INVASIVE SPECIES OCCURRENCE FREQUENCY IS NOT A SUITABLE PROXY FOR ABUNDANCE IN THE NORTHEAST

A Thesis Presented

by

TYLER J. CROSS

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
Of the requirements for the degree of

MASTER OF SCIENCE

May 2016

Environmental Conservation
Forest Resources and Arboriculture

INVASIVE SPECIES OCCURRENCE FREQUENCY IS NOT A SUITABLE PROXY FOR ABUNDANCE IN THE NORTHEAST

A Thesis Presented

by

TYLER J. CROSS

Approved as to style and content by:

____________________________________,
Bethany A. Bradley, Chair

____________________________________,
Kristina Stinson, Member

____________________________________,
John T. Finn, Member

____________________________________,
Julie Richburg, Member

____
Curt Griffin, Department Head,
Environmental Conservation

ACKNOWLEDGEMENTS

First, I would like to thank my advisor Bethany Bradley for her help and support. In the same vein, thank you John Finn, Kristina Stinson, Julie Richburg, and Jenica Allen for your guidance, advice, and troubleshooting help.

Additionally, thank you to the contributors to and managers of the databases Eddmaps, ImapInvasives NY, and IPANE. Without your assistance and goodwill none of this would have been achievable.

Finally, this work would not have been possible without the structure and assistance of the University of Massachusetts's Department of Environmental Conservation.

ABSTRACT

INVASIVE SPECIES OCCURRENCE FREQUENCY IS NOT A SUITABLE PROXY FOR ABUNDANCE IN THE NORTHEAST

MAY 2016

TYLER J. CROSS, B.S., UNION COLLEGE

M.S. UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Bethany Bradley

Spatial information about invasive species abundance is critical for estimating impact and understanding risk to ecosystems and economies. Unfortunately, at landscape and regional scales, most distribution datasets provide limited information about abundance. However, national and regional invasive plant occurrence datasets are increasingly available and spatially extensive. We aim to test whether the frequency of these point occurrences can be used as a proxy for abundance of invasive plants. We compiled both occurrence and abundance data for nine regionally important invasive plants in the northeast US using a combination of herbarium records, surveys of expert knowledge, and various invasive species spatial databases. We integrated all available abundance information based on infested area, percent cover, or qualitative descriptions into abundance rankings ranging from 0 (absent) to 4 (highly abundant). Within equal area grid cells of 800 m, we counted numbers of occurrence points and used an ordinal regression to test whether higher numbers of occurrence points were positively correlated with abundance rankings. We compiled a total 49,341 occurrence points in 18,533 cells, of which 12,183 points (25%) within 4,278 cells (32%) had associated abundance information. In six of nine study species we found slight but significant positive overall relationships between abundance rank and occurrence frequency at high abundance ranks. However, at low abundance rankings the relationship tended to be negative and the magnitude of the overall difference in occurrence frequency was too small to be relevant to management. My results suggest that currently available occurrence datasets are unlikely to serve as effective proxies for abundance, and models derived from invasive plant occurrence datasets should not be interpreted as indicative of plant abundance and associated impact. Increased efforts to collect and report invasive species abundance information, and/or higher densities of occurrence points in heavily infested areas are strongly needed for regional scale assessments of potential abundance and associated impact.

Keywords: *Alliaria petiolata, Celastrus orbiculatus, Centaurea stoebe,* Citizen science, *Cynanchum louiseae, Frangula alnus*, *Fallopia japonica*, *Lythrum salicaria, Microstegium vimineum, Persicaria perfoliata,*

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Invasive species abundance is recognized as an important metric of potential impact on an ecosystem (Daehler, 2003; Parker et al., 1999.; Seabloom et al., 2013; Stohlgren and Schnase, 2006). Unfortunately, spatial data available for invasive species, such as museum/ herbarium records and management records, are typically limited to occurrences only. Including all occurrences rather than just abundant infestations in species distribution models leads to vast overestimation of invasion risk (Bradley, 2013), which is less useful for guiding control efforts aimed at reducing ecological and economic impacts (Hulme, 2006; McDonald et al., 2009). Hence, spatial data and associated spatial models of invasive species abundance at landscape and regional scales are strongly needed for coordinating monitoring and management.

Occurrence data alone are not typically effective for predicting abundance. Several studies have tested whether habitat suitability values based on presence-only or presence/absence data can effectively model abundance, with generally poor results (Jiménez-Valverde et al., 2009; Pearce and Ferrier, 2001; Sakai et al., 2001; VanDerWal et al., 2009). For invasive plants, presence-only models were effective for differentiating presence from absence, but could not predict increasing abundance (Pearce and Ferrier, 2001), particularly when herbarium records were the source of data (Bradley, 2016). However, spatial models trained with abundance data perform reasonably well for predicting invasive species abundance (Bradley, 2016; Kulhanek et al., 2011). Thus, in order to effectively predict invasion risk associated with invasive species abundance, better spatial abundance data are needed.

One approach for estimating abundance in the absence of explicit abundance data uses the frequency of occurrence points as a proxy for local abundance. It is a well-accepted pattern in ecology that a species' abundance is positively correlated with the frequency of its occurrence across a region (He et al., 2000; Holt et al., 2002). Collection of occurrence data for occupancy modeling requires repeated observations of experimental plots to measure frequency of species occurrence using consistent levels of search effort (Royle and Dorazio, 2008). The resulting occupancy and absence data are then used to model abundance. While contributed occurrence datasets do not meet these criteria, increasingly widespread and repeated collections by research, monitoring, and management groups could provide sufficient spatial occurrence information to act as a proxy for local abundance.

In the United States, invasive plant occurrence data are available through herbarium records like those contained in the Global Biodiversity Information Facility (GBIF; www.gbif.org) as well as spatial data compilations like the Invasive Plant Atlas of New England (IPANE; Mehrhoff et al. 2003) or the Early Detection & Distribution MAPping System (EDDMAPS; Bargeron 2016). The latter databases contain data compiled from a range of sources, including both citizen scientists and conservation professionals. In some cases, invasive plant abundance data, either qualitative or quantitative, are included along with occurrence locations.

Botanical records like GBIF have long been accepted as an important source of occurrence data for use in species distribution or habitat modelling. For invasion ecology in particular, management and citizen science databases are also increasingly being used to model habitat suitability (Dickinson et al., 2010). While there is some concern that contributed datasets could contain information recorded by under-trained individuals (Crall et al., 2015), recent research has increasingly shown that data from citizen scientists is reliable (Danielsen et al., 2005; Fowler et al., 2013). Data are even more reliable when contributors are trained and/or when data are professionally verified, which is typical of invasive species datasets. Thus, citizen science and management records provide a robust dataset that increases numbers of occurrence records and broadens regional coverage (Delaney et al., 2008; Fore et al., 2001).

Given the importance of abundance information for modeling invasion risk across landscapes and regions (Daehler, 2003; Parker et al., 1999; Seabloom et al., 2013; Stohlgren and Schnase, 2006), I aim to test whether the spatial frequency of point occurrences can be used as an effective proxy for invasive plant abundance. Here, I compiled a comprehensive database of occurrence and abundance data for nine problematic invasive plant species across the Northeast US. I hypothesize that the number of occurrences within equal area grid cells will be positively related to invasive plant abundance. This analysis provides a first empirical test of the relationship between the frequency of invasive plant occurrences and local plant abundance.

CHAPTER 2

METHODS

Study Species and Area

I selected nine non-native, invasive plants that are of concern to regional managers and have been identified as having negative environmental impacts in the Northeast US (**Table 1**). These species were chosen to encompass a range of current distributions from widespread to emerging invaders within the study area, which included thirteen states and the District of Columbia between Virginia and Maine (**Figure 1**).

Compilation of Existing Data Sources

I compiled existing distribution and abundance data from four online databases that record geolocations of invasive species: the Global Biodiversity Information Facility (GBIF; www.gbif.org), the Invasive Plant Atlas of New England (IPANE; Mehrhoff et al. 2003), the Early Detection & Distribution MAPping System (EDDMAPS; Bargeron 2016), and iMAP Invasives for the state of New York (http://www.nyimapinvasives.org). Additionally, I collected and included invasive species occurrence and abundance information for selected species from several smaller databases compiled by researchers or managers in New England.

Ultimately, the included data ranged from botanical records collected by professional scientists, to citizen-science efforts in which interested individuals collect and enter occurrences of invasive species into online repositories. All data included geographic location, with a subset also containing abundance information reported in a variety of formats. For databases containing polygons rather than points, it was assumed that the polygon extents described the area of the invasive plant infestation. I removed duplicate points, as well as points that fell outside of the study area and points where the spatial precision was lower than one-thousandth of a decimal degree (equivalent to less than 100 m throughout the study area). I also excluded points located at town or country centroids, which were assumed to be problematic due to poor locational accuracy.

For all geographic locations also containing abundance information, I standardized abundance to a qualitative, ranked scale of 0-4 (**Table 2**) ranging from absent to highly abundant. Bins for quantitative cover estimates were arbitrary, but consistent with previous rankings of relative invader abundance and importance (Rouget et al., 2003)**.** Bins for quantitative range extent estimates were based on commonly reported metrics of area (square meters for small areas, acres for larger areas). The break between ranks 3-4 (moderate vs. high abundance) of 40 acres for range extent was chosen to match the break in cover estimate of 25% of a grid cell (see below); a grid cell was approximately 160 acres.

Point Count vs Abundance Comparison

In order to calculate frequency of occurrences, I used the fishnet tool in ArcGIS 10.2 to create an equal area grid of 800 m^2 grid cells covering the study area. This spatial resolution approximates 30 arc seconds, which is a typical gridded resolution for species distribution modeling at regional to continental scales. Polygon layers were transformed into point occurrences with one point per grid cell. I then summed the number of point occurrences within each grid cell. All spatial analyses were performed using ArcGIS 10.2.

In order to calculate abundance for each grid cell, I extracted the maximum abundance ranking associated with all points falling within each grid cell. Maximum abundance ranking, as opposed to the mean abundance ranking, is a better indication of the severity of the infestation. For polygon data, the area of the polygon overlapping each grid cell was calculated and grid cells were ranked according to the area category in **Table 2**.

In order to test whether frequency of point occurrences was related to abundance, I compared ranked abundance estimates at the 800 m^2 grid cell resolution to the number of points falling within each grid cell. Grid cells only contained abundance information for a given species if one or more occurrences within the cell had associated abundance, or if polygon features identifying the extent of an invasion overlapped the grid cell. As a result, only locations containing both abundance and point occurrence information were tested.

Inter-Database Comparisons

I collected data from four major online databases, as well as a few smaller collections provided by invasive species managers or conservation groups. The online databases contained data from a variety of sources, ranging from volunteer groups to state and federal agencies, and used different collection methods. For example, IPANE trained observers to collect data with a specific methodology (Mehrhoff et al., 2003), while EDDMAPS archived occurrence information contributed by any user that passed its veracity testing (Bargeron, 2016). In this analysis, all data sources were treated equally. However, I compared the overall data availability and extents of the datasets for each species. Area extents were calculated using a convex hull, which bounds all occurrence data. Convex hulls were clipped to include only the study area.

Statistical Analysis

Ordinal regression analysis was used to test the hypothesis that frequency of occurrence points was positively related to abundance. I used ordinal regressions because the abundance classification bins were not equally spaced but were increasing in rank order. Ordinal regression tests for an overall linear response across all ranks, as well as directionality between binned ranks based on proportional odds ratios. Ordinal regression was performed using the proportional odds logistic regression function in the MASS package in R (version 2.15.2). For visualization I also created box-and-whisker plots for each study species showing abundance ranking vs occurrence frequency.

The occurrence frequency data were skewed towards low values, mainly ones. To test for whether the abundance of ones biased the results, I repeated the analysis for the subset of grid cells with occurrence point frequency greater than one.

CHAPTER 3

RESULTS

Point Count vs Abundance Comparison

Numbers of occurrences ranged from *C. stoebe* with 516 points (52% with abundance information) spread across 408 cells to *C. orbiculatus*, with 9,104 points (31% with abundance information) spread across 2883 cells. After removing duplicate and non-useable points from the online databases a total of 49,341 occurrence points remained within 18,533 grid cells (**Table 3**). A quarter (25%) of all points collected had associated abundance information related to at least one abundance metric. In most species studied, grid cells with and without abundance information contained similar frequencies of occurrence points, with the mean frequency of occurrences being 2.65 points per grid cell (**Figure 2**).

Ordinal regression revealed significant positive overall relationships between frequency of occurrences and ranked abundance for six of the nine study species **(Table 5)**. However, this relationship was not consistent across abundance rankings. For most species with significant relationships, there was a negative relationship at the lowest rank (rare, rank 1), where grid cells in which the species was rare had slightly more occurrences than higher rankings. The overall positive relationship was driven by a similar slightly larger number of occurrences at the highest rank (many large infestations, rank 4) **(Table 5)**. When grid cells with only one occurrence point were removed from the analysis, similar results were found **(Appendix A)**. When only grid cells with greater than one occurrence point were considered in the dataset I found a similar weak positive relationship between abundance class and frequency, although not in all the species the relationship was observed in when the full dataset was considered **(Table 5)**. As in the full dataset analysis, the weak positive relationship that between abundance class and frequency was joined by a set of stronger trends, indicating that the relationship between abundance class and frequency was negative at low abundance ranks and positive at high abundance ranks **(Appendix B)**.

Inter-Database Comparisons

The plurality of my occurrence points came from the Early Detection and Distribution MAPping System (EDDMAPS), which contributed 23,167 data points. Followed by ImapInvasives (19,698 data points), IPANE (1,836 data points), and GBIF (1572 data points) **(Table 3)**. Only EDDMaps and GBIF targeted the full study region, the other databases were regionally scaled, focusing on smaller sections of the Northeast. Comparing the convex hull areas encompassing occurrence location data showed that the citizen science database EDDMAPS had on average 26% broader range coverage than the botanical records available in GBIF **(Table 6).**

CHAPTER 4

DISCUSSION

Ordinal regression analysis revealed a weak positive relationship in six of nine species **(Table 5)**. While this does support the hypothesis that the frequency of occurrence points in a grid cell would increase with the cell's abundance rank, the weak overall trend is overshadowed by the existence of a strong negative relationship between abundance class and frequency at low abundance classes. Additionally, in three of nine species, there were no significant relationship between frequency and abundance class and frequency at any rank. This, taken with the fact that the median number of occurrences in all species ranged from only one to two, suggests that even statistically significant differences are unlikely to translate into reliable means of distinguishing abundant infestations from occurrences **(Figure 3)**.

With sufficient, regular sampling, it has often been found that numbers of species occurrences are positively related to species abundance (He et al., 2002). Species with larger populations are more likely to be observed, and, thus, higher rates of observation indicate more individuals are occupying the habitat. (Royle and Dorazio, 2008). However, effective occupancy modeling requires consistent, widespread sampling and resampling to measure occurrences (He et al., 2002). Although the invasive plant database I compiled contains substantial, widespread occurrences recorded by citizen scientists, managers, and museum collectors **(Figure 4)**, it does not fulfill these requirements. It is likely that spatial biases in the invasive plant database limit the potential of any attempt to use occurrence data to predict invasive species abundance.

Previous research comparing abundance and occurrence frequency of invasive species has shown no relationship or even a negative relationship, with more occurrences in areas of low abundance (Marvin et al., 2009). Marvin et al. (2009) hypothesized that point data collected by invasive plant managers tended to focus on early detection and rapid response (EDRR), which targets small nascent infestations (Moody and Mack, 1988). As a result, EDRR data collection efforts might, counterintuitively, tend to have more occurrences in areas of low abundance. My results do show evidence of this negative relationship at low abundance ranks; however the directionality of the effect becomes positive as abundance class increases **(Table 5)**. It is possible that the wide range of data sources I used includes both the effect shown in Marvin et al., 2009 as well as clustered occurrences in areas of high abundance. Increasing the numbers of point occurrences recorded and reported in areas with abundant infestations would help to increase the strength of the relationship observed with existing data.

Each database used within this study had a different methodology for collecting, archiving, and validating the authenticity of the data it contained. For instance, IPANE relied on trained observers to collect and input data into the database, while EDDMAPS took occurrence information from a broader group of individuals, with less top-down management or control, and then used a system of affiliated experts and filters to validate the records. Individuals who contributed to the different databases differed widely in the purposes of their individual data collection efforts. Not all groups collecting invasive species population data were focused on early detection and rapid response efforts. Notably, IPANE volunteers were specifically instructed to do complete censuses of their survey areas, rather than focusing on small populations. It is possible that these different collection methodologies led to the duality of the trends I observed through the ordinal regression analysis. Moreover, data coverage and availability varied widely between states (Figure 4). Location specific services like IPANE and Imap Invasives New York, which serve New England and New York State respectively, influenced my ability to compile data across the entirety of the study area.

While there remains a stigma attached to data collected by citizen scientists, there is considerable evidence that it has the potential to be just as accurate as data collected by professionals (Crall et al., 2015; Fitzpatrick et al., 2009; Fore et al., 2001). The inclusion of data products created by citizen scientists ensured that I had data that adequately covered the study area, which would not have been possible had I relied solely on professionally collected herbarium records like those retrieved from the Global Biodiversity Information Facility (GBIF; **Table 6**). Additionally, while the regional data from GBIF was spatially extensive (**Table 6**), its lack of abundance information limits its usefulness for predicting abundant invasions (Bradley, 2016). Contributed management and citizen science datasets are needed to measure and model patterns of regional scale invasive plant abundance **(Table 3; Figure 5)**.

Spatial models depicting invasive species populations have been shown to be more accurate when created using abundance data rather than simple occurrence information (Bradley, 2016; Kulhanek et al., 2011). As habitat suitability models are useful tools for regional and landscape scale invasive species management, there is a clear need for data collectors, professionals and citizen scientists alike, to continue to collect and report abundance information. The contributed databases compiled here report a considerable amount of abundance information **(Table 3)**. As yet, these abundance data are underutilized in modeling efforts, and biogeographers should consider including this important source of data in regional models.

While existing abundance information is an excellent start, data collectors could consider modifications of collection methods to better inform spatial models. For example, abundance data were most often reported as a single metric, either by quantitatively or qualitatively describing cover, or by quantitatively or qualitatively describing extent. The combination of cover and extent information is much more informative for estimating the magnitude of an invasion, and I recommend that collectors consider reporting both pieces of information. While not everyone collecting data on species populations is equipped to make quantitative measurements, I found that qualitative estimates using consistent rankings (e.g., **Table 2**) were very useful for interpreting relative abundance. I urge scientists, managers and citizens collecting invasive plant occurrence data to include extent and cover information when archiving their data to online repositories.

CHAPTER 5

CONCLUSIONS

Overall, I found significant positive relationships between occurrence frequency and abundance in six of nine target invasive plants. However, the magnitude of the relationship was quite small; making it unlikely that frequency of occurrence could be used as an effective proxy for abundance for risk assessments and management planning. Additionally, there was often a slight negative relationship at low abundance, potentially because of the influence of EDRR efforts collecting frequent occurrences in low-abundance locations.

Given the importance of abundance for understanding invasion risk, additional recording and reporting of abundance is needed. A large proportion (25%) of the compiled occurrences contained some abundance information, but this could be improved if both cover and extent were reported.

Table 1: List of target invasive plants

	Name	Growth Habit
1.	*Alliaria petiolata* (Garlic Mustard)	Forb/Herb
2.	*Celastrus orbiculatus* (Oriental Bittersweet)	Vine
3.	*Centaurea stoebe* (Spotted Knapweed)	Forb/herb
4.	*Frangula alnus* (Glossy Buckthorn)	Tree, shrub
5.	*Fallopia japonica* (Japanese Knotweed)	Forb/herb Subshrub
6.	*Lythrum salicaria* (Purple Loosestrife)	Subshrub, forb/herb
7.	*Microstegium vimineum* (Japanese Stilt-Grass)	Graminoid
8.	*Persicaria perfoliata* (Mile-A-Minute Vine)	Forb/herb, Vine
9.	*Cynanchum louiseae* (Black Swallow-wort)	Vine, forb/herb

Table 2: Classification scheme used to combine quantitative and qualitative abundance estimates into abundance rankings.

Abundance Ranking	**Quantitative Cover Estimate**	**Quantitative Range Extents**	**Qualitative Cover Estimate**	**Qualitative Range Extents**
0	0 / Not Present	0	Absent	Absent / Not Present
1	≤1%	≤1m^2	Trace or Single Plant	Rare
2	1-5%	1m^2 - 1 acre	Low or Scattered Plants	Small Patches
3	5-25%	1-40 acres	Moderate or Scattered Dense Patches	Several Small Patches
4	>25%	40+ acres	High or Dense Monoculture	Many Small or Several Large Patches

Table 3: Total numbers of occurrence and abundance points compiled for each of the primary data sources

	Total Points	Abundance Points	% of Points Abundance	Total Cells	Abundance Cells	% of Cells Abundance
EDDMAPS	22888	9316	40%	5904	3108	53%
GBIF	1572	0	0%	1572	0	0%
Imap Invasives	22096	710	3%	5852	371	6%
IPANE	1955	1718	88%	1229	1106	90%
Other	830	439	53%	531	227	43%
All Databases	49341	12246	25%	13486*	4278*	32%

*Because cells can be occupied by points from more than one database, this number is not a simple sum of each databases total

Table 4: Total numbers of occurrence and abundance points compiled for each target species

Species	**Points with Abundance**	**Total Points**	**Percent of Points with Abundance**	**Cells with Abundance**	**Total Cells**	**Percent of Cells with Abundance**
Alliaria petiolata	1849	7610	24%	941	3480	27%
Celastrus orbiculatus	2845	9104	31%	1368	2883	47%
Centaurea stoebe	267	516	52%	229	408	56%
Fallopia japonica	2313	9482	24%	1291	4155	31%
Frangula alnus	1699	4412	39%	665	1078	61%
Microstegium vimineum	1520	7942	19%	522	1883	27%
Lythrum salicaria	1183	7539	16%	619	3503	18%
Persicaria perfoliata	290	1386	21%	164	578	28%
Cynanchum louiseae	280	1350	21%	161	565	29%
All	12246	49341	25%	5960	18533	32%

Table 5: Significant trends between occurrence frequency and abundance class. Significant ($p<.05$) relationships between occurrence frequency and abundance rank discovered through ordinal regressions were noted here. Abundance rankings were grouped and compared by rank-order (rank 1 was compared to 2, 3, and 4, etc). Significant differences are labeled with the direction of the relationship, non-significant comparisons were labeled NS.

All Data	Divisions of Abundance			
	Overall Coefficient	**1 - 2,3,4**	**1,2 - 3,4**	**1,2,3 - 4**
Alliaria petiolata	Positive	Negative	NS	Positive
Celastrus orbiculatus	Positive	Negative	NS	Positive
Centaurea stoebe	Positive	Negative	NS	Positive
Fallopia japonica	Positive	Negative	Positive	Positive
Microstegium vimineum	Positive	Negative	Negative	NS
Lythrum salicaria	Positive	NS	Positive	Positive
Frangula alnus	NS	NS	NS	NS
Persicaria perfoliata	NS	NS	NS	NS
Cynanchum louiseae	NS	NS	NS	NS

*NS indicates that the relationship is non-significant ($p>.05$)

Table 6: Spatial extents (in km^2) within the study area encompassed by each dataset and species based on a convex hull around all occurrence points

	Alliaria petiolata	*Celastrus orbiculatus*	*Centaurea stoebe*	*Frangula alnus*	*Fallopia japonica*	*Microstegium vimineum*	*Lythrum salicaria*	*Persicaria perfoliata*	*Cynanchum louiseae*	All species
EDDMAPS	461,290	491,214	441,627	277,685	541,226	374,242	513,057	313,746	306,613	413,411
IPANE	96,323	365,187	19,017		188,413		112,120			156,212
Imap	163,164	152,213		129,542	168,651	84,859	154,595	387,589	306,613	149,799
GBIF	417,449	275,807	362,130	289,958	461,755		391,499	175,419	236,961	326,372

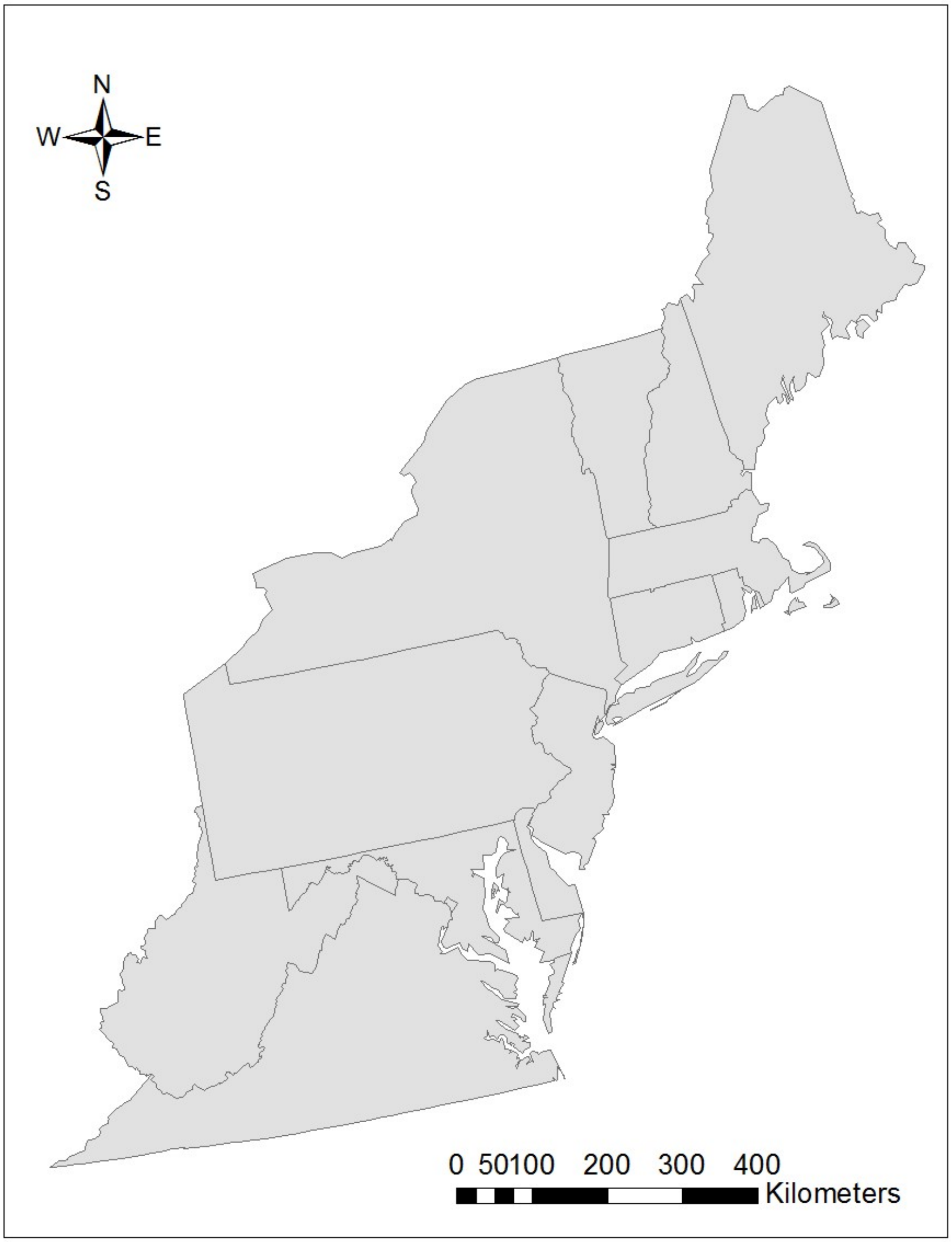


Figure 1: Distribution and abundance data were collected for thirteen states and the District of Columbia between Maine and Virginia

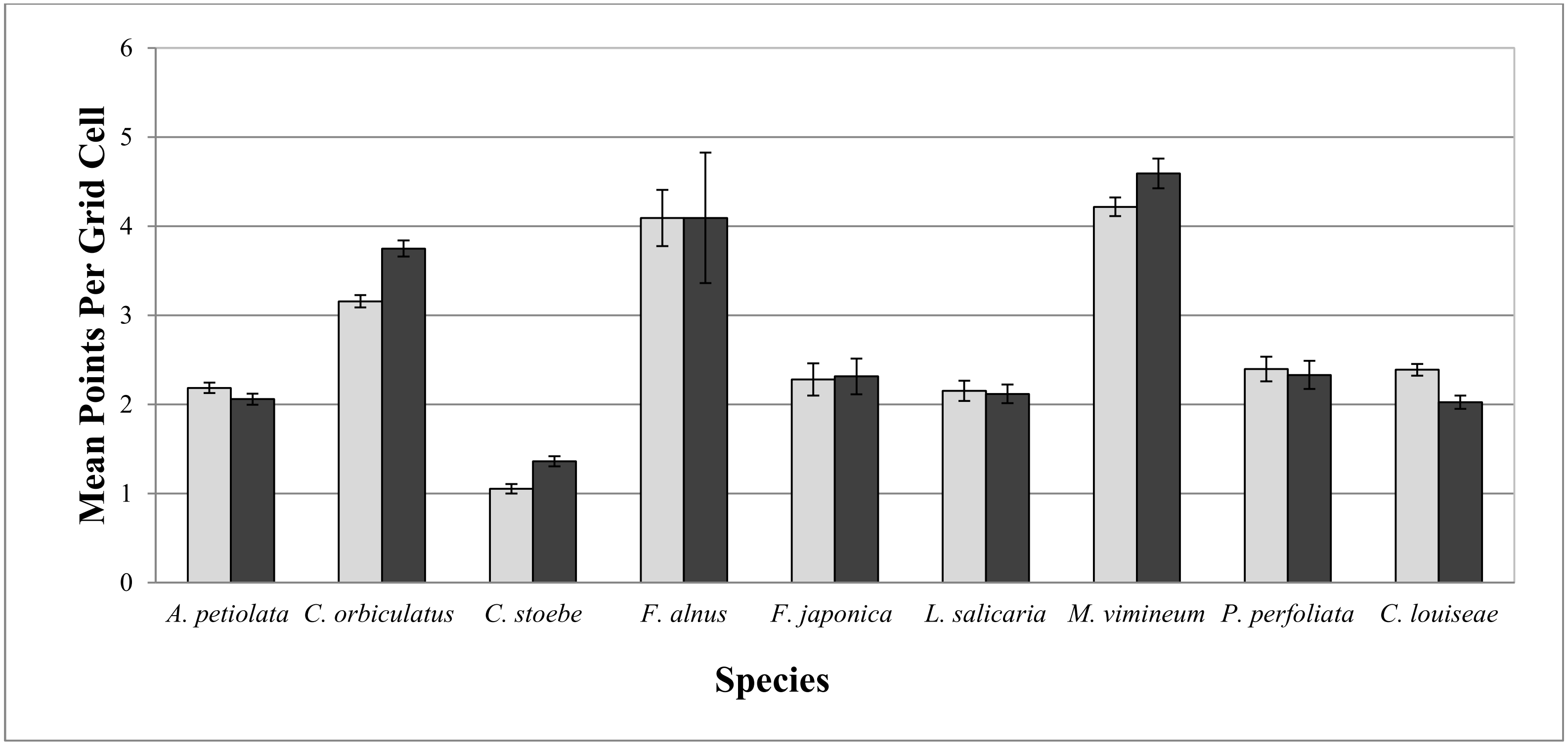


Figure 2: The mean number of points per grid cell in cells where there was abundance information compared to those without abundance information. Grid cells with associated abundance information (black bars) had similar numbers of occurrence points as grid cells that did not contain abundance information (gray bars) for most species. Bars show mean number of occurrence points plus or minus standard error.

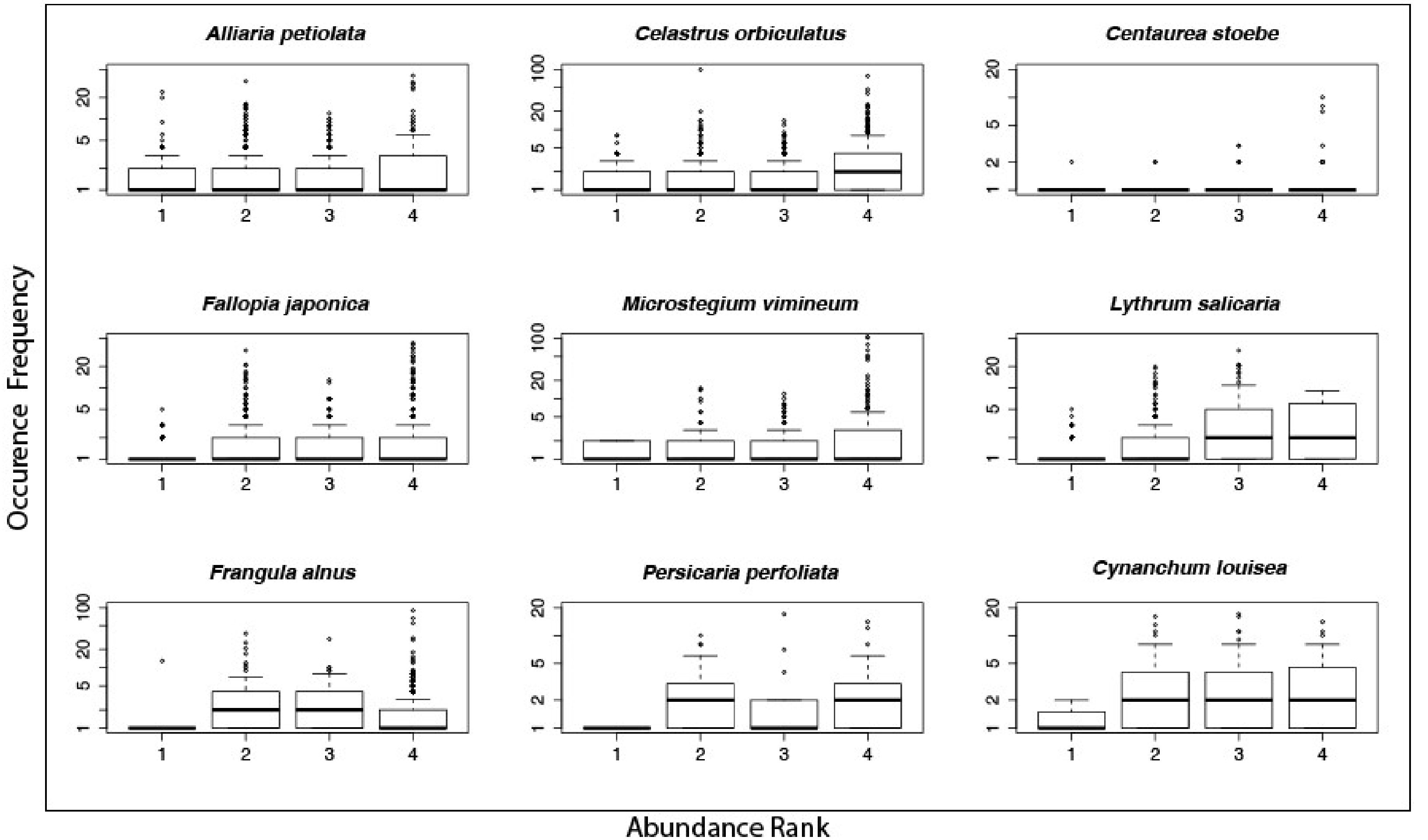


Figure 3: Box plot of abundance ranking vs. the frequency of occurrences within grid cells. The first six plots, *Alliaria petiolata* through *Lythrum salicaria* represent species in which statistically significant trends were shown to exist through the use of ordinal regressions. In all six significant cases there was both a negative relationship between frequency and abundance when abundance values were low and a positive relationship overall.

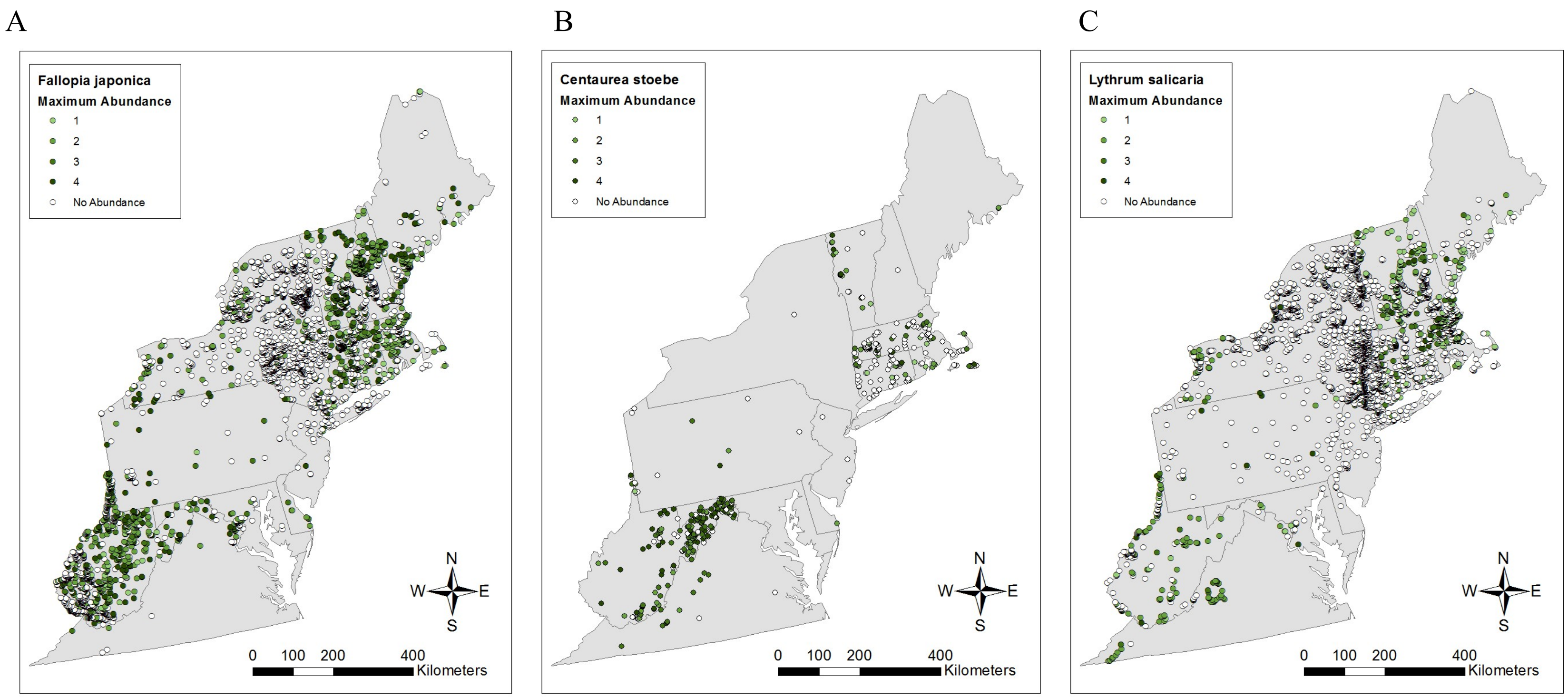


Figure 4: Abundance maps for three example species A) *Fallopia japonica, B) Centaurea stoebe,* and C) *Lythrum salicaria* show that occurrence and abundance data are widespread across the Northeast but data from several mid-Atlantic states were less well reported in the databases included in this analysis. Graduated colors correspond to abundance rankings in Table 2. Maps for the remainder of species examined are located in **Appendix B**

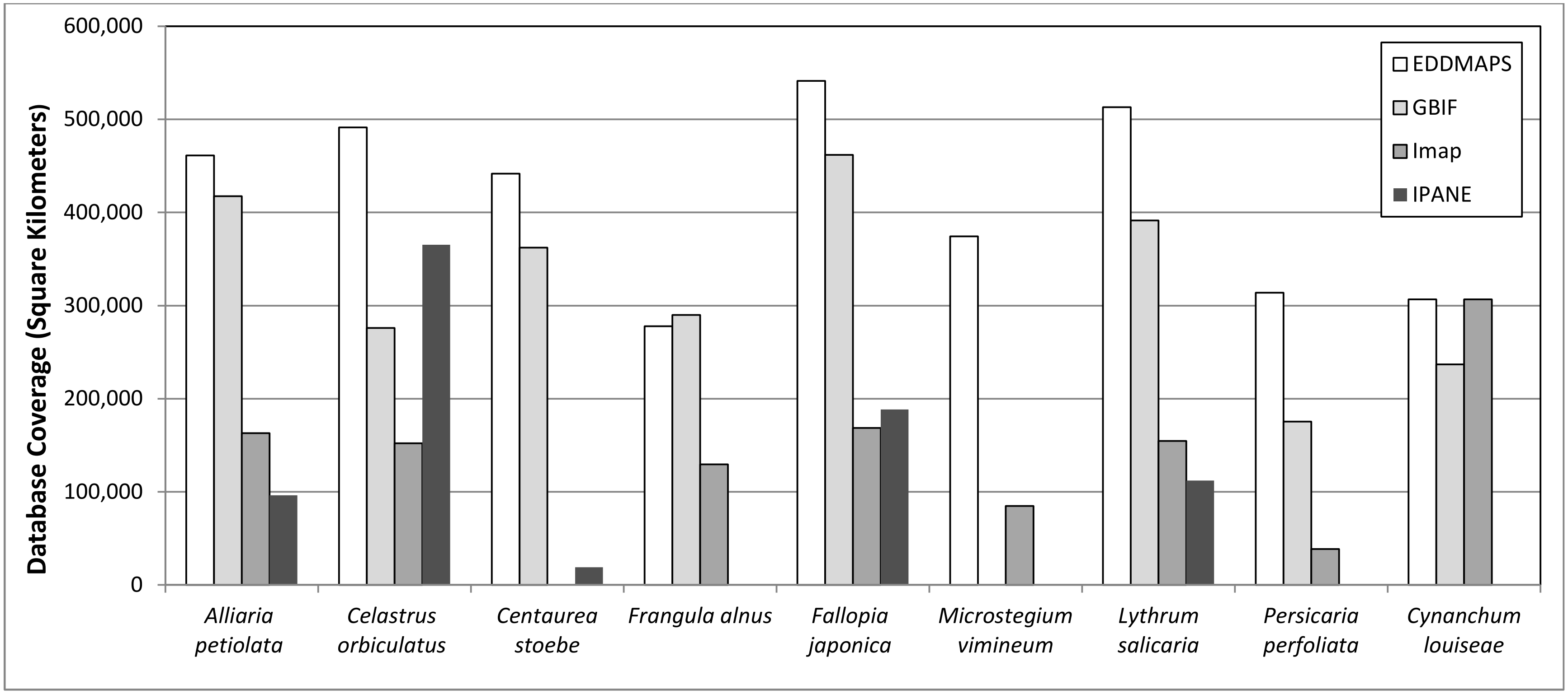
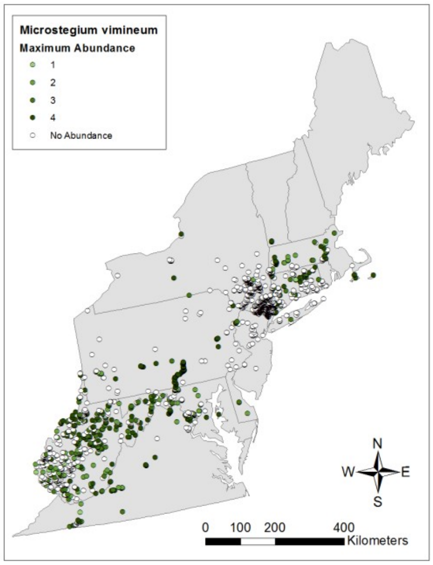
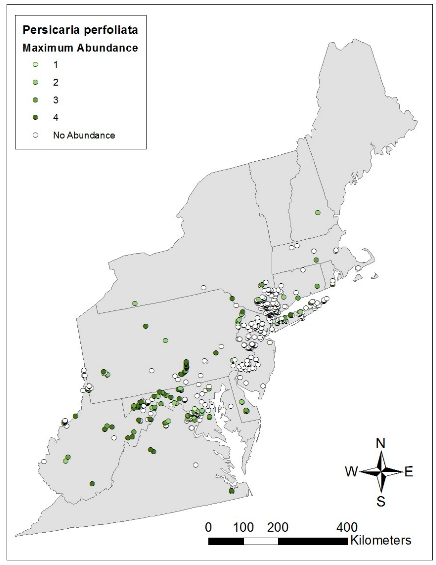


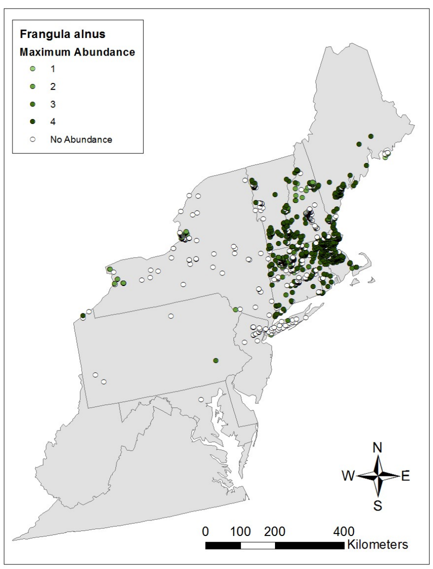
Figure 5: Differing coverage areas associated with different databases, measured in square kilometers. The Eddmaps database covered a larger area than all other databases in eight of nine species. While the Invasive Plant Atlas of New England (IPANE) database was largely constricted to the region from which it took its name, but still managed to cover a greater amount of space on average than the New York focused Imap Invasives database

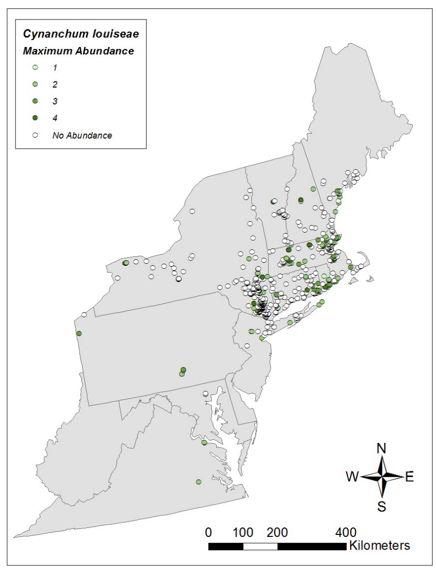
APPENDIX A

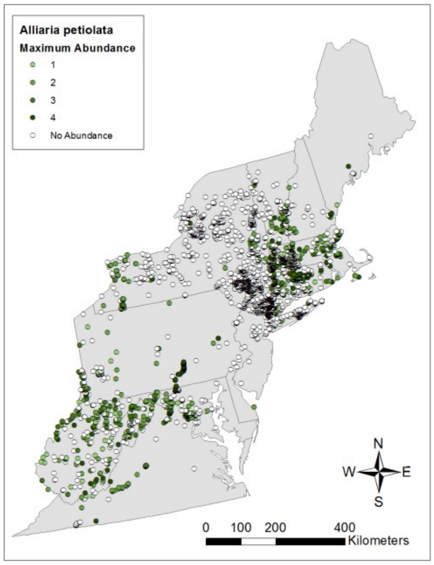
ABUNDANCE MAPS FOR ALL NINE SPECIES

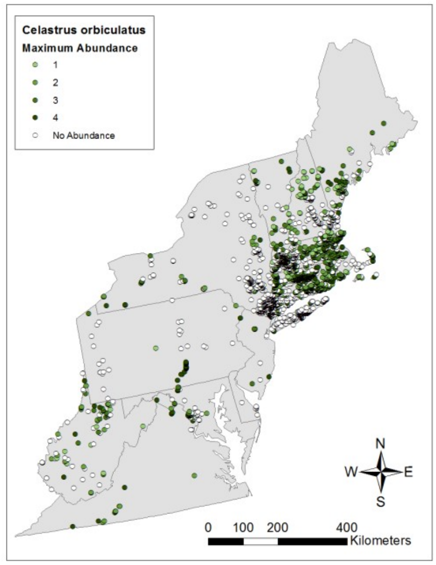









Alliaria petiolata
Maximum Abundance
1
2
3
4
No Abundance
N
W
E
S
0 100 200 400
Kilometers


Celastrus orbiculatus
Maximum Abundance
1
2
3
4
No Abundance
N
W
E
S
0 100 200 400
Kilometers

APPENDIX B:

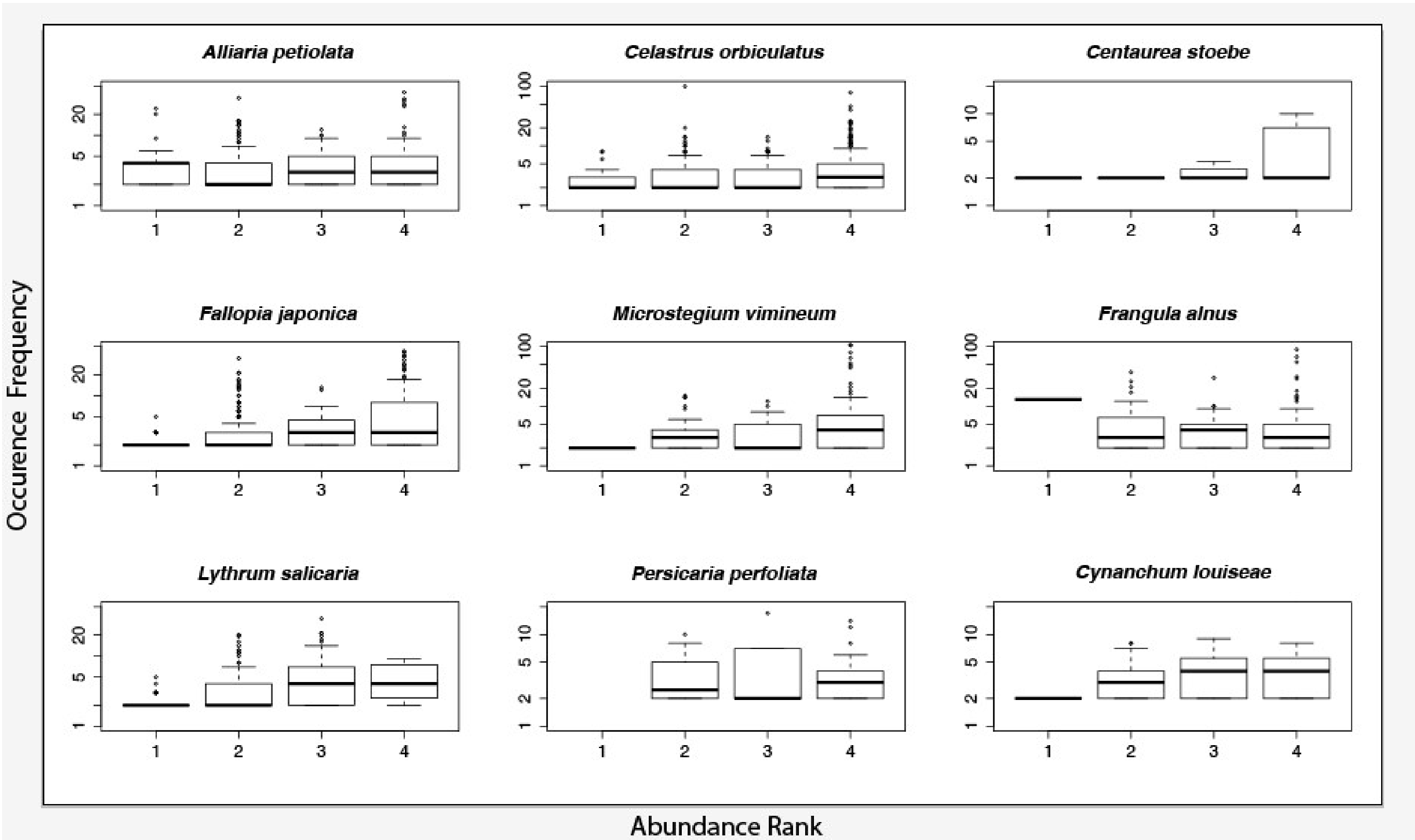
SIGNIFICANT TRENDS BETWEEN OCCURRENCE FREQUENCY AND ABUNDANCE CLASS

Freq >1	Frequency by Abundance Class			
	Overall Coefficient	1 - 2,3,4	1,2 - 3,4	1,2,3 - 4
Alliaria petiolata	Positive	Negative	Negative	NS
Celastrus orbiculatus	Positive	Negative	Negative	Positive
Centaurea stoebe	Positive	Negative	Positive	Positive
Fallopia japonica	Positive	Negative	Positive	Positive
Microstegium vimineum	Postive	Negative	Negative	NS
Lythrum salicaria	NS	NS	NS	NS
Frangula alnus	NS	NS	NS	NS
Persicaria perfoliata	NS	NS	NS	NS
Cynanchum louiseae	NS	NS	NS	NS

* NS indicates that the relationship is non-significant (p>.05)

APPENDIX C

BOX AND WHISKER PLOTS EXAMINING THE RELATIONSHIP BETWEEN FREQUENCY AND ABUNDANCE RANK WHEN GRID CELLS CONTAINING ONLY ONE PRESENCE POINT WERE DISCOUNTED



BIBLIOGRAPHY

Royle, JA., Dorazio, R., 2008. Hierarchical Modeling and Inference in Ecology: The analysis of data from populations, metapopulations, and communities. Academic Press.

Bargeron, C.T.; Swearingen, J.; Douce, G.K.; Moorhead, D.J.; Rawlins, K.A. 2011. EDDMAPS: a collaborative, easy-to-use, early detection and distribution mapping system.

Bradley, B.A., 2016. Predicting abundance with presence-only models. Landscape Ecology 31, 19–30. doi:10.1007/s10980-015-0303-4

Bradley, B.A., 2012. Distribution models of invasive plants over-estimate potential impact. Biological Invasions 15, 1417–1429. doi:10.1007/s10530-012-0380-0

Crall, A.W., Jarnevich, C.S., Young, N.E., Panke, B.J., Renz, M., Stohlgren, T.J., 2015. Citizen science contributes to our knowledge of invasive plant species distributions. Biol.Invasions 17, 2415–2427. doi:10.1007/s10530-015-0885-4

Daehler, C., 2003. Performance comparisons of co-occurring native and alien invasive plants: Implications for conservation and restoration. Annu.Rev.Ecol.Evol.Syst. 34, 183–211. doi:10.1146/annurev.ecolsys.34.011802.132403

Danielsen, F., Burgess, N.D., Balmford, A., n.d. Monitoring Matters: Examining the Potential of Locally-based Approaches. Biodiversity & Conservation 14, 2507–2542. doi:10.1007/s10531-005-8375-0

Delaney, D.G., Sperling, C.D., Adams, C.S., Leung, B., 2008. Marine invasive species: validation of citizen science and implications for national monitoring networks. Biol.Invasions 10, 117–128. doi:10.1007/s10530-007-9114-0

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J., Phillips, T., Purcell, K., 2012. The current state of citizen science as a tool for ecological research and public engagement. Front.Ecol.Environ. 10, 291–297. doi:10.1890/110236

Dickinson, J.L., Zuckerberg, B., Bonter, D.N., 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. Annu.Rev.Ecol.Evol.Syst. 41, 149–172. doi:10.1146/annurev-ecolsys-102209-144636

Fitzpatrick, M.C., Preisser, E.L., Ellison, A.M., Elkinton, J.S., 2009. Observer bias and the detection of low-density populations. Ecol.Appl. 19, 1673–1679. doi:10.1890/09-0265.1

Fore, L.S., Paulsen, K., O'Laughlin, K., 2001. Assessing the performance of volunteers in monitoring streams. Freshwater Biology 46, 109–123. doi:10.1111/j.1365-2427.2001.00640.x

Fowler, A., Whyatt, J.D., Davies, G., Ellis, R., 2013. How Reliable are Citizen-Derived Scientific Data? Assessing the Quality of Contrail Observations Made by the General Public. Trans.GIS 17, 488–506. doi:10.1111/tgis.12034

GBIF Secretariat: GBIF Backbone Taxonomy, 2013-07-01. Accessed via http://www.gbif.org/species/5376075 on 2015-08-01

He, F., Gaston, K.J., Associate Editor: Lenore Fahrig, 2000. Estimating Species Abundance from Occurrence. The American Naturalist 156, 553–559. doi:10.1086/303403

He, F., GASTON, K.J., WU, J., 2002. On species occupancy-abundance models. Écoscience 9, 119–126.

Holt, A.R., Gaston, K.J., He, F., 2002. Occupancy-abundance relationships and spatial distribution: A review. Basic and Applied Ecology 3, 1–13. doi:10.1078/1439-1791-00083

Hulme, P.E., 2006. Beyond control: wider implications for the management of biological invasions. J.Appl.Ecol. 43, 835–847. doi:10.1111/j.1365-2664.2006.01227.x

iMapInvasives: An online data system supporting strategic invasive species management. © 2016, NatureServe. Available at http://www.imapinvasives.org. (Date accessed: Oct 01, 2015.)

Jiménez-Valverde, A., Diniz, F., Azevedo, E.B. de, Borges, P.A.V., 2009. Species Distribution Models Do Not Account for Abundance: The Case of Arthropods on Terceira Island. Annales Zoologici Fennici 46, 451–464. doi:10.5735/086.046.0606

Kulhanek, S.A., Leung, B., Ricciardi, A., 2011. Using ecological niche models to predict the abundance and impact of invasive species: application to the common carp. Ecological Applications 21, 203–213. doi:10.1890/09-1639.1

Marvin, D.C., Bradley, B.A., Wilcove, D.S., 2009. A Novel, Web-based, Ecosystem Mapping Tool Using Expert Opinion. Nat.Areas J. 29, 281–292. doi:10.3375/043.029.0307

McDonald, A., Riha, S., DiTommaso, A., DeGaetano, A., 2009. Climate change and the geography of weed damage: Analysis of U.S. maize systems suggests the potential for significant range transformations. Agriculture, Ecosystems & Environment 130, 131–140. doi:10.1016/j.agee.2008.12.007

Moody, M.E., Mack, R.N., 1988. Controlling the Spread of Plant Invasions: The Importance of Nascent Foci. Journal of Applied Ecology 25, 1009–1021. doi:10.2307/2403762

Parker, I.M., Simberloff, D., Lonsdale, W.M., Goodell, K., Wonham, M., Kareiva, P.M., Williamson, M.H., Von Holle, B., Moyle, P.B., Byers, J.E., Goldwasser, L., 1999. Impact: Toward a Framework for Understanding the Ecological Effects of Invaders. Biological Invasions 1, 3–19. doi:10.1023/A:1010034312781

Pearce, J., Ferrier, S., 2001. The practical value of modelling relative abundance of species for regional conservation planning: a case study. Biological Conservation 98, 33–43. doi:10.1016/S0006-3207(00)00139-7

Rouget, M., Richardson, D.M., Associate Editor: Scott D. Wilson, 2003. Inferring Process from Pattern in Plant Invasions: A Semimechanistic Model Incorporating Propagule Pressure and Environmental Factors. The American Naturalist 162, 713–724. doi:10.1086/379204

Sakai, A., Allendorf, F., Holt, J., Lodge, D., Molofsky, J., With, K., Baughman, S., Cabin, R., Cohen, J., Ellstrand, N., McCauley, D., O'Neil, P., Parker, I., Thompson, J., Weller, S., 2001. The population biology of invasive species. Annu.Rev.Ecol.Syst. 32, 305–332. doi:10.1146/annurev.ecolsys.32.081501.114037

Seabloom, E.W., Borer, E.T., Buckley, Y., Cleland, E.E., Davies, K., Firn, J., Harpole, W.S., Hautier, Y., Lind, E., MacDougall, A., Orrock, J.L., Prober, S.M., Adler, P., Alberti, J., Michael Anderson, T., Bakker, J.D., Biederman, L.A., Blumenthal, D., Brown, C.S., Brudvig, L.A., Caldeira, M., Chu, C., Crawley, M.J., Daleo, P., Damschen, E.I., D'Antonio, C.M., DeCrappeo, N.M., Dickman, C.R., Du, G., Fay, P.A., Frater, P., Gruner, D.S., Hagenah, N., Hector, A., Helm, A., Hillebrand, H., Hofmockel, K.S., Humphries, H.C., Iribarne, O., Jin, V.L., Kay, A., Kirkman, K.P., Klein, J.A., Knops, J.M.H., La Pierre, K.J., Ladwig, L.M., Lambrinos, J.G., Leakey, A.D.B., Li, Q., Li, W., McCulley, R., Melbourne, B., Mitchell, C.E., Moore, J.L., Morgan, J., Mortensen, B., O'Halloran, L.R., Pärtel, M., Pascual, J., Pyke, D.A., Risch, A.C., Salguero-Gómez, R., Sankaran, M., Schuetz, M., Simonsen, A., Smith, M., Stevens, C., Sullivan, L., Wardle, G.M., Wolkovich, E.M., Wragg, P.D., Wright, J., Yang, L., 2013. Predicting invasion in grassland ecosystems: is exotic dominance the real embarrassment of richness? Glob Change Biol 19, 3677–3687. doi:10.1111/gcb.12370

Stohlgren, T., Schnase, J., 2006. Risk analysis for biological hazards: What we need to know about invasive species. Risk Anal. 26, 163–173. doi:10.1111/j.1539-6924.2006.00707.x

VanDerWal, J., Shoo, L.P., Johnson, C.N., Williams, S.E., Associate Editor: Wolf M. Mooij, Editor: Donald L. DeAngelis, 2009. Abundance and the Environmental Niche: Environmental Suitability Estimated from Niche Models Predicts the Upper Limit of Local Abundance. The American Naturalist 174, 282–291. doi:10.1086/600087