Biol Invasions (2014) 16:1043–1053
DOI 10.1007/s10530-013-0556-2

ORIGINAL PAPER

Modeling the secondary spread of viral hemorrhagic septicemia virus (VHSV) by commercial shipping in the Laurentian Great Lakes

Jennifer L. Sieracki · Jonathan M. Bossenbroek · Mohamed Faisal

Received: 7 February 2013 / Accepted: 18 September 2013 / Published online: 26 September 2013

Abstract Researchers have only begun to study the role of shipping in the spread of invasive species in the Laurentian Great Lakes despite a well-documented history of introductions in these lakes due to ballast water release. Here, we determine whether ballast water discharge was a likely vector of spread of the fish disease, viral hemorrhagic septicemia virus genotype IVb (VHSV-IVb), throughout the Great Lakes and St. Lawrence Seaway. Three models were developed to assess whether the spread of VHSV was due to (1) chance (random model), or (2) ballast water discharge (location model), and whether (3) increased propagule pressure, as measured by the number of visitations by ships carrying ballast water from VHSV infected areas, increased the likelihood of a discharge location becoming infected with VHSV (propagule pressure model). The third model was also used to assess the probable point of initial introduction of VHSV. Presence and absence accuracies and weighted Cohen's kappa were calculated to determine which models best predicted observed presences and absences of VHSV. Location models explain the patterns of VHSV detections better than random models, and inclusion of "propagule pressure" often improved model fit; however, the relationship is weak likely because of a long lag time between introduction and detection, a high rate of false negatives in reporting, and the possible contribution of other vectors of spread. Montreal was also identified as the more likely introduction site of VHSV, rather than Lake St. Clair, the site where the virus was first detected.

Keywords Ballast water management · Invasive species · Model uncertainty · Organism detection · Spatial modeling · Viral hemorrhagic septicemia virus (VHSV)

Introduction

Commercial ship ballast water has been identified as a major component of non-native species spread globally (Molnar et al. 2008). For example, in the Laurentian Great Lakes, 55 % of non-native species found are believed to have been introduced by ballast water since the opening of the St. Lawrence Seaway in 1959 (NOAA Undated). Commercial ships can carry between millions and billions of living organisms (i.e.

Electronic supplementary material The online version of this article (doi:10.1007/s10530-013-0556-2) contains supplementary material, which is available to authorized users.

J. L. Sieracki (✉) · J. M. Bossenbroek
Department of Environmental Sciences, Lake Erie Center, University of Toledo, 6200 Bayshore Road, Oregon, OH 43616, USA
e-mail: sier0010@gmail.com

M. Faisal
Aquatic Animal Health Laboratory, Department of Pathobiology and Diagnostic Investigation, Michigan State University, East Lansing, MI 48824, USA

propagules) in just 1 L of their ballast water (Drake et al. 2007; Leichsenring and Lawrence 2011; Ruiz et al. 2000). Even ships defined as carrying "no ballast on board" (NOBOB) may contain residual water and sediments harboring microorganisms (Drake et al. 2007). Not only can ships bring new species into the Great Lakes, but they have moved these species within the Great Lakes basin (Griffiths et al. 1991). "Secondary spread" of an invasive species, or the spread that occurs after the introduction of a species to a new region, can be a major contributor to dispersal within a region (Rup et al. 2010). Herein, we examine the role of shipping as a vector of secondary spread of viral hemorrhagic septicemia virus (VHSV) within the Great Lakes.

VHSV is a fish rhabdovirus that infects a wide range of fish species in North America, Europe, and Asia and is believed to have been introduced to the Great Lakes either via ballast water or migratory fish (Bain et al. 2010). VHSV has led to large fish kills, both in aquaculture and the wild (Kim and Faisal 2011; World Organisation for Animal Health 2011) and was first identified in eastern Lake Ontario in 2005 (Lumsden et al. 2007). Subsequent review of a rhabdovirus previously isolated from muskellunge in 2003 places the first verified record of VHSV in Lake St. Clair in 2003 (Elsayed et al. 2006; Faisal et al. 2012). The Great Lakes genotype of the virus was identified as being related to the North American and Japanese genotype (IVa); however, was distinct enough to be placed in its own sublineage (IVb) (Elsayed et al. 2006; Faisal et al. 2012). Since 2005, VHSV-IVb has spread rapidly across all five Great Lakes, with detections in Lakes Erie and Huron in 2006, Lake Michigan in 2007, and as far west as Duluth/Superior harbors in Lake Superior in 2009 (Fig. 1). Despite a lack of detections prior to 2003, recent genetic research suggests that the virus may have been in the freshwaters of the Laurentian Great Lakes much earlier (Pierce and Stepien 2012). Moreover, there are eleven genetically distinct populations, or isolates, of the IVb strain found only in the Great Lakes and a few nearby inland waters (Pierce and Stepien 2012; Thompson et al. 2011). One of the isolates, U13653 (or vcG002), was originally found in eastern Lake Ontario and is the second most prevalent and widespread isolate as compared to the one originally found in Lake St. Clair (MI03GL) in 2003 (Pierce and Stepien 2012; Thompson et al. 2011). The prevalence of the U13653 isolate suggests the initial

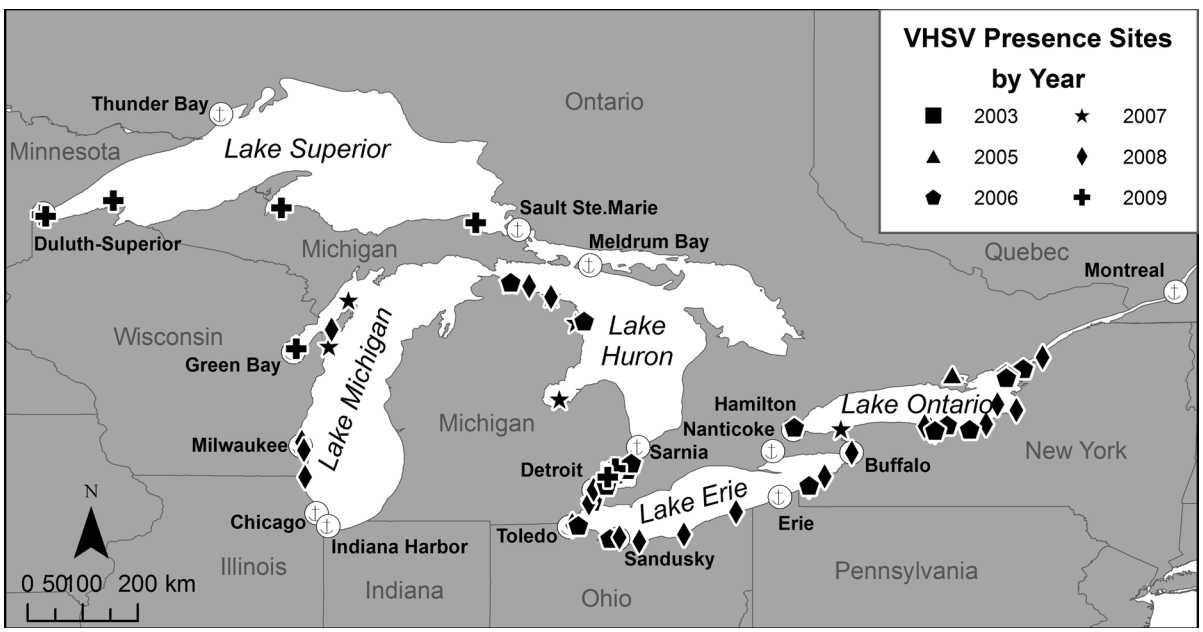

Fig. 1 VHSV presence locations in the Great Lakes for 2003–2009. No known occurrences of VHSV are found east of the Thousand Islands area until the Atlantic coast. *Squares*, *triangles*, *pentagons*, *stars*, *diamonds*, and *crosses* represent VHSV occurrences for 2003 and 2005–2009 respectively

introduction of VHSV to the Great Lakes occurred via the St. Lawrence River. Since MI03GL and U13653 only diverge by one mutational step and both have been isolated from fish in eastern Lake Ontario, this hypothesis seems plausible (Pierce and Stepien 2012; Thompson et al. 2011). Regardless of genetic sequence, VHSV-IVb has become rapidly widespread in the Great Lakes.

One of the reasons VHSV-IVb has been successful in invading the Great Lakes is because of the presence of environmental conditions that are favorable for the transmission of the virus. VHSV-IVb is particularly likely to spread in fish populations with high densities that are experiencing stress, which usually occur when fish come together during spawning (Kane-Sutton et al. 2010). In the Great Lakes, many fish spawn in the spring and early summer, when temperatures are ideal for the transmission of VHSV-IVb (Eckerlin et al. 2011; Kane-Sutton et al. 2010; Kim and Faisal 2011). Additionally, VHSV has been found to survive in freshwater for up to 14 days at 15 °C and 20 days at 10 °C under controlled conditions (Hawley and Garver 2008; Kim and Faisal 2011), indicating that the virus may be carried by water currents for several days in the spring and fall.

While the actual dispersal capabilities of VHSV are relatively unknown, it is unlikely that it was able to invade the full length of the Great Lakes in such a short period without a human-mediated, long-distance vector of spread. On the other hand, Bain et al. (2010) found no relationship between VHSV occurrences and locations identified as "shipping centers". We thus hypothesize that commercial shipping may have been a vector of spread throughout the Great Lakes for VHSV. Ships in the Great Lakes generally draw in and discharge ballast water at ports as they unload and load cargo (Eames et al. 2008). They may also adjust their ballast mid-lake during bad weather and when entering connecting channels and rivers (Cangelosi and Mays 2006). This allows for many opportunities to pick up, move, and discharge invasive species. Moreover, because ships travelling exclusively within the Great Lakes make trips that happen over a short period of time, survival of invasive species may be greater than in those ships coming from outside the Great Lakes (Rup et al. 2010).

Here we set out to assess whether shipping played a role in the secondary spread of VHSV and whether we could use shipping spread models to identify the most likely location of initial VHSV introduction. To assess the role of Great Lakes shipping in the secondary spread of VHSV, we developed two primary questions: (1) Are VHSV occurrences related to the location and amount of ballast water being discharged throughout the Great Lakes?; (2) Is it possible to identify the site of initial introduction of VHSV based on ballast water discharge patterns? To answer the first question we developed three dynamic spatial models. The first two models, a random model and a location model, were built to determine if VHSV is related to ballast water discharge locations. The third model, a propagule pressure model, was built to determine if the number of visits from possibly infected ships increases the likelihood VHSV will become established at a discharge location. To answer the second question, the initial introduction location was changed in the propagule pressure model to identify the infection source that best fits the observed VHSV occurrences. Lake St. Clair was chosen as an initial introduction location, since it is the earliest detection of VHSV-IVb. Montreal was selected as a second possibility in order to determine if VHSV may have been introduced via the St. Lawrence River instead. By answering these questions, we hope to establish if Great Lakes shipping has been responsible for secondary spread of VHSV throughout the Great Lakes, and if Lake St. Clair was the first site of introduction in the Great Lakes.

Methods

Site description

The Laurentian Great Lakes and the St. Lawrence Seaway are the areas of interest for the study. We defined the St. Lawrence Seaway as being the portion of the St. Lawrence River from the western edge of Anticosti Island west to its source at Lake Ontario.

Spatial modeling

We developed three competing models to assess the role that shipping plays in the spread of VHSV. Each of the models were run to simulate the spread of the virus from 2003 to 2009 and had the same basic structure for each year of the model: (1) the number of VHSV introductions and their locations were selected

using different stochastic processes, (2) each new introduction location was converted to an infection area based on the assumption that VHSV occurs in an area and not at a given point as identified by the presence data, and 3) the area of infection was further increased in each year to simulate the spread of the virus via natural means, such as by currents and fish hosts. The results of the models were areas of "predicted" infection, which were compared to the observed VHSV presence and absence data to assess model fit.

Our three models primarily differed in the way annual infection locations were chosen (i.e. Step 1 from above). The "random model" identified annual infection locations by randomly selecting locations throughout the entire study area. The number of infection locations was randomly drawn from a Poisson distribution with λ, the mean and variance of Poisson distributions, equal to the mean number of actual VHSV infections reported for the years 2003–2009. The total number of VHSV detections was 56, so $\lambda = 8$. For the "location model", the number of infections per year was selected from a Poisson distribution as above; however, the infection locations were selected randomly only from known ballast water discharge locations.

The third model, the "propagule pressure model", was more complex and included data on VHSV sources, destinations, and number of trips made between source and discharge sites. As the first known location of VHSV-IVb was Lake St. Clair in 2003 (Elsayed et al. 2006), our first models initiated VHSV infection at that location. If a ship was identified as picking up ballast water in an area known to have VHSV, that ship was identified as carrying infected ballast water. Discharge locations receiving water from those infected ships in that year were next selected as possible locations of new VHSV infections. The total number of infected ships discharging at each location was calculated for each destination location. To determine if the discharge locations receiving at least one visit from an infected ship would become infected with VHSV that year we used a binomial distribution to identify if, for each ship visit, the discharge location became infected with VHSV. The number of binomial trials was equal to the total number of infected ships that discharged at the location in that year. The probability of infection for each binomial trial was calculated for each port based on a decay curve of virus-like particles (VLPs):

$$p(VLP) = 1 - e^{-0.11x}$$

where $p(VLP)$ = proportion of VLPs remaining and x = day of the trip (Lovell and Drake 2009). Because of the lack of data on niche availability or the probability of establishment at each port, $p(VLP)$ served as both the probability of infestation and the probability of establishment. Additional single probability values of 0.50 and 0.01 were tested as representing the probabilities of infestation and establishment; however, little improvement in model fit was detected and model ranks were unchanged. The number of days a trip took was determined by calculating the mode of the number of days for each trip between ballast water source and discharge locations. If one or more of the infected ship visits at each discharge location resulted in infection (i.e., at least one binomial trial = 1), then that location was identified as being infected.

The random and location models were built to test the hypothesis that VHSV occurrences are related to discharge locations, while the propagule pressure model was built to test the hypothesis that infection locations are related to the amount of ballast water discharge being released at each location. The propagule pressure model was also revised to identify if another location besides Lake St. Clair may have been a likely initial source of VHSV.

All three models include parameters that simulate the possible area of infection due to natural spread once VHSV has been introduced to a particular location (i.e. steps 2 and 3 above). It has been estimated that at least one strain of VHSV is capable of being moved outside of a host in seawater for up to 2-km (Meyers and Winton 1995). This distance might be somewhat arbitrary, as it depends on water current and wind which vary spatially and temporally; however, it was used as a reasonable estimate for identifying how far from a presence location VHSV may actually be found. The area created by a 2-km radius from the presence location was identified as the initial area of infection. Beyond the initial area of infection, it is unknown how far fish or currents carry the virus in any given year, so three distance values were tested to simulate the distance VHSV would travel per year. Buffers of 10-, 20-, or 30-km radii

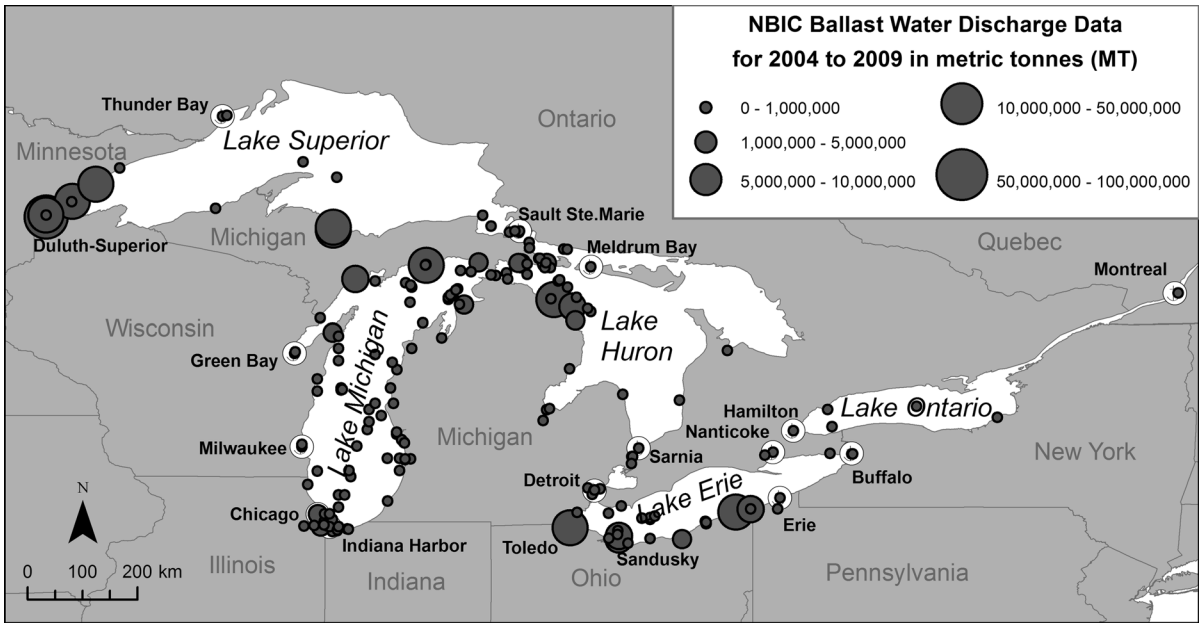

Fig. 2 Location and amount of ballast water discharged by ships arriving at US ports between 2004 and 2009 (from NBIC). Only discharge events involving ballast water picked-up in the Great Lakes are included. *Circles* of increasing size represent the amount of ballast water discharged at a location. Despite only receiving small amounts of ballast discharge from within the Great Lakes, the ports and river in and around Montreal receive large amounts of ship traffic (National Research Council 2008)

were added to the infection areas every year to simulate the natural spread of the virus. Distances beyond 30-km were not considered, as VHSV would be predicted to have spread to the entire Great Lakes within the 7 years of infection modeled.

All models require VHSV occurrence data and all but the random model requires ballast water source and/or discharge location data. The VHSV occurrence data was collected from a variety of sources, including the Nonindigenous Aquatic Species (NAS) database (USGS 2009), Department of Pathobiology and Diagnostic Investigation in the College of Veterinary Medicine at Michigan State University (2011), Cornell University (2010), and Minnesota and Wisconsin Department of Natural Resources (2010; Fig. 1). Other occurrence data were either unattainable or unidentified. Unattainable data included more recently published occurrences of VHSV in *Diporeia* spp. in Lakes Michigan and Ontario and in piscicolid leeches (*Myzobdella lugubris*) collected from Lakes St. Clair and Erie (Faisal et al. 2012; Faisal and Winters 2011). Both presences and absences were collected for the years 2003–2009 and were identified in all five Great Lakes, Lake St. Clair and its connecting waterways, and the St. Lawrence River in the Thousand Islands area. Ballast water source and discharge locations and number of trips were obtained from the National Ballast Information Clearinghouse (NBIC) data for 2004–2009 (Smithsonian Environmental Research Center and USCG 2009; Fig. 2). The NBIC requires the reporting of the last location of ballast water pick-up (i.e. source information) and the location where that ballast water and potential propagules were then discharged for each individual ship. Source and discharge information was recorded at the US port of arrival based on the NBIC data. All records containing source and/or discharge locations outside the Great Lakes were deleted. Remaining source and discharge locations were mapped using coordinates when available and location descriptions. Coordinates were obtained for location descriptions that included port and city names where possible. All other discharge and source points were located using topographic maps and aerial photographs. Four source locations (27 ship visits) were excluded from the data due to unclear location descriptions.

To identify the possibility of another likely location for the introduction of VHSV to the Great Lakes, the

propagule pressure model was modified to initiate VHSV infection of the Great Lakes from Montreal. Due to recent genetic research by Thompson et al. (2011) and Pierce and Stepien (2012), we hypothesized that it was possible that VHSV may have initially been introduced to the St. Lawrence River. We chose Montreal as a possible introduction location since it is located on a part of the river that receives a large amount of ship traffic (National Research Council 2008). In particular, Montreal receives a large amount of traffic from the Atlantic coast of Canada, where VHSV-IVc, a closely related strain to VHSV-IVb was identified in 2000, 2002, and 2004 (Pierce and Stepien 2012). All strains of VHSV are hypothesized to have originated from a marine reservoir in the North Atlantic Ocean (Thompson et al. 2011; Pierce and Stepien 2012), and Strain IV appears to have originated specifically in the Northwest Atlantic Ocean (Pierce and Stepien 2012). Despite not receiving a large amount of ballast water sourced within the Great Lakes from ships visiting US ports (Fig. 2), Montreal receives numerous ship visits from areas where VHSV-IVb potentially could have originated.

All models were run for each natural spread distance (10-, 20-, and 30-km). Each model was built in the ArcGIS Model Builder and run for 100 iterations. A single iteration was comprised of a seven-year simulation (i.e. 2003–2009) with each year adding to the spread of the virus identified in the previous year. The predictions of the models were compared to the actual VHSV presence/absence locations for 2003–2009. The models have been exported to Python and included in Appendix 1 (Online Resource 1).

Analyses of model performance

In order to analyze the performance of the models, presence accuracy (i.e., sensitivity), absence accuracy (i.e., specificity), and weighted kappa were calculated for each iteration of each model. A confusion matrix was built for each iteration to identify the number of true positives and negatives and false positives and negatives produced by each model and to calculate the above measurements (Fielding and Bell 1997; Manel et al. 2001). The models' abilities to accurately predict presences and absences were calculated for each model iteration (Fielding and Bell 1997; Manel et al. 2001).

Table 1 Level of agreement for each range of kappa values (Landis and Koch 1977; Gilchrist 2009)

Range of kappa value	Level of agreement
≥0.81	Almost perfect
0.61–0.80	Substantial
0.41–0.60	Moderate
0.21–0.40	Fair
0.00–0.20	Slight
<0.00	None

To determine the level of agreement between model predictions and actual VHSV presences and absences while correcting for chance we used a weighted Cohen's kappa statistic (Cohen 1968; Warrens 2011). The weighted kappa allows for weights to be applied to each cell in a confusion matrix, so that those cells calculated with data that is more uncertain than others will have less affect on the kappa statistic. We used a weighted kappa, as opposed to other calculations of fit (e.g. Cohen's kappa and AUC), due to the high false negative rate of the cell culture technique most frequently used in testing for VHSV. Despite cell culture being useful for identifying VHSV in fish that are carrying the active (positive-strand) virus (i.e. most likely to shed the disease), it was important that we identify all VHSV locations, even where the virus was inactive. A high false negative rate reduced our confidence in any reported absences. In experiments testing human viruses, cell culture was found to have false negative rates of 66–76 % (Covalciuc et al. 1999; Wald et al. 2003). While not all of the VHSV presence/absence data were identified using cell culture tests, Hope et al. (2010) found even the more sensitive qRT-PCR test that was used on the remaining data did not detect VHSV in all fish exhibiting clinical signs of the infection. Because of this, an estimated false negative rate of 66 % was used for our analysis.

Weighted kappa is calculated from the weighted proportions of observed and chance data for each cell of the confusion matrix. For our data, $w_{11} = 1.00$ (true positives) and $w_{22} = 0.33$ (true negatives). The weight for true negatives was based on the range of cell culture false negative rates. In order to test the sensitivity of the estimated false negative rate, weighted kappas were also calculated with $w_{22} = 0.50$ and 0.67. A level of agreement was

assigned to each range of kappa values (Table 1; Gilchrist 2009; Landis and Koch 1977).

Presence accuracy, absence accuracy, and weighted kappa were calculated for each iteration, and averaged for comparison. Standard deviations were calculated for all means. In total, fifteen models were tested: random, location, Lake St. Clair only propagule pressure, Montreal only propagule pressure, and Lake St. Clair and Montreal propagule pressure models each run with 10-, 20-, and 30-km spread distances.

Results

The results of the weighted Cohen's kappa statistics indicate that VHSV spread is not random and that VHSV occurrences are related to ballast water discharge locations (Table 2) although the strength of inference was slight. The location models tended to have higher presence accuracy at each spread distance than the random models (Table 3), indicating that the location models were better able to predict the presence of VHSV than the random models. Random models were better at predicting absences (Table 4); however, the location models were found to perform better overall with weighted kappas of 0.03, 0.04, and 0.05 at the 10-, 20-, and 30-km spread distances respectively (Table 2). All random models had weighted kappas between −0.04 and 0.00, suggesting that these models performed worse or equal to what would be expected by chance. Sensitivity analyses of the weighted w_{22} parameter only produced slight changes in the weighted kappa results with location models still performing better than random models. These results indicate that the spread of VHSV is related to ballast water discharge locations.

Further, locations that receive ballast water from infected ships were more likely to become infected with VHSV (Table 2). Most of the propagule pressure models performed better than the random and location models (Table 2). Sensitivity analyses of weighted kappas produced slightly higher measures of fit for most of the propagule pressure models, still resulting in better performance than the random and location models. Also, even though absence accuracies were generally lower than what was calculated for the random and location models (Table 4), presence

Table 2 Mean weighted Cohen's kappa for each model tested

Model	Natural spread distance		
	10-km	20-km	30-km
Random	−0.04 (0.12)	−0.02 (0.12)	0.00 (0.09)
Location	0.03 (0.07)	0.04 (0.08)	0.05 (0.07)
Propagule pressure			
Lake St. Clair only	−0.14 (0.01)	0.03 (0.00)	0.11 (0.00)
Montreal only	0.07 (0.01)	0.13 (0.01)	0.12 (0.00)
Lake St. Clair and Montreal	0.09 (0.01)	0.12 (0.00)	0.11 (0.00)

Weighted Cohen's kappa is the proportion of agreement corrected by chance between the model predictions and actual presence/absence data (Cohen 1968; Warrens 2011). Numbers in parentheses are standard deviations

Table 3 Mean presence accuracies for each model tested

Model	Natural spread distance		
	10-km	20-km	30-km
Random	0.23 (0.10)	0.50 (0.15)	0.71 (0.15)
Location	0.39 (0.07)	0.61 (0.08)	0.73 (0.10)
Propagule pressure			
Lake St. Clair only	0.38 (0.01)	0.64 (0.00)	0.77 (0.00)
Montreal only	0.65 (0.00)	0.73 (0.00)	0.79 (0.00)
Lake St. Clair and Montreal	0.65 (0.00)	0.72 (0.00)	0.77 (0.00)

Presence accuracies were calculated as the total number of actual presences that were accurately identified each year. Numbers in parentheses are standard deviations

Table 4 Mean absence accuracies for each model tested

Model	Natural spread distance		
	10-km	20-km	30-km
Random	0.80 (0.07)	0.53 (0.11)	0.30 (0.10)
Location	0.58 (0.08)	0.34 (0.09)	0.20 (0.07)
Propagule pressure			
Lake St. Clair only	0.72 (0.00)	0.28 (0.00)	0.02 (0.00)
Montreal only	0.25 (0.01)	0.08 (0.01)	0.02 (0.00)
Lake St. Clair and Montreal	0.18 (0.01)	0.07 (0.00)	0.02 (0.00)

Absence accuracies were calculated as the total number of actual absences that were accurately identified each year. Numbers in parentheses are standard deviations

accuracies were typically higher (Table 3). Additionally, propagule pressure models resulted in less variation overall (Tables 2, 3, and 4), since it repeatedly selected those locations that received large numbers of ship visits.

Not only do the results support the hypothesis that ports receiving more visits by infected ships are more likely to become infected, but they also indicate that Montreal is a more likely initial introduction location for VHSV (Table 2). The best performing model was the Montreal only 20-km model, even when considering the results of the weighted kappa sensitivity analysis. Additionally, combining Lake St. Clair and Montreal as simultaneous initial introduction locations produced very little change in the weighted kappas achieved by the Montreal only models (Table 2).

Discussion

The spread of VHSV within the Great Lakes has been aided by the secondary spread of ballast water. Though our model fit was only "slight" (based on the kappa scale used), the best fit model that we compared included the location, source, and amount of ballast water discharged, suggesting that these parameters are important indicators for identifying future VHSV infections. Furthermore, the results of our models also reveal that Lake St. Clair is a less likely initial introduction location of VHSV to the Great Lakes than Montreal. We did not test other locations due to lack of information indicating alternatives; however, our results show that it is possible to use the model to identify locations that tend to be areas of initial introduction to the Great Lakes.

The performance of our models may have been limited in part by the data used for model validation and the quality of the data included in our model. For one, the tests that were used to detect VHSV have a high false negative rate (Chico et al. 2006; Covalciuc et al. 1999; Hope et al. 2010; Miller et al. 1998; Wald et al. 2003). Additionally, many absences that were identified were in areas where VHSV had been identified previously, suggesting that the potential to infect existed, but VHSV was not detected in the individual fish that was tested. For instance, the Minnesota Department of Natural Resources had no positive tests for VHSV in the St. Louis River estuary between 2006 and 2010; however, researchers from Cornell University detected the virus in 2009. While we attempted to overcome this issue by measuring model fit using a weighted kappa statistic, absences that are not actually absences may have still been overly considered in the model. Incorrectly identified absences also would have been incorrectly identified for all remaining years in the model run. Error propagation would have affected both absence accuracy and weighted kappa statistics. The location from which infected fish were collected may also have added uncertainty to the presence/absence data. While many fish were collected live during monitoring efforts, others were collected during fish kills. Fish collected during fish kills would have mostly been found washed up on shore and likely far from the location where VHSV was actually contracted. Finally, the lack of Canadian ballast water data led to an incomplete dataset. This prevented us from establishing the complete pattern of ballast water movement in the Great Lakes. Whereas the limitations in the data used may not have been biased towards reducing the fit of any particular model over the other, it did prevent the accurate assessment of each model's ability to capture the past spread of VHSV.

Despite the limitations of the data used and the "slight" fit of even the best performing model, the pattern of secondary spread in the Great Lakes still indicates that shipping has played a role in the long-distance dispersal of VHSV. This is indicated by the Montreal models' abilities to capture VHSV presences at a higher rate than all of the other models. Further, at the best fit spread distance of 20-km, models that included ballast water discharge as a component of spread were able to explain the occurrence of VHSV at Duluth/Superior harbor at a much higher rate than the random models. In fact, the Montreal only model was the only model that correctly identified it 100 % of the time at the 20-km spread distance. The only presences that the Montreal only model fails to predict with regularity are located in eastern Lake Ontario, a part of the Great Lakes that receives very little ballast discharge. However, if the St. Lawrence River is the actual source of VHSV, the virus has potentially persisted in eastern Lake Ontario longer than in other parts of the Great Lakes, leading to greater localized spread of the virus due to natural vectors. Other vectors of spread that have been identified are bait fishing and fish stocking, which potentially could

contribute to long-distance spread along with ballast discharge. Nevertheless, we hypothesize that if bait and fish stocking were larger contributors to the long-distance spread of VHSV, more inland occurrences of the virus would have been detected. To date, only four inland waters that are not connected to the Great Lakes have positive occurrences of VHSV. Our conclusion that ballast water is a vector of spread for VHSV is contradictory to findings by Bain et al. (2010) who suggested there is no relationship between VHSV occurrences and centers of shipping. Their research only included shipping harbors as areas of shipping activity, and did not include actual ballast water discharge locations. Several locations that were identified as recreational boating or open shoreline by Bain et al. (2010) were identified by us as being close enough to ballast water discharge locations to become infected by discharged VHSV.

Even though we were not able to determine how much of a role ballast water plays in spreading VHSV, it was still identified as a vector that should be managed so as not to undermine other efforts that have been undertaken, such as restrictions on bait and fish stocking (APHIS 2008). If ships had been required to treat their ballast water prior to entering the Welland Canal, VHSV could have potentially been isolated to Lake Ontario and the St. Lawrence River. However, most ballast water management systems that are currently being tested for oceangoing ships would be inefficient for use by ships in the Great Lakes, since much of the US fleet have larger ballast tanks and higher pumping rates (Cangelosi and Mays 2006; USEPA Science Advisory Board 2011). Ships within the Great Lakes also tend to take shorter trips between ports, which may not allow enough time for chemical or physical treatments to sufficiently reduce propagule pressure (Cangelosi and Mays 2006). Without available ballast water treatment systems, there are a number of voluntary best management practices that ships in the Great Lakes may apply, such as drawing in water during the day or avoiding drawing in water where sediments are churned up (Shipping Federation of Canada 2000). However, these practices may not be effective in preventing the further spread of VHSV if not applied in the most suitable locations.

Our model can be used to identify the locations where the most promising best management practices would effectively be applied. One approach proposed by the shipping industry involves moving water uptake offshore, analogous to the requirements for ocean BWE outside the 200 nautical mile limit (Shipping Federation of Canada 2000). It is possible that invasive species may not be able to survive if released in deep waters offshore, far from required habitats and food resources. On the other hand, releasing invasive species in the deeper, offshore parts of the Great Lakes will only be effective if water currents do not carry the invasive species to more favorable habitats prior to mortality. Locations and times of year when water currents will not aid in the survival of invasive species will need to be identified. For example, our results could be combined with water circulation models that have been created by Beletsky and Schwab (2008) in order to identify those locations and times where and when ballast water may be released to reduce the probability of invasive species surviving. Further, our model can be used to identify those ports where the pick-up or discharge of ballast water should be avoided, or should be followed by ballast water exchange offshore.

Natural resource managers may also use our model to identify hotspots for invasive species. We expect to further validate our model by backcasting the secondary spread of zebra mussels (*Dreissena polymorpha*), an invasive bivalve, and ruffe (*Gymnocephalus cernuus*), an invasive fish. Both species are believed to have been introduced to the Great Lakes via ballast water (Grigorovich et al. 2003; Hebert et al. 1989; Simon and Vondruska 1991; Stepien et al. 2005). Once we parameterize our model for these species, predictions for the future spread of ruffe and other invasive species can be made. For example, managers are concerned about the introduction of killer shrimp (*Dikerogammerus villosus*), which has not yet been detected in the Great Lakes, but has been identified as a species that is likely to invade if ballast water management proves ineffective (Grigorovich et al. 2003). Our model can identify those areas where invasive species may occur next or may already occur, but may not be detected using conventional methods. Management practices can then be directed to those locations.

In summary, commercial ship ballast water movement and discharge patterns are likely contributing to the secondary spread of VHSV in the Great Lakes. Discharge locations that receive increasing visits from ships carrying ballast water from sources infected with

VHSV are more likely to become infected with the virus itself. Additionally, Montreal is the more likely location of initial VHSV introduction, not Lake St. Clair. Because ballast water is a component of long-distance spread in the Great Lakes, it is important that this vector be regulated along with bait and fish stocking. Our best fit model may be a tool that can aid managers and policy-makers in identifying locations where ballast water may best be managed.

Acknowledgments We would like to thank Ling Shen of the Minnesota Department of Natural Resources and Drs. Mark Bain and Paul Bowser from Cornell University for providing VHSV occurrence information. We would also like to thank Dr. David Reid for providing technical knowledge on ballast water management. This project was funded by grants from United States Department of Agriculture, National Institute of Food and Agriculture, Cooperative State Research, Education, and Extension Service, USDA-NIFA (CSREES) #2010-38927-21048 and National Oceanic and Atmospheric Agency, Center for Sponsored Coastal Ocean Research, NOAA-CSCOR #NA10NOS4780218. This is publication #2013-21 from the Lake Erie Research Center.

References

APHIS (2008) Amended federal order: viral hemorrhagic septicemia (VHS). Animal and Plant Health Inspection Service (APHIS), Washington, DC

Bain MB, Cornwell ER, Hope KM, Eckerlin GE, Casey RN, Groocock GH, Getchell RG, Bowser PR, Winton JR, Batts WN, Cangelosi A, Casey JW (2010) Distribution of an invasive aquatic pathogen (viral hemorrhagic septicemia virus) in the Great Lakes and its relationship to shipping. PLoS ONE 5:e10156

Beletsky D, Schwab D (2008) Climatological circulation in Lake Michigan. Geophys Res Lett 35:L21604. doi:10.1029/2008GL035773

Cangelosi A, Mays N (2006) Great ships for Great Lakes? Commercial vessels free of invasive species in the Great Lakes-St. Lawrence Seaway System. Northeast Midwest Institute, Washington, DC

Chico V, Gomez N, Estepa A, Perez L (2006) Rapid detection and quantitation of viral hemorrhagic septicemia virus in experimentally challenged rainbow trout by real-time RT-PCR. J Virol Methods 132:154–159. doi:10.1016/j.jviromet.2005.10.005

Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 70:213–220. doi:10.1037/h0026256

Covalciuc KA, Webb KH, Carlson CA (1999) Comparison of four clinical specimen types for detection of influenza A and B viruses by optical immunoassay (FLU OIA test) and cell culture methods. J Clin Microbiol 37:3971–3974

Drake LA, Doblin MA, Dobbs FC (2007) Potential microbial bioinvasions via ships' ballast water, sediment, and biofilm. Mar Pollut Bull 55:333–341. doi:10.1016/j.marpolbul.2006.11.007

Eames I, Landeryou M, Greig A, Snellings J (2008) Continuous flushing of contaminants from ballast water tanks. Mar Pollut Bull 56:250–260. doi:10.1016/j.marpolbul.2007.10.032

Eckerlin GE, Farrell JM, Casey RN, Hope KM, Groocock GH, Bowser PR, Casey J (2011) Temporal variation in prevalence of viral hemorrhagic septicemia virus type IVb among Upper St. Lawrence River smallmouth bass. Trans Am Fish Soc 140:529–536. doi:10.108/00028487.2011.581975

Elsayed E, Faisal M, Thomas M, Whelan G, Batts W, Winton J (2006) Isolation of viral haemorrhagic septicaemia virus from muskelluge, *Esox masquinongy* (Mitchill), in Lake St Clair, Michigan, USA reveals a new sublineage of the North American genotype. J Fish Dis 29:611–619. doi:10.1111/j.1365-2761.2006.00755.x

Faisal M, Winters AD (2011) Detection of viral hemorrhagic septicemia virus (VHSV) from *Diporeia* spp. (Pontoporeiidae, Amphipoda) in the Laurentian Great Lakes, USA. Parasit Vectors 4:1–4. doi:10.1186/1756-3305-4-2

Faisal M, Shavalier M, Kim RK, Millard EV, Gunn MR, Winters AD, Schulz CA, Eissa A, Thomas MV, Wolgamood M, Whelan GE, Winton J (2012) Spread of the emerging viral hemorrhagic septicemia virus strain, Gentoype IVb, in Michigan, USA. Viruses 4:734–760. doi:10.3390/v4050734

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49. doi:10.1017/S0376892997000088

Gilchrist JM (2009) Weighted 2 X 2 kappa coefficients: recommended indices of diagnostic accuracy for evidence-based practice. J Clin Epidemiol 62:1045–1053. doi:10.1016/j.jclinepi.2008.11.012

GLANSIS: Great Lakes aquatic nonindigenous species information system (Undated) NOAA. http://www.glerl.noaa.gov/res/Programs/glansis/glansis.html. Accessed 11 Sept 2012

Griffiths RW, Schloesser DW, Leach JH, Kovalak WP (1991) Distribution and dispersal of the zebra mussel (*Dreissena polymorpha*) in the Great Lakes region. Can J Fish Aquat Sci 48:1381–1388. doi:10.1139/f91-165

Grigorovich IA, Colautti RI, Mills EL, Holeck K, Ballert AG, MacIsaac HJ (2003) Ballast-mediated animal introductions in the Laurentian Great Lakes: retrospective and prospective analyses. Can J Fish Aquat Sci 60:740–756. doi:10.1139/F03-053

Hawley LM, Garver KA (2008) Stability of viral hemorrhagic septicemia virus (VHSV) in freshwater and seawater at various temperatures. Dis Aquat Org 82:171–178. doi:10.3354/dao01998

Hebert PDN, Muncaster BW, Mackie GL (1989) Ecological and genetic studies on *Dreissena polymorpha* (Pallas): a new mollusc in the Great Lakes. Can J Fish Aquat Sci 46:1587–1591. doi:10.1139/f89-202

Hope KM, Casey RN, Groocock GH, Getchell RG, Bowser PR, Casey JW (2010) Comparison of quantitative RT-PCR with cell culture to detect viral hemorrhagic septicemia

virus (VHSV) IVb infections in the Great Lakes. J Aquat Anim Health 22:50–61. doi:10.1577/H09-028.1

Kane-Sutton M, Kinter B, Dennis PM, Koonce JF (2010) Viral hemorrhagic septicemia virus infection in Yellow Perch, *Perca flavescens*, in Lake Erie. J Great Lakes Res 36:37–43. doi:10.1016/j.jglr.2009.11.004

Kim R, Faisal M (2011) Emergence and resurgence of the viral hemorrhagic septicemia virus (*Novirhabdovirus, Rhabdoviridae, Mononegavirales*). J Adv Res 2(1):9–23. doi:10.1016/j.jare.2010.05.007

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174

Leichsenring J, Lawrence J (2011) Effect of mid-oceanic ballast water exchange on virus-like particle abundance during two trans-Pacific voyages. Mar Pollut Bull 62:1103–1108. doi:10.1016/j.marpolbul.2011.01.034

Lovell SJ, Drake LA (2009) Tiny stowaways: analyzing the economic benefits of a US Environmental Protection Agency permit regulating ballast water discharges. Environ Manag 43(3):546–555. doi:10.1007/s00267-008-9215-2

Lumsden JS, Morrison B, Yason C, Russell S, Young K, Yazdanpanah A, Huber P, Al-Hussinee L, Stone D, Way K (2007) Mortality event in freshwater drum *Aplodinotus grunniens* from Lake Ontario, Canada, associated with viral haemorrhagic septicemia virus, Type IV. Dis Aquat Organ 76:99–111. doi:10.3354/dao076099

Manel S, Williams HC, Ormerod SJ (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. J Appl Ecol 38:921–931. doi:10.1046/j.1365-2664.2001.00647.x

Meyers TR, Winton JR (1995) Viral hemorrhagic septicemia virus in North America. Annu Rev Fish Dis 5:3–24. doi:0959-8030/95

Miller TA, Rapp J, Wastlhuber U, Hoffmann RW, Enzmann PJ (1998) Rapid and sensitive reverse transcriptase-polymerase chain reaction based detection and differential diagnosis of fish pathogenic rhabdoviruses in organ samples and cultured cells. Dis Aquat Organ 34:13–20. doi:10.3354/dao034013

Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of invasive species to marine biodiversity. Front Ecol Environ 6(9):485–492. doi:10.1890/070064

NAS-Nonindigenous aquatic species database (2009) USGS. http://nas.er.usgs.gov. Accessed 30 July 2011

National Ballast Information Clearinghouse (2009) Smithsonian Environmental Research Center (SERC) and USCG. http://invasions.si.edu/nbic/search.html. Accessed Aug 2012

National Research Council (2008) Great Lakes shipping, trade, and aquatic invasive species. National Academy of Sciences, Washington, DC

Pierce LR, Stepien CA (2012) Evolution and biogeography of an emerging quasispecies: diversity patterns of the fish viral hemorrhagic septicemia virus (VHSV). Mol Phylogenet Evol 63:327–341. doi:10.1016/j.ympev.2011.12.024

Ruiz GM, Rawlings TK, Dobbs FC, Drake LA, Mullady T, Huq A, Colwell RR (2000) Global spread of microorganisms by ships. Nature 408:49–50. doi:10.1038/35040695

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. Can J Fish Aquat Sci 67:256–268. doi:10.1139/F09-180

Shipping Federation of Canada (2000) Code of best practices for ballast water management. Shipping Federation of Canada, Montreal

Simon TP, Vondruska JT (1991) Larval identification of the ruffe, *Gymnocephalus cernuus* (Linnaeus) (Percidae: Percini), in the St. Louis River Estuary, Lake Superior drainage basin, Minnesota. Can J Zool 69:436–442. doi:10.1139/z91-068

Stepien CA, Brown JE, Neilson ME, Tumeo MA (2005) Genetic diversity of invasive species in the Great Lakes versus their Eurasian source populations: insights for risk analysis. Risk Anal 25(4):1043–1060. doi:10.1111/j.1539-6924.2005.00655.x

Thompson TM, Batts WN, Faisal M, Bowser P, Casey JW, Phillips K, Garver KA, Winton J, Kurath G (2011) Emergence of *Viral hemorrhagic septicemia virus* in the North American Great Lakes region is associated with low viral genetic diversity. Dis Aquat Organ 96:29–43. doi:10.3354/dao02362

USEPA Science Advisory Board (2011) Efficacy of ballast water treatment systems: a report by the EPA Science Advisory Board. USEPA, Washington, DC

Wald A, Huang M-L, Carrell D, Selke S, Corey L (2003) Polymerase chain reaction for detection of herpes simplex virus (HSV) DNA on mucosal surfaces: comparison with HSV isolation in cell culture. J Infect Dis 188(9):1345–1351. doi:10.1086/379043

Warrens MJ (2011) Cohen's linearly weighted kappa is a weighted average of 2×2 kappas. Psychometrika 76(3):471–486. doi:10.1007/s11336-011-9210-z

World Organisation for Animal Health (2011) Chapter 2.3.9: viral haemorrhagic septicaemia. Manual of diagnostic tests for aquatic animals 2011: summary. World Organisation for Animal Health, Paris