Identifying hotspots for plant invasions and forecasting focal points of further spread

Inés Ibáñez*[1], John. A. Silander Jr[2], Jenica M. Allen[2], Sarah A. Treanor[2] and Adam Wilson[2]

[1]School of Natural Resources and Environment, University of Michigan, 440 Church St, Ann Arbor, MI 48109 1041, USA; and [2]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA

Summary

1. To ensure the successful detection, control and eradication of invasive plant species, we need information that can identify areas prone to invasions and criteria that can point out which particular populations may become foci of further spread. Specifically, our work aimed to develop statistical models that identify hotspots of invasive plant species and evaluate the conditions that give rise to successful populations of invasive species.

2. We combined extensive data sets on invasive species richness and on species per cent ground cover, together with climate, local habitat and land cover data. We then estimated invasive species richness as a function of those environmental variables by developing a spatially explicit generalized linear model within a hierarchical Bayesian framework. In a second analysis, we used an ordinal logistic regression model to quantify invasive species abundance as a function of the same set of predictor variables.

3. Our results show which locations in the studied region, north-eastern USA, are prone to plant species invasions given the combination of climatic and land cover conditions particular to the sites. Predictions were also generated under a range of climate scenarios forecasted for the region, which pointed out at an increase in invasive species incidence under the most moderate forecast. Predicted abundance for some of the most common invasive plant species, *Berberis thumbergii*, *Celastrus orbiculatus*, *Euonymus alata*, *Elaeagnus umbellata* and *Rosa multiflora*, allowed us to identify the specific conditions that promote successful population growth of these species, populations that could become foci of further spread.

4. *Synthesis and applications.* Reliable predictions of plants' invasive potential are crucial for the successful implementation of control and eradication management plans. By following a multivariate approach the parameters estimated in this study can now be used on targeted locations to evaluate the risk of invasions given the local climate and landscape structure; they can also be applied under different climate scenarios and changing landscapes providing an array of possible outcomes. In addition, this modelling approach can be easily used in other regions and for other species.

Key-words: *Berberis*, *Celastrus*, *Elaeagnus*, *Euonymus*, hierarchical Bayes, invasive species, IPANE, *Rosa*

Introduction

The spread of invasive species has developed into a major ecological and economic issue for many regions of the world (Pimentel 2002). As a consequence, inference on where to target early control and eradication efforts has become critical in the management of those species. Identifying which particular areas may be likely to host large numbers of invasive species and recognizing which specific populations may become a hub of further spread are common goals for ecologists and natural resources managers. We evaluated how variables related to a region's climate, landscape and habitat are associated with the establishment and further spread of invasive plant species with

*Correspondence author. E-mail: iibanez@umich.edu

the objective of providing information to better target eradication and control efforts in locations where future invasions are likely to take place over time.

Much research on invasive species has focused on species traits that promote invasive species success (e.g. Rejmanek *et al.* 2005; Herron *et al.* 2007). While some organism-based generalizations are possible, for example the positive relationship between high reproductive output and invasive potential, characteristics of the recipient environment are also important determinants of invasive species success (Richardson & Pysek 2006; Eschtruth & Battles 2009). Studies of community invasibility have been recently reviewed elsewhere and suggest that a single set of attributes cannot be assigned to areas harbouring many invasive species (Fridley *et al.* 2007; Melbourne *et al.* 2007; Compagnoni & Halpern 2009). Perhaps most importantly, the interaction of species' traits and conditions in the recipient area are critical to determining invasive species success, such that a species with invasive tendencies will not succeed in all recipient areas (Richardson & Pysek 2006; Eschtruth & Battles 2009; Milbau *et al.* 2009), nor during all phases of the invasion process (Milbau & Stout 2008; Milbau *et al.* 2009). Studies attempting to synthetically predict invasive species success generally suggest that plant invasions are particularly idiosyncratic and often only vague generalizations about patterns and processes are possible (Hayes & Barry 2008; but see Milbau & Stout 2008).

Such difficulties are attributed not only to variation in plant trait interactions with environmental conditions, but are also due to the fact that plant introductions take place under a wide range of particular circumstances in time and space (McDonald, Motzkin & Foster 2008). For example, some plant species only became invasive after being extensively planted, as was the case with kudzu, *Pueraria lobata* (Forseth & Innis 2004) in the southeast USA. This perennial vine was planted to control soil erosion and afterwards became one of the most aggressive invasive plants around forest edges (Blaustein 2001). In other cases the species' spread was aided by a particular set of circumstances (i.e. 'windows of opportunity') as seems to have happened with Japanese barberry, *Berberis thunbergii*, in the Northeast USA. *Berberis thunbergii* was first introduced as a horticultural planting to replace the European barberry; its spread into natural areas took place once agricultural fields were abandoned in the early to mid-20th century (Silander & Klepeis 1999; DeGasperis & Motzkin 2007; Mosher, Silander & Latimer 2009). Would kudzu have become the major problem it is now, with $500 million spent per year on its control (Miller 1991), without initial planting campaigns? Or would Japanese barberry have become so prevalent if old fields have not been available to such an extent and its planting encouraged early on? The set of conditions that allowed a species to become invasive may not be replicated at different sites or points in time; these idiosyncrasies of species invasions hamper our ability to predict future events.

In particular, the structure of the landscape (e.g. corridors for dispersal, occurrence of disturbed areas) has played a very important role in the establishment and spread of invasive plant species. Many plant invasions have been facilitated by human-altered landscapes (e.g. Mack & D'Antonio 1998; Levine & D'Antonio 1999; Cadenasso & Pickett 2001; Vila *et al.* 2007). In fact, recent research suggests that identification of the ecological neighbourhoods that benefit the spread and establishment of particular species increases the accuracy of predictions of species' invasion (Ibáñez *et al.* 2009).

Our approach considers invasions at the regional scale (i.e. hundreds of kilometres). Although climate is undoubtedly the major driver of a species distribution (Woodward 1987) at this scale, we also incorporated data sources in addition to climate, specifically landscape structure and habitat type, that have been key to predicting current patterns of plant invasions (Soberon & Peterson 2005; Chytry *et al.* 2008; Ibáñez *et al.* 2009). We ask: on the basis of climate and landscape structure, what makes a particular region prone to plant invasions? If propagules are available, can we identify areas where the potential for successful establishment and spread of alien species is high? Many of the invasive plant species we find today can grow in a large array of settings (e.g. climates, habitats), but only under particular ecological conditions do they become aggressive invaders (Pino *et al.* 2005).

Another consideration is the exacerbation of invasive species impact predicted with global change (Hobbs & Mooney 2005; Meyerson & Mooney 2007). Natural systems will be challenged not only by human-mediated species movement but also via the effects of global climate change. The stresses imposed by multifaceted environmental change may make natural areas more susceptible to invasions, creating more opportunities for the spread of alien species. Therefore, when evaluating the susceptibility of a region to invasive plants, we should operate in the context of a dynamic environment where several climate scenarios may be taken into account.

In the case of our study system, the New England region of eastern North America, one-third of the vascular plant flora is non-indigenous, of which 3–5% are invasive (Mehrhoff 2000). These alien species constitute a major threat to the preservation of the region's natural vegetation (Farnsworth 2004) and significantly affect the local economy (Barton *et al.* 2004; Wang *et al.* 2006). The region is mainly characterized by secondary forests regenerated after cropland abandonment in southern and central New England (Foster 1992) and after logging activities in the north (Dibble, Brissette & Hunter 1999). Such disturbances, together with the introduction of alien species since the 17th century, have promoted the establishment of many plant species. In recent years there has been a major effort to collect data on the extent of plant invasions across the region (Mehrhoff *et al.* 2003). The Invasive Plant Atlas of New England project (IPANE: http://www.IPANE.org) was initiated in 2001 to evaluate the status of invasive species. We can now use these data to evaluate the susceptibility of the region to plant invasions.

The specific questions we aim to answer here are: (i) What are the climate and landscape attributes associated with the abundance of invasive plants in New England? (ii) Are there areas in the region with the potential to become hot spots for invasive plant species? Will these change under the forecasted climate scenarios? (iii) For a particular species, what are

the environmental conditions associated with prolific populations?

Materials and methods

SPECIES DATA

Plot-based surveys on the presence/absence and abundance of invasive species have been gathered by a group of volunteers across New England as part of the IPANE project since 2001 (Mehrhoff *et al.* 2003). Volunteers are initially trained by IPANE staff on plant identification and field data collection methods, and there is periodic follow-up training. For those species that are challenging to identify, verification is undertaken by submitting voucher specimens or photographs. The validity of volunteer-collected data has been assessed in numerous studies (e.g. Galloway *et al.* 2006; Brooks *et al.* 2008) and makes us confident in the reliability of this data set. To date over 5000 field plots have been inventoried and the location of some 11 000 individual incursions of more than 100 invasive species has been documented (Fig. 1). From those records we calculated the total number of invasive plant species at each location (i.e. richness). We also have measures of species abundances, defined as the percent of ground cover occupied by each species present in the plots (10 m radius circles), coded as ordinal categories: *1*: less than 1%; *2*: 1–5%; *3*: 5–25%; *4*: 25–50%; *5*: 50–75%; *6*: 75–100%. The specific species analysed for this study are: *Berberis thunbergii* DC Japanese barberry, a shade-tolerant shrub, *Celastrus orbiculatus* Thunb. oriental bittersweet, a woody vine, *Elaeagnus umbellata* Thunb. autumn olive, a nitrogen-fixing shrub, *Euonymus alata* Thunb. burning bush, a perennial shrub, and *Rosa multiflora* Thunb. ex Murr. multiflora rose, a perennial shrub. These species are frequently found throughout the region, with percentage presence across IPANE sites ranging from 7% to 23%. Of the ~5000 data points available, we used 4200 in our analysis and withheld the remainder records (~900) to validate the different models examined. For more detail on the history of the invasions, the species, and on data collection and protocols see the IPANE web site http://www.IPANE.org.

CLIMATE DATA

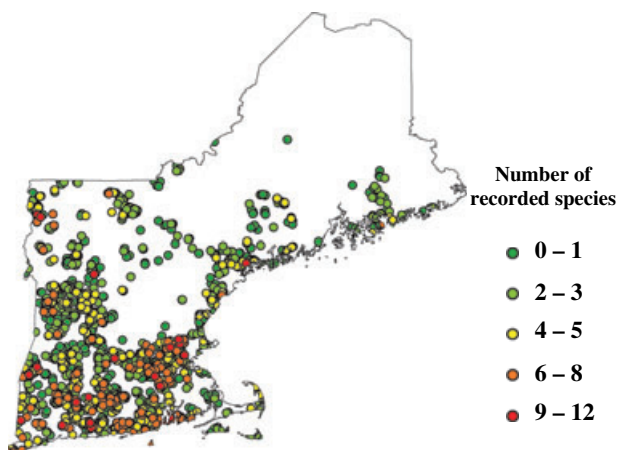Current climate data are averages for 1950–2000 downloaded from the WorldClim data base (Hijmans *et al.* 2005). Each of the IPANE sample locations was intersected with climate data at a 30 arcsec (~1 km) resolution. We selected five climate variables that seemed ecologically most relevant in affecting the degree of invasion (i.e. plant growth, survival and reproduction; Pino *et al.* 2005). We included: (i) mean maximum temperature of the warmest month; (ii) mean minimum temperature of the coldest month (both of which inform us about the tolerance limits of each species with respect to temperature); (iii) mean annual precipitation; (iv) mean precipitation seasonality (i.e. coefficient of variation in precipitation along the year, which measures how evenly rainfall is distributed throughout the year); and (v) mean precipitation of the warmest quarter (which provides a proxy for dryness during the summer months). Maps predicting potential invasive species richness were generated using climate data from the same source at a 5 min spatial scale (~10 km).

Predictions of species richness were also generated under two climate models forecasted for the region for the years 2070–2100 (Wang 2005). Changes in temperature and soil moisture will vary among different climate models making the impact of predicted climate change on invasive species model dependent (Beaumont, Hughes & Pitman 2008). To account for such variability in outcomes we used forecasts from the two most extreme climate predictions from the more than 20 global climate models participating in the IPCC AR4 (Wang 2005). Specifically, we considered outcomes from two widely used general circulation climate models under the B1 scenario (where carbon dioxide forcings are maintained at current levels): CCSM3 and UKMO-HadCM3. These forecasts cover from least to most extreme temperature and precipitation change projections for the Northeast USA. CCSM3 predicts a slight increase in temperature with higher summer precipitation, while HadCM3 forecasts a greater rise in temperature and lower precipitation rates (Fig. S1, Supporting Information).

LAND USE LAND COVER (LULC) DATA

Landscape data for the region are available through the C-CAP program which uses 30-m resolution Landsat Thematic Mapper and Landsat Enhanced Thematic Mapper satellite imagery (NOAA 1995–present) and classifies the images into LULC data. For our analysis, the 12 terrestrial classes identified by the C-CAP analysis were collapsed into six categories: (i) developed, (ii) deciduous and mixed forests, (iii) evergreen forests, (iv) crops, (v) scrub/shrubland and (vi) pastures and grasslands (Table S1, Supporting Information). Road density was obtained from the Census 2000 Topologically Integrated Geographic Encoding and Referencing (TIGER) data base of the United States Census Bureau (2007). Using ArcGIS v9.2 (ESRI 2006) and Hawth's Tools (Beyer 2004) we estimated per cent coverage of each category and total length of roads (km) in 1 and 10 km radii circles around each sampled location. The 10 km radius buffer summaries performed best for the richness analysis, while the 1 km buffers were best for the species percentage cover analyses. As with the climate data, we used 5 min (~10 km grid cells) summaries of the LULC data to generate maps predicting the invasive outcome as a function of the local landscape (Fig. S2).

SITE DATA (HABITAT TYPE AND CANOPY CLOSURE)

Although we were interested in regional-scale predictions based on climate and landscape structure (~10 × 10 km grid cells), we also included information on local site characteristics (i.e. habitat type and canopy closure). Inclusion of site variables allowed us to differentiate the effects of climate and LULC from those due to local characteristics such as habitat type and light availability. As part of the

Number of recorded species

- 0 – 1
- 2 – 3
- 4 – 5
- 6 – 8
- 9 – 12

Fig. 1. Region of New England with number of invasive species recorded at sampled locations (*n* = 4207).

IPANE sampling protocol, volunteers gathered data on habitat and per cent canopy closure at each sampled location (Mehrhoff *et al.* 2003). Habitat types were originally classified into 31 categories, but to simplify the analysis we grouped them into eight general types: edge, deciduous and mixed forests, evergreen forests, open wetlands, closed canopy wetlands, marine influence, open/abandoned and other (this would include sites that were not defined by any of the original types) (Table S1). At each location, an estimate of canopy closure was recorded as an ordinal four-level variable, where the canopy closure ranged from 0 to 100 in increments of 25%.

MODEL DEVELOPMENT

To accommodate the multilevel structure of our data, where we combined several sources of information collected at different scales, we developed hierarchical models (Clark 2005) for the analyses of invasive species richness and species per cent cover data. We developed our models within a Bayesian framework because Bayesian methods are well suited for analyses of hierarchical models (Gelman & Hill 2007). Also, using hierarchical Bayesian techniques allowed us to take into consideration the different sources of uncertainty inherent in our data, process models and the predictions (Clark and Gelfand 2006a,b).

RICHNESS ANALYSIS

We analysed invasive species richness data for the sample locations in New England (n = 4200 points, Fig. 1) as a function of the predictive environmental variables described above. This analysis identified the environmental conditions that may predispose a particular area to be successfully invaded by alien species. The number of invasive species (count data) at location 'i', N_i, was considered to be drawn from a Poisson distribution with mean λ_i that was then estimated as a function of the predictive variables using a ln link function:

$$N_i \sim \text{Poisson}(\lambda_i) \qquad \text{eqn 1}$$

$$\ln(\lambda_i) = X_i\beta + w_i + e_i \qquad \text{eqn 2}$$

X_i is the set of environmental variables at location i (see descriptions above). We ran three different models to test the value of adding different types of environmental data. We first modelled with climate data only (model 1), then climate and LULC data (model 2) and finally, climate, LULC and site data (model 3). The vector of fixed effect coefficients associated with each of the environmental variables, β, also included an intercept term and each term had the following vague prior distribution:

$$\beta_k \sim \text{Normal}(0, 10\,000) \qquad \text{eqn 3}$$

In the case of the site variable 'habitat', a factor, we estimated a different coefficient for each habitat type, h, and each coefficient was centred around the mean ($\beta_{\text{habitat } h} - \beta_{\text{habitat mean}}$).

To account for any spatial correlation in the data we included spatially explicit random effects, w_i, that were estimated as a function of the neighbouring locations within a 10 km radius. In other words, locations nearby were likely to be more similar than those farther away. We tried several neighbourhood radii (1, 5, 10 and 20 km) in a previous analysis and 10 km radius represented best the spatial structure of the data (Ibáñez *et al.* 2009). Spatial effects were assigned to a conditional autoregressive prior distribution (CAR), where $b_{ij} = 1$ if location j is a neighbour of i within that radius, and $b_{ij} = 0$ otherwise:

$$w_i|w_{j,j\neq 1} \sim \text{Normal}\left(\sum_j b_{ij}w_j, \sigma_w^2\right) \qquad \text{eqn 4}$$

$$\sigma_w^2 = 1/\tau_w \text{ where } \tau_w \sim \text{Gamma}(0\cdot5, 0\cdot0005)$$

After initially running this model with only spatial random effects, we also included random error terms, ε_i, with an exchangeable normal prior:

$$\varepsilon_i \sim \text{Normal}(0, \sigma_\varepsilon^2)$$

$$\sigma_\varepsilon^2 = 1/\tau_\varepsilon \text{ where } \tau_\varepsilon \sim \text{Gamma}(0\cdot5, 0\cdot0005) \qquad \text{eqn 5}$$

The random effects at each location were then the sum of the spatial term, w, and the error term, ε, constituting a convolution prior (Besag, York & Mollie 1991; Mollie 1996), which adds flexibility in the allocation of the residuals. The value of the hyperparameter priors for the precision terms, $\tau\sim\text{Gamma}(0\cdot5,0\cdot0005)$, were those commonly used in CAR models (Kesall & Wakefield 1999) and implied a standard deviation centred around $0\cdot05$, which ranged between $0\cdot01$ and $2\cdot5$ with a prior probability of 1%. Given the ln link we used to model the number of invasive species, the constraints we set on the precision terms cover all the variability we observed in the data and still aided the model runs.

PERCENTAGE GROUND COVER ANALYSIS

In a separate model we analysed percentage ground cover data for the study species, given the species was observed at a sample location. Here we used cover data as a proxy for population success and as a measure of potential for further spread. We estimated ground cover as a function of the same environmental variables available: climate, LULC and site characteristics. In this case per cent cover, PC, was recorded in six ordinal categories, from ≤1% to 100% (see Species data). At location i we assigned a probability π_{ik} of falling in each of the six, k, categories (1–6), using an ordered logit model (or proportional odds model) (McCullagh 1980; Cox 1995). Here the coverage categories are distributed as:

$$\text{PC}_i \sim \text{Multinomial}(\pi_{i,1:6}, 1) \qquad \text{eqn 6}$$

The probability of falling within a specific category k site i, π_{ik}, is described as follows:

$$P(\text{PC}_i = 1) = \pi_{i1} = \frac{1}{1 + e^{X_i\beta + w_i - \text{Upperlimit}_1}}$$

$$P(\text{PC}_i = 2) = \pi_{i2} = \frac{1}{1 + e^{X_i\beta + w_i - \text{Lowerlimit}_2}} - \frac{1}{1 + e^{X_i\beta + w_i - \text{Upperlimit}_2}}$$

...

$$P(\text{PC}_i = 6) = \pi_{i6} = 1 - \frac{1}{1 + e^{X_i\beta + w_i - \text{Upperlimit}_6}}$$

$$\text{eqn 7}$$

X_i represents the matrix of environmental data at each location. We ran model 1 (only climate), model 2 (climate and LULC) and model 3 (climate, LULC and site variables) to determine which environmental factors best fit the data. β represents the vector of the intercept and fixed effects coefficients associated with the predictor variables and was estimated similarly as in the richness analysis. Spatial correlation in the data was accounted for by also including location-level spatially explicit random effects, w. In this case adding an additional random term, ε, did not significantly improve the fit of the models and was not used. Upper and lower

limits reflect the boundaries of the ordinal categories of the data (see Species data).

Model simulations were run in OpenBUGS 1.4 (Thomas *et al.* 2006). Simulations were run until convergence of the parameters was ensured (~50 000 iterations), after which posterior mean estimates of the parameters were calculated from another 50 000 iterations.

MODEL SELECTION

Model selection among the three variations described (models 1, 2 and 3), was made on the basis of calculating the posterior predictive loss (PPL) function (Gelfand & Ghosh 1998) for the richness analysis [the convolution prior did not allow us to estimate deviance information criterion (DIC), see below]. For the analysis of the species per cent cover data we estimate the DIC (Spiegelhalter *et al.* 2000) to compare among models. In both cases, models that minimized PPL or DIC values were considered to have a better fit to the data.

MODEL EVALUATION

In order to evaluate our predictions with independent data we withheld the last ~900 data points collected in the IPANE data base. These additional data points fall within the geographical and climatic ranges considered in the analyses. We compared our regional predictions with these independent data and evaluated the goodness of fit of each model by calculating the sum of squares of *Predicted-Observed*. In case of having more than one observation per grid cell we used the mean value for species richness and for per cent ground cover. This method is easy to use and does not make any assumptions about the stochasticity implicit in the data (Hilborn & Mangel 1997). Models that minimized the sum of squares were considered to be better at predicting the independent data. In the case of species richness, we also calculated the number of grid cells correctly estimated for number of invasive species, with a two species margin.

VISUALIZING PREDICTIONS

As we are interested in where invasive species may be particularly successful and not only where they already are well-established, we generated maps of the potential number of invasive species (i.e. richness) for the whole New England region. To generate prediction maps we used the posterior parameter values (*β* coefficients), regional data on climate (current and forecasted), current LULC data and the average canopy closure estimated from our data set (37%). This method allows visualization of grid cell predictions on a continuous map surface.

Results

MODEL SELECTION

In the case of the richness data analysis, model 1, climate only, had the lowest PPL, although the other two models did not substantially differ in fit. When looking at particular species analyses for cover, model 3 (climate, LULC and site) showed the best goodness of fit (lowest DIC) for three species, *B. thunbergii*, *C. orbiculatus* and *R. multiflora.* For the other two species, *E. umbellata* and *E. alata*, model 1 preformed best, although models 2 and 3 were not very different.

MODEL EVALUATION

When comparing our predictions with the independent data (Table 1), model 3 (climate, LULC and site) had the best goodness of fit (smallest sum of squares *predicted–observed*) for the richness data and also for four of the five species, *B. thunbergii*, *C. orbiculatus*, *E. alata* and *R. multiflora.* The only species in which model 1 (climate only) fitted the independent data best

Table 1. Model selection criteria: predictive loss for richness and DIC for abundance. Model evaluation: least squares sum (all) and per cent of grid cells correctly identified (richness only), comparisons of independent data with model predictions, lower values indicate better predictions of the independent data for the least squares method, while higher values show higher per cent in correct identification for the grid cells; bold values indicate the best model predicting the independent data

	Richness	Abundance				
		Berberis thunbergii	*Celastrus orbiculatus*	*Elaeagnus umbellata*	*Euonymus alata*	*Rosa multiflora*
Model selection criteria: predictive loss and DIC						
Model 1 (Climate)	**16 770**	3346	2836	**977**	**834**	> > 3000
Model 2 (Climate + LULC)	16 780	2460	2751	997	866	2765
Model 3 (Climate + LULC + Site)	16 780	**2434**	**2736**	988	864	**2725**
Number of data points:	4207	877	879	329	290	982
Model evaluation: using independent data – least squares						
Model 1 (Climate)	480·82	46·5	40·6	**15·96**	16·3	42·5
Model 2 (Climate + LULC)	441·16	44·8	24·9	20·84	1·2	28·2
Model 3 (Climate + LULC + Site)	**371·79**	**43·8**	**17·4**	17·98	**0·82**	**14·5**
Number of data points:	911	164	257	85	49	287
Per cent of grid cells correctly identified for number of invasive species						
Model 1 (Climate)	90	–	–	–	–	–
Model 2 (Climate + LULC)	**92**	–	–	–	–	–
Model 3 (Climate + LULC + Site)	90	–	–	–	–	–
Number of grid cells with independent data	164	–	–	–	–	–

DIC, deviance information criterion; LULC, Land Use Land Cover.

was *E. umbellata*. When looking at the per cent of grid cells for which the number of invasive species were correctly predicted the three models were very similar (Table 1). Given the results of the model selection and evaluation, we are reporting parameter results and regional predictions from model 3 to facilitate comparisons among the two analyses, richness and species' per cent cover (Fig. 2, for specific parameter values, see Table S2).

PARAMETER ESTIMATES

Both the richness and abundance models indicate that climate variables were the most influential in determining the co-occurrences or cover of invasive species (Fig. 2). After climate, site factors were next in importance to model fit, followed by LULC variables. Species richness was higher in areas with warmer temperatures in both summer and winter along with higher summer precipitation and precipitation seasonality. Conversely, high annual precipitation had a negative effect on

the number of invasive plant species. With respect to LULC, areas with a large proportion of evergreen forests or croplands were associated with lower richness, while roads had a positive effect on the number of invasive species. Although not significant, developed areas, deciduous forests and pastures seem to promote the occurrence of invasive species. Edge habitat most favoured a large number of invasive species and evergreen forest habitat had lower numbers. Higher levels of canopy closure were also associated with lower invasive species richness.

Results of the per cent ground cover analyses show differences in the direction and magnitude of the effect of each predictive variable among the five species (Fig. 2, Table S2). Warmer temperatures, mainly in summer, were beneficial for *C. orbiculatus*, *E. alata* and *R. multiflora*, but have a negative effect on *B. thunbergii* and *E. umbellata*. High annual precipitation positively affected *C. orbiculatus*, *E. umbellata* and *E. alata*, but had a negative influence on *B. thunbergii* and *R. multiflora*.
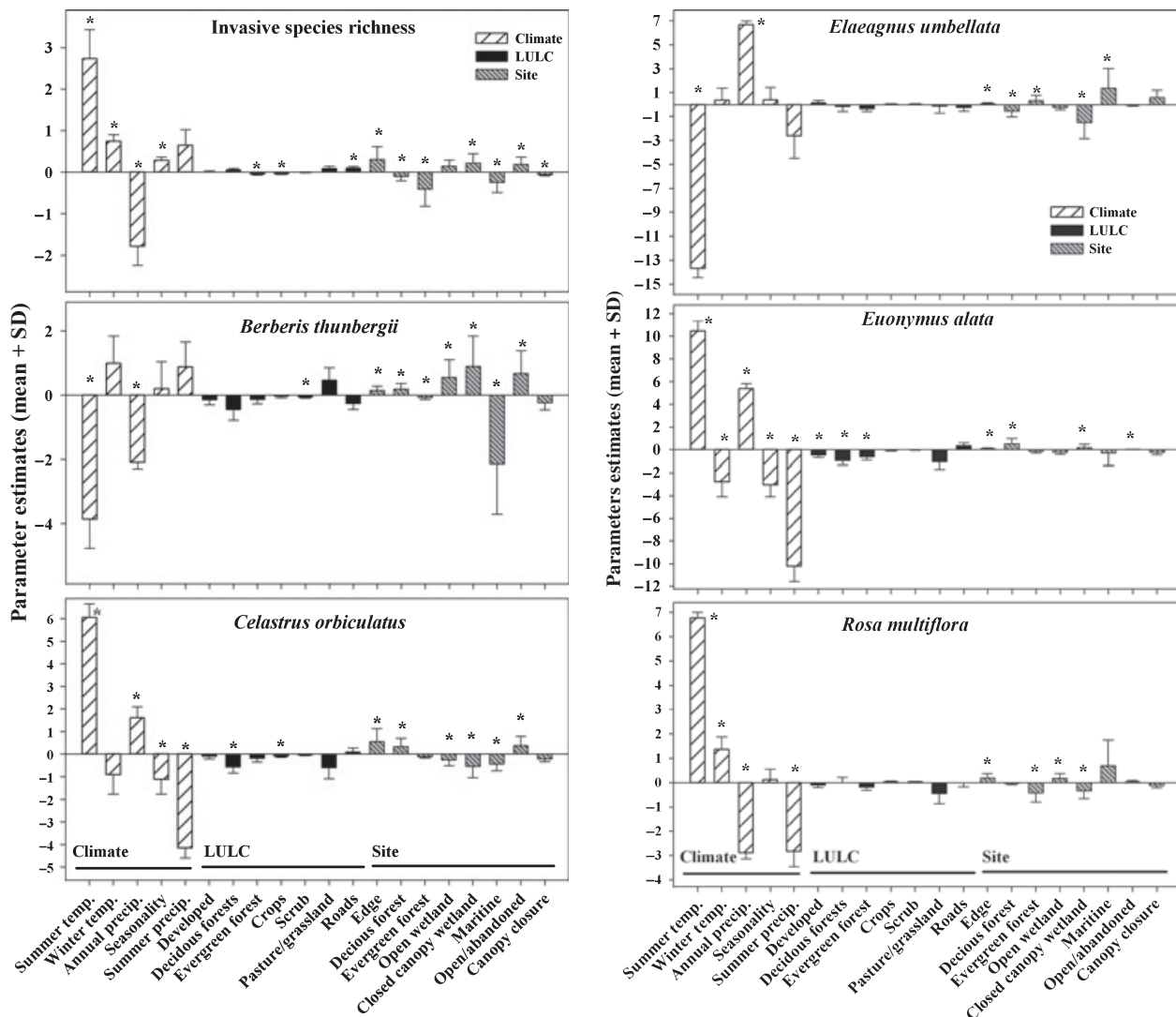

Fig. 2. Posterior means (+SD) of the parameters estimated in the analyses (for comparisons among variables the parameter posterior means have been multiplied by the average variable value). Asterisks indicate that the 95% credible interval around the parameter mean values did not include zero.

The per cent of shrub/scrubland was the only significant LULC variable for *B. thunbergii* and *C. orbiculatus* was negatively associated with deciduous forests and croplands. Developed areas, deciduous and evergreen forests also diminished the per cent cover of *E. alata*. The only general pattern, lower per cent cover of these species, was around landscapes where evergreen forests were dominant.

Habitat-level variables were particularly species specific, though edge habitat seemed to favour the growth of all five species. Deciduous forest sites promoted the abundance of *B. thunbergii*, *C. orbiculatus* and *E. alata*, and decreased *E. umbellata* and *R. multiflora*. Evergreen forest sites had a significant negative effect on *B. thunbergii* and *R. multiflora* abundance, contrary to its effect on *E. umbellata*. Open wetlands had a positive effect on *B. thunbergii* and *R. multiflora* cover; closed canopy wetlands enhanced cover of *B. thunbergii* and *E. alata*, and were detrimental to *C. orbiculatus*, *E. umbellata* and *R. multiflora*. Maritime-influenced sites were only favourable for *E. umbellata* and *R. multiflora*. Open/abandoned habitats had a significantly positive influence on *B. thunbergii*, *C. orbiculatus* and *E. alata*. In all species, except for *E. alata*, the canopy closure coefficient was negative, suggesting that canopy closure could have had a negative effect on the population growth of those species.

POTENTIAL HOT SPOT MAPS

We generated predictions (the number of potential invasive species) based on the region's climate and LULC. In the case of climate, we also generated predictions under the two forecasted climate extremes described in the Methods section (Fig. 3). Invasive species richness is predicted to generally increase under the climate change projection with slight increases in temperature with higher summer precipitation and slightly lower annual precipitation (CCSM3), but decrease under projections with greater increases in temperature and lower precipitation (HadCM3) (Fig. 3).

Comparing predictions among the current and future climate scenarios, we find that all species, except *E. umbellata*, have a higher risk of increased abundance relative to current conditions under moderate climate change (CCSM3) while their incidence may be reduced under the more extreme climate change forecast (HadCM3) (Fig. 4).

Discussion

When management resources are limited, identifying those conditions that promote invasion, ahead of the invasion, will be critical to the success of early detection and rapid response control efforts (Andow 2005). Our model predictions provided explicit spatial locations where early detection efforts should be focused using our network of volunteers as these areas are most likely to become invaded. And what it is more our approach and the models developed in this study can be applied elsewhere and to any species.

The factors promoting a particular introduction are defined by a set of temporal and spatial circumstances that may not apply to other species, times or locations (Davis 2009). Such individuality motivated us to focus on a particular region, New England, and instead explore the regional attributes that may
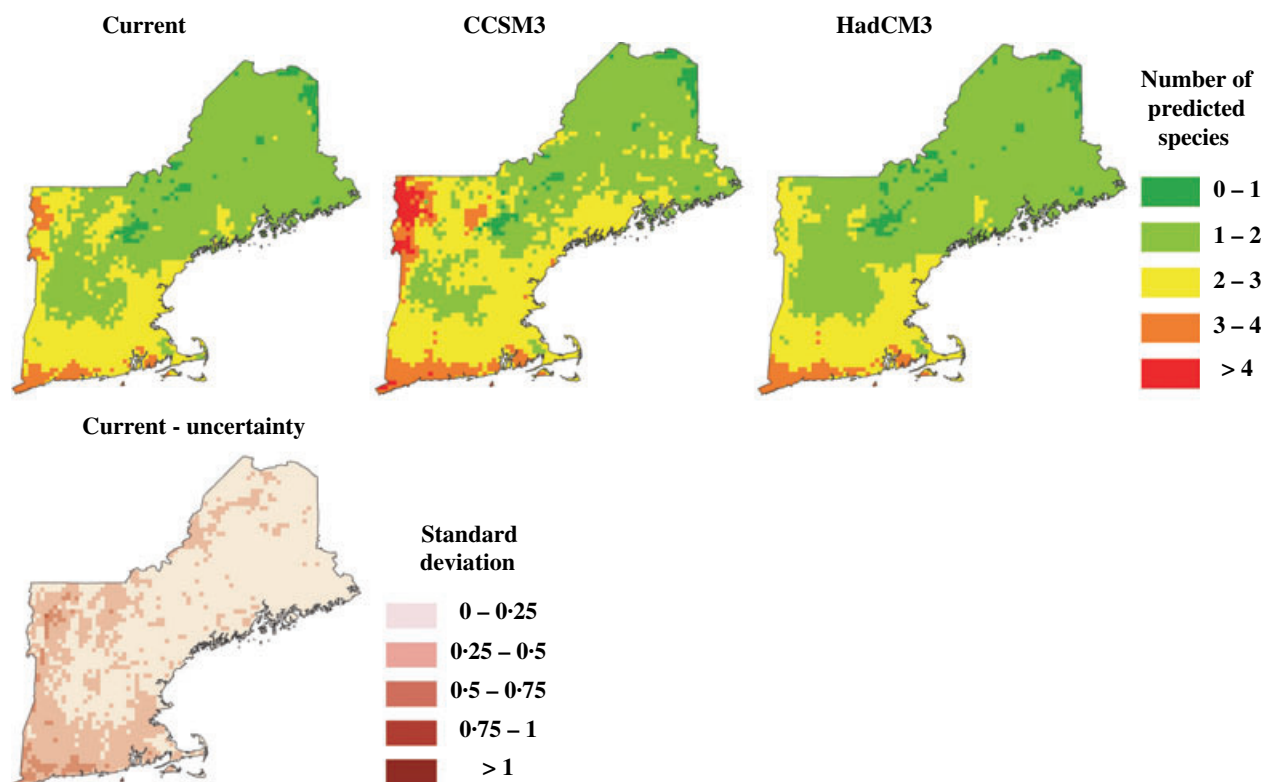

Fig. 3. Predicted invasive species richness under current and two climate scenarios.

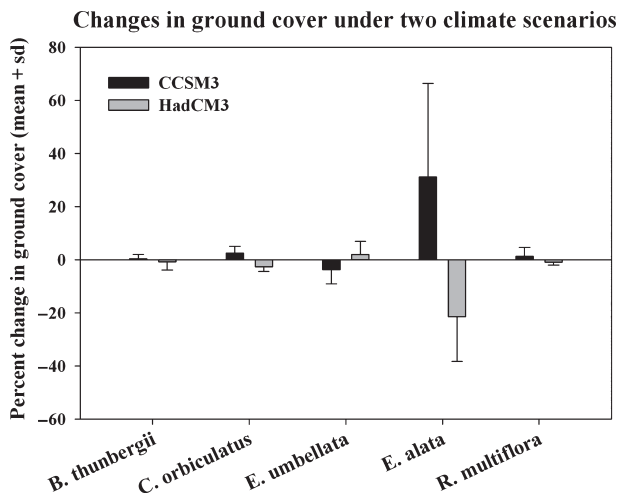## Changes in ground cover under two climate scenarios

Fig. 4. Changes in ground cover (current–forecasted) for the studies species under the two climate scenarios included in the analyses.

have fostered plant invasions in the area. We concentrated on the climatic conditions that may have promoted the establishment of invasives, but viewed the process in the context of the landscape these species encountered as they spread. Including LULC variables in our analysis allowed indirect assessment of many aspects of the invasion process like propagule introduction and build-up and species naturalization, which are highly linked to human development and to the incidence of disturbances. Unique magnitudes, directions and combinations of climatic, LULC and habitat-level factors influencing the distribution of each species in the study region are evident. Our results provide clear evidence for the unique nature of plant invasions when considering individual species, although, some general patterns were also identified given the range of environmental conditions found within the region.

Even with species-specific differences, we were able to identify potential hot spots for plant species invasions. Relatively warmer areas with sufficient summer precipitation appear to be associated with invasive species occurrence. Locations with landscapes dominated by deciduous forests, pastures and grasslands and with a high number of roads seemed to foster higher numbers of invasive species. At the site level, disturbed habitats (i.e. edges and open areas) were also associated with the establishment of these species. These patterns confirm the role of human activities in the occurrence of invasive species. Our regional predictions reflected our data on the number of invasive species, revealing a higher incidence in the southern part of the region. Additionally, the analysis forecasted potentially larger numbers of invasives in north-western New England, where the presence of invasive species is currently relatively low, but the climate and landscape structure point to a high risk for further invasions. Such predictions can also be generated at different spatial scales aided with the implementation of local invasive species management plans.

The wide distributions of many current invasive plant species prompted us to concentrate on the identification of prob-

lematic populations (i.e. potential sources of propagules into uncolonized areas). Many of the currently introduced invasive species are able to establish in a large variety of sites (Ibáñez *et al.* 2009), but will only develop stable and highly fecund populations in areas where the initial availability of propagules is high and where the environmental conditions are optimal for the growth and reproduction of the species. In this case climate and site data informed us on the environmental factors affecting the population, and we used LULC data as a rough proxy for landscapes that encourage establishment and spread of invasive plants (often linked to human activities).

Most of the work done in identifying optimal sites for plant invasions focuses on the climatic and specific habitat characteristics that promote the growth of a particular species (e.g. Buckley, Briese & Rees 2003; Hendrix *et al.* 2008; Gasso *et al.* 2009). However, for four of the five species examined, including neighbourhood LULC data, improved the predictions; this reinforces the value of combining predictive variables of different types and spatial scales (Meentemeyer *et al.* 2008; Ibáñez *et al.* 2009). Land use change has influenced the growth and spread of invasive plants in the past (DeGasperis & Motzkin 2007; Mosher *et al.* 2009) and while we assumed here that landscape relationships will remain stable, we should also keep in mind that future landscape dynamics may alter the prevalence of invasive plants. In these analyses, each species showed a unique pattern that determined which combinations of climate, LULC and site properties promoted the growth of productive populations, confirming the idiosyncrasy of the invasion process and the difficulties of developing generalizations from single species analyses.

In addition, climate change will also affect plant species distributions; therefore, taking into account future climate predictions will be imperative if we want to have a comprehensive understanding of all possible outcomes. An increase in temperatures may open higher latitudes to colonization by plants from warmer regions. And, a world-wide decrease in soil moisture (IPCC 2007) will shape the future distribution of many species in accordance with their tolerance to drought. We then extended our forecasts to incorporate future climates from predictions generated by general circulation models. The moderate climate projections (CCSM3) have the potential to increase plant species invasions. In the case of four of our studied species, it could also have a positive effect on their population growth. On the other hand, predictions under the more extreme temperature and drought projections (HadCM3) seem to indicate a lower risk of plant invasions and population abundance. However, these forecasts are limited to the extent that they do not take into consideration that the native vegetation would also be affected, and that the competitive interactions among species and other factors may also be modified. Nevertheless, having the range of possible outcomes will help us to single out the most likely areas to become major focal points of invasion and spread.

The regional data collected by the IPANE volunteers (S. Treanor, J.A. Silander, J.M. Allen & L.J. Mehrhoff, unpublished data) provided us with the often unavailable, but necessary, large-scale data set to analyse the invasion process

and to generate predictions of future patterns of plant invasions. The link between citizen scientists, academic research and natural resource managers is fully elucidated by the process and data presented here and demonstrates an improved strategy for planning early detection and monitoring strategies. These species-specific predictions, in particular the information about the direction and magnitude of the environmental drivers affecting population growth, can be now used directly by natural resource managers in early detection work and in planning eradication and control campaigns. Furthermore, the multivariate approach followed in this study can be implemented in other locations and for other species using the same or different set of predicted variables with the purpose of generating realistic predictions of plant invasive potential.

Acknowledgements

References

Andow, D.A. (2005) Characterizing ecological risks of introductions and invasions. *Invasive Alien Species. A New Synthesis* (eds H.A. Mooney, R.N. Mack, J.A. McNeely, L.E. Neville, P.J. Schei & J.K. Waage), pp. 84–103. Island Press, Washington, DC.

Barton, A.M., Brewster, L.B., Cox, A.N. & Prentiss, N.K. (2004) Non-indigenous woody invasive plants in a rural New England town. *Biological Invasions*, **6**, 205–211.

Beaumont, L.J., Hughes, L. & Pitman, A.J. (2008) Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters*, **11**, 1135–1146.

Besag, J., York, J. & Mollie, A. (1991) Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, **43**, 1–59.

Beyer, H.L. (2004) *Hawth's Tools for ArcGIS* http://www.spatialecology.com/htools.

Blaustein, R.J. (2001) Kudzu invasion into Southern United States life and culture. *The Great Reshuffling: Human Dimensions of Invasive Species* (ed. J.A. McNeeley), pp. 55–62. IUCN, Gland, Switzerland and Cambridge, UK. The World Conservation Union.

Brooks, W.R., Mellor, D.T., Ehrenfield, E., McGowan, E., Goodell, E. & Jordan, R. (2008) *Using citizen scientists to collect plant distribution data in NY/NJ forest land.* Proceedings for the Northeast Natural History Conference, Albany, NY, 17–18 April 2008.

Buckley, Y.M., Briese, D.T. & Rees, M. (2003) Demography and management of the invasive plant species *Hypericum perforatum*. II. Construction and use of an individual-based model to predict population dynamics and the effects of management strategies. *Journal of Applied Ecology*, **40**, 494–507.

Cadenasso, M.L. & Pickett, S.T.A. (2001) Effect of edge structure on the flux of species into forest interiors. *Conservation Biology*, **15**, 91–97.

Chytry, M., Vojtech, J., Pyseck, P., Hajek, O., Knollova, I., Lubomir, T. & Danihelka, J. (2008) Separating habitat invasibility by alien plants from the actual level of invansion. *Ecology*, **89**, 1541–1553.

Clark, J.S. (2005) Why environmental scientists are becoming Bayesians. *Ecology Letters*, **8**, 2–14.

Clark, J.S. & Gelfand, A.E. (2006a) A future for models and data in ecology. *Trends in Ecology and Evolution*, **21**, 375–380.

Clark, J.S. & Gelfand, A.E.E. (2006b) *Hierarchical Modelling for the Environmental Sciences: Statistical Methods and Applications.* Oxford University Press, Oxford, UK.

Compagnoni, A. & Halpern, C.B. (2009) Properties of native plant communities do not determine exotic success during early forest succession. *Ecography*, **32**, 449–458.

Cox, C. (1995) Location-scale cumulative odds models for ordinal data: a generalized nonlinear model approach. *Statistics in Medicine*, **14**, 1191–1203.

Davis, M.A. (2009) *Invasion Biology*. Oxford University Press, New York.

DeGasperis, B.G. & Motzkin, G. (2007) Windows of opportunity: historical and ecological controls on *Berberis thunbergii* invasions. *Ecology*, **88**, 3115–3125.

Dibble, A.C., Brissette, J.C. & Hunter, M.L. (1999) Putting community data to work: some understory plants indicate red spruce regeneration habitat. *Forest Ecology and Management*, **114**, 275–291.

Eschtruth, A.K. & Battles, J.J. (2009) Assessing the relative importance of disturbance, herbivory, diversity, and propagule pressure in exotic plant invasion. *Ecological Monographs*, **79**, 265–280.

ESRI (2006) *U.S. USGS 1:100,000 Topographic Quadrangle Series Indexes.* ESRI, Redlands, CA.

Farnsworth, E.J. (2004) Patterns of plant invasions at sites with rare plant species throughout New England. *Rhodora*, **106**, 97–117.

Forseth, I. & Innis, A. (2004) Kudzu (*Pueraria montana*): history, physiology, and ecology combined to make a major ecosystem threat. *Critical Reviews in Plant Sciences*, **23**, 401–413.

Foster, D.R. (1992) Land-use history (1730–1990) and vegetation dynamics in Central New England, USA. *Journal of Ecology*, **80**, 753–772.

Fridley, J.D., Stachowicz, J.J., Naeem, S., Seabloom, E.W., Smith, M.D., Stohlgren, T.J., Tilman, D. & Von Holle, B. (2007) The invasion paradox: reconciling pattern and process in species invasions. *Ecology*, **88**, 3–17.

Galloway, A.W.E., Tudor, M.T. & Haegen, M.V. (2006) The reliability of citizen science: a case study of Oregon White Oak stand surveys. *Wildlife Society Bulletin*, **34**, 1425–1429.

Gasso, N., Sol, D., Pino, J., Dana, E.D., Lloret, F., Sanz-Elorza, M., Sobrino, E. & Vila, M. (2009) Exploring species attributes and site characteristics to assess plant invasions in Spain. *Diversity and Distributions*, **15**, 50–58.

Gelfand, A.E. & Ghosh, S.K. (1998) Model choice: a minimum posterior predictive loss approach. *Biometrika*, **85**, 1–11.

Gelman, A. & Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge University Press, New York.

Hayes, K.R. & Barry, S.C. (2008) Are there any consistent predictors of invasion success? *Biological Invasions*, **10**, 483–506.

Hendrix, P.F., Callaham, M.A., Drake, J.M., Huang, C.-Y., James, S.W., Snyder, B.A. & Zhang, W. (2008) Pandora's box contained bait: the global problem of introduced earthworms. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 593–613.

Herron, P., Martine, C., Latimer, A. & Leicht-Young, S.A. (2007) Invasive plants and their ecological strategies: prediction and explanation of woody plant invasion in New England. *Diversity and Distributions*, **13**, 633–644.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hilborn, R. & Mangel, M. (1997) *The Ecological Detective. Confronting Models with Data.* Princeton University Press, Princeton, NJ.

Hobbs, R.J. & Mooney, H.A. (2005) Invasive species in a changing world: interactions between global change and invasives. *Invasive Alien Species. A New Synthesis* (eds H.J. Mooney, R.N. Mack, J.A. McNeely, L.E. Neville, P.J. Schei & J.K. Waage), pp. 310–331. Island Press, Washington, DC.

Ibáñez, I., Silander, J.A., Wilson, A.M., LaFleur, N., Tanaka, N. & Tsuyama, I. (2009) Multivariate forecasts of potential distributions of invasive plant species. *Ecological Applications*, **19**, 359–375.

IPCC (2007) Climate change 2007: the physical science basis. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor & H.L. Miller). Cambridge University Press, Cambridge, UK and New York.

Kesall, J.E. & Wakefield, J.C. (1999) Discussion of "Bayesian models for spatially correlated disease and exposure data," by Best *et al. Bayesian Statistics 6* (eds J.M. Bernardo, J.O. Berger, A.P. Dawid & A.F.M. Smith), p. 151. Oxford University Press, Oxford, UK.

Levine, M.M. & D'Antonio, C.M. (1999) Elton revisited: a review of evidence linking diversity and invasibility. *Oikos*, **87**, 1522–1536.

Mack, M.C. & D'Antonio, C.M. (1998) Impacts of biological invasions on disturbance regimes. *Trends In Ecology & Evolution*, **13**, 195–198.

McCullagh, P. (1980) Regression models for ordinal data. *Journal of the Royal Statistical Society, Series B*, **42**, 109–142.

McDonald, R.I., Motzkin, G. & Foster, D.R. (2008) Assessing the influence of historical factors, cotemporary processes, and environmental conditions on the distribution of invasive species. *Journal of the Torrey Botanical Society*, **135**, 259–270.

Meentemeyer, R.K., Rank, N.E., Anacker, B.L., Rizzo, D.M. & Cushman, J.H. (2008) Influence of land-cover change on the spread of an invasive forest pathogen. *Ecological Applications*, **18**, 159–171.

Melbourne, B.A., Cornell, H.V., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A.L., Hall, R.J., Harrison, S., Hastings, A., Holland, M., Holyoak, M., Lambrinos, J., Moore, K. & Yokomizo, H. (2007) Invasion in a heterogeneous world: resistance, coexistence or hostile takeover? *Ecology Letters*, **10**, 77–94.

Mehrhoff, L.J. (2000) Immigration and expansion of the New England Flora. *Rhodora*, **102**, 280–298.

Mehrhoff, L.J., Silander, J.A., Jr, Leicht, S.A., Mosher, E.S. & Tabak, N.M. (2003) *IPANE: Invasive Plant Atlas of New England.* Department of Ecology & Evolutionary Biology, University of Connecticut, Storrs, CT. http://www.ipane.org.

Meyerson, L.A. & Mooney, H.A. (2007) Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment*, **5**, 199–208.

Milbau, A. & Stout, J.C. (2008) Factors associated with alien plants transitioning from casual, to naturalized, to invasive. *Conservation Biology*, **22**, 308–317.

Milbau, A., Stout, J.C., Graae, B.J. & Nijs, I. (2009) A hierarchical framework for integrating invasibility: experiments incorporating different factors and spatial scales. *Biological Invasions*, **11**, 947–950.

Miller, J.H. (1991) Control kudzu on your land. *Forest Farmer*, **51**, 7–11.

Mollie, A. (1996) Bayesian mapping of disease. *Markov Chain Monte Carlo in Practice* (eds W.R. Gliks, S. Richardson & D.J. Spiegelhalter), pp. 359–379. Chapman & Hall, New York.

Mosher, E.S., Silander, J.A., Jr & Latimer, A. (2009) The role of land-use history in major invasions by woody plant species in the northeastern North American Landscape. *Biological Invasions*, **11**, 2317–2328.

NOAA (1995–present) *National Oceanographic and Atmospheric Administration. Coastal Services Center. The Coastal Change Analysis Program (C-CAP).* NOAA Coastal Services Center, Charleston, SC. http://www.csc.noaa.gov/crs/lca/ccap.html.

Pimentel, D. (ed.) (2002) *Biological Invasions: Economic and Environmental Costs of Alien Plant, Animal, and Microbe Species.* CRC Press, Boca Raton, FL.

Pino, J., Font, X., Carbo, J., Jove, M. & Pallares, L. (2005) Large-scale correlates of alien plant invasions in Catalonia (NE of Spain). *Biological Conservation*, **122**, 339–350.

Rejmanek, M., Richardson, D.M., Higgins, S.I., Pitcairn, M.J. & Grotkopp, E. (2005) Ecology of invasive plants: state of the art. *Invasive Alien Species. A New Synthesis* (eds H.A. Mooney, R.N. Mack, J.A. McNeely, L.E. Neville, P.J. Schei & J.K. Waage), pp. 104–161. Island Press, Washington, DC.

Richardson, D.M. & Pysek, P. (2006) Plant invasions: merging the concepts of species invasiveness and community invisibility. *Progress in Physical Geography*, **30**, 409–431.

Silander, J.A., Jr & Klepeis, D.M. (1999) The invasion ecology of Japanese barberry (*Berberis thunbergii*) in the New England landscape. *Biological Invasions*, **1**, 189–201.

Soberon, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**, 1–10.

Spiegelhalter, D.J., Best, N., Carlin, B.P. & Linde, A.V.D. (2000) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 583–639.

Thomas, A., O'Hara, R., Ligges, U. & Sturts, S. (2006) *Making BUGS Open. R News*, **6**, 12–17.

USCB United States Census Bureau (2007) *Topologically Integrated Geographic Encoding and Referencing (TIGER).* http://www.census.gov/geo/www/tiger/.

Vila, M., Corbin, J.D., Dukes, J.S., Pino, J. & Smith, S.D. (2007) Linking plant invasions to environmental change. *Terrestrial Ecosystems in a Changing World* (eds J. Canadell, D. Pataki & L.F. Pitelka), pp. 115–124. Springer, Berlin.

Wang, G. (2005) Agricultural drought in a future climate: results from 15 global climate models participating in the IPCC 4th assessment. *Climate Dynamics*, **25**, 739–753.

Wang, Q., An, S.Q., Ma, Z.J., Zhao, B., Chen, J.K. & Li, B. (2006) Invasive *Spartina alterniflora*: biology, ecology and management. *Acta Phytotaxonomica Sinica*, **44**, 559–588.

Woodward, F.I. (1987) *Climate and Plant Distributions.* Cambridge University Press, Cambridge.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Table S1. Original LULC and habitat land classifications and their corresponding categories used in the analyses

Table S2. Parameter estimates, posterior means (SD) and 95% credible intervals (CI), for the species Richness analysis and for the Percent Cover analysis of the five species. Values in bold indicate that the CI did not include zero, for habitat coefficients we indicate if the CI for $(\beta_{\text{habitat } k} - \beta_{\text{habitat mean}})$ did not include zero

Fig. S1. Regional climate, last 30 year averages (current) and forecasted, for the region of New England. Temperature (°C), precipitation (mm).

Fig. S2. Land Use Land Cover percentages along the region of New England (% cover).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.