esa ECOSPHERE

Accounting for residential propagule pressure improves prediction of urban plant invasion

Amy J. S. Davis,[1,4,†] Kunwar K. Singh,[2,3,5] Jean-Claude Thill,[1] and Ross K. Meentemeyer[2,3]

[1]*Department of Geography and Earth Sciences, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, North Carolina 28223, USA*
[2]*Center for Geospatial Analytics, North Carolina State University, 5180 Jordan Hall, Raleigh, North Carolina 27695, USA*
[3]*Department of Forestry and Environmental Resources, North Carolina State University, 3122 Jordan Hall, Raleigh, North Carolina 27695, USA*

Citation: Davis, A. J. S., K. K. Singh, J.-C. Thill, and R. K. Meentemeyer. 2016. Accounting for residential propagule pressure improves prediction of urban plant invasion. Ecosphere 7(3):e01232. 10.1002/ecs2.1232.

Abstract. Plant invasions substantially impact the ecosystem services provided by forests in urbanizing regions. Knowing where invasion risk is greatest helps target early detection and eradication efforts, but developing an accurate predictive model of invasive species presence and spread on the basis of habitat suitability remains a challenge due to spatial variation in propagule pressure (the number of individuals released) which is likely conflated with suitability. In addition to neighborhood propagule pressure that originates with propagules dispersing from naturalized populations within invaded habitats, we expect residential propagule pressure arising from the widespread use of exotic plants in the yards of single-family residences to be an important driver of invasions, and to notably improve the predictive accuracy of species distribution models (SDMs). To this end, we collected presence/absence data for a widespread forest invader, *Ligustrum sinense* (Chinese privet), from 400 stratified random plots located along an urban gradient across the Charlotte, North Carolina metropolitan area. We assessed the relative contribution of residential propagule pressure and neighborhood propagule pressure to improving the predictive performance of a probit SDM for Chinese privet that only contains environmental predictors. Our results indicate that, although the environment-only model predicted the highest geographic area to be at risk of invasion by privet, it also had the highest rate of failure to accurately predict observed privet occurrences as indicated by the omission (incorrectly predicted absence) and commission (incorrectly predicted presence) error rates. Accounting for residential propagule pressure substantially improved model performance by reducing the omission error by nearly 50%, thereby improving upon the ability of the model to predict privet invasion in suboptimal habitat. Given that this increase in detection was accompanied by a decrease in the geographic area predicted at risk, we conclude that SDMs for invasive exotic shrubs and potentially for other synanthropic generalist plants may be highly inefficient when residential propagule pressure is not accounted for. Accounting for residential propagule pressure in models of invasive plants results in a more focused and accurate prediction of the area at risk, thus enabling decision makers to feasibly prioritize regional scale monitoring and control efforts.

Key words: Chinese privet; force of invasion; generalist invader; habitat suitability model; human-mediated invasion pressure; invasive shrub; *Ligustrum sinense*; plant invasion; propagule pressure; species distribution model; urban forest.

Received 11 September 2015; revised 8 October 2015; accepted 13 October 2015. Corresponding Editor: D. P. C. Peters.

[4] Present address: United States Environmental Protection Agency, National Exposure Research Laboratory, 109 T.W. Alexander Drive, Durham, North Carolina 27709, USA.
[5] Present address: Northern Plant Ecology Lab, Department of Biology, University of Saskatchewan, Saskatoon, Saskatchewan S7N5E2, Canada.
† **E-mail:** Davis.Amy@epa.gov

Introduction

Urban forests provide several key ecosystem services such as clean air and water, mitigation of the urban heat island effect, and recreational opportunities that are essential to the quality of life and sustainability of cities and metropolitan areas. Yet, the continued provisioning of these ecosystem services is threatened by uncontrolled exotic species invasions. Of particular concern are exotic shrub invasions prevalent in the forest understory along the eastern seaboard of the United States due to their negative impacts on both biodiversity and forest regeneration (Merriam and Feil 2002, Hartman and McCarthy 2008). Several invasive shrub species of Eastern U.S. forests (e.g., *Lonicera maacki*, *Ligustrum sinense*, *Elaeagnus umbellata*) were originally introduced as ornamental plants and widely planted in residential landscapes prior to escaping from yards and gardens to naturalize in forests (Dirr 1998, Martin et al. 2008). The same traits that enable them to grow under a wide variety of conditions in residential yards are also what enables them to be such successful forest understory invaders (Martin et al. 2008).

To date, remote sensing approaches have been ill-suited to the accurate detection and monitoring of understory invasions without the costly acquisition of hyperspectral or LiDAR data (Singh et al. 2015). Instead, the risk of invasion is assessed using species distribution models (SDMs). SDMs statistically relate species occurrences with environmental and/or climatic predictors in a geographic information system to predict the probability of species presence on the basis of habitat suitability (Elith and Leathwick 2009). This approach has been met with mixed success, as SDMs predict invader distribution much more accurately for species with narrow habitat requirements than those that possess wider environmental tolerances (Evangelista et al. 2008). Furthermore, deriving a generalist invader's true ecological niche based solely on habitat suitability data can lead to SDMs with poor predictive performance due to variations in propagule pressure, which as we show below, can be conflated with suitability (Chytrý et al. 2008). Broadly defined, propagule pressure is the cumulative introduction effort of non-native propagules (seeds) to a novel location (Lockwood et al. 2009). Propagule pressure is linked to the probability of arrival and establishment at a site, two necessary conditions beyond habitat suitability that must be met in order for invasion to occur (Shea and Chesson 2002, Tanentzap and Bazely 2009). Previous work has identified that habitat invasibility, defined as the intrinsic susceptibility of a habitat to invasion, can be confounded by propagule pressure (Lonsdale 1999, Chytrý et al. 2008). Under this scenario, locations receiving a heavy influx of propagules will have higher establishment rates and thus be more invaded; conversely, lower invasion rates will be associated with low levels of propagules, regardless of the intrinsic invasibility of a site, or suitability of the site for a particular invader. This phenomenon, coupled with spatial autocorrelation resulting from both a characteristically patchy distribution of the invader across the landscape, and the tendency of environmental predictors to have increasing similarity with decreasing distance, may inflate the effect of environmental factors on the probability of invader presence in SDMs (Legendre 1993, Dormann et al. 2007). Thus, unaccounted propagule pressure can lead to substantial model error (Eschtruth and Battles 2011).

Despite the theoretical motivation to account for propagule pressure in SDMs, it has rarely been done. When propagule pressure is accounted for, typically using a proxy variable, it has been shown to improve the performance of invasive species SDMs (Havel et al. 2002, Meentemeyer et al. 2008, Dullinger et al. 2009). However, these studies have focused on estimating the likelihood of invader dispersal from known invaded locations, which are largely nearby and within the same habitat type, and thus implicitly ignore the potential contribution of propagules from other sources. We refer to this as "neighborhood propagule pressure." In addition to neighborhood propagule pressure, residential propagule pressure arising from the dispersal of propagules from exotic shrubs planted in the yards of single-family residences is also likely an important factor explaining the distribution of exotic shrubs in metropolitan forests (Lockwood et al. 2009). Several studies have linked the presence of exotic species in natural landscapes to anthropogenic variables such as housing density, urbanization, and human population density (Gavier-Pizarro et al. 2010, Pennington et al. 2010, Pyšek et al.

2010). Furthermore, the likelihood that an ornamental plant will escape from cultivation and naturalize has been linked to its popularity in the landscape trade and prevalence in residential landscapes (Kowarik 2003, Krivanek et al. 2006, Dehnen-Schmutz et al. 2007, Hanspach et al. 2008). Given that the most popular and widely used ornamental shrubs tend to be habitat generalists and are among the most prevalent forest invaders, we investigated if a metric reflecting residential propagule pressure can be used to improve the accuracy of SDMs developed for generalist invaders, using the invasive shrub Chinese privet (*Ligustrum sinense*) as a case study.

The proxy used in this study, the residential force of invasion (rFOI) was adapted from the force of invasion described by Havel et al. (2002). Based on the premise that single-family residences serve as external sources of invasive propagules, the potential rFOI at a given location, i is measured by the cumulative sum of the inverse weighted distances from i, to every single-family house, weighted by age, within the study extent. We expect that explicitly weighting single-family residences by distance and age should result in a more realistic and more effective proxy of human-mediated invasion pressure in urban landscapes than a simple measure of housing density. Specifically, we expected to observe an increase in predictive accuracy in SDMs that account for residential propagule pressure in addition to neighborhood propagule pressure, as compared to those that do not. In addition, given that both sources of propagule pressure represent the inherently spatial process of invader spread, we also assessed the level of spatial autocorrelation present in the residuals of each model. If both sources of propagule pressure are important to the distribution of the target species, then the addition of either source of propagule pressure should result in a reduction in spatial autocorrelation as compared to a model that includes only environmental predictors.

Data and Methods

Study system and target species

We conducted our study in the Charlotte-Mecklenburg metropolitan area in North Carolina. Charlotte is one of the ten fastest growing cities in the United States, with a 2010 population of over 730 000, and encompasses the majority of Mecklenburg County (2010 U.S. Population Census). It is located in the Piedmont physiographic province, which is characterized by gently rolling terrain, erosion prone soils, and forests dominated by mixed hardwood and pine. Steep slopes are limited and are primarily located adjacent to streams. Rapid population growth and an expanding human footprint that can be characterized as urban sprawl have consumed much of the forests and agricultural land in the area (Meentemeyer et al. 2013, Delmelle et al. 2014). The remaining forests are largely mixed deciduous, dominated by oaks and hickories, and are highly fragmented.

Chinese privet is an ideal case study species for this research as it is prevalent in forests throughout the southeast and has been reported as invasive throughout the eastern United States. It is also widely utilized as an ornamental shrub due to its tolerance to a wide variety of environmental conditions. Chinese privet was first imported for ornamental use in 1852 (Dirr 1998). It was reported as being naturalized in forests throughout the North Carolina Piedmont as early as the 1930s (Radford et al. 1968). This semi-evergreen to evergreen shrub is still widely used as a hedge, as it tolerates shade, heat, drought and the clay soils that are characteristic of the Piedmont. The shrub produces small bluish-black drupes in the late fall that are consumed by birds (Wilcox and Beck 2007) and deer (Stromayer et al. 1998, Williams et al. 2008). Invasion by Chinese privet is a threat to biodiversity because it is capable of forming dense thickets, which crowd out native vegetation and prevents forest regeneration (Merriam and Feil 2002, Hart and Holmes 2013).

Field data collection

To examine the effects of residential development and environmental factors on the probability of Chinese privet presence, we sampled 345 field plots (100 m^2 in size) in patches of primarily deciduous forests stratified across three classes of building density: urban, suburban and rural in Mecklenburg County during 2009–2012 (Fig. 1). Building density was mapped from 2011 countywide parcel data (Mecklenburg County Geospatial Information Services) using a 1 km circular moving window and was

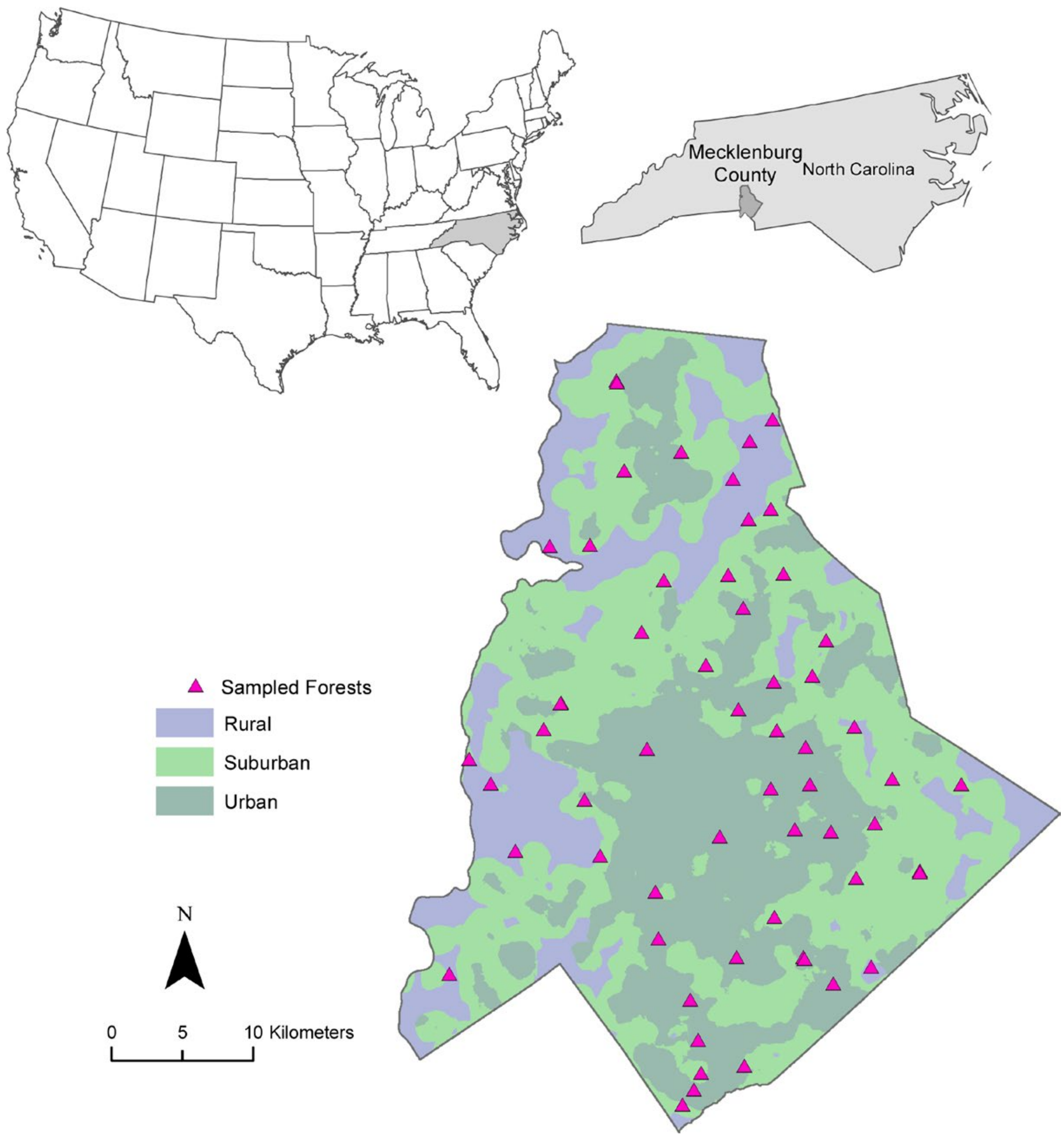

Fig. 1. Study extent showing the distribution of forests sampled for Chinese privet.

assigned to one of three classes using the method described by Theobald (2005).

We used a stratified random sampling method to select forest patches that were a minimum of 2 hectares in each class to ensure that the number of forest patches sampled reflects their spatial distribution by class within the study extent. Each forest sampled contained a minimum of 3 plots and a maximum of 10 plots, as determined by forest patch size to ensure that the spatial heterogeneity within each patch was well represented. Most forests contained 5–6 plots. Plot locations were chosen randomly in a GIS for each forest patch and were located in situ using a Trimble GeoXT global positioning system (GPS). At each plot, we recorded whether privet was present or absent, and if present, assigned scores of one to five to each plot according to the percentage of coverage by Chinese privet: 1–20, 21–40, 41–60, 61–80, or 81–100%. Thirty patches embedded in urban areas were sampled (136 plots), 31 suburban forests (139) and 11 rural forests (50). Due to difficulty in obtaining landowner permission, we augmented our study using an additional 55 plots from 14 rural forests and 4 suburban forests obtained via a random subset from the Mecklenburg County Department of Natural Resources (DNR) systematic invasive survey conducted from 2003 to 2009. The final combined data set has 400 observations.

Environmental data

Several topographic variables and habitat factors were examined as potential predictors of privet presence as indicated by theory, previous research, and based on the authors' field observations. We calculated the topographic moisture index (TMI) as the natural log of the ratio of the upslope contributing area to the slope and delineated convex, concave, and flat areas using the soil curvature method of Moore et al. (1991). The TMI and curvature are derived from a 10 m resolution DEM. Relative slope position (RSP) provides a measure of the relative position of each pixel/geographic location as compared to its neighbors using a 100 m moving window. Aspect was transformed using the method described in Beers et al. (1966). The annual mean potential solar radiation (SI) based on monthly intervals was derived using the area solar radiation tool in the ArcGIS 10. All these metrics were derived from a 10 m resolution DEM. Canopy closure (CC) was obtained from the Multi-Resolution Land Characteristics (MRLC) Consortium and is based on the 2011 National Land Cover Database (NLCD).

Due to Chinese privet's high adaptability to various climatic conditions and present geographic range that spans throughout the eastern United States, we assumed that climate would have little effect on more localized distributions. Therefore, climatic variables were not used in this study.

Residential propagule pressure

Our proxy of residential propagule pressure, the rFOI is based on the premise that the yards of single-family residences serve as external sources of propagules of Chinese privet. Thus, the potential rFOI at a given forest location i is measured by the cumulative sum of the inverse weighted Euclidean distances d_{ik} from i, to every single-family residence (k) within 1500 m of i, weighted by their age (w_k), as recorded in the county's georeferenced property database (Mecklenburg County Geospatial Information Services), using an inverse distance dispersal kernel ($\frac{1}{d_{ik}^{\alpha}}$):

$$r\mathrm{FOI}_i = \sum_{k=1}^{N} \frac{1}{d_{ik}^{\alpha}} w_k. \tag{1}$$

The α parameter modifies the degree to which the likelihood of arrival from residential sites to forest sites decreases as the distance between them increases (Havel et al. 2002). The optimal value for α was identified using likelihood profiling (Appendix A). We weighted the rFOI by housing age as it was assumed that the residential force of invasion would be more established and stronger out of older housing developments where Chinese privet may have been used as an ornamental since the early 19th century.

Neighborhood propagule pressure

To examine the influence of neighborhood propagule pressure on the probability of invader presence, we estimated the neighborhood force of invasion (Havel et al. 2002, Meentemeyer et al. 2008) occurring at each field plot generated by observed presences within 1500 m. Similar to the derivation of the rFOI, the neighborhood force of invasion (nFOI) is estimated as the distance (d_{ik}) from every known invaded cell present in the training data set, k, to every other cell (i) excluding itself in the study extent using an inverse distance power dispersal kernel (Clark et al. 2005), weighted by the invasive cover (W_{ick}):

$$n\mathrm{FOI}_i = \sum_{k=1}^{n} \frac{1}{d_{ik}^{\alpha}} W_{ick} \tag{2}$$

The field-estimated percentage invasive cover was transformed to an ordinal variable, with values ranging from 1 to 5, with 1 representing the lowest cover category of 1–20% and 5 representing the highest cover of 81–100%. We used a neighborhood of 1500 m to account for the scope of spatial dispersion around each site, as 99% of fleshy fruited seeds released from shrubs have been shown to be dispersed within this radius (Vittoz and Engler 2007). Likelihood profiling analysis was used to identify the optimal β parameter for nFOI (β = 1.5) (Appendix A).

Model development

We developed the models using generalized linear modeling with the probit-link function in the "stats" package of R, version 2.15.3 (R Core Team, 2012). Probit models differ from the more commonly used logit-link function in that the probit model assumes a normal distribution of the errors, whereas the logit model assumes a standard logistic distribution of the errors, but both typically yield very similar results (Long and Freese 2006). Probit model errors have the advantage of being analytically more tractable when trying to account for spatial dependence (Anselin 2002, Johnson et al. 2012). In probit models, the probability that a location will be invaded is calculated as:

$$\mathrm{Prob}(Y=1|X)=\Phi X\beta+\epsilon. \tag{3}$$

The link function is indicated by Φ, the cumulative normal distribution, X represents a vector of predictors and β are the parameters to be estimated. We used the overall privet prevalence (0.31) as the threshold to score plots as being invaded (1) or not invaded (0) (Manel et al. 2001).

To avoid overfitting, all possible models were enumerated by the bestglm package in R, and were evaluated using Akaike's information criterion (AIC) to select the best set of environmental predictors. Higher order variable interactions were examined, but resulted in a decrease in prediction, suggesting overfitting (Wenger and Olden 2012). We then investigated the effect of adding (1) rFOI (the "rfoi" model); (2) nFOI ("nfoi" model); and (3) rFOI and nFOI ("nrfoi" model); to the environment-only model, to isolate the relative contribution of rFOI and nFOI to improving model performance. We also compared the rFOI with a simpler metric of residential propagule pressure that is computationally faster to derive, namely single-family housing density (the "sfhd" model). Single-family housing density was calculated for each forest location using a 1500 m neighborhood.

Model evaluation

We used 75% of the data to estimate our models and the remaining 25% for validation (Fielding and Bell 1997). Both data sets contained approximately equal prevalence of Chinese privet. Model performance was evaluated based on omission error and commission error, overall predictive accuracy, and area under the receiver operating characteristic curve (AUC) (Pearce and Ferrier 2000). The observed probability of privet occurrence is plotted against the probability predicted by each model using a loess smooth function to assess model fit and check the assumption of linearity (Jacoby 2000). Model fit was assessed using the AIC.

Assessment of spatial dependence

Spatial autocorrelation (SAC) can be the result of an omitted abiotic or biotic variable or of poor model specification (Austin 2007). If our hypothesis is correct, we expect SDMs that do not include both measures of propagule pressure to have significant spatial autocorrelation, which can lead to an overestimation of the effects of the environmental predictors (Legendre 1993, Dormann et al. 2007). To assess the degree of SAC that may be present, we calculated the Moran's I statistic for the generalized residuals (numerator) of each SDM model, standardized by the square root of the variance (denominator) (Amaral et al. 2013):

$$\frac{yi-\Phi}{\sqrt{\Phi i(1-\Phi)}} \tag{4}$$

where Φ is the cumulative normal distribution of the predicted value given by $x_i\beta$, x is a $N \times k$ matrix of the predictors, and β is a $k \times 1$ vector of coefficients. The distribution of the Moran's I test statistic under the null hypothesis is asymptotically normal and can be used for hypothesis testing (Kelejian and Prucha 2001, Amaral et al. 2013). We standardized the residuals following Amaral et al. (2013) to minimize the effects of heteroskedasticity which is often present in probit residuals and can bias the Moran's *I* test statistic. A binary neighbors (observations were considered as neighboring if they were no further than 1500 meters apart) list was used to construct a row standardized spatial weight matrix (Bivand 2013).

Risk maps

Binary distribution maps were created using the Geospatial Data Abstraction Library as implemented in R (rgdal package, Bivand et al. 2014) to summarize the risk of invasion across the study extent as estimated by each of the four models (environment-only, rfoi, nfoi, and nrfoi). We then used overlay analysis to assess the degree of spatial consistency of the predicted distribution of privet between the environment-only model, and the rfoi, nfoi, and nrfoi models, respectively.

Results

Environment-only model

The results of the five best performing models as denoted by AIC are reported in Table 1. They indicate that the best set of environmental predictors for privet presence is RSP, SI, and CC. RSP and CC are negatively related to the prevalence of privet, not surprisingly, as it prefers lower floodplains and mesic environments that receive some sun, even though it can tolerate shade and upland habitats (Table 2). Despite having the highest omission rate of all the models examined (Table 3), the base model predicts the greatest geographic area (976 km^2) as being vulnerable to invasion by privet (Fig. 2a). Overall, the environment-only model for privet has the lowest predictive performance,

Table 1. Model selection for the best subset of environmental predictors of the distribution of Chinese privet, showing the five best performing environment-only models as indicated by AIC, in comparison to the full model containing all six predictors (last model in table).

Model	AIC
SI + CC + RSP	316.44
SI + ASP + CC + RSP	316.74
SI + CUR + CC + RSP	317.55
SI + CUR + ASP + CC + RSP	317.81
SI + TMI + CC + RSP	318.41
SI + TMI + CUR + ASP + CC + RSP	319.80

Table 2. Model coefficients for each predictor, with standard errors shown in parentheses.

	Model				
Predictor	env†	rfoi	nfoi	nrfoi	sfhd
RSP	−1.6206 (±0.3315)	−1.5356 (±0.3235)	−1.3590 (±0.3236)	−1.328 (±0.3326)	−1.5588 (±2.9967)
SI	0.0018 (±0.0008)	0.0013 (±0.0008)	0.0013 (±0.0008)	0.0013 (±0.0008)	0.0011 (±0.0008)
CC	−1.2061 (±0.3064)	−1.0407 (±0.3160)	−0.9320 (±0.3136)	−0.8387 (±0.3206)	−0.8635 (±0.3211)
rFOI	NA	0.0006 (±0.0001)	NA	0.0005 (±0.0001)	NA
nFOI	NA	NA	1758 (± 361.8)	1444 (±354.7)	NA
SFHD	NA	NA	NA	NA	0.0041 (±0.0008)

† Environment-only model.

Table 3. Evaluation of the species distribution models developed for Chinese privet.

Model	AIC	Threshold	Omission	Commission	Accuracy	AUC	Moran's I‡
env†	320.00	0.31	0.29	0.30	0.70	0.80	0.210***
rfoi	293.45	0.31	0.15	0.20	0.82	0.91	0.138***
nfoi	284.97	0.31	0.24	0.24	0.76	0.87	0.136***
nrfoi	271.78	0.31	0.15	0.14	0.86	0.94	0.091**
sfhd	294.59	0.31	0.18	0.19	0.82	0.89	0.145***

† Environment-only model.
‡ Significance levels: *P < 0.05, **P < 0.01, ***P < 0.001.

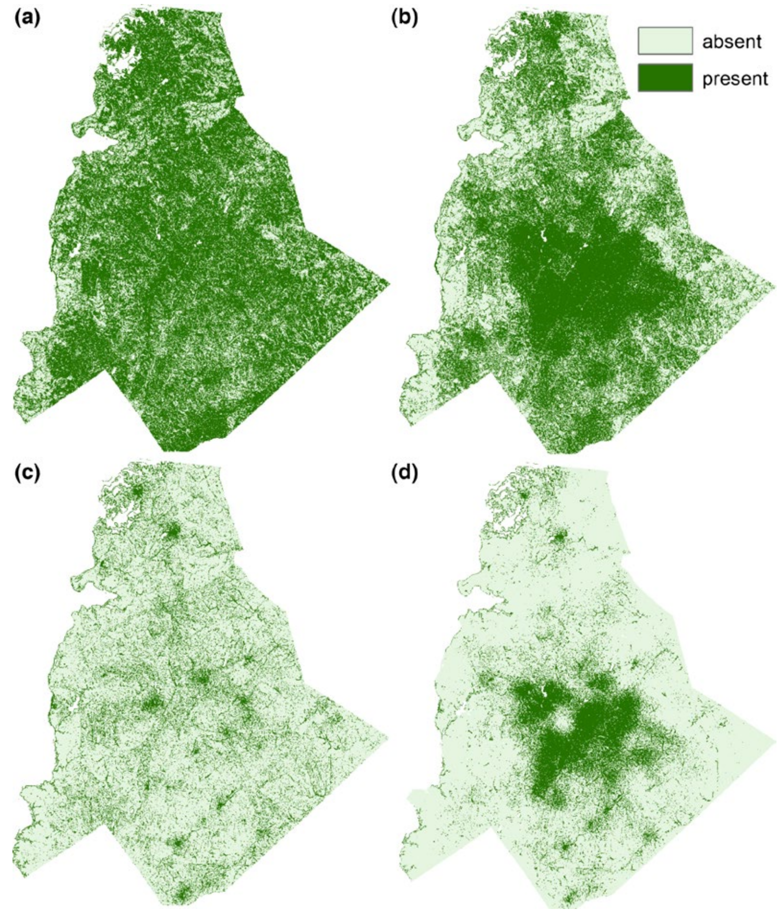

Fig. 2. Risk maps generated by each model: (a) environment-only, (b) rfoi, (c) nfoi, and (d) nrfoi. The area predicted by each model to be at risk of invasion by Chinese privet is listed adjacent to each map.

also exhibiting the highest commission rate and lowest accuracy and AUC, as compared to the other models (Table 3). The environment-only model has the highest residual spatial autocorrelation as indicated by Moran's *I*, suggesting a misspecification error (i.e., the omission of one or both forms of propagule pressure); and the highest AIC, indicating that the environment-only model has the poorest fit of the models examined (Table 3).

Influence of residential propagule pressure on model performance

The addition of rFOI substantially decreased the omission and commission error rates, and increased the AUC (Table 3). Although the overall accuracy of models containing rFOI or sfhd is the same, rFOI appears to be a better predictor of Chinese privet distribution as compared to sfhd, having a lower omission error, slightly lower AIC and higher AUC (Table 3). The addition of rFOI reduced the level of Moran's *I* observed in the environment-only model by nearly half. The rFOI model predicts the second highest geographic area as vulnerable to invasion by Chinese privet (Fig. 2b).

Contribution of neighborhood propagule pressure to model performance

The addition of nFOI to the environment-only model also results in lowered omission and commission error rates and thus increased predictive accuracy. The Moran's *I* of the nfoi model is also substantially lower as compared to the environment-only model. However, as compared to the rfoi model, the omission and commission error rates are higher, and the AUC is lower (Table 3). The nfoi model predicts a much smaller area (274 km^2) to be at risk of invasion by privet than the rfoi model (789 km^2) (Fig. 2c).

Models with residential and neighborhood propagule pressure

As evidenced in Table 3, the best predictive model contains both rFOI and nFOI, having the lowest commission rate, and highest accuracy and AUC. This model also has the lowest Moran's *I*, suggesting that rFOI and nFOI are indeed representing two distinct sources of propagule pressure and that either unaccounted neighborhood propagule pressure or unaccounted residential propagule pressure can generate at least some of the spatial dependence observed in the environment-only model residuals. The nrfoi model predicts the lowest area at risk of privet invasion (Fig. 2d).

Comparison of risk maps

The environment-only and rfoi models have the highest spatial agreement with regard to predicting privet presence (Table 4). The level of agreement between the environment-only and nfoi or nrfoi model is much lower, due to these models predicting a much lower geographic area at risk of invasion by privet. Notably, the location of areas predicted as devoid of privet by the nfoi model but invaded according to the rfoi model (Appendix B) are mostly located in the older, more densely urbanized sections of Charlotte, which suggests that the rFOI is a mechanism distinct from nFOI in influencing the risk of privet invasion. The degree of consistency of predicted absences between the environment-only and each of the other models is very high, ranging from 88 to 99% (Table 4). Maps highlighting the differences between models can be found in Appendix B.

Table 4. Results from overlay analysis showing the percentage spatial agreement of the predicted distribution of privet between the environment-only and the rfoi, nfoi, and nrfoi models, respectively.

	Environment-only	
Model	Presence	Absence
Presence		
rfoi	76% (742 km^2)	12% (47 km^2)
nfoi	12% (273 km^2)	1% (1 km^2)
nrfoi	13% (252 km^2)	3% (14 km^2)
Absence		
rfoi	24% (234 km^2)	88% (337 km^2)
nfoi	88% (703 km^2)	99% (383 km^2)
nrfoi	87% (724 km^2)	97% (371 km^2)

Discussion

We investigated whether the addition of metrics of residential propagule pressure and neighborhood propagule pressure to SDMs developed for Chinese privet better explains invasion patterns in metropolitan forests and improves model accuracy. Our results show that (1) residential propagule pressure does influence the likelihood of invasion by Chinese privet and its inclusion generated a significant decrease in omission errors as compared to either environment-only or nfoi models; and (2) the highest predictive accuracy is obtained when both residential and neighborhood sources of propagule pressure are included in the models. We discuss our results in more detail below.

A comparison of models with and without residential propagule pressure measured as the rFOI, reveals that the addition of rFOI results in substantially lower omission and commission errors. The risk map generated by the environment-only model (Fig. 2a) suggests that the majority of the study extent has suitable habitat for privet. This is not surprising, given that Chinese privet has a broad geographic distribution and tolerates a wide range of environmental conditions. However, the rfoi distribution map revealed that much less habitat is at risk of invasion after accounting for residential propagule pressure via the rFOI (Fig. 2b). More significantly, the results of the overlay analysis identified an additional 47 km^2 of forest area at risk of invasion after accounting for the rFOI (Table 4), which is striking considering that the environment-only model had the highest predicted geographic distribution of Chinese privet of all the models considered. This indicates that SDMs for privet and potentially for other invasive species may be grossly inefficient, both under- and over predicting the distribution of the invader when residential propagule pressure is not accounted for. Furthermore, the superiority of the rfoi model in comparison to the environment-only or nfoi models in correctly identifying habitats already invaded by Chinese privet (as indicated by the omission error, Table 3), indicates that privet is found in habitats predicted as unsuitable by the environment-only or the nfoi models presumably due to high residential propagule pressure. Thus, marginally suitable habitats that receive high amounts of seed rain from residences harboring the invader for ornamental use are more susceptible to invasions than what otherwise would be predicted by models that do not estimate residential propagule pressure. These results suggest that residential propagule pressure can overwhelm abiotic resistance to invasion, as has been observed previously with neighborhood propagule pressure (Von Holle and Simberloff 2005, Thomson 2007, Eschtruth and Battles 2011). Accounting for neighborhood propagule pressure does little to improve the omission error, and instead results in a dramatic reduction in the predicted area at risk of invasion by Chinese privet (Fig. 2c). As compared to the rfoi model, the nfoi model has greater rates of omission and commission and lower AUC, despite having a lower AIC. Although these measures are often consistent, the AIC provides an in-sample measure of the goodness-of-fit of the model compared to the observed data, whereas the AUC is a measure of discriminatory power, and, was derived "out-of sample" using the holdout/testing data.

The addition of either rFOI or nFOI to the environment-only model greatly reduces the level of spatial autocorrelation present in the residuals as indicated by Moran's *I*, and the lowest level is obtained in the nrfoi model, which includes both rFOI and nFOI. The nrfoi model also has the highest overall accuracy and AUC. This suggests that the failure to account for either source of propagule pressure can cause the predictive error and spatial dependence observed in the environment-only model. However, a drawback to accounting for neighborhood propagule pressure using metrics like the nFOI, is that a priori data on known invaded locations is required. The consequence of this is that models that incorporate neighborhood effects are suitable only for interpolation and ill-suited for extrapolative applications such as predicting the potential distribution in a novel geographic range. Furthermore, our knowledge of invader presence is limited to the locations that were sampled, thus if broad geographic areas exist within the study extent where no data were collected or presences were not identified, then neighborhood propagule pressure may be poorly estimated or unknown for these locations. An advantage of the rFOI as compared to the nFOI, is that all the values of the rFOI can readily be determined for the entire study extent both within the model calibration area and if needed, outside the model calibration area, since all of the cells belonging to residences can be derived via GIS. However, the rFOI can also introduce uncertainty and predictive error into the model since it assumes that all residences are a potential source of propagules, and this is difficult, if not impossible, to verify. If the species are missing from residences at random, this can result in over-estimation of residential propagule pressure, which is not likely to affect model performance unless the overall prevalence of species presence in residential landscapes is very low. If the errors are spatially clustered, this may reflect spatial variability in residential propagule pressure due to a mismatch of the age of the housing development and the popularity of the plant, resulting in more substantial model error. However, this can be addressed by examining if model accuracy varies with housing age and/or by incorporating assumptions based on the invader's popularity over time in the nursery trade. We conclude from this that although our results point to the most robust model as including both residential and neighborhood propagule pressure, our results demonstrate that accounting only for residential propagule pressure can result in models with good accuracy. Thus, the inclusion of rFOI in extrapolative models developed for predicting the spread of invasive plants within the context of land cover-or climate change scenarios may improve their performance and warrants future investigation.

When dealing with invasive species it is highly likely that the response variable is spatially autocorrelated owing to dispersal and other biotic processes (Legendre 1993, Dirnböck and Dullinger 2004, Bahn et al. 2008). We have shown that residential and neighborhood propagule pressures are sources of residual SAC in a SDM for Chinese privet and can be utilized to improve prediction. This marks a radical departure from the dominant approach to species distribution modeling that advocates accounting for residual SAC using an autocovariate or removing it using spatial filters, thus overlooking the opportunity to investigate the origins of the spatial dependence (Van Teeffelen and Ovaskainen 2007, De Knegt et al. 2010, Miller and Franklin 2010,).

One thing to keep in mind when examining the classification accuracy of SDMs is that commission errors may not be the result of model error, but due to the factors distinct from environmental suitability, such as that the organism has not colonized the site because it has yet to arrive there, or it has arrived there, but has failed to establish a reproducing population due to demographic stochasticity (Taylor and Hastings 2005, Barbosa et al. 2013). Other difficult to measure factors such as biotic resistance, competition, and predation may also prevent establishment (Theoharides and Dukes 2007). The failure to colonize all suitable sites has also been attributed to presumed lack of equilibrium of the invader with the environment due to an insufficient residence time (Václavík and Meentemeyer 2012). Accounting for residential propagule pressure has improved our model commission error by identifying suitable habitat that is estimated to receive a relatively low number of propagules and thus highly likely to experience dispersal or establishment failure. A much greater reduction in omission error was observed, revealing the potential importance of source-sink dynamics in driving invader distributions in human-dominated landscapes. However, in order to maximize the likelihood that residential propagule pressure will have a measurable effect on the performance of SDMs, this approach is best suited to improving the prediction of invasions by species that are widely planted, have long residence times, and are not at the early stage of invasion (Pyšek et al. 2009).

A potential caveat to this work is that we only sampled from a subset of Chinese privet's large geographic range (Jarnevich et al. 2015). Other environmental or climactic predictors may be important for explaining the distribution of Chinese privet and other invasive species at the continental or global extent. However, our conclusions are drawn from a stratified random sampling design yielding high-quality presence-absence data, the majority of which was collected by us, with additional sampling points obtained from Mecklenburg Department of Natural Resources that were collected by seasoned, knowledgeable professionals. A key disadvantage of using data from throughout Chinese privet's range in the United States is that we would have had to rely on presence-only data from a multitude of secondary (or tertiary) sources with varying accuracy with regard to spatial location and species identification. A drawback common to most SDMs, including ours, is that they are developed using a cross section of data points obtained from only a single visit to a location, thus, the species could have been present in the past, but has not persisted, due to demographic stochasticity, or the species has not yet arrived at a suitable habitat (Sinclair et al. 2010, Barbosa et al. 2013). The latter presents a more significant source of error, but likely only when extrapolating models outside the current known distribution (Sinclair et al. 2010).

The main thrust of this study was to determine whether the seed rain of propagules dispersing from Chinese privet grown as ornamental shrubs in residential landscapes generates a measurable propagule pressure measured as the rFOI. A follow-up question was whether it can be used to explain the presence of Chinese privet in addition to a metric that takes into account dispersal from known invaded field plots. As such, we expected the dispersal kernel used in the residential and neighborhood force of invasion models to be the same. Although we applied a 1500 m neighborhood to both rFOI and nFOI, different optimal values were identified, suggesting that although propagules are arriving from both sources, the predominant dispersal vectors are different. The rFOI α value of 0.5 implies that more propagules are dispersing farther from their source as compared to the β value of nFOI. This may be due to deer, birds, and other urban-adapted wildlife preferentially foraging in residential yards for food as opposed to forest interiors (Williams et al. 2008).

This is the first study that explicitly investigates the potential links between propagule pressure and the performance of SDMs in urban landscapes. We have shown that high residential propagule pressure increases the risk of invasion in habitats that would otherwise be identified as unsuitable for invasion. This suggests that residential propagule pressure can surmount unfavorable environmental conditions and result in the establishment of the invader. Our study has also demonstrated that omission and commission errors of the environment-only model are associated with unaccounted residential propagule pressure and to a lesser degree, neighborhood propagule pressure. As observed in Fig. 2a, invasive species with wide environmental tolerances will likely have a vast potential distribution predicted by models that consider only environmental factors and such large risk areas can overwhelm regional scale monitoring and control efforts (Evangelista et al. 2008). This work has shown the potential for accounting for rFOI and/or nFOI to derisk areas of suitable habitat that have low propagule pressure, thereby enabling land managers to focus on a feasible number of forest locations for monitoring and control of invaders. Accounting for residential propagule pressure has the potential to improve the accuracy of SDMs developed for other prevalent invasive plants that were originally introduced to the United States as ornamental species and continue to be widely planted in residential landscapes (e.g., *Lonicera maacki*, *Elaeagnus umbellata*). Our approach is not limited to ornamental plants and can also be applied to the improved mapping of any problematic synanthropic generalist species whose current and/or future distribution under given climate change or land use change scenarios threatens species of conservation concern (McKinney 2006). Our results suggest that SDMs that do not consider residential propagule pressure or similar mechanistic proxies of human-mediated propagule pressure, will be wildly inaccurate when predicting the risk of invasion by exotic species that have strong ties to human settlement patterns.

Acknowledgments

Funding for this research was provided by the UNC Charlotte Graduate School in the form of a fellowship to A. Davis. J. Murrey Atkins Library at UNC Charlotte provided partial funding for the publication fees associated with this study. We thank Christa Rogers from the Mecklenburg Department of Natural Resources for providing us with their invasive survey data. We also thank two anonymous reviewers, whose comments greatly enhanced the manuscript.

Literature Cited

Amaral, P. V., L. Anselin, and D. Arribas-Bel. 2013. Testing for spatial error dependence in probit models. Letters in Spatial and Resource Sciences 6:91–101.

Anselin, L. 2002. Under the hood issues in the specification and interpretation of spatial regression models. Agricultural Economics 27:247–267.

Austin, M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological Modelling 200:1–19.

Bahn, V., W. B. Krohn, and R. J. O'Connor. 2008. Dispersal leads to spatial autocorrelation in species distributions: a simulation model. Ecological Modelling 213:285–292.

Barbosa, A. M., R. Real, A.-R. Muñoz, and J. A. Brown. 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. Diversity and Distributions 19:1333–1338.

Beers, T. W., P. C. Dress, and L. C. Wensel. 1966. Aspect transformation in productivity research. Journal of Forestry 64:691–692.

Bivand, R. 2013. spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-68. http://CRAN.R-project.org/package=spdep.

Bivand, R., T. Keitt, and B. Rowlingson. 2014. rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.9-1. http://CRAN.R-project.org/package=rgdal.

Chytrý, M., V. Jarosik, P. Pysek, O. Hajek, I. Knollová, L. Tichý, and J. Danihelka. 2008. Separating habitat invasibility by alien plants from the actual level of invasion. Ecology 89:1541–1553.

Clark, C. J., J. R. Poulsen, B. M. Bolker, E. F. Connor, and V. T. Parker. 2005. Comparative seed shadows of bird-, monkey-, and wind-dispersed trees. Ecology 86:2684–2694.

De Knegt, H. J., F. Van Langevelde, M. B. Coughenour, A. K. Skidmore, W. F. De Boer, I. M. A. Heitkönig, N. M. Knox, R. Slotow, C. Van der Waal, and H. H. T. Prins. 2010. Spatial autocorrelation and the scaling of species-environment relationships. Ecology 91:2455–2465.

Dehnen-Schmutz, K., J. Touza, C. Perrings, and M. Williamson. 2007. A century of the ornamental plant trade and its impact on invasion success. Diversity and Distributions 13:527–534.

Delmelle, E. C., Y. Zhou, and J.-C. Thill. 2014. Densification without growth management? Evidence from local land development and housing trends in Charlotte, North Carolina, USA. Sustainability 6:3975–3990.

Dirnböck, T., and S. Dullinger. 2004. Habitat distribution models, spatial autocorrelation, functional traits and dispersal capacity of alpine plant species. Journal of Vegetation Science 15:77–84.

Dirr, M. A. 1998. *Manual of Woody Landscape Plants: Their Identification, Ornamental Characteristics, Culture, Propagation and Uses*, Fifth edition. Stipes Publishing Co., Champaign, Illinois, USA.

Dormann, C., et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628.

Dullinger, S., I. Kleinbauer, J. Peterseil, M. Smolik, and F. Essl. 2009. Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. Biological Invasions 11:2401–2414.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40:677–697.

Eschtruth, A. K., and J. J. Battles. 2011. The importance of quantifying propagule pressure to understand invasion: an examination of riparian forest invasibility. Ecology 92:1314–1322.

Evangelista, P. H., S. Kumar, T. J. Stohlgren, C. S. Jarnevich, A. W. Crall, J. B. III Norman, and D. T. Barnett. 2008. Modelling invasion for a habitat generalist and a specialist plant species. Diversity and Distributions 14:808–817.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38–49.

Gavier-Pizarro, G. I., S. I. Stewart, C. Huebner, N. S. Keuler, and V. C. Radeloff. 2010. Housing is positively associated with invasive exotic plant species richness in New England, USA. Ecological Applications 20:1913–1925.

Hanspach, J., I. Kuhn, P. Pysek, E. Boos, and S. Klotz. 2008. Correlates of naturalization and occupancy of introduced ornamentals in Germany. Perspectives in Plant Ecology Evolution and Systematics 10:241–250.

Hart, J. L., and B. N. Holmes. 2013. Relationships between *Ligustrum sinense* invasion, biodiversity, and development in a mixed bottomland forest. Invasive Plant Science and Management 6:175–186.

Hartman, K. M., and B. C. McCarthy. 2008. Changes in forest structure and species composition following invasion by a non-indigenous shrub, Amur honeysuckle (*Lonicera maackii*). Journal of the Torrey Botanical Society 135:245–259.

Havel, J. E., J. B. Shurin, and J. R. Jones. 2002. Estimating dispersal from patterns of spread: spatial and local control of lake invasions. Ecology 83:3306–3318.

Jacoby, W. G. 2000. Loess: a nonparametric, graphical tool for depicting relationships between variables. Electoral Studies 19:577–613.

Jarnevich, C. S., T. J. Stohlgren, S. Kumar, J. T. Morisette, and T. R. Holcombe. 2015. Caveats for correlative species distribution modeling. Ecological Informatics 29:6–15.

Johnson, D. S., P. B. Conn, M. B. Hooten, J. C. Ray, and B. A. Pond. 2012. Spatial occupancy models for large data sets. Ecology 94:801–808.

Kelejian, H. H., and I. R. Prucha. 2001. On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics 104:219–257.

Kowarik, I. 2003. Human agency in biological invasions: secondary releases foster naturalisation and population expansion of alien plant species. Biological Invasions 5:293–312.

Krivanek, M., P. Pysek, and V. Jarosik. 2006. Planting history and propagule pressure as predictors of invasion by woody species in a temperate region. Conservation Biology 20:1487–1498.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659–1673.

Lockwood, J. L., P. Cassey, and T. M. Blackburn. 2009. The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. Diversity and Distributions 15:904–910.

Long, J. S., and J. Freese. 2006. *Regression models for categorical dependent variables using Stata*. Stata Press, College Station, Texas, USA.

Lonsdale, W. M. 1999. Global patterns of plant invasions and the concept of invasibility. Ecology 80:1522–1536.

Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. Journal of Applied Ecology 38:921–931.

Martin, P. H., C. D. Canham, and P. L. Marks. 2008. Why forests appear resistant to exotic plant invasions: intentional introductions, stand dynamics, and the role of shade tolerance. Frontiers in Ecology and the Environment 7:142–149.

McKinney, M. L. 2006. Urbanization as a major cause of biotic homogenization. Biological Conservation 127:247–260.

Meentemeyer, R. K., B. L. Anacker, W. Mark, and D. M. Rizzo. 2008. Early detection of emerging forest disease using dispersal estimation and ecological niche modeling. Ecological Applications 18:377–390.

Meentemeyer, R. K., W. W. Tang, M. A. Dorning, J. B. Vogler, N. J. Cunniffe, and D. A. Shoemaker. 2013. FUTURES: Multilevel simulations of emerging urban-rural landscape structure using a stochastic patch-growing algorithm. Annals of the Association of American Geographers 103:785–807.

Merriam, R. W., and E. Feil. 2002. The potential impact of an introduced shrub on native plant diversity and forest regeneration. Biological Invasions 4:369–373.

Miller, J. A., and J. Franklin. 2010. Incorporating spatial autocorrelation in species distribution models. Pages 685–702 *in* M. M. Fischer, and A. Getis, editors. *Handbook of applied spatial analysis*. Springer, Berlin, Heidelberg.

Moore, I. D., R. B. Grayson, and A. R. Landson. 1991. Digital Terrain Modelling: A review of hydrological, geomorphological, and biological applications. Hydrological Processes 5:3–30.

Pearce, J., and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling 133:225–245.

Pennington, D. N., J. R. Hansel, and D. L. Gorchov. 2010. Urbanization and riparian forest woody communities: diversity, composition, and structure within a metropolitan landscape. Biological Conservation 143:182–194.

Pyšek, P., M. Křivánek, and V. Jarošík. 2009. Planting intensity, residence time, and species traits determine invasion success of alien woody species. Ecology 90:2734–2744.

Pyšek, P., et al. 2010. Disentangling the role of environmental and human pressures on biological invasions across Europe. Proceedings of the National Academy of Sciences 107:12157–12162.

R Core Team. 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radford, A. E., H. E. Ahles, and C. R. Bell. 1968. *Manual of the vascular flora of the Carolinas*. University of North Carolina Press, Chapel Hill, North Carolina, USA.

Shea, K., and P. Chesson. 2002. Community ecology theory as a framework for biological invasions. Trends in Ecology and Evolution 17:170–176.

Sinclair, S. J., M. D. White, and G. R. Newell. 2010. How useful are species distribution models for managing biodiversity under future climates? Ecology and Society 15:8 http://www.ecologyandsociety.org/vol15/iss1/art8/.

Singh, K. K., A. J. Davis, and R. K. Meentemeyer. 2015. Detecting understory plant invasion in urban forests using LiDAR. International Journal of Applied Earth Observation and Geoinformation 38:267–279.

Stromayer, K. A. K., R. J. Warren, A. S. Johnson, P. E. Hale, C. L. Rogers, and C. L. Tucker. 1998. Chinese privet and the feeding ecology of white-tailed deer: The role of an exotic plant. Journal of Wildlife Management 62:1321–1329.

Tanentzap, A. J., and D. R. Bazely. 2009. Propagule pressure and resource availability determine plant community invasibility in a temperate forest understorey. Oikos 118:300–308.

Taylor, C. M., and A. Hastings. 2005. Allee effects in biological invasions. Ecology Letters 8:895–908.

Theobald, D. M. 2005. Landscape Patterns of Exurban Growth in the USA from 1980 to 2020. Ecology and Society 10: 32. http://www.ecologyandsociety.org/vol10/iss1/art32/.

Theoharides, K. A., and J. S. Dukes. 2007. Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. New Phytologist 176:256–273.

Thomson, D. M. 2007. Do source-sink dynamics promote the spread of an invasive grass into a novel habitat? Ecology 88:3126–3134.

Václavík, T., and R. K. Meentemeyer. 2012. Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Diversity and Distributions 18:73–83.

Van Teeffelen, A. J., and O. Ovaskainen. 2007. Can the cause of aggregation be inferred from species distributions? Oikos 116:4–16.

Vittoz, P., and R. Engler. 2007. Seed dispersal distances: a typology based on dispersal modes and plant traits. Botanica Helvetica 117:109–124.

Von Holle, B., and D. Simberloff. 2005. Ecological resistance to biological invasion overwhelmed by propagule pressure. Ecology 86:3212–3218.

Wenger, S. J., and J. D. Olden. 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. Methods in Ecology and Evolution 3:260–267.

Wilcox, J., and C. W. Beck. 2007. Effects of *Ligustrum sinense* Lour (Chinese privet) on abundance and diversity of songbirds and native plants in a southeastern nature preserve. Southeastern Naturalist 6:535–550.

Williams, S. C., J. S. Ward, and U. Ramakrishnan. 2008. Endozoochory by white-tailed deer (*Odocoileus virginianus*) across a suburban/woodland interface. Forest Ecology and Management 255:940–947.

Supporting Information

Additional supporting information may be found online at http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1232/supinfo