Land-Use Proxies for Aquatic Species Introductions in the Laurentian Great Lakes

A THESIS
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Elon Martin O'Malia

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Dr. Joel C. Hoffman, Advisor

May 2015

Acknowledgements

Thank you to Rochelle Sturtevant and Matt Cannister for providing me with thousands of nonindigenous aquatic species sightings records from GLANSIS and the USGS NAS databases. Thank you to Valerie Brady and Lucinda Johnson for supplying nonindigenous aquatic invertebrate sightings from the Great Lakes Indicator Consortium's Coastal Wetlands Monitor project. Thank you David Allan, Sigrid Smith, Christine Joseph, and Sarah Bailey for providing marina size and ballast water discharge data from the Great Lakes Environmental Assessment and Mapping (GLEAM) project. I would also like to acknowledge the Integrated Biosciences program at the University of Minnesota and the University of Minnesota Duluth – U.S. EPA Cooperative Training Partnership for both funding my graduate education and also putting me into direct contact with working research scientists that were not affiliated with the University. Additionally, thank you to Donn Branstrator and Lucinda Johnson for serving on my committee, and of course many thanks to my advisor Joel Hoffman.

Dedication

This thesis is dedicated to my wife.

Abstract

Many nonindigenous aquatic species (NAS) adversely impact ecosystems, human health, and the economy of the Laurentian Great Lakes region. Targeted prevention and eradication efforts in response to early detection of NAS can be both cost advantageous and effective at preventing further spread or establishment. To help inform the process of priority site selection for early detection monitoring, I developed and evaluated land-use metrics of three prominent anthropogenic introduction pathways (commercial maritime traffic, recreational maritime traffic, and live release from urban areas). Logistic and linear regression analyses were conducted between species presence or species richness and introduction pathway intensity for 23 NAS over a five-decade period (1970 – 2013) to explain the apparent spatio-temporal patterns of historic aquatic invasions. The probability of NAS sightings increased with increasing city size, commercial maritime trips, and marina size for all NAS, decade, and pathway combinations. Of the land-use metrics evaluated, city population size was the best model factor and potential proxy of both NAS presence and richness, even for NAS introduced through ballast water discharge. The importance of commercial maritime traffic to NAS presence and richness may have been underrepresented due to rapid secondary spread of planktonic NAS away from port locations prior to detection. Nonetheless, city population size, total commercial maritime trips, and marina size may be reasonable proxies for propagule pressure given the significant relationships between these specific pathway metrics and NAS sightings and richness, and as such, are applicable to the development of early detection monitoring programs in the Laurentian Great Lakes.

Table of Contents

List of Tables

List of Figures

Introduction

The spread of nonindigenous aquatic species (NAS) is a significant and persistent threat to biodiversity, the economy, and human health in the Laurentian Great Lakes region (Lodge et al. 1998, Butchart et al. 2010). The eradication of established invaders is often infeasible, but targeted efforts in response to early detection can be cost effective (Leung et al. 2002, Vander Zanden et al. 2010) and successful at preventing further spread or establishment (Anderson 2005). Because of the considerable time and resources that are necessary to detect new aquatic species introductions (Hoffman et al. 2011), predictive models designed to identify areas with high risk of introduction can be beneficial to guide early detection efforts. The development and evaluation of proxies that represent introduction potential can help prioritize sampling locations, thus maximizing the success of early detection programs and thereby aiding the prevention of NAS spread (Bossenbroek et al. 2001).

NAS can alter ecosystems through a variety of mechanisms (Simberloff 2010). Zebra Mussels (*Dreissena polymorpha)* clog water intakes for industrial and municipal usage, alter food webs by rapidly consuming large quantities of phytoplankton, and bioaccumulate anthropogenic toxins which can be transferred up trophic levels (MacIsaac 1996, Higgins and Vander Zanden 2010). Rusty Crayfish (*Orconectes rusticus*) can replace native species as the dominant crayfish in introduced ecosystems (Olden et al. 2006). Substantive declines in water quality have been attributed to introduction of Common Carp (*Cyprinus carpio*) (Weber and Brown 2011). Through increased predation and competition, transmission of pathogens, or by altering food webs and nutrient cycles

(Ricciardi and MacIsaac 2011), nonindigenous species introductions are considered to be the second greatest factor driving the decrease in biodiversity for both terrestrial and aquatic ecosystems, second only to habitat loss (Wilcove 1998, U.S. EPA 2014). In addition to ecological impacts, the global cost of damages and management of nonindigenous species was estimated to be $314 billion annually (in 2005 dollars), with the U.S. accounting for over one third of that (Pimentel et al. 2001 and 2005).

Nonindigenous Aquatic Species in the Laurentian Great Lakes

Human activities, such as global trade and recreational water sports, facilitate the transport, introduction, and spread of NAS (Mills et al. 1993, Sala et al. 2000, Holeck et al. 2004). Many NAS unintentionally bypass natural barriers to dispersal by using anthropogenic pathways between geographically isolated populations and novel environments. The speed and breadth of NAS dispersal has increased with human activity (Ricciardi 2006), leaving only 16% of all marine ecoregions worldwide without a NAS present (Spalding et al. 2007, Molnar et al. 2008).

The Laurentian Great Lakes are particularly vulnerable to NAS invasions, with over 180 species currently present in the region (USGS 2014). The introduction of the Sea Lamprey (*Petromyzon marinus*) in Lake Ontario in the 1830s is the earliest record of a NAS in the Great Lakes; Sea Lamprey are thought to have naturally, or via shipping, migrated up the Erie Canal after its construction (Mills et al. 1993). The rate of NAS introduction has increased with human activity, and a new NAS has been reported in the Great Lakes on average of every 8 months since 1959 (Ricciardi 2006), but appears to be slowing (USGS 2014). Invasions into the Great Lakes do not only represent a regional problem, they threaten North America because the Great Lakes serve as a ‘beachhead’ for NAS spread to the interior of the continent (Rothlisberger and Lodge 2013).

The process of NAS invasions can be divided into four stages separated by physical and biological barriers. The stages are transportation, introduction, establishment, and spread (Blackburn et al. 2011). Transportation is the facilitated movement of nonindigenous organisms between locations, followed by introduction which occurs when organisms are released into novel geographic regions. The invaders are established when the introduced individuals begin reproducing. Spread occurs when the nonindigenous species disperse within the newly invaded region. Since introduction and spread are similar processes in which organisms are released to new locations, they are often referred to as primary and secondary spread respectively. Since transportation and introduction are the first two stages of an invasion, they are often the focus of preventative actions and regulations. The progression from transportation to introduction suggests that quantifiable metrics of human land-use (pathway intensity) may be useful as proxies for introduction potential, i.e. more transportation leads to greater risk of introduction.

There are barriers within and between the stages of invasion that influence how it proceeds, such as survival, reproduction, dispersal, and environmental conditions. If individuals or populations cannot surmount the barriers, then invasions slow or stop entirely (Colautti and MacIsaac 2004, Blackburn et al. 2011). For example, the nonindigenous Killer Shrimp (*Dikerogammarus villosus*) requires highly oxygenated water (Boets et al. 2010), and thus if it were to be introduced into anoxic waters, individuals may fail to survive due to the unfavorable environmental conditions, and the invasion would halt prior to establishment.

The most common pathway for historic introduction and spread of NAS was via ballast water discharge from commercial vessels, accounting for 65% of the introductions in the Great Lakes, and 80% worldwide (Ricciardi 2006). To add stabilizing weight (ballast), water from departing ports is pumped into ship hulls and subsequently released once a vessel reaches its destination and the ballast is no longer required. Aquatic organisms that are native to departing ports are often unintentionally loaded with the ballast water and introduced to novel regions when the water is discharged.

In response to growing concerns, governmental regulation began in 1989 in Canada and 1993 in the United States. Inbound ships were thereafter required to exchange freshwater ballast with saline ocean water before entering the Great Lakes - St. Lawrence Seaway, with the intention of killing transient organisms either through flushing out of the ballast tanks or through osmotic shock from the rapid change in salinity (Morris 2009). The efficacy of this method is still in question for some taxonomic groups (Briski et al. 2011), since many successful invaders into the Great Lakes have high salinity tolerances, especially those with a diapausing egg stage (Bailey et al. 2004). Indeed, Grigorovich et al. (2003) concluded that in spite of the new regulations, the rate of NAS discovery in the 1990s was higher than in the 30 years prior.

Even though ballast water discharge was the primary pathway of introduction into the Laurentian Great Lakes, there are numerous other anthropogenic pathways that have facilitated species introductions. In Lake Superior alone, there are eight identified pathways by which NAS may be introduced or spread: maritime commerce, agency activities, organisms in trade, illegal activities, fishing and aquaculture, water recreation, tourism and development, and canals and diversions (Lake Superior Binational Program 2014). Organisms in trade and water recreation, in particular, have been identified as significant and under-regulated risks for introduction of NAS (Johnson 2001, McDowall 2004). Historically, the bulk of live releases (organisms in trade pathway) have been due to intentional fish stocking activities of legitimate hatcheries (Mills et al. 1993), but increasingly, attention is focused on private owner releases from aquaria and accidental introductions from garden ponds (Rixon et al. 2005). Recreational water activities also pose a high risk of spreading NAS, particularly when organisms 'hitchhike' on propellers and hulls or in bilge water tanks between boat launches, or when nonindigenous live bait is disposed of improperly or accidentally into receiving waters (Kerr et al. 2005, Rothlisberger et al. 2011).

The potential risk of introduction and spread of NAS varies by pathway. Ballast water discharge has been identified as the most significant pathway of introduction into the Great Lakes region (Ricciardi 2006), but given that the vast majority of ballast water discharge and commercial trips are domestic (U.S. and Canada) within the Great Lakes, the potential for within and among lake spread may be even greater (Rup et al. 2010). Recreational boating on the other hand is rarely the cause of introductions of foreign species (Buchan and Padilla 1999), but is widely understood to be the predominant cause of introductions to inland lakes throughout the United States (Rothlisberger et al. 2011, Kelly et al. 2012). It is important to note that mechanisms by which NAS are introduced might vary from how they spread, and that this may lead to different spatio-temporal patterns of species sightings for each pathway. The prominence of ballast water discharge for primary spread suggests that introduction occurs mainly in and around ports where discharge is greatest. However, the same invader may be found in marinas soon after introduction due to recreational boat activity near ports that transport the NAS via this secondary spread pathway; multiple stages of the invasion process can be found in close proximity, facilitated by separate pathways and constituting different NAS sightings patterns.

Modeling Nonindigenous Aquatic Species Presence and Risk

Although the major pathways of introduction and spread have been identified, it is difficult to directly predict if specific NAS will establish a population once introduced. Propagule pressure (PP; the frequency and abundance of introductions) is broadly accepted as a meaningful indicator of species establishment success (Kolar and Lodge 2001, Leung et al. 2004, von Holle and Simberloff 2005, Lodge et al. 2006), but this relationship has rarely been quantified (Bradie et al. 2013). It is extremely difficult to calculate PP for aquatic organisms, so investigators have used proxies such as discharge volume and number of animal imports (Lockwood et al. 2009) to model the risk of introduction and spread. Net cargo tonnage was proposed to be a reasonable proxy for PP for most commercial vessels (Ricciardi 2006, Lo et al. 2012), but PP proxies for this pathway in particular have had only mixed success at explaining sightings patterns

(Wonham et al. 2013). The quantity and proximity of recreational boat traffic was successfully used as a proxy for NAS presence in inland lakes and at small scales (Bossenbroek et al. 2001, Herborg et al. 2009), while live release events have been estimated using the number of aquaria stores within urban areas, population size, and ownership survey data (Duggan et al. 2006, Gertzen et al. 2008). Even though these proxies have been developed and successfully used to understand NAS patterns, there is a lack of broad spatial scale studies that incorporate multiple pathways of introduction and spread in the Great Lakes proper.

Aquatic introduction risk assessment models have focused on three general areas within invasion biology: common characteristic, ecological niche, and introduction pathway. Common characteristic models analyze traits of successful invaders, along with their invasion histories, to predict which species may be the next to invade and which may prove detrimental if introduced (Kulhanek et al. 2011). This modeling approach successfully singled out invaders that are currently in the Great Lakes, and identified 17 others species that are posed to invade via ballast water discharge (Ricciardi and Rasmussen 1998).

Alternatively, ecological niche models (ENM) compare abiotic parameters of an invader's native region with abiotic characteristics of novel ecosystems to determine if introduction may lead to establishment based on the similarity of environmental conditions (Peterson and Vieglais 2001); ENMs target the establishment stage of an invasion. This approach was used by Herborg et al. (2007) to model and predict the risk of introduction of the nonindigenous Chinese Mitten Crab (*Eriocheir sinensis*) based on the species' reproduction requirement of high saline waters.

Pathway introduction models use measurable land-use quantities to assess potential locations where introduction and spread may occur. Similar to ENMs, pathway models can potentially identify specific locations where new NAS introductions may occur and establish, and thus where preventative action may be beneficial. Unlike ENMs, pathway models are applicable to multiple species, since a single pathway may facilitate various species, while ENMs are designed for species-specific abiotic criteria. The pathway modeling approach successfully predicted which inland lakes were susceptible to Zebra Mussel invasions from the recreational boat pathway. In this case, each lake was treated as an experimental unit with an attributed value of pathway intensity (amount of boat traffic) and a presence/absence value for Zebra Mussel. Lakes with higher pathway intensity had a greater chance of Zebra Mussel presence (Schneider et al. 1998, Gertzen and Leung 2011).

Integrated into many modeling approaches is the usage of a geographic information system (GIS). Haltuch et al. (2000) used spatial data in a GIS framework to determine dispersal rates and range expansion of nonindigenous *Dreissena* mussels in Lake Erie. The spread of the Eurasian Round Goby (*Neogobius melanostomus*) throughout Wisconsin tributaries of Lake Michigan was successfully explained by modeling spatially explicit abiotic factors of invaded streams (Kornis and Vander Zanden 2010). Glardon et al. (2008) used GIS to identify the risk of introduction and spread of an aquarium strain of the green alga *Caulerpa taxifolia* in coastal waters of Florida. The use of GIS in predictive and historic modeling has been paramount to the success of many projects.

Study Objective

Based in the pathway modeling approach, the objective of this study was to develop and evaluate three land-use metrics (commercial boat traffic, city population size, and recreational boat traffic) as potential proxies of PP associated with three major pathways (maritime commerce, organisms in trade, and water recreation). The metrics were tested as model factors that could explain the apparent spatio-temporal relationship between pathway intensity (land-use metric to represent propagule pressure) and either historic NAS sightings or cumulative NAS richness. That is, model factors that are significant and that can account for substantial variation in NAS sightings might be good proxies of future introduction potential. These land-use metrics could then be applied at a basin-wide scale to direct early detection efforts to locations where risk of introduction and spread are highest based on the historic sightings patterns. I hypothesize that increasing levels of pathway intensity (i.e. propagule pressure) correspond to greater probability of NAS presence and also higher NAS richness. Also, species-specific relationships do exist between NAS and their introduction and spread pathways.

Methods

Fundamental to the design of this project was the inherent diversity of commercial maritime traffic, recreational marinas, and city population size within the Laurentian Great Lakes region, which spans roughly 1,300km of longitude and 850km of latitude. An average of 250 million tons of cargo is moved by water through the region annually between major U.S., Canadian, and foreign ports (U.S. Army Corps of Engineers 2014). Over 500 recreational marinas have been documented (Allan et al. 2013) and roughly 36 million people live within the region in a variety of sized cities (U.S. Census Bureau 2010, Statistics Canada 2010). This variation in the combinations of spatial overlap and intensity of introduction pathways was exploited to conduct a ‘natural’ experiment to evaluate land-use proxies for PP as a means to inform where future early detection efforts should be focused. Two dependent variables (NAS presence and NAS richness) were tested using logistic and linear regression analyses, while a nonmetric multidimensional scaling (NMDS) ordination was used to explore patterns in species composition near invaded cities of varying size.

Relatively sessile study organisms were the focus of this project because the goal was to isolate anthropogenic introduction and spread patterns from patterns associated with naturally dispersing species. In a similar effort to limit this study’s scope to the fundamental impacts and patterns of anthropogenic pathway influence, this project was designed to omit all consideration of environmental conditions. In the broadest sense, I aimed to detect and test only human influence on NAS introduction and spread in the Great Lakes. Even though the inclusion of environmental characteristics has increased the accuracy of smaller scale and inland lake presence/absence studies (Leung and Mandrak 2007, Herborg et al. 2009), for this geographically and environmentally broad project, the goal was to find a human impact signal regardless of variation in the chemical and physical qualities of aquatic habitats.

Development of Proxy Metrics for Propagule Pressure

Proxy metrics for propagule pressure were developed using land-use data for three prominent pathways of aquatic introduction and spread: commercial boat traffic, recreational boat traffic, and live release from urban areas (Table 1). Each pathway was associated with a set of spatially-explicit locations in a geographic information system (GIS). Commercial boat traffic was based in ports, recreational boat traffic was centered in marinas, and live release from urban areas was associated with cities with populations above 2,500 people. All spatial analyses were conducted using ESRI ArcMap v10.1 (www.esri.com), while statistical analyses were done in SigmaPlot (www.sigmaplot.com) and the R statistics program (www.r-project.org).

Three metrics were developed for commercial boat traffic: total number of commercial boat trips to a port, average annual cargo imports and exports (metric tons), and annual average ballast water discharge (m^3) within a port. These three commercial boat traffic metrics were further categorized by trip origination and source location of the ballast water or cargo, as well as whether a trip resulted in discharged ballast water or not.

Data for total commercial boat trips and the ballast discharge history of those trips were gathered from the National Ballast Water Information Clearinghouse (NBIC) database (http://invasions.si.edu/nbic). The discharge history of a commercial vessel trip included designation as discharging or non-discharging, and the origin of the discharged ballast water as domestic (i.e. within U.S. or Canadian waters of the Great Lakes) or from overseas. Vessel trips were designed as having an overseas ballast water origin if the vessel passed through waters further than 200 nautical miles from the U.S. or Canadian landmasses. Data was compiled for 67 U.S. ports from 2004 – 2013 (Table 1).

Annual average tonnage of cargo imported and exported per port was gathered from the U.S. Army Corps of Engineers principle ports data files (U.S. ACE 2014). Cargo summaries for 45 ports on the Laurentian Great Lakes were gathered, including descriptions of cargo origin as domestic or foreign. The dataset spanned 15 years from 1996 to 2011.

Annual average ballast water discharge data were gathered primarily from Bailey et al. (2012), which were augmented with data from the NBIC database. The dataset was available from a published source (Allan et al. 2013) and was limited to the annual average discharge values for 2005 – 2007. The dataset from Bailey et al. (2012) was more comprehensive than the NBIC because it also included data from the Canadian Coast Guard's Information System on Marine Navigation (INNAV). To add to the dataset, ports with zero ballast discharge according to the dataset that had ballast records in the NBIC database for 2005 – 2007, were changed to the NBIC values; 7 ports were altered in this manner. This hybrid of Bailey et al. (2012) and the NBIC database was used as the primary ballast water discharge metric for 97 U.S. and Canadian ports (Table 1).

In addition to the designations of ballast water origins as domestic or foreign, volume of ballast water discharge from high-risk ports was calculated. Current NAS richness was determined for each port API (area of pathway influence – see methods section) and the top quartile of ports (>7 species) was designated as high-risk. The current NAS richness was used to be consistent with the time range of the data. Even though the constituents of the top quartile of NAS richness ports changed some from the 1970s to the present, the ranking of ports by NAS richness did not fluctuate that much for 97 sample ports; the average standard deviation of the NAS richness rank of a port was 1.95, with a maximum standard deviation of 4.5, indicating that the rank of a port changed little over the study time period. The hybrid dataset of Bailey et al. (2012) and the NBIC database was used to calculate total volume of ballast arriving from each of the identified high-risk ports.

One inherent problem with the commercial boat traffic metrics was the limited range of historical information available. In order to test the assumption that the relative activity at a port did not change much over the study time period, I did a preliminary study that ranked the annual amount of cargo imported and exported from a port for each year from 1996 to 2011; the cargo dataset was used because it had the greatest temporal range of any of the commercial boat traffic metrics. The average standard deviation for each port was 2.05, with a maximum standard deviation of 4.1 for 45 ports. Based on these results, I accepted the assumption that the relative amount of shipping activity at each port remained fairly constant through time. Thus, all commercial boat traffic metrics were reasonably represented with recent data as a necessity of data range limitations.

The experimental proxy metric for recreational boat traffic was marina size as estimated by the number of boat slips within a marina. The locations of the marinas and the number of slips were available from Allan et al. (2013), and provided by the Great Lakes Environmental Assessment and Mapping project (www.greatlakesmapping.org).

Slip data were available for one year (2010) for 516 marinas (Table 1). Again, the limitations of historical data necessitated the assumption that current marina size adequately represented past land-use patterns. Unfortunately, there was not an opportunity to verify this assumption given that the data was limited to only one year.

The proxy metric for the live release pathway was population size within metropolitan and micropolitan statistical areas of greater than 2,500 individuals as delineated by the U.S. Office of Management and Budget and the Canadian Census Bureau. Population size was used, as opposed to population density, because NAS introduction potential is better represented as a function of total intensity (number of individuals), not population density (individuals/area). Introduction potential can be modeled as a constant rate per capita because propagule pressure increases with the presence of aquaria stores and live food markets (Duggan 2006), based on the assumption that the prevalence of aquaria stores and live food markets increase with population size. Thus, the total potential of a single location is a function of number of individuals (population size).

Both the spatial extent for the statistical areas (urban areas/cities) and their corresponding populations from 1970 - 2013 were obtained from the U.S. and Canadian Census Bureaus (U.S. Census Bureau 2010, Statistics Canada 2010). Population data were divided into five decadal periods (1970s, 1980s, 1990s, 2000s, and 2010s). If population values were unavailable (from the online sources) before a given decade, the most historic city population values were attributed to all decades prior to the last recorded decade. This correction was applied to 35 Canadian cities for the 1990s, 1980s, and 1970s, as well as 6 U.S. cities for the 1980s and 1970s. Population size was determined for 75 U.S. and 41 Canadian urban areas (Table 1). Over 77% of the population of the Great Lakes region (~27 million people) was represented in these urban area delineations.

Area of Pathway Influence (API)

Spatial pathway data were gathered as GIS-based point locations and polygon areas from their corresponding data sources (Table 1). Marinas were modeled as points located in the waters of the Great Lakes proper, while cities were polygon areas on the land adjacent to the Great Lakes. Aerial photography was used to digitize port polygons by incorporating all physical structures directly related to port activity, e.g. loading docks, ship berths, etc. For each pathway point or polygon, a larger spatial polygon, i.e. a buffer, was generated to represent the area of pathway influence (API) of each location into the waters of the Great Lakes. An API is the spatial representation of the extent of influence, impacts, and introduction and spread potential of a given pathway location. Buffer areas were generated directly around marinas and ports to create APIs, while the maximum extent of a city's waterfront was used to delineate the shoreline boundaries of each city API. All buffers were limited to the waters of the Great Lakes, which excluded tributaries and near shore lakes. Over 81% of all NAS sightings were within one or more of these APIs.

Buffer distance for the three pathways varied. Port and city pathway locations had 10km buffers applied, while marinas had a 1km buffer. Buffer distances were based on a combination of the distances between pathway locations and the physical characteristics of the Great Lakes. I assumed that introduction and spread are facilitated and limited by near shore physical processes; given that physical processes in the Laurentian Great Lakes separate coastal and offshore waters (Rao and Schwab 2007, Yurista and Kelly 2009), a 10km distance was set as the zone of greatest influence with respect to NAS spread into the waters of the Great Lakes. If overlap occurred between buffers in the same pathway (Figure 1A), the two pathway locations were considered not independent, and a single API was created from the combination of the maximum extents of the original overlapping buffers and the sum of the pathway intensity metrics were attributed to the new larger API (Figure 1B). When the port buffers were set to the maximum distance of 10km, the number of independent port locations was decreased by about 30% on average for each metric. For marinas, a 1km buffer was used to represent the small physical nature of marinas as structures designed to separate nearshore waters from open water. A 1km buffer applied around marinas produced a similar reduction in the number of spatially-independent marinas, from 516 to 304 (~40%) (Table 1).

For city APIs, a10km buffer was used to match the port buffer distance, but city APIs were not combined where overlap occurred; rather, the mid-point between two city APIs was used as a division (Figure 1C). This was done because numerous cities were immediately adjacent to one another, which would result in a only a few large continuous APIs accounting for the majority of the Great Lakes, especially in Lake Michigan, Erie, and Ontario.

I tested each set of pathway APIs (i.e. city APIs, port APIs, and Marina APIs) for spatial autocorrelation using Moran's Index (Moran 1948) calculated using the Spatial Autocorrelation Tool in ArcMap v10.1. Spatial autocorrelation is the tendency of objects that are spatially close to be more similar to one another than to objects further away. Spatially autocorrelated data inherently violates assumptions of statistical independence, because some inference can be made about the values at a location based on neighboring values.

In addition to the individual APIs (port, city, marina), composite APIs were created to test NAS presence and richness using multivariate logistic and linear regression analyses. The univariate APIs treated each set of pathway APIs as independent feature sets on the landscape (Figures 2 and 3). Bivariate (port/city, port/marina, and marina/city) and multivariate (all three pathways) APIs were assumed to influence one another. Wherever spatial overlap of different individual pathway APIs occurred, a new spatial API combination was created with the limited extent of the overlapping APIs, which was attributed with all metrics of the underlying pathways (Figure 4). When spatial overlap did not occur in the bivariate or multivariate APIs, zero values were attributed for the non-overlapping pathway, i.e. in the bivariate APIs where only a city was present, the port metric for that given API would be zero.

Nonindigenous Aquatic Species Sightings

The primary dataset for NAS sightings was the U.S. Geological Survey Nonindigenous Aquatic Species database (USGS NAS) (www.nas.er.usgs.gov). The database is comprised of voluntarily reported NAS sightings from individuals and government agencies, as well as from literature reviews by the USGS. Before inclusion in the database, taxonomists confirm questionable sightings reported from non-scientific sources. Each record included date of sighting, location, and assumed introduction pathway. This dataset may be subject to pseudo-absence errors since reporting is voluntary and participation in, and knowledge of this database, are unknown. Some limitations of using presence only data are that prevalence (commonness) cannot be determined (Ward et al. 2009), and observational bias may affect analysis (Phillips et. al 2009). Determining prevalence was not an objective of this project, but observational bias may be present in this dataset. Unfortunately, this bias cannot be identified or isolated since the dataset does not have detailed information about the source of each record. In addition to the USGS NAS database, incidental NAS sightings from the Great Lakes Indicator Consortium's (GLIC) Coastal Wetland Monitoring (CWM) project for 2011 and 2012 (unpublished data) were included as well. Given the spatial scale and temporal range of this approach, this is the best available dataset for historical NAS sightings across the Laurentian Great Lakes region back through 1970.

Sightings included for analysis were limited to invertebrate NAS and five select fish species sightings in the Great Lakes from 1970 to 2013. Invertebrate NAS were selected because of their relative inability to propagate great distances independent of anthropogenic pathways, while the fish species were included specifically for their direct association with the aquaria trade and live release pathway (Duggan 2006, Gertzen and Leung 2008) or commercial boat traffic (Mills et al. 1993). Only species sightings within

10km of the shore were used; sightings beyond 10km from shore were not assigned to an API.

The study included 23 species and ~3300 sightings (Table 2 and Figure 5). Eighteen of those species had greater than 10 sightings and thus were of the greatest utility for analysis. All analyses were considered presence only; complete absence data did not exist for either source of NAS sightings.

Quantifying the Relationship between Human Activity and NAS Presence

I used a GIS-based spatial approach to attribute each API with a presence (1) or absence (0) score using species sighting locations. If a species sighting occurred within an API, the NAS sighting was associated with that specific API. Each API had an associated pathway-specific metric that quantified the amount of human activity at the given location; these metrics were the pathway intensity of an API. Presence relationships for each of the three pathways over five decadal periods were tested using univariate, bivariate, and multivariate logistic regression analysis between pathway intensity and NAS presence (Figure 6). To address each of the three pathways equally, two group species models and three individual species were analyzed for presence relationships. Individual species models were conducted for Quagga Mussel, Round Goby, and Zebra Mussel, because these species had the greatest number of sightings, but also because of their specific documented associations with the commercial and recreational boat pathways (USGS 2014). The organism in trade pathway was modeled explicitly by developing a group aquaria species model comprised of seven NAS

(Goldfish (*C. auratus*), Freshwater Jellyfish (*C. sowerbyi*), Chinese Mystery Snail (*C. chinensis)*, European Ear Snail (*R. auricularia)*, the freshwater hydroid *C. caspia*, the freshwater bryozoan *L. carteri*, and Pacu (*C. macropomum*)) that had documented linkages with that pathway (USGS 2014). A group model was used to increase the sample size of the sightings because all the aquaria species had fewer than 100 records in the database, and five of the six had fewer than 10 records. The second group model was an all species model that included all NAS records in an effort to detect broad introduction and spread patterns (Table 2). To assess temporal changes, each model was divided into decades (i.e. 1970s are sightings between Jan. 1, 1970 and Dec. 31, 1979).

To address temporal variability of a species sighting at a location, a correction for pseudo-absence was applied where each species sighting was determined to persist from one decade to the next from the earliest sighting onward, e.g., a sighting in an API in 1991 would result in sightings for the 1990s, 2000s, and 2010s for that given species and API. Temporal variation in sighting records may reflect inconsistent reporting practices, the discontinuation of reporting once a species was well established in an area, or the database-side practice of non-entry of reports after a species was well documented in a specific location (USGS 2014).

Statistical Analyses

Univariate, bivariate, and multivariate logistic regression analyses were conducted for each time period (decade) and NAS or NAS group using presence/absence data. These regressions represent a continuous probability of presence for a specific NAS in an

API of given pathway intensity. The odds ratio value was calculated for bivariate and multivariate models. The odds ratio is the percentage that the response variable (y-axis) will change due to a one-unit change of the independent variable (x-axis). Positively correlated relationships have positive values, while negatively correlated relationships have negative values. To determine whether there was a change in the species-specific association with each pathway after initial introduction, the odds ratio value was calculated for the first decade of an invasion for each NAS and NAS group, and for the most current decade (2010s). The first decade of an invasion varied by NAS and was based on the year of the first sighting since 1970. Model significance was tested using the likelihood ratio statistic and its associated p-value. The model fit was assessed using the -2*log(likelihood) value (-2LL). Values of -2LL closer to zero indicate a better model fit (Hair et al. 2006).

An additional multivariate logistic regression analysis was conducted to test for relationships between pathway intensity of ports, cities, and marinas, and NAS first sightings locations; the first sightings records were determined for those NAS with more than 10 sightings (n=18) since 1970. First sightings were grouped based on the year of the record, and 10 NAS had multiple first sightings occur in the same year at different locations. Both Round Goby and the Freshwater Jellyfish *Craspedacusta sowerbyi* had greater than 10 first sightings. There were a total of 74 first sightings records for the 18 primary species; the average number of first sightings was 2.5.

Since numerous analyses were conducted on the same datasets (species sightings and pathway intensity), the Bonferroni inequality method (Bonferroni 1936) was applied and a significance level of $\alpha = 0.01$ was used to determine if results were significant. The presence of multicollinearity was tested for in every multivariate model using the variance inflation factor (VIF). Multicollinearity is when two or more independent variables in a multivariable analysis have a linear correlation prior to the multivariate analysis. Multicollinearity can result in spurious and inflated associations between independent and dependent variables.

Nonindigenous Aquatic Species Richness and Composition

NAS richness was calculated as the cumulative total of unique NAS associated with an API over the entire study time range (1970 - 2013). Species richness was calculated for univariate and multivariate APIs, and then analyzed using linear regression. An $\alpha = 0.01$ was used to determine if the results were significant and R^2 value were used to assess model fits.

NAS composition patterns within city APIs were tested using a 2-axis nonmetric multidimensional scaling (NMDS) ordination. NMDS analyses were done for two size classes of cities: big cities (>35,000) and small cities (<35,000). The division between big and small was determined from the results of the NAS richness analysis (see results section) that suggested a difference in NAS richness trends above and below the population level of 35,000. Cities with zero NAS richness were not included in the analyses. To generate overlay vectors to the ordinations, a Pearson's correlation was calculated between each NAS presence dataset and the two NMDS axis values; these scores were used as the end points (x and y coordinates) for each vector leading from the origin. Overlay vectors in similar alignments would suggest that NAS were often found in similar locations, and thus differences in the composition of NAS in various city size groups could be assessed. The application of the Pearson's Correlation to compare a binary and a continuous variable is known as a Point-Biserial Correlation (Tate 1954).

Management Application – Marina Monitoring

To explore the application of the multivariate models to early detection monitoring in marinas in Lake Superior, the historic probability of NAS sightings at 31 sites was back-calculated using the 2010s all species multivariate logistic regression equation. Ordering the probabilities from highest to lowest created a priority monitoring location list, where the assumption was made that locations with the highest historic probabilities of presence were the most likely to be invaded again. To compare the difference between monitoring in marinas, or in the ports and cities around the marinas, a second list of probabilities was calculated by setting all marina values to zero; discrepancies between the two lists could be attributed to monitoring in marinas specifically (i.e. within the 1km marina buffer zone vs. the adjacent 10km port or city buffer zone). The U.S. side of Lake Superior was chosen as the application area because the U.S. Fish and Wildlife Service, Ashland, WI branch is currently developing a plan for an invertebrate NAS early detection network in Lake Superior.

Results

From 1970 forward, there was at least one NAS sighting in every port API except one (Sliver Bay, MN), in all but seven city APIs (94%), and in 148 (~50%) marina APIs. The intensity of the pathway metrics varied widely both between and within pathways.

When tested for spatial autocorrelation; ports, cities, and marinas were not distributed on the landscape significantly different than randomly (Moran's Index: p-value = 0.304; p-value = 0.957; p-value = 0.35). Though city APIs were spatially associated with both ports and marinas, 80% of port APIs and 70% of marina APIs overlapped city APIs, low variance inflation factor (VIF) scores for the independent variables in the multivariate logistic regression models (port=1.058, city=1.055, marina=1.005) did not support the presence of multicollinearity (i.e., it was not the case that large ports and marinas were generally associated with large cities, or vice versa).

Nonindigenous Aquatic Species Presence

First Sightings

All three pathway metrics were significant model factors with respect to locations of first NAS sightings (since 1970). Notably, city population size and commercial vessel trips had roughly the same odds ratio values, 58.2% and 54.1% respectively ($p= 0.007$ and 0.002). Marina size was also a significant factor, but had a negative odds ratio, -52.1% ($p <0.001$). The multivariate model was significant when tested using the likelihood ratio statistic ($p<0.001$) and the model fit (-2LL) was 162.3.

Univariate Models

All NAS or NAS groups analyzed using univariate logistic regression were more likely to be present at larger or more active locations then at smaller locations. This trend was consistent from decade to decade, and the pathway intensity value at p(50) (50% probability of presence) decreased over time in all models (Table 3, Figures 7 and 8). A decreasing pathway intensity value at p(50) over time indicated that probability of NAS presence increased at pathway intensity levels above the p(50). For example, the p(50) pathway intensity value for the 1980s all species city pathway univariate logistic regression model was 5.15 (city population size of ~141,000); in the 1990s when the p(50) pathway value decreased to 3.4 (~2500), the probability of presence for the pathway intensity of 5.15 and above, had increased to 0.95 (Figure 7B)

Of the three metrics for the commercial boat traffic pathway (total vessel trips, average annual ballast water discharge, and average annual cargo), only the vessel trips metrics (total vessel trips and non-discharging vessel trips) had more than one significant relationship with NAS presence (Table 4). The differences in the model fits (-2LL) between the two trips metrics were minimal. There was a significant linear correlation between the two metrics ($R^2 = 0.5$, $p<0.001$), suggesting that both factors had a similar relationship with NAS presence. Since total trips was more inclusive as a metric, it was designated as the best commercial boat traffic metric. Total vessel trips was the only significant univariate factor of those evaluated for commercial boat traffic; it notably had only significant relationships in the Round Goby models (Table 4).

There were significant relationships between city size and NAS presence for all NAS and decade combinations except for the aquaria species model in 1980s (Figures 7 and 8). The probability of presence was always lower for smaller cities compared to larger cities. Each NAS or NAS group had a species-specific relationship with city population size (Figure 7 B, E, H).

The Round Goby commercial boat traffic model had a large decrease in the pathway value at p(50) from the 1990s to the 2000s, and no change from the 2000s to the 2010s (Figure 7D); the probability of presence increased above the p(50) pathway value and decreased below the p(50) pathway value from the 2000s to the 2010s as the sigmoidal shape varied slightly. The regression curve shapes were similar to those in the city size univariate models. The model fit improved from the 1990s to the 2010s (Table 3), suggesting continued spread facilitated by commercial boat traffic.

For the Quagga Mussel city size models, from the 1990s to the 2000s, the probability of presence sharply increased as city size increased, with a maximum increase of 30% in cities around 10^6 (1,000,000) population size. In the 2000s, the p(50) pathway intensity value was found to occur near 10^5 (~80,000) population size, while there was only a 1% increase in cities with less than 10^4 (10,000) population size (Figure 7B). The pathway value at p(50) was the same for the 2000s and the 2010s. Quagga Mussel sightings appear to have been limited to large and medium sized cities. The model fit was similar from the 1990s to the 2010s.

Alternatively, the Round Goby city size model from the 1990s to the 2000s had a decrease in the p(50) pathway value from a city size of $10^{5.3}$ (~200,000) to $10^{3.8}$ (~6,300) with only a 10% increase in the probability of presence in the largest cities and a 30% increase in the smallest cities (Figure 7E). After the 2000s, the p(50) pathway value decreased to a city size of $10^{3.4}$ (~2,500). From the 1990s to the 2010s, the model fit initially became worse and then got better.

The Zebra Mussel univariate logistic regression models had a roughly equal increase in probability of presence at both extremes of the regression curves, with probability increases of around 20% at $10^{3.5}$ (~2,500) and 10^7 (10,000,000) city sizes (Figure 7H). There was a dramatic shift in the pathway intensity at p(50) and the rate of change of the probabilities of presence from the 1980s to 1990s, which suggests a rapid increase in sightings in medium sized cities (10^3 to 10^4), while small and large cities had a similar sightings rate post-introduction. The p(50) value in the 1990s, 2000s, and 2010s was the same. The Zebra Mussel model fit was similar among decades.

The all species model indicated a rapid increase in sightings from the 1970s to the 1980s for all but the largest cities, and a larger increase from the 1980s to the 1990s. By the 1990s, all cities approaching a population of ~$10^{4.5}$ (31,500) had a ~100% chance of NAS presence; the minimum probability was 40% in the smallest cities (Figure 8B). The p(50) pathway value decreased substantially from the 1970s to the 1990s. The probability of presence for any sized city was greater than 0.5 for the 2000s and 2010s. Zebra Mussel and Round Goby sightings in the 1980s and beyond strongly influenced this model because they accounted for over 35% of total sightings records. The all species univariate city size model fit declined in the 1980s then improved for each decade after the 1980s.

Notably, the aquaria species model had significant relationships with city population size beginning in the 1990s. From the 1990s to the 2010s, the p(50) pathway value decreased from a city size of $10^{6.6}$ (~4 million) to $10^{5.7}$ (~500,000). The probability of presence was always highest in large cities and the increase over time was greater in large cities than small cities (Figure 8E). The largest probability increase was from the 1990s to the 2000s (30%). The model fit was slightly worse after the 1990s.

Similar to the city pathway models, the marina pathway models had significant relationships with most NAS for most decades (Figures 7 and 8). Increasing marina size was associated with greater probability of presence for Round Goby, Zebra Mussel, and all species, and not significant for Quagga Mussel and aquaria species. Again, species-specific variation was present. The Zebra Mussel and all species models had decreases in the p(20) pathway value (marina size where probability of presence was 0.2; p(20) was used when there was no p(50) value) from a marina size of $\sim 10^{3.4}$ (~2,700) to $\sim 10^{2}$ (100) slips between the 1990s and 2000s, with a nearly 50% increase in probability of presence in the largest marinas and a 25% or less increase in marinas with below 10^{2} (100) slips (Figures 7I and 8C). Conversely, Round Goby probability of presence was similar for all decades (Figure 7F). The model fit for the all species and Zebra Mussel marina models declined from the 1980s to the 1990s, and thereafter remained roughly constant. The Round Goby marina size model fit declined from the 1990s to the 2010s.

The temporal changes in the probability of NAS presence varied across pathways (Figures 7 and 8). Even though there was species-specific variation in the regression curves for each pathway, all the city models and all the marina models had similar pathway specific temporal shifts. The city models consistently had a higher maximum probability of presence as compared to the marina models. The largest changes in probability for the marina models were near the largest marinas, whereas changes in the city models' inflection points and rate varied by species. The aquaria species city models resembled the progression of the marina models more than the city models. The two significant commercial boat traffic models were similar to the trend for the city pathway models.

Multivariate Models

Total and non-discharging commercial vessel trips had the greatest number of significant relationships of any of the commercial boat traffic metrics in the multivariate logistic regression models (Table 5). The model fits were the same when using either the total or non-discharging metric; thus, total commercial vessel trips, as the most inclusive of the vessel trips metrics, was determined to be the primary commercial boat traffic metric in multivariate modeling. All further reported multivariate models include only this metric to represent the commercial boat traffic pathway.

Notably, commercial boat traffic (ports) had the smallest odds ratio value in all models except for the Round Goby multivariate logistic regression model. It was the least important factor in the multivariate NAS presence models (Figure 9). However, it did have the largest odds ratio for the Round Goby model for both the first and current decade (76% and 59% respectively). The odds ratio of commercial boat traffic for both the Round Goby and Zebra Mussel models decreased from the first decade to the 2010s.

In contrast, city population size had the largest single odds ratio and was a significant model factor for all NAS and all decades except the 1980s aquaria species model (Figure 9). The odds ratio for the all species model in the 1980s was 129%, nearly double the next closest value. Similar to the commercial boat traffic metric, the odds ratios from this pathway decreased from the first decade to the current decade, indicating that a substantial increase in presence of NAS sightings had occurred near cities. The only exception was for the Quagga Mussel model, which had an increase over time in the odds ratio for the city population metric. Marina size had negative odds ratio values, which was consistent with results from the first sightings multivariate model. The odds ratio values became more negative from the first decade to the 2010s, which was opposite of the trend for both vessel trips and city population size. This suggests that the negative correlation with marina size increased over time. The model fit for each model got worse from the first decade to the last.

Bivariate Models

Bivariate logistic regression models for each combination of pathways (port/city, port/marinas, and marina/city) were also evaluated (Table 6). Consistent with the multivariate models, increasing commercial vessel trips and city population size resulted in increased probability of NAS presence, while increasing marina size decreased the probability.

When both commercial boat traffic and population size were considered, only three of 10 models resulted in both metrics being significant factors to NAS presence

(Table 6); the all species, Round Goby, and Zebra Mussel models had both metrics as significant factors in the 2010s. For these three models, the odds ratio for ports was higher than for cities. For all but two of the bivariate port and city models, city size was a significant factor, and for most of those models, city size was the only significant factor.

For the bivariate logistic regression port and marina models, the commercial boat traffic metric was the sole significant factor for 75% of the models. Alternatively, when city and marina size were modeled together, in all but one model, both factors were significant. VIF values for the bivariate models were all below 1.05, suggesting that no multicollinearity was present.

Model Fit and Logistic Regression Results Summary

The -2*log(likelihood) statistic indicates a perfect model fit at value 0, but many of the calculated -2LL values ranged between 100 to >600. With each subsequent addition of a variable from univariate to bivariate to multivariate, the maximum -2LL value increased. However, the likelihood ratio statistics indicated that many of the relationships were significantly better than models that did not have the independent variables included (pathway intensity metrics). The most conservative interpretation of the univariate data is to use only the pathway value at p(50), which inherently discards much of the species-specific interpretations of the data. Nonetheless, the same conclusions can be made that NAS presence occurred preferentially in large ports, cities, and marinas, and then spread to smaller locations, because the p(50) pathway values decreased as a function of time. For the Round Goby, Zebra Mussel, and the all species marinas univariate logistic regression models where the p(50) pathway value was not reached for any of the decades, the conservative interpretation would suggest that those decades did not have any significant NAS presence until the p(50) value was reached.

Overall, city population size was the best univariate and multivariate factor of those evaluated. The metric had significant relationships with all NAS and NAS groups for the majority of decades and often had the largest odds ratio values. Number of commercial vessel trips was the best metric for commercial boat traffic in univariate and multivariate logistic regression models, but none of the commercial boat traffic metrics increased model fit more than the others. Marina size was also a significant proxy of NAS presence. Increasing marina size in the univariate models equated to increased probability of presence, but when included in multivariate models, the odds ratios were negative, indicating a negative relationship.

Nonindigenous Aquatic Species Richness

The highest NAS richness was 13 species for both the commercial boat traffic and live release pathways, and it occurred in the Duluth-Superior port and city APIs respectively. The highest species richness for marinas was 6 and occurred in three locations: Chicago, Il; Erie, PA; and Port Colborne, Ontario. The average NAS richness was 4.6 in port APIs, 3.7 in city APIs, and 0.9 in marina APIs.

Univariate linear regression analysis demonstrated that both the marina and city pathway metrics were significant positive factors of NAS richness (Figure 10), and that a linear relationship was the best fit for the models. Commercial boat traffic was not a significant factor of NAS richness. City population size had a larger R^2 value ($R^2 = 0.39$) than marina size ($R^2 = 0.10$).

All three pathway metrics were significant factors in the multivariate linear regression models of NAS richness and pathway intensity (Table 7). The relative magnitude of the regression coefficients of commercial boat traffic and city size was fairly even among decades, with the coefficient for commercial boat traffic averaging only 0.035 more than city size in each decade. The marina size coefficient was always negative and the greatest magnitude of the three. Notably, all coefficients increased sequentially from one decade to the next. Because the marina, port, and city metrics in the multivariate linear regression models were based on modern data (i.e., constant over time), the regression coefficients were responsive to the cumulative NAS richness. Thus, the relatively steady increases in magnitude, both positive and negative, of the regression coefficients are not unexpected given the increase in NAS richness over time. Because the changes in the regression coefficients were similar among decades, this suggests that NAS richness increased at a constant rate.

There was a significant linear relationship between API area and NAS richness for city APIs ($p<0.001$, $R^2=0.32$); there was no relationship between port ($p=0.35$) or marina ($p=0.08$) API size and NAS richness. Nonetheless, city population size was the best model factor to explain NAS richness near cities ($R^2=0.39$).

Species Composition in City APIs

Only the NMDS ordination for large cities resulted in a convergent solution (Table 8 and Figure 11). There was not enough data within small cities to generate an ordination with any stress (a measure of goodness of fit). Ordination axis 1 for large cities was defined by Bloody Red Shrimp (*Hemimysis anomala)*, Quagga Mussel (*Dressena rostriformis bugensis),* and Fishhook Waterflea (*Cercopagis pengoi)* on the positive side, and Common Carp (*Cyprinus carpio*) and Spiny Waterflea (*Bythotrephes longimanus*) on the negative side (Table 8 and Figure 11). Axis 2 for large cities was defined by Goldfish (*Carassius auratus*), New Zealand Mud Snail (*Potamopyrgus antipodarum*), and Scud (*Echinogammarus ischnus*).

Management Application – Marina Monitoring

The lowest historic probability of a NAS sighting in a marina in Lake Superior was 40% (Table 9); the greatest probability was 86%. Two Harbors, MN and Sault St. Marie, MI-Ontario were the top two most likely locations to find a NAS in a marina, while Duluth-Superior, which had the largest city and port metric values, was third. The highest probability marina in Duluth-Superior was the smallest one. However, when the multivariate logistic regression equation was rerun with all marinas set to zero (i.e. a search conducted within city or port 10km buffer and not in the marina 1km buffer), Duluth-Superior was the most likely location for NAS presence, with a 91% probability of a historic sighting; second was Marquette, MI, and third was Two Harbors, MN. The minimum probability of presence increased from 0.4 to 0.62 when marinas were set to zero; the maximum increased by only 0.05. The greatest increase was 24% in Washburn, WI, which has two large marinas. Seven of the top 10 locations were the same in both lists; Silver Bay, MN; Ashland, WI; and Ontonagon, WI were the three locations that moved into the top ten when the probability of presence was calculated for only ports and cities. The negative correlation between marina size and NAS presence explains why Two Harbors, MN and Sault St. Marie, MI-Ontario were the top two most likely locations to find a NAS, since neither had an associated marina value and each had the second highest metric for commercial boat traffic or city size respectively. Similarly, Silver Bay, Ashland, and Ontonagon moved into the top ten after the negative relationship represented by marina size was removed.

Discussion

In the Laurentian Great Lakes, the probability of NAS sightings increase with increasing city size, commercial vessel trips, and marina size. City population size was the best land-use metric to explain primary and secondary spread patterns. The prominence of city size as a factor of introduction and spread may be attributed partially to rapid secondary spread leading to detection of NAS in city APIs and not port APIs, where introduction actually occurred. Once introduced, spread throughout the Great Lakes basin occurred at varying rates and was pathway- and species-specific. Nonetheless, by 2013, nearly all locations with high human activity were occupied by at least one of the studied NAS. The evaluated land-use metrics (potential proxies for propagule pressure) not only influenced introduction and spread, but also cumulative responses such as NAS richness; NAS richness near cities increased as a function of city size, and the rate of increase was fairly constant throughout the study time period. City population size, total commercial vessel trips, and marina size may be reasonable proxies for propagule pressure given the significant relationships between these specific pathway metrics and NAS sightings. I will discuss the results of this study as they apply to the differences in primary and secondary spread patterns of NAS in the Great Lakes, NAS richness and composition, and the development of an early detection marina monitoring program in Lake Superior.

Primary Spread – First Sightings

All three pathway metrics were significant factors of first NAS sighting patterns. Both commercial boat traffic and city population size had similar odds ratios, suggesting that there was an equal association of NAS first sightings with each metric. This near equal split of association with each metric is consistent with the ballast water discharge mediated introduction history of NAS in the Laurentian Great Lakes where 65% of NAS introductions are attributed to ballast water discharge (Mills 1993, Ricciardi 2006). Additionally, ~50% of NAS included in this study were introduced through ballast water discharge and not the organisms in trade or recreational traffic pathways (USGS 2014), suggesting that the land-use metrics reasonably explained the pattern of first sightings in the Laurentian Great Lakes.

Secondary Spread and Change over Time

Unlike the results for the primary spread model, spatio-temporal patterns of species presence (primary and secondary spread) did not align well with the anthropogenic history of introductions in the Great Lakes. The strong association between species presence and city size is disproportionate to the prominence of the introduction history of the maritime commerce pathway, suggesting that introduction and detection locations were not always the same; secondary spread may have occurred rapidly, perhaps even before detection in many situations. Nonindigenous species populations often experience a lag phase in population growth during the establishment stage of invasions due to environmental limitations or low genetic diversity of a founder population leading to reduced fitness (Crooks 2005). Because of this lag phase, NAS may be present and undetected for years before their populations are large enough to be sighted by either scientific projects or causal observation (Beranek 2012). Before detection and during the lag phase, secondary spread from recreational and commercial boat traffic may occur that confounds clear associations between introduction and detection locations.

Further complicating the spatial associations between introduction and detection locations from the commercial boat traffic pathway is the physical influence of ballast discharge events. Ballast water can travel upwards of 7.5km in the first day after discharge (Wells et al. 2011), potentially carrying or influencing the movement of planktonic NAS away from introduction locations. Given the buffer size, dispersion may have moved NAS out of port APIs and into city APIs prior to detection, and thus the strong association with the live release pathway metric.

The very different life history strategies of Zebra Mussel and Round Goby has likely influenced their distribution patterns and is useful to help explain why commercial boat traffic was not the clear dominant factor for modeling introduction and spread. Ultimately, the Round Goby models appear to represent an introduction pattern that is strongly associated with its primary spread pathway, while the Zebra Mussel models are indicative of NAS introduction with heavy influence from secondary spread pathways.

The dramatic change in odds ratio (decreased by one half) for the city size metric in the Zebra Mussel multivariate model is evidence of the rapid increase in sightings of this species since its introduction. The spread of Zebra Mussel in the 1990s may be explained by Zebra Mussel’s high fecundity, the presence of a free-swimming larval stage (veliger) (Mackie and Schlosser 1996), and the abundance of secondary spread vectors (Rup et al. 2010).

Alternatively, Round Goby, also a ballast-introduced organism, engages in little natural dispersal as individuals remain in the same small area throughout their adult lives (Ray and Corkum 2001). Larval Round Goby however, display a diurnal vertical migration behavior where the larvae move up the water column during the night to feed, and return to the depths during the day to avoid predation (Hensler and Jude 2007). This local migratory behavior may have facilitated the spread of Round Goby throughout the Great Lakes via the secondary spread pathway of domestic ballast water discharge (Hensler and Jude 2007).

Once discharged, Round Goby larvae are likely to swim to the lake bottom, whereas Zebra Mussel veligers can drift with discharged ballast water away from their introduction location. The Round Goby's behavior may promote a stronger spatial association with initial introduction pathways than with secondary spread pathways, and thus introduction and detection locations are the same for this species. This may be the reason for the large port odds ratio in the Round Goby multivariate logistic regression models, as this species represents the direct relationship between introduction and detection locations.

NAS sighting records demonstrate that the Laurentian Great Lakes have become a highly invaded system. By the end of the 1990s, there was a 40% chance or better of finding a NAS near any sized city; by the end of 2013, the probability of sighting at least one NAS (of those examined) was greater than 70%. In multivariate models, change in odds ratio can be interpreted as the variability inherent in a study ecosystem. For commercial boat traffic and city population size, the decrease in the change in odds ratio values from the first decade of an invasion to the present and the decrease in the pathway intensity associated with the p(50), implies a homogenization of the probability of presence as the system became more invaded by NAS. For all the NAS or NAS groups, except Quagga Mussel, the declining odds ratios and the pathway intensity at p(50) in ports and cities indicates the continuing spread throughout medium and small land-use areas. Quagga Mussel near cities had the opposite trend, where the odds ratio increased over time.

Quagga Mussel, which has a very similar life history to Zebra Mussel (Mills et al. 1996), did not have the population and range expansion explosion soon after its introduction in the 1990s that marked the Zebra Mussel invasion, nor was there any significant relationships with the marina pathway. Mills et al. (1996, 1999) demonstrated that although Quagga Mussel seems to occupy a similar native niche as Zebra Mussel, based on their native ranges, there is a spatial dichotomy in the pattern of dreissenid domination in the Laurentian Great Lakes, where co-occupation of a single location is often not observed; this may be due to Quagga Mussel replacing the Zebra Mussel in many locations (Wilson et al. 2006). This transition from Zebra to Quagga Mussel may be occurring more rapidly in large cities where propagule pressure is highest; the increased competition may cause the local extirpation of Zebra Mussel in favor of Quagga Mussel. The change over time in the odds ratio for Quagga Mussel in the multivariate logistic regression models is indicative of more sightings in larger cities. The absence of any significant relationship with marina size may be due to a limitation of suitable habitat in marinas because Quagga Mussel prefer deeper waters where wave action is absent (Spidle et al. 1995).

Similar to the Quagga Mussel univariate logistic regression models, the aquaria species models suggest that aquaria species may be an emerging concern in large and medium cities. The relationship between city size and aquaria species presence developed only in the 1990s and has increased in magnitude since. The potential of this group to be introduced and spread is disproportionally based in large city (greater propagule pressure). Especially in the context of the other pathways (multivariate models), aquaria species have the greatest potential for spread of all the NAS in the current decade, having the highest odds ratio value for the 2010s. This finding is consistent with current literature documenting the growing participation in the aquaria trade, and the importance of this under-regulated pathway to the overall NAS burden in the Great Lakes (Padilla and Williams 2004, Rixon et al. 2005, Duggan et al. 2006).

Invasibility - Richness and Composition

NAS richness may provide more information than presence/absence because it may represent a cumulative response to propagule pressure over time. The shape of the relationship between pathway intensity and NAS richness might provide information regarding the trajectory of NAS invasions in each of the pathway locations in the Great Lakes. For example, a non-linear correlation that approaches a horizontal asymptote suggests that a NAS carrying capacity has been reached. In contrast, a positive exponential function suggests that additional factors beyond pathway intensity influenced NAS richness, as in the facilitated invasion hypothesis of Ricciardi (2001), which theorizes that past invasions facilitate future introductions through the alteration of ecosystems to favor new invaders. For example, the Zebra Mussel introduction to the Great Lakes in the mid-1980s may have provided an abundant and familiar food source for Round Goby, which followed in the early 1990s (Ricciardi and MacIsaac 2000). Further, an exponential relationship suggests more invasions are yet to come in the Great Lakes. Alternatively, a linear relationship suggests that invasion rates are constant and scale in proportion to pathway intensity, and thus pathway intensity is an important component of invasibility.

Ricciardi (2001) demonstrated that the cumulative introductions of new NAS into the Great Lakes increased logarithmically as a function of time from 1810 to 1999. My results suggest a change in trajectory for the more recent decades given the best-fit function for NAS richness versus city size was a linear relationship, not an exponential one. The multivariate linear regression analysis also supports a linear relationship because of the roughly constant rate increase over time.

Of the three pathway metrics, city population size had the greatest, positive association with NAS richness. Propagule pressure has been indicated as a factor of establishment success (Lockwood et al. 2009), but my results demonstrate an association with NAS richness as well, indicating that propagule pressure has a cumulative effect. Copp et al. (2010) came to a similar conclusion based on a study of population density and nonindigenous fish species richness in England.

Unlike the city population metric, which was the best factor of those tested for NAS richness ($R^2 = 0.39$), the number of boat slips per marina explained very little of the pattern of NAS richness in marinas ($R^2 = 0.1$). Many large marinas had 1 or 0 sightings, which may have influenced any potential relationship between NAS richness and marina size. This may be the result of a pseudo-absence error in the sightings dataset originating from an observational bias against monitoring in marinas, as exemplified by the lack of published work about NAS in Laurentian Great Lakes' marinas. However, given that significant relationships in both univariate and multivariate models were abundant and the univariate logistic regression models were consistent with the other pathways (i.e. increased probability of presence with increasing size), the consideration of marina size to influence presence, and potentially richness, should not be discarded. Ultimately, more marina monitoring projects are needed.

Alone, commercial boat traffic was not a significant factor of NAS richness within ports. However, when modeled in a multivariate linear regression for NAS richness, the influence of commercial boat traffic was roughly equal to the odds ratio of city size, if not slightly larger. This suggests that busy ports near large cities have the highest NAS richness throughout the Great Lakes.

The NMDS ordination suggests that the primary factor (NMDS axis 1) for NAS composition in large cities is a geographic gradient between the lower and upper Great Lakes. Bloody Red Shrimp, Quagga Mussel, and Fishhook Waterflea defined one end of NMDS axis 1; these species are abundant in the lower Great Lakes, and less so in Lake Superior, Lake Michigan, and Lake Huron (Mills et al. 1999, Grigorovich 2008, Marty et al. 2010, USGS 2014). Alternatively, Spiny Waterflea and Common Carp defined the other end of NMDS axis 1. Spiny Waterflea (with a preference for cooler water temperatures) (Garton et al. 1990), and Common Carp, are mainly found in the upper Great Lakes (USGS 2014); high-density populations of Spiny waterflea have been noted in Lake Erie as well (Brown and Branstrator 2004), and Common Carp is widely distributed throughout the tributaries of the Great Lakes (Bailey and Smith 1981), but sightings in the Great Lakes proper are mainly in western Lake Michigan and Lake Superior (USGS 2014). The alignment of NAS along an abiotic gradient suggests that environmental characteristics did impact NAS sightings patterns in the Great Lakes. Additionally, when each point in the NMDS was labeled by lake and designated as upper or lower lake (Figure 11B), the lower lakes (Erie and Ontario) points were roughly clustered together in the middle with the upper lakes around the exterior. Compositions in Lake Michigan were divided into two distinct groups at either end of axis 1 suggesting a potential variation in northern and southern distributions. This pattern in the NMDS further supports the presence of a geographic influence on the compositions of NAS in the Great Lakes. Future studies should include measurements of ecosystem parameters along with metrics of pathway intensity to improve overall model fit.

NMDS Axis 2 was defined by three NAS associated with the live-release pathway (Goldfish, New Zealand Mud Snail, and Scud). This further indicates the importance of large population centers as introduction pathways for live release organisms (Rixon et al. 2005, Duggan et al. 2006). Attempts to further classify the geospatial patterns and life history traits of these three organisms did not produce any additional explanation. Ultimately, these species distributions may be similar only in their dissimilarity to other NAS distributions as a result of very strong associations with large cities. No significant trends were observed in small cities primarily because of the lack of NAS richness and the high variability of specific NAS presence at any given location caused by limited sightings for over half of the NAS.

Land-use Metrics as Proxies for Propagule Pressure

Overall, city population size was the best model factor of NAS presence. This metric was a significant factor for all NAS during the study time period (1970s – 2013) for both first sightings and spread. This result is complementary to previous studies that found population density was a significant factor of terrestrial and aquatic nonindigenous species presence and richness (Gilbert et al. 2004, Duggan et al. 2006, Copp et al. 2010, Spear et al. 2013).

All three metrics of commercial boat traffic had highly significant ($\alpha < 0.001$), linear relationships with one another (total vessel trips and average annual cargo: $R^2 = 0.77$, total trips and annual average ballast discharge: $R^2 = 0.49$, and ballast discharge and cargo: $R^2 = 0.41$). This explains the similarities between the multivariate logistic regression model fits when anyone of the three commercial boat traffic metrics was included, but it does not explain why total trips provided the best fits or why it was the only commercial boat traffic metric that produced a significant univariate logistic regression model.

The weak relationship between total trips and ballast discharge volume ($R^2 = 0.49$) may be due to vessels that claim no-ballast-on-board (NOBOB) status; NOBOB vessels do not have ballast water to discharge. However, many NOBOB vessels have residual amounts of ballast water and sediment in their tanks from previous trips that can be unintentionally discharged when ballast water is taken on in Great Lake ports (Doblin et al. 2001); these residuals may harbor viable NAS resting eggs (Bailey et al. 2005). As a metric, total vessel trips captures the arrival of NOBOB vessels which may explain the significance of this metric as compared to ballast water volume, which does not account for NOBOB vessels. Additionally, the significance of the non-discharging vessel trips metric may suggest that NOBOB vessels, which do not officially discharge, are a major factor of the commercial boat traffic pathway, and thus are of significant concern for stopping future NAS introductions.

Unlike the commercial boat traffic metric, marina size was a significant factor in both multivariate and univariate logistic regression models. Large marinas had a greater probability of NAS sightings than smaller marinas when modeled independently (univariate), but when analyzed in the multivariate models, increasing size resulted in less than expected NAS presence in a marina. In light of the dearth of published research on NAS presence in marinas globally, and specifically in the Laurentian Great Lakes, I propose two hypotheses to explain these results.

First, there was a statistical detection limit in small marinas when considering the influence of cities and ports. In smaller marinas, where species richness was often low or zero, there was no change in the probability of presence with regard to surrounding city or port size. Conversely, in larger marinas, there was a detectable decrease in the probability of NAS presence as compared to the surround cities and ports. Thus, it is an issue of signal detection, where a true signal is only detectable in large marinas and undetectable in small marinas where there is not enough information given the complete absence of species presence in many small marinas. This limit in signal detection produces the negative correlation present in all the multivariate logistic regression models.

A second hypothesis is that ecological conditions are less favorable in large marinas than in small marinas due to scaling anthropogenic alteration, and thus habitat suitability is lowest in highly modified marinas. Recent rapid marina-based NAS monitoring projects in England and Scotland indicate that at least for marine NAS, marinas can provide adequate habitat for NAS (Arenas et al. 2006, Ashton et al. 2006), though size was not a consideration in those studies. However, given the significant results of the univariate logistic regression marina models, this hypothesis is not supported. Ultimately, more marina specific monitoring for NAS in the Great Lakes is needed.

Management Application – Marina Monitoring

The development of a monitoring program focused on Lake Superior marinas, as opposed to the ports and cities around marinas, would result in a 12% average reduction in the probability of a historic NAS sighting, suggesting that marinas are not the best locations to search for new NAS introductions in Lake Superior. However, the application of my findings to develop a marina monitoring project was inconclusive about the role of marinas as locations of introduction and spread, and future dedicated monitoring of marinas is warranted because of the strong univariate relationships. Additionally, given the documented presence of over 20 marine NAS in English and Scottish marinas (Arenas et al. 2006, Ashton et al. 2006), the potential for rapid colonization in Australian marinas (Bax et al. 2002), and the potential of the recreational pathway to spread NAS (Leung et al. 2004), marinas should not be overlooked.

Ultimately, projects like the one proposed by the U.S. Fish and Wildlife Service to monitor marinas for new NAS introductions, are needed to understand the importance of these locations as pathways for introduction and spread.

Conclusion

Development and evaluation of land-use metrics to explain historic NAS sightings patterns may help inform future actions to prevent or slow new NAS invasions. In the Laurentian Great Lakes, probability of historic NAS sightings was greatest in large cities, ports, and marinas, and has increased as a function of time in medium and small pathway locations. City population size had a strong univariate and multivariate association between all the NAS studied and first sightings, NAS presence, and NAS richness. Commercial boat traffic was an important factor for first sightings and less so in multivariate models, and may have been underrepresented as a factor due to confounding sightings patterns caused by rapid secondary spread and natural dispersal. The recreational pathway (marinas) also had significant relationships with most NAS studied, though multivariate logistic regression models were inconclusive for this metric potentially due to a statistical detection limit. Marinas in the Great Lakes, as centers of introduction and spread, need to be studied further. The historically increasing importance of the live release pathway is notable given that human populations (proxy for the live release pathway) continue to increase as well. Additionally, the detection of a broad geographic influence on species composition near large cities suggests that future studies should include environmental factors as variables to better understand invasion patterns.

Modeling propagule pressure by developing land-use metrics as proxies, succeed in finding large-scale spatial and temporal patterns associated with NAS sightings. This study highlighted the interplay between city population size and commercial boat traffic as factors to inform early detection efforts, and quantified and confirmed some of the working assumptions about NAS introduction and spread patterns in the Laurentian Great Lakes, specifically, that NAS sightings are strongly associated with various forms of human activity.

Table 1 – Spatial data and pathway attributes associated with the three pathways evaluated. When categorizing the metrics, domestic was defined as both U.S. and Canadian ports, while foreign included any trip, cargo, or ballast water that passed through waters greater than 200 nautical miles from the U.S. or Canadian landmasses. Unique locations are the number of spatially distinct sites with data; locations were combined based on the distance between sites (for ports, if two ports were within 10km they were merged into one site with the sum of both attributes; marinas were combined if within 1km; city locations were not combined).

Pathway	Spatial Data (metric)	Unique Locations	Locations after Spatial Combination	Data Time Range	Source
Commercial Boat Traffic	Commercial Vessel Trips - Total - Discharging - Non-discharging	67	50	2004 - 2013	National Ballast Water Clearinghouse (2015)
	Annual Avg. Cargo (metric tons) - Total - Domestic - Foreign	45	34	1996 – 2011	U.S. Army Corps of Engineers (2014)
	Annual Avg. Ballast water discharge (m^3) - Total - From highly invaded domestic ports	97	65	2005 – 2007	National Ballast Water Clearinghouse (2015), Bailey et al. (2012)
Recreational Boat Traffic	Marina size (boat slips)	516	304	2010	Allan et al. (2013)
Live Release from Urban Areas	Population size (full time residents)	116	116	1970 – 2010	U.S. Census Bureau (2010), Statistics Canada (2010)

Table 2 – Number of species sightings within 10km of shore from 1970 to 2014. Lake identifies recorded species presence by lake. Lake abbreviations: S - Lake Superior; H - Lake Huron; M - Lake Michigan; E - Lake Erie; O - Lake Ontario. Model abbreviations: AS – All species model; IND – Individual species model; AQ – Aquaria species model; FS – First sightings model. Regression analyses were conducted using subsets of this dataset in various models: three individual species models, two group species models, and one first sightings model.

Scientific name	Common name or description	Number of Records	Year Range	Lake	Regression Model
Dreissena polymorpha	Zebra Mussel	1088	1986 – 2013	S,H,M,E,O	AS, IND, FS
Neogobius melanostomus	Round Goby	670	1990 – 2013	S,H,M,E,O	AS, IND, FS
Cyprinus carpio	Common Carp	350	1971 – 2013	S,H,M,E,O	AS, FS
Gymnocephalus cernua	Eurasian Ruffe	272	1987 – 2013	S,H,M	AS, FS
Dreissena rostriformis bugensis	Quagga Mussel	243	1989 – 2011	S,H,M,E,O	AS, IND, FS
Bythotrephes longimanus	Spiny Waterflea	117	1984 – 2010	S,H,M,E,O	AS, FS
Carassius auratus	Goldfish	85	1975 - 2013	S,H,M,E,O	AS, AQ, FS
Bosmina coregoni	Waterflea	71	1994 - 2006	S,H,M,E,O	AS, FS
Hemimysis anomala	Bloody Red Shrimp	59	2006 - 2011	H,M,E,O	AS, FS
Echinogammarus ischnus	Scud	57	1994 - 2008	S,H,M,E,O	AS, FS
Eurytemora affinis	Calanoid copepod	56	1994 - 2006	S,H,M,E,O	AS, FS
Cercopagis pengoi	Fishhook Waterflea	52	1998 - 2006	S,H,M,E,O	AS, FS
Orconectes rusticus	Rusty Crayfish	47	1981 - 2011	S,H,M	AS, FS
Potamopyrgus antipodarum	New Zealand Mud Snail	40	1991 - 2008	S,H,E,O	AS, FS
Bithynia tentaculata	Faucet Snail	35	1974 - 2012	S,H,M,E,O	AS, FS
Craspedacusta sowerbyi	Freshwater Jellyfish	28	1971 - 2003	H,M,E	AS, AQ, FS
Corbicula fluminea	Asian Clam	14	1980 - 2012	S,H,M,E	AS, FS
Daphnia lumholtzi	Waterflea	10	1999 - 2007	S,H,E	AS, FS
Cipangopaludina chinensis	Chinese Mystery Snail	9	1974 - 2008	H,M,E,O	AQ
Radix auricularia	European Ear Snail	5	1974	M,H,E	AQ
Cordylophora caspia	Freshwater hydroid	2	2001 - 2007	S, M	AQ
Lophopodella carteri	Freshwater bryozoan	1	1995	M	AQ
Colossoma macropomum	Pacu	1	2014	E	AQ
Total		3294			

Table 3 – Model fit (-2*log(likelihood)), p(50), and significance for univariate logistic regression models. Values of -2*log(likelihood) closer to ‘0’ indicate a better fitting model. The p(50) is the pathway intensity where the probability of presence is 0.5. The p(50) values are the exponent on a logarithmic scale, i.e. 2.65 = $10^{2.65}$.; pathway intensity ranges: port $10^{0.3} – 10^{3.9}$; city $10^{3.4} – 10^{6.9}$; marina $10^{1} – 10^{3.4}$. The p-value was calculated using the likelihood ratio test statistic to test for significance. Significance was determined at $\alpha = 0.01$. ‘-’ denotes that the relationship was not significant. ‘+’ denotes that the relationship was significant, but the regression curve did not include 0.5.

Pathway		-2*log(likelihood)					p(50)					p-value				
Metric	Model	1970s	1980s	1990s	2000s	2010s	1970s	1980s	1990s	2000s	2010s	1970s	1980s	1990s	2000s	2010s
Commercial Boat Traffic	All Species	-	62.8	-	-	-	-	2.85	-	-	-	0.220	0.010	0.614	0.614	0.614
	Aquaria Species		-	-	-	-		-	-	-	-		0.370	0.532	0.308	0.169
	Quagga Mussel			-	-	-			-	-	-			0.717	0.562	0.335
	Round Goby			63.7	46.8	39.4			2.75	1.4	1.35			0.010	0.009	0.005
	Zebra Mussel		-	-	-	-		-	-	-	-		0.051	0.913	0.913	0.878
Population Size	All Species	50.8	137.5	83.3	67.6	61.7	6.15	5.15	3.4	+	+	<0.001	<0.001	<0.001	0.001	0.007
	Aquaria Species		-	108.1	117.3	123.9		-	6.6	5.7	5.4		0.173	0.009	<0.001	<0.001
	Quagga Mussel			122.1	131.3	135.0			6	4.9	4.9			0.004	<0.001	<0.001
	Round Goby			133.0	157.6	146.2			5.3	3.8	3.4			<0.001	0.009	0.001
	Zebra Mussel		89.6	93.6	93.0	90.4		5.7	3.5	3.5	3.5		<0.001	<0.001	<0.001	<0.001
Recreational boat traffic	All Species	-	147.4	369.0	384.4	393.3	-	+	2.7	2.4	2.3	0.100	0.001	<0.001	<0.001	<0.001
	Aquaria Species		-	-	-	-		-	-	-	-		0.230	0.085	0.052	0.011
	Quagga Mussel			-	-	-			-	-	-			0.103	0.179	0.134
	Round Goby			152.4	220.3	261.8			+	+	+			<0.001	<0.001	<0.001
	Zebra Mussel		101.1	329.4	330.4	344.4		+	2.9	2.9	2.9		<0.001	<0.001	<0.001	<0.001

Table 4 – Comparison of commercial boat traffic univariate logistic regression models. Significance was determined at α = 0.01. There were no significant relationships in the aquaria species models (omitted). Model fit values (-2*log(likelihood)) are displayed for significant results (*); p-values are displayed for non-significant results. 'NV' indicates that the analysis had no variation. Trips were categorized as discharging and non-discharging ballast water. Cargo was categorized into domestic (U.S. and Canada) and foreign origins. Highly invaded ports were defined as the top quartile ports for NAS richness (>7 species).

A. Commercial Vessel Trips	Total					Discharging					Non-discharging				
Model	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s
All Species	0.220	69*	0.464	NV	NV	0.677	0.990	0.750	NV	NV	0.020	51*	0.230	NV	NV
Quagga Mussel			0.710	0.560	0.330			0.680	0.700	0.220			0.605	0.030	0.080
Round Goby			64*	47*	39*			0.050	0.075	0.011			62*	43*	0.011
Zebra Mussel		0.051	0.910	0.913	0.878		0.228	0.328	0.328	0.474		52*	0.072	0.072	0.172

B. Commercial Cargo	Total					Domestic					Foreign				
Model	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s
All Species	0.680	0.431	0.550	NV	NV	0.100	0.429	0.510	NV	NV	0.460	0.440	0.300	NV	NV
Quagga Mussel			0.410	0.139	0.120			0.305	0.329	0.357			0.670	0.043	0.183
Round Goby			0.560	0.480	0.098			0.298	0.303	0.342			0.755	0.213	0.450
Zebra Mussel		0.200	0.170	0.230	0.249		0.314	0.658	0.658	0.768		0.570	0.850	0.250	0.130

C. Volume of Ballast Water	Total					Highly invaded domestic ports				
Model	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s
All Species	0.640	0.980	0.043	NV	NV	0.560	0.050	0.131	NV	NV
Quagga Mussel			0.760	0.490	0.910			0.190	0.210	0.520
Round Goby			0.044	0.293	0.123			0.017	0.288	0.154
Zebra Mussel		0.410	0.120	0.120	0.310		0.100	0.660	0.660	0.540

Table 5 – Comparison of commercial boat traffic metrics in the multivariate logistic regression models. The significance and overall model fit was determined for each of the commercial boat traffic metrics. Only commercial boat traffic metric results are displayed. Significance was determined at α = 0.01. There were no significant relationships in the aquaria species models (omitted). Model fit values (-2*log(likelihood)) are displayed for significant results (*); p-values are displayed for non-significant results. Trips were categorized as discharging and non-discharging ballast water. Cargo was categorized into domestic (U.S. and Canada) and foreign origins. Highly invaded ports were defined as the top quartile ports for NAS richness (>7 species).

A. Commercial Vessel Trips	Total					Discharging					Non-discharging				
Model	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s
All Species	0.086	339*	632*	610*	613*	0.145	341*	638*	0.055	0.014	0.086	338*	629*	607*	610*
Quagga Mussel			0.443	0.900	0.310			0.509	0.944	0.379			0.400	0.800	0.330
Round Goby			324*	473*	514*			326*	480*	520*			324*	471*	515*
Zebra Mussel		210*	615*	612*	621*		215*	620*	617*	625*		216*	615*	612*	620*

B. Commercial Cargo	Total					Domestic					Foreign				
Model	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s
All Species	0.048	337*	637*	0.100	0.035	0.052	337*	637*	0.094	0.033	109*	337*	0.010	0.142	0.047
Quagga Mussel			0.778	0.935	0.460			0.759	0.924	0.462			0.864	0.974	0.391
Round Goby			326*	488*	530*			326*	488*	529*			321*	485*	527*
Zebra Mussel		212*	0.080	0.054	626*		212*	0.075	0.051	627*		213*	0.082	0.059	626

C. Volume of Ballast Water	Total					Highly invaded domestic ports				
Model	70s	80s	90s	00s	10s	70s	80s	90s	00s	10s
All Species	0.034	0.019	664*	0.044	0.017	115*	361*	0.015	0.145	0.037
Quagga Mussel			0.300	0.433	0.151			0.542	0.574	0.840
Round Goby			334*	489*	533*			333*	489*	528*
Zebra Mussel		0.013	0.013	0.012	623*		211*	0.180	0.149	0.025

Table 6 – Bivariate logistic regression model results including, odds ratio, model fit (-2*log(likelihood) (-2LL)), and significance for each combination of commercial boat traffic (port), population size (city), and recreational boat traffic (marina) metrics; the significance of each metric within the model was calculated and assessed as an odds ratio. Odds ratios were calculated as percent change for each bivariate model regression coefficient. The 1st decade varied by species and was based on the year of the first sightings record. Values of -2LL closer to ‘0’ indicate a better fitting model. Significance was determined for each variable at $\alpha = 0.01$ (‘-’ denotes that the relationship was not significant).

	1st Decade Odds Ratio					2010s Odds Ratio				
Model	Port	City	-2LL	p-value		Port	City	-2LL	p-value	
All Species	-	190	75	0.100	<0.001	81	54	155	0.008	0.007
Aquaria Species	-	91	117	0.357	0.005	-	122	168	0.091	<0.001
Quagga Mussel	-	-	-	0.573	0.012	-	71	195	0.319	0.001
Round Goby	74	-	178	<0.001	0.015	88	53	212	<0.001	<0.001
Zebra Mussel	-	123	112	0.024	<0.001	93	78	194	<0.001	<0.001
Model	Port	Marina	-2LL	p-value		Port	Marina	-2LL	p-value	
All Species	192	-	68	0.010	0.020	55	-	460	<0.001	0.190
Aquaria Species	128	-	92	0.002	0.110	83	-	195	<0.001	0.790
Quagga Mussel	-	-	-	0.863	0.075	-	-46	208	0.078	<0.001
Round Goby	87	-32	236	<0.001	0.010	80	-42	335	<0.001	<0.001
Zebra Mussel	87	-	156	<0.001	0.400	60	-	434	<0.001	0.397
Model	Marina	City	-2LL	p-value		Marina	City	-2LL	p-value	
All Species	-62	200	85	0.001	<0.001	-42	40	508	<0.001	<0.001
Aquaria Species	-	-	-	0.012	0.511	-47	55	252	<0.001	<0.001
Quagga Mussel	-50	65	221	<0.001	<0.001	-54	68	272	<0.001	<0.001
Round Goby	-36	79	291	<0.001	<0.001	-49	43	442	<0.001	<0.001
Zebra Mussel	-39	111	195	0.003	<0.001	-37	41	519	<0.001	<0.001

Table 7 – Regression coefficients from the multivariate linear regression of NAS richness by decade. ‘-’ denotes that the variable was not significant in the regression analysis. Significance was determined at α = 0.01.

	Regression Coefficient					p-value				
	1970s	1980s	1990s	2000s	2010s	1970s	1980s	1990s	2000s	2010s
Commercial boat traffic	-	0.084	0.182	0.236	0.298	0.034	<0.001	<0.001	<0.001	<0.001
Population size	-	0.043	0.143	0.228	0.224	0.014	<0.001	<0.001	<0.001	<0.001
Recreational boat traffic	-	-0.076	-0.293	-0.471	-0.548	0.011	<0.001	<0.001	<0.001	<0.001

Table 8 – Correlation values for nonmetric multidimensional scaling (NMDS) ordination of NAS composition for large population cities (>35,000). The correlation between NAS presence and both NMDS Axis 1 and 2 scores were calculated using Pearson's Correlation equation; values closet to 1 and -1 indicate the greatest correlations. The tables are sorted from largest to smallest values for Axis 1 and Axis 2 respectively. 'NA' indicates that there were no NMDS scores; *C. sowerbyi* -no sightings in large cities; *D. polymorpha* - every large city had a sighting.

Species	NMDS Axis 1 Correlation	Species	NMDS Axis 2 Correlation
Hemimysis anomala	0.49	*Carassius auratus*	0.69
Dreissena rostriformis	0.40	*Potamopyrgus antipodarum*	0.62
Cercopagis pengoi	0.36	*Echinogammarus ischnus*	0.61
Potamopyrgus antipodarum	0.24	*Cyprinus carpio*	0.54
Gymnocephalus cernua	0.10	*Neogobius melanostomus*	0.47
Echinogammarus ischnus	0.10	*Corbicula fluminea*	0.45
Neogobius melanostomus	0.02	*Gymnocephalus cernua*	0.40
Daphnia lumholtzi	-0.07	*Bosmina coregoni*	0.40
Corbicula fluminea	-0.18	*Bithynia tentaculata*	0.34
Bosmina coregoni	-0.18	*Dreissena rostriformis*	0.32
Carassius auratus	-0.19	*Daphnia lumholtzi*	0.32
Eurytemora affinis	-0.19	*Cercopagis pengoi*	0.30
Bithynia tentaculata	-0.20	*Eurytemora affinis*	0.24
Orconectes rusticus	-0.37	*Hemimysis anomala*	0.11
Cyprinus carpio	-0.51	*Orconectes rusticus*	0.06
Bythotrephes longimanus	-0.54	*Bythotrephes longimanus*	-0.03
Craspedacusta sowerbyi	NA	*Craspedacusta sowerbyi*	NA
Dreissena polymorpha	NA	*Dreissena polymorpha*	NA

Table 9 – Priority marina monitoring locations derived from multivariate logistic regression equation. For 31 sites in U.S. Lake Superior, the historic percent probability of any NAS sighting was calculated using the equations generated from the 2010s all species multivariate logistic regression model. To compare the difference between the probability of presence in marinas or in ports and cities, marina intensity was set to zero for all locations and the probability was calculated again using the regression equation. A priority monitoring list was generated by sorting the historic probability of presence from largest to smallest.

Pathway Location			Anthropogenic Activity by Pathway Location			Probability of any NAS Presence:	
City	Port	Marina	Population	Vessel Trips	Boat Slips	Marinas	Ports and Cities
Two Harbors, MN	Two Harbors		3,548	2494		0.86	0.86
Sault Ste Marie, MI-Ontario	Sault Ste. Marie		80,790	102		0.85	0.85
Duluth-Superior, MN-WI	Duluth-Superior	Spirit Lake Marina	117,489	8297	70	0.80	0.91
Marquette, MI	Presque Island-Marquette	Marina by ore Dock	26,303	4089	60	0.77	0.89
Duluth-Superior, MN-WI	Duluth-Superior	Harbor Cove Marina	117,489	8297	215	0.77	0.91
Marquette, MI	Presque Island-Marquette	Cinder Pond Marina	26,303	4089	101	0.75	0.89
Houghton, MI			15,136			0.74	0.74
Duluth-Superior, MN-WI	Duluth-Superior	Barker's Island Marina	117,489	8297	420	0.74	0.91
Taconite Harbor, MN	Taconite Harbor		200	46		0.73	0.73
Laurium, MI			7,244			0.72	0.72
Munising, MI			2,951			0.70	0.70
Grand Marais, MN			1,340			0.69	0.69
Silver Bay, MN	Silver Bay, MN	Silver Bay Marina	1,850	1499	108	0.66	0.84
Ontonagon, WI	Ontonagon	Ontonagon Village Marina	1,400	53	36	0.62	0.78
L'Anse, MI		L'Anse Marina	1,965		13	0.57	0.70
Ashland, WI	Ashland	Ashland Marina	7,413	20	150	0.57	0.79
Baraga, MI		Baraga Marina	2,000		24	0.54	0.70
Bayfield, WI		Bayfield City Dock	500		20	0.51	0.66
Copper Harbor, MI		Copper Harbor State Dock	110		10	0.51	0.63
Grand Portage, MN		Grand Portage Marina	565		30	0.50	0.67
Bayfield, WI		Schooner Bay Marina	500		30	0.49	0.66
Bayfield, WI		Roy's Pointe Marina	500		42	0.48	0.66
Bayfield, WI		Erickson's Marina	500		50	0.47	0.66
Skanee, MI		Witz Marina	450		60	0.46	0.66
Washburn, WI		Washburn Marina	2,100		139	0.45	0.70
Cornucopia, WI		Siskiwit Bay Marina	100		28	0.45	0.62
Bayfield, WI		Apostle Islands Marina	500		110	0.43	0.66
Port Wing, WI		Port Wing Marina	165		70	0.42	0.64
Knife River, MN		Knife River Marina	230		104	0.41	0.64
Bayfield, WI		Port Superior Marina	500		200	0.40	0.66
Bayfield, WI		Pike's Bay Marina	500		208	0.40	0.66

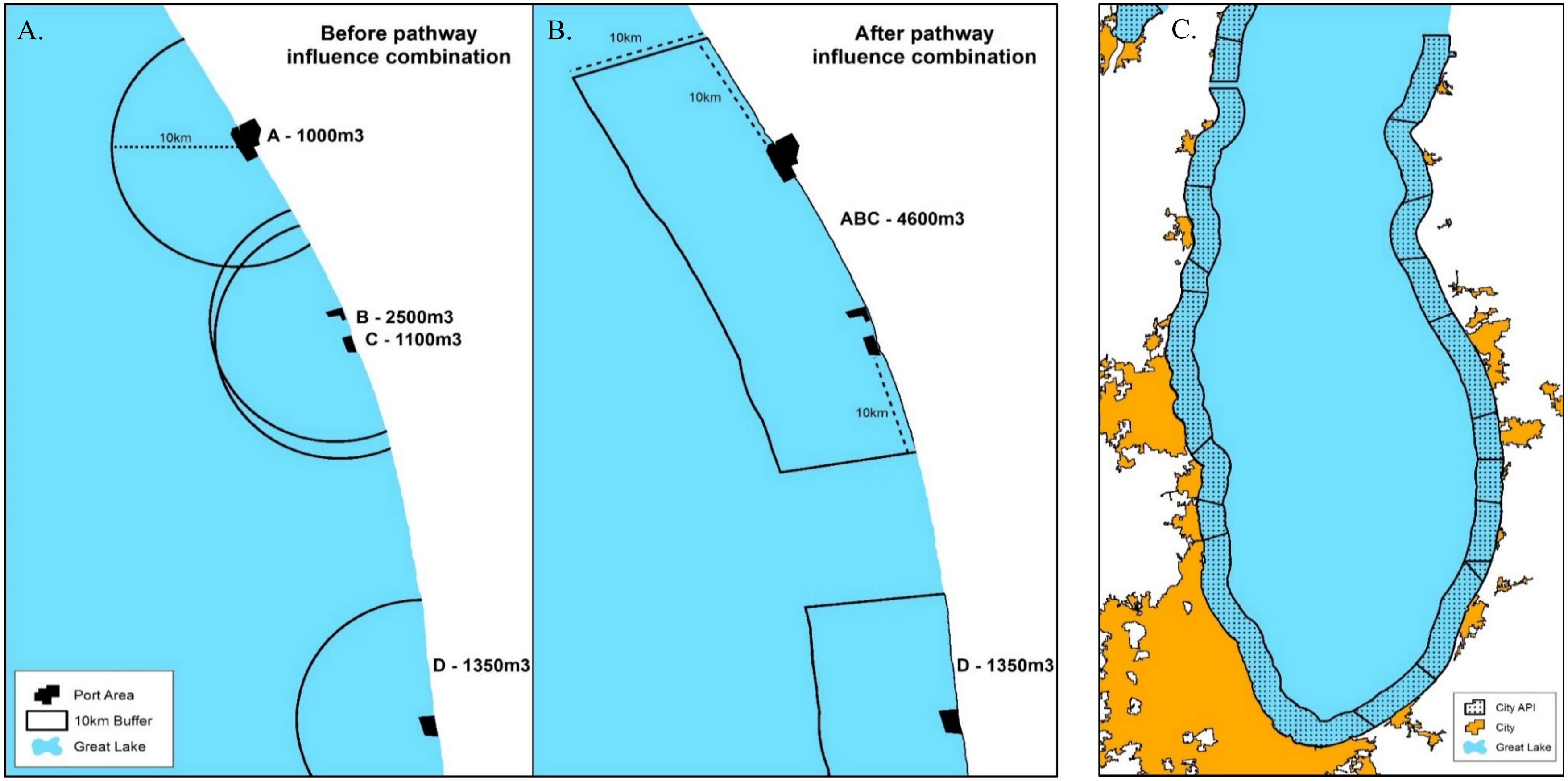


Figure 1 – Method for delineating areas of pathway influence (API). Port and marina pathway locations were merged based on buffer distances. Buffers were applied to pathway locations, and if over lap occurred (A), the areas of pathway influence (API) were merged, and all associated attributes of pathway intensity were summed together (B). The buffer distance for ports and cities was 10km; the marina buffer was 1km. To maintain discrete areas, cities APIs were not merged under any circumstances (C). If overlap would occur for two city API, the border between the two became the dividing line.

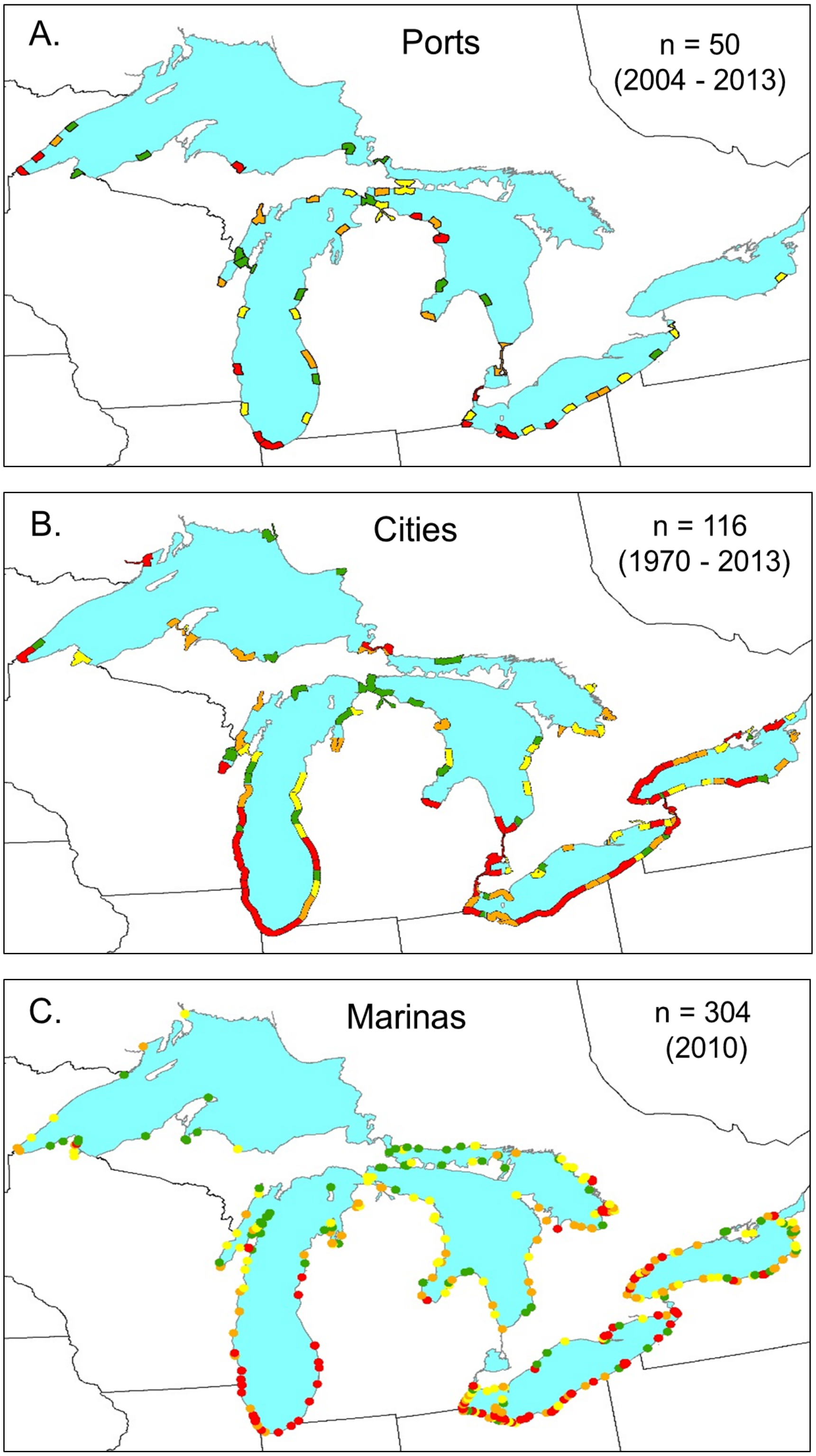


Figure 2 – Areas of pathway influence (API) for the three pathways. Pathway intensity metrics were assigned to each API, quartiles of the pathway intensity metrics are displayed by color: 1st – green, 2nd – yellow, 3rd – orange, 4th – red; display purposes only. Sample size and data range are displayed in the upper right of each panel.

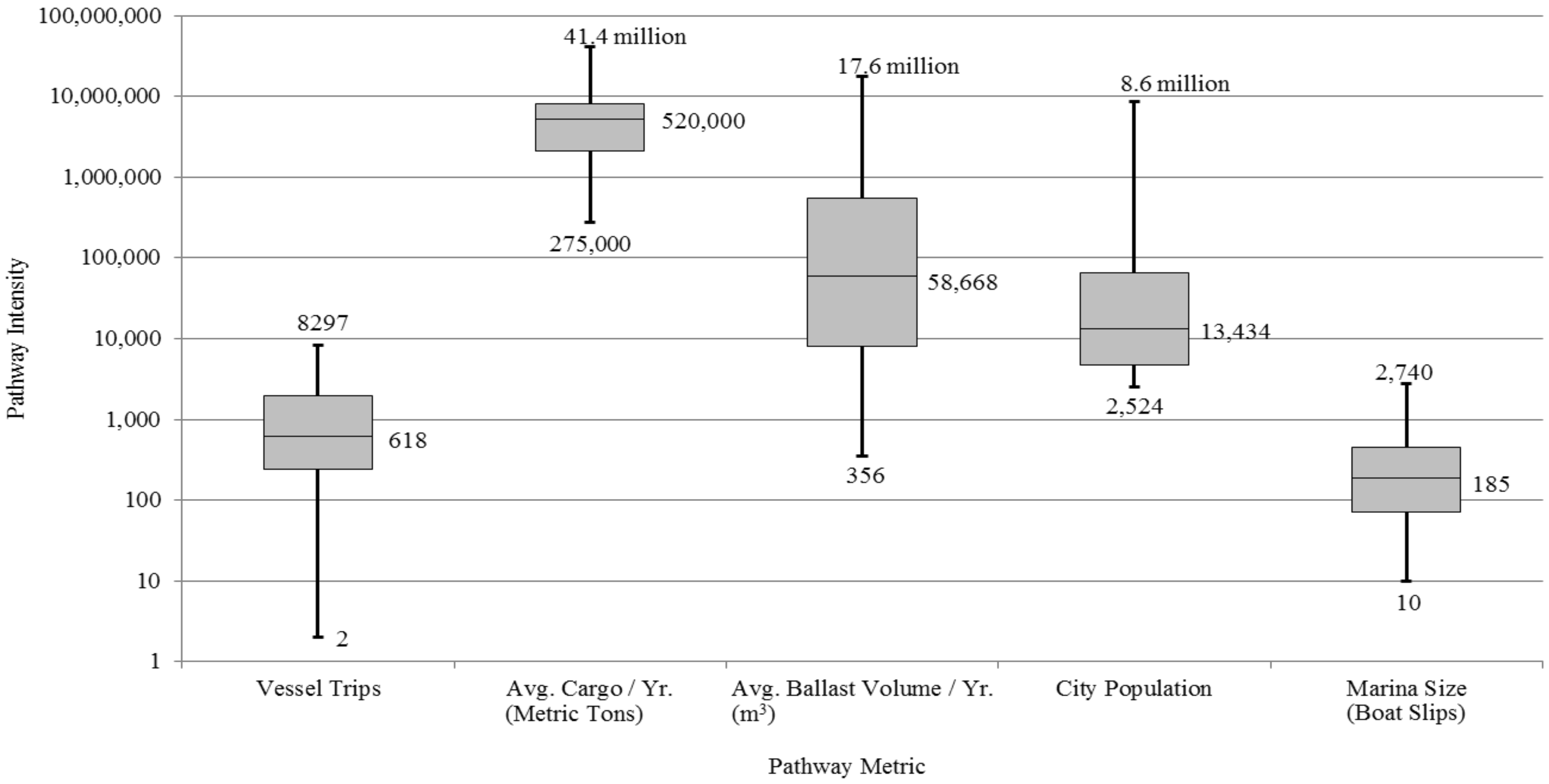


Figure 3 – Distribution of pathway intensity data for each of the three metrics of commercial boat traffic, and the organisms in trade and recreational boat traffic pathways. Data represent pathway intensity after API combinations. Minimum and maximum dataset values are represented by the bottom and top bars respectively; median values are displayed as the centerline in the boxes with the upper portion of the box defining the 3rd quartile, and the lower portion of the box defining the 2nd quartile. Above the box to the upper bar, and below the box to the lower bar are the 4th and 1st quartile respectively. Data sources provided in Table 1.

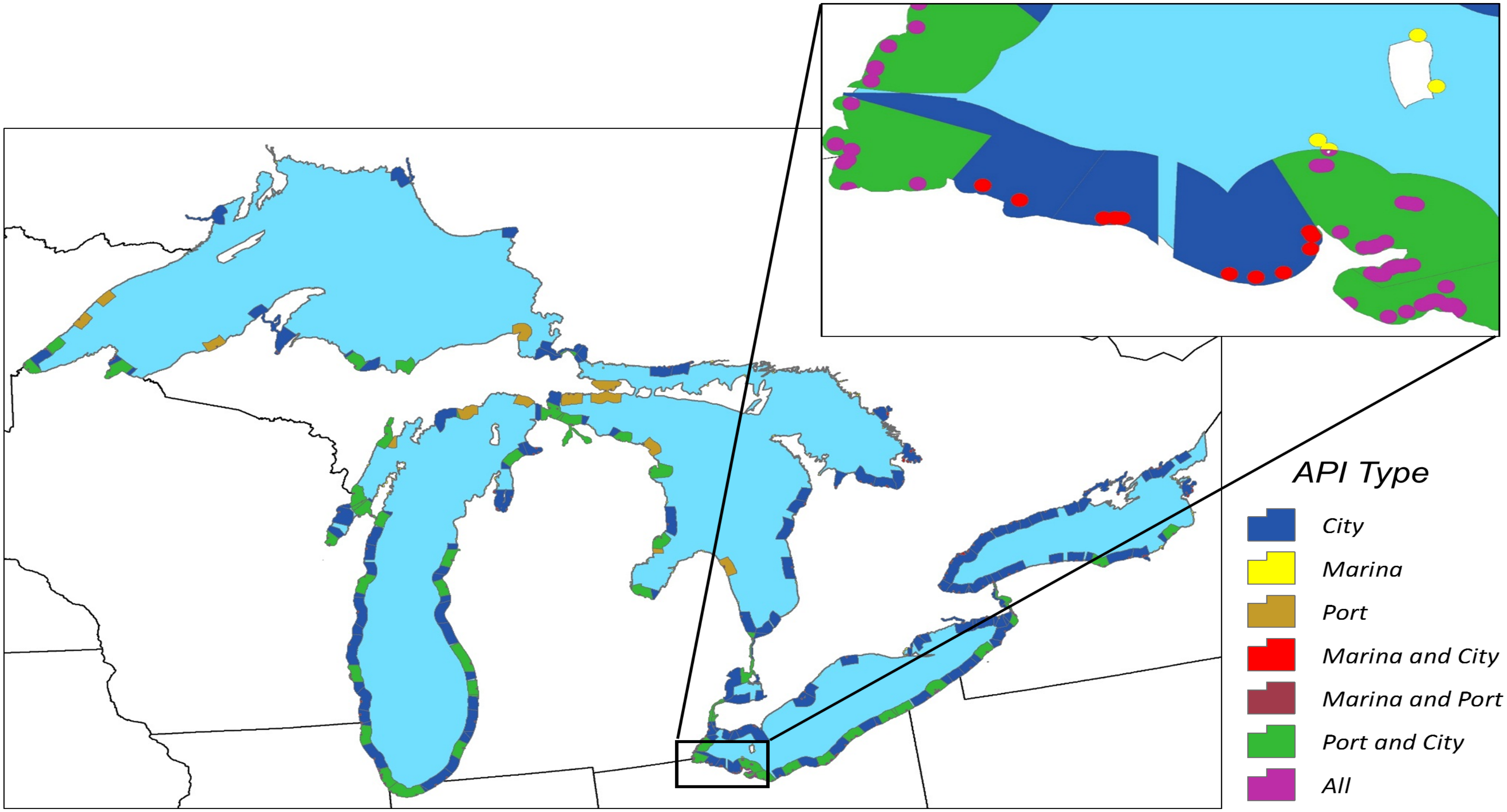


Figure 4 – Method for delineated multivariate APIs. For multivariate analyses, APIs had to be categorized into all possible combinations of the three pathway APIs. When spatial overlap of the individual APIs occurred, a new multivariable API was created with all the associated intensity metrics of the underlying APIs. If only two pathway APIs overlapped, the metric for the third pathway was assigned a zero value.

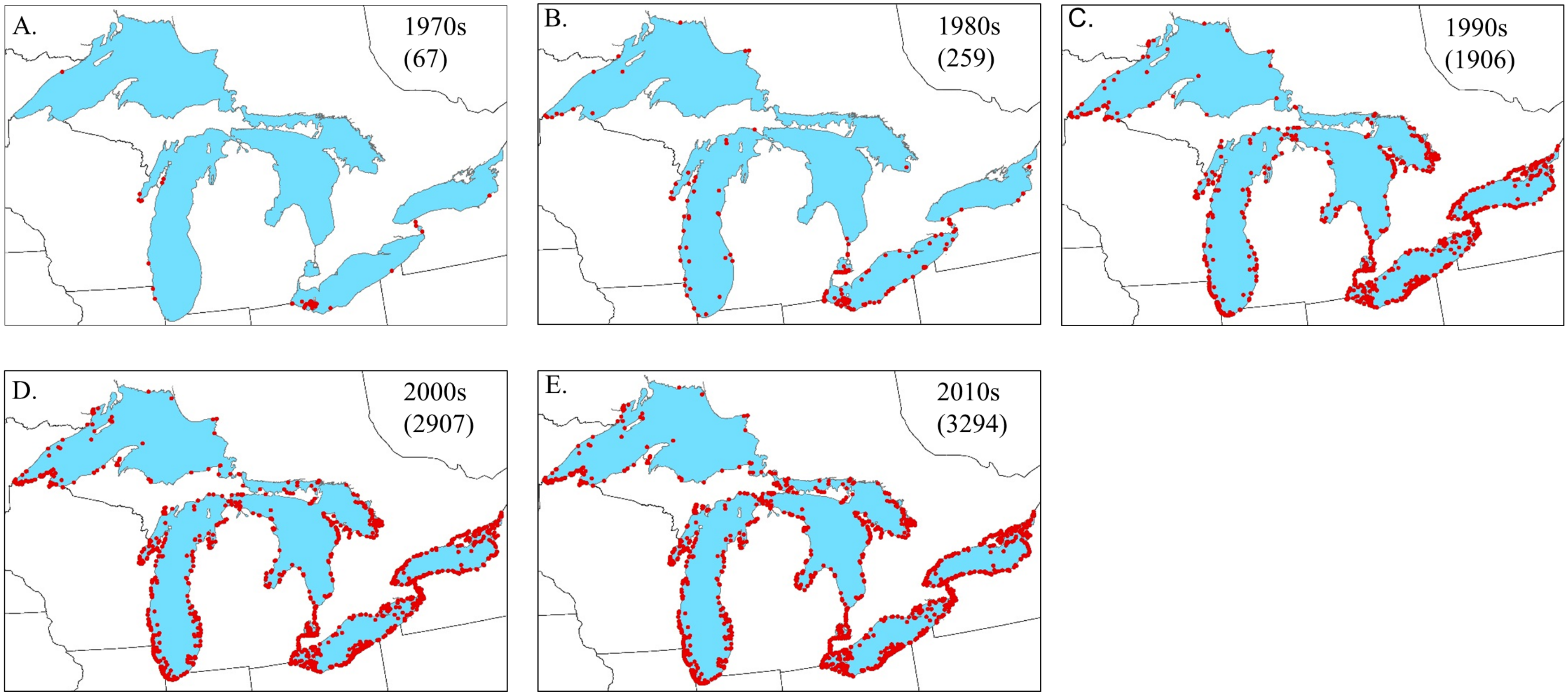


Figure 5 - Cumulative NAS sightings grouped by decade from 1970 to 2013. Total sightings per decade are displayed in the upper right corner of each panel. Each sighting is represented by a closed circle. Sightings were included in a decade if they occurred between the 1st of January and the 31st of December, i.e. sightings in the 1970s decade were from Jan. 1 1970 to Dec. 31 1979. A correction for pseudo-absence was applied where each species sighting was treated as if it persisted from one decade to the next from the earliest sighting onward.

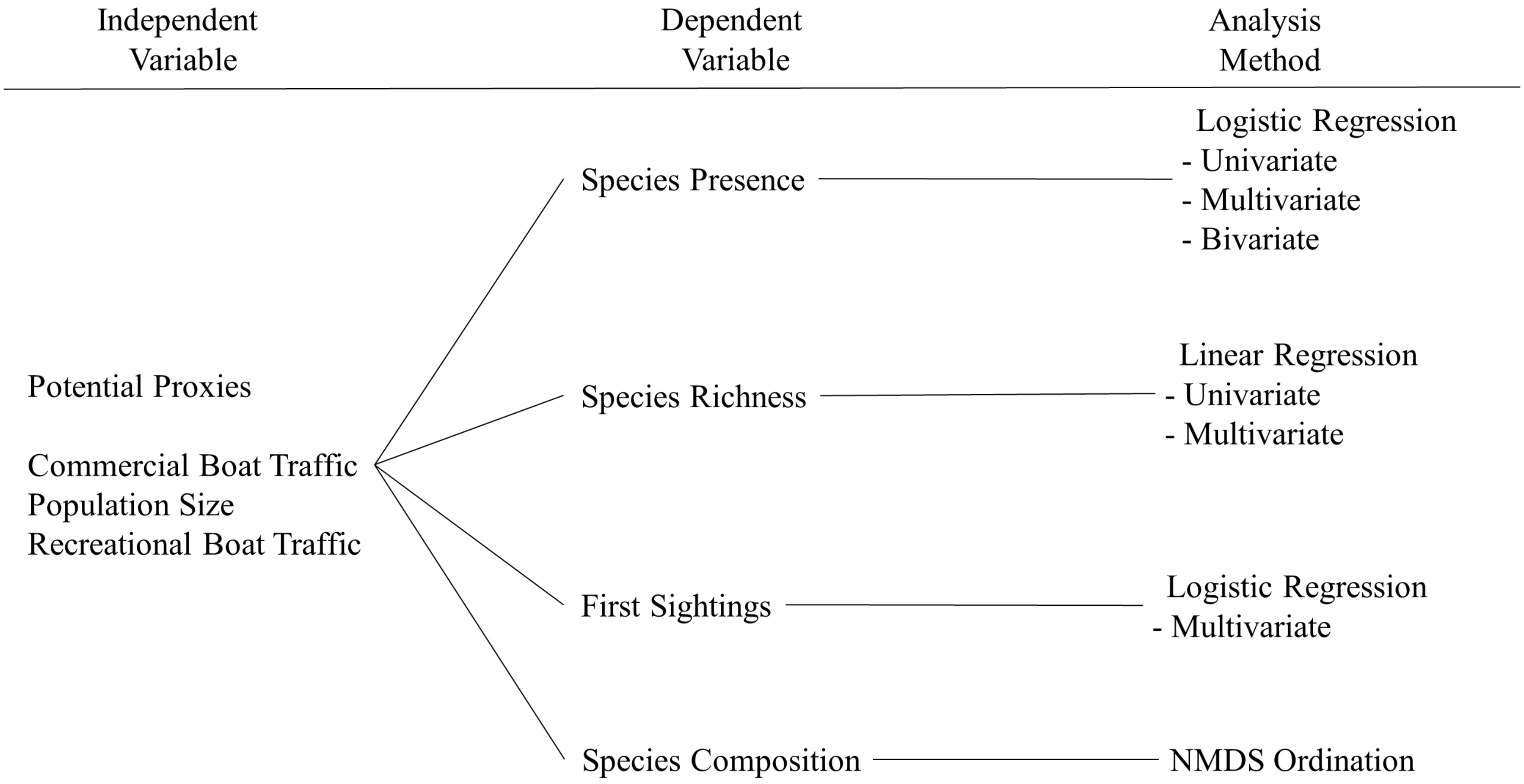


Figure 6 – Summary of regression analyses conducted. Pathway intensity for the three anthropogenic introduction and spread pathways was evaluated with respect to four variables to characterize NAS sightings data. Logistic regression analyses were done using presence/absence data of NAS sightings. Species richness was calculated as the cumulative species richness for the study time period (1970-2013). A nonmetric multidimensional scaling (NMDS) ordination was done to compare NAS composition in large and small cities.

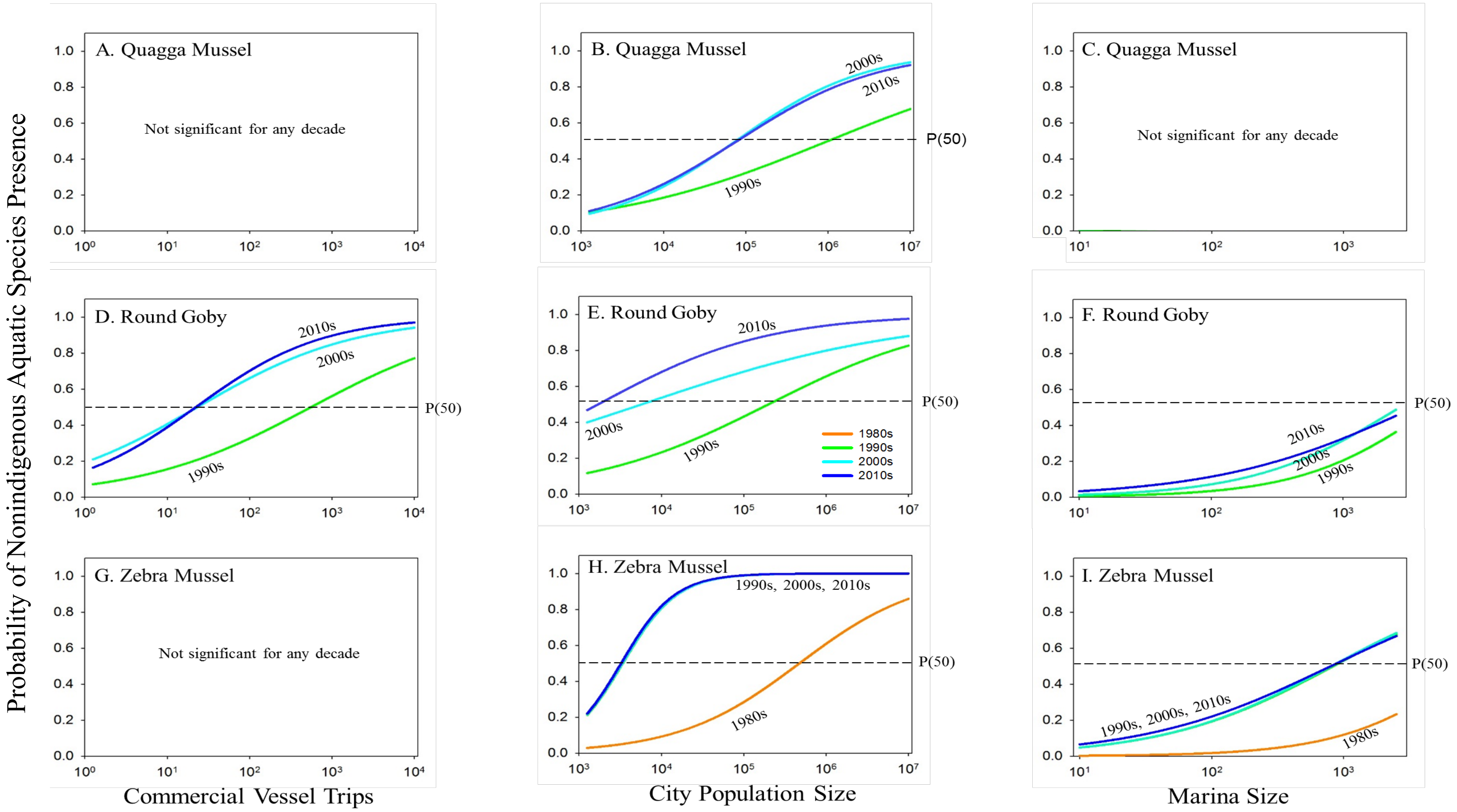


Figure 7 – Probability of NAS presence from univariate logistic regression models. All displayed relationships were significant at α < 0.01. The absence of a probability curve for a given decade indicates that the relationship was not significant for that given decade. Quagga Mussel and Round Goby models were done for three decades (1990s-2010s); the Zebra Mussel models were done for four decades (1980s – 2010s). The p(50) marks the point where the probability of presence equals 0.5.

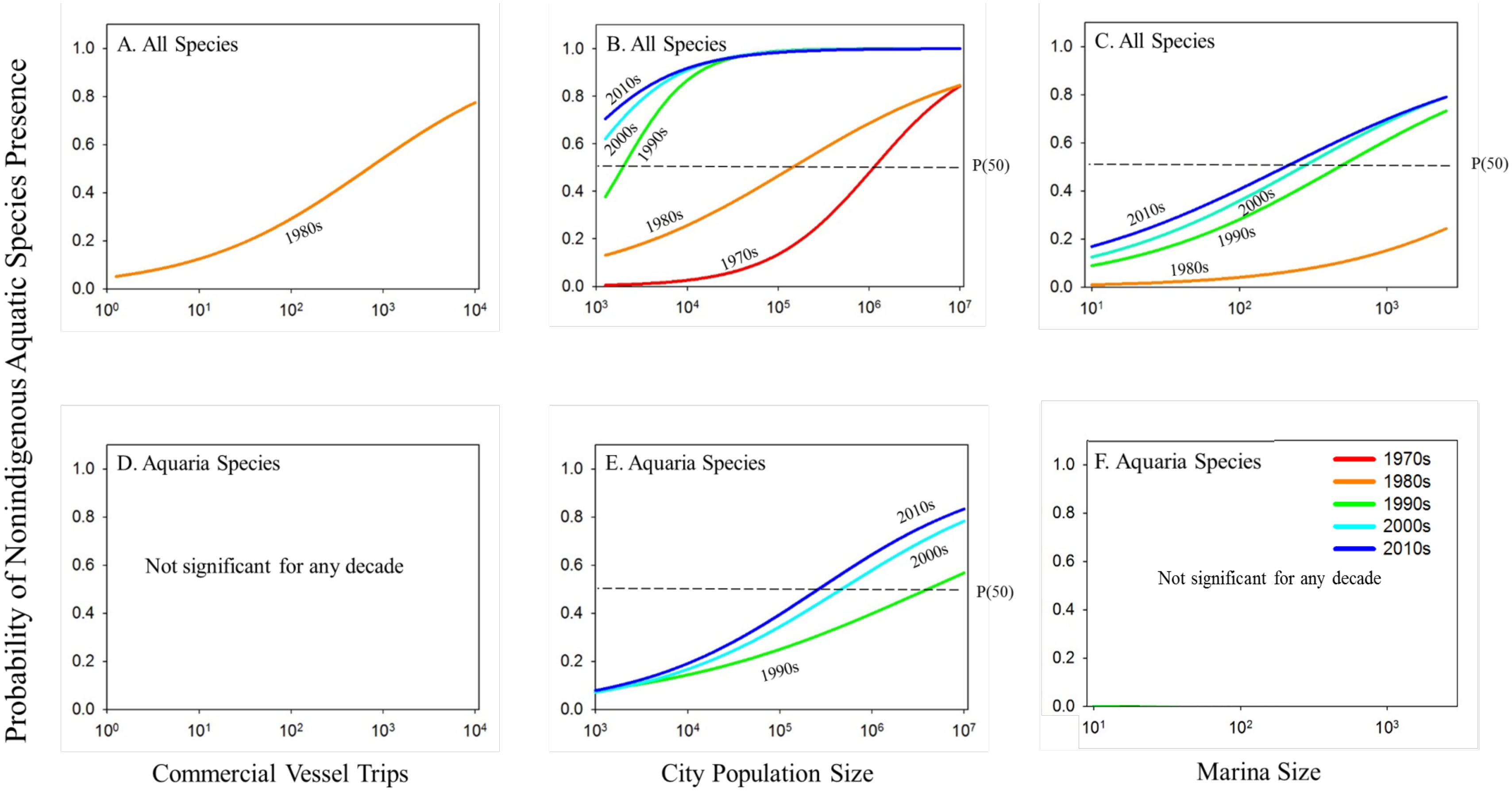


Figure 8 – Probability of NAS group univariate logistic regression models. The all species model consisted of the 18 primary NAS; the aquaria species model included five NAS associated with the aquaria trade. All displayed relationships were significant at $\alpha < 0.01$. The absence of a regression curve for a given decade indicates that the relationship was not significant for that given decade. The all species models were conducted for all five decades (1970s-2010s); the aquaria species models were done for four decades (1980s – 2010s). The p(50) marks the point where the probability of presence equals 0.5.

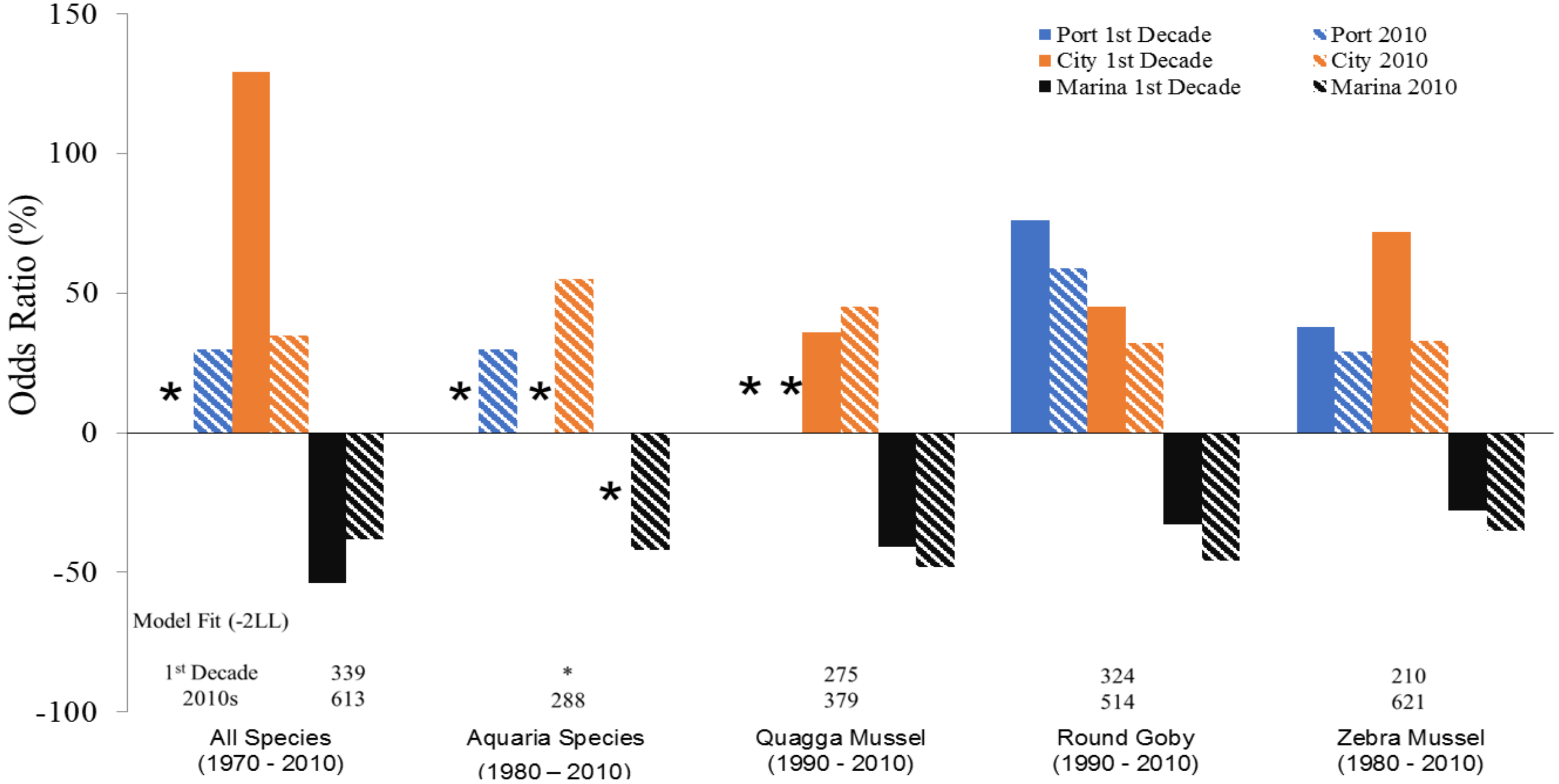


Figure 9 – Percent change in odds ratio for the multivariate logistic regression NAS presence models. The change in odds ratio value is the percent (%) change of probability (y-axis) given a one-unit increase on the x-axis (pathway intensity). Higher odds ratios translate to a greater influence from a given pathway. Negative values indicate that for a change on the x-axis, the probability of presence decreases. Solid bars represent the odds ratio for the first decade that a species was sighted. '*' denotes that the relationship was not significant for the given species / pathway / decade. Commercial boat traffic – port; population size – city; recreational boat traffic – marina.

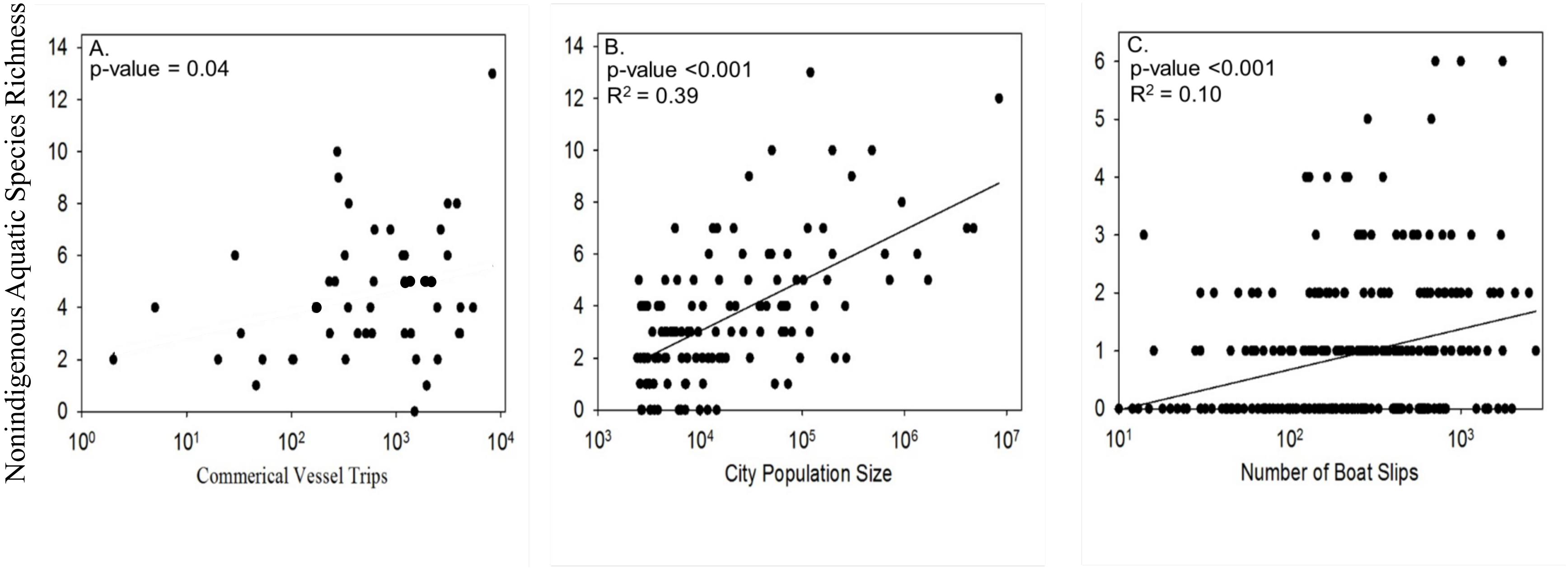
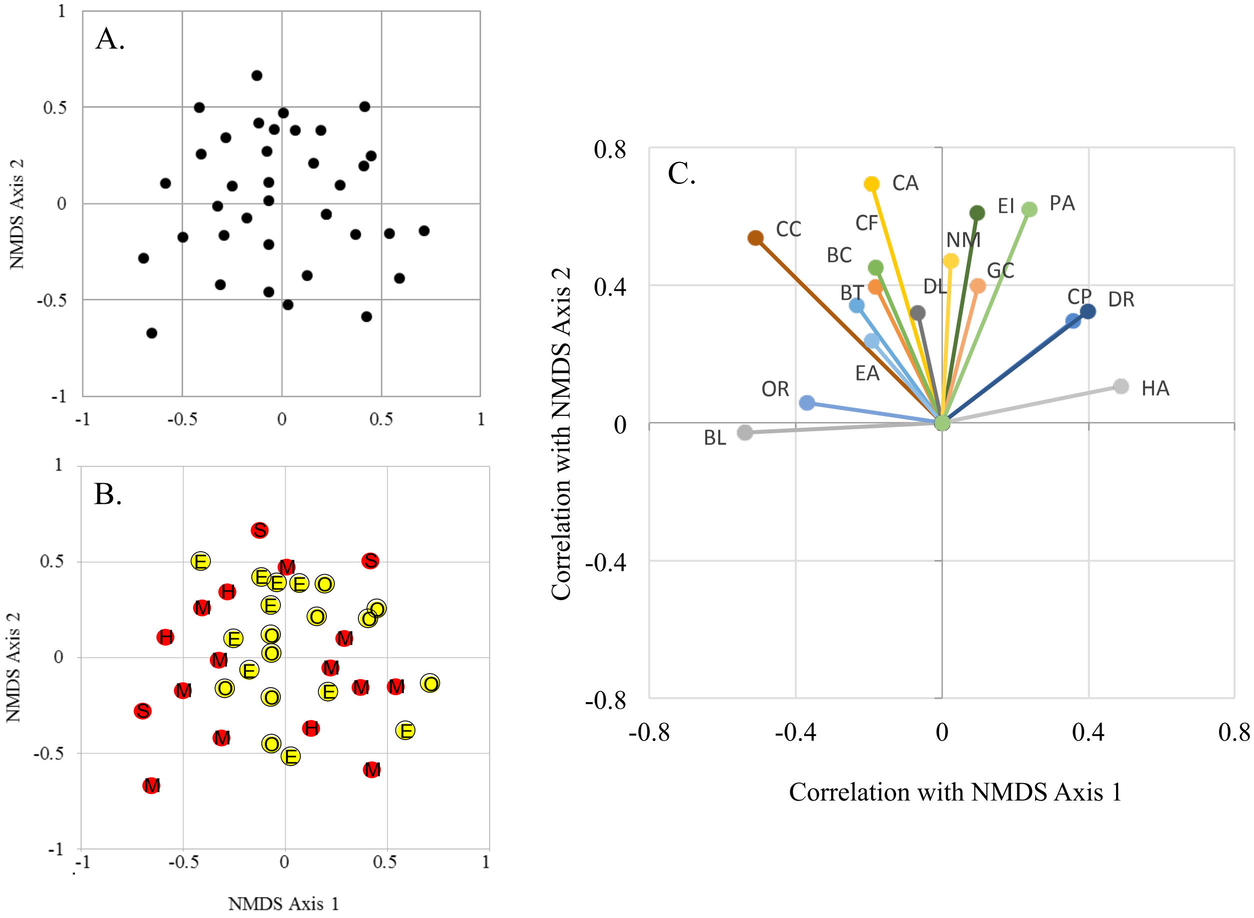


Figure 10 – Cumulative NAS richness as a function of pathway intensity (line represents best-fit linear regression). Significance was determined at $\alpha = 0.01$.



Abbreviations:
BT - *Bithynia tentaculata* (Faucet Snail);
BC - *Bosmina coregoni* (Waterflea);
BL – *Bythotrephes longimanus* (Spiny Waterflea);
CA – *Carassius auratus* (Goldfish);
CP – *Cercopagis pengoi* (Fishhook Waterflea);
CF – *Corbicula fluminea* (Asian Clam);
CS – *Craspedacusta sowerbyi* Freshwater Jellyfish);
CC – *Cyprinus carpio* (Common Carp);
DL – *Daphnia lumholtzi* (Waterflea);
DP - *Dreissena polymorpha* (Zebra Mussel);
DR - *Dreissena rostriformis* (Quagga Mussel);
EI - *Echinogammarus ischnus* (Scud);
EA - *Eurytemora affinis* (Calanoid copepod);
GC - *Gymnocephalus cernua* (Eurasian Ruffe);
HA - *Hemimysis anomala* (Bloody Red Shrimp);
NM - *Neogobius melanostomus* (Round Goby);
OR - *Orconectes rusticus* (Rusty Crayfish);
PA - *Potamopyrgus antipodarum* (New Zealand Mud Snail).

Figure 11 – Nonmetric multidimensional scaling (NMDS) ordination of NAS composition for large cities (>35,000). NMDS plot (A), NMDS plot labeled by lake (letters: S-Superior, M-Michigan, H-Huron, E-Erie, O-Ontario) and geographic region (upper lakes – red without border, lower lakes – yellow with border) (B), and individual species distribution correlation with NMDS axis (C). Vector overlays were calculated using Pearson’s Correlation between NAS presence and NMDS Axis 1 and 2; the correlation values were plotted as the end points of the vectors from the origin.

References

Allan, D., McIntyre, P., Smith, S., Halpern, B., Boyer, G., Buchsbaum, A., Burton, A., Campbell, L., Chadderton, L., Ciborowski, J., Doran, P., Eder, T., Infante, D., Johnson, L., Joseph, C., Marino, A., Prusevich, A., Read, J., Rose, J., Rutherford, E., Sowa, S., and Steinman, A. 2013. Joint Analysis of Stressors and Ecosystem Services to Enhance Restoration Effectiveness. PNAS 110(1): 372–377.

Anderson, L. 2005. California's Reaction to Caulerpa taxifolia: A Model for Invasive Species Rapid Response. Biological Invasions 7(6): 1003–1016.

Arenas, F., Bishop, J., Carlton, J., Dyrynda, P., Farnham, W., Gonzalez, D., Jacobs, M., Lambert, C., Lambet, G., Nielsen, S., Pederson, J., Porter, J., Ward, S., and Wood, C 2006. Alien Species and Other Notable Records from a Rapid Assessment Survey of Marinas on the South Coast of England. Journal of the Marine Biological Association of the United Kingdom 86(06): 1329-1337.

Ashton, G., Boos, K., Shucksmith, R., and Cook, E. 2006. Rapid Assessment of the Distribution of Marine Non-native Species in Marinas in Scotland. Aquatic Invasions 1(4): 209-213.

Bailey, S., Duggan, I., van Overdijk, C., Johengen, T., Reid, D., and MacIsaac, H. 2004. Salinity Tolerance of Diapausing Eggs of Freshwater Zooplankton. Freshwater Biology 49(3): 286–295.

Bailey, S., Duggan, I., Jenkins, P., and MacIsaac, H. 2005. Invertebrate Resting Stages in Residual Ballast Sediment of Transoceanic Ships. Canadian Journal of Fisheries and Aquatic Sciences 62(5): 1090-1103.

Bailey, S., Chan, F., Ellis, S., Bronnenhuber, J., Bradie, J., and Simard, N. 2012. Risk Assessment for Ship-mediated Introductions of Aquatic Nonindigenous Species to the Great Lakes and Freshwater St. Lawrence River. DFO Canadian Science Advisory Secretariat Doc. 2011/104 vi + 224: 1–230.

Bax, N., Hayes, K., Marshall, A., Parry, D., and Thresher, R. 2002. Man-made Marinas as Sheltered Islands for Alien Marine Organisms: Establishment and Eradication of an Alien Invasive Marine Species. Turning the tide: the eradication of invasive species. IUCN SSC Invasive Species Specialist Group. IUCN, Gland, Switzerland and Cambridge, UK: 26-39.

Beck, K., Zimmerman, K., Schardt, J., Stone, J., Lukens, R., Reichard, S., Randall, J., Cangelosi, A., Cooper, D., and Thompson, J.P. 2008. Invasive Species Defined in a Policy Context: Recommendations from the Federal Invasive Species Advisory Committee. Invasive Plant Science and Management 1(4): 414–421.

Beranek, A. 2012. An Assessment of the Long-term Phenology and Impact of Bythotrephes longimanus in Island Lake Reservoir, Minnesota, using Sediment Records. Retrieved from the University of Minnesota Digital Conservancy, http://purl.umn.edu/121007. Accessed 3/25/2015.

Bonferroni, C. 1936. Teoria Statistica delle Classi e Calcolo delle Probabilita. Libreria internazionale Seeber.

Blackburn, T., Pyšek, P., Bacher, S., Carlton, J., Duncan, R., Jarošík, V., Wilson, J., and Richardson, D. 2011. A Proposed Unified Framework for Biological Invasions. Trends in Ecology & Evolution 26(7): 333–339.

Boets, P., Lock, K., Messiaen, M., and Goethals, P. 2010. Combining Data-driven Methods and Lab Studies to Analyse the Ecology of *Dikerogammarus villosus.* Ecological Informatics 7(2): 133-139.

Bossenbroek, J., Kraft, C., and Nekola, J. 2001. Prediction of Long-distance Dispersal using Gravity Models: Zebra Mussel Invasions of Inland Lakes. Ecological Applications 11(6): 1778–1788.

Bradie, J., Chivers, C., and Leung, B. 2013. Importing Risk: Quantifying the Propagule Pressure–establishment Relationship at the Pathway Level. Diversity and Distribution 19(8): 1020–1030.

Briski, E., Bailey, S., and MacIsaac, H. 2011. Invertebrates and their Dormant Eggs Transported in Ballast Sediments of Ships Arriving to the Canadian Coasts and the Laurentian Great Lakes. Limnology and Oceanography. 56(5): 1929–1939.

Brown, M. and Branstrator, D. 2004. A 2001 Survey of Crustacean Zooplankton in the Western Arm of Lake Superior. Journal of Great Lakes Research, 30(1): 1-8.

Buchan, L., and Padilla, D. 1999. Estimating the Probability of Long-distance Overland Disersal of Invading Aquatic Species. Ecological Applications 9(1): 254 – 265.

Butchart, S., Walpole, M., Collen, B., van Strien, A., Scharlemann, J., and Watson, R. 2010. Global Biodiversity: Indicators of Recent Declines. Science 328(5982): 1164.

Colautti, R. and MacIsaac, H. 2004. A neutral terminology to define ‘invasive’ species. Diversity and Distributions 10(2):135-141.

Copp, G., Vilizzi, L., and Gozlan, R. 2010. The demography of introduction pathways, propagule pressure and occurrences of non-native freshwater fish in England. Aquatic Conservation: Marine and Freshwater Ecosystems 20(5): 595–601.

Counts III, C. 1986. The zoogeography and history of the invasion of the United States by *Corbicula fluminea* (Bivalvia: Corbiculidae). American Malacological Bulletin, Special Edition 2:7-39.

Crooks, J. 2005. Lag Times and Exotic Species: The Ecology and Management of Biological Invasions in Slow-motion. Ecoscience 12(3): 316-329.

Doblin, M., Reid, D., and Dobbs, F. 2001. Assessment of Transoceanic Nobob Vessels and Low-Salinity Ballast Water as Vectors for Nonindigenous Species Introductions to the Great Lakes. In Proceedings of the Second International Conference on Marine Bioinvasions, New Orleans, La (pp. 34-35). Available online: http://massbay.mit.edu/exoticspecies/conferences/2001/abstracts03.html. Accessed 3/25/2015.

Duggan, I., Rixon, C., and MacIsaac, H. 2006. Popularity and Propagule Pressure: Determinants of Introduction and Establishment of Aquarium Fish. Biological Invasions 8(2): 377–382.

Garton, D., Berg, D., and Fletcher, R. 1990. Thermal Tolerances of the Predatory Cladocerans *Bythotrephes cederstroemi* and *Leptodora kindtii*: Relationship to Seasonal Abundance in Western Lake Erie. Canadian Journal of Fisheries Aquatic Sciences 47(4):731-738.

Gertzen, E., Familiar, O., and Leung, B. 2008. Quantifying Invasion Pathways: Fish Introductions from the Aquarium Trade. Canadian Journal of Fisheries and Aquatic Science 65(7): 1265–1273.

Gertzen, E., and Leung, B. 2011. Predicting the Spread of Invasive Species in an Uncertain World: Accommodating Multiple Vectors and Gaps in Temporal and Spatial Data for Bythotrephes longimanus. Biological Invasions 13(11): 2433–2444.

Gilbert, M., Grégoire, J., Freise, J., and Heitland, W. 2004. Long-distance Dispersal and Human Population Density allow the Prediction of Invasive Patterns in the Horse Chestnut leafminer Cameraria ohridella. Journal of Animal Ecology 73(3): 459–468.

Glardon, C., Walters, L., Quintana-Ascencio, P., McCauley, L., Stam, W., and Olsen, J. 2008. Predicting Risks of Invasion of Macroalgae in the Genus Caulerpa in Florida. Biological Invasions 10(7): 1147–1157.

Grigorovich, I., Colautti, R., and MacIsaac, H. 2003. Ballast-mediated Animal Introductions in the Laurentian Great Lakes: Retrospective and Prospective Analyses. Canadian Journal of Fisheries and Aquatic Science 60(6): 740 – 756.

Grigorovich, I., Kelly, J., Darling, J., and West, C. 2008. The Quagga Mussel Invades the Lake Superior Basin. Journal of Great Lakes Research *34*(2): 342-350.

Hair, J., Black, W., Babin, B., Anderson, R., and Tatham, R., 2006. Multivariate data analysis (Vol. 6). Upper Saddle River, NJ: Pearson Prentice Hall.

Haltuch, M., Berkman, P., and Garton, D. 2000. Geographic Information System (GIS) Analysis of Ecosystem Invasion: Exotic Mussels in Lake Erie. Limnology and Oceanography 45(8): 1778–1787.

Hensler, S. and Jude, D. 2007. Diel Vertical Migration of Round Goby Larvae in the Great Lakes. Journal of Great Lakes Research 33(2): 95-302.

Herborg, L., Jerde, C., Lodge, D., Ruiz, G., and MacIsaac, H. 2007. Predicting Invasion Risk Using Measures of Introduction Effort and Environmental Niche Model. Ecological Applications 17(3): 663–674.

Herborg, L., O'Hara, P., and Therriault, T. 2009. Forecasting the Potential Distribution of the Invasive Tunicate Didemnum vexillum. Journal of Applied Ecology 46(1): 64–72.

Higgins, S., and Vander Zanden, M. 2010. What a Difference a Species Makes: a Meta–analysis of Dreissenid Mussel Impacts on Freshwater Ecosystems. Ecological Monographs 80(2): 179–196.

Hoffman, J., Kelly, J., Trebitz, A., Peterson, G., and West, C. 2011. Effort and Potential Efficiencies for Aquatic Non-native Species Early Detection. Canadian Journal of Fisheries and Aquatic Science 68(12): 2064–2079.

Holeck, K., Mills, E., MacIsaac, H., Dochoda, M., Colautti, R., and Ricciardi, A. Bridging Troubled Waters: Biological Invasions, Transoceanic Shipping, and the Laurentian Great Lakes. BioScience 54(10): 919–929.

Johnson, L., Brawley, S., and Adey, W. 2012. Secondary Spread of Invasive Species: Historic Patterns and Underlying Mechanisms of the Continuing Invasion of the European Rockweed Fucus Serratus in Eastern North America. Biological Invasions 14(1): 79–97.

Kelly, N., Wantola, K., Weisz, E., and Yan, N. 2012. Recreational Boats as a Vector of Secondary Spread for Aquatic Invasive Species and Native Crustacean Zooplankton. Biological Invasions 15(3): 509–519.

Kerr, S., Brousseau, C., and Muschett, M. 2005. Invasive Aquatic Species in Ontario: A Review and Analysisof Potential Pathwaysfor Introduction. Fisheries 30(7): 21–30.

Kolar, C., and Lodge, D. 2001. Progress in Invasion Biology: Predicting Invaders. TRENDS in Ecology and Evolution 16(4): 199–204.

Kornis, M., and Vander Zanden, M. 2010. Forecasting the Distribution of the Invasive Round Goby (Neogobius melanostomus) in Wisconsin Tributaries to Lake Michigan. Canadian Journal of Fisheries and Aquatic Science. Sci. 67(3): 553–562.

Kulhanek, S., Ricciardi, A., and Leung, B. 2011. Is Invasion History a Useful Tool for Predicting the Impacts of the World's Worst Aquatic Invasive Species? Ecological Applications 21(1): 189–202.

Lake Superior Binational Program. Lake Superior Aquatic Invasive Species Complete Prevention Plan. 2014. Available at http://www.epa.gov/glnpo/lakesuperior/index.html. Accessed 2/5/2015.

Leung, B., Lodge, D., Finnoff, D., Shogren, J., Lewis, M., and Lamberti, G. 2002. An Ounce of Prevention or a Pound of Cure: Bioeconomic Risk Analysis of Invasive Species. Proceedings: Biological Sciences 269(1508): 2407–2413.

Leung, B., Drake, J., and Lodge, D. 2004. Predicting Invasions: Propagule Pressure and the Gravity of Allee Effects. Ecology 85(6): 1651–1660.

Leung, B., and Mandrak, N. 2007. The Risk of Establishment of Aquatic Invasive Species: Joining Invasibility and Propagule Pressure. Proceedings of the Royal Society B 274(1625): 2603–2609.

Lo, V., Levings, C., and Chan, K. 2012. Quantifying Potential Propagule Pressure of Aquatic Invasive Species from the Commercial Shipping Industry in Canada. Marine Pollution Bulletin 64(2): 295–302.

Lockwood, J., Cassey, P., and Blackburn, T. 2009. The More you Introduce the More you Get: The Role of Colonization Pressure and Propagule Pressure in Invasion Ecology. Diversity and Distributions 15(5): 904–910.

Lodge, D., Stein, R., Kenneth, B., Covich, A., Bronmark, C., Garvey, J., and Klosiewski, S. 1998. Predicting Impact of Freshwater Exotic Species on Native Biodiversity: Challenges in Spatial Scaling. Australian Journal of Ecology 23(1): 53–67.

Lodge, D., Williams, S., MacIsaac, H., Hayes, K., Leung, B., Reichard, S., Mack, R., Moyle, P., Smith, M., Andow, D., Carlton, J., and McMichael, A. 2006. Biological Invasions: Recommendations for U.S. Policy and Management. Ecological Applications 16(6): 2035–2054.

MacIsaac, H. 1996. Potential Abiotic and Biotic Impacts of Zebra Mussels on the Inland Waters of North America. American Zoologist. 36(3): 287–299.

Marvier, M., Kareiva, P., and Neubert, M. 2004. Habitat Destruction, Fragmentation, and Disturbance Promote Invasion by Habitat Generalists in a Multispecies Metapopulation. Risk Analysis 24(4): 869–878.

Mackie, G. and Schlosser, D. 1996. Comparative Biology of Zebra Mussels in Europe and North America: An Overview. American Zoologist 36(3): 244-258

Marty, J., Bowen, K., Koops, M., and Power, M. 2010. Distribution and Ecology of Hemimysis anomala, the Latest Invader of the Great Lakes Basin. Hydrobiologia 647(1): 71-80

McDowall, R. 2004. Shoot First, and then Ask Questions: A look at Aquarium Fish Imports and Invasiveness in New Zealand. New Zealand Journal of Marine and Freshwater Research 38(3): 503–510.

Mills, E., Leach, J., Carlton, J., and Secor, C. 1993. Exotic Species in the Great Lakes: A History of Biotic Crises and Anthropogenic Introductions. Journal of Great Lakes Research 19(1): 1–54.

Mills, E., Rosenberg, G., Spidle, A., Ludyanskiy, M., Pligin, Y., and May, B. 1996. A Review of the Biology and Ecology of the Quagga Mussel (*Dreissena bugensis*), a Second Species of Freshwater Dreissenid Introduced to North America. American Zoology 36(3): 271-286.

Mills, E., Chrisman, J., Baldwin, B., Owens, R., O'Gorman, R., Howell, T., Roseman, E., and Raths, M. 1999. Changes in the Dressenid Community in the Lower Great Lakes with Emphasis on Southern Lake Ontario. Journal of Great Lakes Research 25(1):187-197.

Molnar, J., Gamboa, R., Revenga, C., and Spalding, M. 2008. Assessing the Global Threat of Invasive Species to Marine Biodiversity. Frontiers in Ecology and the Environment 6(9): 485–492.

Moran P. 1948. The Interpretation of Statistical Maps. Journal of the Royal Statistical Society 2(10): 243-251.

Morris, J. 2009. Setting a Course for Ballast Water Management. U.S. Coast Guard. Available online: http://www.uscg.mil/proceedings/Winter2008-09/articles/80_Morris_Ballast%20Water%20Management.pdf. Accessed 2/5/2015.

National Ballast Information Clearinghouse 2015. NBIC Online Database. Electronic publication, Smithsonian Environmental Research Center & United States Coast Guard. Available from http://invasions.si.edu/nbic/search.html. Accessed 2/5/2015.

Olden, J., McCarthy, J., Maxted, J., Fetzer, W., and Vander Zanden, M. 2006. The Rapid Spread of Rusty Crayfish (Orconectes rusticus) with Observations on Native Crayfish Declines in Wisconsin (U.S.A.) over the past 130 years. Biological Invasions 8(8): 1621–1628.

Padilla, D., and Williams, S. 2004. Beyond Ballast Water: Aquarium and Ornamental Trades as Sources of Invasive Species in Aquatic Ecosystems. Frontiers in Ecology and the Environment 2(3): 131–138.

Peterson, T., and Vieglais, D. 2001. Predicting Species Invasions using Ecological Niche Modeling: New Approaches from Bioinformatics Attack a Pressing Problem. BioScience 51(5): 363–371.

Phillips, S., Dudik M., Graham, C., Lehmann, A., Leathwick, J., and Ferrier, S. 2009. Sample Selection Bias and Presence-only Distribution Models: Implications for Background and Pseudo-absence Data. Ecological Applications 19(1): 181-197.

Pimentel, D., McNair, S., Janeka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J., Aquino, T., and Tsomondo, T. 2001. Economic and Environmental Threats of Alien Plant, Animal, and Microbe invasions. Agriculture, Ecosystems and Environment 84(1): 1–20.

Pimentel, D., Zuniga, R., and Morrison, D. 2005. Update on the Environmental and Economic Costs Associated with Alien-invasive Species in the United States. Ecological Economics 52(3): 273–288.

Rao, Y. and Schwab, D. 2007. Transport and Mixing between the Coastal and Offshore Waters in the Great Lakes: a Review. Journal of Great Lakes Research 33(1): 202-218.

Ray, W. and Corkum, L. 2001. Habitat and Site Affinity of the Round Goby. Journal of Great Lakes Research 27(3): 329-334.

Ricciardi, A., and MacIsaac, H. Recent Mass Invasion of the North American Great Lakes by Ponto-Caspian Species. Trends in Ecology and Evolution 15(2): 62-65.

Ricciardi, A. 2001. Facilitative Interactions among Aquatic Invaders: Is an "Invasional Meltdown" occurring in the Great Lakes? Canadian Journal of Fisheries and Aquatic Science 58(12): 2513–2525.

Ricciardi, A. 2006. Patterns of Invasion in the Laurentian Great Lakes in Relation to Changes in Vector Activity. Diversity and Distributions 12(4): 425–433.

Ricciardi, A., and MacIsaac, H. 2011. Fifty Years of Invasion Ecology: The Legacy of Charles Elton. Wiley - Blackwell.

Ricciardi, A., and Rasmussen, J. 1998. Predicting the Identity and Impact of Future Biological Invaders: A Priority for Aquatic Resource Management. Canadian Journal of Fisheries and Aquatic Science 55(7): 1759–1765.

Rixon, C., Duggan, I., Bergeron, N., Ricciardi, A., and MacIsaac, H. 2005. Invasion Risks Posed by the Aquarium Trade and Live Fish Markets on the Laurentian Great Lakes. Biodiversity and Conservation 14(6): 1365–1381.

Rothlisberger, J., and Lodge, D. 2013. The Laurentian Great Lakes as a Beachhead and a Gathering Place for Biological Invasions. Aquatic Invasions 4: 361-374.

Rothlisberger, J., Chadderton, L., McNulty, J., and Lodge, D. 2011. Aquatic Invasive Species Transport via Trailered Boats: What is Being Moved, Who is Moving it, and What Can Be Done. Fisheries 5(3): 121–132.

Rup, M., Bailey, S., Wiley, C., Minton, M., Miller, A.W., Ruiz, G., and MacIsaac, H. 2010. Domestic Ballast Operations on the Great Lakes: Potential Importance of Lakers as a Vector for Introduction and Spread of Nonindigenous Species. Canadian Journal of Fisheries and Aquatic Science 67(2): 256–268.

Sala, O., Chapin III, S., Armesto, J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huennek, L., Jackson, R., Kinzig, A., Leemans, R., Lodge, D., Mooney, H., Oesterheld, M., Poff, N.L., Sykes, M., Walker, B., Walker, M., and Wall, D. 2000. Global Biodiversity Scenarios for the Year 2100. Science 287(5459): 1770–1774.

Schneider, D., Ellis, C., and Cummings, K. 1998. A Transportation Model Assessment of the Risk to Native Mussel Communities from Zebra Mussel Spread. Conservation Biology 12(4): 788–800.

Simberloff, D. "Invasive Species". Conservation Biology for All. Ed. Sodhi, N. and Erhlich, P. Oxford: Oxford University Press 2010, 131-138. Web. Accessed 2/5/2015.

Spaling, M., Fox, H., Allen, G., Davidson, N., Ferdaña, Z., Finlayson, M., Halpern, B., Jorge, M., Lombana, A., Lourie, S., Martin, K., McManus, E., Molnar, J., Recchia, A., Cheri A., and Robertson, J. 2007. Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. BioScience 57(7): 573–583.

Spear, D., Foxcroft, L.C., Bezuidenhout, H., and McGeoch, M.A. 2013. Human Population Density Explains Alien Species Richness in Protected Areas. Biological Conservation, 159: 137–147.

Spidle, A., Mills, E., and May, B. 1995. Limits to Tolerance of Temperature and Salinity in the Quagga Mussel (Dreissena bugensis) and the Zebra Mussel (Dreissena polymorpha). Canadian Journal of Fisheries and Aquatic Sciences 52(10):2108-2119.

Statistics Canada. Available online: http://www12.statcan.gc.ca/census-recensement/index-eng.cfm. Accessed 2/5/2015.

Tate, R. 1954. Correlation between a Discrete and a Continuous Variable. Point-Biserial Correlation. The Annals of Mathematical Statistics 25(3): 603-607.

U.S. Army Corps of Engineers 2014. Available online: http://www.navigationdatacenter.us/data/datappor.htm. Accessed 2/5/2015.

U.S. Census Bureau. Available online:_www.census.gov. Accessed 2/5/2015

U.S. Environmental Protection Agency 2014. Invasive Species. Available online: http://www.epa.gov/owow/invasive_species/ Accessed: 2/5/2015.

U.S. Geological Survey. Non-indigenous Aquatic Species database. Available online: http://nas.er.usgs.gov Accessed 2/5/2015.

Vander Zanden, J., Hansen, G., Higgins, S., and Kornis, M. 2010. A Pound of Prevention, Plus a Pound of Cure: Early Detection and Eradication of Invasive Species in the Laurentian Great Lakes. Journal of Great Lakes Research 36(1): 199–205.

von Holle, B., and Simberloff, D. 2005. Ecological Resistance to Biological Invasion Overwhelmed by Propagule Pressure. Ecology 86(12): 3212–3218.

Ward, G., Hastie, T., Barry, S., Elith, J., and Leathwick, R. 2009. Presence-only Data and the EM Algorithm. Biometrics 65(2): 554-563.

Weber, M. and Brown, M. 2009. Effects of Common Carp on Aquatic Ecosystems 80 Years after "Carp as a Dominant": Ecological Insights for Fisheries Management. Reviews in Fisheries Science 17(4): 524–537.

Wells, M, Bailey, S., and Ruddick, B. 2011. The Dilution and Dispersion of Ballast Water Discharge in Goderich Harbor. Marine Pollution Bulletin 62(6): 1288-1296.

Wilcove, D., Rothstein, D., and Dubow, J. 1998. Quantifying Threats to Imperiled Species in the United States. Bioscience 48(8): 607–615.

Wilson, K., Howell, E., and Jackson, D. 2006. Replacement of Zebra Mussels by Quagga Mussels in the Canadian Nearshore of Lake Ontario: the Importance of Substrate, Round Goby Abundance, and Upwelling Frequency. Journal of Great Lakes Research 32(1): 11-28.

Wonham, M., Byers, J., Grosholz, E., and Leung, B. 2013. Modeling the Relationship between Propagule Pressure and Invasion Risk to Inform Policy and Management. Ecological Applications 23(7): 1691-1706.

Yurista, P. and Kelly, J. 2009. Spatial patterns of water quality and plankton from high-resolution continuous *in situ* sensing along a 537-km nearshore transect of western Lake Superior, 2004. State of Lake Superior. Edited by M. Munawar and R. Heath. Aquatic Ecosystem Health and Management Society, Burlington, Ont., Canada, 439-471.