Projecting terrestrial species invasion spread using commodity flow pathways

Ashish Gauli,[1,2] Nathan Wikle,[1,3] Ryan Yan,[1,4]
Dr. Louis Gross,[1,5] Dr. Daniel Simberloff,[5] Angela Chuang,[5] Cedric Landerer[5]

[1]National Institute for Mathematical and Biological Synthesis (NIMBioS)
[2]Computer Science, Fisk University
[3]Mathematics, Truman State University
[4]Mathematical Biology, College of William and Mary
[5]Ecology and Evolutionary Biology, University of Tennessee

July 31, 2015

Abstract: Models to create projections of invasive species spread are critical in identifying high priority shipping centers and pathways of bioinvasion, essential information in the prevention of environmental and economic harm caused by invasive species. Research into the development of models of biological invasion across broad spatial regions have been targeted towards network models of international shipping pathways. The majority of models for terrestrial invasions have been confined to basic diffusive spread. Nevertheless, many species of harmful pests are known to travel long distances, in part due to transport through commodity shipments. For example, *Solenopsis invicta*, the red imported fire ant, can travel via turf shipments within the United States. The availability of commodity flow data suggests that a model utiliz-

1

ing domestic freight as the dominant vector of dispersal may be a promising, yet unexplored, avenue of research in invasive species ecology. To this end, we utilize data from the Freight Analysis Framework (FAF) to create projections of invasive species spread. The FAF data provides tonnage information on commodities shipped within the United States at state level resolution. Weight values from four sets of data are used in a resampling method to generate the transition matrix at the core of a stochastic matrix model used to create projections of terrestrial invasive species spread. Additionally, we incorporate elements of environmental difference and logistic growth in our projections to achieve greater biological soundness in the results. Furthermore, we introduce ComFlo, an interactive web-based tool designed to aid in the visualization of commodity shipments within the United States.

Introduction

Modern societies are increasingly dependent upon a complex network of global and national trade which facilitates the quick transport of all types of goods. Although this interconnectedness has led to the flourishing of today's global economy, the increased volume of trade has had serious consequences for ecological health worldwide. In particular, human commerce is the foremost vector for the global dissemination of invasive species, which has resulted in ecosystem modification and potentially large economic costs to the invaded regions (Mack et al., 2000). Understanding the role of commodity movement in the spread and introduction of invasive species is a crucial step in their prevention and control. In the long term, we aim to supply management agencies with the information needed to control the spread of invasive species.

Due to its prevalence and outsized role as a conduit for invasive species, the global maritime shipping network has been thoroughly analyzed to identify common hotspot ports and pathways

of invasion. Maritime shipping alone accounts for over 90 percent of the world's trade (Kaluza et al., 2010), and careful consideration has been taken to characterize its role in the spread of foreign species. Data from the Automated Identification System (AIS) technology, a real-time tracking system installed on most large ships, has allowed the construction and analysis of the global shipping network by Kaluza et al. (2010). This network's characterization as a small-world and scale-free network suggests that it is particularly prone to the rapid dispersal of invasive species along its edges (Pastor-Satorras and Vespignani, 2001). Indeed, ballast water discharge and hull fouling of freight vessels are directly responsible for the transfer of many marine organisms (Ruiz et al., 2000). And yet, transported species are not limited to marine species. Nonindigenous plant pest species have also been found regularly in cargo entering the United States (McCullough et al., 2006; Work et al., 2005), and even more species enter undetected (Work et al., 2005).

As a result, shipping pathway risk assessments are needed for efficient prevention and management of foreign species invasions (Hulme, 2009). Research in this area has largely been devoted to studying the global shipping network. For example, Drake and Lodge (2004) applied a gravity model to identify probable pathways of marine invasion and identified specific hotspot ports most likely to distribute invasive species. Similarly, Seebens et al. (2013) identified hotspots of bioinvasion using maritime shipping data, and extended their model to consider port environmental conditions and biogeography when identifying probable pathways. In another study, Paini and Yemshanov (2012) used a Markov chain model to identify the most probable countries of origin for the invasion of the Khapra beetle to Australian ports. Increasingly, network models have been used to investigate international species transport, however, the expansion of these networks to domestic modes of transportation has been limited (Hulme, 2009).

The role of domestic freight shipping as a method of terrestrial invasive species dispersal has

been largely confined to an analysis of the domestic movement of imports entering the United States and Canada (Koch et al., 2011; Colunga-Garcia et al., 2009; Yemshanov et al., 2013). Using domestic shipping data in conjunction with import data, the risk of foreign invasion to urban forest areas was assessed for specific regions in the United States. Establishment rates were found to be low overall, except in a small subset of urban areas with high population density and a high volume of imports (Colunga-Garcia et al., 2009; Koch et al., 2011). Similar analysis of Canadian imports reveals that establishment rates are lower, on average, than those found in the United States (Yemshanov et al., 2013). Although key points of entry and invasion pathways were established for exotic pests, in all cases, the focus of these studies was on the establishment of new foreign species via imports; the dispersal of already established invasive species across the United States was never discussed.

Research into domestic freight as a mode of species dispersal seems promising for multiple reasons. First, there exists a substantial database on domestic commodity shipping within the United States. Based largely on data collected by the U.S. Census Bureau in their Commodity Flow Survey (CFS), the Freight Analysis Framework (FAF) database contains aggregated data on freight shipments in the United States, separated by commodity type and mode of transportation (Southworth et al., 2011; FHWA, 2015). The data are regularly used by state and regional departments of transportation in the creation of freight forecasting models (Board, 2008; Chow et al., 2010; Sivakumar and Bhat, 2002), however, their potential value in bioinvasion literature remains untapped. Second, it is widely believed that particular species are transported through domestic commodity shipments. For example, the USDA holds strict regulations on the shipment of soil from metropolitan areas known to contain the Red Imported Fire Ant (USDA, 2010). Thus, a network model of commodity flow in the United States could be utilized to identify hotspots and pathways of invasion and provide a method to assess the impact of enhanced resources devoted to inspection.

4

This idea has motivated our work, and has led to the development of the following goals:

1. Develop a tool to visualize the flow of commodities within the United States using data obtained from the Federal Highway Administration's Freight Analysis Framework.

2. Create projections of terrestrial invasive species spread using a stochastic model where dispersal occurs through a commodity freight network.

3. Use the aforementioned model to perform risk analyses to better assist managers in considering alternative plans for controlling invasive species spread.

Methodology

Freight Data

Shipping data were obtained from the Freight Analysis Framework (FAF), a database that combines data sources to create a complex freight movement summary of terrestrial commodity shipments between geographic regions within the United States. In particular, flow matrices have been created for 1997, 2002, 2007, and 2012, with additional flow forecasts from 2015 to 2040. The 2007 and 2012 data contain 123 domestic regions, 74 of which are metropolitan areas, 33 the remainders of particular states, and 16 entire states, where no FAF metropolitan regions exist. There are 43 commodity classes, as classified by the Standard Classification of Transported Goods (SCTG). We retrieved FAF data for the following years: 1997, 2002, 2007, and 2012. Each line of data records an origin and destination, a particular commodity, mode of transportation, value of goods shipped, weight in kilotons, as well as metrics involving distance traveled. An extensive review of FAF methodology can be found in Southworth et al. (2011).

An effort was made to characterize the large-scale domestic commodity shipping networks obtained from the FAF database. Basic network topological features have been used to characterize many existing real-world networks, including the internet, the World Wide Web, bi-

ological networks (intracellular biochemical pathways, genetic regulatory networks, etc.), and a variety of transportation networks (Barrat et al., 2008). These complex networks are often defined as having small-world and scale-free properties, an indication of the presence of important high-degree nodes known as hubs. In particular, the worldwide air-transportation network (WAN) has been found to have heavy-tailed, power-law degree and strength distributions (Barrat et al., 2004; Guimerá and Amaral, 2004), while the global cargo-ship network (GCSN) has a right-skewed degree distribution, and power-law weight and strength distributions (Kaluza et al., 2010). Because the WAN and GCSN are large-scale transportation networks, it was expected that the domestic shipping networks obtained from the FAF would have similar network characteristics - in particular, that the FAF networks would have small-world and scale-free properties.

However, an initial analysis of the FAF networks reveals a less structured topology. As an example, consider the network of agricultural products transported by truck between the 123 FAF designated metropolitan regions of the United States. A small-world network has the property that any two vertices can be traveled between by passing through a small number of intermediate vertices, while also requiring the existence of highly-clustered nodes. This property can be measured by examining the relationship between the average shortest path length $\langle l \rangle$ and $M(l)$, the average number of nodes within a distance less than or equal to l from any given vertex. $M(l)$ is expected to increase exponentially in a small-world network (Barrat et al., 2008). As shown in Figure 1, any two nodes in the agricultural commodity network can be reached within two steps. However, because nodes in this network are highly connected ($\langle k \rangle = 80$), it is unclear if the observed relationship is a result of node clustering.

The agricultural network has certain properties that conform with scale-free assumptions, but not others. The degree distribution $P(k)$ for the network is shown in Figure 2a. Notice that there is an increase in $P(k)$ as k increases, and the network shows no evidence of a heavy-

6

tail degree distribution. However, when strength and edge weight distributions are considered, we observe that each is heavy-tailed and approximates a power-law distribution (Figure 2b and Figure 2c). Thus, there appears to be a correlation between the weighted properties of the network and the network topologies. Although these results are specific to the agricultural commodity network, they are have been seen in many of the commodity flow networks found in the FAF database that we have analyzed. We believe that the difference in the degree and weighted property distributions is a characteristic of the highly aggregated node structure of the network, the spatial properties of the existing domestic infrastructure, and variation in the data itself. A more thorough exploration of the statistical and topological properties, as well as dynamics, is needed to correctly further characterize the shipping networks.

A Domestic Shipping Pathway Model

Using a stochastic matrix model framework, we developed a methodology to project terrestrial invasive species spread using commodity pathways. For this model, two important assumptions were made. First, the invasive species of interest must be transported via a particular commodity or set of commodities and a specific mode of transportation. Second, an initial domestic region has already been invaded by the species, and additional invasion can only occur via transport within the United States. This means that additional establishment from foreign imports is no longer considered in the propagation of the species. Thus, the model is restricted to domestic freight movement. However, it could potentially be coupled with a global model if desired.

Because the FAF data were regionally aggregated to individual U.S. states in 1997 and 2002, our model considered the spread of a hypothetical invasive species within 51 geographic areas: the fifty states and Washington D.C. If higher resolution data are available, this can be easily extended to m regions. A state vector $\mathbf{s_t}$ represents the fraction of land area within each region in which that species is present, in year t. More specifically, s_{t_i} is the fraction of land area in

7

region i, in year t, in which the particular invasive species is present, for $i = 1, 2, ..., m$. Thus, at time $t = 0$, the invaded regions should be initialized to 1, and all susceptible regions to 0.

In a given year t, it is assumed that the invading population will continue to grow in each region i until the species has become fully established. Thus, a logistic growth function g is applied element-wise to $\mathbf{s_t}$. This is defined as

$$g(s_{t_i}) = \frac{s_{t_i} e^r}{1 + s_{t_i}(e^r - 1)}$$

where $i = 1, 2, ..., m$ and r is the annual growth rate. Function g is a monotonically increasing function which reaches a horizontal asymptote at one. We have arbitrarily chosen r to be 0.01, however, it is recommended that an estimated growth rate be used for the species in question.

Next, a transition matrix $\mathbf{P_t}$ is created for year t from the FAF data. Each element $p_{t_{ij}}$ represents the proportion of a given commodity produced in state j that is being exported to state i. More specifically,

$$p_{t_{ij}} = \frac{w_{t_{ij}}}{\sum_{i=1}^m w_{t_{ij}}},$$

where $w_{t_{ij}}$ is the weight of commodity shipped from state j to state i in year t. Because the FAF data consists of four sampling years (1997, 2002, 2007, and 2012), variation in the amount of commodity produced in each state is added to the model using a resampling method. For each year t, the weights of the commodity produced in state j are randomly chosen from the four sets of available data. Therefore, $\mathbf{P_t}$ is most likely not equal to $\mathbf{P_{t+1}}$ and the transition matrix no longer remains constant from one year to the next. However, it maintains its Markov property. To account for this added variability and to generate the most likely projections of invasive spread, we average the simulation results over many repeated runs of the simulation.

To account for the reduced establishment potential of a species outside of their natural climatic range, we incorporate an environmental similarity matrix $\mathbf{E}$ into our model. For this preliminary model, we utilize a linear combination of two environmental variates, temperature

8

and precipitation, when calculating the similarity matrix.

Environmental data were retrieved from the NOAA FTP service on 30 year normals of mean temperature and precipitation recorded at stations across the United States (NOAA). Each row of data, representing one station, contains one value (the 30 year normal) for each month of the year. We aggregated the data to the state level, resulting in a matrix with 51 rows, representing all states and Washington D.C. The Euclidean distance between each of the vectors was calculated, and used to populate a 51 x 51 distance matrix $\mathbf{D}$. Each element D_{ij} signifies the distance in temperature or precipitation vector space between two states, and as a result $\mathbf{D}$ is symmetric. This matrix is used to create the environmental similarity matrix $\mathbf{E}$. Each element E_{ij} was calculated by standardizing the distance matrix to values between 0 and 1, thus creating a dissimilarity matrix, and then subtracting the resulting value from one, creating the similarity matrix. Specifically,

$$E_{ij} = \frac{D_{ij} - min\left(D_{ij}\right)}{max\left(D_{ij}\right) - min\left(D_{ij}\right)},$$

where $min(D_{ij})$ and $max(D_{ij})$ refer to the minimum and maximum-valued elements in matrix $\mathbf{D}$, respectively. This was performed using both the temperature and precipitation data, and a linear combination of the resulting temperature environmental matrix $\mathbf{E^T}$ and precipitation environmental matrix $\mathbf{E^P}$ gives $\mathbf{E} = \alpha\mathbf{E^T} + (1-\alpha)\mathbf{E^P}$. We chose $\alpha = 0.5$, however this value can be adjusted to better reflect the relative importance of temperature and precipitation to the environmental suitability of a region.

Thus, combining these elements, we have

$$\mathbf{s_{t+1}} = (\mathbf{E} \circ \mathbf{P_t})g(\mathbf{s_t}).$$

We define $\mathbf{E} \circ \mathbf{P_t}$ as the Hadamard product, or entry-wise multiplication, of $\mathbf{E}$ and $\mathbf{P_t}$. This process should be repeated for $t = 0, 1, ..., N$, at which point $\mathbf{s_N}$ represents the projected invasive spread after N years.

Risk curves were generated from the model output. Specifically, results were collected from 100 simulations, s_{t_i} was evaluated at each time step for each simulation, and the proportion of simulations in which s_{t_i} surpassed a given threshold level was recorded. This proportion was calculated at each time step for multiple threshold levels, in order to plot a family of risk curves.

Data Visualization

ComFlo, our web-based commodity flow visualization tool, was created in an effort to project the movement of invasive species through domestic commodity shipping pathways in the United States. However, its applications extend to anyone interested in visualizing the commodity shipping data found in the Freight Analysis Framework (FAF) database. The application runs under Shiny Server (Chang et al., 2015), a web application framework developed by RStudio (RStudio Team, 2012), and utilizes R packages implementing relational database systems and mapping. The commodity flow data were obtained from the FAF website, encompassing all available sampling years - 1997, 2002, 2007, and 2012 - each of which contains flow data representing 43 commodity classes and seven modes of transportation. For more information on the FAF methodology, and to obtain data, please visit `http://ops.fhwa.dot.gov/Freight/freight_analysis/faf/index.htm`.

ComFlo features an interactive graphic representation of commodity flow as its centerpiece. The political map of the United States is shown, with individual states outlined in grey. After selecting a state, commodity, mode of transportation, and year, the map will show imports and exports specific to those selections as blue and red lines, respectively, with nodes centered in the geographical center of each state. The width of the line represents the flow of the selected commodity relative to the total flow of the commodity in the United States. Similarly, the size of a node represents the total flow of the commodity within the state, relative to the total flow of the commodity in the United States. Line width, l_{ij}, for a line from state j to state i, is calculated

as

$$l_{ij} = \frac{w_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{m} w_{ij}},$$

where w_{ij} refers to the kiloton weight of the selected commodity shipped from region j to region i. Thus, line width is the relative proportion of the selected commodity shipped throughout the United States.

The user interface of ComFlo consists of different widgets that allow users to interact with the system. SelectBox allows the users to select year, multiple commodity types, multiple transportation modes, multiple states, and the flow direction of the commodities. A leaflet-based map displays the results based on the user selection. A sliding tool allows the users to adjust the width of the lines on the map.

ComFlo's server consists of the FAF database, a geo-location database, an R-based relational database system, and rendering expressions. The FAF database consists of over 10 million data records, including forecasts for 2015 and 2020, obtained from FAF with information on origin state, destination, commodity type, mode of transportation, and shipment weight. When ComFlo is in use, user input is retrieved, and the corresponding dataset is retrieved. Because multiple selection options are included, the user input is parsed as strings, a table or a list, which is then used to make queries over the FAF database using the relational database system. For example, if the user selects 1997, Alcoholic Beverage, Truck, Tennessee, and Export, the output will be exports, in kilotons, originating in Tenessese, heading to all destination states. With respect to the origin and destination, the output dataset is connected with a geo-location database that contains the latitude and longitude of each U.S. state. Appropriate polylines are drawn on the leaflet map and rendered to the user interface, to reflect the real-time data input provided by the user.

Results

In addition to making projections of invasion spread via the commodity pathway model, we developed several distinct but complementary products utilizing the model projections and the data itself.

1. We developed ComFlo, a web tool available for use by the scientific community as well as the general public. This site is accessible through the NIMBioS webpage at `http://rweb.nimbios.org:3838/sreuser/faf-data-visualization-system/`. ComFlo allows users to select origins and destinations within the United States, and displays commodity flows based on the user's choice of commodity and mode of transport. This tool is a significant advancement beyond those currently available freely online and through the FAF website, as similar services do not allow selection of specific commodities or modes of transport. Additionally, ComFlo grants the user the ability to view the source data easily through the table feature, as well as allowing them to export a .csv formatted data file of their selection. These features, used in conjunction, provide users the ability to visualize significant pathways of biological invasion, quantify these pathways, and export the data for further analysis. This is a streamlined research tool beyond what is offered elsewhere. We anticipate the web tool will be useful to researchers wishing to quickly identify high volume shipping pathways that might act as significant pathways of species invasions and hope it will aid them in their work.

2. We created choropleth maps (example: Figure 3) to show the spread of an invasive species over time, given a set of initial conditions. Each map plots the state vector, representing the fraction of land area within each region in which the species is present, for each year in the simulation. The final output is a series of choropleth maps showing the change in the percentage of establishment in each state over the whole simulation runtime. These maps

provide an immediate visual interpretation of the threat and spread rate of a particular species.

Results from these maps are highly dependent on the chosen species, mode of transport, commodity of interest, and species growth rate r. Thus, general trends across species are hard to infer. However, we found that projected establishment of a particular species did not vary greatly when many simulations were averaged, despite the stochastic transition matrix. In addition, states with high population densities were often projected to be invaded. For example, in 3, a choropleth map showing the projected spread of the red invasive fire ant in the United States, California and Texas are shown to be likely destinations of invasion, despite being a large geographic distance from Alabama, the state of origin. This may imply that the likelihood of invasion is dependent upon the domestic freight network of the United States, and a restructuring of distance that includes network topology and dynamics may be better able to project invasion than a standard diffusion model.

3. We developed families of risk curves (Figure 4) for the invasion of a given destination state within a standard time frame. The y-axis measures the probability that an invasive species becomes established beyond a certain threshold level in the state within a given number of years, shown along the x-axis. Each curve charts the risk associated with a different threshold level of establishment. Thus, an environmental manager can assess the risk of increased establishment over a period of time.

Discussion

The ultimate goal of our research is to provide useful information to policymakers to allow for the effective management of invasive species. In our research, we explored the role of

domestic shipping in the transport of invasive species within the United States, and developed projections based on this relationship. This method represents a novel approach to projecting invasive species distributions forward in time. Our model of invasion spread projects both local dissemination as well as long-distance dispersal events via human transport vectors, an attribute unseen in classical diffusive spread models.

It is important to note that we have constructed a general model of spread, which may be applied to a large number of species, given that our stated assumptions hold. The model is used to make projections for specific species, dependent on the input data, the growth parameters, and the initial state vector, which are tuned to make projections for specific species. Thus, confidence in the model's projections should be expected to increase with more precise parameter estimation. This is perhaps most evident in the environmental similarity matrix $\mathbf{E}$. Although the included methodology provides some environmental consideration in the spread of the invasive, a maximum entropy ecological niche model would provide more robust E_{ij} entries and can be expected to contribute more robust projections as a result.

As a result of the model projections, we have produced many tools for projecting and identifying future biological invasions. ComFlo serves as a powerful tool in identifying high volume pathways of commodity flows which may serve as highways of biological invasions. The risk curves of establishment potential can be used to determine a time window in which an environmental manager would want to take action. Similarly, the choropleth maps serve as a visual interpretation of the timeline of spread of an invasive species. And additionally, future network analysis, searching for nodes of centrality, will aid in identifying hotspots of invasion and highly significant pathways - locations where resources should be focused. It is our aim to provide these tools to be used in partnership with each other to identify and prevent the spread of invasive species before they can do damage.

14

Future Plans

The current implementation of certain features in the model can be improved. While the environmental similarity matrix serves as a reasonable basis for comparison of climates between various regions in the study, it fails to restrict species from expanding outside of a strict environmental range. More sophisticated models have been developed which would resolve this issue. Where available, it would be best to replace $\mathbf{E}$ by fitting the simulation with a comprehensive environmental niche model. These types of models predict the probable expansion range of a species and analyses have been performed for several notable pest species.

Similarly, the growth rate r included in the logistic growth term in the model will not be the same for each species. Where establishment data are available, or through a literature search, a growth rate should be estimated for the species of interest. One future direction for this research includes incorporating existing distribution data for high impact species as test cases, from which model parameters can be better estimated.

In addition, network analysis will be used to identify nodes of high connectivity and edges with increased volume of traffic. After identifying a list of candidate nodes, simulations will be modified to replicate the effects of various management strategies, and the associated change in spread rate and risk to susceptible areas over time will be assessed. From this assessment, a list of the nodes most responsive to regulation will be created, which can be used by environmental managers seeking to devote resources to more efficient inspection. Further plans would include the association of costs of inspections to the model, in order to produce optimized cost strategies for inspection.

An additional direction for research is the incorporation of the FAF predictions of future flows into both ComFlo and the matrix model. A comparison between the results from the model using resampling with results from the model using the FAF's commodity flow predic-

tions may be worthwhile.

Acknowledgments

References and Notes

A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.

Alain Barrat, Allesandro Barthelemy, and Alessandro Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008.

Transportation Research Board. *NCHRP Report 606: Forecasting Statewide Freight Toolkit*. National Cooperative Highway Research Program, 2008.

Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie, and Jonathan McPherson. *shiny: Web Application Framework for R*, 2015. URL `http://CRAN.R-project.org/package=shiny`. R package version 0.12.1.

Joseph Y. J. Chow, Choon Heon Yang, and Amelia C. Regan. State-of-the art of freight forecast modeling: lessons learned and the road ahead. *Transportation*, 37(6):1011–1030, 2010.

Manuel Colunga-Garcia, Robert A. Haack, and Adesoji O. Adelaja. Freight transportation and the potential for invasions of exotic insects in urban and periurban forests of the united states. *Journal of Economic Entomology*, 102(1):237–246, 2009.

John M. Drake and David M. Lodge. Global hot spots of biological invasions: evaluating options for ballast water management. *Proceedings of the Royal Society of London B: Biological Sciences*, 271(1539):575–580, 2004.

FHWA. Freight analysis framework, May 2015. URL `http://ops.fhwa.dot.gov/Freight/freight_analysis/faf/index.htm`.

R. Guimerá and L.A.N. Amaral. Modeling the world-wide airport network. *The European Physical Journal B - Condensed Matter and Complex Systems*, 38(2):381–385, 2004.

Philip E. Hulme. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, 46(1):10–18, 2009.

Pablo Kaluza, Andrea Kolzsch, Michael T. Gastner, and Bernd Blasius. The complex network of global cargo ship movements. *Journal of The Royal Society Interface*, 7(48):1093–1103, 2010.

Frank H. Koch, Denys Yemshanov, Manuel Colunga-Garcia, RogerD. Magarey, and WilliamD Smith. Potential establishment of alien-invasive forest insect species in the united states: where and how many? *Biological Invasions*, 13(4):969–985, 2011.

Richard N. Mack, Daniel Simberloff, Mark W. Lonsdale, Harry Evans, Michael Clout, and

Fakhri A. Bazzaz. Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications*, 10(3):689–710, June 2000.

Deborah G. McCullough, Timothy T. Work, Joseph F. Cavey, Andrew M. Liebhold, and David Marshall. Interceptions of nonindigenous plant pests at us ports of entry and border crossings over a 17-year period. *Biological Invasions*, 8(4):611–630, 2006.

NOAA. U.s. climate normals. URL `https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/climate-normals/1981-2010-normals-data`.

Dean R. Paini and Denys Yemshanov. Modelling the arrival of invasive organisms via the international marine shipping network: A khapra beetle study. *PLoS ONE*, 7(9), 2012.

Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86(14):3200–3203, Apr 2001.

RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2012. URL `http://www.rstudio.com/`.

Gregory M. Ruiz, Tonya K. Rawlings, Fred C. Dobbs, Lisa A. Drake, Timothy Mullady, Anwarul Huq, and Rita R. Colwell. Global spread of microorganisms by ships. *Nature*, 408 (6808):49–50, 2000.

H. Seebens, M. T. Gastner, and B. Blasius. The risk of marine bioinvasion caused by global shipping. *Ecology Letters*, 16(6):782–790, 2013.

Aruna Sivakumar and Chandra Bhat. Fractional split-distribution model for statewide commodity-flow analysis. *Transportation Research Record: Journal of the Transportation Research Board*, 1790:80–88, 2002.

Frank Southworth, Bruce E. Peterson, Ho-Ling Hwang, Shih-Miao Chin, and Diane Davidson. *The Freight Analysis Framework Version 3 (FAF3): A Description of the FAF3 Regional Database and How It Is Constructed*. Oak Ridge National Laboratory, June 2011.

USDA. *Imported Fire Ant Program Manual*. United States Department of Agriculture, 1 edition, 2010.

Timothy T. Work, Deborah G. McCullough, Joseph F. Cavey, and Ronald Komsa. Arrival rate of nonindigenous insect species into the united states through foreign trade. *Biological Invasions*, 7(2):323–332, 2005.

Denys Yemshanov, Frank Koch, Mark Ducey, Robert Haack, Marty Siltanen, and Kirsty Wilson. Quantifying uncertainty in pest risk maps and assessments: adopting a risk-averse decision maker's perspective. *NeoBiota*, 18:193–218, Sep 2013.
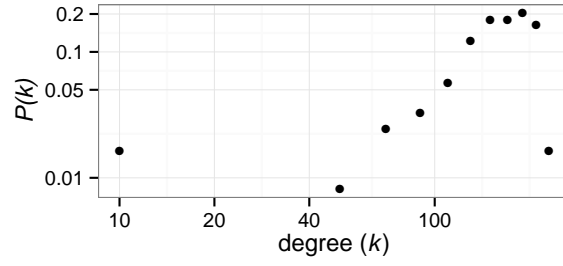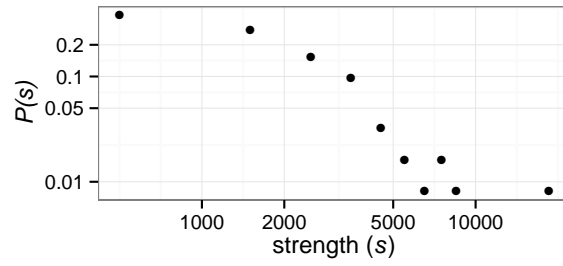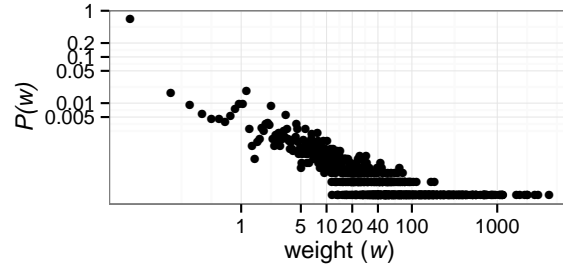
Figures
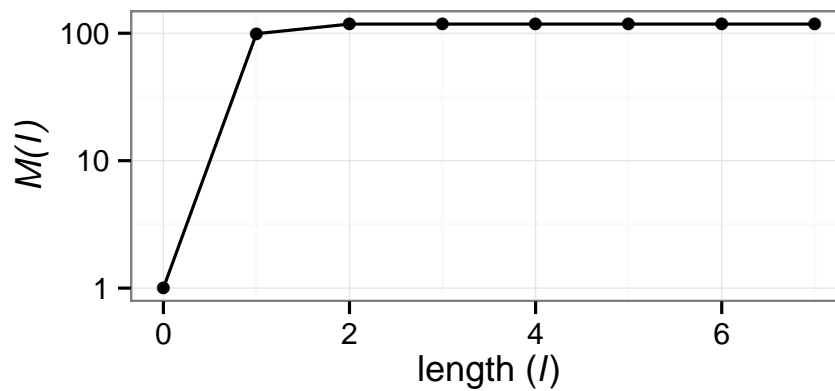
Figure 1: Average number of nodes $M(l)$ within a distance less than or equal to l from any given vertex, for the 2012 agricultural commodity shipping network.

(a) Degree distribution $P(k)$

(b) Strength distribution $P(s)$

(c) Weight distribution $P(w)$

Figure 2: Network characteristic distributions for the 2012 agricultural commodity shipping network.
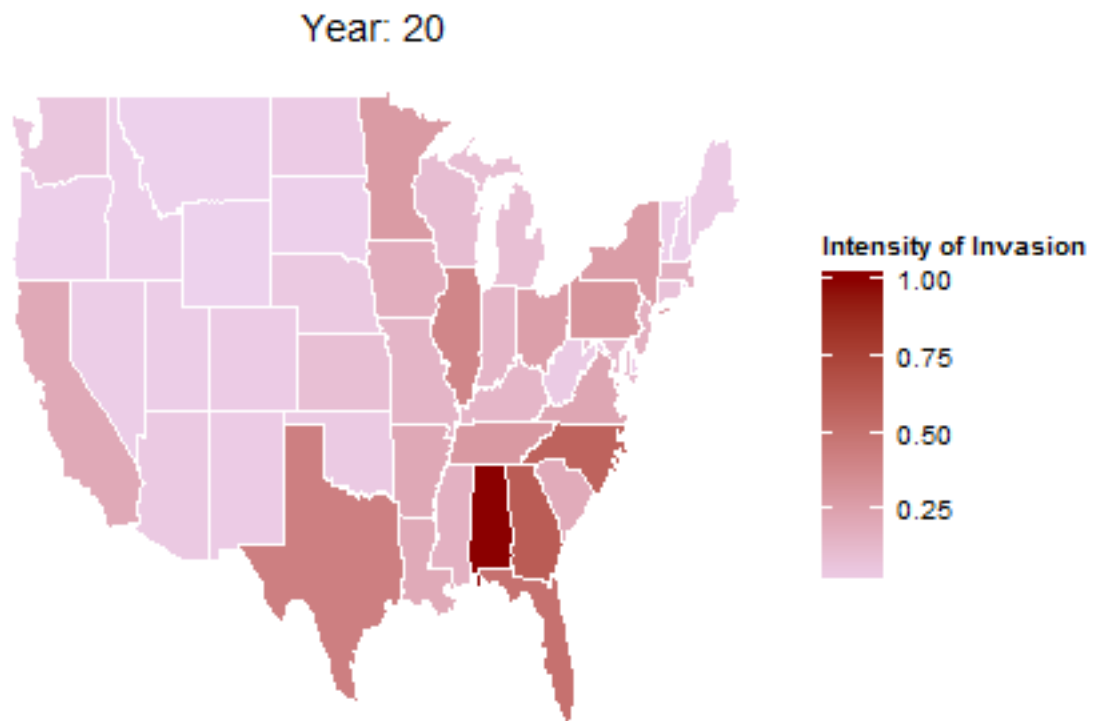
Figure 3: Year 20 choropleth map of projected red invasive fire ant spread from Alabama, transported via agricultural products, with growth rate $r = 0.01$. The plot shows the average result from 100 simulations.
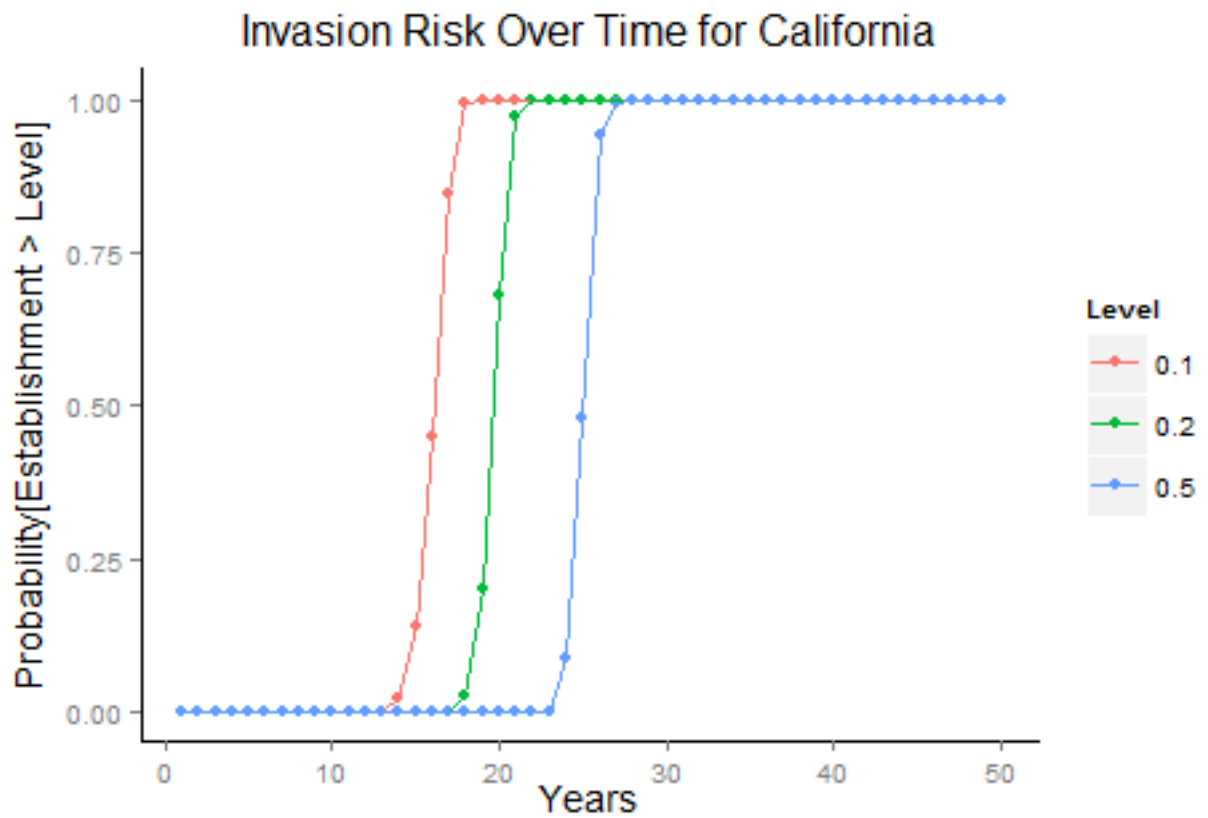
Figure 4: Risk curves for the probability of passing establishment thresholds in California. Plot shows aggregate result of 150 runs with 50 years in each. The simulations were initialized with Alabama fully established and growth rate $r = 0.01$, using data on agricultural products shipped by truck.