**Running head:** Herbaria records inform observations of lag phases

**Title of paper:** Observations of extended lag phase of non-native invasive *Frangula alnus* may be spatial-scale dependent

**Author:** Matthew E. Aiello-Lammens

**Affiliation:** Pace University, Department of Environmental Studies and Science

**Address:** 861 Bedford Road, Pleasantville, NY 10570

**Email and phone number of corresponding author:** maiellolammens@pace.edu; 631-327-2371

**Abstract**

Analysis of herbaria records allows for an examination of patterns of spatial spread of non-native plants in novel ranges, aiding in understanding the processes that govern non-native species invasions. In particular, herbarium records have been used to estimate extended lag phase durations for non-native invasive species. I used herbaria records to investigate the rate of spread and pattern of establishment for the invasive plant *Frangula alnus* in northeastern and central North America. I collected records spanning a temporal range from ca. 1880 to the present and a spatial range covering the entire invaded area in northeast North America. To address unequal sampling effort in specimen collection, I compared temporal and spatial patterns of *F. alnus* accessions to patterns in a group of ecologically similar native species. *Frangula alnus* likely had multiple initial introductions into North America, which were geographically separated, ranging from southern Ontario to the coastal Mid-Atlantic region. Trends in record collection in time and space show the rate of spread of *F. alnus* was initially slow, then increased rapidly during the early 20th century, and reached a relatively constant rate of spread in the later 20th century. Examining the spread of this species at the continental scale, there is little evidence that it experienced an extended lag phase between phases of establishment and rapid spatial spread, in contradiction to previous reports. Rather, it appears that *F. alnus* has steadily increased in area of occupancy since ca. 1920. These results raise the question of whether extended lag phases may be a spatial-scale specific pattern. The analytical methods presented here provide one way to investigate this question further.

**Keywords:** species invasion; *Frangula alnus*; lag phase; herbarium; spatial spread

**Introduction (850 words – target is 1000)**

Trade and travel have contributed to the introduction of numerous species to places where they have no evolutionary or ecological history. Many introduced species perish, some establish self-sustaining populations, and a few grow to such high numbers that they negatively impact native species. The population dynamics of a non-native species transitioning from one established in a new region to one rapidly growing in abundance and occupied area are rarely directly observed. We must piece together this history by examining emerging patterns taken from snapshots in time.

For most non-native invasive species there is relatively little population growth in number and area from the time of establishment to the time of rapid spread. This period is commonly referred to as a lag phase (Crooks 2005, Crooks and Soulé 1999, Kowarik 1995, Theoharides and Dukes 2007). Because the lag phase is early in the invasion process when population sizes are small, management actions taken during this time have the potential to be most effective at mitigating the negative impacts of a species invasion (Pluess et al. 2012). However, for these same reasons, this phase often goes unobserved.

Many invasive species show evidence of extended lags in growth of population size in the number of individuals and in range size, but factors causing extended lags are not well understood. Several ecological and evolutionary processes may be involved (Gurevitch et al. 2011, Mack et al. 2000, Pyšek and Hulme 2005, Sakai et al. 2001). Our lack of understanding is a consequence of relatively little empirical work examining lags (Aikio et al. 2010a), despite much theoretical work (as reviewed by Crooks 2005). Historical biodiversity data from herbaria and museums can partly address this lack of empirical support, providing data to re-construct patterns of spatial spread of invasive species. While lags are often discussed through the lens of population sizes, spatial patterns of species occurrence are directly related, as they are emergent properties of population processes (Merow et al. 2011).

The specimen holdings of herbaria provide a rich data source for spatial analyses of plant species. In fact, analysis of the information from herbarium records have been used to estimate species rates of spread through time and space (e.g., Aikio et al. 2010a, Crawford and Hoagland 2009, Delisle et al. 2003, Lavoie 2012, Miller et al. 2009, Newbold 2010, Salo 2005). One limitation of this approach is that herbarium records are sometimes collected with unequal sampling effort in time and/ or space, resulting in biases. For example, temporal trends may reflect variation in collection rate (Prather et al. 2004). Similarly, spatial bias may result from geographically limited collecting, as happens when collections are done close to a home institution, or other developed area, as a matter of convenience (Hijmans et al. 2000, Kadmon et al. 2004). These potential biases make it difficult to determine if observed trends in collections, such as increased number of records through time, are associated with changes in the population size or range size of the species of interest, or instead representative of trends in specimen collection. We must distinguish these patterns to use these data for understanding ecological processes of species invasions.

A small number of methods address the problems associated with unequal sampling effort. Most involve comparison of trends in the distribution of records for a species of interest to other species with similar habitat requirements – i.e. associated species (Miller et al. 2009). For example, Catling and Porebski (1994) studying the spread of *Frangula alnus* in Ontario, compared patterns of collection records of *F. alnus* and native *Rhamnus alnifolia*, showing that specimens were collected well outside of urban areas during the time frame of the introduction and early spread of *F. alnus*. They thus concluded that *F. alnus* was primarily located near urban areas early in its invasion based on habitat, and not on unequal sampling effort. Their approach was a visual qualitative assessment, but recent developments extend this idea quantitatively (Aikio et al. 2010a, Delisle et al. 2003), and have been used to reconstruct patterns of range expansion for both non-native and native plants (e.g., Miller et al. 2009; Larkin 2011). The method developed by Aikio and colleagues (2010a) was specifically developed and applied to identify the existence, and estimate the duration, of extended lag phases for invasive plants.

In this study, I reconstructed the spread of the non-native invasive species *Frangula alnus* throughout its novel range in northeastern North American. I compiled a dataset of historical presence observations of *F. alnus* based on herbarium records and the scientific literature. I calculated measures related to its rate of spatial expansion, while accounting for potential unequal sampling efforts. I hypothesized that I would find support for previous claims that *F. alnus* experienced an approximately 100 year extended lag, lasting from the time of its introduction, ca. 1860, to the early 1970s. The approach I employ here could be particularly useful when examining lag phases of non-native species, because at present, lags most likely occurred during a time when natural history collections were spatially sparse and technology did not allow for rapid digitization, resulting in limited availability for continental scale analyses.

**Materials and Methods (1813) words – target – cut ~200)**

Study Species. *Frangula alnus* Miller (Rhamnaceae, syn. *Rhamnus frangula*), common name Glossy Buckthorn, is a shrub or small tree, 3 to 7 m tall at maturity, native to Eurasia. It produces fleshy fruits from mid-summer to mid-fall, each with *ca.* 2-3 seeds. Large plants yield as many as *ca.* 430 to 1560 seeds per year (Medan 1994). Fruits can be consumed by birds and small mammals, contributing to seed dispersal. It is a habitat generalist, but occurs in particularly high density in wetlands, riparian areas, and old fields (Godwin 1943, Medan 1994). In North America, it is common in many areas in the northeastern region and has been shown to negatively impact native flora (Fagan and Peart 2004, Frappier et al. 2003), making it a species of concern for land-managers (Catling and Porebski 1994, Cunard and Lee 2008).

Collection of presence records. I collected historical presence records for *F. alnus* throughout northeast North America by requesting record information from herbaria curators and collection managers and by searching on-line herbaria databases, on-line biodiversity databases, and the scientific literature using the keywords “Frangula alnus OR Rhammus frangula” in the ISI Web of Science database. Collection managers from Rutgers Chrysler Herbarium, Miami University of Ohio Herbarium, Carnegie Mellon Herbarium, and Cleveland Museum of Natural History Herbarium provided record information for all *F. alnus* or *R. frangula* in their collections. I searched the Global Biodiversity Information Facility (GBIF), the Canada Biodiversity Information Facility (CBIF), and the Invasive Plant Atlas of New England (IPANE) for records of *F. alnus,* and excluded records accessed in other searches. Similarly, I excluded any presences from the literature search corresponding with herbarium specimens. I retained all presence records with information on the year and location (county level or finer) of observation.

Georeferencing presence records. For records lacking latitude and longitude location information, I used descriptive location data (e.g., town or building names, etc.) and GoogleEarth (Google Inc.) to georeference records (*sensu* Garcia-Milagros & Funk 2010). I assigned location uncertainty values following guidelines of the Biogeomancer Consortium (Chapman and Wieczork 2006).

Number of records through time. Trends in the number of records added to herbaria through time can be used to infer rates of spread for non-native invasive species (*sensu* Aikio et al. 2010a, Pyšek and Prach 1993). I calculated the *log* of the cumulative number of *F. alnus* presence records per year and plotted these values against time. I fit linear, quadratic, and cubic regression models to these data, and used a likelihood ratio test to determine the best-fit model. A linear trend would imply a constant rate of increase through time, and these comparisons allow for a test of deviations from such a trend. I calculated an annual rate of growth for the cumulative number of presence records by dividing the cumulative number of records at year *t +* 1 by those at year *t*. I then calculated the geometric mean growth rate and 10-year moving window geometric mean growth rates, the later which minimizes the influence of extreme inter-annual fluctuations in growth rates.

Area of occupancy through time. To examine spatial spread of *F. alnus*, I overlaid 5 x 5 arc minute grid (each cell approximately 57 km2 to 67 km2 in area) onto a map of presence locations. The grid was constructed using the QGIS software (QGIS Development Team 2012). I used the R statistical programming environment (R Core Team 2012) with additional functions from the “raster”, “rgdal”, and “dismo” packages (Hijmans 2012, Hijmans et al. 2013, Keitt et al. 2012, respectively) to carry out all other analyses. Each presence record was assigned to a grid cell based on their latitude and longitude values. If a grid cell contained at least one *F. alnus* presence, I considered it occupied. I then calculated the total number of grid cells occupied per decade and the cumulative number of grid cells occupied from time of first introduction to the present. For the latter, I assumed once a grid cell was occupied, it remained so. These measures correspond to those used in other studies (e.g., Delisle et al. 2003, Pyšek and Prach 1995, Weber 1998). I calculated the rate of growth for the area of occupancy following the method I used for the number of records, using cumulative number of occupied grid cells instead of cumulative number of records. I plotted the square root of the cumulative number of grid cells versus time (years) and tested for deviations from linearity.

Occupied counties through time. Many records could only be georeferenced to the county in which they occurred. Similar to the calculations ofarea of occupancy through time, I calculated the cumulative number of counties occupied by *F. alnus* through time. I assumed that once *F. alnus* was found in a county, it was thereafter occupied. I calculated the growth rate for the cumulative number of counties occupied as above.

Accounting for unequal sampling effort in time and space. Comparing trends in records of a non-native species to those of native species can aid in distinguishing patterns resulting from a species expansion versus unequal sampling effort (Delisle et al. 2003). I examined patterns in herbarium records for *Alnus incana* (Speckled Alder), *A. serrulata* (Smooth Alder), *Rhamnus alnifolia* (Alderleaf Buckthorn), *Salix peiolarisi* (Meadow Willow)*, Hamamelis virginica* (syn. *macrophylla*) (Witch Hazel), and *Fraxinus americana* (White Ash), and compared these to those of *F. alnus*. These plants have similar habitat requirements as *F. alnus* and often co-occur with it (personal observation, Catling and Porebski 1994, Houlahan and Findlay 2004, Little et al. 1980). I collected records of these species reported in the University of Wisconsin, Ohio State University, University of Minnesota, Morton Arboretum Herbarium, Michigan State University, Cleveland Museum of Natural History, and Brooklyn Botanic Gardens herbaria databases, as well as searching GBIF. Records were georeferenced to the county level as needed. Grouping these records, I calculated the metrics for the number of records through time, the number of grid cells occupied through time, and the number of counties occupied through time, as described above.

I compared the trends for *F. alnus* and the group of associated species for both number of grid cells occupied through time and counties occupied through time using a modified version of the method of Delisle et al. (2003). They compared trends in area of occupancy through time by dividing the cumulative number of occupied 10 x10 km grid cells for each non-native plant by the cumulative number of cells occupied by a group of five native species, yielding a proportion of non-native to native occupied cells for each year of their study period. Assuming that spatial growth of native plants in the herbaria represents the increase in spatial coverage of herbaria records (i.e., an increase in the number of locations where samples are collected), and not the spread of these species *per se*, a number of conclusions can be drawn. An increasing proportion through time suggests periods during which the spatial coverage of records for the non-native plant outpaces the background increase in spatial coverage. This can be interpreted as the result of the non-native plant rapidly spreading. A decreasing proportion suggests a period of relatively slow spatial spread for the non-native species (Delisle et al. 2003).

I examined spatial spread in a larger region than that studied by Delisle and colleagues (2003) and used records from multiple herbaria. Further, several presence records for the group of associated species were located in areas where *F. alnus* has not been observed, likely due to the larger niche breadth of some of the associated species compared to *F. alnus*. Including these records increased the area of occupancy measures for the group of associated species compared to the possible area of occupancy for *F. alnus*. Similarly, I collected *F. alnus* presence records from institutions I did not collect associated species records from (i.e., Miami University Herbarium and Rutgers Herbarium). These locations were not always well represented in the associated species dataset. Therefore, including these records increased the area of occupancy of *F. alnus* compared to the possible area of occupancy for the group of associated species. To account for both of these issues, I examined the ratio of cumulative area of occupancy of *F. alnus* to the cumulative area of occupancy of the associated group of species, limiting the records used to an area of coarse spatial overlap for both datasets. I defined the spatial overlap by creating a 30 x 30 arc minute grid (i.e., approximately 50 x 50 km) and overlaying this on a map of all presence locations. Each record was assigned membership to one 30 arc minute grid cell. For subsequent analyses I only used records that occurred in 30 arc minute grid cells occupied by at least one record from both datasets during the study period. Using these restricted data, I calculated the ratio of the increase in the cumulative area of occupancy of *F. alnus* to the group of associated species. I also compared the cumulative number of counties occupied through time, using occurrences where only records from counties that were occupied at some time by both *F. alnus* and one of the associated plants. I compared growth rates for the cumulative number of grid cells occupied and the cumulative number of counties occupied between *F. alnus* and the group of associated species by dividing the annual growth rate of *F. alnus* records by the annual growth rate of the entire group of associated species.

Comparison of number of records through time. Similar to the comparisons of areal increase described above, I compared trends in the cumulative number of records through time. For each year of the study period, I divided the cumulative number of *F. alnus* records by the cumulative number of associated species records. A similar correction method has been used in several studies to account for bias due to unequal sampling effort (e.g., Mihulka & Pyšek 2001; Aikio et al. 2010a; Larkin 2011). To compare rates at which *F. alnus* and associated species presences were collected, I calculated the ratio of growth rates using both the annual and 10-year moving window average estimates. A ratio value less than 1.0 indicated that the rate at which *F. alnus* presences were being collected was lower than that of the associated species, while a ratio value of greater than 1.0 indicated that it was greater than that of the associated species.

NULL MODEL. Because herbarium collections are relatively sparse during most of the 19th century, when *F. alnus* was introduced, it is difficult to be confident that the correction methods described above are adequately accounting for the potential effects of increased collection effort. To address this, I used a null model analysis approach similar to that of (Aikio 2010b). The *F. alnus*  and associated species data sets were combined into a single data set consisting of columns for species identification, grid cell identification, and date of collection. I then permuted both the grid cell identification and date of collection columns. This has the effect of shuffling both when and where occurrences were collected, but maintaining the total number of occurrences associated with specific years and locations independently. After these permutations, the cumulative area of occupancy through time values and ratio of the cumulative area of occupancy through time values between *F. alnus* and the associated species were calculated. In total, 999 permutations were carried out.

**Results**

I compiled a dataset of 809 *F. alnus* presence records and 6057 associated species presence records from multiple sources (Table 1). These records show that *F. alnus* was collected across northeast and middle North America (Fig. 1). The associated species collection locations were largely inclusive of where *F. alnus* was collected. The earliest dated presence of *F. alnus* was in 1879 in Hudson County, New Jersey (accessed from CHRB). The earliest dated presence of one of the associated species was for *Hamamelis virginiana* in 1836 in Richland County, Ohio (accessed from CM). Only 36 records of associated species pre-dated the first *F. alnus* record. A total of 14 *F. alnus* specimens were deposited in 4 separate herbaria prior to 1900; 12 were georeferenced to the metropolitan New York City region (9 in the state of New Jersey) and 2 were collected in southern Ontario (as reported in Catling & Porebski 1994). The proportion of the total number of herbarium specimens collected in each decade for both *F. alnus* and the group of associated species tended to increase through time (Fig. 2). The associated species showed a substantial increase in the proportion of the number of records collected during the first half of the 20th century, followed by another increase after 1950. There was a steady increase in the proportion of the number of records collected per decade from the time of the first recorded presence of *F. alnus* in the United States to the end of the 20th century. For both groups, there was a decline in the number of records at the start of 21st century, which is consistent with herbarium collection in general (Prather et al. 2004).

The trend of the log cumulative number of records through time indicates a deviation from exponential growth for both *F. alnus* and the group of associated species (Fig. 3A). The best-fit regression models were a cubic polynomial regression for *F. alnus* (R2 = 0.989, df = 124, P << 0.05) and a quadratic polynomial regression for the associated species (R2 = 0.979, df = 174, P << 0.05). The cumulative number of records for the group of associated species increased steadily from the late 1800s to the present. The cumulative number of records for *F. alnus* increased rapidly from the late 1800s to approximately 1930. Ratios of the annual growth rates and ten year moving average growth rates show that the rate for *F. alnus* collections was generally slower than that of the associated species prior to approximately 1920, but faster during most of the 20th century (Fig. 3B). The ratio of the cumulative increase in records of *F. alnus* to the associated species corroborates this pattern, indicating that the increase of *F. alnus* records was greater than the associated species from 1920 to the present (Fig. 3C).

The cumulative number of occupied 5 arc minute grid cells for *F. alnus* and the associated species, constrained to overlap within 30 arc minute grid cells, increased through time for both datasets (Fig. 4). Trends in the square root of cumulative occupied grid cells versus time supports a departure from a simple diffusion model of spatial spread in historical presence records (Fig. 4A). The best-fit regression models were a quadratic polynomial regression for *F. alnus* (R2 = 0.996, df = 129, P << 0.05) and a cubic polynomial regression for the group of associated species (R2 = 0.990, df = 140, P << 0.05), both indicating a departure from linearity. The rate of increase in presence records was low for both *F. alnus* and the associated species until 1890, after which the number of occupied cells rapidly increased for the group of associated species. In contrast, the number of occupied cells continued to increase slowly for *F. alnus* until approximately 1920, as demonstrated by the ratio between the two growth rates (Fig. 4B). Prior to 1915, with the exception of the earliest years of the *F. alnus* invasion, the rate at which new grid cells were considered occupied was lower than the corresponding rate for the associated species. From 1915 to the present, the rate at which grid cells were considered occupied by *F. alnus* was consistently higher than the corresponding rate of growth for the associated species. The ratio of the cumulative number of occupied grid cells of *F. alnus* to the associated species shows that after approximately 1910 the rate at which *F. alnus* increased in area of presence was greater than that of the group of associated species (Fig. 4C).

Trends in the cumulative number of counties occupied were similar to the results of the analysis of the cumulative number of occupied grid cells. The number of counties where herbarium samples were collected increased rapidly during the late 19th to early 20th century. In contrast, the number of counties where *F. alnus* was found increased slowly during the early 20th century, but rapidly after 1940 (Fig. 5A). Assuming counties were selected for specimen collections randomly, the relationship between the square root of the cumulative number of occupied counties and time (years) should be linear for the associated species. However, this was not the case for either the associated species or *F. alnus* (Fig. 5A). A cubic polynomial regression model was the best-fit model for both the associated species (R2 = 0.967, df = 166, P < 0.001) and *F. alnus* (R2 = 0.997, df = 130, P < 0.001). The difference in the rate of growth of cumulative occupied counties between *F. alnus*­ and the associated species shows a similar pattern to that of the cumulative occupied grid cells (Fig. 5B). The rate at which counties are considered occupied is slower for *F. alnus* than the associated species early in the invasion history (prior to 1900) and faster during most of the 20th century. The ratio of occupied counties of *F. alnus* to those occupied by the associated species shows that early in the invasion history, *F. alnus* was less frequently collected in newly sampled counties, but that for most of the 20th century the number of counties occupied by *F. alnus* increased more rapidly than the number of counties occupied by associated species (Fig. 5C). *Frangula alnus* specimens were collected on average 48 years after at least one of the associated species had been collected in a county.

Examining the patterns in the null model analysis of cumulative occupied grid cells through time and the ratio of occupied grid cells through time, an extended lag phase lasting from its introduction until approximately 1920 seems more apparent for *F. alnus*. Comparing the trends in the cumulative number of grid cells occupied for *F. alnus* and the associated species under the null model with the trends observed in the real data shows two different patterns (Figure 6A). The associated species null model pattern matches the shape of the real data, including showing an abrupt change in the rate of collections at approximately 1920. The *F. alnus* null model pattern, on the other hand, differs from the pattern in the real data. The null model pattern resembles that of the associated species, including a concave down pattern indicative of a slowing of the species spread, whereas the actual data show a concave up pattern, indicative of an acceleration of the species spread. The ratio of occupied grid cells of *F. alnus* to those occupied by the associated species based on the null model results is similar to the actual ratio, in that there is a period of decline in the ratio early in its invasion history, followed by a transition to an increasing ratio (Figure 6B). However, this transition occurs far earlier in the invasion history in the null model analysis compared to the actual data. Furthermore, neither the decline in ratio prior to 1900 nor the increasing ratio after 1900 are as substantial as that observed in the real data. In fact, relative to the observed ratios, the null model ratios vary very little over the length of *F. alnus* invasion history.

**Discussion (2288 words – target – cut ~500)**

Early observations and likely region of first introduction. The exact introduction event for *F. alnus* is unknown, but its long history of use as a medicinal plant (United States Pharmacopeial Convention 1910) and evidence that it was planted as an ornamental shrub (Sudworth and Fernow 1891, Sherff 1912), suggest it was likely purposely planted in many locations in North America. Based on the historic presence records, the location of first introduction was likely the metropolitan New York City area and areas of New Jersey near New York Harbor. Specimens for 12 of 14 records dating from before 1900 were collected in this region. New York Harbor was a large shipping and port area, raising the possibility that shipping may also have played a role in an accidental introduction. Many non-native plants have been introduced to port areas by the unloading of solid ballast, with seeds mixed in with rocks and other materials used as ballast (Barney 2006, Sorrie 2005). Locations in southern Ontario account for the other two pre-1900 records, indicating that introduction into that region was early in the spread of *F. alnus* as well.

Accounting for unequal sampling effort in historical biodiversity collections. Unequal sampling effort in space or time is common in historical biodiversity collections (Boakes et al. 2010, Kadmon et al. 2004). Calculating the ratio of the cumulative number of records in space and time of non-native to native plants provides a way to account for such biases and to distinguish periods of relatively slow versus rapid spread (Delisle et al. 2003). The primary assumption in this method is that native species are in equilibrium with their environment prior to the collection of any records. This might not be the case in my study region given that ecological conditions in northeast North America have not been stable over the last 130 years. Land use changes have caused changes in plant communities (Wright and Fridley 2010), and affected invasion dynamics (Mosher et al. 2009). However, given that the associated species have similar ecological requirements as *F. alnus*, I expect that any such changes would affect trends in these species and *F. alnus* equally.

Assuming the associated species are in ecological equilibrium, the rate of the cumulative number of records, grid cells, or counties occupied for these species is representative of the effort of specimen collectors, rather than the spread of these plants *per se*. If the ratio of cumulative records of *F. alnus* to the associated species is increasing in time, this indicates a period when *F. alnus* sampling is increasing more rapidly than background sampling, likely do to its rapid growth. If the ratio is constant, *F. alnus* may be increasing, however indistinguishably from sampling effort. If the ratio is decreasing, *F. alnus* may still be increasing, but more slowly than sampling effort. Because the *cumulative* number of records was used in all three temporal trends I investigated, the absolute rate of change in samples cannot show a decline. Thus, a period of decreasing or stable ratio values preceding periods of increasing ratio values may indicate lag phases.

Increases of *Frangula alnus* through time. The ratios between cumulative records through time, occupied grid cells through time, and occupied counties through time of *F. alnus* and the associated species each indicate that *F. alnus* has increased rapidly since ca. 1920 (Figs. 4C, 5C, and 6C). For each, the ratio of ten-year average growth rates was greater than 1.0 after 1920, indicating that the number of *F. alnus* samples increased more rapidly than those of the associated species (Fig. 3B, 5B, and 6B). Prior to 1920, both the ratio values of cumulative number of records and growth rates fluctuated greatly. This is likely the result of the small number of records for both *F. alnus* and the associated species during this time. The addition of only a few records to either dataset would dramatically change the ratio between them. For example, *F. alnus* records increased more slowly than the associated species from 1879 to the 1920s, but there are anomalous years in which its growth rate was much larger than that of the associated species (Fig. 3B, 5B, and 6B). When the cumulative number of records for *F. alnus* increased from 4 to 10 from 1893 to 1894, the resulting growth rate was *R* = 2.5. Comparatively, during this time the number of records for the associated species increased from 171 to 191, resulting in a growth rate of *R* = 1.12. It is difficult to determine if *F. alnus* did in fact start its invasion with a very high growth rate, then immediately slowed, or rather this result is an artifact of calculating growth rates with small sample sizes. Based on the relative consistency of the proportion of total records collected in each decade (Fig. 2), there is more support for the latter interpretation.

Trends of the square root of the cumulative number of grid cells (or counties) occupied by the associated species through time did not fit linear relationships (Fig. 4A and 5A). Assuming these species were already present in most suitable environments (i.e., at equilibrium), this suggests that rates of specimen collection changed through time. Both trends can be described as initially concave up, followed by a steep linear trend, followed by a concave down curve, followed by another period of linearity. This suggests the rate of collections increased rapidly in the early 20th century, was high throughout the second half of the 20th century, and then decelerated near the end of the 20th century and beginning of the 21st century. Combined with the increase in the number of records during the earlier part of the study period (Fig. 2 and 3A), these patterns suggest that the early 20th century was a period of high sampling effort, coinciding with the beginning of an intense effort to collect specimens for herbaria as previously documented (Prather et al. 2004).

The earliest presence records of *F. alnus* were collected during the period of high sampling effort near the end of the 19th century. However, it remained uncommon across the study region until at least the 1920s. Assuming that cumulative presence curves for the associated species represent the spatial spread of collection effort, and that a collector would collect *F. alnus* if it were present during a survey, it appears that the rate of spatial spread of *F. alnus* during the early 20th century was slow. This is most clearly exemplified by the trends in cumulative number of counties occupied through time. On average *F. alnus* was first observed 48 years after at least one of the associated species was observed in a county. By 1900, records for at least one of the associated species was collected in 42% (73 of 172) of counties, where as *F. alnus* was collected in 2% (4 of 172) of counties. This strongly suggests that *F. alnus* was not common at this time. However, the ratio of presence records shows a consistent rate of spread throughout the novel range during the following 100 years.

Evidence for an extended lag phase. Examining the relative rates of increase in occupied grid cells and occupied counties at this spatial scale, there is no compelling evidence for an extended lag phase persisting beyond the early 20th century. All three analyses suggest that *F. alnus* increased at a rate greater than sampling effort from at least 1920 on. Given that the rate of increase in record number and spatial presence for *F. alnus* was less than that of the associated species from 1880 to 1920, it is plausible that this period of time represents an extended lag phase. However, trends calculated for this period are based on a small number of *F. alnus* records, as is evident in the high variability in both the ratio of growth rates and the ratio of records of occupied area.

At least three previous studies examined aspects of the range expansion of *F. alnus* using herbarium records (Catling and Porebski 1994, Howell and Blackwell 1977, Larkin 2011). Each focused on a part of the invaded range. Howell & Blackwell (1977) examined its spread into, and throughout, Ohio. They found evidence that it likely entered Ohio in the 1920’s, first observed in Lake County (northeastern Ohio). The authors speculate that the spread of *F. alnus* westward throughout Ohio was facilitated by the range-expansion of the non-native European Starling (*Sturnus vulgaris*). Many bird species are noted to eat the fruit of *F. alnus* and defecate undigested seeds, and *S. vulgaris* in particular have been known to eat these fruit (Howell & Blackwell 1977). The authors did not account for potential unequal sampling effort in herbarium records, but rather simply reported when and where *F. alnus* first appeared in Ohio. Catling & Porebski (1994) examined the historical spread of *F. alnus* in southern Ontario, Canada and found that it was first recorded in London, Ontario in 1898 and Ottawa in 1899. Their data suggested that *F. alnus* spread to other urban centers, but primarily remained confined to these areas until the 1970’s. From the 1970’s to the early 1990’s *F. alnus* spread into natural areas outside of urban areas at an increased rate compared to the previous 70 years. The authors suggest that this observation is indicative of a lag phase in the spread of *F. alnus*, though no quantitative analysis was carried out. More recently, Larkin (2011) examined the lengths of lag phases for multiple (>200) species of non-native invasive species in the Wisconsin and the southern Lake Michigan region, applying the methods of (Aikio et al. 2010a). The earliest record of *F. alnus* was collected in 1908 in the southern Lake Michigan region, and based on quantitative analysis had a 31-year lag phase. Comparatively, it was found in southern and northern Wisconsin in 1927 and 1941, and had 36- and 15-year lag phases, respectively. The timings of these region-specific introductions suggest that *F. alnus* spread north from the southern Lake Michigan region into more northern parts of Wisconsin. The length of the lag phase calculated for southern Lake Michigan and southern Wisconsin generally agrees with my observations for the whole range. Further, the pattern of *F. alnus* spreading from urban areas into more rural regions corresponds with the results of recent genetic analyses (De Kort et al. 2016).

Caveats and potential sources of bias. I assumed the cumulative number of occupied grid cells, and counties, as representative of the area of extent. This may not be valid if *F. alnus* invades an area, but later goes locally extinct. However, this is an unlikely occurrence. With one notable exception (Cunard and Lee 2008), there are no documented examples of the local extinction of *F. alnus*, either by natural processes or management actions.This supports the notion that once a location is occupied, it remains so. Nevertheless, it is likely that some locations of historical *F. alnus* presences no longer have the species present because of changes in land use (e.g., development of once-woodland plots). This is the case for several of the records observed in the metropolitan New York region, the site of many of the earliest records. Disturbances, or removal of populations from the landscape, due to anthropogenic influences may have a substantial effect on the spread of *F. alnus* throughout its novel region, but it can still be successful in very small, isolated, and disturbed plots within an urban or suburban land-use matrix (personal observation, Del Tredici 2010). Given that the finest spatial resolution used in this study was 5 x 5 arc minutes (approximately 10 x 10 km), a complete local extinction of *F. alnus* is unlikely.

Most herbaria do not have electronic databases of their holdings that are easily accessed by the public (i.e., a web-based search interface) and many do not have a complete electronic database of records within the institution (Lavoie 2012). Therefore, there are certainly several herbaria records absent from the compiled dataset. However, given the extent of the data compilation in space and across institutions, I am confident that the patterns and trends reported are generally representative of the spread of *F. alnus*. Future analyses that involve even greater effort in gathering herbarium specimens could increase the accuracy of my results, particularly providing a more accurate area of extent.

**Conclusions**

These results suggest *F. alnus* experienced a rapid rate of spreadthroughout its invasion history. Examining the ratio of the cumulative number of grid cells and counties occupied by *F. alnus* to those occupied by associated species, I found this invasive expanded rapidly throughout its new range since the mid to late 1920s to the present. Patterns of spatial spread and estimates of lag phase duration likely vary between regions.

However, an analysis from one region yielded an extended lag phase length consistent with what was observed throughout the entire invaded range (Larkin 2011). The patterns and processes of range expansion, particularly of lag phase dynamics, vary depending on scale and local ecological conditions (Theoharides and Dukes 2007). Future analyses that compare regional patterns to those for the whole range described here could further elucidate the role of spatial scale in processes of species invasions. As access to herbaria records and other historical biodiversity increases, techniques to combine information from desperate sources, such as those I presented here, will help scientists use these data to investigate important standing questions regarding species range expansions.

**Literature Cited**

Aikio, S., R. P. Duncan, and P. E. Hulme. 2010a. Lag-phases in alien plant invasions: separating the facts from the artefacts. Oikos 119:370–378.

Aikio, S., R. P. Duncan, and P. E. Hulme. 2010b. Herbarium records identify the role of long-distance spread in the spatial distribution of alien plants in New Zealand. J. Biogeogr. 37:1740-1751.

Barney, J. N. 2006. North American history of two invasive plant species: phytogeographic distribution, dispersal vectors, and multiple introductions. Biol. Invasions 8:703–717.

Boakes, E. H., P. J. K. McGowan, R. A. Fuller, D. Chang-qing, N. E. Clark, K. O’Connor, and G. M. Mace. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. PLoS Biol. 8:e1000385.

Catling, P. M., and Z. S. Porebski. 1994. The history of invasion and current status of glossy buckthorn, *Rhamnus frangula*, in southern Ontario. Can. Field-Naturalist 108:305–310.

Chapman, A. D. and J. Wieczorek [eds.]. 2006. Guide to Best Practices for Georeferencing. Global Biodiversity Information Facility, Copenhagen. 90 p.

Crawford, P. H. C., and B. W. Hoagland. 2009. Can herbarium records be used to map alien species invasion and native species expansion over the past 100 years? J. Biogeogr. 36:651–661.

Crooks, J. 2005. Lag times and exotic species: the ecology and management of biological invasions in slow-motion. Ecoscience 12:316–329.

Crooks, J. A., and M. E. Soulé. 1999. Lag times in population explosions of invasive species: causes and implications. Pages 103–125 *In* O. T. Sandlund, P. J. Schei, and A. Viken [eds.], Invasive species and biodiversity management. Kluwer Academic Dordrecht, The Netherlands.

Cunard, C., and T. D. Lee. 2008. Is patience a virtue? Succession, light, and the death of invasive glossy buckthorn (*Frangula alnus*). Biol. Invasions 11:577–586.

Delisle, F., C. Lavoie, M. Jean, and D. Lachance. 2003. Reconstructing the spread of invasive plants: taking into account biases associated with herbarium specimens. J. Biogeogr. 30:1033–1042.

Fagan, M. E., and D. R. Peart. 2004. Impact of the invasive shrub glossy buckthorn (*Rhamnus frangula* L.) on juvenile recruitment by canopy trees. For. Ecol. Manage. 194:95–107.

Frappier, B., R. T. Eckert, and T. D. Lee. 2003. Potential impacts of the invasive exotic shrub *Rhamnus frangula* L. (glossy buckthorn) on forests of southern New Hampshire. Northeast. Nat. 10:277–296.

Garcia-Milagros, E., and V. A. Funk. 2010. Improving the use of information from museum specimens: using Google Earth© to georeference Guiana Shield specimens in the US National Herbarium. Front. Biogeogr. 2:71–77.

Godwin, H. 1943. *Frangula alnus* Miller. J. Ecol. 31:77–92.

Gurevitch, J., G. A. Fox, G. M. Wardle, Inderjit, and D. Taub. 2011. Emergent insights from the synthesis of conceptual frameworks for biological invasions. Ecol. Lett. 14:407–418.

Hijmans, R. J. 2012. Introduction to the ‘raster’package (version 2.0-08). R. <http://probability.ca/cran/web/packages/raster/vignettes/Raster.pdf>.

Hijmans, R. J., K. A. Garrett, Z. Huaman, D. P. Zhang, M. Schreuder, and M. Bonierbale. 2000. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. Conserv. Biol. 14:1755–1765.

Hijmans, R. J., S. J. Phillips, J. R. Leathwick, and J. Elith. 2013. dismo: Species distribution modeling. R. < https://cran.r-project.org/web/packages/dismo/vignettes/sdm.pdf>

Houlahan, J. E., and С. S. Findlay. 2004. Effect of invasive plant species on temperate wetland paint diversity. Conserv. Biol. 18:1132–1138.

Howell, J. A., and W. H. J. Blackwell. 1977. The history of *Rhamnus frangula* (glossy buckthorn) in the Ohio flora. Castanea 42:111–115.

Kadmon, R., O. Farber, and A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. Ecol. Appl. 14:401–413.

Keitt, T. H., R. Bivand, and B. Rowlingson. 2012. rgdal. R. <http://cran.r-project.org/web/packages/rgdal/rgdal.pdf>.

De Kort, H., J. Mergeay, H. Jacquemyn, and O. Honnay. 2016. Transatlantic invasion routes and adaptive potential in North American populations of the invasive glossy buckthorn, *Frangula alnus*. Ann. Bot. 118:1089–1099.

Kowarik, I. 1995. Time lags in biological invasions with regard to success and failure of alien species. Pages 15-38 *In* P. Pyšek, K. Prach, M. Rejmánek, and M. Wade [eds.], Plant invasions: General aspects and special problems. SPB Adademic Publishing, Amsterdam.

Larkin, D. J. 2011. Lengths and correlates of lag phases in upper-Midwest plant invasions. Biol. Invasions 14:827–838.

Lavoie, C. 2012. Biological collections in an ever changing world: herbaria as tools for biogeographical and environmental studies. Perspect. Plant Ecol. Evol. Syst. 15:68–76.

Little, E. L., S. Bullaty, and A. Lomeo. 1980. National Audubon Society Field Guide to North American Trees: Eastern Region. Alfred A. Knopf, New York, NY. 714 p.

Mack, R. N., D. Simberloff, W. Mark Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. 2000. Biotic invasions: causes, epidemiology, global consequences, and control. Ecol. Appl. 10:689–710.

Medan, D. 1994. Reproductive biology of *Frangula alnus* (Rhamnaceae) in southern Spain. Plant Syst. Evol. 193:173–186.

Mihulka, S., and P. Pyšek. 2001. Invasion history of Oenothera congeners in Europe: a comparative study of spreading rates in the last 200 years. J. Biogeogr. 28:597–609.

Miller, R. J., A. D. Carroll, T. P. Wilson, and J. Shaw. 2009. Spatiotemporal analysis of three common wetland invasive plant species using herbarium specimens and geographic information systems. Castanea 74:133–145.

Moe, D. 1984. The late quaternary history of *Rhamnus frangula* in Norway. Nord. J. Bot. 4:655–660.

Merow, C., N. LaFleur J. A. Silander Jr, A. M. Wilson, and M. Rubega. 2011. Developing dynamic mechanistic species distribution models: predicting bird-mediated spread of invasive plants across northeastern North America. Am. Nat. 178:30–43.

Mosher, E. S., J. A. Silander Jr, and A. M. Latimer. 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.

Newbold, T. 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. Prog. Phys. Geogr. 34:3–22.

Pluess, T., R. Cannon, V. Jarošík, J. Pergl, P. Pyšek, and S. Bacher. 2012. When are eradication campaigns successful? A test of common assumptions. Biol. Invasions 14:1365–1378.

Prather, L. A., O. Alvarez-Fuentes, M. H. Hayfield, and C. J. Ferguson. 2004. The decline of plant collecting in the United States: a threat to the infrastructure of biodiversity studies. Syst. Bot. 29:15–28.

Pyšek, P., and P. E. Hulme. 2005. Spatio-temporal dynamics of plant invasions: linking pattern to process. Ecoscience 12:302–315.

Pyšek, P., and K. Prach. 1993. Plant invasions and the role of riparian habitats: a comparison of four species alien to central Europe. J. Biogeogr. 20:413–420.

Pyšek, P., and K. Prach. 1995. Invasion dynamics of *Impatiens glandulifera* - a century of spreading reconstructed. Biol. Conserv. 74:41–48.

QGIS Development Team. 2012. QGIS Geographic Information System. <http://qgis.osgeo.org>.

R Core Team. 2012. R: A language and environment for statistical computing. Vienna, Austria. <http://www.r-project.org>.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O’Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. Annu. Rev. Ecol. Syst. 32:305–332.

Salo, L. F. 2005. Red brome (*Bromus rubens* subsp. *madritensis*) in North America: possible modes for early introductions, subsequent spread. Biol. Invasions 7:165–180.

Sherff, E. E. 1912. Range extenstions of *Rhamnus frangula* and *Sporobolus asperifolius*. Rhodora 14:227–229.

Sorrie, B. A. 2005. Alien vascular plants in Massachusetts. Rhodora 107:284–329.

Sudworth, G. B. and B. E. Fernow. 1891. Trees of Washington, D.C.: compliments of the Forestry Division. Bell Brothers Lithographing Co., Washington, D.C. 15 p.

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

Del Tredici, P. 2010. Wild Urban Plants of the Northeast: A Field Guide. Comstock Publishing Associates, Ithica, NY.

United States Pharmacopeial Convention. 1910. The Pharmacopeia of the United States of America, 9th Decennial Revision. P. Blakiston’s Son & Company, Philadelphia, PA.

Weber, E. 1998. The dynamics of plant invasions: a case study of three exotic goldenrod species (*Solidago* L.) in Europe. J. Biogeogr. 25:147–154.

Wright, J. P., and J. D. Fridley. 2010. Biogeographic synthesis of secondary succession rates in eastern North America. J. Biogeogr. 37:1584–1596.

**Table 1.** Historical presence record sources and counts. Abbreviations for “Accession Method” column: CBIF = Canadian Biodiversity Information Facility, GBIF = Global Biodiversity Information Facility, IPANE = Invasive Plant Atlas of New England, and vPlants = Virtual Herbarium of the Chicago Region.

|  |  |  |  |
| --- | --- | --- | --- |
| **Source** | **Herbarium Code** | **Record Count** | **Accession Method** |
| Harvard University Arnold Arboretum | A | 2 | IPANE |
| Acadia University Herbarium | ACAD | 8 | GBIF |
| Botanischer Garten und Botanisches Museum Berlin-Dahlem | B | 1 | GBIF |
| Brooklyn Botanical Garden | BKL | 35 | Institute website |
| Chicago Botanic Garden | CHIC | 3 | Institute website / vPlants |
| Rutgers University Chrysler Herbarium | CHRB | 10 | Provided by curator |
| Cleveland Museum of Natural History | CLM | 69 | Provided by curator |
| Carnegie Museum of Natural History | CM | 68 | Provided by curator |
| University of Connecticut Torrey Herbarium | CONN | 85 | IPANE / GBIF |
| Field Museum of Natural History | F | 28 | Institute website / vPlants |
| Harvard University Herbaria | GH | 5 | IPANE |
| Royal Botanical Gardens | HAM | 7 | CBIF |
| University of Kansas | KANU | 3 | GBIF |
| Forest Products Laboratory | MAD | 1 | Institute website (U. of Wis.) |
| University of Massachusetts | MASS | 22 | IPANE |
| University of Minnesota | MIN | 48 | Institute website |
| Missouri Botanical Garden | MO | 2 | Institute website |
| Morton Arboretum | MOR | 43 | Institute website |
| Michigan State University | MSC | 25 | Institute website |
| Université de Montréal | MT | 22 | GBIF |
| Miami University | MU | 57 | Provided by curator |
| Yale University Connecticut Botanical Society | NCBS | 3 | IPANE |
| New England Botanical Club | NEBC | 63 | IPANE |
| New York Botanical Garden | NY | 5 | Institute website |
| Ohio State University | OS | 81 | Institute website |
| Queen’s University | QK | 2 | Institute website |
| University of Wyoming Rocky Mountain Herbarium | RM | 2 | Institute website |
| Smithsonian Institution | US | 3 | Institute website |
| University of Wisconsin – Green Bay | UWGB | 16 | Institute website (U. of Wis.) |
| University of Wisconsin – Stevens Point | UWSP | 6 | Institute website (U. of Wis.) |
| Naturhistorisches Museum Wien | W | 2 | GBIF |
| University of Wisconsin – Madison | WIS | 85 | Institute website (U. of Wis.) |
| Yale University Peabody Museum of Natural History | YU | 4 | IPANE / GBIF |
| Literature Search |  | 5 | ISI Web of Science |

Fig. 1. Geographic locations of collected records for *F. alnus* and a group of associated species. Triangles represent records from the compiled historical presence records for *F. alnus* and crosses represent records from the compiled historical presence records for the group of associated species. Some locations were assigned geographic locations based on the latitude and longitude values of counties as defined by the US Census Bureau. The dashed black line delineates the study region.

Fig. 2. Proportion of the total number of records collected in each decade for *Frangula alnus* (black bars) and all associated species combined - *Alnus incana*, *A. serrulata*, *Rhamnus alnifolia*, *Salix peiolarisi, Hamamelis virginica* (syn. *macrophylla*), and *Fraxinus americana* (grey bars).

Fig. 3. (A) Log cumulative number of records through time. Linear (solid) and cubic polynomial (dot-dash) regression predictions are plotted over the cumulative increase curves. (B) Ratio of growth rates of cumulative number of records of *F. alnus* versus associated species calculated annually (crosses) and by10 year moving window average (geometric mean) (triangles). Note that some extreme data points are not shown (those >1.2 or <0.8, but contribute to the moving window average values. (C) Ratio of the cumulative number of records of *F. alnus* to those of associated species.

Fig. 4. (A) Square root of the cumulative number of grid cells through time. Shown here are the linear and polynomial regression lines for models using year as a predictor variable and the square root of the cumulative number of grid cells occupied as the response variable. Linear regression predictions are shown for both *F. alnus* and the group of associated species (solid black lines). The best-fit polynomial regression fit is shown for each set (quadratic for *F. alnus* and cubic for the group of associated species, dash and dot-dash lines, respectively) (B) Ratio of growth rates of cumulative occupied grid cells calculated annually (crosses) and by 10 year moving window average (geometric mean) (triangles). Outlier data not shown (>1.2 or <0.8), but do contribute to the moving window average values. (C) Ratio of square root of the 5 arc min grid cells occupied by *F. alnus* and associated species at a given time step. Occupied 5 arc min grid cells were constrained to be within 30 arc min grid cells occupied by both *F. alnus* and associated species at least once during the study period.

Fig. 5. (A) Square root of the cumulative number of counties occupied through time. Shown here are the linear and polynomial regression lines for models using year as a predictor variable and the square root of the cumulative number of counties occupied as the response variable. Linear regression predictions are shown for both *F. alnus* and the group of associated species (solid black lines). The best-fit polynomial regression fit is shown for each set (cubic for both *F. alnus* and the group of associated species, dot-dash and dash lines, respectively) (B) Ratio of the rate of growth for cumulative occupied counties calculated annually (crosses) and by 10 year moving window average (geometric mean) (triangles). Outlier data not shown (those >1.2 or <0.8, but contribute to the moving window average values. (C) Square root of the ratio of the cumulative number of counties occupied by *F. alnus* to those occupied by the group of associated species.