**Intro**

~~I linked demographic and species distribution models for the invasive plant~~ *~~Frangula alnus~~* ~~throughout its North American range, and compared simulation results to observed patterns of spatial spread to determine the demographic properties and processes most likely responsible for this species’ success~~.

*I removed this from the intro (last paragraph) because I didn’t actually test these two hypotheses directly. Further, the starling hypothesis does not have support, because buckthorn spread west prior to starlings, and the frappier hypothesis is hard to test. Might be useful for the discussion section though.*

“While some authors have proposed that the extended lag phase for *F. alnus* is a result of the time required for European starling to spread through North America (Howell and Blackwell 1977) or the time required for adaptation (Frappier et al. 2003b), these hypotheses have not been examined in detail.”

*Here’s all the lag-phase specific stuff I’ve cut out …*

“One of the remaining gaps in our knowledge is a full understanding of the processes governing population dynamics during invasive species lag phases (Pyšek and Hulme 2005, Blackburn et al. 2011, Gurevitch et al. 2011). The lag phase is the period of time between the establishment of self-sustaining populations and rapid expansion in abundance and area of occupancy characteristic of invasive species (Kowarik 1995, Crooks and Soulé 1999, Sakai et al. 2001, Pyšek and Hulme 2005, Theoharides and Dukes 2007). This lag may simply be the result of the time required for a population to grow large enough that emigration leads to colonization of new populations. However, lags that extend longer than such constraints are commonly reported for invasive plants in varying geographic locations (Kowarik 1995, Crooks and Soulé 1999, Aikio et al. 2010a, Larkin 2011).

Several explanations for why extended lags occur have been proposed. A species might only begin rapidly expanding after it has adapted to the novel range (e.g., evolution via hybridization (Ellstrand and Schierenbeck 2000)). It may expand following landscape disturbance, such as logging or a change in ecosystem conditions (e.g., *F. alnus* and logging (Lee and Thompson 2012)). It may only expand after an invader that facilitates its expansion enters the ecosystem (e.g., *Celastrus orbiculatus* (Oriental bittersweet) and *Sturnus vulgaris* (European starlings) (Merow et al. 2011)).

”

**Methods**

Across all six sites, 815 individual plants were measured and tagged with a unique identifier, 182 first tagged in 2010, 567 in 2011, and 65 in 2012.

At four of the six sites, I counted the number of fruit on tagged plants during the month of July of 2012 to estimate differences in timing of fruit set within a season and to estimate annual maximum fruit per plant values.

Another advantage to using the IPM framework is that the regression analyses allowed me to estimate parameter uncertainty bounds to use in a GSA (described below in **Exploring parameter uncertainty space via global sensitivity analysis**).

Other records were reported in the Early Detection and Distribution Mapping System database (EDDMapS n.d.), iMap Invasives – New York (© 2013 The Nature Conservancy) database, and the Great Lakes Indian Fish and Wildlife Commission invasive species survey data (Garske and Falck 2007, Garske 2010).

Approximately 75% of occurrences were collected after the year 2000.

The climate/ weather variables matched a set of widely applied bioclimatic variables (Hijmans et al. 2005), and were calculated as 10-year and 50-year aggregated climate variables. All variables were provided as raster layers (WGS84 projection with a decimal degree coordinate reference system) at a spatial resolution of 5 x 5 arc minutes. The variables were continuous measurements with the exceptions of potential vegetation, pH top, and pH sub. While each predictor variable was chosen based on hypothesized importance in plant performance, some variables were not used in the final SDMs because they were highly correlated with other variables. Eliminations were based both on correlations with other variables and results from preliminary SDMs (see appendix for details).

**Methods – Supp Info**

I could have optionally used height as a structuring variable, however I collected height data as categorical values above 400 cm. DAH and height for plants less than 400cm were highly correlated (*r =* 0.885, *t*1484 = 73.149, *P* < 0.001), therefore I expect both variables would lead to similar results when used as a structuring variable.

Examining summary values for the fruit count data show that mean plot density is generally larger, or nearly equivalent, for plants that did not fruit versus those that did (Table A6). Plot density had a weak, but significant, effect on whether a plant was observed with fruit or not (Logistic Regression, P < 0.001, coefficient estimate -0.006). Because this effect was weak, I chose to focus only on the effect of plot density on the total number of fruit observed on a plant.

The inconsistent effects of density dependence estimated using the eight DAH size classes above makes it difficult to use these results to parameterize the effects of density dependence in my demographic model. Why do some size classes appear to have positive density dependent growth effects, while others have negative effects? Negative density dependent growth affects have been observed for *F. alnus*, specifically relating to fruit number (Medan 1994). It is possible that positive density dependence effects are the result of the fact that plants will grow in clumps, i.e. localized areas of high density. Godwin (1936) observed that the majority of fruit (1268 of 1804) on one particularly high fruiting individual plant dropped to the ground before being eaten by birds. Similar observations have been made for the functionally closely related *Rhamnus cathartica* (Common buckthorn; Knight et al. 2007). Such minimal amounts of dispersal may result in very clumped distributions of *F. alnus* on a landscape. However, while my analysis may suggest positive density dependence for some DAH size classes, the pattern appears to be primarily driven by plants with no observed fruit at all density values. The majority of plants greater than 0.5 cm DAH produced no fruit (647 not fruiting versus 203 fruiting). There are several reasons, not related to density, that may explain why I observed no fruit on a plant. First, some observations were taken later in the season than others, and fruit may have already dropped or been taken by birds. Second, some sites, and within sites, some plots, may have conditions more conducive to fruit production. Third, there may be year-to-year variability within individuals not accounted for here. That is, if a plant produces a large amount of fruit one year, it may not produce as much the following year.

**Model Evalutation**

**LOOK THROUGH ALL OF THE BELOW TO MAKE SURE IT FLOWS!!!**

A confusion matrix is 2 x 2 matrix of counts of true positive, true negative, false positive, and false negative predictions.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Actual Status | |
|  |  | Present | Absent |
| Predicted status | Present | True positive (A) | False positive (B) |
| Absent | False absence (C) | True absence (D) |

In this analysis, I considered the historical occurrence observations of *F. alnus* as the actual presences (i.e., A + C) and the simulated occurrence observations as predicted presences (i.e., A + B). False absences were counted as the historical occurrence locations where the simulation did not predict *F. alnus* to occur (C). False positives (B) and true absences (D) were challenging to estimate because the data collected for *F. alnus* occurrences was presence only. It did not include verified absence data, and thus the values of B and D may not be accurate (Boyce et al. 2002). However, these values can be important metrics for estimating simulation over prediction. I accounted for the lack of verified absence data by considering simulation predictions in only a subset of patches where either *F. alnus* or an associated species were observed historically (974 of 3423 patches; associated species are defined in AIELLO-LAMMENS IN REVIEW). It is not possible to be certain that a patch in which one of the associated species was observed but *F. alnus* was not accurately represents an absence. However, these patches are more likely to represent absences than a randomly selected patch in which no occurrence records (either *F. alnus* or an associated species) are recorded. Using historic and simulated occurrences in this subset of patches, I calculated the **positive predictive power**, A / (A + B), which is the probability that a patch that is predicted as being occupied is in fact occupied (Fielding and Bell 1997). In terms of the area of occurrence, positive predictive power is equal to:

where is the number of grid cells predicted occupied in the simulation, is the number of grid cells occupied based on historical observations, and is the number of grid cells both predicted occupied and observed occupied. As false positives (B) decrease, and converge, and the positive predictive power increases. Such a result would occur if more patches were in fact occupied than were observed as being so.

I also calculated model **sensitivity**, A / (A+C), as a measures of how well a simulation model predicted observed historic occurrences (Fielding and Bell 1997). In terms of area of occurrence, this measure is:

The grid cells in that are *not* in are omissions (C), or false absences, and sensitivity = 1 – omission rate. If the is relatively small, compared to , then the sensitivity is small and the omission rate is high. Likewise, as the sensitivity approaches a value of 1, the omission rate approaches 0. This measure does not include any information concerning absences, and is often used as a measure of model fit in simulations using presence only data (Boyce et al. 2002, Merow et al. 2011).

**Discussion (maybe)**

(THIS PASSAGE COMES FROM MY INTRO, BUT SEEMS MORE LIKE DISCUSSION TO ME NOW?) This is similar to patterns found by Larkin (2011) for this species regionally in Michigan and Wisconsin (i.e., a lag of approximately 30 – 40 years).

I THINK THIS PASSAGE IS REDUNDANT WITH OTHER ASSERTIONS MADE IN THE DISCUSSION

The patterns I found here are indicative of many ornamental plants introduced into North America during the 20th century. In a recent study, Larkin (2011) examined correlations between plant characteristics and the length of lag phases for almost 100 plants in three distinct regions in Michigan and Wisconsin. He did not find any strong predictive relationships between these characteristics and lag phase length, suggesting that dynamics governing lag-phases may be specific to species and circumstance. However, there was evidence that species introduced for ornamental purposes (65% of the species examined, including *F. alnus*), experienced comparatively shorter lag phases than other invaders, providing further support for the importance of human assisted dispersal.