EPA Public Access

Author manuscript

Hydrobiologia. Author manuscript; available in PMC 2019 July 02.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Hydrobiologia. 2018 July 2; 817(1): 23–40. doi:10.1007/s10750-018-3551-x.

Pathways and places associated with nonindigenous aquatic species introductions in the Laurentian Great Lakes

Elon M. O'Malia¹, Lucinda B. Johnson², and Joel C. Hoffman³

¹University of Minnesota Duluth, Integrated Bioscience Graduate Program, 207 Swenson Science Building, 1035 Kirby Drive, Duluth, MN 55812

²Natural Resource Research Institute, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth, MN 55811

³U.S. Environmental Protection Agency, Office of Research and Development, Mid-Continent Ecology Division, 6201 Congdon Blvd, Duluth, MN 55804

Abstract

Propagule pressure (i.e., the frequency and abundance of introductions) is a common indicator of the likelihood of nonindigenous aquatic species (NAS) establishment success. Evaluating propagule pressure associated with multiple introduction pathways relative to present NAS distribution patterns may identify which pathway presents the greatest risk. Our objective was to develop and evaluate three geospatial metrics for the Laurentian Great Lakes as proxies of propagule pressure associated with three major introduction pathways: maritime commerce, organisms in trade, and water recreation. Logistic and linear regression analyses were conducted between NAS presence and introduction pathway intensity (e.g., number of vessel trips received by a port) for 23 NAS over a five-decade period (1970 – 2013). Notably, city population size was the best predictor of NAS presence, even for NAS introduced through ballast water discharge. Moreover, through time, city population size was an increasingly significant predictor of the presence of organisms in trade, signaling a change in both the types of organisms introduced and places where introductions are occurring. Nonetheless, all three metrics are reasonable proxies for propagule pressure and as such are applicable for risk assessment, monitoring, and control strategies.

Keywords

Ballast water; live release; marinas; population size; propagule pressure

Introduction

Human activities, such as global trade and recreational water sports, facilitate the transport, introduction, and spread of nonindigenous aquatic species (NAS; Mills et al., 1993, Sala et al., 2000, Holeck et al., 2004, Ricciardi 2016). Many NAS 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) and affects both freshwater and marine ecosystems, leaving only 16% of all marine ecoregions worldwide without a NAS present (Spalding et al., 2007, Molnar et al., 2008). Globally, international maritime commerce and commercial aquaculture remain the major pathways for nonindigenous aquatic species introductions, primarily affecting the temperate regions of Europe, Asia, Australia, and North America (Molnar et al., 2008; Seebens et al., 2016).

Propagule pressure (PP; the frequency and abundance of introductions) is broadly accepted as a meaningful indicator of the likelihood of species establishment success (Kolar & Lodge, 2001, Leung et al., 2004, von Holle & Simberloff, 2005, Lodge et al., 2006), but this relationship has rarely been quantified (Bradie et al., 2013). It is difficult to calculate PP for nonindigenous aquatic species, so investigators have used proxies such as ballast water discharge volume and number of animal imports (Lockwood et al., 2009) or shipping intensity (Seebens et al. 2016) to model the risk of introduction, establishment, 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 mixed success at explaining detection patterns (Wonham et al., 2013). The quantity and proximity of recreational maritime traffic was successfully used as a proxy for NAS presence in inland lakes and at regional scales (Bossenbroek et al., 2001, Herborg et al., 2009). 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). Although these proxies have been developed and successfully used to understand NAS patterns, multiple pathways generally operate simultaneously over broad spatial scales (e.g., Clarke Murray et al., 2014), such that studies like this one, which incorporate multiple introduction and spread pathways over large geographic areas can reveal which pathways best correspond to NAS distribution patterns, and thus likely represent the greatest risk.

The development and evaluation of proxies that represent introduction potential can increase our understanding of how and where NAS are introduced, but can also prioritize sampling locations, thus maximizing the success of preventing NAS introduction and spread (Bossenbroek et al., 2001). Although major pathways of introduction and spread have been identified, it is difficult to predict whether or not specific NAS will establish a population once introduced (Blackburn et al., 2011). The present distribution of NAS reflects all stages of the invasion process, from transport to introduction and establishment to secondary spread (Colautti and MacIsaac 2004). Here, we use the Laurentian Great Lakes as a case study region to explore how a variety of proxies for propagule pressure are associated with NAS introductions, establishment, and spread by examining their relationship to NAS distribution. The most common pathway for historical introduction and spread of NAS in the Great Lakes was via ballast water discharge from commercial vessels, accounting for 65% of the introductions (Ricciardi, 2006). These introductions threaten a large part of North America because the Great Lakes serve as a beachhead for NAS spread to the interior of the continent (Rothlisberger & Lodge, 2013).

Broadly, our goal was to evaluate what kinds of places have been susceptible to NAS introduction and establishment. Specifically, the objective of this study was to develop and

evaluate three geospatial metrics (commercial vessel trips, city population size, and marina size) as potential proxies of PP associated with three major pathways (maritime commerce, organisms in trade, and water recreation, respectively) over a large geographic area. Although the case study focus is on the Great Lakes, these pathways are common to NAS introductions generally (e.g., Clarke Murray et al. 2014). These geospatial metrics were tested as model factors that could explain the apparent spatial and temporal relationship between pathway intensity (i.e., a geospatial metric to represent propagule pressure) and NAS detection records (presence). Our first hypothesis was that increasing pathway intensity (i.e. propagule pressure) corresponds to greater probability of NAS presence, and that the relationship would change by NAS or NAS group (e.g., aquaria species). To test this hypothesis, we examined the relationship between NAS first detection locations and the intensity of each pathway, considering pathway intensity individually (i.e., simple regression analysis) and collectively (i.e., multiple regression analysis). Our second hypothesis was that the relationship with respect to individual pathway intensity or multiple pathway intensities would change through time as the relative strength of those pathway intensities changed. To test this hypothesis, we examined how the relationship between NAS detections and pathway intensity changed over time. To explore how cumulative NAS introductions relate to the three pathways, we used multiple linear regression to analyze the relationships between species richness and pathway intensity.

Methods

Fundamental to our study design was the inherent range of variability and magnitude of maritime commerce, city population size, and recreational marina size within the Laurentian Great Lakes region. That is, the variation in location and magnitude of introduction pathways was exploited to conduct a natural experiment to evaluate geospatial proxies for PP. Among aquatic ecosystems globally, the Laurentian Great Lakes are among the most vulnerable to NAS invasions, with over 180 NAS currently present in the region (USGS 2014). The Great Lakes receive international shipping vessels via the Great Lakes-Saint Lawrence Seaway System, which connects all five Great Lakes to the Atlantic Ocean through a series of canals and locks. 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).

Development of Proxy Metrics for Propagule Pressure

Proxy metrics for propagule pressure were developed using geospatial data for three prominent pathways of aquatic introduction and spread: maritime commerce within ports, live release from urban areas with populations above 2,500, and recreational maritime traffic in marinas (Table 1 and Figure 1). All spatial analyses were conducted using ESRI ArcMap™ v10.1 (ESRI, 2014), while statistical and regression analyses were done in SigmaPlot® V 14.0 and the R statistics program (R Core Team, 2014) using the Vegan ordination package (Oksanen, et al., 2013).

Each proxy metric was associated with a specific geographic boundary such as a port or marina (hereafter referred to as an area of pathway influence, or API), which is described below. Maritime commerce was categorized into eight proxy metrics (Table 1); each was analyzed independently to explore the relationship with NAS presence and NAS richness. The experimental proxy metric for the live release pathway was city population size, which was divided into five decadal periods: 1970s, 1980s, 1990s, 2000s, and 2010s. For live releases, we assume that the likelihood of introduction within an API increases with population density. The proxy metric for recreational maritime traffic was marina size, measured by the number of boat slips within a marina. We assumed the likelihood of introduction within a marina increases with marina size owing to an increase in traffic.

Nonindigenous Aquatic Species Detections

We modeled the introduction and spread patterns of three NAS: quagga mussel (*Dreissena bugensis*, Andruzov 1897), zebra mussel (*Dreissenapolyomorpha*, Pallas 1771) and round goby (*Neogobius melanostomus*, Pallas 1814), along with two group models, aquaria-specific species (NAS_{Aq}) and all NAS detections combined (NAS_{all}) (Table 2). The primary dataset for NAS detections 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 detections from individuals and government agencies, as well as from literature reviews by USGS. This dataset is presence only data, and represents detections from both primary (initial introduction into a novel area) and secondary spread (dispersal throughout the newly invaded area). In addition, incidental NAS detections from the Great Lakes Indicator Consortium's (GLIC) Coastal Wetland Monitoring (CWM) Program for 2011 and 2012 (V. Brady, unpublished data) were included to improve coastal wetland habitat coverage throughout the Great Lakes (Uzarski et al., 2017). The NAS detection records were used to determine the presence of invertebrate NAS and five select fish species (common carp, eurasian ruffe, goldfish, round goby, and scud) from 1970 to 2013; we included fish species directly associated with the aquaria trade and live release (Duggan, 2006, Gertzen & Leung, 2011) or maritime commerce pathways (Mills et al., 1993). Only species detections within 10 km of the shoreline were used (see below). The study included 23 species and approximately 3,300 detections (Table 2 and Supplemental Figure 1).

Area of Pathway Influence (API)

An API is the spatial representation of the extent of influence of a given pathway location, specifically, ports, cities, and marinas. Each API was assigned a pathway-specific metric that quantified the amount of human activity at the given location (pathway intensity). By buffering points and polygon areas that represented the pathway locations, APIs were generated for each pathway location (Figure 2). Port polygons were determined using aerial photography by incorporating all physical structures directly related to port activity, e.g. loading docks, ship berths, etc. To delineate the shoreline boundaries of city APIs, we used the maximum extent of a city's waterfront. In contrast, marinas were modeled as point locations. Buffers were applied to the pathway locations; a buffer of 10 km was applied to port and city APIs because alongshore currents in the Great Lakes separate coastal and offshore waters at approximately 10 km from shore (Rao & Schwab, 2007, Yurista & Kelly, 2009). For marina APIs, a 1 km buffer was used to represent the small physical nature of

marinas as structures designed to separate marina waters from nearshore open water. Buffers were limited to Great Lakes waters, excluding tributaries and near shore lakes. If overlap occurred between buffers in the same pathway (Supplemental 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 individual pathway intensity metrics (Supplemental Figure 1B). Spatial combination was not done for city APIs where overlap occurred, rather, the mid-point between two city APIs was used as a division (Supplemental Figure 1C).

In addition to the individual APIs, we created composite APIs; wherever spatial overlap of different individual pathway APIs occurred (i.e. a marina within a city API), the smallest extent of the overlapping APIs was classified as a new API and was attributed with all metrics of the underlying pathways (Supplemental Figure 2). Pathway values of zero were attributed for non-overlapping APIs, i.e., in the composite APIs where only a city was present, the port and marina metric for that given API would be zero.

If a species detection occurred within an API, the NAS detection was associated with that specific API. The presence-only data found in the NAS database potentially is subject to pseudo-absence error, especially through time (i.e., the propensity to not report a long-established NAS). To address temporal inconsistency of a species detection at a given location, we assumed that after the initial detection, the species was present thereafter, e.g., if a NAS was sighted in an API in 1991, we assumed the NAS was present in that API during the 1990s, 2000s, and 2010s. Although the effect of this assumption is to maximize NAS richness at a given API, our analysis focused on NAS that became broadly established throughout the Great Lakes region through time. Thus, it is improbable these successful invaders routinely failed to successfully establish; we do acknowledge, however, that some successful invaders do fail to establish after introduction (e.g., introduced ruffe, *Gymnocephalus cernus*, in Lake Huron; Gutsch & Hoffman, 2016).

Statistical Analyses

NAS Presence Analyses.—We first analyzed whether there was a relationship between where any NAS was first detected (first presence; FP) and the intensity of each pathway, then analyzed how detections of select individual species and NAS groups were related to pathway intensity, and finally examined how the relationship between NAS detections and pathway intensity changed over time using a set of individual species models as well as a set of pathway-specific models (Table 2).

To test for relationships between NAS FP (response) and pathway intensity of ports, cities, and marinas (factors), we used logistic regression, including both simple (each pathway as a single factor) and multiple (all three pathways as factors logistic models). All FP records were pooled across species for those NAS with more than 10 detections ($n=18$) since 1970. First presences were grouped based on the year of the record, and 10 NAS had multiple first presences occur in the same year at different locations.

To test for relationships between NAS detections and pathway intensity, we used select individual species and species groups (pooling all detections from 1970 to present).

Individual species models (IND) were analyzed for quagga mussel, round goby, and zebra mussel because these species had the greatest number of detections and are associated with the maritime commerce and recreational maritime pathways (USGS, 2014). To test the organism in trade pathway, we developed a group aquaria species (NAS_{AQ}) model comprised of seven NAS (goldfish, freshwater jellyfish, chinese mystery snail, european ear snail, freshwater hydroid, freshwater bryozoan, and pacu) (Table 2) associated with the live release pathway (USGS, 2014); we used a group model to increase the sample size of aquaria species detections (n=131). In addition, to evaluate whether there were patterns among all NAS detections, we developed a group model that included all NAS records (NAS_{ALL}; Table 2). We used both simple (each pathway as a single factor) and multiple (all three pathways as factors) logistic regression.

We tested temporal change in the relationships between NAS presence and pathway intensity for each of the three pathways over five decadal periods using logistic regression analysis, again including both simple and multiple models. To do so, with respect to each species or group, we developed a model for each of the five decades (i.e. 1970s are detections between Jan. 1, 1970 and Dec. 31, 1979). To determine whether there was a change in the species-specific association with each pathway after initial introduction, the odds ratio value from a multiple linear analysis 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 FP since 1970. Model significance was tested using the likelihood ratio statistic and its associated p. The model fit was assessed using the $-2 \times \log(\text{likelihood})$ value (i.e., $-2LL$).

The Bonferroni inequality method (Bonferroni 1936) was applied and a significance level of $\alpha = 0.01$ was used to determine significant with more confidence than the standard $\alpha = 0.05$ level. The presence of multicollinearity was tested for in every multiple logistic model using the variance inflation factor (VIF). Low variance inflation factor (VIF) scores for the independent variables (API pathway intensity) in the multiple 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). The absence of multicollinearity indicates there were no interactions between the pathway metrics (independent variables).

Each set of pathway APIs was tested for spatial autocorrelation using Moran's Index (Moran, 1948). No significant spatial autocorrelation was present for cities and marinas (Moran's Index for cities and marinas respectively: $p = 0.47$ and 0.35). Ports were significantly clustered, but to a very small degree ($p = 0.041$; Moran's $I = 0.10$; Moran's I values near 0 suggests no spatial autocorrelation while values near 1 suggest clustering).

NAS Richness Analyses.—To explore how cumulative NAS introductions relate to the three pathways, species richness was calculated for composite APIs, and then relationships between species richness and pathway intensity were analyzed using multiple linear regression (Poisson regression). NAS richness was calculated as the cumulative total of unique NAS associated with an API over the entire study time range (1970 – 2013). We recognize that NAS richness patterns are potentially skewed by the search effort associated

with specific areas (Hoffman et al., 2011). While we do not have independent effort estimates, a recent continental-scale analysis of NAS richness patterns that included the Great Lakes found that spatial “hotspots” in NAS richness were the same pre- and post-data rarefaction, suggesting an effort effect is not sufficiently large at this Great Lakes scale to alter interpretation of richness patterns (Davis & Darling, 2017).

Results

Nonindigenous Aquatic Species Presence

From 1970 forward, there was at least one NAS present in every port API except one (Silver 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.

First Presence (FP).—We found a significant relationship between all three pathway metrics and locations of NAS FP since 1970. 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 ; regression coefficient 0.45 and 0.43). Marina size was also a significant factor, but had a negative odds ratio, -52.1% ($p < 0.001$; coefficient -0.736). The multiple logistic model was significant when tested using the likelihood ratio statistic ($p < 0.001$) and the model fit ($-2LL$) was 162.3.

Individual Species and Species Groups.—In general, all NAS or NAS groups analyzed using simple logistic regressions were more likely to be present at locations with greater pathway intensity (i.e. more commercial vessel trips, larger city population size, and more marina slips) (Figure 3 and Figure 4).

Commercial Vessel Trips (Maritime Commerce).—To determine the best metric to use for maritime commerce, we tested eight pathway metrics using multiple (Table 3) and simple (Supplemental Table 1) logistic modeling. Commercial vessel trips (CVT) and non-discharging vessel trips had the greatest number of significant relationships with species presence over the five decadal period (11 of 19 for multiple regression and 4 of 19 for simple regression models). The differences in the model fits ($-2LL$) between these two trips metrics were minimal, and there was a significant linear regression between them ($R^2 = 0.5$, $p < 0.001$), suggesting that both factors had a similar relationship with NAS presence. Since total CVT was more inclusive as a metric, it was designated as the best maritime commerce metric. Hereafter, we focus on the outputs with respect to CVT.

Surprisingly, the only species for which its presence was significantly related to CVT in the simple regression models was the round goby; the probability of presence increased for all ports after the 1990s (Figure 3D). The CVT metric was not significantly related to the presence of quagga mussel, zebra mussel, or the pooled aquaria species group (NAS_{Aq}) ($p > 0.01$); there was only one decade (1980s) in which we found a significant relationship between CVT and the all species (NAS_{All}) group (Table 4, Figure 3 and 4).

City Population Size.—We found significant relationships between NAS presence and city population size for all NAS and decade combinations except for the NAS_{Aq} model in

the 1980s (Table 4, Figure 3 and 4). The probability of NAS presence was always greater for larger population cities than for smaller population cities.

The probability of quagga mussel presence sharply increased as population size increased from the 1990s to the 2000s, with a maximum increase of 30% in cities around 10^6 (1,000,000) population size. However, in cities below 10^4 (10,000) population size, there was only a 1% increase in the probability that quagga mussel was present (Figure 3B). The probability of presence did not change between the 2000s and the 2010s. More generally, quagga mussel detections were limited to large and medium sized population cities throughout all decades.

The probability of zebra mussel presence was substantially greater in the largest population cities (approximately 80%) than the smallest population cities (<5%) during the 1980s. From the 1980s to 1990s there was a dramatic increase (average of 60%) in medium-sized population cities (10^3 to 10^4 ; Figure 3H), suggesting a rapid increase in zebra mussel introductions near medium sized population cities while small and large population cities had a similar probability post-introduction. From the 1990s onward, the probabilities were the same among cities of different size with population greater than 10^4 .⁷

Alternatively, the probability of round goby presence increased 10% in the largest population cities and 30% increase in the smallest population cities from the 1990s to the 2000s (Figure 3E). After the 2000s, the probability in smaller population cities (< $10^{3.5}$) increased slower than those cities with larger populations.

Presence of the NAS_{Aq} species was not significantly related to city population size until the 1990s (Figure 4E). From the 1990s to the 2010s, the probability was highest in large population cities, and the increase over time was greater in large population cities than small population cities (an approximate 30% increase from the 1990s to the 2000s).

For the NAS_{All} group, the probability of presence increased rapidly from the 1970s to the 1980s in all but the largest population cities. We found a larger increase in the probability from the 1980s to the 1990s in large- and medium-sized population cities (Figure 4B). By the 1990s, all cities approaching a population of $\sim 10^{4.5}$ (31,500) had an approximately 100% chance of NAS presence; the minimum probability was 40% in the smallest population cities. The probability of presence for any size city was greater than 50% for the 2000s and 2010s. Zebra mussel and round goby detections in the 1980s and beyond strongly influenced this model, accounting for over 35% of total detections records.

Marina Size.—Similar to the city population pathway analysis, the probability of NAS presence was significantly related to marina size for most decades (Table 4, Figure 3 and 4). Increasing marina size was associated with greater probability of presence with respect to round goby, zebra mussel, and the NAS_{All} group, but was not significant with respect to quagga mussel and the NAS_{Aq} group. The probability of round goby presence in marinas changed little over time (Figure 3F). In contrast, the probability of zebra mussel and NAS_{All} presence increased 25% (or less) in marinas with below 10^2 (100) slips, but increased nearly 50% in the largest marinas (Figures 3I and 4C).

Multiple Logistic Pathway Analysis.—Using multiple logistic regression, we found a significant relationship for at least one decade between the probability of four of the five NAS being present (response) and CVT (factor) (Figure 5). However, except for round goby, this metric had the smallest odds ratio value and was the least important factor in multiple logistic NAS presence models. The CVT metric did have the largest odds ratio in the round goby model for both the first and current decade (76% and 59% respectively). Yet, the odds ratio for CVT decreased from the first decade to the 2010s for round goby, zebra mussel, and NAS_{All} models. In contrast, city population size had the largest single odds ratio and was a significant model factor for all NAS models and all decades except for the 1980s NAS_{Aq} group model. Similar to the CVT metric, the odds ratios for three of the five species models decreased from the first decade to the current decade, indicating that a substantial increase in presence of NAS had occurred near cities. The only exception was for quagga mussel, for which the city population size odds ratio increased over time. Marina size had negative odds ratio values for all NAS and NAS group models, consistent with the results from the FP simple linear regression analyses. The odds ratio values were increasingly negative through time which was opposite of the trend for both CVT and city population size. That is, the negative influence of marine size on NAS presence increased over time even as the positive influence of both city population size and CVT increased.

Nonindigenous Aquatic Species Richness

The highest NAS richness was 13 species, found in the Duluth-Superior port and city APIs, and associated with the maritime commerce and live release pathways, respectively. The highest species richness for marina APIs was 6, found in three locations: Chicago, IL; Erie, PA; and Port Colborne, Ontario. We found a significant relationship between NAS richness and pathway intensity; all three pathway metrics were significant factors in the multiple linear regression model (Table 5). The relative magnitude of the maritime commerce and city population size regression coefficients were similar among decades except for during the 1970s. The marina size coefficient within the multiple logistic model was always negative (i.e., an inverse relationship between marina size and NAS richness). Notably, both population and marina size coefficients decreased through the 1980s then stabilized for the next three decades; maritime commerce coefficients fluctuated between the 1970s and 1980s, then stabilized as well. The changes in the regression coefficients were minimal after the 1980s, demonstrating that NAS richness increased at a constant rate.

Discussion

The probability of past NAS presence increased with increasing commercial vessel trips (CVT), city population size, and marina size; however, city population size was the best metric to explain primary and secondary spread patterns as represented by reported detections. While city population size is recognized to be an important aspect of NAS spread, especially with respect to aquaria species (Duggan et al., 2006, Gertzen et al., 2008), the finding is important because international shipping is generally considered to be the dominant vector for NAS spread among coastal ecosystems, both globally and in the Laurentian Great Lakes (Carlton & Geller, 1993; Ricciardi 2016). We did find a relationship between NAS_{Aq} species presence and city population size. It first was apparent in the 1990s,

and has increased in magnitude since. In contrast, the strength of the relationship between NAS presence and maritime commerce declined over time. This indicates there have been subtle shifts in introduction pathways and locations through time. Here, we discuss the quality of the geospatial metrics, as well as the implications of our results for how NAS spread (as inferred from presence or absence determined from detections) has occurred.

Geospatial Metrics

We found significant associations between NAS presence and city population size; the overall importance of this pathway is disproportionate with respect to the historical prominence of the maritime commerce pathway (Ricciardi, 2006). In contrast, Seebens et al. (2016), based on a dataset of marine algae, identified global hotspots for invasion using only ship traffic data. Yet, we emphasize that the scale (regional or global) and species of interest are fundamental considerations. Further, the city size metric may represent multiple potential pathways of NAS introduction beyond the live-release avenue, including a disturbance effect (Leprieur et al. 2008), making it a more general metric than either maritime commerce or marina size.

The prominence of city population size as a predictor of NAS presence, even for ballast water introduced species, implies that introduction and detection locations were not always the same. That is, an explanation for the apparent mismatch between the metric and the introduction pathway is that secondary spread often occurred rapidly, prior to first detection. Nonindigenous species populations generally experience a lag phase in population growth during the establishment stage of invasions, which is an effect of 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 detected (Crooks, 2005, Rilov et al., 2004, Branstrator et al., 2017). It is plausible that secondary spread by recreational and commercial maritime traffic occurred before first detection during the lag phase, which would confound associations between introduction and FP locations. Ballast water can travel upwards of 7.5 km in the first day after discharge (Wells et al., 2011), which can aid dispersal of planktonic NAS and account for rapid secondary spread prior to first detection. Moreover, in some cases, current-mediated dispersion may have transported NAS out of port APIs and into city APIs prior to detection. As such, maritime commerce and the CVT metric may still be associated with the spread of NAS within the Great Lakes (Sieracki et al., 2014).

Although city population was a good predictor with respect to overall NAS distribution, the first detection (FP) results corroborate previous findings regarding the source of NAS introductions to the Great Lakes. While all three pathway metrics were significant factors of NAS FP patterns, both CVT and city population size had similar odds ratios, suggesting that there was an equal association with NAS FP. This near equal split of association with each metric is consistent with the ballast water discharge (maritime commerce) 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, approximately 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); evidence that the geospatial metrics reasonably explained the first presence pattern across the basin.

Marina size was positively related to the probability of NAS presence when modeled independently (simple regression), but in multiple logistic models, increasing marina size reduced the probability of NAS detection relative to what was expected based on the associated CVT and marina size. The results from the simple regression models are consistent with Davis & Darling (2017), who found a strong association between NAS richness and the number of recreational vessels and slips in their examination of patterns throughout the 48 contiguous US states. The finding merits further investigation as to why the probability of NAS had a negative relationship to marina size for the multiple logistic results. As disturbed areas, marinas attract and harbor NAS (Ashton et al., 2003, Foster et al., 2016). One possible explanation is that there was an observer effect favoring detection in small marinas over large marinas owing to the reduced effort and greater ease of survey within small marinas.

Limitations

Public NAS detections can potentially contribute to research on the patterns and potential mechanisms behind NAS introduction and spread, but with a few caveats. We recognize that the data are subject to observational bias given the non-systematical nature of the detections data, which can affect analysis (Phillips et al., 2009). Nevertheless, these data sets are among the best we currently have to document large-scale distributions of NAS. Further, recent studies of NAS distribution based on stratified, random sampling broadly corroborate our findings. Trebitz et al. (2009) analyzed benthic invertebrate NAS detections and richness patterns within the Duluth-Superior port (the location with the highest NAS richness in this study), and found that NAS were more likely to be found near either city or near port facilities than in areas upstream of the cities and port facilities. Further, Hoffman et al. (2016) analyzed nonindigenous fish species detections from the three largest ports in Lake Superior and found that the rank order of nonindigenous fish species richness corresponded to the rank order of CVT to the three ports.

A specific, related concern is that increasing population size may increase the likelihood of NAS detection, rather than introduction (presence). Ultimately, we cannot fully evaluate the strength of such an effect because it would require survey effort data. However, observations based on the data conflict with this interpretation. First, in the simple logistic regression analyses, increasing pathway intensity for all three metrics were associated with increased NAS presence, suggesting pathway intensity broadly is related to NAS detection. Second, the zebra mussel analysis revealed that NAS detections increased in medium cities faster than large cities shortly after introduction; this finding contradicts the notion that a greater population yields increased detection. Third, we found a strong association between round goby and CVT (its primary spread pathway - adult round goby does not disperse, but larvae are susceptible to ballast water intake; Ray & Corkum, 2001, Hensler & Jude, 2007), yet a strong association between zebra mussel and both city population size and marina size (i.e., secondary spread pathways (Rup et al., 2010); zebra mussels are readily spread on infested docks, boats, or other items (Mackie & Schlosser, 1996)) but not CVT. Thus, our approach

identified species-specific relationships to APIs consistent with the known dispersal pathways.

Of additional note, is that our assumption of persistent NAS presence for decades following first presence, potentially made the dataset non-independent. However, with only five data points (decades) we could not conduct rigorous statistical analysis testing for temporal autocorrelation. As an alternative, we performed a repeated measures logistic regression analysis using the three pathway metrics as fixed effects and NAS presence for each decade as a random effect using the R-project package lme4 (Bates et al., 2015). We found that decade was indeed a significant effect ($p < 0.01$) for each iteration, suggesting each decade dataset was behaving independently even if the entire dataset was non-independent.

NAS Spread Patterns

Taken as a whole, our results demonstrate a general pattern of secondary spread from larger ports, cities, and marinas to medium and small locations as NAS were increasingly found throughout the Great Lakes and associated with an increasing number of pathway locations. Spread was pathway- and species-specific, and occurred at varying rates. By 2013, nearly all locations with high human activity were occupied by at least one of the evaluated NAS. By the end of the 1990s, there was at least a 40% chance of a NAS being present near any size city; by the end of 2013, this value had increased to greater than 70%. With respect to NAS richness, the influence of maritime commerce was roughly equal to that of city population size, such that busy ports near large population cities have the highest NAS richness throughout the Great Lakes. Moreover, the probability of NAS presence was homogenized through time and across pathways as the system became more invaded and the pathway intensity values associated with the p50 (e.g., the city size at which there is a 50% probability of NAS presence) decreased for most NAS and NAS groups. This is an important result, revealing that the probability of NAS detection in a kind of place (i.e., a medium city) changes through time as NAS spread continues, and thus both the association between a NAS and its original introduction pathway declines through time, and the need to manage NAS in a greater range of kinds of places increases.

Our results indicate that aquaria species are a relatively recent concern, particularly for large and medium size cities. Whereas the association between NAS and maritime commerce weakened through time, the association between city size and aquaria species emerged only in the 1990s and thereafter strengthened. The finding is consistent with the rapid increase in the international export value of ornamental fish during the 1980s and 1990s (Padilla and Williams, 2004). The potential of this group to be introduced and spread is disproportionately based in large population cities; when evaluated using multiple logistic regression to account for the commercial and recreational maritime activities pathways, 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. Note that most of the aquaria species analyzed are invertebrates and not fish; although aquaria fish present a relevant threat, especially in trophic climates (Padilla and Williams, 2004), many are tropical species that likely will not survive once introduced to temperate climates (Lapointe et al., 2016).

The multiple linear regression analysis between NAS richness and pathway intensity reveal an initial rapid increase in species richness in the 1970s and 1980s, then a constant rate of increase in NAS richness from the 1980s onward. This suggests that introductions have slowed in recent decades when compared to the analyses conducted by Ricciardi (2001), who found that the cumulative introductions of new NAS into the Great Lakes increased exponentially from 1810 to 1999. Additionally, our analysis indicates that propagule pressure has a measurable cumulative response (i.e., cumulative NAS richness was significantly related to pathway intensities through time). That is, NAS richness was greatest in those areas subject to sustained, high propagule pressure as indicated by population density and CVT. Copp et al. (2010) found that population density and nonindigenous fish species richness are correlated in England. In a regional study of coastal British Columbia, Canada, Clarke Murray et al. (2014) found that shipping activity, population density, and recreational boating could explain NAS richness of benthic invertebrates. Whether propagule pressure from multiple pathways can enhance NAS richness above and beyond a single pathway via the introduction of species with a variety of life histories remains a frontier for research.

In conclusion, by contrasting varying combinations of pathway intensities (by proxy) expressed in cities, ports, and marinas throughout our case study area, the Great Lakes region, we were able to examine the past and current relationship between major NAS introduction pathways and NAS detections. Developing geospatial metrics as proxies for specific NAS introduction pathways revealed large-scale spatial and temporal patterns associated with NAS detections. Despite the prominent role of commercial maritime activities in NAS introduction and spread, our study revealed that city population size is an important metric for explaining NAS presence and introduction patterns in the Great Lakes, and thus a potentially important metric for other aquatic systems globally. At present, nearly all medium and large size cities in the region harbor at least one NAS. Our findings suggest the possibility of transport of newly introduced NAS to areas around cities prior to first detection; future studies modelling nearshore currents around ports and cities or investigating patterns in both commercial and recreational vessel traffic would be valuable, in this regard. Further, we identified a temporal shift in the relative strength of specific pathway and NAS associations; city population size was an increasingly significant predictor of the presence of organisms in trade. This shift should inform future studies regarding which pathways are most relevant to current and future NAS introductions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Rochelle Sturtevant and Matt Cannister for providing nonindigenous aquatic species detection records from GLANSIS and the USGS NAS databases, Valerie Brady for providing nonindigenous aquatic invertebrate detections from the Great Lakes Indicator Consortium's Coastal Wetlands Monitoring project, and 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. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection

Agency. This is Contribution # 622 of the University of Minnesota's Natural Resources Research Institute publication series.

References

- Allan JD, McIntyre PB, Smith SDP, Halpern BS, Boyer GL, Buchsbaum A, Burton GA, Campbell LM, Chadderton WL, Ciborowski JJH, Doran PJ, Eder T, Infante DM, Johnson LB, Joseph CA, Marino AL, Prusevich A, Read JG, Rose JB, Rutherford ES, Sowa SP & Steinman AD, 2013 Joint Analysis of Stressors and Ecosystem Services to Enhance Restoration Effectiveness. *PNAS* 110: 372–377. [PubMed: 23248308]
- Ashton G Boos VK, Shucksmith R, & J. E and Cook, 2006). Rapid assessment of the distribution of marine non-native species in marinas in Scotland. *Aquatic Invasions* 1: 209–213.
- Bailey SA, Chan F, Ellis SM, Bronnenhuber JE, Bradie JN & 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.
- Bates D, Maechler M, Bolker B & Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67: 1–48.
- Bonferroni CE 1936 *Teoria Statistica delle Classi e Calcolo delle Probabilita*. Libreria internazionale Seeber.
- Blackburn TM, Pysek P, Bacher S, Carlton JT, Duncan RP, Jarosik V, Wilson JRJ & Richardson DM, 2011 A Proposed Unified Framework for Biological Invasions. *Trends in Ecology & Evolution* 26: 333–339. [PubMed: 21601306]
- Bossenbroek JM, Kraft CE & Nekola JC, 2001 Prediction of Long-distance Dispersal using Gravity Models: zebra mussel Invasions of Inland Lakes. *Ecological Applications* 11: 1778–1788.
- Bradie J, Chivers C & Leung B, 2013 Importing Risk: Quantifying the Propagule Pressure-establishment Relationship at the Pathway Level. *Diversity and Distribution* 19: 1020–1030.
- Branstrator DK, Beranek AE, Brown ME, Hembre LK & Engstrom DR, 2017 Colonization dynamics of the invasive predatory cladoceran, *Bythotrephes longimanus*, inferred from sediment records. *Limnology and Oceanography* 62: 1096–1110.
- Brown ME & Branstrator DK, 2004 A 2001 Survey of Crustacean Zooplankton in the Western Arm of Lake Superior. *Journal of Great Lakes Research*, 30: 1–8.
- Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW & Watson REA, 2010 Global Biodiversity: Indicators of Recent Declines. *Science* 328: 1164. [PubMed: 20430971]
- Carlton JT & Geller JB, 1993 Ecological Roulette: The Global Transportation of Nonindigenous Marine Organisms. *Science* 261: 78–82. [PubMed: 17750551]
- Colautti RI, & MacIsaac HJ, 2004 A neutral terminology to define 'invasive' species. *Diversity and Distributions* 10: 135–141.
- Copp GH, Vilizzi L & Gozlan RE, 2010 The demography of introduction pathways, propagule pressure and occurrences of non-native freshwater fish in England. *Aquatic Conservation: Marine and Freshwater Ecosystems* 20: 595–601.
- Clarke Murray C, Gartner H, Gregr EJ, Chan K, Pakhomov E & Therriault TW, 2014 Spatial distribution of marine invasive species: environmental, demographic, and vector drivers. *Diversity and Distributions* 20: 824–836.
- Crooks JA 2005 Lag Times and Exotic Species: The Ecology and Management of Biological Invasions in Slow-motion. *Ecoscience* 12: 316–329.
- Davis AJS & Darling JA, 2017 Recreational freshwater fishing drives non-native aquatic species richness patterns at a continental scale. *Diversity and Distributions* 23:692–702. [PubMed: 30147430]
- Duggan IC, Rixon CAM & MacIsaac HJ, 2006 Popularity and Propagule Pressure: Determinants of Introduction and Establishment of Aquarium Fish. *Biological Invasions* 8: 377–382.
- ESRI, 2014 ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.
- Foster V, Giesler RJ, Meriwether A Wilson W, Nall CR & Cook CJ, 2016 *Marine Biology* 163: 173. [PubMed: 27512234]

- Garton DW, Berg DJ & Fletcher RJ, 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* 47:731–738.
- Gertzen E, Familiar O & Leung B, 2008 Quantifying Invasion Pathways: Fish Introductions from the Aquarium Trade. *Canadian Journal of Fisheries and Aquatic Science* 65: 1265–1273.
- Gertzen E & 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: 2433–2444.
- Grigorovich IA, Kelly JR, Darling JA & West CW, 2008 The quagga mussel Invades the Lake Superior Basin. *Journal of Great Lakes Research* 34: 342–350.
- Gutsch M & Hoffman JC, 2016 A review of Ruffe (*Gymnocephalus cernua*) life history in its native versus non-native range. *Reviews in Fish Biology and Fisheries* 26:213–233.
- Hensler SR & Jude DJ, 2007 Diel Vertical Migration of round goby Larvae in the Great Lakes. *Journal of Great Lakes Research* 33: 95–302.
- Herborg LM, O'Hara P, P. & Therriault TW, 2009 Forecasting the Potential Distribution of the Invasive Tunicate *Didemnum vexillum*. *Journal of Applied Ecology* 46: 64–72.
- Holeck KT, Mills EL, MacIsaac HJ, Dochoda MR, Colautti RI & Ricciardi A, 2014 Bridging Troubled Waters: Biological Invasions, Transoceanic Shipping, and the Laurentian Great Lakes. *BioScience* 54: 919–929.
- Hoffman JC, Schloesser J, Trebitz AS, Peterson GS, Gutsch M, Quinlan H & Kelly JR, 2016 Sampling design for early detection of aquatic invasive species in Great Lakes ports. *Fisheries* 41:26–37.
- Hoffman JC, Kelly JR, Trebitz AS, Peterson GS & West CW, 2011 Effort and potential efficiencies for aquatic non-native species early detection. *Canadian Journal of Fisheries and Aquatic Sciences* 68:2064–2079.
- Kerr SJ, Brousseau CS & Muschett M, 2005 Invasive Aquatic Species in Ontario: A Review and Analysis of Potential Pathways for Introduction. *Fisheries* 30: 21–30.
- Kolar CS & Lodge DM, 2001 Progress in Invasion Biology: Predicting Invaders. *Trends in Ecology and Evolution* 16: 199–204. [PubMed: 11245943]
- 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.
- Lapointe NWR, Fuller PL, Neilson M, Murphy BR & Angermeier PL, 2016 Pathways of fish invasion in the Mid-Atlantic region of the United States. *Management of Biological Invasions* 7:221–233.
- Leprieux F, Beauchard O, Blanchet S, Oberdorff T & Brosse S, 2008 Fish Invasions in the World's River Systems: When Natural Processes Are Blurred by Human Activities. *PLOS Biology* 6(2): e28. [PubMed: 18254661]
- Leung B, Drake JM & Lodge DM, 2004 Predicting Invasions: Propagule Pressure and the Gravity of Allee Effects. *Ecology* 85: 1651–1660.
- Lo VB, Levings CD, & Chan KMA, 2012 Quantifying Potential Propagule Pressure of Aquatic Invasive Species from the Commercial Shipping Industry in Canada. *Marine Pollution Bulletin* 64: 295–302. [PubMed: 22206724]
- Lockwood JL, Cassey P & Blackburn TM, 2009 The More you Introduce the More you Get: The Role of Colonization Pressure and Propagule Pressure in Invasion Ecology. *Diversity and Distributions* 15: 904–910.
- Lodge DM, Stein RA, Brown KM, Covich AP, Bronmark C, Garvey JE & Klosiewski SP, 1998 Predicting Impact of Freshwater Exotic Species on Native Biodiversity: Challenges in Spatial Scaling. *Australian Journal of Ecology* 23: 53–67.
- Lodge DM, Williams S, MacIsaac HJ, Hayes KR, Leung B, Reichard S, Mack RN, Moyle PB, Smith M, Andow DA, Carlton JT & McMichael A, 2006 Biological Invasions: Recommendations for U.S. Policy and Management. *Ecological Applications* 16: 2035–2054. [PubMed: 17205888]
- Mackie GL & Schlosser DW, 1996 Comparative Biology of zebra mussels in Europe and North America: An Overview. *American Zoologist* 36: 244–258.
- Marty J, Bowen K, Koops MA & Power M, 2010 Distribution and Ecology of *Hemimysis anomala*, the Latest Invader of the Great Lakes Basin. *Hydrobiologia* 64: 71–80.

- McDowall RM, 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: 503–510.
- Mills EL, Leach JH, Carlton JT & Secor CL, 1993 Exotic Species in the Great Lakes: A History of Biotic Crises and Anthropogenic Introductions. *Journal of Great Lakes Research* 19: 1–54.
- Mills EL, Chrisman JR, Baldwin B, Owens RW, O’Gorman R, Howell T, Roseman EF & Rath MK, 1999 Changes in the Dreissenid Community in the Lower Great Lakes with Emphasis on Southern Lake Ontario. *Journal of Great Lakes Research* 25:187–197.
- Molnar JL, Gamboa RL, Revenga C & Spalding MD, 2008 Assessing the Global Threat of Invasive Species to Marine Biodiversity. *Frontiers in Ecology and the Environment* 6: 485–492.
- Moran PAP, 1948 The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society* 2: 243–251.
- 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.
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH & Wagner H, 2013 vegan: Community Ecology Package. R package version 2.0–10. <http://CRAN.R-project.org/package=vegan>
- Padilla DK & Williams SL, 2004 Beyond Ballast Water: Aquarium and Ornamental Trades as Sources of Invasive Species in Aquatic Ecosystems. *Frontiers in Ecology and the Environment* 2: 131–138.
- Phillips SJ, Dudik M, Graham CH, Lehmann A, Leathwick J & Ferrier S, 2009 Sample Selection Bias and Presence-only Distribution Models: Implications for Background and Pseudo-absence Data. *Ecological Applications* 19: 181–197. [PubMed: 19323182]
- R Core Team, 2014 R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria URL <https://www.R-project.org/>.
- Rao YR & Schwab DJ, 2007 Transport and Mixing between the Coastal and Offshore Waters in the Great Lakes: a Review. *Journal of Great Lakes Research* 33: 202–218.
- Ray WJ & Corkum LD, 2001 Habitat and Site Affinity of the round goby. *Journal of Great Lakes Research* 27: 329–334.
- 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: 2513–2525.
- Ricciardi A, 2006 Patterns of Invasion in the Laurentian Great Lakes in Relation to Changes in Vector Activity. *Diversity and Distributions* 12: 425–433.
- Rilov G, Benayahu Y & Gasith A, 2004 Prolonged lag in population outbreak of an invasive mussel: a shifting-habitat model. *Biological Invasions* 6: 347–364.
- Rixon CAM, Duggan IC, Bergeron NMN, Ricciardi A, & MacIsaac HJ, 2005 Invasion Risks Posed by the Aquarium Trade and Live Fish Markets on the Laurentian Great Lakes. *Biodiversity and Conservation* 14: 1365–1381.
- Rothlisberger JD & Lodge DM, 2013 The Laurentian Great Lakes as a Beachhead and a Gathering Place for Biological Invasions. *Aquatic Invasions* 4: 361–374.
- Rothlisberger JD, Chadderton WL, McNulty J & Lodge DM, 2011 Aquatic Invasive Species Transport via Trailered Boats: What is Being Moved, Who is Moving it, and What Can Be Done. *Fisheries* 5: 121–132.
- Rup MP, Bailey SA, Wiley CJ, Minton MS, Miller AW, Ruiz GM, & MacIsaac HJ, 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: 256–268.
- Sala OE, Chapin FS III, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huennek LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes MT, Walker BH, Walker M, & Wall DH, 2000 Global Biodiversity Scenarios for the Year 2100. *Science* 287: 1770–1774. [PubMed: 10710299]
- Sieracki JL, Bossenbroek JM, & Chadderton WL, 2014 A Spatial Modeling Approach to Predicting the Secondary Spread of Invasive Species Due to Ballast Water Discharge. *PLOS One*, Available online: 10.1371/journal.pone.0114217

- Statistics Canada. Available online: <http://www12.statcan.gc.ca/census-recensement/index-eng.cfm>. Accessed 2/5/2015.
- Tate RF, 1954 Correlation between a Discrete and a Continuous Variable. Point-Biserial Correlation. *The Annals of Mathematical Statistics* 25: 603–607.
- Trebitz AS Kelly JR, Hoffman JC, Peterson GS, & West CW, 2009 Exploiting habitat and gear patterns for efficient detection of rare and non-native benthos and fish in Great Lakes coastal ecosystems. *Aquatic Invasions* 4: 651–667.
- Trebitz AS & Hoffman JC, 2015 Coastal Wetland Support of Great Lakes Fisheries: Progress from Concept to Quantification. *Transactions of the American Fisheries Society* 144: 352–372
- 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. Geological Survey. Nonindigenous Aquatic Species database. Available online: <http://nas.er.usgs.gov> Accessed 2/5/2015.
- Uzarski DG, Brady VJ, Cooper MJ, Wilcox DA, Albert DA, Axler RP, Bostwick P, Brown TN, Ciborowski JJH, Danz NP, Gathman JP, Gehring TM, Grabas GP, Garwood A, Howe RW, Johnson LB, Lamberti GA, Moerke AH, Murry BA, Niemi GJ, Normant CJ, Ruetz CR III, Steinman AD, Tozer DC, Wheeler R, O'Donnell TK, & Schneider JP, 2017 Standardized Measures of Coastal Wetland Condition: Implementation at a Laurentian Great Lakes Basin-Wide Scale. *Wetlands* 37: 15–32.
- von Holle BV & Simberloff D, 2005 Ecological Resistance to Biological Invasion Overwhelmed by Propagule Pressure. *Ecology* 86: 3212–3218.
- Wells MG, Bailey SA & Ruddick B, 2011 The Dilution and Dispersion of Ballast Water Discharge in Goderich Harbor. *Marine Pollution Bulletin* 62: 1288–1296. [PubMed: 21440268]
- Wonham M, Byers JE, Grosholz ED & Leung B, 2013 Modeling the Relationship between Propagule Pressure and Invasion Risk to Inform Policy and Management. *Ecological Applications* 23: 1691–1706. [PubMed: 24261049]
- Yurista PM & Kelly JR, 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 Munawar M and Heath R. *Aquatic Ecosystem Health and Management Society*, Burlington, Ont., Canada, 439–471.

Fig 1 –.

Distribution of pathway intensity data for each of the three metrics of commercial maritime traffic, and the organisms in trade and recreational maritime 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.

Fig 2 –.

Areas of pathway influence (API) for the three pathways. Pathway intensity metrics were assigned to each API. Quartiles of the pathway intensity metrics' data are displayed on a gradient from white to black: 1st quartile (least intensity) – white; 4th quartile (highest intensity) – black; each API is also labeled based on the quartile. Sample size and data range are displayed in the upper right of each panel.

Fig 3 –.

Probability of NAS presence from simple logistic regression models. All displayed relationships were significant at $\alpha < 0.01$. 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 and is provided as a reference point for comparison within the figures.

Fig 4 –.

Probability of NAS group simple 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 and is provided as a reference point for comparison within the figures.

Fig 5 –.

Percent change in odds ratio for the multiple 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. Maritime commerce – port; population size – city; water recreation–marina.

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
Maritime Commerce	Commercial Vessel Trips	67	50	2004 – 2013	National Ballast Water Clearinghouse (2015)
	- Total				
	- Discharging				
	- Non-discharging				
	Annual Avg. Cargo (metric tons)	45	34	1996 – 2011	U.S. Army Corps of Engineers (2014)
	- Total				
	- Domestic				
	- Foreign				
	Annual Avg. Ballast water discharge (m ³)	97	65	2005 – 2007	National Ballast Water Clearinghouse (2015), Bailey et al. (2012)
	- Total				
	- From highly invaded domestic ports				
Water Recreation	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 detections within 10km of shore from 1970 to 2013. 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: IND - Individual species model; NAS_{Aq} - Aquaria species model; * - included in both First presence (FP) and all NAS (NAS_{All}) models; all NAS listed were included in FP and NAS_{All} models even if not noted with an *. Regression analyses were conducted using subsets of this dataset in various models: three individual species models, two group species models, and one first presence model.

Scientific name	Common name or description	Number of Records	Year Range	Lake	Regression Model
<i>Dreissena polymorpha</i> (Pallas, 1771)	zebra mussel	1088	1986 – 2013	S,H,M,E,O	IND
<i>Neogobius melanostomus</i> (Pallas, 1814)	round goby	670	1990 – 2013	S,H,M,E,O	IND
<i>Cyprinus carpio</i> (Linnaeus, 1758)	common carp	350	1971 – 2013	S,H,M,E,O	*
<i>Gymnocephalus cernua</i> (Linnaeus, 1758)	eurasian ruffe	272	1987 – 2013	S,H,M	*
<i>Dreissena bugensis</i> (Andrusov, 1897)	quagga mussel	243	1989 – 2011	S,H,M,E,O	IND
<i>Bythotrephes longimanus</i> (Leydig, 1860)	spiny waterflea	117	1984 – 2010	S,H,M,E,O	*
<i>Carassius auratus</i> (Linnaeus, 1758)	goldfish	85	1975 – 2013	S,H,M,E,O	NAS _{Aq}
<i>Bosmina coregoni</i> (Baird, 1857)	waterflea	71	1994 – 2006	S,H,M,E,O	*
<i>Hemimysis anomala</i> (G. O. Sars, 1907)	bloody red shrimp	59	2006 – 2011	H,M,E,O	*
<i>Echinogammarus ischnus</i> (Stebbing, 1899)	scud	57	1994 – 2008	S,H,M,E,O	*
<i>Ewytenora affinis</i> (Poppe, 1880)	calanoid copepod	56	1994 – 2006	S,H,M,E,O	*
<i>Cercopagis pengoi</i> (Ostroumov, 1891)	fishhook waterflea	52	1998 – 2006	S,H,M,E,O	*
<i>Orconectes rusticus</i> (Girard, 1852)	rusty crayfish	47	1981 – 2011	S,H,M	*
<i>Potamopyrgus antipodarum</i> (J.E. Gray, 1853)	New Zealand mud snail	40	1991 – 2008	S,H,E,O	*
<i>Bithynia tentaculata</i> (Linnaeus, 1758)	faucet snail	35	1974 – 2012	S,H,M,E,O	*
<i>Craspedacusta sowerbyi</i> (Lankester, 1880)	freshwater jellyfish	28	1971 – 2003	H,M,E	NAS _{Aq}
<i>Corbicula fluminea</i> (O.F. Muller, 1774)	asian clam	14	1980 – 2012	S,H,M,E	*
<i>Daphnia lumholzi</i> (G. O. Sars, 1885)	waterflea	10	1999 – 2007	S,H,E	*
<i>Cipangopaludina chinensis</i> (Gray, 1834)	Chinese mystery snail	9	1974 – 2008	H,M,E,O	NAS _{Aq}
<i>Radix auricularia</i> (Linnaeus, 1758)	european ear snail	5	1974	M,H,E	NAS _{Aq}
<i>Cordylophora caspia</i> (Pallas, 1771)	freshwater hydroid	2	2001 – 2007	S, M	NAS _{Aq}
<i>Lophodolabella carteri</i> (Hyatt, 1865)	freshwater bryozoan	1	1995	M	NAS _{Aq}
<i>Colossoma macropomum</i> (Cuvier, 1816)	pacu	1	2013	E	NAS _{Aq}
Total		3294			

Table 3 –

Comparison of eight maritime commerce metrics within the multiple logistic regression models to determine which metric of maritime commerce was most explanatory. Supplemental Table 1 provides the simple logistic regression analyses results for each of the eight maritime commerce metrics. The significance ($\alpha = 0.01$) and overall model fit was determined for each of the maritime commerce metrics; only the results for the maritime commerce metric are displayed. There were no significant relationships in the aquaria species models (omitted). Model fit values ($-2 \times \log(\text{likelihood})$) are displayed for significant results (*); p-values are displayed for non-significant results. Trips were categorized as total trips, 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). Simple logistic regression modeling (Supplemental Table 1) produced similar results.

A. Commercial Vessel Trips		Total			Discharging			Non-discharging			
Model		70s	80s	90s	00s	10s	70s	80s	90s	00s	10s
All species	0.086		339*	632*	610*	613*	0.145	341*	638*	629*	610*
Quagga mussel				0.443	0.900	0.310			0.509	0.944	0.330
Round goby				324*	473*	514*			326*	480*	520*
Zebra mussel			210*	615*	612*	621*		215*	620*	617*	625*
B. Commercial Cargo		Total			Domestic			Foreign			
Model		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
Quagga mussel				0.778	0.935	0.460			0.759	0.924	0.462
Round goby				326*	488*	530*			326*	488*	529*
Zebra mussel			212*	0.080	0.054	0.626*		212*	0.075	0.051	0.627*
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	0.115*	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	0.623*		211*	0.180	0.149	0.025

Table 4 –

Model fit ($-2 \times \log(\text{likelihood})$), regression coefficient, and standard error for simple logistic regression models. Values of $-2 \times \log(\text{likelihood})$ closer to '0' indicate a better fitting model. The p-value was calculated using the likelihood ratio test statistic to test for significance. Significance was determined at $\alpha = 0.01$. 'NS' denotes that the relationship was not significant.

Pathway Metric Model	1970s	1980s	1990s	2000s	2010s
Total Commercial Vessel Trips	All species	62.8 / 1.0 (0.4)	NS	NS	NS
	Aquaria species	NS	NS	NS	NS
	Quagga mussel	NS	NS	NS	NS
	Round goby	63.7 / 0.9 (0.4)	46.8 / 1.0 (0.4)	39.4 / 1.3 (0.5)	NS
	Zebra mussel	NS	NS	NS	NS
Population Size	All species	50.8 / 1.7 (0.4)	83.3 / 2.6 (0.8)	67.6 / 2.0 (0.8)	61.7 / 1.7 (0.8)
	Aquaria species	NS	108.1 / 0.6 (0.2)	117.3 / 0.9 (0.2)	123.9 / 1.0 (0.2)
	Quagga mussel	NS	122.1 / 0.7 (0.2)	131.3 / 1.2 (0.3)	135 / 1.1 (0.2)
	Round goby	NS	133 / 0.9 (0.2)	157.6 / 0.6 (0.2)	146.2 / 0.8 (0.3)
	Zebra mussel	89.6 / 1.3 (0.3)	93.6 / 3.0 (0.7)	93 / 3.0 (0.7)	90.4 / 3.0 (0.8)
	All species	NS	369 / 1.3 (0.2)	384.4 / 1.3 (0.2)	393.3 / 1.2 (0.2)
Marina Size	Aquaria species	NS	NS	NS	NS
	Quagga mussel	NS	NS	NS	NS
	Round goby	152.4 / 1.9 (0.4)	220.3 / 1.7 (0.3)	261.8 / 1.3 (0.3)	NS
	Zebra mussel	101.1 / 2.0 (0.6)	329.4 / 1.5 (0.2)	330.4 / 1.5 (0.2)	344.4 / 1.4 (0.2)

Table 5 –

Regression coefficients from the multiple linear regression of NAS richness by decade. All coefficients were statistically significant at $p < 0.01$; 513 degrees of freedom.

	Regression Coefficient			
	1970s	1980s	1990s	2010s
Commercial Vessel Trips	0.287	0.337	0.171	0.153
Population size	0.612	0.372	0.211	0.240
Marina Size	-0.621	-0.411	-0.326	-0.365
				-0.376