11-15-2012

# Challenges in predicting the future distributions of invasive plant species

Chad C. Jones
*Connecticut College*, cjones8@conncoll.edu

## Recommended Citation

Jones, Chad C. 2012. Challenges in predicting the future distributions of invasive plant species. Forest Ecology and Management, 15 November 2012, Vol. 284, pp. 69-77. The views expressed in this paper are solely those of the author.

# Challenges in predicting the future distributions of invasive plant species

Chad C. Jones[a]

[a]Department of Botany, 270 Mohegan Ave, Connecticut College, New London, CT 06320, USA Phone:860-439-5304; Fax: 860-439-2519; cjones8@conncoll.edu

# Abstract

Species distribution models (SDMs) are increasingly used to predict distributions of invasive species. If successful, these models can help managers target limited resources for monitoring and controlling invasive species to areas of high invasion risk. Model accuracy is usually determined using current species distributions, but because invasive species are not at equilibrium with the environment, high current accuracy may not indicate high future accuracy. I used 1982 species distribution data from Bolleswood Natural Area, Connecticut, USA, to create SDMs for two forest invaders, *Celastrus orbiculatus* and *Rosa multiflora*. I then used more recent data, from 1992 and 2002, as validation data sets to determine how model accuracy changed over time and if current and future accuracy were related. I also tested if three alternative approaches – iterative modeling, alternative methods of choosing suitability thresholds and using a risk assessment framework – improved accuracy in predicting future distributions. Model accuracy declined over time with greater declines for models of the species (*Celastrus*) with the higher initial accuracy. By 2002, 49% of *Celastrus* and 85% of *Rosa* new occurrences were correctly predicted by models. Neither iterative modeling nor alternative thresholds improved accuracy of predicting 2002 occurrences, but a risk assessment framework showed promise for guiding monitoring efforts.  These results suggest that measures of current accuracy may not indicate a model's predictive accuracy and must be used cautiously. Distinguishing between predictions of current and future distributions is critical. While iterative models were not successful in this study, I argue that using models in a risk assessment framework closely tied to monitoring will greatly increase the utility of SDMs for managing invasive species.

*Keywords:* Celastrus orbiculatus*, equilibrium assumption*, *Maxent*, *model accuracy*, Rosa multiflora

## 1. Introduction

Exotic plant species can greatly impact natural communities around the world (Vitousek *et al.*, 1997). Even forests, which have often been considered resistant to the spread of exotic species, have been invaded (Martin *et al.*, 2009). If managers of forest reserves and other natural areas can determine where invasions are most likely to occur in the future, they can focus monitoring and control efforts on those areas. Monitoring invasive species is resource intensive (Rew *et al.*, 2005), so being able to focus limited resources is highly beneficial. Species distribution models (SDMs) have been increasingly used to predict potential distributions of invasive species (Elith *et al.*, 2010), which can facilitate early detection of invasive species and maximize monitoring efficiency. In recent years, there has been a proliferation of modeling approaches that can be used to model invasive species (Elith *et al.*, 2006).

There are important limitations to using SDMs to model potential distributions of invasive species. The methods used to develop SDMs assume that species are in equilibrium with their environment (Guisan and Zimmermann, 2000), meaning that they occur in all suitable habitats. Invasive species, especially in early stages of invasion, may occur only in a subset of habitats in which they can thrive, simply due to lack of dispersal (Václavík and Meentemeyer, 2012). As a result, models may not predict the full distribution of potential invasions. While species might be expected to approach equilibrium as time passes following initial invasion, it is usually not possible to tell how close an invasive species is to equilibrium. In addition, many species are not in equilibrium with climate even after centuries (Svenning and Skov, 2004) and given ongoing changes in climate and land use, equilibrium may never occur. The degree to which an invasive species is in equilibrium is also scale dependent. Even if a species is at equilibrium in the invaded range as a whole, it may be far from equilibrium in smaller areas within the invaded range, and vice-versa.

The problem of non-equilibrium in SDMs for invasive species has been widely acknowledged (Thuiller *et al.*, 2005; Václavík and Meentemeyer, 2009; Elith *et al.*, 2010; Václavík and Meentemeyer, 2012) and several approaches have been developed to address it, such as using native range data (Ibanez *et al.*, 2009) or mechanistic niche models (Kearney and Porter, 2009). Several recent studies have tested the effects of lack of equilibrium by using only a subset of available records in a limited geographic area to simulate earlier invasion stages (Thuiller *et al.*, 2004; Dupin *et al.*, 2011) or by using simulated species (Václavík and Meentemeyer, 2012). All of these studies suggest that models created during early stages of invasion will poorly predict future distributions of invasive species as they expand. For example, Sutherst and Bourne (2009) showed qualitatively that early models did not predict new areas of expansion of an invasive tick in Africa, despite high model accuracy at the time models were created. However, no study has quantitatively assessed how accuracy of a given model changes over time, which would help us to understand how current models may fare in the future.

The implications of lack of equilibrium are further compounded by the methods used to assess model accuracy. Most studies assess model accuracy by testing model results with data left out during model creation (Fielding and Bell, 1997). This provides an assessment of accuracy in predicting current distributions of the species. Predicting current distributions differs significantly from predicting future distributions, but these are often not clearly distinguished (see Václavík and Meentemeyer, 2009 for discussion). Predicting future distributions is often the stated objective of the models, but seldom can be directly assessed (Elith *et al.*, 2010). Because invasive species are not in equilibrium with the environment, accuracy in predicting current distributions and future distributions may not correspond. The implicit assumption in many studies seems to be that models that best predict the current distribution will also best predict the future distribution, but this assumption has not been tested. If this is not the case, then measures of current accuracy are not useful for determining which models will best predict future distributions.

When historical data on the spread of invasive species are available, the relationship between current and future predictions can be evaluated directly (Elith *et al.*, 2010). Earlier data can be used to create models and more recent data can be used to test the ability of the models to predict new occurrences over time. The use of historical data can also allow evaluation of possible methods to improve predictions of future distributions. There are many potential approaches for improving predictions including: (i) iterative modeling, (ii) alternative methods for setting suitability thresholds, and (iii) using a risk assessment framework rather than threshold based predictions.

An iterative modeling framework may improve the ability of models to predict future distributions of species that are expanding their range. In this framework, models are developed based on available data, then model predictions are used to inform future monitoring (Aitken *et al.*, 2007). Additional data collected during monitoring is then used to update models. This iterative process has been recommended for both recovering species (Cianfrani *et al.*, 2010) and for invasive species (Jones *et al.*, 2010), but has not been widely evaluated (but see Bromberg *et al.*, 2011).

Quantitatively evaluating model predictions typically requires setting a threshold for suitability (Liu *et al.*, 2005). This threshold allows continuous suitability scores (which are the typical model output) to be converted to predictions of presence or absence, which are simpler to interpret. Locations with predicted suitability scores above the threshold are predicted presences, while those with lower scores are predicted absences. However, there are many methods for determining thresholds, and the specific suitability threshold selected may greatly affect the model's ability to accurately predict future invasions (Liu *et al.*, 2005). If thresholds are set to maximize current accuracy (a recommended method) they may poorly reflect future distributions. Threshold independent measures of current accuracy can also be used to assess models (Fielding and Bell, 1997), but only allow probabilistic predictions of future distributions.

An alternative approach is to use models to produce risk maps rather than specific predictions of potential distributions (e.g., Venette *et al.*, 2010). Areas with higher habitat suitability scores indicate higher risk. This approach involves using multiple thresholds to define risk categories – usually based on set proportions of the total area – rather than a single threshold to distinguish between predicted presence and absence. This approach reframes the question from where a species is predicted to occur in the future, to identifying the parts of the area of interest that are at greatest risk of invasion. Intensity of monitoring could then be based on the risk level in each area. This provides a more flexible framework for understanding future species distributions. If future occurrences are more common in areas with higher suitability scores, this approach may be successful even when the strict threshold approach is not.

In this paper I used historical records of two invasive plant species in the Bolleswood Natural Area, Connecticut, USA, to produce SDMs predicting potential distributions of the species. I then used more recent data to determine how well the models predicted new occurrences over time. I asked the following questions:

1. Does model accuracy decline with time?
2. Does model accuracy measured when the model is created correlate with future predictive accuracy?
3. Can future predictive accuracy be improved by iterative modeling or by using different methods to determine the suitability threshold?
4. Can a risk assessment framework better depict future distributions than a single threshold method?

The goal of this study is to examine the behavior of models over time with real rather than simulated species. Because this study was conducted on a small scale due to the constraints of the historical dataset, results should not be used to infer the potential distributions of the target species in other areas. Nevertheless, this study can provide an example of how SDMs for invasive species may behave over time and of the challenges that can be faced in trying to predict the future distributions of non-equilibrium species.

# 2. Methods

## *2.1 Study site*

The Bolleswood Natural Area is a 65 ha section of the Connecticut College Arboretum in New London and Waterford, Connecticut, USA. The western portion was cultivated until 1951 while the rest of the natural area was too rocky for cultivation (Fig. 1; Goslee *et al.*, 2005). The natural area was established in 1952 and has since been protected from cutting and fires. Most of the natural area is currently forested, with stands dominated by oak (*Quercus* spp.), oak-hemlock (*Quercus-Tsuga*) and hemlock-hardwood stands (Hemond *et al.*, 1983). The terrain varies from rocky ridges and ravines to flat former agricultural fields. Elevation ranges from 34 to 70 m. Temperatures average -2.3 °C in January and 21.8 °C in July. Annual precipitation averages 123.1 cm (www.worldclim.org; Hijmans *et al.*, 2005)

## *2.2 Study species*

I selected two common plants that invade forests for modeling; *Celastrus orbiculatus* Thunb. and *Rosa multiflora* Thunb. *Celastrus* is a woody vine, native to eastern Asia, first introduced to the United States as an ornamental around 1860 (Leicht-Young *et al.,* 2007). It is an aggressive invader in fields and open forests, particularly in disturbed habitats (Silveri *et al.*, 2001) and has spread throughout eastern North America (USDA-NRCS, 2010). *Celastrus* can damage trees and can form dense thickets, negatively impacting native vegetation (McNab and Meeker, 1987).

*Rosa* is a deciduous shrub, also native to eastern Asia that was intentionally introduced into the United States in the late 1800s (Banasiak and Meiners, 2009; Mosher *et al.*, 2009). It has spread across most of North America (USDA-NRCS, 2010). *Rosa* occurs in both open and forested habitats, but is limited by light availability in dense forests (Banasiak and Meiners, 2009). *Rosa* forms dense thickets and has been associated with decreased species richness (Meiners *et al.*, 2001).

## *2.3 Species data and habitat variables*

Eight hundred ninety 3 × 3 m plots were established in the natural area in 1952. These plots were contiguously arranged in four 6 m wide transects of varying length spaced 122 m apart (Fig. 1; Niering and Goodwin, 1962). All plant species have been recorded in these plots every 10 years since 1952. *Celastrus* first occurred in these plots and *Rosa* first became common in 1982 (Fig. 2), so I used presence/absence data for these two species in 1982 to construct models. I excluded plots that were entirely open water or open bog, leaving 857 plots.

I used nine environmental variables that reflected topography, vegetation structure, moisture availability, and spatial patterns (Table 1). The plots were separated into four forest age classes based on categories from Niering & Goodwin (1962). By 1982 (the first year of this study) all areas were forested, but these four classes represented differences in time since reforestation and land use history. Additionally, I used percent canopy cover (measured for each plot in 1992 because 1982 data were not available at the plot level). Each plot was assigned a drainage class based on Goslee *et al*. (2005). Elevation, slope and the cosine of aspect were calculated from a DEM using ArcMap 9.2

(ESRI, Redlands, California, USA). Finally, I measured the distance from the western edge of the natural area to each plot using ArcMap 9.2. Records indicate most of the invasion of the natural area has come from the western edge (Fike and Niering, 1999). I tested for multicollinearity among habitat variables, but I found no strong correlations ($|r|<0.57$) so all variables were included in model development.

One challenge with SDMs is finding environmental data from the same time period as the species data. This can be especially problematic when considering model predictions over time, as in this study. The topographic variables (drainage class, elevation, aspect and slope) are unlikely to have changed much over the timescale of this study. While the actual forest age changed over the course of the study, differences remained among the four classes. However, if species colonization is based on absolute forest age, changes in distributions over the course of the study could be due to changes in age. Likewise, canopy cover likely changed over the course of the study and this could influence results. However, nearly all plots had developed a closed canopy prior to 1982, so changes in canopy cover over the 20 year time period were modest. Finally, while proximity to the edge of the natural area did not vary over time, its importance may change as the species spread into the natural area.

### *2.4 Model development and assessment*

Model development and assessment involves two steps. First a modeling algorithm is used to calculate suitability scores for each plot. Then threshold values can be determined to separate plots into predicted presences or absences (or risk categories) based on their suitability scores.

I used Maxent to calculate suitability scores for each plot. Maxent is a machine learning modeling method that creates a probability distribution across the study area that maximizes entropy (i.e., is as close to uniform as possible) given the constraints of the input data (Phillips *et al.*, 2004; Phillips *et al.*, 2006). Maxent only requires presence points and selects absences randomly from background point. In an extensive review of different modeling methods for native species, Maxent was among the best performing methods (Elith *et al.*, 2006). I used Maxent 3.3.0 (http://www.cs.princeton.edu/~schapire/maxent/) with default options (regularization =1, maximum iterations =500). Maxent estimates the relative contributions of each variable based on the increase in regularized gain in each iteration, thus providing the percent contribution for each variable. I also created models using logistic regression (LR) and Genetic Algorithm for Rule-set Predictions (GARP), but patterns were very similar to those for Maxent, so they are not reported here.

Methods of model assessment are modified from Jones *et al.* (2010). I used 5-fold cross validation to assess models by randomly partitioning 1982 presence and absence points for each species into five equal subsets.  I ran five replicate models of each modeling technique for each species; each replicate used 80% of the data for model construction with a different set of 20% reserved to test model accuracy.

I used three measures to assess model accuracy. First, I used Sensitivity, or the proportion of true positives accurately predicted by a model. For invasive species, it is generally more important to accurately predict where a species will occur than where it will not. Thus, Sensitivity is an important metric of a model’s ability to predict future occurrences of an invasive species, especially when combined with more complete measures of model accuracy.

Second, I used the true skill statistic (TSS), a threshold-dependent measure of model accuracy which is similar to the commonly used Kappa statistic (Allouche *et al.*, 2006). TSS incorporates both the rate of false positive and false negative predictions, but unlike Kappa is not sensitive to frequency of presence points (Allouche *et al.*, 2006). TSS is calculated as *S*ensitivity + *S*pecificity – 1, where Specificity is the proportion of true negatives accurately predicted by the model. As with Kappa, values of TSS >0.6 are considered good, 0.2-0.6 fair to moderate, and <0.2 poor (Landis and Koch, 1977).

Both Sensitivity and TSS need predictions of presence/absence at each point in the landscape to compare with data not used in model creation (Fielding and Bell, 1997). This required selection of a threshold value above which the model is considered to predict presence (Fielding and Bell, 1997). I tested four different methods to select threshold values. The first three are methods recommended by Liu *et al.* (2005). First, I selected the threshold that maximized TSS using the data from 1982 (equivalent to the sensitivity-specificity sum maximization of Liu *et al.* [2005]). I calculated TSS for each model replicate across the full range of possible threshold values (0-100) and selected the threshold that maximized TSS. This method provides an optimal threshold for evaluating the current accuracy of each model replicate (Robertson *et al.*, 2004) but may not be optimal for future accuracy. Second, I calculated the threshold as the proportion of plots in the original 1982 data where the species occurred (prevalence). Third, I used the average suitability scores of all plots as calculated by the models. In addition to these three recommended methods I also set a fourth threshold as the minimum suitability score in a plot where the species occurred in 1982 (thus maximizing Sensitivity in 1982).

This last threshold measure is likely to produce many false positives in measuring current accuracy but may maximize the ability of models to predict future occurrences. Because this threshold measure is the least likely to decline in accuracy over time I used this threshold as a conservative measure of the effects of time on model accuracy. Each of these thresholds was applied to suitability scores calculated using Maxent with data from 1982. Model TSS and Sensitivity were calculated in each year for each model using thresholds generated by each of the four methods.

Finally, I assessed accuracy using the area under the curve (AUC) of Receiver Operating Characteristic (ROC) plots, which do not require selection of a habitat-suitability threshold (Fielding and Bell, 1997). AUC measures the accuracy of a model over the entire range of threshold values. Values of AUC generally range from 0.5 (equivalent to that due to chance) to 1.0 (perfect performance). Values >0.9 are considered good, 0.7-0.9 moderate, and <0.7 poor (Pearce and Ferrier, 2000).

To assess model predictions of current accuracy, I used the 1982 data for the 20% of the plots not included in model development for each replicate. To measure accuracy in predicting future distributions I used all previously non-invaded plots in 1992 and 2002. Thus in 1992 I used all plots where the species had never been recorded prior to 1992 and in 2002 I used all plots not invaded prior to 2002. This includes 20 new occurrences in 1992 and 36 in 2002 for *Celastrus* and 36 and 15 respectively for *Rosa*. I calculated the AUC, Sensitivity, and TSS in 1992 and 2002 based on Maxent models created from the 1982 data.

I compared model accuracy among years and species using two-way ANOVA with year and species as the factors (SPSS, 2006). ANOVA tests with significant ($p \leq 0.05$)

main effects of year were followed by Tukey’s HSD *post-hoc* tests of means.

To determine if predictive accuracy declined with time (question 1), I examined the main effect of year in the models. When there were significant interactions between year and species, I used one-way ANOVA tests to compare accuracy among years for each species. Significant ANOVA tests ($p \leq 0.05$) were followed by *post-hoc* tests as for the two-way ANOVA (or Tamhane’s T2 tests if variances were unequal). To evaluate if models with higher accuracy in 1982 also had higher accuracy in subsequent years (question 2), I looked for significant year by species interactions and visually evaluated results. If models with initially high accuracy declined more than models with lower initial accuracy, and this difference led to a significant interaction, this indicated poor correspondence between current and future accuracy.

## *2.5 Effects of threshold measures and iterative modeling on model accuracy*

I tested two approaches to improve predictive accuracy. First, I compared the four methods for determining thresholds to determine if certain methods provide better accuracy over time. Each of the four threshold values were applied to Maxent models using data from 1982. For each year and species I used one-way ANOVA followed by Tukey’s HSD *post-hoc* tests of means to compare model accuracy (TSS and Sensitivity) using the four methods to determine thresholds.

Second, I used an iterative modeling approach by using data from 1992 to predict occurrences in 2002. I contructed Maxent models using data from 1992, with five replicate models per species. I then calculated the AUC, TSS, and Sensitivity in 2002 for each replicate using the minimum suitability of occurrences threshold. I compared accuracy in 2002 between the Maxent models using 1982 data and those using 1992 data using t-tests.

### *2.6 Prioritizing monitoring based on relative risk*

I developed an alternate method of assessing risk of future invasion based on habitat suitability scores. Using the suitability scores from the Maxent models using 1982 data, I separated plots into risk categories as follows: Very High = the 10% of plots with the highest suitability scores, High = the next 20% of plots, Moderate = the next 20% and Low = the 50% of plots with the lowest habitat suitability scores. These categories are arbitrary but could allow managers to focus on smaller areas with higher risk. This was done separately for each model replicate. For each species and year, I compared the observed number of occurrences in each risk category with the expected number assuming new occurrences were spread randomly across risk categories. For example, of the 36 new occurrences for *Celastrus* in 2002, 3.6 would be expected in the Very High category, 7.2 in each of the High and Moderate categories and 18 in the Low category. I used one-sample t-tests (SPSS, 2006) for each species and year to compare the observed occurrences in each risk category for the five replicate models with the expected number of occurrences in that same risk category. This risk assessment approach will be useful if occurrences are more common in areas of higher risk than expected by chance. For visual comparison among years and species, numbers of occurrences in each year were converted to percentages observed in each risk category.

## **3. Results**

The majority of the occurrences of both species in 1982 were on the western end of the transects (Fig. 3). *Celastrus* spread eastward over time, while *Rosa* filled in gaps in its distribution. Model results for both species showed high predicted suitability on the western edge of the study area and low suitability to the east (Fig. 3, Appendix A).

Presence of young forest (Age 1) was the most important variable for both species with a relative contribution of 70.9% for *Celastrus* and 59.3% for *Rosa* (Appendix B). Drainage class (17.7%) and distance from the western edge of the natural area (7.9%) also contributed for *Celastrus*, while distance from the western edge (30.1%) and canopy cover (4.2%) contributed for *Rosa*.

All measures of model accuracy (using Maxent models with 1982 data and minimum occurrence threshold) showed a significant decline over time (main effect: Year $p$<0.001; Table 2). Accuracy in 2002 was significantly lower than in 1982 in almost all cases (Fig. 4). AUC declined for both species over time, while TSS declined dramatically for *Celastrus* but not for *Rosa*. The ability of models to predict new occurrences (Sensitivity) in 2002 was poor for *Celastrus* (49%) and moderate for *Rosa* (85%) even using the most conservative threshold measure. Patterns were similar regardless of the threshold measure used.

AUC and Sensitivity were higher overall for *Rosa* than for *Celastrus*, but there was a significant Year by Species interaction for all three accuracy measures ($p$<0.001; Table 2). All three measures were higher for *Celastrus* than *Rosa* in 1982, but were much lower in 2002 (Fig. 4). Accuracy for *Celastrus* began to decline in 1992; while for *Rosa* it increased or stayed the same. Full one-factor ANOVA results for Year are found in Appendix C.

### *3.1 Effects of threshold measures and iterative modeling*

Thresholds values varied among selection methods (Appendix D). Threshold values that maximized TSS in 1982 were more variable among replicates and often much higher than with other selection methods. The relative accuracy of threshold selection methods depended on species and year (Fig. 5), but all measures showed the same general patterns of decline over time and stronger decline for *Celastrus*. There was generally little difference in accuracy among selection methods in 1982. By 2002, the method that maximized TSS in 1982 generally performed more poorly than the other methods because it maximizes current rather than future accuracy. Full ANOVA results for threshold comparisons are in Appendix E.

Using updated information from 1992 to create Maxent models did not improve model accuracy in 2002 over models using 1982 data only (Fig. 6). Model accuracy in 2002 was never significantly higher for models using 1992 data than for models using 1982 data, and AUC for *Celastrus* was actually lower.

### *3.2 Prioritizing monitoring based on relative risk*

After controlling for different numbers of plots, significantly more occurrences than expected by chance were found in areas of Very High risk for both species in all years while fewer than expected were found in areas of Low risk, even in 2002 (Fig. 7). For *Rosa*, the distribution of new occurrences was fairly consistent among years with 60-73% in areas of Very High risk (only 10% of the land area) and 24-26% in areas of High risk. For *Celastrus,* >78% of occurrences in 1982 and 1992 were in areas of Very High risk. In

2002 this dropped to 24%, but even then, 75% of new occurrences were in either the Very High or High risk category (30% of the land area). These patterns indicate that these categories are effectively describing both current and future risk of invasion.

## 4. Discussion

For both species, predictive accuracy declined after 20 years, despite high initial accuracy. This suggests that, at least in this case, models are not accurately predicting the long-term potential distribution, but instead model the current and near-term distributions. That accuracy declines over time is not surprising, and fits with evidence from a number of recent studies (Elith *et al.*, 2010; Dupin *et al.*, 2011; Václavík and Meentemeyer, 2012).

This decline in accuracy over time could have been caused by changes in habitat variables over time that were not incorporated into the models. For example, using canopy cover data from 1982 (had it been measured) and updating model results based on changing canopy cover in each time period may have improved model accuracy. However, because canopy cover contributed very little to model results (Appendix B), the effect was likely modest. Although forest age was important, *Celastrus* spread from the youngest forest to older forest classes over time, even though all forests were increasing in age. It is important to consider how habitat variables changing over time impacts model results, but this did not appear to be a major problem in this study.

There is some evidence that *Celastrus* is gradually spreading away from the sources of colonization, which might negatively impact model accuracy over time, since proximity to the edge of the natural area was included in the model. However, this variable contributed only 8% to the model results for *Celastrus*. Nevertheless, the use of variables representing dispersal patterns needs to be carefully considered if accuracy over time is an important goal, as including these variables will tend to focus results on current and near term distributions rather than potential equilibrium distributions.

A major reason that the models failed to accurately predict future distributions is that the species (particularly *Celastrus*) shifted the habitat it invaded from primarily young open forest in 1982 to much older forest in 2002, with the shift mostly occurring after 1992. A shift in habitat such as this one can occur for two major reasons. First, it may be the result of rapid evolution in the invasive species in its new range (Broennimann *et al.*, 2007), but this is unlikely in this case because of the scale of the study. Second, a distribution shift may also occur during an invasion if the initial invasion only includes a subset of suitable habitat or if initial distributions include sink habitat (Anderson *et al.*, 2006). In this case, the shift is the result of non-equilibrium conditions. The small size of the study area and single initial invasion focus may have accentuated this shift in this study. However, even over larger spatial scales, invasion patterns can lead to similar shifts in habitat (Crossman and Bass, 2008; Jones *et al.*, 2010).

More critical than the general decline in accuracy, however, is that models that more accurately predicted the current distribution in 1982 (*Celastrus*), performed much worse in predicting the future distribution. In retrospect, this is likely because *Rosa* was closer to equilibrium with the environment than *Celastrus*, which allows for better predictions of future distributions (Václavík and Meentemeyer, 2012). However, this difference would not have been apparent in 1982 as both appeared to be in initial stages of spreading. Thus, current model accuracy measures do not necessarily indicate the ability of models to predict future distributions and must be interpreted with care.

The method used to select the threshold did impact accuracy somewhat, but the overall pattern of decline in accuracy over time was similar for all threshold measures. Using the threshold that maximizes current accuracy, however, generally performed the worst over time. This method may be useful if the goal is maximize current accuracy, but not if the goal is long term predictions.

Iterative modeling only slightly increased predictive ability in 2002 for *Rosa* and not at all for *Celastrus*. Iterative modeling should reduce the impact of distribution shifts on predictive ability, but the ten year interval may have been too long. In 1992 *Celastrus* still had largely invaded the youngest forests, while in 2002 the majority of new occurrences were in older forest. More frequent sampling and model updates may have been able to better predict this shift. For iterative modeling to be effective, some sampling must occur in areas predicted to be less suitable by the original models.

Using the model results in a risk assessment framework without setting thresholds showed more promise.  Even after 20 years, very few occurrences were found in areas rated as Low risk. In general, new occurrences progressively spread from plots with higher to lower suitability scores. This suggests that a stratified monitoring plan based on risk categories could successfully detect most new occurrences. This type of approach would allow managers to prioritize areas for monitoring and control of invasive species, with the most intense monitoring occurring in areas of higher risk. This approach would have successfully detected almost all of the new occurrences in the current study.

Other approaches are being developed that may improve the ability of SDMs to predict future distributions. Adding data from the species' native range or from larger spatial scales in the invaded range shows potential (Broennimann and Guisan, 2008; Ibanez *et al.*, 2009; Jones *et al.*, 2010). However, scale and resolution issues may come into play. For example, a model for *Celastrus* using data from both the native and invaded range indicated that all of the study area in the current study was suitable at a 1 $km^2$ resolution (Ibanez *et al.*, 2009). While *Celastrus* can likely occur somewhere in each square kilometer in the region, this conclusion is not particularly helpful for management of the species in local natural areas such as the Bolleswood.

Another new approach that may help to overcome the lack of equilibrium issue is to use mechanistic or physiological information about the species to describe the species' niche and then use this information to create a niche-based model (Kearney and Porter, 2009). These mechanistic models do not require the assumption of equilibrium, but they do require considerable knowledge about the physiology and ecology of the species. Nevertheless this approach may provide complimentary predictions that will help to evaluate the predictions of SDMs (Elith *et al.*, 2010)

## 5. Conclusions

This study demonstrates that it is critical to clearly differentiate between predicting the actual current distribution and the potential future distribution of an invasive species. These two types of predictions are often confused in SDM studies of invasive species (see Václavík and Meentemeyer, 2009 for discussion). High model accuracy using data from the time of model creation does not ensure that the model will accurately predict the future distribution. The ability of models to predict new occurrences will decline with time and without data from the future it is impossible to directly assess a model's future accuracy. The small scale and spatial autocorrelation in the current study may exacerbate this problem, and in some circumstances models can accurately predict both current and future distributions (e.g., Jarnevich *et al.*, 2010). However, spatial autocorrelation and invasion patterns can influence even large-scale models (Crossman and Bass, 2008; Jones *et al.*, 2010) leading to similar patterns as found in this study. The key message is that high current model accuracy does not always indicate high future accuracy, and these should be explicitly distinguished.

Nevertheless, there are ways to better use SDMs to help control invasive species. Model development should be more closely and explicitly tied to monitoring. Initial Models can be used to map the risk of potential invasion rather than to predict future presence or absence. These maps can then be used to design monitoring programs with intensity of sampling based on invasion risk. Additional information about the species native range and/or physiology can also be included. Results from monitoring can then be used to iteratively model the potential distribution. The combination of frequent iterative modeling in a risk assessment framework with the use of additional mechanistic or native range data is most likely to successfully detect future spread of invasive species.

## Acknowledgements

I thank William Niering and Richard Goodwin who initially set up the permanent plots used in this study. Glenn Dreyer, Pamela Hine, Christine Small and many students contributed to sampling over the years.

## References

Aitken, M., Roberts, D.W., Shultz, L.M., 2007. Modeling distributions of rare plants in the Great Basin, western North America. West. N. Am. Nat. 67, 26-38.

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223-1232.

Anderson, R.P., Peterson, A.T., Egbert, S.L., 2006. Vegetation-index models predict areas vulnerable to purple loosestrife (*Lythrum salicaria*) invasion in Kansas. Southwest. Nat. 51, 471-480.

Banasiak, S.E., Meiners, S.J., 2009. Long term dynamics of *Rosa multiflora* in a successional system. Biol. Invasions 11, 215-224.

Broennimann, O., Guisan, A., 2008. Predicting current and future biological invasions: Both native and invaded ranges matter. Biol. Lett. 4, 585-589.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T., Guisan, A., 2007. Evidence of climatic niche shift during biological invasion. Ecol. Lett. 10, 701-709.

Bromberg, J.E., Kumar, S., Brown, C.S., Stohlgren, T.J., 2011. Distributional changes and range predictions of downy brome (*Bromus tectorum*) in Rocky Mountain National Park. Invasive Plant Sci. Manag. 4, 173-182.

Cianfrani, C., Le Lay, G., Hirzel, A.H., Loy, A., 2010. Do habitat suitability models reliably predict the recovery areas of threatened species? J. Appl. Ecol. 47, 421-430.

Crossman, N.D., Bass, D.A., 2008. Application of common predictive habitat techniques for post-border weed risk management. Divers. Distrib. 14, 213-224.

Dupin, M., Reynaud, P., Jarosik, V., Baker, R., Brunel, S., Eyre, D., Pergl, J., Makowski, D., 2011. Effects of the training dataset characteristics on the performance of nine species distribution models: application to *Diabrotica virgifera virgifera*. PloS One 6, e20957.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129-151.

Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species. Methods Ecol. Evol. 1, 330-342.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38-49.

Fike, J., Niering, W.A., 1999. Four decades of old field vegetation development and the role of Celastrus orbiculatus in the northeastern United States. J. Veg. Sci. 10, 483-492.

Goslee, S.C., Niering, W.A., Urban, D.L., Christensen, N.L., 2005. Influence of environment, history and vegetative interactions on stand dynamics in a Connecticut forest. J. Torrey Bot. Soc. 132, 471-482.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135, 147-186.

Hemond, H.F., Niering, W.A., Goodwin, R.H., 1983. Two decades of vegetation change in the Connecticut Arboretum Natural Area. Bull. Torrey Bot. Club 110, 184-194.

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. Int. J. Clim. 25, 1965-1978.

Ibanez, I., Silander, J.A., Wilson, A.M., LaFleur, N., Tanaka, N., Tsuyama, I., 2009. Multivariate forecasts of potential distributions of invasive plant species. Ecol. Appl. 19, 359-375.

Jarnevich, C.S., Holcombe, T.R., Barnett, D.T., Stohlgren, T.J., Kartesz, J.T., 2010. Forecasting Weed Distributions using Climate Data: A GIS Early Warning Tool. Invasive Plant Sci. Manag. 3, 365-375.

Jones, C.C., Acker, S.A., Halpern, C.B., 2010. Combining local- and large-scale models to predict the distributions of invasive plant species. Ecol. Appl. 20, 311-326.

Kearney, M., Porter, W., 2009. Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. Ecol. Lett. 12, 334-350.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159-174.

Leicht-Young, S.A., Silander, J.A., Latimer, A.M., 2007. Comparative performance of invasive and native Celastrus species across environmental gradients. Oecologia 154, 273-282.

Liu, C.R., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28, 385-393.

Martin, P.H., Canham, C.D., Marks, P.L., 2009. Why forests appear resistant to exotic plant invasions: intentional introductions, stand dynamics, and the role of shade tolerance. Front. Ecol. Environ. 7, 142-149.

McNab, W.H., Meeker, M., 1987. Oriental bittersweet: a growing threat to hardwood silviculture in the Appalachians. North. J. Appl. For. 4, 174-177.

Meiners, S.J., Pickett, S.T.A., Cadenasso, M.L., 2001. Effects of plant invasions on the species richness of abandoned agricultural land. Ecography 24, 633-644.

Mosher, E.S., Silander, J.A., Latimer, A.M., 2009. The role of land-use history in major invasions by woody plant species in the northeastern north american landscape. Biol. Invasions 11, 2317-2328.

Niering, W.A., Goodwin, R.H., 1962. Ecological studies in the Connecticut Arboretum Natural Area I. Introduction and a survey of vegetation types. Ecology 43, 41-54.

Pearce, J., Ferrier, S., 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecol. Model. 133, 225-245.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231-259.

Phillips, S.J., Dudik, M., Schapire, R.E., 2004. A maximum entropy approach to species distribution modeling. in:  Proceedings of the Twenty-First International Conference on Machine Learning, Banff, Canada, pp. 655-662.

Rew, L.J., Maxwell, B.D., Aspinall, R., 2005. Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. Weed Sci. 53, 236-241.

Robertson, M.P., Villet, M.H., Palmer, A.R., 2004. A fuzzy classification technique for predicting species' distributions: applications using invasive alien plants and indigenous insects. Divers. Distrib. 10, 461-474.

Silveri, A., Dunwiddie, P.W., Michaels, H.J., 2001. Logging and edaphic factors in the invasion of an Asian woody vine in a mesic North American forest. Biol. Invasions 3, 379-389.

SPSS, 2006. SPSS 15.0 for Windows. SPSS Inc., Chicago, Illinois, USA

Sutherst, R.W., Bourne, A.S., 2009. Modelling non-equilibrium distributions of invasive species: A tale of two modelling paradigms. Biol. Invasions 11, 1231-1237.

Svenning, J.-C., Skov, F., 2004. Limited filling of the potential range in European tree species. Ecol. Lett. 7, 565-573.

Thuiller, W., Brotons, L., Araujo, M.B., Lavorel, S., 2004. Effects of restricting environmental range of data to project current and future species distributions. Ecography 27, 165-172.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O., Rouget, M., 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. Glob. Chang. Biol. 11, 2234-2250.

USDA-NRCS, 2010. The PLANTS Database (http://plants.usda.gov). Accessed 16 June 2010. National Plant Data Center, Baton Rouge, Louisiana, USA.

Václavík, T., Meentemeyer, R.K., 2009. Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecol. Model. 220, 3248-3258.

Václavík, T., Meentemeyer, R.K., 2012. Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Divers. Distrib. 18, 73-83.

Venette, R.C., Kriticos, D.J., Magarey, R.D., Koch, F.H., Baker, R.H.A., Worner, S.P., Raboteaux, N.N.G., McKenney, D.W., Dobesberger, E.J., Yemshanov, D., Barro, P.J.D., Hutchison, W.D., Fowler, G., Kalaris, T.M., Pedlar, J., 2010. Pest risk maps for invasive alien species: A roadmap for improvement. Bioscience 60, 349-362.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmanek, M., Westbrooks, R., 1997. Introduced species: A significant component of human-caused global change. N. Z. J. Ecol. 21, 1-16.

**Table 1.** Predictor variables used in models

| Variable | Range | Description |
|---|---|---|
| Age 1[a] | 0,1 | Areas dominated by old fields in 1952 |
| Age 2[a] | 0,1 | Areas dominated by shrubs in 1952 |
| Age 3[a] | 0,1 | Areas dominated by transitional (young) forest in 1952 |
| Percent canopy cover | 0 – 100% | Visual estimate of canopy cover in 1992 |
| Drainage class[b] | 1 – 6 | Assessment of soil drainage from saturated (1) to excessively drained (6) |
| Distance from west | 12 – 471 m | Distance from western edge of the natural area (the most likely source of invasion) |
| Elevation | 35 – 68 m | Elevation above sea level |
| Slope | 0 – 58° | Slope angle in degrees |
| Aspect | -1 – 1 | Cosine of aspect (1=north facing, -1=south facing) |

[a] from Niering and Goodwin (1962). All areas not included in Age 1-3 were mature forest in 1952.

[b] from Goslee et al. (2005)

Table 1. Results of two-way ANOVA tests for three measures of model accuracy.

| Source | Sum of Squares | df | Mean Square | F | P |
|---|---|---|---|---|---|
| **A. AUC** (adjusted $R^2$ = 0.925) | | | | | |
| Corrected Model | 0.08 | 5 | .02 | 72.04 | <0.001 |
| Intercept | 25.18 | 1 | 25.18 | 117723.10 | <0.001 |
| Year | 0.06 | 2 | 0.03 | 143.82 | <0.001 |
| Species | 0.00 | 1 | 0.00 | 10.70 | 0.003 |
| Year * Species | 0.01 | 2 | 0.01 | 30.92 | <0.001 |
| Error | 0.01 | 24 | 0.00 | | |
| Total | 25.26 | 30 | | | |
| Corrected Total | 0.08 | 29 | | | |
| | | | | | |
| **B. Sensitivity** (adjusted $R^2$ = 0.900) | | | | | |
| Corrected Model | 0.81 | 5 | 0.16 | 53.04 | <0.001 |
| Intercept | 21.28 | 1 | 21.28 | 6971.25 | <0.001 |
| Year | 0.43 | 2 | 0.21 | 70.19 | <0.001 |
| Species | 0.13 | 1 | 0.13 | 43.95 | <0.001 |
| Year * Species | 0.25 | 2 | 0.12 | 40.43 | <0.001 |
| Error | 0.07 | 24 | 0.00 | | |
| Total | 22.17 | 30 | | | |
| Corrected Total | 0.88 | 29 | | | |
| | | | | | |
| **C. TSS** (adjusted $R^2$ = 0.836) | | | | | |
| Corrected Model | 0.61 | 5 | 0.121 | 30.58 | <0.001 |
| Intercept | 13.85 | 1 | 13.85 | 3497.06 | <0.001 |
| Year | 0.32 | 2 | 0.16 | 39.94 | <0.001 |
| Species | 0.00 | 1 | 0.00 | 0.10 | 0.751 |

|  |  |  |  |  |  |
|---|---|---|---|---|---|
| Year * Species | 0.29 | 2 | 0.14 | 36.47 | <0.001 |
| Error | 0.10 | 24 | 0.01 |  |  |
| Total | 14.55 | 30 |  |  |  |
| Corrected Total | 0.70 | 29 |  |  |  |

**Fig. 1.** Location of the sample plots in Bolleswood Natural Area. Triangle in inset shows the location of the study site in southeastern Connecticut, USA. Solid white lines indicate 6 m wide transects of contiguous 3 X 3 m plots. Dashed line indicates the boundary of the natural area. Aerial photo shows forest cover in 1934 to illustrate variation in forest age.

**Fig. 2.** Frequency of *Celastrus orbiculatus* and *Rosa multiflora* in the sample plots at each sample date

**Fig 3.** Species distribution over time and model results for *Celastrus orbiculatus* and *Rosa multiflora*. For species distributions (A, C), black squares indicate species occurrence in 1982, medium gray in 1992 and light gray in 2002. Model results (B,D) are from a single replicate (results from all replicates are found in Appendix S1). Darker shading indicates greater predicted suitability.

**Fig 4.** Model accuracy (AUC, Sensitivity, and TSS) over time by species. Within each species different letters indicate significant differences in accuracy among years based on Tukey's HSD *post-hoc* tests following significant ANOVA tests.

**Fig. 5**. Effect of threshold selection method on Sensitivity and TSS. Within each species, different letters indicate significant differences in accuracy among thresholds based on Tukey's HSD *post-hoc* tests following significant ANOVA tests.

**Fig. 6.** Model accuracy in 2002 comparing models created in 1982 and 1992. Within each species, stars indicate significant differences between the 1982 and 1992 models based on t-tests. * 0.05>p>0.01, **0.01>p>0.001, ***p<0.001. Patterns are similar regardless of threshold measure used, results using the mean suitability threshold are shown here.

**Fig. 7.** Percentage of new occurrences in each risk category by year. Percentages in a given year add to 100%. Very high = 10% of plots with highest suitability scores, High = 10-30%, Moderate = 30-50%, Low = 50% of plots with lowest suitability scores. Gray line indicates expected percentage of occurrences in each risk category if new occurrences were distributed randomly among categories. * = significant differences between observed and expected percentages based on one-sample t-tests (p<0.001). + = no t-statistic calculated because no variation among model replicates.

Figure 1

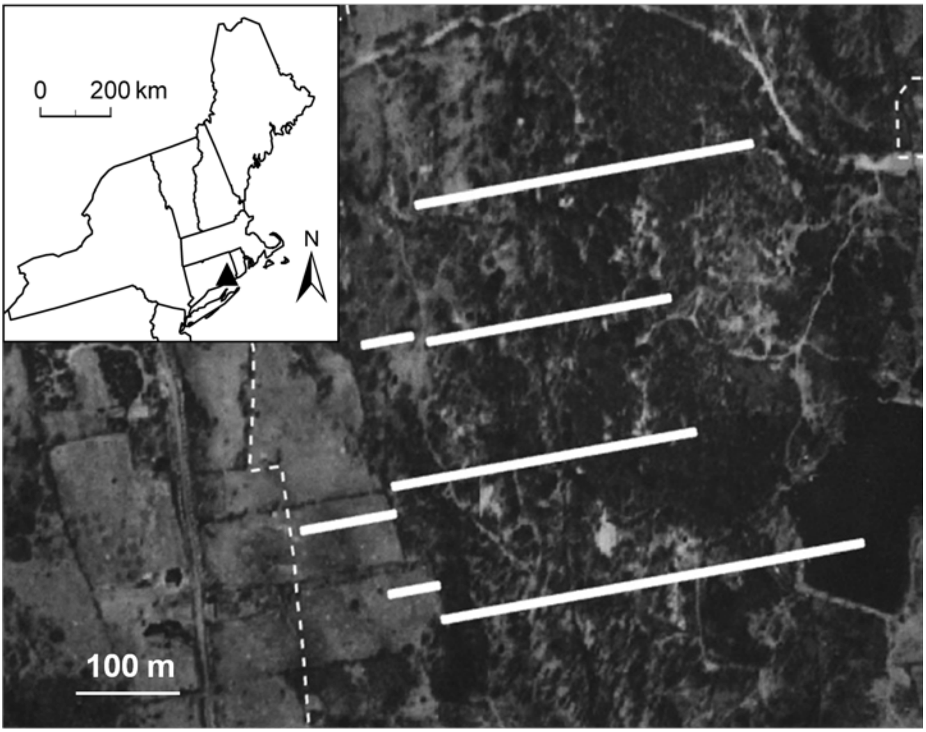
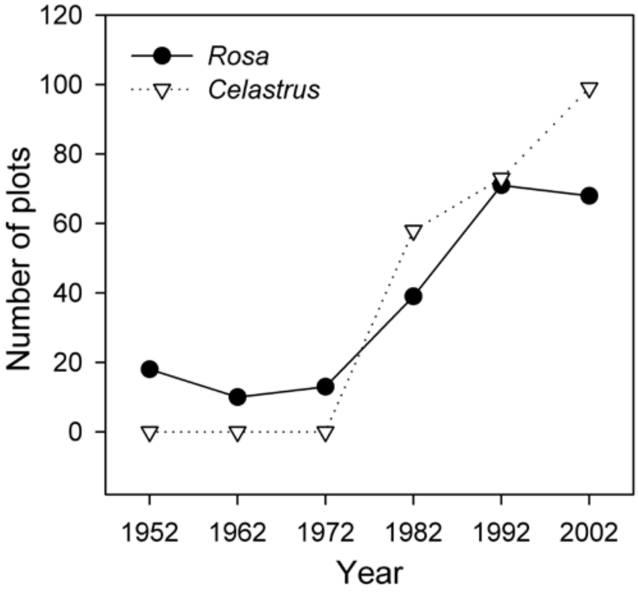


Figure 2



Rosa
Celastrus
Number of plots
120
100
80
60
40
20
0
1952
1962
1972
1982
1992
2002
Year

Figure 3

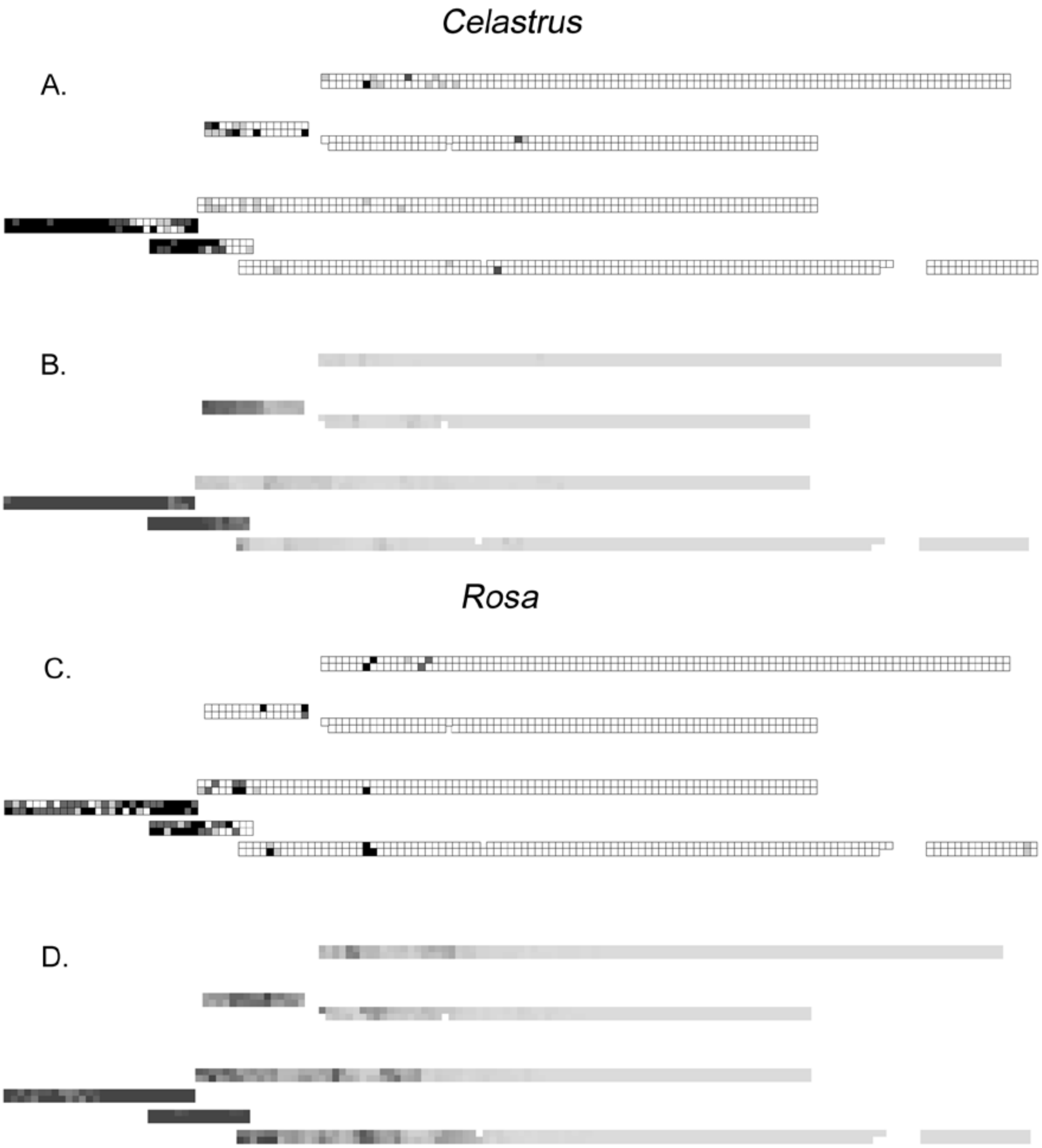


Figure 4

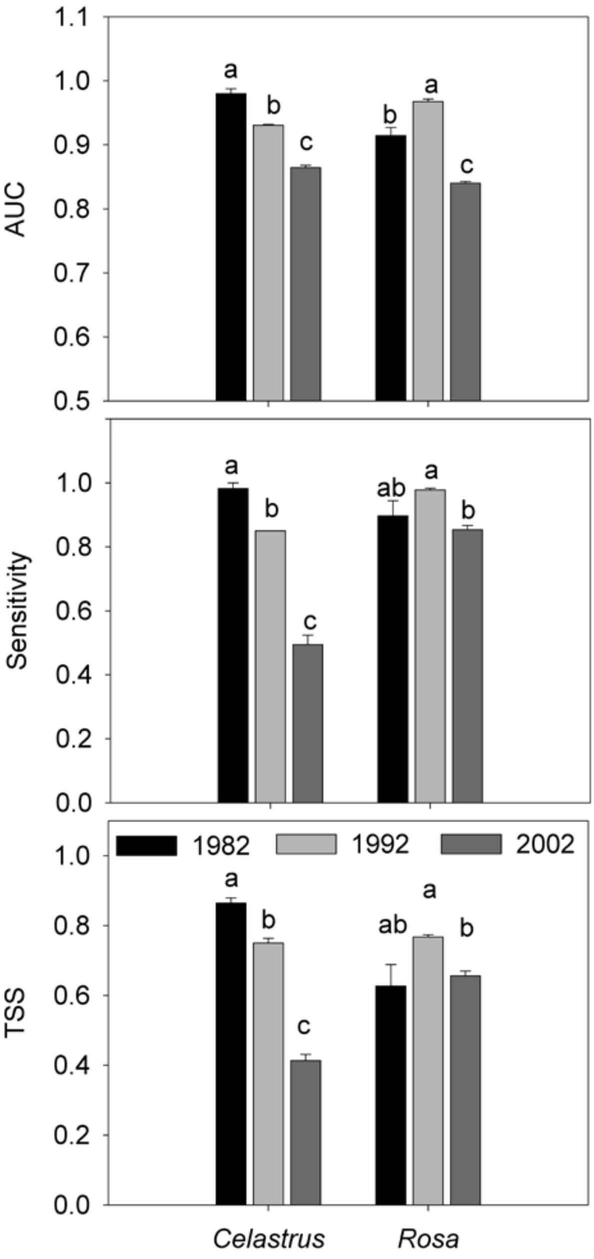

AUC
1.1
1.0
0.9
0.8
0.7
0.6
0.5
a
b
c
b
a
c
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
a
b
c
ab
a
b
TSS
1.0
0.8
0.6
0.4
0.2
0.0
1982
1992
2002
a
b
c
ab
a
b
Celastrus
Rosa

Figure 5

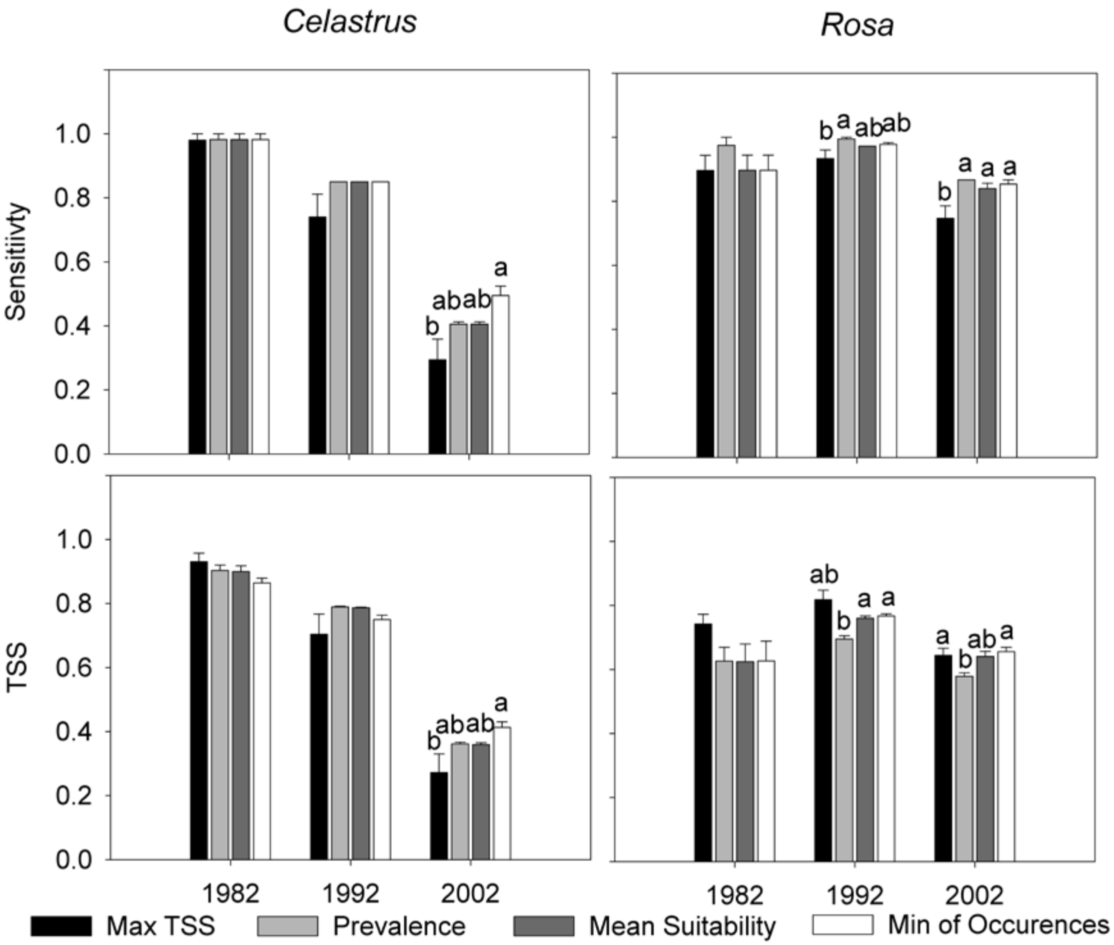

Celastrus
Rosa
Sensitiivty
TSS
0.0
0.2
0.4
0.6
0.8
1.0
b
abab
a
b a abab
b a a a
ab
b a a
a b ab a
1982
1992
2002
Max TSS
Prevalence
Mean Suitability
Min of Occurences

Figure 6

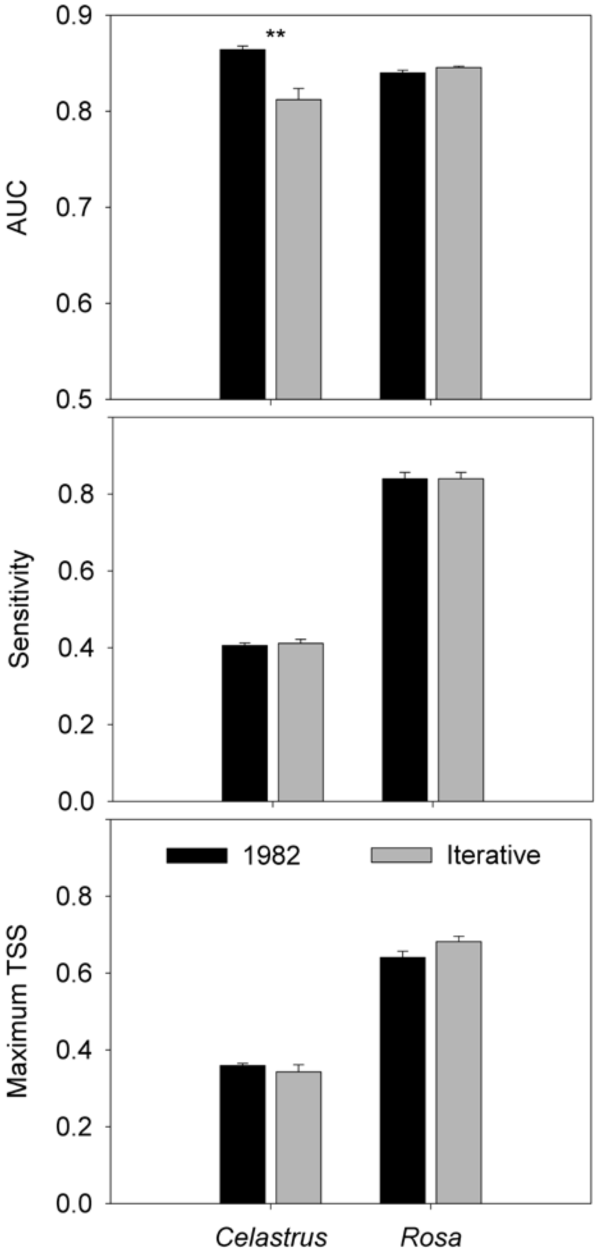
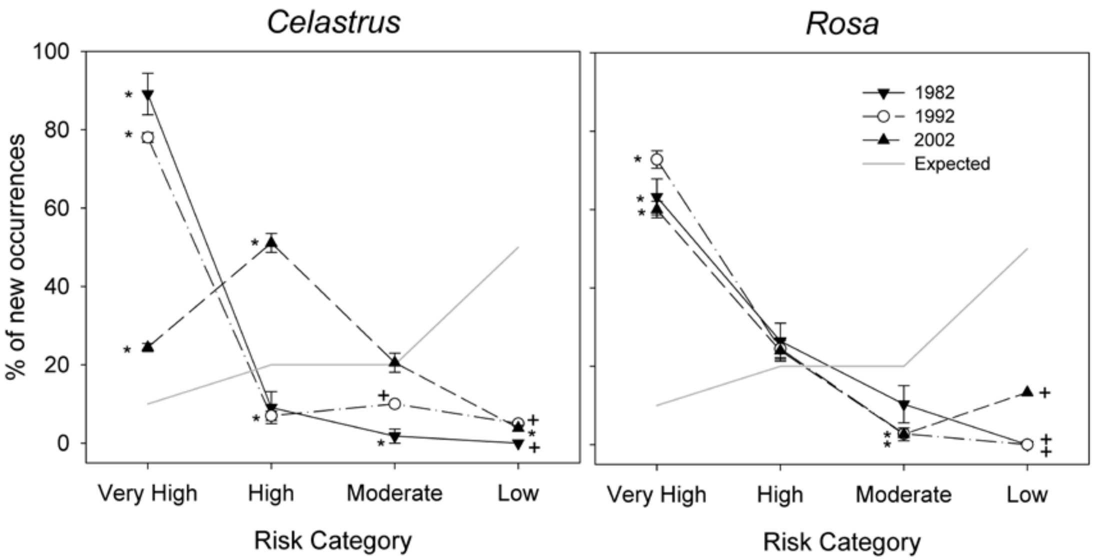

0.9
0.8
0.7
0.6
0.5
AUC
**
0.8
0.6
0.4
0.2
0.0
Sensitivity
1982
Iterative
0.8
0.6
0.4
0.2
0.0
Maximum TSS
*Celastrus*
*Rosa*

Figure 7


Celastrus
Rosa
% of new occurrences
100
80
60
40
20
0
Very High
High
Moderate
Low
Risk Category
1982
1992
2002
Expected
*
+

ty and red indicates high predicted suitability.

## *Celastrus* Models

plicate 1

plicate 2

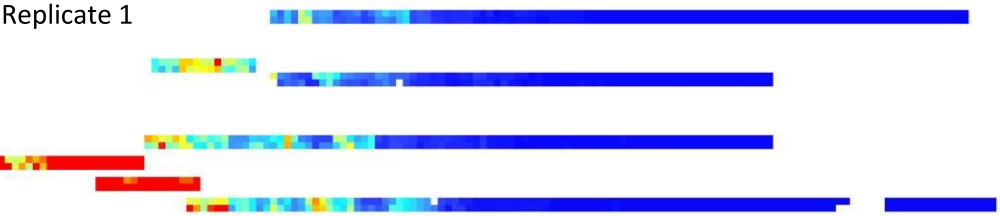
plicate 3

plicate 4

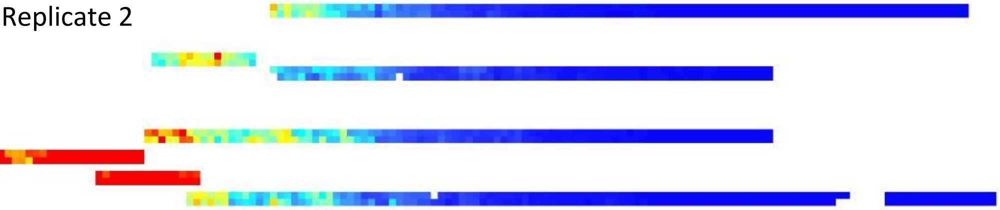
plicate 5

## 2. *Rosa* Models

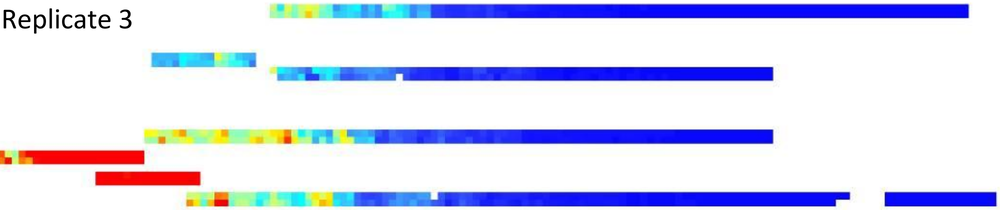
Replicate 1



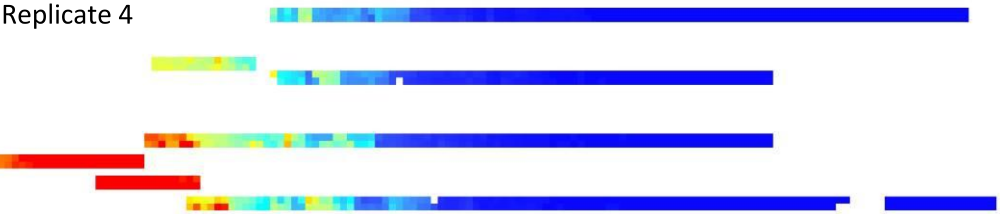
Replicate 2



Replicate 3



Replicate 4



Replicate 5

tions of each variable based on the increase in regularized gain in each iteration, thus providing ent contribution for each variable. Values indicate the average percent contribution among es (SE in parentheses).

| | *Celastrus* | SE | *Rosa* | SE |
|---|---|---|---|---|
| Age1 | 70.88 | (3.27) | 59.32 | (2.17) |
| Age2 | 1.48 | (0.17) | 0.34 | (0.15) |
| Age3 | 0.06 | (0.02) | 1.84 | (0.27) |
| Canopy | 0.86 | (0.31) | 4.22 | (0.66) |
| Aspect | 0.1 | (0.05) | 0.5 | (0.18) |
| Distance from west | 7.94 | (0.75) | 30.1 | (2.92) |
| Drainage | 17.68 | (2.31) | 1.08 | (0.52) |
| Elevation | 0.42 | (0.18) | 0.3 | (0.23) |
| Slope | 0.52 | (0.31) | 2.3 | (0.52) |

ased on all previously uninvaded plots. For sensitivity and TSS the threshold was set at the lowest ty score for species presences in the training data.

| | *Celastrus orbiculatus* | | | | | *Rosa multiflora* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Sum of Squares | df | Mean Square | *F* | *p* | Sum of Squares | df | Mean Square | *F* | *p* |
| Groups | 0.034 | 2 | 0.017 | 128.55 | <0.001 | 0.041 | 2 | 0.021 | 69.21 | <0.001 |
| roups | 0.002 | 12 | 0.000 | | | 0.004 | 12 | 0.000 | | |
| | 0.035 | 14 | | | | 0.045 | 14 | | | |
| **y** | | | | | | | | | | |
| Groups | 0.636 | 2 | 0.318 | 157.54 | <0.001 | 0.040 | 2 | 0.020 | 4.88 | 0.028 |
| roups | 0.024 | 12 | 0.002 | | | 0.049 | 12 | 0.004 | | |
| | 0.660 | 14 | | | | 0.089 | 14 | | | |
| Groups | 0.550 | 2 | 0.275 | 232.61 | <0.001 | 0.055 | 2 | 0.028 | 4.085 | 0.044 |
| roups | 0.014 | 12 | 0.001 | | | 0.081 | 12 | 0.007 | | |
| | 0.564 | 14 | | | | 0.136 | 14 | | | |

thods to distinguish between plots predicted as presences and absences. Max TSS: the threshold ximized TSS using the test data from 1982. Prevalence: the proportion of plots in the original 1982 ere the species occurred. Mean suitability: the average suitability scores of all plots as calculated nodels. Lowest occurrence: the minimum suitability score in a plot where the species occurred in

| *Celastrus orbiculatus* | | | | *Rosa multiflora* | | | |
|---|---|---|---|---|---|---|---|
| Max TSS | Prevalence | Mean suitability | Lowest occurrence | Max TSS | Prevalence | Mean suitability | Lowest occurrence |
| 0.083 | 0.064 | 0.058 | 0.033 | 0.230 | 0.044 | 0.100 | 0.095 |
| 0.056 | 0.064 | 0.055 | 0.061 | 0.184 | 0.044 | 0.115 | 0.136 |
| 0.408 | 0.064 | 0.065 | 0.021 | 0.246 | 0.044 | 0.114 | 0.108 |
| 0.173 | 0.064 | 0.062 | 0.022 | 0.295 | 0.045 | 0.119 | 0.112 |
| 0.387 | 0.064 | 0.066 | 0.037 | 0.102 | 0.045 | 0.101 | 0.097 |
| *0.221* | *0.064* | *0.061* | *0.035* | *0.211* | *0.044* | *0.110* | *0.109* |

e methods for calculating thresholds. Methods included maximizing TSS, setting the threshold at alence value, setting the threshold at the mean suitability value and at the lowest suitability score ccurrence in 1982.

**.1**. One-factor ANOVA comparing Sensitivity among threshold methods

| | *Celastrus orbiculatus* | | | | | *Rosa multiflora* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Sum of Squares | df | Mean Square | *F* | *p* | Sum of Squares | df | Mean Square | *F* | *p* |
| Groups | 0.000 | 3 | 0.000 | 0.00 | 1.000 | 0.023 | 3 | 0.008 | 0.84 | 0.492 |
| roups | 0.028 | 16 | 0.002 | | | 0.147 | 16 | 0.009 | | |
| | 0.028 | 19 | | | | 0.170 | 19 | | | |
| Groups | 0.045 | 3 | 0.015 | 2.37 | 0.109 | 0.010 | 3 | 0.003 | 3.35 | 0.046 |
| roups | 0.102 | 16 | 0.006 | | | 0.016 | 16 | 0.001 | | |
| | 0.147 | 19 | | | | 0.026 | 19 | | | |
| Groups | 0.100 | 3 | 0.033 | 5.32 | 0.010 | 0.044 | 3 | 0.015 | 6.05 | 0.006 |
| roups | 0.101 | 16 | 0.006 | | | 0.039 | 16 | 0.002 | | |
| | 0.201 | 19 | | | | 0.083 | 19 | | | |

**.2:** One-factor ANOVA comparing TSS among threshold methods

| | *Celastrus orbiculatus* | | | | | *Rosa multiflora* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Sum of Squares | df | Mean Square | *F* | *p* | Sum of Squares | df | Mean Square | *F* | *p* |
| Groups | 0.011 | 3 | 0.004 | 1.93 | 0.166 | 0.051 | 3 | 0.017 | 1.40 | 0.278 |
| roups | 0.031 | 16 | 0.002 | | | 0.194 | 16 | 0.012 | | |
| | 0.042 | 19 | | | | 0.245 | 19 | | | |
| Groups | 0.024 | 3 | 0.008 | 1.52 | 0.247 | 0.038 | 3 | 0.013 | 9.99 | 0.001 |
| roups | 0.083 | 16 | 0.005 | | | 0.020 | 16 | 0.001 | | |
| | 0.107 | 19 | | | | 0.059 | 19 | | | |
| Groups | 0.051 | 3 | 0.017 | 3.67 | 0.035 | 0.018 | 3 | 0.006 | 4.58 | 0.017 |
| roups | 0.074 | 16 | 0.005 | | | 0.021 | 16 | 0.001 | | |
| | 0.125 | 19 | | | | 0.039 | 19 | | | |