remaining metrics were used to explain variation in foliar N, K, and B concentrations. The standardized regression model coefficients for all models are summarized in Tables 5–7.

Table 4. Summary of UAS-derived canopy metrics.

	19 June		3 July		17 July	
Metric	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
Crop height mean (m)	0.725	0.089	0.873	0.123	1.004	0.183
Crop height std. dev. (m) *	0.235	0.028	0.364	0.042	0.336	0.049
Crop height histogram kurtosis	−0.743	0.471	−1.111	0.241	−0.258	1.319
Crop height histogram skewness *	−0.508	0.286	−0.264	0.257	−0.611	0.469
Rumple index (RI)	2.478	0.333	3.615	0.358	3.142	0.560
Canopy relief ratio (CRR) median	0.547	0.028	0.480	0.051	0.530	0.026
Canopy relief ratio (CRR) IQR	0.218	0.057	0.277	0.038	0.218	0.049
Moran's I	0.635	0.024	0.649	0.024	0.658	0.043
Triangular greenness index (TGI) mean	0.131	0.020	0.146	0.026	0.155	0.019
Triangular greenness index (TGI) st. dev.	0.044	0.013	0.049	0.010	0.061	0.013

Note: * Collinear variable, omitted from regression modeling.

Table 5. Summary of standardized coefficients from regression models using canopy structural metrics to predict nutrient concentrations.

	UAS Survey Date	Nutrient Sample Date	Crop Mean Height	Crop Height Histogram Kurtosis	Rumple Index	Canopy Relief Ratio Median	Canopy Relief Ratio IQR	Moran's I	Water Flow Index	Adj. r^2
N	19 June	20 June				0.609 ***	−0.328 **	0.338 **	0.222	0.428
	3 July	20 June			0.431 ***		−0.258 *	0.390 ***	0.243 *	0.398
		1 July	0.527 ***							0.259
	17 July	20 June	0.777 ***		0.485 ***				0.181	0.554
		1 July	1.242 ***	−0.675 **	0.322 **					0.592
		18 July	1.116 ***	−0.772 ***						0.301
K	19 June	20 June			−0.319 **	0.533 ***				0.259
	3 July	20 June				0.516 ***				0.248
		1 July	0.394 ***				−0.301 **			0.255
	17 July	20 June	0.697 **	−0.430 *		0.380 **				0.423
		1 July				0.588 ***				0.329
		18 July	0.275			0.474 ***	0.290 **			0.454
B	19 June	20 June			−0.499 ***	0.381 **				0.256
	3 July	20 June		0.448 ***						0.181
		1 July		0.378 **						0.121
	17 July	20 June	0.336 **		−0.383 *		0.577 ***	−0.456 ***		0.408
		1 July					0.381 ***	−0.404 ***		0.257
		18 July	0.267 *							0.048

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 6. Summary of standardized coefficients from regression models using spectral metrics to predict nutrient concentrations.

	UAS Survey Date	Nutrient Sample Date	TGI Mean	TGI Std. Dev.	Adj. r^2
N	19 June	20 June	−0.818 ***		0.661
	3 July	20 June	−0.818 ***		0.660
		1 July	−0.822 ***		0.667
	17 July	20 June	−0.235 **	−0.734 ***	0.570
		1 July	−0.203 *	−0.691 ***	0.491
		18 July	−0.292 **	−0.616 ***	0.434
K	19 June	20 June	−0.354 **	−0.317 *	0.313
	3 July	20 June		−0.655 ***	0.414
		1 July	−0.610 ***		0.356
	17 July	20 June		−0.592 ***	0.334
		1 July		−0.501 ***	0.231
		18 July		−0.621 ***	0.370
B	19 June	20 June	−0.279 *		0.054
	3 July	20 June		−0.278 *	0.054
		1 July			0.000
	17 July	20 June	0.248 *	−0.300 *	0.108
		1 July	0.287 *		0.059
		18 July			0.000

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 7. Summary of standardized coefficients from regression models using a combination of canopy structural and spectral metrics to predict nutrient concentrations.

	UAS Survey Date	Nutrient Sample Date	Crop Mean Height	Crop Height Histogram Kurtosis	Rumple Index	Canopy Relief Ratio Median	Canopy Relief Ratio IQR	Moran's I	Water Flow Index	TGI Mean	TGI Std. Dev.	Adj. r^2
N	19 June	20 June						0.203 **	0.280 ***	−0.760 ***	−0.177	0.760
	3 July	20 June	0.494 ***			−0.589 ***			0.154 *	−0.939 ***		0.756
		1 July	0.623 ***			−0.843 ***				−1.037 ***		0.805
	17 July	20 June	1.053 ***		0.529 ***	−0.267 **				−0.402 ***		0.704
		1 July	1.269 ***	−0.429 *	0.505 ***	−0.203				−0.359 ***		0.711
		18 July								−0.292 **	−0.616 ***	0.434
K	19 June	20 June					−0.384 ***		0.254 *		−0.760 ***	0.450
	3 July	20 June									−0.655 ***	0.414
		1 July								−0.610 ***		0.356
	17 July	20 June				0.417 ***				−0.186	−0.381 ***	0.475
		1 July				0.450 ***					−0.271 *	0.369
		18 July				0.417 ***					−0.408 ***	0.488
B	19 June	20 June			−0.499 ***	0.381 **						0.256
	3 July	20 June		0.448 ***								0.181
		1 July		0.508 ***		−0.467 **					−0.394 ***	0.194
	17 July	20 June	0.336 **		−0.383 *		0.577 ***	−0.456 ***				0.408
		1 July	0.784 ***				0.806 ***	−0.318 **			0.737 ***	0.368
		18 July	0.267 *									0.048

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

3.2.1. Mean Crop Height

Mean crop height provided information about the central tendency of height within a plot. It was the most important predictor of N and K concentrations at the last two survey dates (Table 5). Surprisingly, mean height was not important at the earliest growth stage observed. It was also important and positively correlated with B concentrations at the last survey date. These results are consistent with observations made by other researchers that increases in N and B concentrations can lead to an accumulation of plant biomass [27,52,53].

3.2.2. Crop Height Histogram Kurtosis

Kurtosis measured the shape of the distribution of canopy heights with increased histogram kurtosis indicating a sharper peak and more outliers in height compared to a normal distribution of heights. Plots with lower kurtosis tended to have a more uniform distribution of heights with narrow plants and wider gaps between rows (Figure 4). Plots with more canopy closure and fewer ground pixels tended to have high kurtosis. Kurtosis was an important predictor and negatively correlated with foliar concentration of N and K at the last date (Table 5). Kurtosis was important and positively correlation with B at the second survey date.

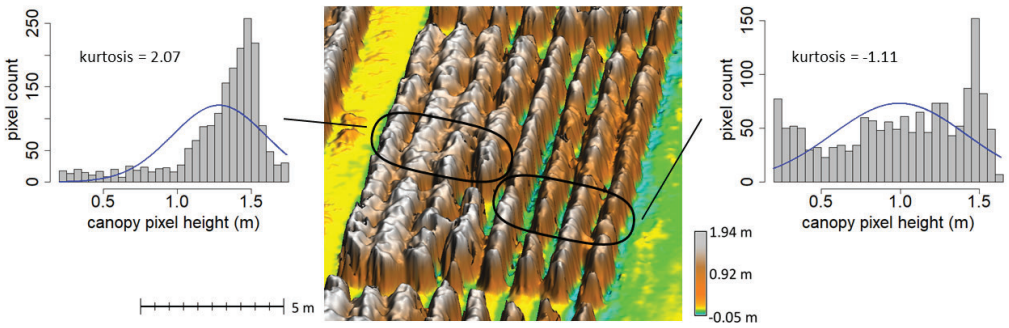


Figure 4. Tobacco canopy surface model with examples high and low height histogram kurtosis. Plots on the left side have high kurtosis, which corresponded primarily to more canopy closure. Plots on the right side have low kurtosis, which corresponded to more narrow plants with gaps between rows. The blue line in each histogram provides a comparison to a normal distribution.

3.2.3. Rumple Index

Rumple index measured the ratio of the 3D surface area of a volume to its 2D area, with an increase in the index indicating more surface complexity. As RI was computed for each plot of equal 2D area, the index can be interpreted as simply the 3D surface area of the canopy. RI was closely related to the slope of the canopy surface within each plot. Plots with steep canopy slopes (narrow plants with many ground pixels) had high RI and plots with gentle slopes (wide plants with canopy closure) had low RI (Figure 5). RI was one of the most important predictors of N concentrations at the second and third survey dates (Table 5). The positive relationship indicates that as surface area and slopes of the canopy increased, N concentrations increased. Conversely, the pyramid canopy shape symptomatic of N deficiency tended to have gentle slopes with wider plant base and lower RI. RI was negatively correlated with B and K at the first survey date, with more gentle slopes associated with lower concentrations.

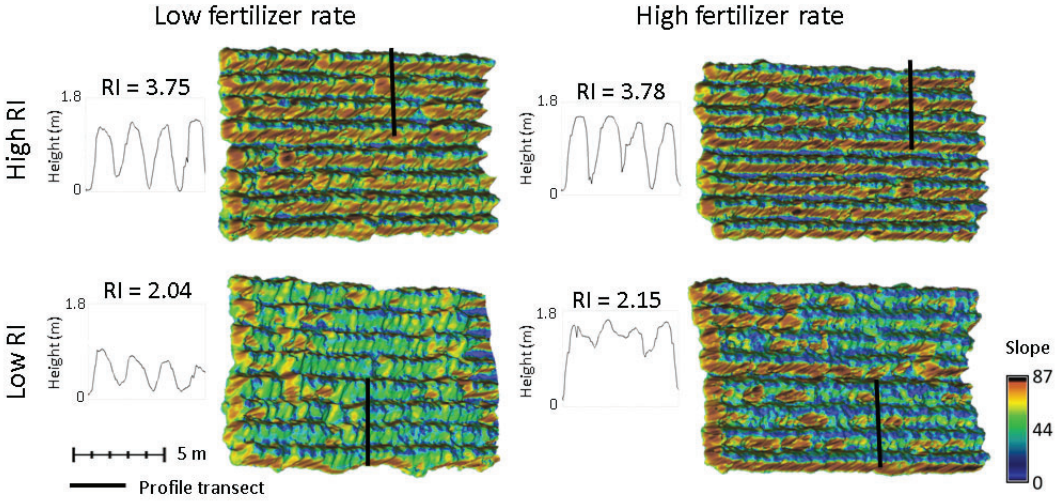


Figure 5. Slope and RI of tobacco canopy surface models with high and low fertilizer rates. Plots with high RI values are shown in the top row and low RI in the bottom row. Plots with low fertilizer rates are in the left column and high fertilizer rates in the right column. High RI values are associated with steeper canopy slopes (red and yellow), while low RI values are associated with gentle canopy slopes (blue and green). Respective vertical profiles through the canopy are shown with each model.

3.2.4. Canopy Relief Ratio

Canopy relief ratio measured the relative shape of the canopy and was useful for characterizing the ratio of plant width to height and identifying gaps in the canopy. Lower median CRR indicated more narrow plants with less canopy closure (Figure 6). High CRR median was associated with fewer gaps between rows and well defined canopy peaks, in contrast to a rounded canopy. CRR median was important and positively correlated with all nutrients observed at one or more survey dates (Table 5). In particular, it was the most important predictor for K at all dates. It was also the most important structural predictor of N at the first date. These relationships indicate that as nutrient concentrations increase, plants get wider and fill gaps between rows. The consistent positive relationship with K may also be related to K deficiency, which is characterized by leaves that curve downward making the canopy more rounded without a defined peak [42]. When spectral information was added, CRR median was negatively correlated with N at the last two survey dates (Table 7). This may be related to N deficiency symptoms, which include stunting with erect leaves that form a pyramid shape with a sharp canopy peak [42]. CRR interquartile range was associated with the presence of defined peaks as well as gaps between rows. It was important and negatively correlated with N at the first two dates and K at the second date. It was very important and positively correlated with prior B concentrations at the last date.

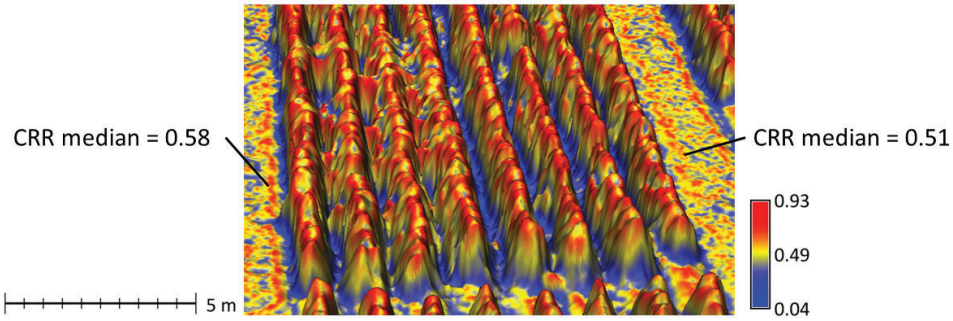


Figure 6. Tobacco canopy surface model with examples high and low canopy relief ratio (CRR) median. Plots on the left had higher CRR median values and generally had fewer valley pixels (blue) and more slope pixels (yellow). Plots on the right side had more gaps between rows (blue pixels), well-defined peaks (red), and fewer slope pixels (yellow).

3.2.5. Spatial Autocorrelation of Crop Height

Moran's I measured spatial autocorrelation of crop heights, which characterized canopy surface smoothness or likeness of crop heights within a neighborhood of surface model pixels (Figure 7). Conversely, it also detected the presence of peaks or pits in the canopy. The 0.55 m $\times$ 0.55 m neighborhood used to compute spatial autocorrelation corresponded to the approximate size of one healthy plant. Many of the plants, however, were stunted with gaps between rows. The patches of ground areas with similar height between rows resulted in increased Moran's I values in plots with low canopy closure. Overall, Moran's I explained the limited variability in nutrient concentrations. Moran's I was important, and positively correlated with N at the first two dates (Table 5). As N concentrations increased, the plot had more clusters of similar heights and was smoother with fewer peaks or pits. It was important and negatively correlated with prior B concentrations at the last survey date. As B concentrations increased, the canopy was less smooth with more gaps or spikes and has fewer patches of similar heights.

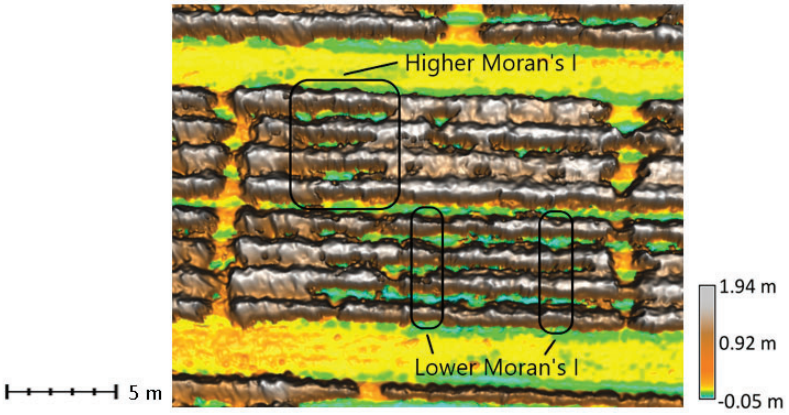


Figure 7. Canopy surface model of tobacco plots with range of Moran's I values. Plots in the upper left portion of the block have relatively high Moran's I (0.74). The plots identified in the lower right have relatively low Moran's I (0.57).

3.2.6. Water Flow Index

An overland water flow index was included in the regression models to account for topography and how water flow across the field might affect nutrient availability or uptake (Figure 8). The water index was positively correlated with N at the first two survey dates, especially when combined with the spectral metrics (Tables 5 and 7). There was considerable rainfall over the two weeks between fertilizer application and the first UAS and nutrient survey dates (97 mm). Over the remaining four weeks of the study, there was much less rainfall (30 mm). Although the field was mostly flat, even small changes in elevation can influence the movement of water and soluble nutrients. It is possible that the rainfall early in plant growth influenced the spatial distribution of nutrient and water availability in the field.

Figure 8. Derived water flow index based on simulated overland water flow.

3.2.7. Triangular Greenness Index

Triangular greenness index measured the similarity of canopy reflectance to the spectral features of chlorophyll and therefore estimated plant chlorophyll content and vigor. TGI mean was important and negatively correlated with all nutrients, although it accounted for almost no variation in B

(Table 6). TGI mean was positively correlated with B at the last survey date. Standard deviation of TGI was important and negatively correlated with all of the nutrients, especially at the last survey date, indicating that a wider range of TGI values in a plot is related to lower concentrations of foliar nutrients.

3.3. Overall Relationship between the Crop Canopy Shape, Reflectance, and Nutrient Concentration

The amount of variability explained by the canopy metrics differed by nutrient and developmental stage as shown in Figure 9. Overall, structure had a stronger relationship with N, K, and B concentrations at later growth stages. Visible range spectral reflectance consistently explained variability in N and K concentrations. For nearly all cases, combining spectral and structural information explained more variability in nutrient concentrations than either set of metrics did alone. Moreover, canopy structure and reflectance were more closely related to N and B nutrient concentrations at earlier stages of growth than nutrient concentrations measured at the time of the UAS survey.

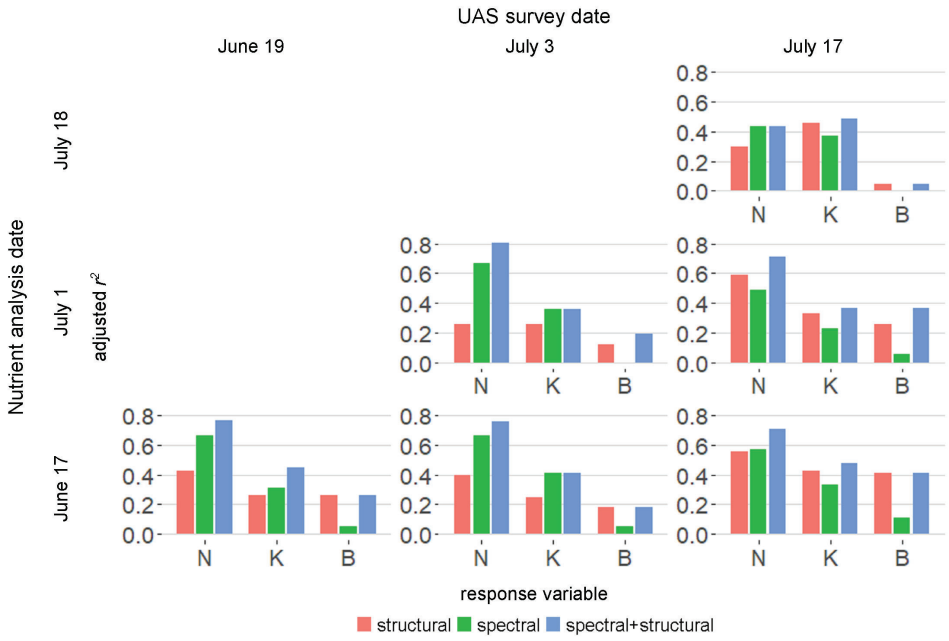


Figure 9. Adjusted r^2 for regression models using canopy structural and spectral metrics to explain foliar concentrations of elemental nutrients in tobacco plants.

3.3.1. Nitrogen

N is the most important nutrient for flue-cured tobacco yield and quality. It is typically applied in time for immediate plant uptake to avoid leaching beyond the root zone. N-deficient plants are characterized by stunting, chlorosis, and upward pointing leaves forming an acute angle with the stalk [39,42]. The structural metrics explained 40% of the variation in N concentrations at the first two growth stages observed (Figure 9). The amount of variability explained increased to as much as 60% at the last observation date. The canopy structure better explained N concentrations measured at the previous date (approximately 2 weeks prior). The visible range spectral metrics explained more variation in N concentrations than structure did overall, with approximately two-thirds of the variability explained at the first two survey dates. At the last survey date, the relationship was slightly weaker and better explained N concentrations measured at previous dates. Combining spectral and structural information strengthened the relationship and explained 75–80% of the variation in N during

the first two survey dates. Again, the relationship decreased slightly as the plant matured. This may be explained by the movement of N from leaves to reproductive tissues as the plant matures.

3.3.2. Potassium

K is also an important macronutrient needed for tobacco production. Symptoms of K deficiency include chlorosis that begins on lower leaves and quickly progresses up the plant and leaves with a distinctive downward cupping or umbrella-like shape [39,42]. The canopy structural metrics were able to explain variation in foliar K levels across all dates, with the relationship increasing slightly as the crop matured (Figure 9). Canopy structure explained 26% of K variation at the first date and up to 45% at the last survey date. K deficiency was not achieved until sometime after the first survey date with visible symptoms developing around the same time. This may explain why predictability of K increased over time. K was also the only nutrient that was better predicted using nutrition measurements taken at the same time as the UAS survey. The canopy was able to explain variation in current concentrations of foliar K as well or better than at previous stages of plant development. Visible-range spectral reflectance was able to explain similar amounts of variation in K as canopy structure. The delay in visible deficiency symptoms is reflected in the increase in variation explained by the spectral metrics over time. Combining spectral and structural predictors increased the variation in K explained at the first and last survey dates, with 45% explained at the first date and 49% variation explained at the last date. There was a slight decrease in model fit when predicting K concentrations measured at the second date. All of the other models, however, were able to explain similar amounts of variation ranging from 42% to 51%. The predictability of K using the UAS-derived canopy metrics was the most consistent of the nutrients studied.

3.3.3. Boron

B is an important nutrient for proper development of growing points in tobacco plants. It is needed in very small amounts, however, and can result in reduced leaf quality if over applied [10]. Visible symptoms of B toxicity include a very slow spreading chlorosis on lower leaves, undulating leaf deformity, and curling of leaf edges upward [39]. The canopy structural metrics were able to explain moderate variation in foliar B concentrations (Figure 9). The amount of variation explained at the earliest survey date was 26% and increased to as much as 41% at the latest date when predicting B concentration from the earliest date. B predictability dropped to as low as 5% at the last date when using nutrient concentrations measured at the time of the flight. The structural metrics were much more closely related to B concentrations measured earlier in plant development. This is consistent with other observations that symptoms of B toxicity can take 33 days to become visible [39]. Visible-range spectral reflectance explained almost no variability in B concentrations. Model fit was not improved by combining spectral and structural predictors since the spectral metrics did not explain variability in B.

4. Discussion

In this study, we investigated the relationship between structural and spectral canopy-level metrics derived from a UAS-mounted RGB camera and nutrient concentrations in flue-cured tobacco. The aim of this research was to determine if the plant response to nutrient disorders can be detected from photogrammetric crop surface models and to evaluate the usefulness of structural metrics describing canopy shape and uniformity for improving remote monitoring of crop nutritional status, particularly when used in combination with widely accepted spectral remote sensing techniques. This research demonstrates that information about crop canopy shape and uniformity can provide useful information related to crop health that can be leveraged to improve remote monitoring and modeling for precision agriculture.

This research confirms that a relationship exists between crop canopy structure and nutrient concentrations in plant tissues. Not surprisingly, the strongest relationship observed was between canopy shape and N foliar concentration (adj. $r^2 = 0.59$, increasing to adj. $r^2 = 0.81$ when combined with

TGI). Many studies have explored methods for remotely assessing N status and other crop variables closely correlated with N concentration, like leaf area index, above ground biomass, and yield. Work by Li et al. [25,26] found that RI and CRR had high correlation with LAI. These results are consistent with our study, which found that RI, CRR, histogram kurtosis, and mean height were some of the most important predictors of foliar N concentrations. The metrics used in this research and other similar structural measures that quantify the spatial arrangement of heights could be useful for remote approximations of other crop variables correlated with N.

The visible-band spectral vegetation index used in this study—Triangular Greenness Index—provides an approximation of chlorophyll content. The close relationship between N and chlorophyll makes it unsurprising that TGI was limited in explaining concentrations of the other nutrients. Chlorosis, however, is characteristic of K deficiency [39,42], so TGI was able to explain moderate variation in foliar K concentration. As expected, variation in B was not explained by TGI as B toxicity is not closely associated with chlorosis.

Several important decisions about how the metrics were computed affected their ability to explain nutrient concentrations. For example, CRR was more effective when a subplant neighborhood size was used. This suggests that capturing variability in relative shape within a plant influences the usefulness of this metric for characterizing canopy structure. Additionally, the weights matrix used to compute Moran's I also made this metric more or less useful. The size of the weights matrix should be chosen based on the area in which heights are expected to be spatially autocorrelated. It is also important, however, to consider features that are not of interest, like bare ground, that fall within the matrix and influence the computation. The effectiveness of Moran's I may be improved by removing soil pixels and using a directional, anisotropic weights matrix to limit the computation to along plant rows for future studies. Nonetheless, lower Moran's I was useful for identifying plots with B toxicity. This was likely due to plots with toxic B levels being associated with higher overall fertilizer rates and very full, closed canopies with no clusters of ground pixels. Furthermore, plots with closed canopies lacked features with high contrast, like bare earth, that could be used as tie points in the SfM process, which may have increased point cloud errors (unrealistic peaks and pits) further reducing spatial autocorrelation. Further research should be done to understand how this approach is sensitive to decisions about the window sizes and weight matrices used.

The regression models predicting N and B foliar concentrations all improved when using nutrient concentrations measured 2–4 weeks prior to the UAS survey. While the direct cause of this effect remains unclear, a few possible explanations warrant discussion. N is very mobile within the plant, so this preference for past nutrient measurements could be explained by the movement of these nutrients to other, unsampled parts of the plants, particularly reproductive tissue [54]. The sharp decrease in correlation between nutrient concentrations at the second and third foliar sample dates could also be related to nutrient mobility (Table 3). Another potential explanation for the increased relationship between the canopy and past nutrient concentrations is an accumulation of symptomlogy over time making the effects of a disorder more apparent. There also could be a delay in appearance of symptoms measurable by the structural and spectral metrics. B toxicity in particular has been shown to take over a month to develop visible symptoms [39]. As this delay pattern was observed for two of the three nutrients studied, further investigation should be completed to understand the cause and practical implications this has for precision agriculture or rapid phenotyping applications.

Symptoms of N and K deficiencies in flue-cured tobacco described by McMurtrey and Henry et al. [39,42] were well captured by the structural metrics used in this study and the distinct features of the deficiencies can be used to differentiate each. N deficiencies are associated with chlorosis, stunting, and erect leaves that form an acute angle with the stalk. K deficiencies are also associated with chlorosis and stunting, but in contrast to N deficiencies, leaf deformation is characterized by a downward cupping or umbrella-like shape. This explains why the most important structural metric for predicting K concentrations was CRR. K-deficient plants with downward curving leaves are limited in width and less likely to fill spaces between rows, resulting in more valleys in the canopy

shape, which have CRR close to zero and lower median CRR values. Furthermore, the downward curved leaves result in a rounded canopy with fewer peaks, which further lowers median CRR values. Whereas high height histogram kurtosis was one of the most important predictors of N deficiency due to leaves that stand at an acute angle relative to the stalk creating a relatively short, pyramid shaped canopy with fewer gaps between rows. Figure 10 illustrates the differences in canopy shape for N and K deficiencies. Kurtosis and CRR in combination with mean height, which captures stunting, were effective in distinguishing N and K deficiencies. If spectral reflectance information is also considered, it is possible that N and P deficiencies could also be distinguished due to the lack of chlorosis associated with P deficiencies [39,42]. Future experiments to observed P deficiencies in field-grown tobacco are needed to test this.

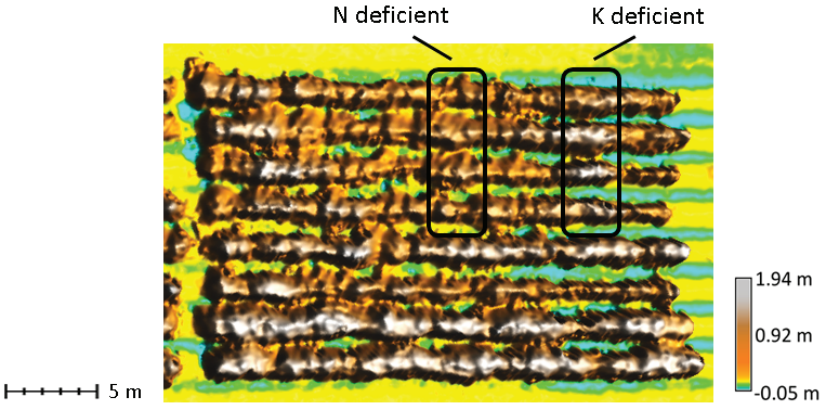


Figure 10. Canopy surface model with nutrient deficient plots. The K-deficient plot in the upper right has a rounded canopy shape with minimal closure between rows resulting in lower median CRR. The N-deficient plot to the left has a pyramid shape with sharper peaks and is generally stunted resulting in high height histogram kurtosis and low mean height.

Many studies have focused on validating the accuracy of measurements taken from photogrammetric surface models, like crop height and volume, through field measurements and precise ground control [53,55–57]. While the importance of accuracy for many measurements taken from canopy surface models is clear, the measurements of interest in this study focus on relative shape of the canopy rather than absolute quantities. Point cloud quality, however, is important to ensure that the reconstruction of the canopy shape is representative of reality. Decisions made during image acquisition and processing, like camera angle, image overlap, flight time of day, altitude, point filtering, and surface interpolation, all have an impact on feature reconstruction. For example, the challenge faced by the SfM algorithm when distinguishing between low-contrast features like plants, shadows, and dark, moist soil was apparent in the June 17 CSM. The canopy appears more closed with fewer soil pixels compared to the 3 July CSM, which was created from images with high contrast between features due to better lighting and lower soil moisture. Furthermore, oblique imagery, which has been shown to improve reconstruction of vertical features, like buildings, forests, and steep terrain [58,59], was not collected for use in this study. Future implementation of this approach should incorporate oblique imagery during point cloud generation to improve representation of canopy features, especially along plant row edges. Several studies [55,58,60–63] have evaluated the impact of these considerations on feature reconstruction, but further research is needed to understand how sensitive the canopy structural metrics used in this study are to data acquisition and processing parameters. However, the relatively strong, consistent relationships between canopy structure and nutrients observed across all dates suggest that trends in canopy shape are being represented well enough in the data to provide meaningful, useful information.

5. Conclusions

In this study, we developed a methodology for extracting crop canopy characteristics from high-resolution, photogrammetric canopy surface models and demonstrated how these characteristics can be used to improve remote monitoring of crop nutrient status. The strength of this approach is its use of both 3D canopy geometry and spectral reflectance derived from very low cost, ubiquitous UAS with a consumer-grade RGB camera. Combining spectral and structural canopy information improved model fit in nearly every case observed, making a strong case for including structural metrics in models used to remotely estimate crop variables. The relationship between canopy structure measured from point clouds and crop stress should be further explored as we may be able to leverage this relationship as technologies for producing more precise point clouds become more readily available (e.g., active LIDAR sensors optimized for use with UAS). In the meantime, useful information about canopy structure can be extracted from photogrammetric point clouds to improve remote estimates of nutritional status and identify areas that warrant further ground-based observation and evaluation. Further investigation should also be done to explore how canopy structure varies in response to different types of stress to attempt to remotely distinguish the cause of plant stress without destructive, specialized testing. High spatial and temporal resolution data on 3D canopy geometry opens up new opportunities to explore patterns in plant responses to stress that could provide useful evidence to support management decisions or rapid phenotyping.

Author Contributions: Conceptualization, K.M. and H.M.; Methodology, K.M., H.M., J.B.H., B.W., and M.C.V.; Validation, K.M. and H.M.; Formal analysis, K.M.; Investigation, K.M. and J.B.H.; Resources, K.M., J.B.H., B.W., and M.C.V.; Writing—original draft preparation, K.M.; Writing—review and editing, K.M., H.M., J.B.H., M.C.V., and A.S.H.; visualization, K.M.; funding acquisition, K.M, J.B.H., B.W., and M.C.V. All authors have read and agreed to the published version of the manuscript.

Funding: Development of the research methodology was partially funded by the North Carolina Space Grant. The field experiment and data collection were funded by the North Carolina Tobacco Foundation, Inc.

Acknowledgments: Thanks to Vick Family Farms of Wilson, NC for donating the field space and planting and maintaining the research site. Thanks to Norman Harrell and the Wilson County Cooperative Extension Center for their help working with the farm. Thanks to technicians Jeremy Machacek, Ingram McCall, and Scott Whitley for managing research preparation. Last, thank you to the graduate and undergraduate students for their contributions applying treatments, collecting data, and more: Jeb Bullard, Drew Clapp, Camden Finch, Paige Herring, Nathan Jahnke, Nick Manning, Channie Renn, Maggie Short, Drake Stevens, and Patrick Veazie.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

UAS	Unmanned aerial system
N	Nitrogen
P	Phosphorus
K	Potassium
B	Boron
NDVI	Normalized Difference Vegetation Index
VARI	Visible Atmospheric Resistance Index
TGI	Triangular Greenness Index
RGB	Red Green Blue
SfM	Structure from Motion
LAI	Leaf area index
CEC	Cation exchange capacity
DSM	Digital surface model
DEM	Digital elevation model

RMSEz	Root mean squared error in z direction
CSM	Crop surface model
RI	Rumple index
CRR	Canopy relief ratio
IQR	Interquartile range
VIF	Variable inflation factor
AIC	Akaike's information criteria

References

1. Stafford, J.V. Implementing Precision Agriculture in the 21st Century. *J. Agric. Eng. Res.* **2000**, *76*, 267–275. [CrossRef]
2. Hunt, E.R.; Daughtry, C.S.T. What good are unmanned aircraft systems for agricultural remote sensing and precision agriculture? *Int. J. Remote Sens.* **2018**, 39, 5345–5376. [CrossRef]
3. Zhang, C.; Kovacs, J. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [CrossRef]
4. Corti, M.; Cavalli, D.; Cabassi, G.; Vigoni, A.; Degano, L.; Marino Gallina, P. Application of a low-cost camera on a UAV to estimate maize nitrogen-related variables. *Precis. Agric.* **2019**, *20*, 675–696. [CrossRef]
5. Cilia, C.; Panigada, C.; Rossini, M.; Meroni, M.; Busetto, L.; Amaducci, S.; Boschetti, M.; Picchi, V.; Colombo, R. Nitrogen Status Assessment for Variable Rate Fertilization in Maize through Hyperspectral Imagery. *Remote Sens.* **2014**, *6*, 6549–6565. [CrossRef]
6. Maresma, A.; Lloveras, J.; Martínez-Casasnovas, J.A. Use of Multispectral Airborne Images to Improve In-Season Nitrogen Management, Predict Grain Yield and Estimate Economic Return of Maize in Irrigated High Yielding Environments. *Remote Sens.* **2018**, *10*, 543. [CrossRef]
7. Berger, K.; Verrelst, J.; Féret, J.B.; Wang, Z.; Wocher, M.; Strathmann, M.; Danner, M.; Mauser, W.; Hank, T. Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. *Remote Sens. Environ.* **2020**, *242*, 111758. [CrossRef]
8. Mitra, G. Essential plant nutrients and recent concepts about their uptake. In *Essential Plant Nutrients*, 1st ed.; Springer: Cham, Switzerland, 2017; pp. 3–36. [CrossRef]
9. Henry, J.B.; Vann, M.C.; Lewis, R.S. Agronomic Practices Affecting Nicotine Concentration in Flue-Cured Tobacco: A Review. *Agron. J.* **2019**, *111*, 3067–3075. [CrossRef]
10. Ruiz, J.M.; Lopez-Lefebre, L.R.; Sanchez, E.; Rivero, R.M.; Garcia, P.C.; Romero, L. Preliminary studies on the influence of boron on the foliar biomass and quality of tobacco leaves subjected to NO3- fertilisation. *J. Sci. Food Agric.* **2001**, *81*, 739–744. [CrossRef]
11. Tariq, M.; Akbar, A.; Lataf-ul-Haq.; Khan, A. Comparing Application Methods for Boron Fertilizer on the Yield and Quality of Tobacco (*Nicotiana tabacum* L.). *Commun. Soil Sci. Plant Anal.* **2010**, *41*, 1525–1537. [CrossRef]
12. Römheld, V. Diagnosis of deficiency and toxicity of nutrients. In *Marschner's Mineral Nutrition of Higher Plants*; Elsevier Science & Technology: San Diego, CA, USA, 2011.
13. Stroppiana, D.; Migliazzi, M.; Chiarabini, V.; Crema, A.; Musanti, M.; Franchino, C.; Villa, P. Rice yield estimation using multispectral data from UAV: A preliminary experiment in northern Italy. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 4664–4667. [CrossRef]
14. Bolton, D.K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* **2013**, *173*, 74–84. [CrossRef]
15. Stanton, C.; Starek, M.J.; Elliott, N.; Brewer, M.; Maeda, M.M.; Chu, T. Unmanned aircraft system-derived crop height and normalized difference vegetation index metrics for sorghum yield and aphid stress assessment. *J. Appl. Remote Sens.* **2017**, *11*, 026035. [CrossRef]
16. Severtson, D.; Callow, N.; Flower, K.; Neuhaus, A.; Olejnik, M.; Nansen, C. Unmanned aerial vehicle canopy reflectance data detects potassium deficiency and green peach aphid susceptibility in canola. *Precis. Agric.* **2016**, *17*, 659–677. [CrossRef]
17. Henry, J.B. Characterization of Tobacco Nutrient Disorders via Remote Sensing. Ph.D. Thesis, North Carolina State University, Raleigh, NC, USA, 2019.

18. Adão, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sens.* **2017**, *9*, 1110. [CrossRef]
19. Hunt, E.R.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.T.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 103–112. [CrossRef]
20. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [CrossRef]
21. Liew, O.W.; Chong, P.C.J.; Li, B.; Asundi, A.K. Signature Optical Cues: Emerging Technologies for Monitoring Plant Health. *Sensors* **2008**, *8*, 3205–3239. [CrossRef]
22. Carrivick, J.; Smith, M.; Quincey, D. *Structure from Motion in the Geosciences*, 1st ed.; Number Book, Whole in New Analytical Methods in Earth and Environmental Science; Wiley-Blackwell: Hoboken, NJ, USA, 2016.
23. Hunt, E.R.; Rondon, S.I. Detection of potato beetle damage using remote sensing from small unmanned aircraft systems. *J. Appl. Remote Sens.* **2017**, *11*, 026013. [CrossRef]
24. Parker, G.G.; Russ, M.E. The canopy surface and stand development: assessing forest canopy structure and complexity with near-surface altimetry. *For. Ecol. Manag.* **2004**, *189*, 307–315. [CrossRef]
25. Li, W.; Niu, Z.; Chen, H.; Li, D. Characterizing canopy structural complexity for the estimation of maize LAI based on ALS data and UAV stereo images. *Int. J. Remote Sens.* **2017**, *38*, 2106–2116. [CrossRef]
26. Li, W.; Niu, Z.; Chen, H.; Li, D.; Wu, M.; Zhao, W. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecol. Indic.* **2016**, *67*, 637–648. [CrossRef]
27. Bendig, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. *Remote Sens.* **2014**, *6*, 10395–10412. [CrossRef]
28. Chu, T.; Chen, R.; Landivar, J.A.; Maeda, M.M.; Yang, C.; Starek, M.J. Cotton growth modeling and assessment using unmanned aircraft system visual-band imagery. *J. Appl. Remote Sens.* **2016**, *10*, 036018-17. [CrossRef]
29. Fisher, L.R. *Flue-Cured Tobacco Guide*, rev. ed.; North Carolina Cooperative Extension: Raleigh, NC, USA, 2019.
30. Mitasova, H.; Mitas, L.; Harmon, R.S. Simultaneous spline approximation and topographic analysis for lidar elevation data in open-source GIS. *IEEE Geosci. Remote Sens. Lett.* **2005**, *2*, 375–379. [CrossRef]
31. Mitas, L.; Mitasova, H.; Kosinovsky, I.; McCauley, D.; Hofierka, J.; Zubal, S.; Lacko, M. *GRASS GIS: V.surf.rst Module*; GRASS: Raleigh, NC, USA, 2018.
32. North Carolina Geographic Information Coordinating Council. *NC OneMap Orthoimagery*; North Carolina Geographic Information Coordinating Council: Raleigh, NC, USA, 2017.
33. North Carolina Emergency Management. *NC Floodplain Mapping Program QL2 LiDAR*; North Carolina Emergency Management: Raleigh, NC, USA, 2015.
34. Jensen, J.; Mathews, A. Assessment of image-based point cloud products to generate a bare earth surface and estimate canopy heights in a woodland ecosystem. *Remote Sens.* **2016**, *8*, 50. [CrossRef]
35. Jenness, J.S. Calculating landscape surface area from digital elevation models. *Wildl. Soc. Bull.* **2004**, *32*, 829–839. [CrossRef]
36. Pike, R.J.; Wilson, S.E. Elevation-Relief Ratio, Hypsometric Integral, and Geomorphic Area-Altitude Analysis. *Geol. Soc. Am. Bull.* **1971**, *82*, 1079. [CrossRef]
37. Moran, P.A.P. Notes of continuous stochastic phenomena. *Biometrika* **1950**, *37*, 17–23. [CrossRef]
38. Mitasova, H.; Thaxton, C.; Hofierka, J.; McLaughlin, R.; Moore, A.; Mitas, L. Path sampling method for modeling overland water flow, sediment transport, and short term terrain evolution in Open Source GIS. In *Developments in Water Science*; Elsevier B.V.: Amsterdam, The Netherlands, 2004; Volume 55, pp. 1479–1490. [CrossRef]
39. Henry, J.B.; Vann, M.; McCall, I.; Cockson, P.; Whipker, B.E. Nutrient Disorders of Burley and Flue-Cured Tobacco: Part 2—Micronutrient Disorders. *Crop. Forage Turfgrass Manag.* **2018**, *4*, 1–7. [CrossRef]
40. Henry, J.B.; Vann, M.; McCall, I.; Cockson, P.; Whipker, B.E. Nutrient Disorders of Burley and Flue-Cured Tobacco: Part 1—Macronutrient Deficiencies. *Crop Forage Turfgrass Manag.* **2018**, *4*, 1–8. [CrossRef]
41. Raper, C.D.; Smith, W.T. Factors Affecting the Development of Flue-cured Tobacco Grown in Artificial Environments. V. Effects of Humidity and Nitrogen Nutrition1. *Agron. J.* **1975**, *67*, 307–312. [CrossRef]

42. McMurtrey, J.E. *Symptoms on Field-Grown Tobacco Characteristic of the Deficient Supply of Each of Several Essential Chemical Elements*; Technical Report 612; United States Department of Agriculture: Washington, DC, USA, 1938.
43. Peterson, B.G.; Carl, P. *PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis*; CRANE Co.: Stamford, CT, USA, 2020.
44. Roussel, J.R.; Auty, D. *lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications*; CRANE Co.: Stamford, CT, USA, 2019.
45. Hijmans, R.J. *Raster: Geographic Data Analysis and Modeling*; CRANE Co.: Stamford, CT, USA, 2017.
46. Anselin, L. Local Indicators of Spatial Association—LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [CrossRef]
47. Mitasova, H.; Hofierka, J.; Thaxton, C.*GRASS GIS: R.sim.water Module*; GRASS: Raleigh, NC, USA, 2019.
48. Hebbali, A. *olsrr: Tools for Building OLS Regression Models*; R package version 0.5.3; GRASS: Raleigh, NC, USA, 2020.
49. Hebbali, A. *Collinearity Diagnostics, Model Fit & Variable Contribution*; GRASS: Raleigh, NC, USA, 2020.
50. Wada, Y.; Wisser, D.; Eisner, S.; Flörke, M.; Gerten, D.; Haddeland, I.; Hanasaki, N.; Masaki, Y.; Portmann, F.T.; Stacke, T.; et al. Multimodel projections and uncertainties of irrigation water demand under climate change. *Geophys. Res. Lett.* **2013**, *40*, 4626–4632. [CrossRef]
51. Campbell, C.R. Tobacco, Flue-cured. In *Reference Sufficiency Ranges for Plant Analysis in the Southern Region of the United States*; Number Bulletin #394 in Southern Cooperative Series; North Carolina Department of Agriculture and Consumer Services Agronomic Division: Raleigh, NC, USA, 2000.
52. López-Lefebre, L.R.; Rivero, R.M.; García, P.C.; Sánchez, E.; Ruiz, J.M.; Romero, L. Boron Effect on Mineral Nutrients of Tobacco. *J. Plant Nutr.* **2002**, *25*, 509–522. [CrossRef]
53. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *39*, 79–87. [CrossRef]
54. Ohyama, T. Nitrogen as a major essential element of plants. In *Nitrogen Assimilation in Plants*; Research Signpost: New York, NY, USA, 2010; pp. 1–18.
55. Kameyama, S.; Sugiura, K. Estimating Tree Height and Volume Using Unmanned Aerial Vehicle Photography and SfM Technology, with Verification of Result Accuracy. *Drones* **2020**, *4*, 19. [CrossRef]
56. Fawcett, D.; Azlan, B.; Hill, T.C.; Kho, L.K.; Bennie, J.; Anderson, K. Unmanned aerial vehicle (UAV) derived structure-from-motion photogrammetry point clouds for oil palm (Elaeis guineensis) canopy segmentation and height estimation. *Int. J. Remote Sens.* **2019**, *40*, 7538–7560. [CrossRef]
57. Jayathunga, S.; Owari, T.; Tsuyuki, S. Evaluating the Performance of Photogrammetric Products Using Fixed-Wing UAV Imagery over a Mixed Conifer–Broadleaf Forest: Comparison with Airborne Laser Scanning. *Remote Sens.* **2018**, *10*, 187. [CrossRef]
58. Díaz, G.M.; Mohr-Bell, D.; Garrett, M.; Muñoz, L.; Lencinas, J.D. Customizing unmanned aircraft systems to reduce forest inventory costs: can oblique images substantially improve the 3D reconstruction of the canopy? *Int. J. Remote Sens.* **2020**, *41*, 3480–3510. [CrossRef]
59. Nesbit, P.R.; Hugenholtz, C.H. Enhancing UAV–SfM 3D Model Accuracy in High-Relief Landscapes by Incorporating Oblique Images. *Remote Sens.* **2019**, *11*. [CrossRef]
60. Dandois, J.P.; Baker, M.; Olano, M.; Parker, G.G.; Ellis, E.C. What is the Point? Evaluating the Structure, Color, and Semantic Traits of Computer Vision Point Clouds of Vegetation. *Remote Sens.* **2017**, *9*, 355. [CrossRef]
61. Dandois, J.; Olano, M.; Ellis, E. Optimal Altitude, Overlap, and Weather Conditions for Computer Vision UAV Estimates of Forest Structure. *Remote Sens.* **2015**, *7*, 13895–13920. [CrossRef]
62. James, M.R.; Robson, S. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surf. Process. Landforms* **2014**, *39*, 1413–1420. [CrossRef]
63. Cunliffe, A.M.; Brazier, R.E.; Anderson, K. Ultra-fine grain landscape-scale quantification of dryland vegetation structure with drone-acquired structure-from-motion photogrammetry. *Remote Sens. Environ.* **2016**, *183*, 129–143. [CrossRef]

CHAPTER 5: CONCLUSION

This dissertation presents three manuscripts that describe novel, interdisciplinary approaches to support data-driven decisions about the management of agricultural pests and diseases across the biosecurity continuum. Throughout this dissertation, applied research has been emphasized by collaborating with pest management practitioners, who are tasked with rapid, efficient responses to emerging plant pests, to develop the approaches and case studies used.

Chapter 2 presented PoPS Global, a simulation framework that uses core drivers of biological invasions and global open data to forecast global spread of invasive plant pests through bridgehead populations. This multiscale framework represents the first effort to integrate processes of terrestrial plant pest biological invasions across scales, including global propagule movement through trade, national phytosanitary capacity, and local population establishment and spread. The modular design of the framework supports early, proactive responses for emerging pests even when limited data are available and enables forecasts at different spatial and temporal resolutions. The simulation provides insights into bridgehead population pathways and how these local populations impact the global transmission network. In particular, the case study of spotted lanternfly (*Lycorma delicatula*) showed that bridgehead populations are likely to form in Europe and become additional sources of propagule in the future. This information can be used by phytosanitary agencies to target trade pathways for increased surveillance. Future work will test the framework's capacity for modeling multiple pest species at once, grouped by pathway or taxonomy, to support pest prioritization decisions.

Chapter 3 presented PoPS Border, a phytosanitary consignment inspection simulator that enables experiments and protocol testing that would not otherwise be possible. Past approaches have attempted to use inspection data not collected with statistical methods to characterize risk levels for trade pathways and design risk-based sampling protocols. The tool developed in the dissertation avoids that common pitfall by computing inspection outcomes on synthetic consignments based on parameters set by the user, informed by high quality inspection data, other analyses, or expert opinion. The tool provides a way to directly measure inspection effectiveness and the work required for hypothetical scenarios, allowing agencies to test proposed inspection approaches to identify vulnerabilities before deploying them in an operational environment. The use cases demonstrated that inspection effectiveness will vary by

sampling approach, contaminant arrangement, and packaging. This information can be used by phytosanitary agencies to design inspection strategies that vary based on cargo pathway and risk levels to balance trade-offs in efficiency and effectiveness. Future work will include a workshop to train analysts from USDA APHIS to use the tool and a pilot study of a risk-based sampling program for cut flower inspections.

Chapter 4 presented a methodology for extracting crop canopy characteristics from high-resolution, photogrammetric canopy surface models. The results provided evidence that these characteristics can be used to improve remote monitoring of crop nutrient stress. The strength of this approach is its use of both 3D canopy geometry and spectral reflectance derived from very low cost, ubiquitous unmanned aerial system (UAS) with a consumer-grade RGB camera. The results demonstrated that crop canopy geometry can provide information about nutrient stress, especially when combined with RGB spectral reflectance. Further investigation should also be done to explore how canopy structure varies in response to different types of stress to attempt to remotely distinguish the cause of plant stress without destructive, specialized testing. High spatial and temporal resolution data on 3D canopy geometry opens up new opportunities to explore patterns in plant responses to stress that could provide useful evidence to support management decisions or rapid phenotyping.

The approaches presented in this dissertation support efforts at multiple scales across the biosecurity continuum and provide valuable information that can inform other models and risk assessments (Figure 1). PoPS Global provides information to help managers prioritize emerging pests, and allocate resources for offshore, border, and post-border biosecurity efforts. It also provides an assessment of trade pathway risk levels that can be used in PoPS Border to vary contamination rates. PoPS Border supports the development of more effective and efficient inspection protocols to reduce entry of pests. It also can provide an estimate of propagule pressure (i.e., the number of pest individuals making past inspections) to inform on the ground surveillance efforts and be used as a driver in PoPS Global or other spread models. Finally, the UAS crop monitoring technique presented in Chapter 4 has potential to provide early detection of invasive plant pests. These tools leverage diverse data to model complex systems spanning ecology, economics, agronomy, and operations research and have immediate, relevant use cases to support pest management practitioners.

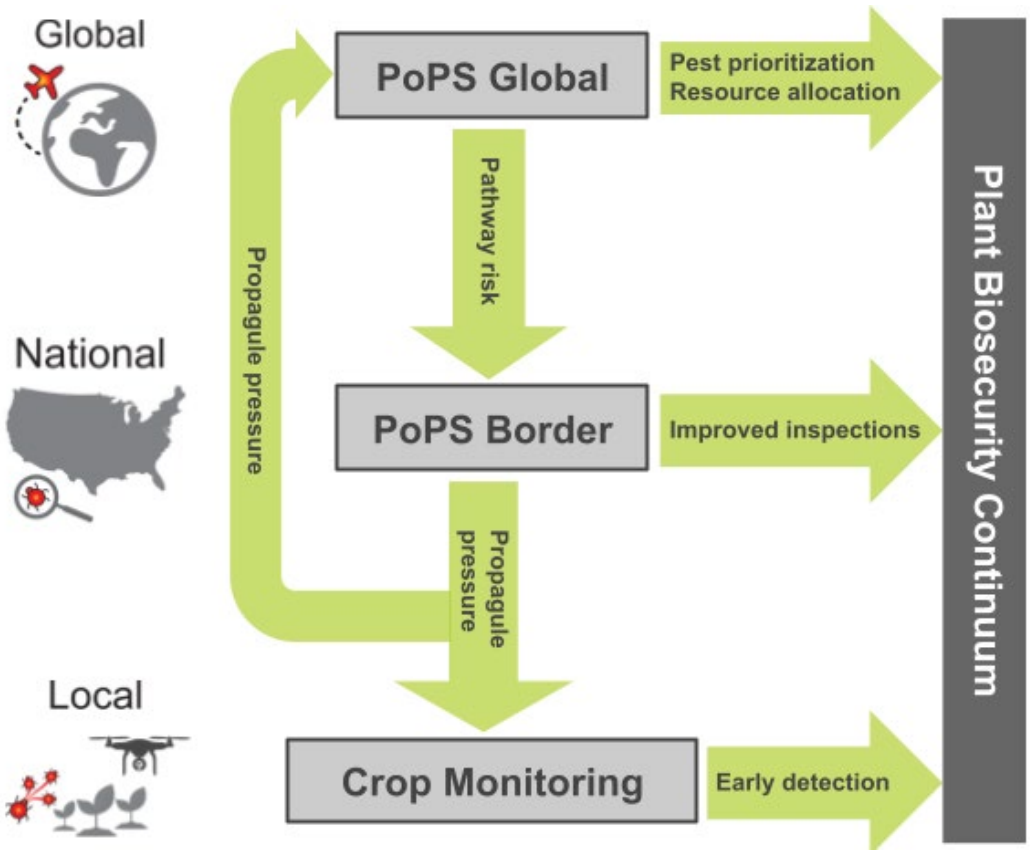


Figure 1. The three research chapters of this dissertation -- PoPS Global, PoPS Border, and crop monitoring with UAS -- support efforts at multiple scales across the plant biosecurity continuum and output data that can inform one another and other quantitative assessments.

APPENDICES

APPENDIX A: SUPPORTING INFORMATION ACCOMPANYING CHAPTER 2: FORECASTING GLOBAL SPREAD OF INVASIVE PESTS AND PATHOGENS THROUGH INTERNATIONAL TRADE

S1. Data and Model Assumptions

S1.1 General Assumptions

The PoPS Global model was adapted from the model presented by Seebens et al. (2013; 2016) for aquatic invasive species. We used different terminology for the probabilities computed to align with the phytosanitary language used by USDA APHIS. The terms entry, establishment, and introduction used in PoPS Global correspond to introduction, establishment, and invasion, respectively, in the Seebens et al. model.

Several optional modules are available in the PoPS Global framework. Species survival can be incorporated into the probability of entry computation by setting the mortality rate to a value less than one. This parameter is used if the pest or pathogen is expected to experience some level of mortality during transport. The mortality rate is weighted by the distance between nodes and decreases the probability of entry exponentially as distance increases. Species flexibility can be used if a pest or pathogen is polyphagous to increase the probability of establishment. The species flexibility parameter can be parameterized as the number of host families, or any other metric of flexibility, and is weighted to increase or decrease the parameter's influence on the probability of establishment.

When the pest or pathogen being modeled has host plants that include agricultural crops, host area data can be obtained as a total harvested area per country from the FAOSTAT database that varies by year (FAO, 2020). Alternatively, a spatially explicit raster of average harvested area from 1997-2003 can be obtained for most crops from EarthStat (Monfreda et al., 2008). Using the crop host maps omit annual variability but does allow the analysis to be constrained to areas likely to be commodity destinations using the Human Influence Index. The crop host maps also permit increasing the spatial resolution of the network to sub-country units, which is not possible with aggregated data from FAOSTAT.

We assumed the UN Comtrade data (DESA/UNSD, 2019) was correct as reported and did not adjust, except for converting nominal values to inflation-adjusted values. The US Bureau of Labor Statistics Consumer Price Index (Series CUUR0000SA0L1E) was used to adjust the trade values. The trade data were not always available for both annual and monthly timesteps.

We first checked the data availability and if the monthly timestep desired was not available, we downloaded the annual timestep and divided equally to get monthly values or summed monthly timesteps to get an annual value if the annual value was not available. When using multiple trade commodities in the model, data for each commodity (Harmonized Commodity Description and Coding System) are downloaded separately and can be summed to use a total traded goods value to compute the probability of entry. Alternatively, separate probabilities of entry can be computed for each commodity to vary the commodity importance parameter (lambda). The UN country codes used in the trade data were converted to ISO3 country codes to match the trade data to geographic areas. The scripts un_to_iso.py and get_comtrade.py in the project Github repository contain details on the country code crosswalk and how historical data are handled. Since the UN Comtrade database does not separately report trade values for Taiwan and it was unclear how data for Taiwan were reported in other databases used, Taiwan was not modeled separately.

The national phytosanitary response capacities data used in the case study were index values ranging from 0 to 3 in increments of 0.5 (Early et al., 2016). The phytosanitary capacity score acted as a risk reduction factor to reduce the probability of entry. We assumed that nearly all countries would have some minimal phytosanitary protocols as a requirement of participating as an international trading partner and that countries with the highest phytosanitary capacity would not achieve perfect prevention of pest and pathogen entry. We conceptualized the phytosanitary capacity parameter as the fraction of species propagule that would survive after phytosanitary measures are taken (inspections, treatments). We rescaled the index so that the lowest scoring countries were assigned 0.3, representing a maximum propagule slippage of 70% and the highest scoring countries were assigned 0.8, representing a minimum propagule slippage of 20%.

S1.2 Case Study Assumptions

We estimated the global distribution of tree of heaven using Maximum Entropy (MaxEnt) modeling (Figure S1; Phillips et al., 2006, 2004). We obtained 40,067 presence data points from BIEN (Enquist et al., 2016; Maitner et al., 2018) and GBIF (GBIF, 2021). We then split the presence points into training and testing datasets using an 80/20 split. We selected the most important climatic variables from an initial MaxEnt model using all 19 Worldclim bioclimatic

variables (Fick & Hijmans, 2017) at a resolution of 10 arc-minutes (approximately 18 km^2). Our final climatic variables were mean annual temperature, mean temperature of the coldest quarter, annual precipitation, and precipitation of the coldest quarter. These 4 variables accounted for 85.9% of the variation across all of the data.

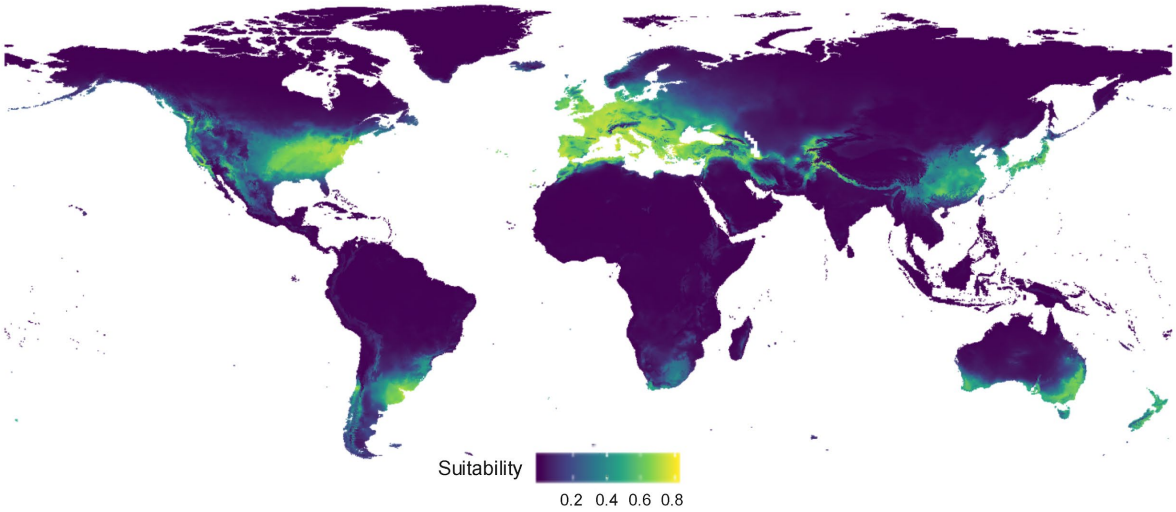


Figure S1. Tree of heaven suitability map created using MaxEnt

We used China and Vietnam as the native countries of spotted lanternfly in our case study. There are reports, however, that the native range may extend to India, Taiwan, or North Korea. We used the native range specified in the CABI Invasive Species Compendium (CABI, 2021), which was consistent across several studies (Lee et al., 2019; Wakie et al., 2020). We also limited transport of spotted lanternfly to the egg laying months. In the Northern Hemisphere, the model allowed transmission from September through April. This is based on several reports stating that eggs are laid from September to December and will begin hatching in May (Dara et al., 2015; USDA-APHIS-PPQ-CPHST, 2018; Wakie et al., 2020). The model allowed transmission from May through October in the Southern Hemisphere. In the management scenarios presented in the case study, stone trade was reduced by 100% for each country and timestep included in the management design.

S2. Parameter Calibration

There is an expected lag time between when the pest was introduced and the first recorded observation, so we treated the pest observations as fuzzy data points and considered accurate any simulated predicted introductions up to five years prior to the observation date. We calibrated the unknown parameter values (i.e., parameters without values based on species biology or literature), alpha (i.e., baseline establishment probability) and lambda (i.e., commodity importance weight), by conducting a value grid search resulting in 372 parameter sample sets, each with 50 stochastic realizations. The parameter sets were ranked using a F_β score computed for each run and averaged per parameter sample set. F_β score is a metric for evaluating binary classifications (e.g., pest introduced or not) based on precision and recall. Precision is the percentage of correctly predicted introductions (i.e., three observed introductions within the fuzzy time frames) out of all predicted introductions (i.e., true positives plus false positives) between the model start year (i.e., 2006) and the last full year of observed data minus the fuzzy window (i.e., 2019 - 5 = 2014). Recall is the percentage of correctly predicted introductions out of all actual introductions (i.e., true positives plus false negatives). High precision and recall corresponded to low commission and omission error, respectively. The F_β score placed more weight on recall (i.e., $\beta = 2$) to emphasize correctly predicting the three known introductions and permit predictions of unobserved occurrences not included in the calibration data. The model was run with the top five ranking parameter sets 1,000 times each and the highest-ranking parameter set was selected using the highest average F_β score.

S3. Data Pipeline and Model Workflow

The PoPS Global workflow is implemented in a series of Jupyter Notebooks available in the code repository for acquiring and formatting the model input data and running the model. To use the data acquisition and formatting notebook, the user must provide as input a raster of the Köppen-Geiger Climate Classification (Beck et al., 2018), a comma-separated values file of phytosanitary capacity scores (if applicable), a global, binary raster of host presence and absence, and an environmental file with information on where to store the model outputs. All other data are acquired and formatted within the notebook workflow. Another notebook is used to configure the desired model parameters, scenario configurations, and number of iterations. The model is

run within the notebook and the results are saved locally. Additional notebooks are also available for reading the model outputs and creating result summaries and plots.

S4. Visualization Dashboard

The PoPS Global visualization dashboard (available at https://pops-global.ngrok.io/) provides a set of interactive visualizations to help users gain a better understanding of the mechanics and output of the PoPS Global model (Figure S2). Users can select from geographic, network or temporal visualizations to explore the quantity, location, and timing of predicted pest introductions between scenarios. These visualization types are accessible from tabs across the top of the dashboard and include options to explore both individual stochastic realizations and aggregations of all stochastic realizations. The dashboard linked above provides a demonstration of the case study presented in this report and is not intended for making pest management decisions.

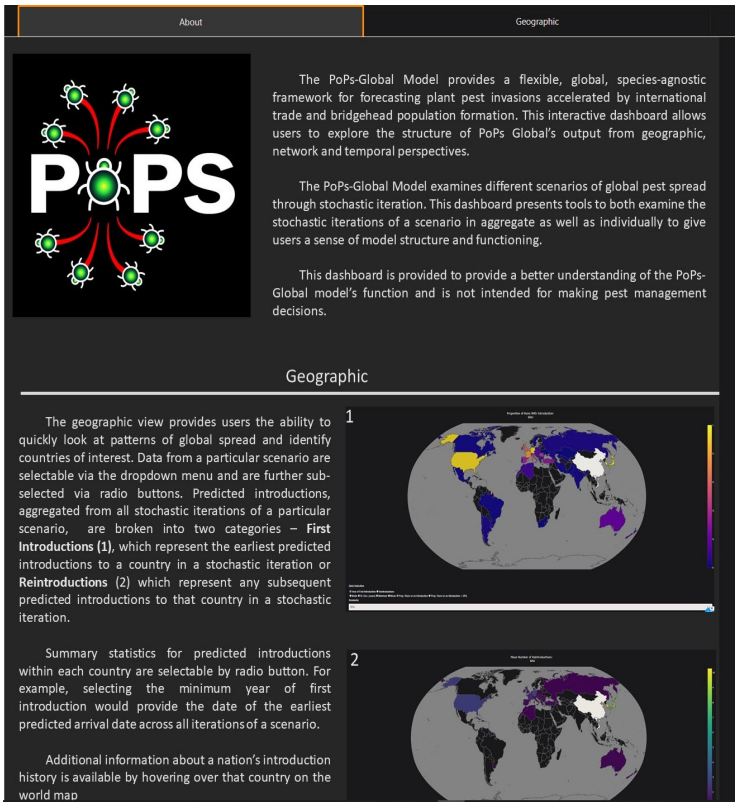


Figure S2. PoPS Global dashboard About tab describing the modeling framework and dashboard functionality.

S4.1 Geographic Tab

The geographic view provides users the ability to explore the patterns of global spread and identify countries of interest. Data and results, including those specific to different scenarios, are accessible via drop-down menus and filtered using radio buttons. Predicted introductions, aggregated from all stochastic iterations of a particular scenario, are broken into two categories accessible from the Data Selection filter under the map: Year of First Introduction (Figure S3), which represent the earliest predicted introductions to a country in a stochastic iteration or Reintroductions (Figure S4) which represent any subsequent predicted introductions to that country in a stochastic iteration.

Summary statistics for predicted introductions within each country are accessible through radio buttons in the Data Selection filter. For example, selecting the minimum year of first introduction would provide the date of the earliest predicted arrival date across all iterations of a scenario. Additional information about a nation's introduction history is available by hovering over that country on the world map.

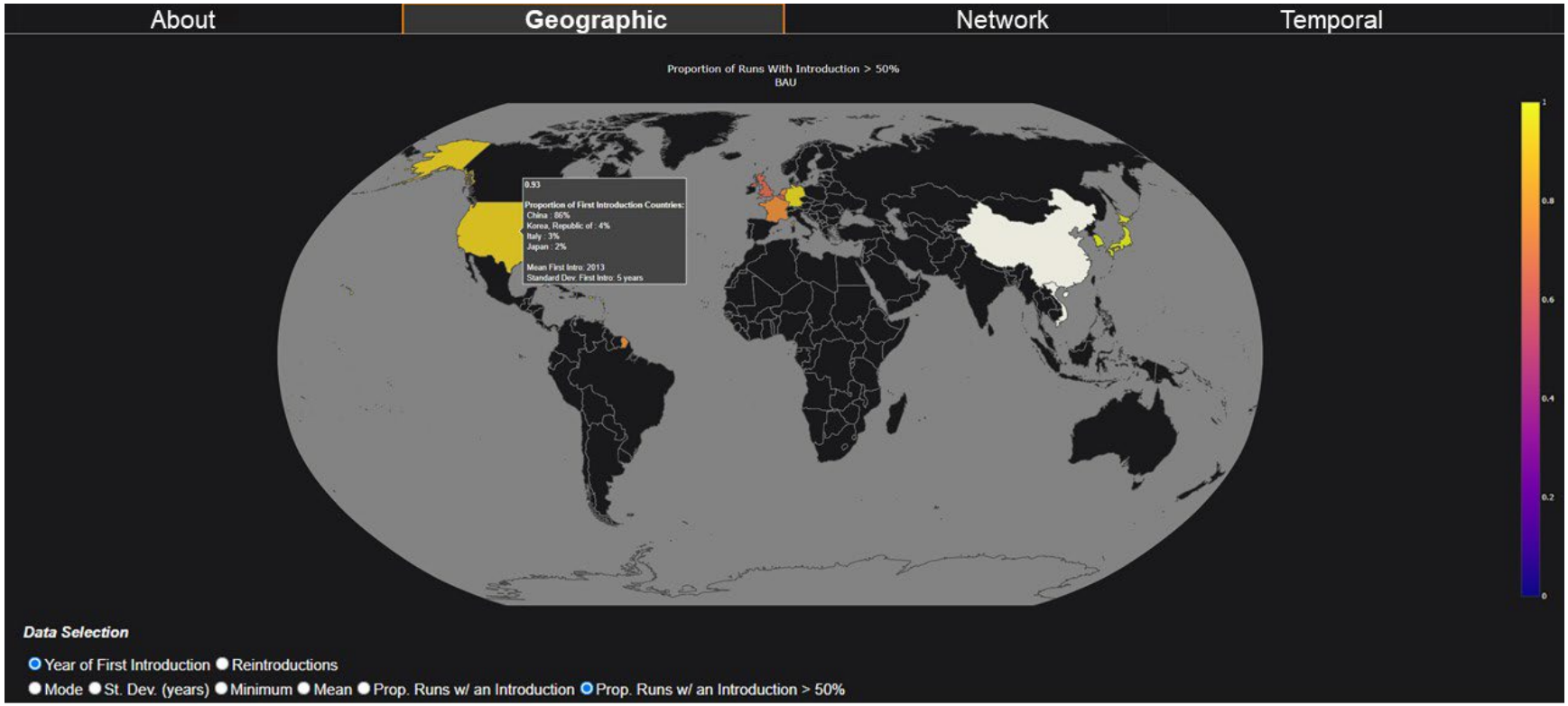


Figure S3. The year of first introductions visualization accessible from the geographic tab of the PoPS Global dashboard.

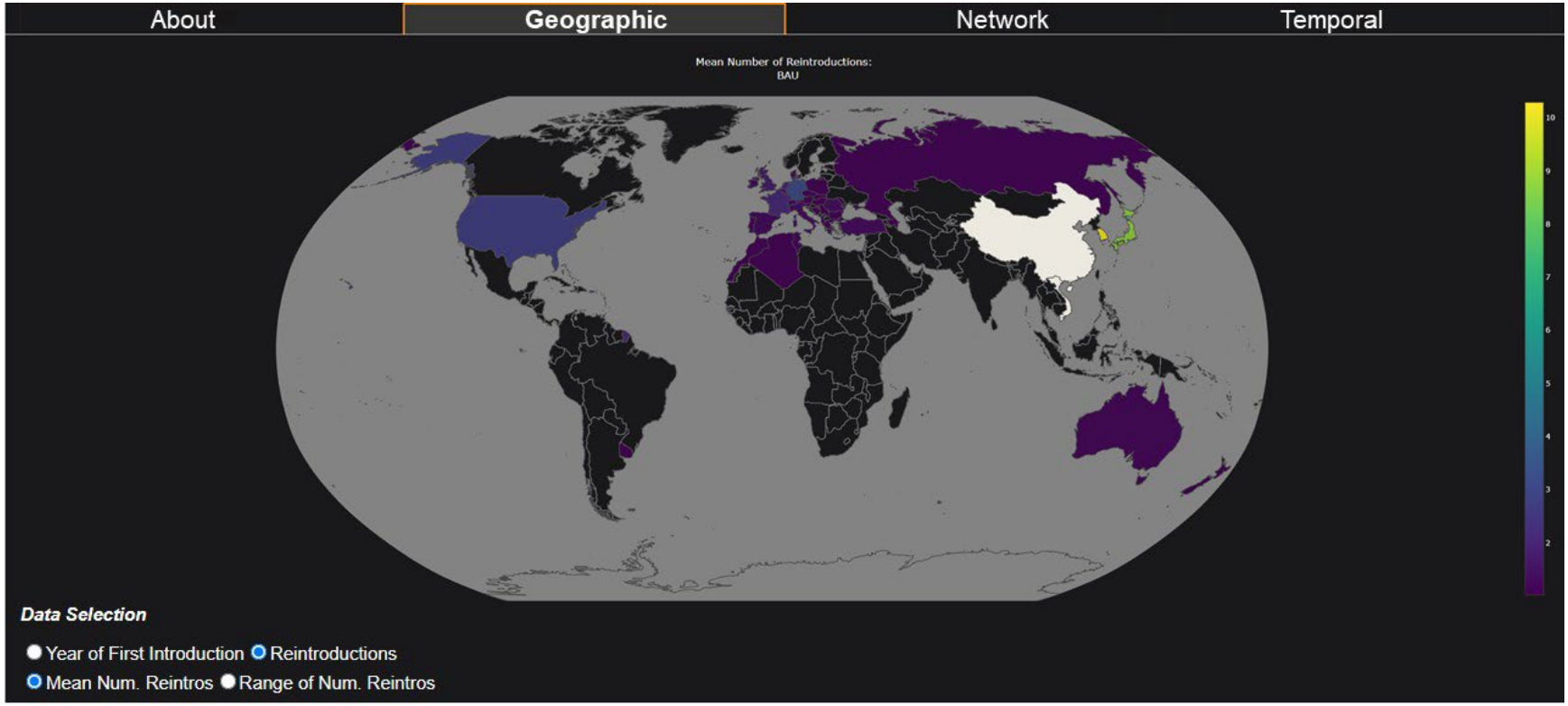


Figure S4. The mean number of reintroductions visualization accessible from the geographic tab of the PoPS Global dashboard.

S4.2 Network Tab

The network graph view removes the geographic component of the data and displays predicted introductions as a graph, with nodes (i.e., circles) representing predicted locations of pest arrival and edges (i.e., lines) representing predicted introductions between countries. A secondary row of tabs provides the ability to select visualizations associated with either All Introductions or First Introductions.

S4.2.1 All Introductions Sub-Tab

The All Introductions sub-tab provides a networked approach to viewing all stochastic runs of a scenario in aggregate (Figure S5). Edges are weighted based on the total number of predictions along that edge in all stochastic iterations of a scenario. Nodes in the default view are organized in shells, with the innermost shells having the highest degree centrality, or the greatest degree of connectedness between all predicted nodes in the network. Nodes are also colored by degree centrality, with the most central nodes having the brightest colors. Nodes bordered in white were used as source countries for the simulation.

Because of the stochastic nature of the PoPS Global model, many countries appear infrequently, representing lesser impact with regards to the predicted global spread of a pest. To focus on more certain or nodes more likely to receive and/or transmit the pest, methods for filtering countries by edge weight are available below the visualization: node centrality and the proportion of runs in a scenario in which that node appears. Figure S6 illustrates the network when nodes with fewer than 30 edges (across all 1000 stochastic realizations) are filtered from the display and Figure S7 is further filtered to only nodes with introductions predicted in at least 50% of the stochastic realizations.

Clicking on any node using the highlight display style (Figure S8) (default) highlights the edges relative to the selected node in the context of the global network. Orange edges represent predicted import to the selected node and purple edges represent predicted transmissions from the selected node. While the node is selected, histograms also display on the right showing the timing of all predicted introductions and transmissions from all sources for the selected node. Clicking on any edge in this view filters the histogram to predicted introductions along that edge (i.e., selected node and node connected by the selected edge).

Clicking on a node using the focus display style (Figure S9) removes other network connections, with the ability to click on remaining edges to view a histogram of predicted introductions along that edge. Clicking on a peripheral node in focus mode highlights the relationship between the original focus node and the peripheral node, with the ability to view the peripheral node's connections (Figure S10).

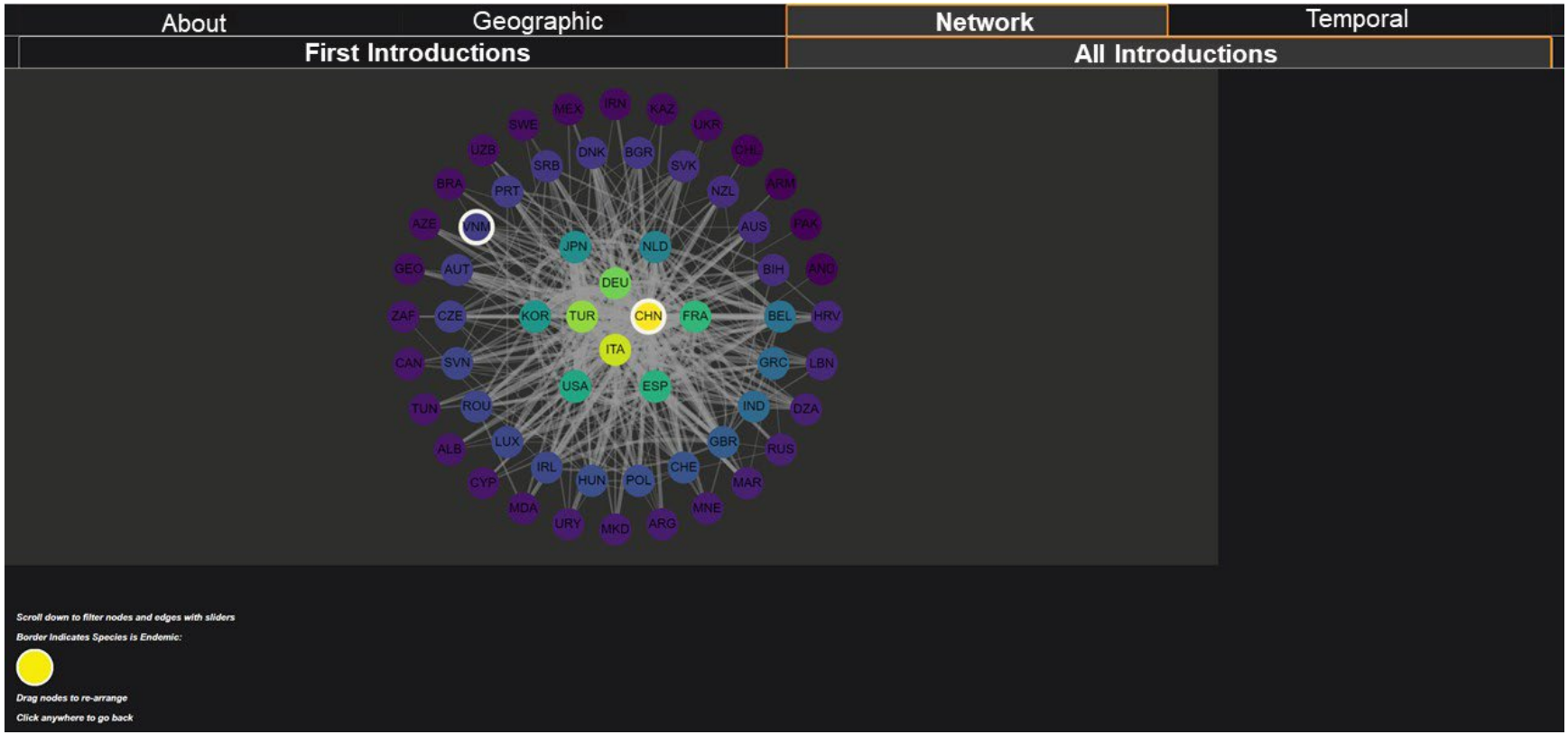


Figure S5. Network graph visualization of the PoPS Global framework. Width of connections between nodes are based on the total number of predictions. Node colors and arrangement represent degree of centrality, or connectedness within the network. Nodes bordered in white are source countries for the simulation.

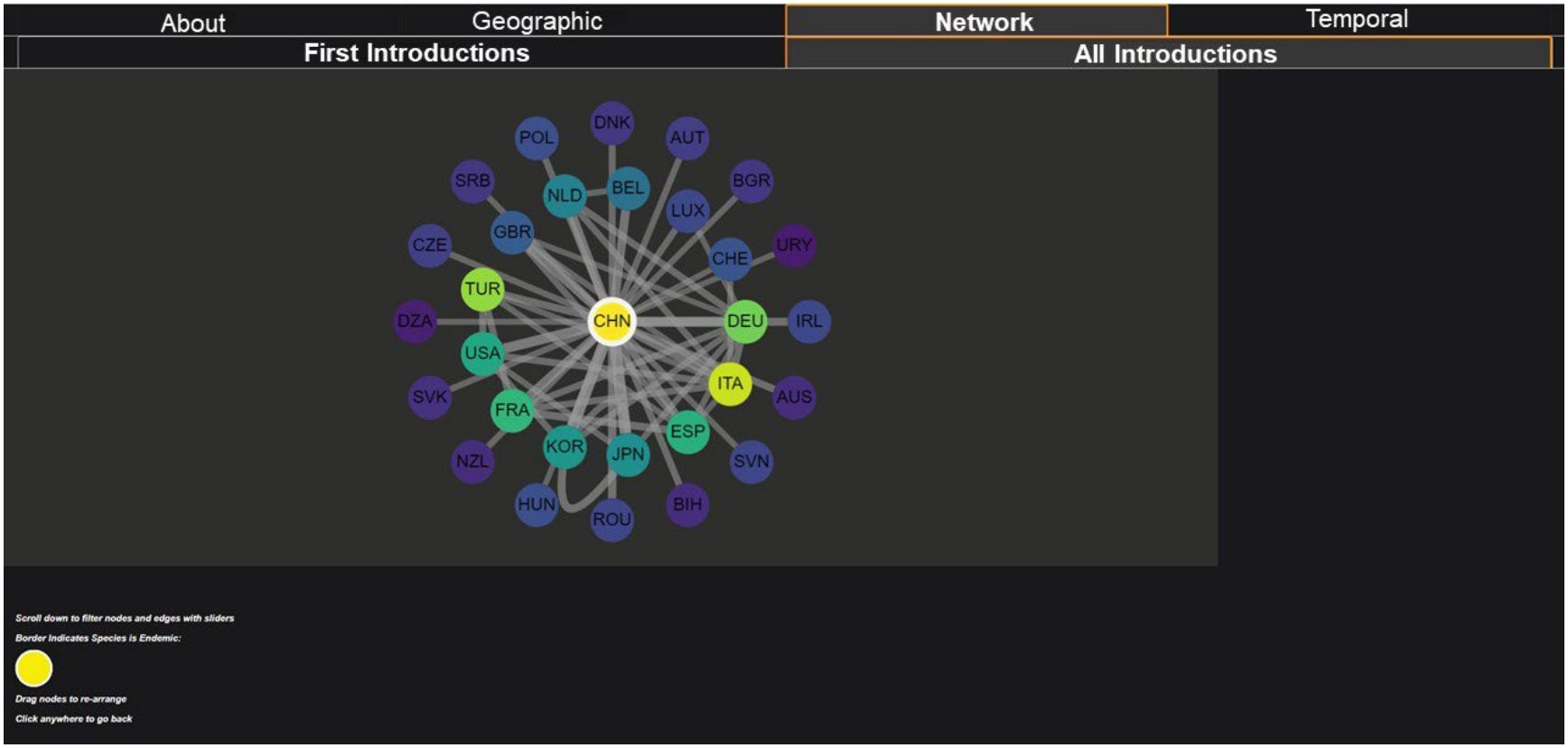


Figure S6. Network graph visualization of the PoPS Global framework, filtered to nodes with at least 30 edges. Width of connections between nodes are based on the total number of predictions. Node colors and arrangement represent the degree of centrality, or connectedness within the network. Nodes bordered in white are source countries for the simulation.

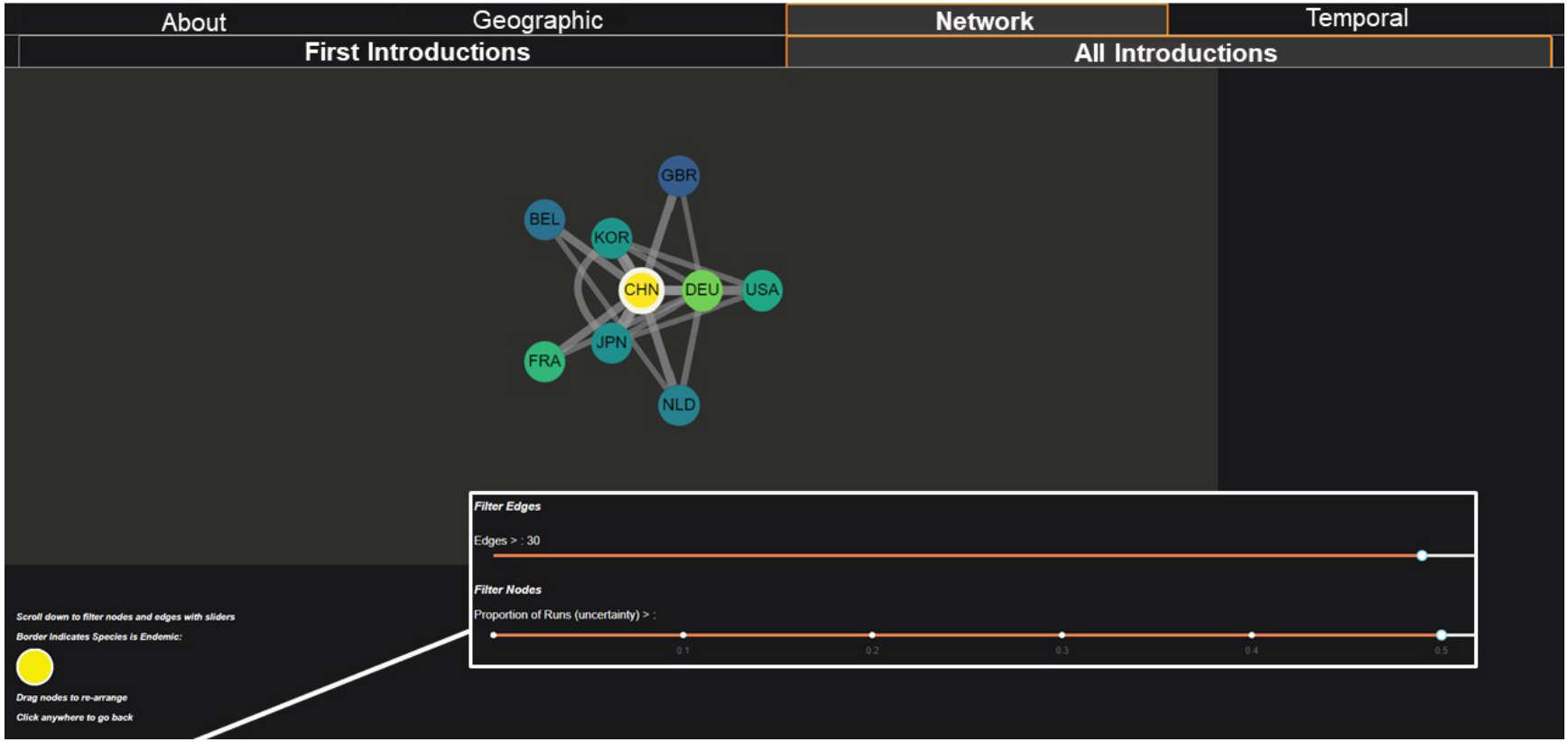


Figure S7. Network graph visualization of the PoPS Global framework, filtered to edges with at least 30 edges and with a predicted introduction in at least 50% of the stochastic realizations. Width of connections between nodes are based on the total number of predictions. Node colors and arrangement represent a degree of centrality, or connectedness within the network. Nodes bordered in white are source countries for the simulation.

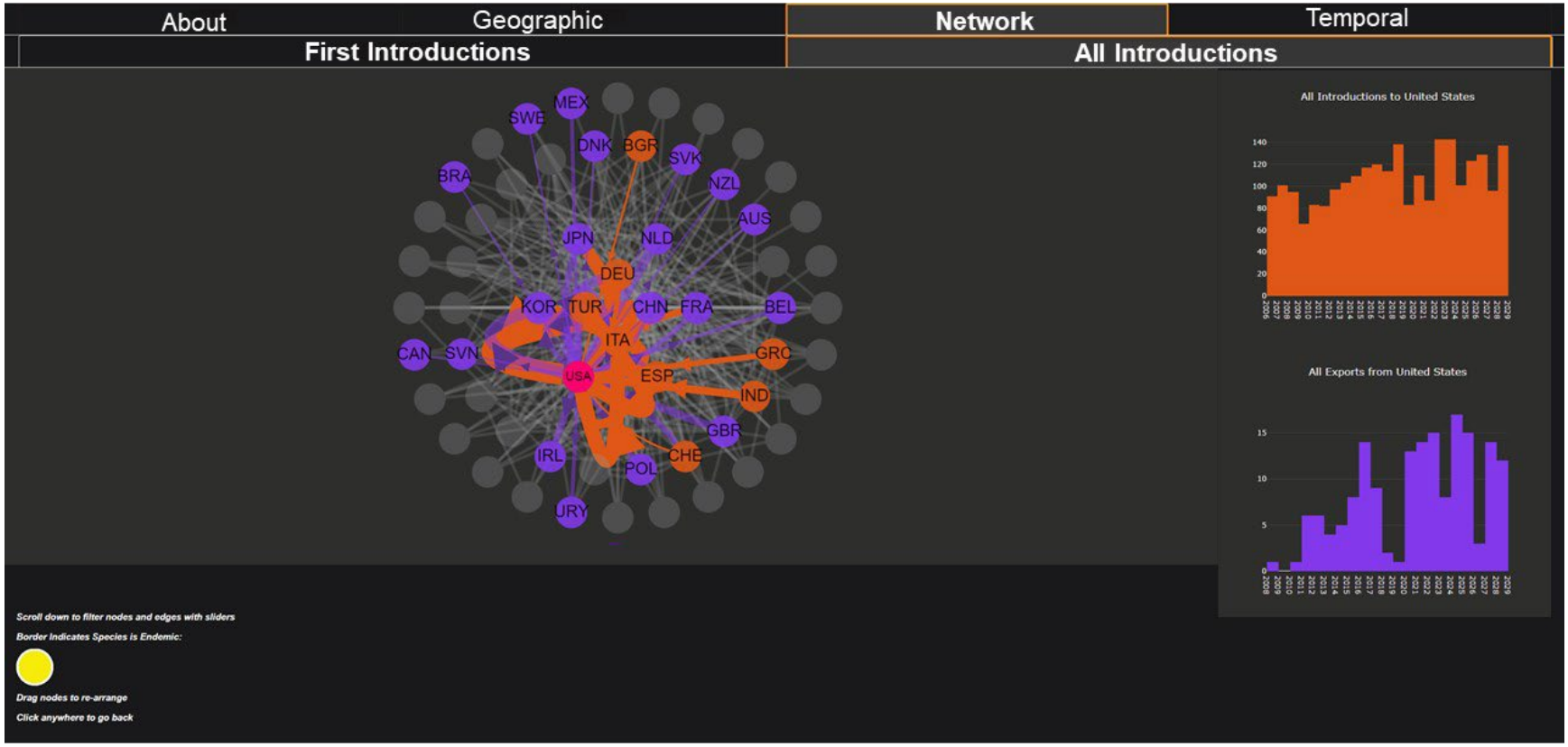


Figure S8. Example graph view using the highlight display style that shows predicted introductions to the selected node (identified in pink) in orange, and predicted transmissions from the selected node in purple. Histograms on the right display timing of predicted movement.

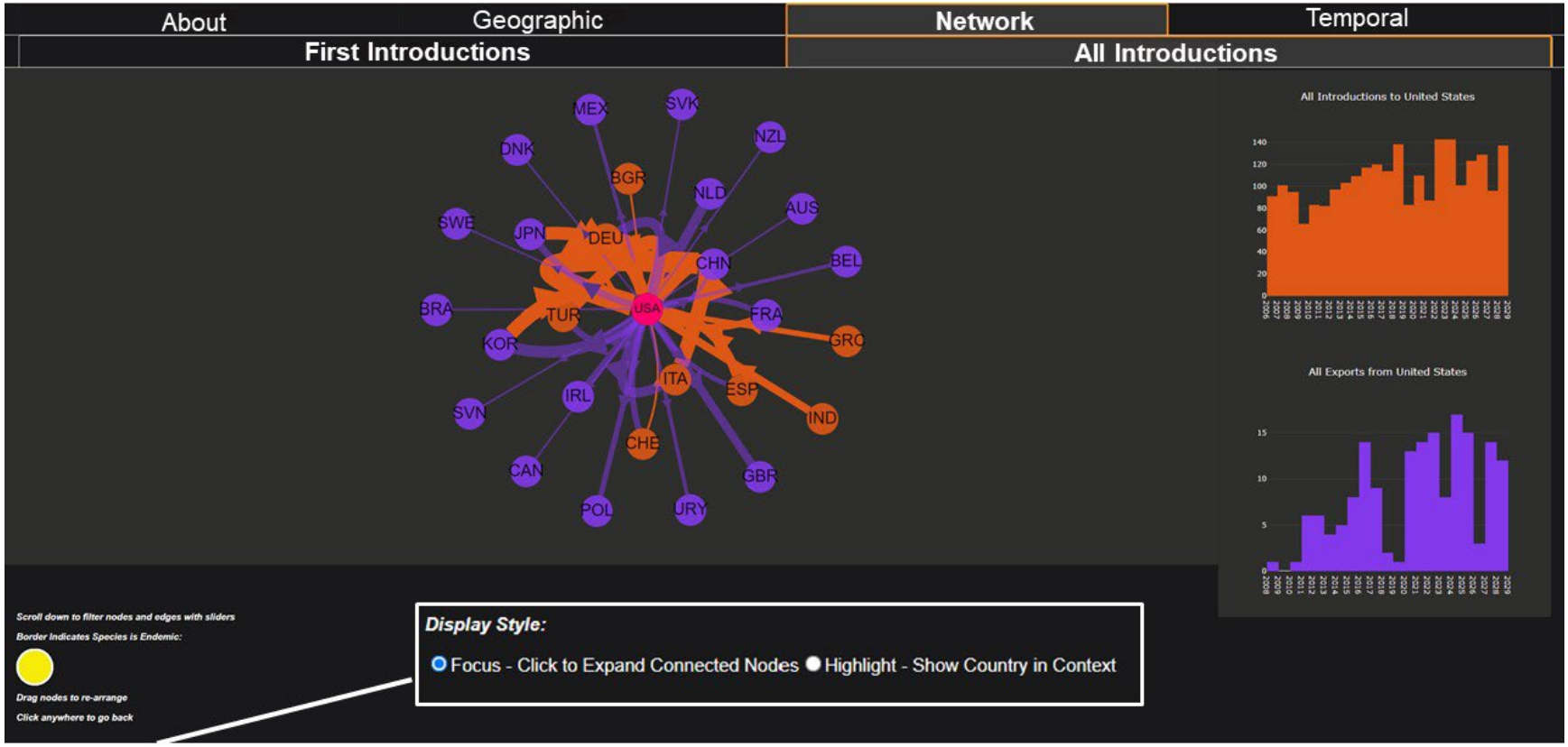


Figure S9. Example graph view using the focus display style, which retains only nodes with a connection to the selected node (identified in pink). Histograms highlight timing of predicted introductions to (in orange) and transmissions from (in purple) the selected node.

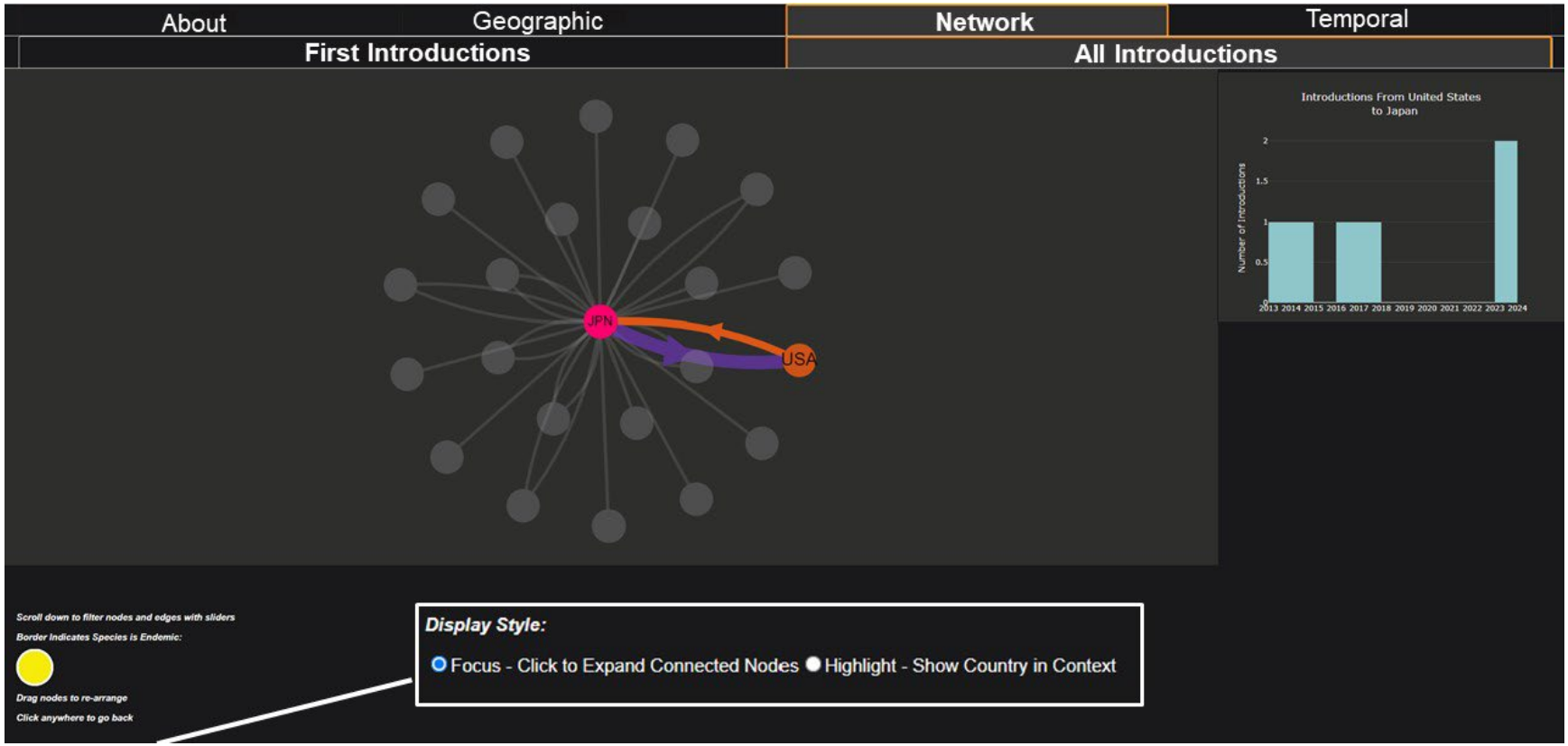


Figure S10. Example of focus display style drilled down to a specific edge and node pair. The orange node is the original selected node and the pink node is now the peripheral node of interest. Gray edges and nodes represent connections to the peripheral node of interest. A histogram on the right depicts the timing of introductions from the originally selected node to the peripheral node of interest.

S4.2.2 First Introductions Sub Tab

In addition to aggregated views of all stochastic realizations, it is possible to visualize individual runs of the model output through the First Introductions sub-tab. The first introductions trees offer a subset of iterations from each scenario, filtered to the direct pathways from source nodes at the top of the graph, to all predicted countries in that realization. In this view, nodes are colored by the predicted probability of introduction. These individual trees allow the user to quickly visualize the stochasticity between iterations and compare the general predicted patterns of spread using the multiple iterations (Figure S11) view, with a slider to progress temporally through that realization. More information about nodes is available on hover.

The individual iterations view (Figure S12) allows the user to analyze specific iterations from the model to understand how the earliest scenario of spread might unfold, with a series of descriptive statistics available at the bottom of the dashboard.

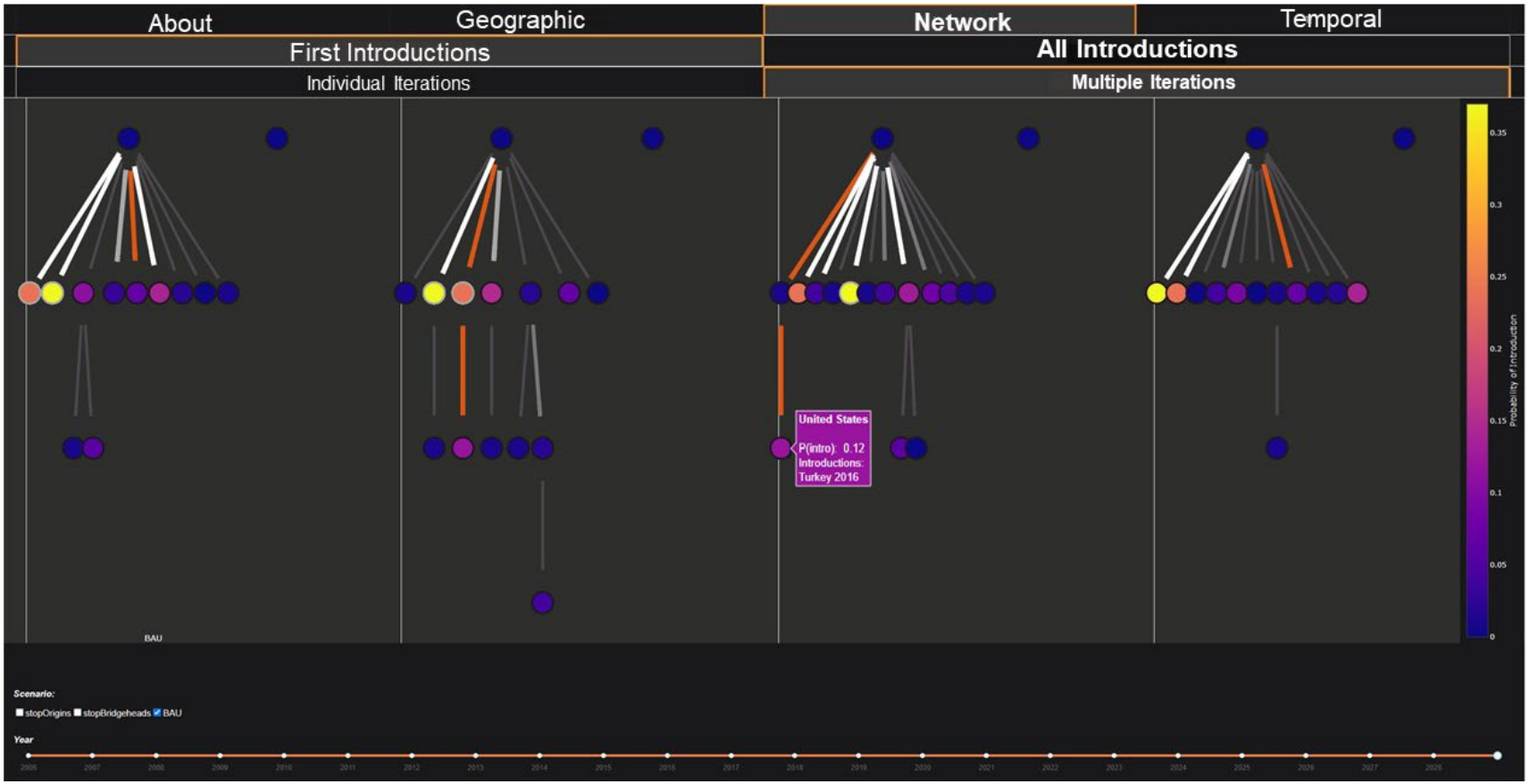


Figure S11. Example of first introduction trees comparing a subset of stochastic realizations. Nodes are symbolized by probability of introduction and the path to the United States is highlighted in orange. Hovering over a node displays additional information.

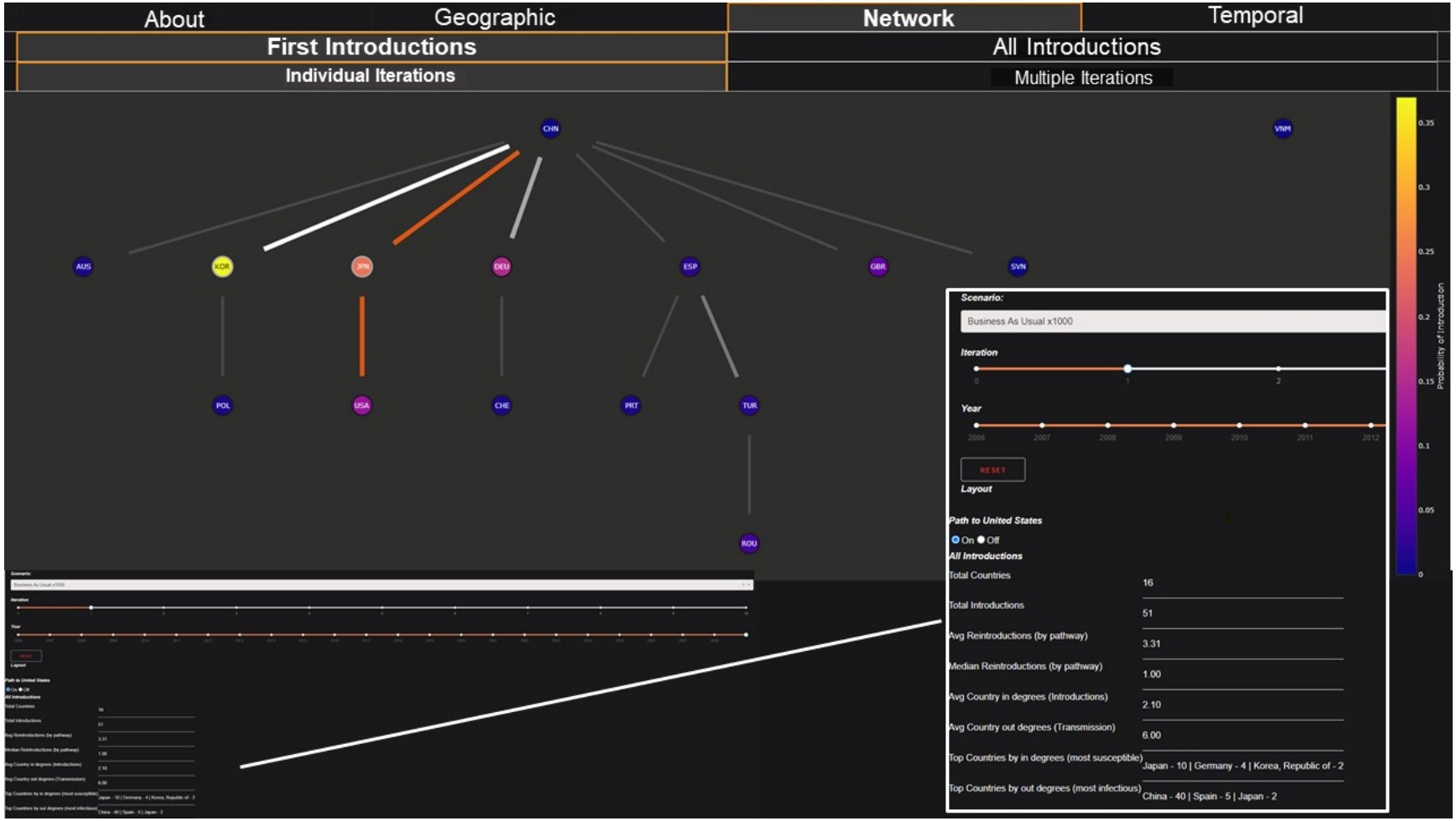


Figure S12. Example of the individual iteration view, which drills down to a single stochastic realization. Nodes are symbolized by probability of introduction, with the path to the United States highlighted in orange. Below the graph are radio sliders to progress temporally or switch to a different stochastic realization, as well as network statistics.

S4.3. Temporal Tab

The temporal view offers comparisons between scenarios, so the user can visualize the impact of scenario-based interventions (e.g., trade restrictions). on the timing and extent of global pest spread. Available views include Introductions per Timestep (Figure S13) and average number of introduced Countries per Timestep (Figure S14).

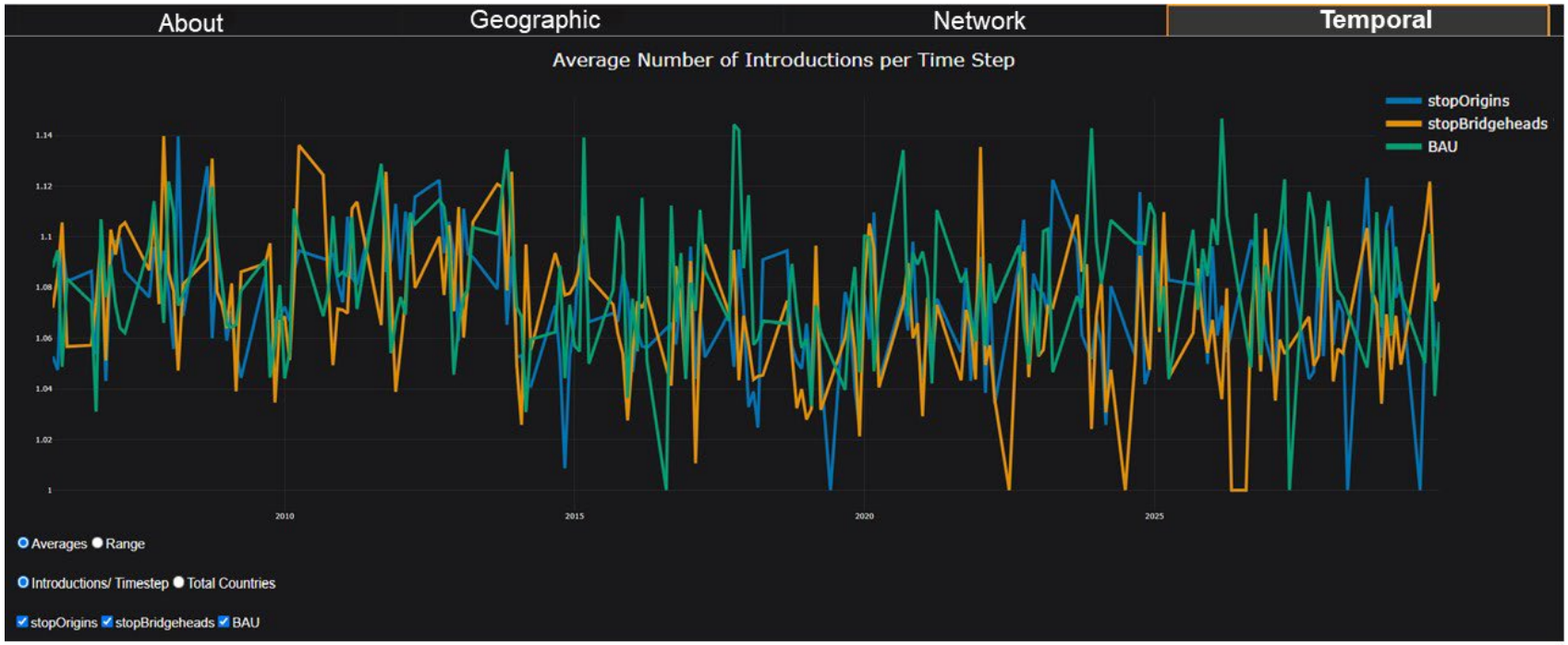


Figure S13. Example visualization of the average number of introductions per time step and management intervention scenario.

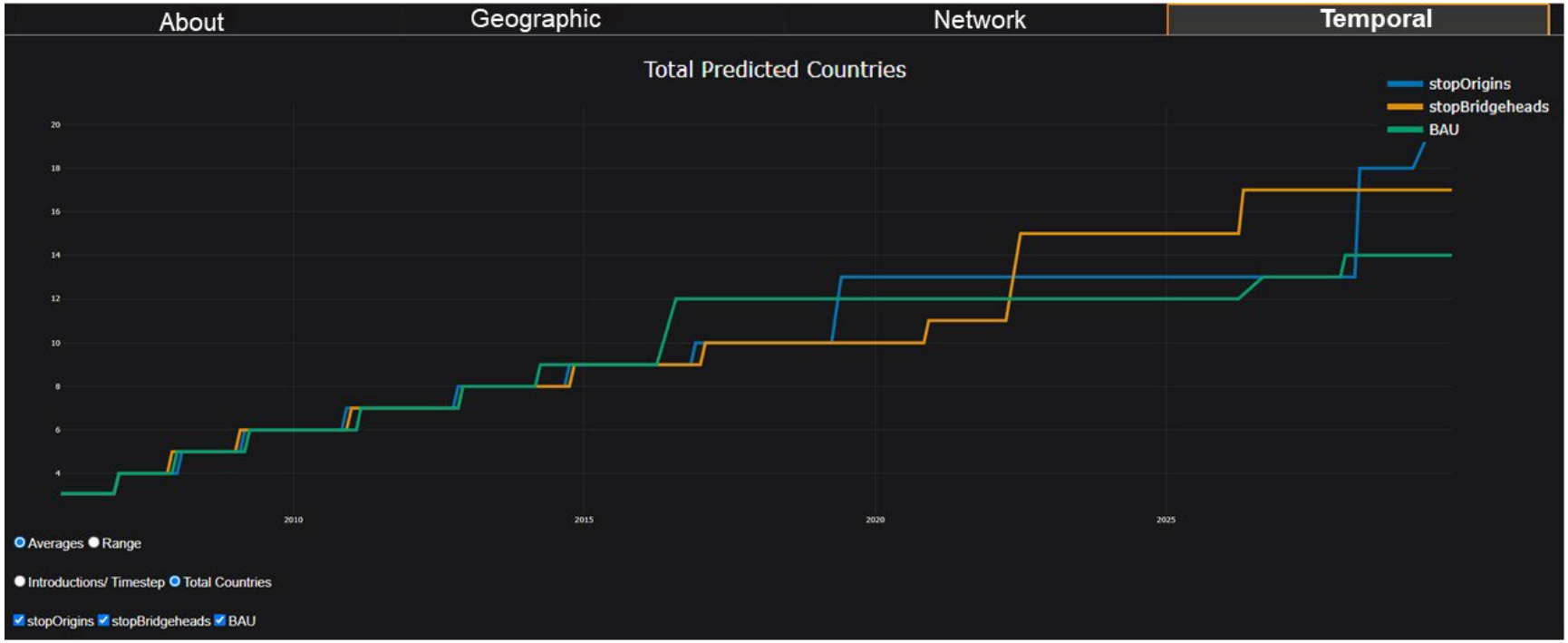


Figure S14. Example visualization of the average total number of countries with a predicted introduction over time and by scenario.

REFERENCES

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 180214. https://doi.org/10.1038/sdata.2018.214

CABI. (2021). *Invasive Species Compendium*. www.cabi.org/isc

Dara, S. K., Barringer, L., & Arthurs, S. P. (2015). *Lycorma delicatula* (Hemiptera: Fulgoridae): A New Invasive Pest in the United States. *Journal of Integrated Pest Management*, *6*(20). https://doi.org/10.1093/jipm/pmv021

DESA/UNSD. (2019). *United Nations Comtrade database*.

Early, R., Bradley, B. A., Dukes, J. S., Lawler, J. J., Olden, J. D., Blumenthal, D. M., Gonzalez, P., Grosholz, E. D., Ibañez, I., Miller, L. P., Sorte, C. J. B., & Tatem, A. J. (2016). Global threats from invasive alien species in the twenty-first century and national response capacities. *Nature Communications*, *7*(1), 12485. https://doi.org/10.1038/ncomms12485

Enquist, B. J., Condit, R., Peet, R. K., Schildhauer, M., & Thiers, B. M. (2016). *Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity* [Preprint]. PeerJ Preprints. https://doi.org/10.7287/peerj.preprints.2615v2

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Food and Agriculture Organization of the United Nations. (2020). *FAOSTAT Database*. FAO. http://www.fao.org/faostat

GBIF. (2021). *What is GBIF?* https://www.gbif.org/what-is-gbif

Lee, D.-H., Park, Y.-L., & Leskey, T. C. (2019). A review of biology and management of *Lycorma delicatula* (Hemiptera: Fulgoridae), an emerging global invasive species. *Journal of Asia-Pacific Entomology*, *22*(2), 589–596. https://doi.org/10.1016/j.aspen.2019.03.004

Maitner, B. S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S. M., Guaderrama, D., Hinchliff, C. E., Jørgensen, P. M., Kraft, N. J. B., McGill, B., Merow, C., Morueta-Holme, N., Peet, R. K., Sandel, B., Schildhauer, M., Smith, S. A., Svenning, J.-C., Thiers, B., … Enquist, B. J. (2018). The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods in Ecology and Evolution*, *9*(2), 373–379. https://doi.org/10.1111/2041-210X.12861

Monfreda, C., Ramankutty, N., & Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, *22*(1). https://doi.org/10.1029/2007GB002947

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S. J., Dudík, M., & Schapire, R. E. (2004). A maximum entropy approach to species distribution modeling. *Proceedings of the Twenty-First International Conference on Machine Learning*, 83. https://doi.org/10.1145/1015330.1015412

Seebens, H., Gastner, M. T., Blasius, B., & Franck Courchamp. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters*, *16*(6), 782–790. https://doi.org/10.1111/ele.12111

Seebens, H., Schwartz, N., Schupp, P. J., & Blasius, B. (2016). Predicting the spread of marine species introduced by global shipping. *From the Cover*, *113*(20), 5646–5651. https://doi.org/10.1073/pnas.1524427113

USDA-APHIS-PPQ-CPHST. (2018). *Technical working group summary report spotted lanternfly, Lycorma delicatula (White, 1845)*.

Wakie, T. T., Neven, L. G., Yee, W. L., & Lu, Z. (2020). The Establishment Risk of *Lycorma delicatula* (Hemiptera: Fulgoridae) in the United States and Globally. *Journal of Economic Entomology*, *113*(1), 306–314. https://doi.org/10.1093/jee/toz259