



Potential distribution of the viral haemorrhagic septicaemia virus in the Great Lakes region

L E Escobar^{1,2}, G Kurath³, J Escobar-Dodero⁴, M E Craft¹ and N B D Phelps^{1,2,5}

1 Veterinary Population Medicine, College of Veterinary Medicine, University of Minnesota, St. Paul, MN, USA

2 Minnesota Aquatic Invasive Species Research Center, University of Minnesota, St. Paul, MN, USA

3 U.S. Geological Survey, Western Fisheries Research Center, Seattle, WA, USA

4 Facultad de Ecología y Recursos Naturales, Universidad Andres Bello, Santiago, Chile

5 Veterinary Diagnostic Laboratory, College of Veterinary Medicine, University of Minnesota, St. Paul, MN, USA

Abstract

Viral haemorrhagic septicaemia virus (VHSV) genotype IVb has been responsible for large-scale fish mortality events in the Great Lakes of North America. Anticipating the areas of potential VHSV occurrence is key to designing epidemiological surveillance and disease prevention strategies in the Great Lakes basin. We explored the environmental features that could shape the distribution of VHSV, based on remote sensing and climate data via ecological niche modelling. Variables included temperature measured during the day and night, precipitation, vegetation, bathymetry, solar radiation and topographic wetness. VHSV occurrences were obtained from available reports of virus confirmation in laboratory facilities. We fit a Maxent model using VHSV-IVb reports and environmental variables under different parameterizations to identify the best model to determine potential VHSV occurrence based on environmental suitability. VHSV reports were generated from both passive and active surveillance. VHSV occurrences were most abundant near shore sites. We were, however, able to capture the environmental signature of VHSV based on the environmental variables employed in our model, allowing us to identify patterns of VHSV potential occurrence. Our findings suggest that VHSV is not at an ecological equilibrium and

more areas could be affected, including areas not in close geographic proximity to past VHSV reports.

Keywords: Great Lakes, Maxent, niche modeling, remote sensing, viral haemorrhagic septicaemia.

Introduction

The viral haemorrhagic septicaemia virus (VHSV; genus *Novirhabdovirus*; family *Rhabdoviridae*) is the causative agent of a globally significant infectious disease that has devastated wild and farmed fish species in freshwater and marine ecosystems since the mid-1900s (Horlyck *et al.* 1984; Meyers, Short & Lipson 1999; Dale *et al.* 2009; Faisal *et al.* 2012). In North American freshwaters, VHSV genotype IVb emerged dramatically in the Great Lakes basin since 2003, affecting more than 30 fish species to date (Cornwell *et al.* 2011; Fratini *et al.* 2011; Thompson *et al.* 2011; Faisal *et al.* 2012). The Great Lakes are the largest freshwater body in the world and are also known for their rich biodiversity and economic importance for the northern region of North America (Mills *et al.* 1993). However, the Great Lakes region has shown its ecological fragility to biological invasions, and specifically to VHSV-IVb since its discovery in Lake St. Clair in 2003, St. Lawrence River in 2004 and Lake Ontario in 2005 (Elsayed *et al.* 2006; Groocock *et al.* 2007; Lumsden *et al.* 2007). In view of the wide host range of the virus, pathogenicity, mortality rates, risk to farm-raised

Correspondence L E Escobar, Veterinary Population Medicine, College of Veterinary Medicine, University of Minnesota, 1365 Gortner Avenue, St. Paul, MN 55108, USA
(e-mail: lescobar@umn.edu)

fish, and high dispersal potential, VHSV is considered one of the most serious pathogens of fish worldwide (Skall *et al.* 2004; Dale *et al.* 2009; Faisal *et al.* 2012).

Evidence suggests that VHSV is not distributed at random in the landscape; in fact, the agent has an environmental signature that can be detectable. For instance, *in vitro* experiments show that VHSV virions can only survive in specific ranges of suitable temperature, pH, salinity and water hardness (Hawley & Garver 2008). In this context, understanding the environmental range required by the agent to persist may be the key to predicting potential areas where ideal environmental conditions exist (Peterson 2006). In ecology, estimating the set of environmental conditions in which a pathogen population could persist (subject to other fine-scale conditions, such as host presence) is termed ecological niche modelling (Drake 2015). Ecological niche modelling is based on the delimitation of the environmental space used by organisms, including pathogenic agents (Peterson 2014). According to Drake (2015), the niche modelling process attempts to discriminate between the environmental conditions occupied and not occupied by the organism of interest in the area of interest. This process provides a set of areas with conditions tolerable by the organism (Drake 2015). Thus, observed pathogen occurrence can be used to empirically estimate the ecological niche of the agent under natural (i.e. detection in the field) or experimental conditions (i.e. physiological tolerances in laboratory), allowing for posterior identification of areas environmentally suitable for the agent (Estrada-Peña *et al.* 2014). Different methods exist to estimate ecological niches (Qiao, Soberón & Peterson 2015); however, presence-background methods are preferred in view of their capacity to delimitate the species' environmental range in multiple environmental dimensions, their mathematical and computational feasibility and their biological realism (Drake 2015). In view of these features, the ecological niche modelling approach can be used successfully to model areas for the occurrence of pathogens at coarse scale, including those in aquatic environments. For example, a recent study explored the potential distribution of the waterborne bacterium *Vibrio cholerae* in marine ecosystems (Escobar *et al.* 2015b). In addition, ecological niche modelling approaches have been used also to successfully forecast the distribution

of aquatic invasive organisms including the plants *Hydrilla verticillata* (Peterson, Papes & Kluza 2003) and *Codium fragile* (Tyberghein *et al.* 2012), the fish common carp *Cyprinus carpio* and tilapia *Oreochromis niloticus* (Zambrano *et al.* 2006), the golden mussel *Limnoperna fortunei* (Oliveira, Hamilton & Jacobi 2010) and the freshwater diatom *Didymosphenia geminata* (Kumar *et al.* 2009). The distribution of freshwater parasites has been estimated at very fine scale (i.e. parasites position along the gastrointestinal tract) with consistent measurable patterns (Kennedy 2009). In general terms, however, the use of ecological niche modelling at macroscale to anticipate aquatic pathogens distribution has been poorly explored (Peterson 2014), but potential exists in these methods considering the abilities of niche modelling to successfully anticipate outbreaks of viral diseases of public health concern (Peterson *et al.* 2006).

Among the main concerns of the distributional ecology of VHSV is its capacity to mutate, spread to novel geographic areas and infect novel host species (Einer-Jensen *et al.* 2004; Thompson *et al.* 2011; Faisal *et al.* 2012). Modelling the suitable habitat for VHSV is a novel and valuable approach for anticipating future epidemics. As a first exploration, we developed an ecological niche model from VHSV detection reports, aimed to identify potential areas for VHSV-IVb occurrence within the Great Lakes region of North America based solely on environmental constraints. In view of the large number of novel fish host species identified for VHSV in the Great Lakes, and the lack of coevolutionary history of this recently introduced virus with native fish in the region (Peterson 2014), fish species were not included in this exploratory modelling process, but could be incorporated in the future as a finer scale variable. Here, based on environmental factors, identification of non-sampled areas hosting ideal environmental conditions for VHSV outbreaks would be the first step in the design of aggressive plans for disease prevention and policy, as well as to guide active epidemiological surveillance for management agencies.

Methods

Our model workflow included three phases: (i) collection and management of environmental variables and virus occurrence data, (ii) ecological

niche modelling and (iii) projection of the model to a geographic scenario (Fig. 1).

Environmental variables and occurrences

The choice of the region used for model calibration is critical in ecological niche modelling (Barve *et al.* 2011); therefore, we focused on an area including the Great Lakes ecosystem. The Great Lakes region in the United States extends from Minnesota to New York, and given the extension, geomorphology and physiography, the area is characterized by a diversity of landscapes including bedrock cliffs, clay bluffs, marsh wetlands, sand beaches and stream mouths (Herdendorf 1990). Specifically, our study area covered $\sim 46 \times 10^4 \text{ km}^2$ and was situated between -97.4° and -75.5° longitude and 39.6° and 50.6° latitude which includes all the VHSV geolocations of the Great Lakes available to us (Fig. 2). As this area included both land and aquatic habitats, we used remote sensing imagery and precipitation derived from climate stations as environmental variables to calibrate the model. Remote sensing data were generated from the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Shuttle Radar Topography Mission (SRTM) sensors, capturing patterns of land surface

temperature (LST), soil humidity in the form of a topographic wetness index and bathymetry capturing patterns of the geomorphology and physiography of the Great Lakes basin (Herdendorf 1990), and primary productivity as an index of vegetation 'greenness' or photosynthetic activity found in land and water (LP DAAC 2009; GEBCO 2014; Hengl *et al.* 2015). Precipitation variables were generated from interpolation of climate stations on land (Hijmans *et al.* 2005). These 18 remote sensing and four precipitation variables provided information on landscape features and seasonality available in the study area (Table 1). Variables were standardized to the same area extent (Fig. 2), file format (ASCII) and spatial resolution (1 km) to fit with modelling software requirements. Characteristics of the individual Great Lakes (i.e. Lake Superior, Lake Michigan, Lake Huron, Lake Erie and Lake Ontario) for these variables were extracted for their characterization. Characteristics of the geomorphology of the Great Lakes was also explored from a Global Relief Model based on SRTM 30+ and ETOPO DEM at $\sim 1\text{-km}$ resolution (Hengl *et al.* 2015), average lake depth and values of water volume contained in each lake (NOAA 2000).

VHSV occurrences in the form of geographic coordinates were derived from both active and

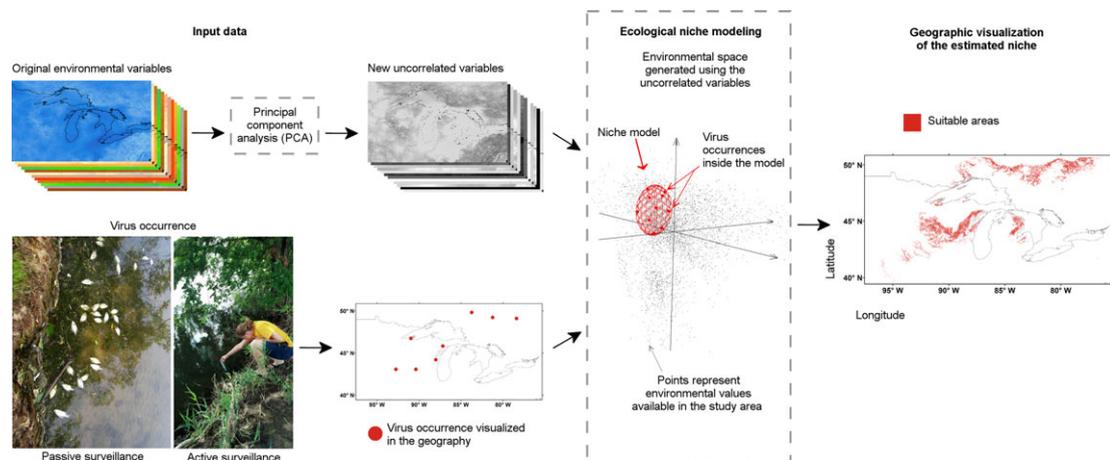


Figure 1 Workflow for the ecological niche modelling of the viral haemorrhagic septicaemia virus. The process is divided into three steps: input data (left), ecological niche modelling (centre) and geographic visualization of the estimated niche (right). The input process includes the generation of environmental variables of importance to the biology of the agent. These variables are then compressed to reduce their number and collinearity to reduce model overfitting. Input data also consider the occurrence of VHSV from field conditions. Occurrences can be derived from samples collected from targeted research projects (active surveillance) or as response of an evident outbreak (passive surveillance). The ecological niche model step considers delimitation of the species distribution on a multidimensional environmental space and can be developed using the NicheA software (Qiao *et al.* 2016). The ecological niche is then transferred from the environmental space to the geographic space to identify those areas in which the suitable environmental conditions exist. Some data management processes require multivariate modelling (dashed line).

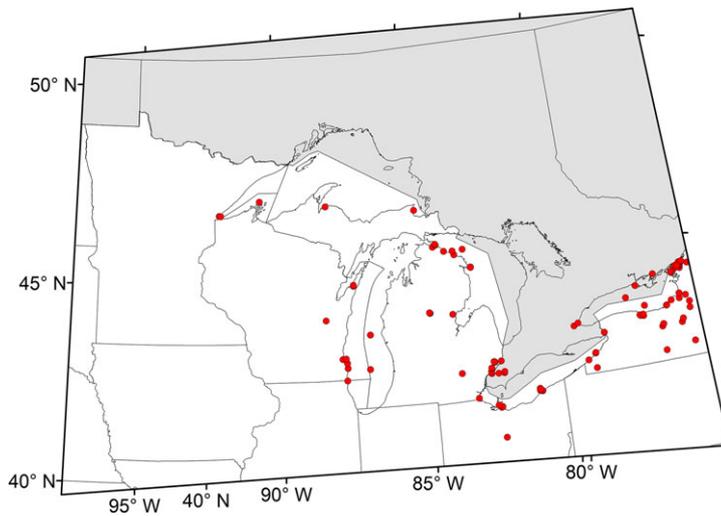


Figure 2 Viral haemorrhagic septicaemia virus (VHSV) occurrences and study area considered in this study. Location of VHSV (red points) in the Great Lakes region of the United States (white) and Canada (grey).

passive surveillance (Fig. 1). Active surveillance is related to research and monitoring efforts focused on identifying the agent in biological or environmental samples from guided surveys and routine fish health inspection, while passive surveillance is related to identification of the virus in response to an outbreak—in other words, the opportunistic analysis of dead or moribund fish from field samples submitted to laboratories for diagnosis (Fig. 1). Our VHSV occurrences were collected from the Molecular Epidemiology of Aquatic Pathogens-VHSV (MEAP-VHSV) database, containing the most extensive compilation of VHSV geolocations in the Great Lakes region from 2003 to 2013 (USGS 2013). Additional occurrences were collected from scientific literature (Kane-Sutton *et al.* 2010; Cornwell *et al.* 2011, 2014; Eckerlin *et al.* 2011; Faisal & Winters 2011; Frattini *et al.* 2011; Faisal *et al.* 2012; Millard & Faisal 2012) and official reports (MDNR 2010; USGS 2013). Inclusion criteria of occurrences were i) VHSV-IVb virus, confirmed by cell culture and/or quantitative PCR and ii) availability of geographic coordinates of sample location. We mitigated spatial pseudoreplication (Hurlbert 1984) and model overfitting (Peterson *et al.* 2011) by removing duplicate occurrences at the same site (i.e. only one occurrence was retained from VHSV reports with identical geographic coordinates).

Ecological niche modelling

To simplify our estimation of suitable environmental conditions where VHSV may occur in the

long term, we based our analysis principally on abiotic variables (e.g. temperature, precipitation, bathymetry; i.e. abiotic favours of the species niche, *sensu* Soberón & Peterson 2005), to identify areas with plausible positive population growth (Holt 2009), in this case, given the presence of fish hosts. Unsuitable conditions were defined as those that do not allow positive population growth of the species. We reduced the number and correlation of environmental variables via a principal component analysis (PCA; Fig. 1) to reduce overfitting of the ecological niche model (Peterson *et al.* 2011). The PCA was developed based on all the environmental variables (Table 1), using the Spatial Analysis tool of the ArcGIS software version 10.2 (ESRI 2015). The correlation matrix revealed that 27% of the 231 pair comparisons between the 22 original variables were correlated above $r (\pm) 0.5$ (Material S1). We included the first components summarizing >95% of the variation from the original 22 variables for model calibration. The first three principal components summarized 96.54% of the overall information contained in the original 22 variables and were highly influenced by bimonthly temperature, primary productivity and bathymetry (Table 1). These components were used to generate the multivariate environmental space used as background for model calibration (Fig. 3). The multivariate environmental space was able to capture geomorphological and physical features in the study area. The resulting uncorrelated components were employed to construct a multivariate environmental space, \mathbf{P} , where the ecological niche model \mathbf{N} was estimated, assuming that $\mathbf{N} \subseteq \mathbf{P}$.

Table 1 Environmental variables used for the ecological niche modelling of viral haemorrhagic septicaemia virus (VHSV)

Abbreviation	Variable	Original resolution	Period	Unit	Sources
NDVI	Mean of annual Normalized Difference Vegetation Index	16-day, 250 m	2005	Index (0–250)	Carroll, DiMiceli, Sohlberg & Townshend (2005)
Bathymetry	Bathymetry	30 arc-second	2009–2014	Metre	GEBCO (2014)
Radiation	Mean potential incoming solar radiation	8-day, 1 km	2012	$\times 365/8$ kWh/m ²	Hengl et al. (2015)
LST-md	Maximum value of the daytime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-ls	Minimum value of the daytime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-hd	Mean value the daytime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-sdd	Standard deviation of the daytime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-hn	Maximum value of the nighttime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-ln	Minimum value of the nighttime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-mn	Mean value the nighttime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-sdn	Standard deviation of the nighttime LST	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
Wetness	Topographic Wetness	8-day, 1 km	2009	Index ($\times/10 + 10$)	Hengl et al. (2015)
LST-12-1	Mean value of the daytime LST for Dec/Jan	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-2-3	Mean value of the daytime LST for Feb/Mar	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-4-5	Mean value of the daytime LST for Apr/May	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-6-7	Mean value of the daytime LST for Jun/Jul	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-8-9	Mean value of the daytime LST for Aug/Sep	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
LST-10-11	Mean value of the daytime LST for Oct/Nov	8-day, 1 km	2011–2012	Degree Celsius	Hengl et al. (2015)
PP-11-1	Long-term precipitation Nov/Dec/Jan	Monthly, 30 arc-second	1950–2000	Millimetres	Hijmans et al. (2005)
PP-2-4	Long-term precipitation Feb/Mar/Apr	Monthly, 30 arc-second	1950–2000	Millimetres	Hijmans et al. (2005)
PP-5-7	Long-term precipitation May/June/July	Monthly, 30 arc-second	1950–2000	Millimetres	Hijmans et al. (2005)
PP-8-10	Long-term precipitation Aug/Sep/Oct	Monthly, 30 arc-second	1950–2000	Millimetres	Hijmans et al. (2005)

The ecological niche model (i.e. correlating VHSV occurrences with environmental conditions) was generated using Maxent version 3.3.3.k (Phillips, Anderson & Schapire 2006). Maxent is a machine learning program that uses the principle of maximum entropy (Phillips, Dudik & Schapire 2004). Maxent first analyses the data in the geographic space and then in the environmental space. In geographic terms, Maxent searches for the probability of occurrences across the grid of the study area. Once cell probabilities are assigned, Maxent generates an analysis in environmental terms via comparing the probability densities of cells with occurrences with the environmental combination in each cell of the grid (for a detailed explanation see Elith *et al.* 2011; Merow, Smith & Silander 2013). In Maxent, the fit of the ecological niche model with the occurrences is strongly influenced by the regularization coefficient, penalizing the level of overfit during model calibration. Considering the vast number of configurations of ecological niches (Qiao *et al.* 2015), the regularization coefficient should be defined in each study case (Warren & Seifert 2011); thus, we explored the fit of each model under nine different coefficient values above and below the default value 1 (i.e. 0.3, 0.5, 1, 1.5, 2, 2.5, 3, 3.5 and 4). We selected the regularization coefficient based on the Akaike

information criterion score (AIC) corrected by sample sizes (AICc) (Burnham, Anderson & Huyvaert 2011). AICc is the preferred metric to accurately identify the ecological niche model with appropriate complexity and that best represents the relative environmental suitability of a species (here VHSV) in a specific calibration area (Warren & Seifert 2011). We ranked the AICc values (i.e. Δi) to identify the model hypothesis with the regularization coefficient generating the lowest AICc, interpreted here as the model that generates cell values (i.e. predictions) closer to cell values with VHSV occurrence (i.e. observations). AIC analyses were developed using the ENMTools software version 1.4.3 (Warren, Glor & Turelli 2010). Specific settings in Maxent included regularization coefficient with best model fit (see above), 1000 bootstrap replicates based on a random selection of occurrence points, 75% of occurrences for model calibration and 25% for model evaluation of each replicate, and raw values of the median replicate as model output. Maxent generates a continuous model output. Maps of prediction in the form of continuous values are useful for qualitative exploratory analysis (Merow *et al.* 2013). Because Maxent predictions are based on multivariate probability densities of predictors across the study area (Elith *et al.* 2011), continuous predictions represent the environmental

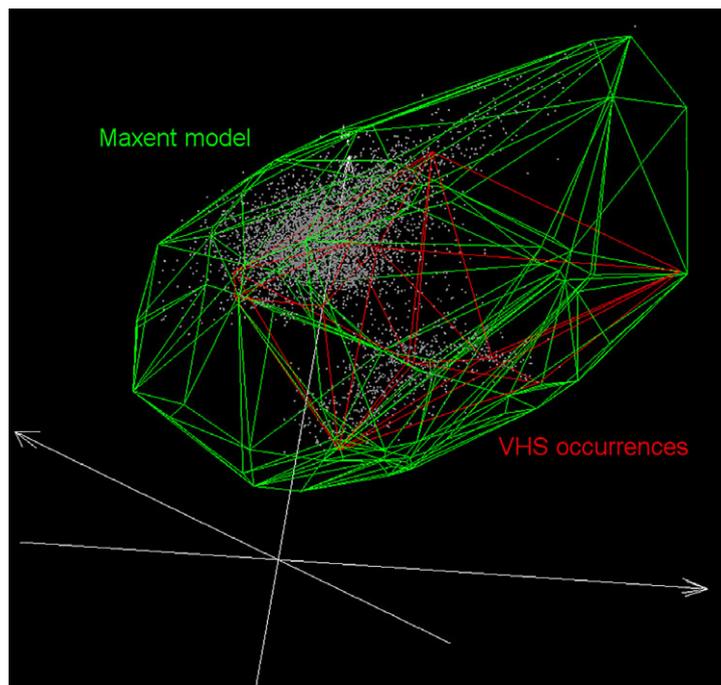


Figure 3 Binary ecological niche model of viral haemorrhagic septicaemia virus (VHSV) in the environmental space. VHSV occurrences representing 94 different environmental conditions (i.e. points in the environmental space) allowed a robust characterization of the environments suitable for virus occurrence (red polyhedron) and were used for model calibration in Maxent. The ecological niche model (green polyhedron generated based on the available occurrences) shows predictions in a broad range of the environmental values resulting from extrapolative estimations of the model. The environmental background (grey points) was generated using the first three principal components from the original environmental variables.

similarity to cells with occurrences. Thus, we interpreted continuous predictions as an index of environmental suitability for VHSV survival.

To reduce the influence of the sampling effort on the model output and to capture the information in the form of an environmental range where VHSV occurs (Drake 2015), we generated a binary model from the continuous surface, using as a threshold the minimum value of predictions from the occurrences used during model calibration. This allowed us to capture the tolerance limits (minimum and maximum) of environmental values from Maxent predictions (Peterson *et al.* 2011; Drake 2015). Once the final model was converted to binary, we compared occurrences with the model extent in a multidimensional environmental scenario. This analysis was conducted with occurrences and the binary model in the software NicheA version 3.0 (Qiao *et al.* 2016). NicheA compared the hypervolume of both the model and the occurrences with an expected precision of 0.1. This allowed us to establish the level of extrapolation provided by the model.

Model projection

We identified areas in P hosting the combination of environments for N and explored the environmental conditions of the model and occurrences based on the original values of the variables and compared areas with virus occurrence vs. areas predicted unsuitable. These analyses were conducted using NicheA and R (R Core Team 2015). We assume that environments outside N have $R_0 < 1$, meaning that the virus is unlikely to spread further (Peterson 2014). The ecological niche model was then transferred from the environmental space to the geographic space in order to compare areas suitable and unsuitable for VHSV occurrence. To generate a product useful for management and surveillance purposes, we went a step beyond classic printed maps in scientific manuscripts. Using ArcGIS, the continuous model was converted to binary based on the minimum training presence threshold (Peterson *et al.* 2011). The binary model was (i) constrained to water bodies found in the study area according to a land water mask derived from MODIS and SRTM Water Body Dataset (SWBD) at 250-m resolution (LP DAAC 2009) and (ii) converted to a KML file for broad access to the information by the community. Spatial analyses and data

management were carried out using the software ArcGIS version 10.2 (ESRI 2015).

Results

Environmental variables and virus occurrences

Lake Ontario was unique in some aspects when compared with the other four Great Lakes (i.e. Lake Superior, Lake Michigan, Lake Huron and Lake Erie). For example, according to the characteristics of geomorphology, the water volume of Lake Ontario according to NOAA (2000) was considerably below the expected values in a linear regression model (Material S2). Reports of solar radiation and topographic relief from NOAA (2000) were similar in all Great Lakes, but Lake Ontario differed, with considerably lower relief. Average values of bathymetry of Lake Ontario were slightly below the other Great Lakes, but these values fell in the range of depth contained in the other lakes. Lakes Superior, Michigan, Huron and Ontario had similar values of primary productivity, while Lake Erie showed less primary productivity. In terms of direct precipitation received by the lake, Lake Superior showed less precipitation for winter and spring, while Lake Huron was the driest during summer. The westernmost Great Lake (i.e. Lake Superior) had the lowest temperatures compared to the eastern Great Lakes (i.e. Lake Ontario and Erie) in terms of both LST at day and night. Lake Erie showed the lowest variability of LST during day of all the Great Lakes.

In terms of virus detections, we captured a total of 189 occurrences (Material S3), resulting in 109 unique occurrences once duplicates were removed from individual collection sites (Fig. 2). These 109 occurrences were used for model calibration and then compared with 109 random points in areas predicted unsuitable. Occurrences in the geographic space represented 94 unique environmental conditions; thus, even when some occurrences were geolocated in different geographic areas, they contained identical conditions (Fig. 3). Furthermore, our sample size covered a considerable portion of the environmental background included in the model.

Ecological niche modelling

The model provided ~360% more information than the occurrences in terms of the hypervolume

measured in the environmental space (Fig. 3), resulting in a model with extrapolative prediction following the fit with the data. The best model fit resulted from a regularization coefficient = 1 (Table 2). Because low temperatures are linked to VHSV disease outbreaks (Goodwin & Merry 2011), we focused on patterns of LST as indication of water temperature. The ecological niche model reflected patterns of VHSV occurrence in a broad range of temperatures and filled the gaps of reports in the range of minimum LST measured at night (LST-n) and extrapolated to temperatures above and below observations (Fig. 4). According to the model, VHSV may find suitable temperatures above those found in the study area; thus, the model was restricted in predictions due to limitations on the environmental values available in the study area (Fig. 4). For instance, for the minimum LST-n variable the model estimated a range of temperature from -45 to 2 °C, but the temperature for this variable in the study area ranged between -46 and 2 °C. Thus, the model did not predict suitability in the lowest range of temperatures (i.e. -46 °C), but was truncated for predictions above 2 °C as there were no cells with values above 2 °C in the minimum LST-n variable (A panel in Fig. 4). On the other hand, the entire study area had a range of maximum LST-n between 10 and 33 °C, with the model predicting VHSV occurrence between 11 and 33 °C; again, the model was truncated to anticipate suitability above values of temperatures outside the calibration area. According to the occurrences, an ideal temperature for VHSV ranges from 1 to 10 °C based on mean nighttime LST (B in Fig. 4) and from 7 to 18 °C based on mean daytime LST (data not shown). The range of LST measurement (i.e. maximum LST-d *minus* minimum LST-d) at

day sites with VHSV occurrence is usually smaller when compared to areas unsuitable for VHSV ($t = -14.78$; $df = 128$; $P < 0.001$; Fig. 5a), resulting in VHSV occurrences in areas with less seasonal temperature variation. We found that vegetation surrounding sites of VHSV occurrence may be an important determinant for viral presence; the 1-km cells with VHSV occurrences had less primary productivity in the form of NDVI values than areas predicted unsuitable for the agent ($t = -12.38$; $df = 128$; $P < 0.001$; Fig. 5b). Additionally, VHSV was often reported in water bodies with low depth (bathymetry, $t = -12.38$; $df = 128$; $P < 0.001$; Fig. 5c).

In general terms, VHSV was influenced by seasonal temperature variation and was restricted by temperature ranges (Fig. 6). VHSV suitability was found in temperatures close to zero degree during winter while unsuitable conditions were found in cold (<0 °C) and hot areas (~ 25 °C). Low precipitation, resulting in low run-off, showed an effect on non-suitability for VHSV in winter; however, seasonality of VHSV suitability was less evident when compared to temperature.

Model projection

The continuous model showed specific areas of high suitability (Fig. 7), discriminating from areas with environmental conditions non-similar to areas with VHSV occurrence. For instance, broad areas of environmental suitability were found across all Lake Erie, southern portions of Lake Michigan, and Lake Winnebago in Wisconsin, most of Georgian Bay and Saginaw Bay in Lake Huron in the United States, and Lake Nipissing in Canada. Green Bay of Lake Michigan and Lake St. Clair also have high environmental suitability

Table 2 Regularization coefficients based on 109 single occurrences of the viral haemorrhagic septicaemia virus

Regularization	Log likelihood	Parameters	AIC	AICc	BIC	Δi
1	-1440.94	26	2933.88	2951	3003.86	0
2	-1468.67	16	2969.35	2975.26	3012.41	24.26
1.5	-1456.18	25	2962.37	2978.03	3029.65	27.03
2.5	-1481.37	11	2984.73	2987.46	3014.34	36.46
3	-1485.71	9	2989.42	2991.24	3013.64	40.24
3.5	-1489.9	10	2999.80	3002.04	3026.71	51.04
4	-1492.27	9	3002.54	3004.35	3026.76	53.35
0.5	-1410.84	50	2921.67	3009.6	3056.24	58.6
0.3	-1378.61	84	2925.22	3520.22	3151.3	569.22

AIC, Akaike information criterion; AICc, AIC bias correction based on sample size; BIC, Bayesian information criterion; Δi , information distance to the best model.

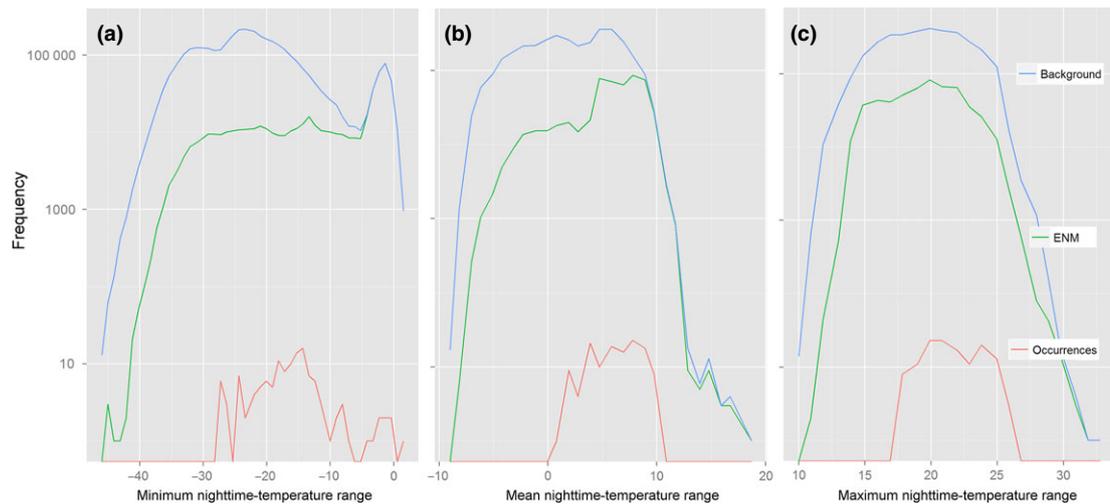


Figure 4 Distribution of viral haemorrhagic septicaemia virus occurrences in terms of nighttime land surface temperature. Minimum (a), mean (b) and maximum (c) value of land surface temperature of the nighttime series data according to the available occurrences (red line), ecological niche model of Maxent (green line) and background values of the study area (blue line). Frequency of pixels with each temperature values is in log₁₀ scale. Notice that Maxent extrapolates temperature values from field observations.

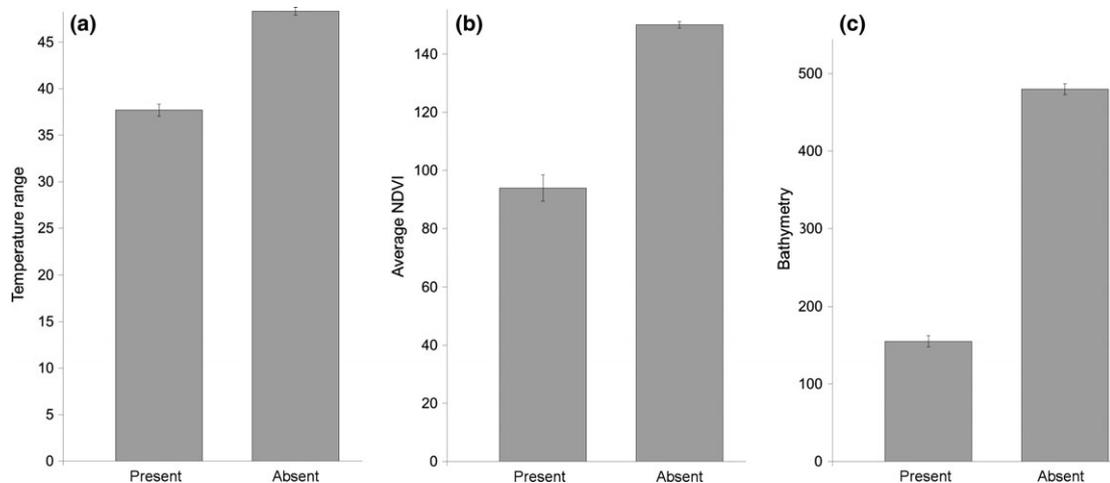


Figure 5 Presence and absence of viral haemorrhagic septicaemia virus (VHSV) according to abiotic variables. Presence of VHSV occurrences (left bar) contrasted against areas predicted as absent by the model (right bar). Temperature in the form of range value the 8-day MODIS nighttime LST time series data (a), primary productivity in the form of NDVI (b) and underwater depth of lakes in the form of bathymetric values (c) were compared. Vertical thin lines are bars of standard error of the mean.

for VHSV establishment (see Material S4). The southern coast of Lake Superior in the USA and the northern coast of Lake Superior and Lake Winnipeg in Canada have ideal environments, suggesting that suitable conditions may exist for VHSV establishment in such water bodies. As expected, rivers were predicted to have suitability to support VHSV occurrence (Material S5). An interesting pattern was found around Lake Ontario, with high

suitability values around the entire shoreline and in small water bodies surrounding the lake (e.g. Finger Lakes in New York, USA), but low suitability in areas throughout the central area of the lake. This pattern could be linked to the greater depth of this lake, compared with, for example, Lake Erie, which is shallower and shows high suitability throughout the entire area of the lake (Fig. 7). Alternatively, the pattern may be influenced by the higher

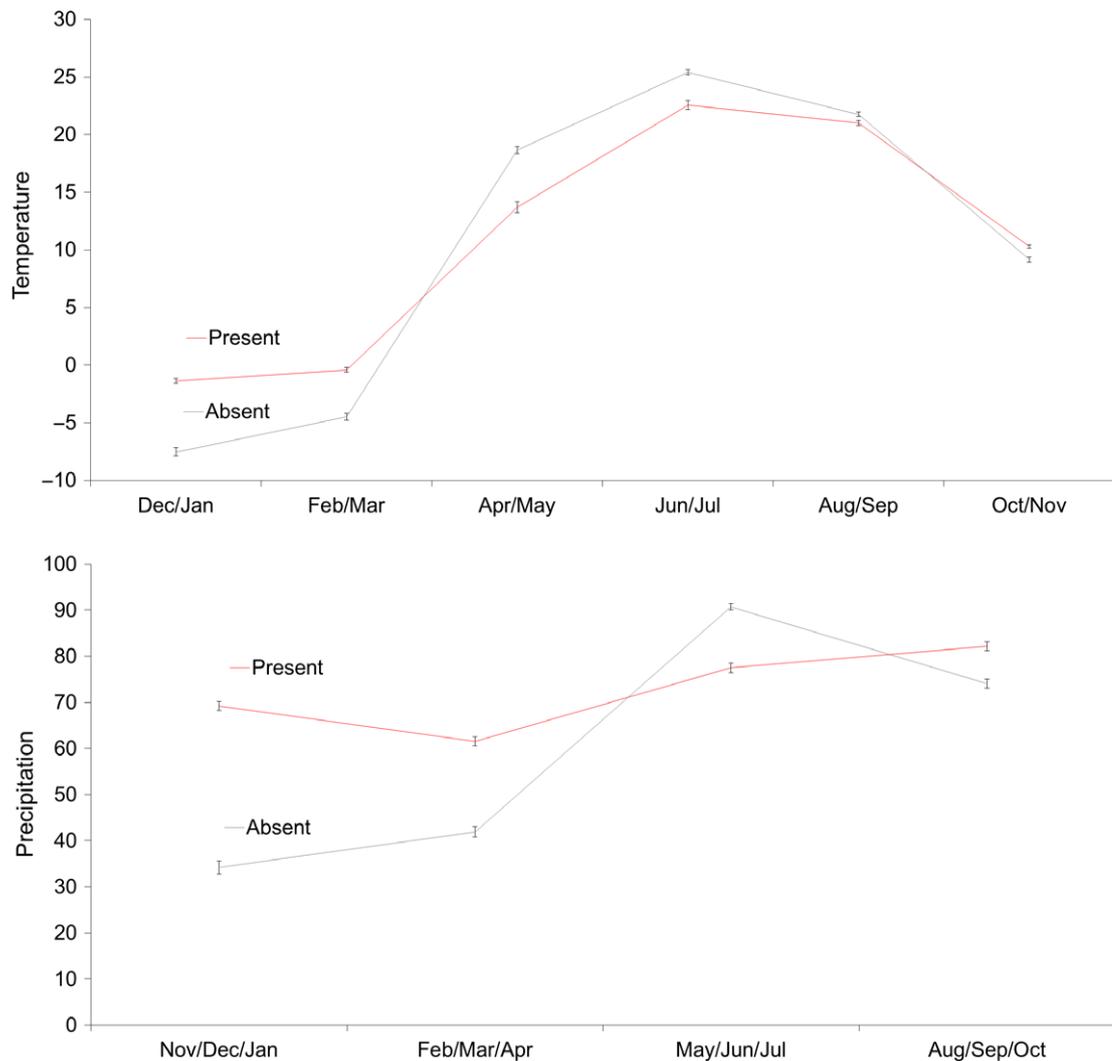


Figure 6 Seasonal occurrence of viral haemorrhagic septicaemia virus. Top: mean seasonal land surface temperature of the daytime series data. Bottom: mean seasonal precipitation from long-term series data. Red lines denote locations where VHSV is present. Grey line denotes areas predicted absent of VHSV according to the ecological niche model. Vertical thin lines are bars of standard error of the mean.

sampling density in Lake Ontario, which may allow for differentiation of shorelines from internal lake areas, as described in more detail below. Another interesting feature was areas of high suitability adjacent to Eastern Lake Ontario, in areas that are not within the lake itself but that have high density of other water bodies. To the Northeast of Lake Ontario, this included the Saint Lawrence Seaway, and towards both the Northeast and Southeast, these areas appeared to have larger tracts of small lakes.

When the continuous model was converted to binary based on the minimum training presence

threshold to capture the entire environmental range (Material S4), we found that more water bodies have environmental conditions appropriate for VHSV occurrence as a consequence of the ecological plasticity of this virus. The binary map allowed us to visualize water bodies with VHSV suitability that were difficult to perceive in a continuous model; thus, we identified several lakes and rivers with low suitability, but with values above the selected threshold (Material S4). The binary map of VHSV environmental suitability is also available as supplementary material for use with user-friendly software (e.g. Google Earth;

climate stations around lakes (e.g. Gallardo, Zu Ermgassen & Aldridge 2013; Zengeya *et al.* 2013). However, such models often fail to accurately capture the environmental conditions in the central areas of the lakes, lack the necessary spatial resolution and also suffer from high spatial autocorrelation (Peterson 2014). Here, we applied recent satellite derived environmental variables from values captured on top layers of water and land, to capture patterns of abiotic factors in and surrounding the water bodies in the Great Lakes basin.

Satellite imagery data are in its infancy to describe variables (e.g. temperature) across the water column in freshwaters through large study areas such as used here. However, land surface temperature may provide robust patterns of surface temperature in water bodies with less spatial autocorrelation than that from interpolations of climate stations (Escobar & Peterson 2013). Besides temperature, freshwater ecosystems require different model designs than those applied to niche modelling in terrestrial habitats. For example, topography is usually a confounding factor in ecological niche modelling in terrestrial ecosystems due to its strong association with temperature; thus, topography is usually avoided during model calibration when modelling terrestrial organisms (Peterson *et al.* 2011). Here, we included bathymetry data derived from geomorphology information in view of the importance of the water column depth on the distributional ecology of aquatic organisms (Ruby, Greenberg & Hastings 1980) and its ability to explain temperature stratification (Hondzo & Stefan 1993). Water depth results from the lakes' topography; thus, our model considered bathymetry and other variables associated with the land morphology (e.g. Wetness Index). This variable may have limited accuracy due to the spatial resolution, but variables characterizing hydrodynamics of freshwaters associated with water depth are limited (but see Nekouee *et al.* 2015). Additionally, considering the important effects of run-off on nutrient supplies in lakes and its consequent effects on biotic communities (Schindler & Nighswander 1970), we included precipitation in our analyses. Precipitation plus temperature, allowed us to explore temporal patterns of suitability that may provide information on the seasonality of ideal conditions for VHSV.

Our results suggest that VHSV often occurs in shallow waters; however, this pattern may be

influenced by sampling bias, considering that dead fish could be more often reported and/or sampled near lakeshores due to waves generated by boats or wind, even if fish are infected in central areas of lakes and due to easiness in active surveillance in the shore (see pictures in Fig. 1). Indeed, we found a high representation of VHSV occurrence in coastal and limited reports in centre areas of lakes (Fig. 2). Thus, the water depth pattern should be considered with caution in view of the nature of our data. Yet, environments predicted suitable for VHSV appear to be consistent and predictable in space and time (Figs 3 & 6). In fact, our ecological niche model exploration revealed that coastal areas in important water bodies of the Great Lakes basin would suffer VHSV outbreaks if susceptible fish species occur at critical densities when the virus is present. While small and medium water bodies are currently VHSV-free in northern (e.g. southern Canada) and western (e.g. Minnesota, USA) areas of the Great Lakes region, invasion is feasible in these areas in ecological terms (Material S4). We could expect VHSV to be reported to a broader extent in the short term, at least for Lake Superior, Lake Michigan, Lake Huron and Georgian Bay considering their connection with areas currently positive for VHSV. Suitability was detected in southeaster of our study area, in the Appalachian Basin in south-western Pennsylvania, in response to geology, land morphology and moisture availability of the area, predicted given the extrapolation of the model (Fig. 4), but when a mask of water bodies was used in our binary map (Fig. S3), only the areas linked to the Susquehanna River remained suitable.

Temperature was a key factor limiting the potential areas for VHSV distribution in the Great Lakes region, as revealed by the temperature values captured from remote sensing variables. The strong effect of temperature on VHSV-IVb has been reported previously in laboratory conditions (e.g. Vo *et al.* 2015). Experimentally, suitable temperatures reported for VHSV-IVb ranged from 10 to 18 °C (Goodwin & Merry 2011), 10 to 20 °C (Winton, Kurath & Batts 2007) and 4 to 20 °C (Vo *et al.* 2015). These values are similar to our findings of 1 to 18 °C based on available occurrences solely and temperature data collected via daytime satellite imagery, suggesting that remote sensing data could be a valuable source of information to characterize field conditions potentially associated with the presence of

infectious agents, especially when our abilities for experimentation in laboratory facilities are limited.

The continuous (Fig. 7) and binary maps of environmental suitability (Material S4) identified in detail areas where VHSV occurrence would be expected considering abiotic conditions. Findings of high suitability were consistent with the areas with high sampling effort, especially in the eastern portion of the Lake Ontario where high virus detection and disease impact have been reported previously (Groocock *et al.* 2007; Lumsden *et al.* 2007; Bain *et al.* 2010; Eckerlin *et al.* 2011; Cornwell *et al.* 2012a,b, 2014). The ability of the ecological niche model to capture the environmental signature of VHSV at a landscape scale supports the idea of potential model transferability from the Great Lakes to other water bodies with potential for VHSV invasion, for example Lake Baikal in southern Siberia, with a latitude similar to that used in this study (Fig. 2). Our niche modelling exploration revealed areas with high suitability for VHSV occurrence allowing identification of priority areas for monitoring; areas currently without VHSV reports would be of high interest to develop future active surveillance, local awareness, and may be candidate areas to develop permanent monitoring based on sentinel fish. Our results can inform fish management plans for an effective use of funding and reduce non-informed opportunistic sampling strategies.

Statistical models summarizing the environmental distribution of pathogens of fish species have an enormous potential to increase our understanding on both the spatial epidemiology of fish diseases and spatial ecology in lake ecosystems, especially when applied to disease prevention and freshwater biogeography. Thus, ecological niche modelling is a key tool to respond to questions related to the evolutionary history of host–pathogen coevolution under a landscape-ecology approach (Escobar *et al.* 2014). The model showed low suitability in the central area of Lake Ontario (Fig. 7), but high suitability around the shorelines. The latter is consistent with previous reports of VHSV occurrence in coastal areas of the same lake with prevalence reaching ~28% (Cornwell *et al.* 2012a,b). The non-continuous suitability predicted in Lake Ontario may be result of the impact of the bathymetric complexity of this lake (Material S2). Another plausible explanation is overfitting of the model to the abundant VHSV occurrences reported in coastal and shallow

water in this lake and surrounding water bodies, resulting in ability to resolve differing suitability between the coast and the central areas of the lake. This idea is supported by the extensive surveillance efforts in the coasts of Lake Ontario (Cornwell *et al.* 2012a).

We also attempted to identify plausible effects of primary productivity on virus prediction, considering that the distribution of other rhabdoviruses has been successfully forecasted using ecological niche modelling calibrated using primary productivity data (e.g. Escobar *et al.* 2015a). However, while several sources of primary productivity exist for terrestrial and marine areas (Res *et al.* 2011; Tyberghein *et al.* 2012), we were unable to find data on primary productivity in freshwaters of the study areas at fine scale (i.e. phytoplankton). Thus, we explored the vegetation index NDVI representing photosynthetic activity – as this variable may provide patterns of vegetation surrounding sites with VHSV reports. NDVI, however, may fail to capture low spectral reflectance from primary productivity in water.

We assessed the regularization coefficient for the best model fit on the association of occurrences and the environmental variables employed and found that regularization parameter 1 generated the best model performance in terms of model fit and model complexity. This finding provides empirical information needed to verify Maxent performance based on different regularization coefficients (Dudik, Phillips & Schapire 2007). Here we also mitigated model overfit by elimination of duplicate occurrences and resampling occurrences by cell, but in the future a more rigorous alternative reducing the number of occurrences based on spatial autocorrelation might provide even less redundant information during modelling calibration (Peterson 2014).

During early stages of the VHSV-IVb invasion, dramatic mortality events were reported across the Great Lakes region between 2005 and 2008 (Elsayed *et al.* 2006; Groocock *et al.* 2007; Lumsden *et al.* 2007). More recently, the epidemiological status of VHSV appears to be in a declining or stable stage (Cornwell *et al.* 2012b; Throckmorton *et al.* 2015). However, our suitability maps suggest that novel areas may be affected by the virus if translocation is allowed by anthropogenic or natural movement of the virus between lakes. Future research at fine scale should aim to determine if VHSV can be detected across the entire geographic

distribution of host species in infected lakes. Additionally, water flow may have important effects on pathogen distribution (Foreman *et al.* 2015), but complexities associated with characterization of water flow limit the availability of such variables at coarse geographic scale (but see Nekouee *et al.* 2015). Thus, our model is static, assuming that environmental patterns captured by the remote sensing time series data are consistent and not strongly influenced by flow. Inclusion of water flow information in pathogen dispersal is an issue that deserves future exploration. Additionally, water flow may have a critical role in VHSV dispersal to VHSV-free areas (VHSV Expert Panel and Working Group 2010). For example, the Mississippi river is connected to many water bodies in the Great Lakes region, making this river a potential VHSV superspreader in view of its high connectivity outside the region. Also Lake Michigan could be 'hyperconnected' to other water bodies through the Chicago canals. Assessing how water body connection may impact Great Lakes VHSV dispersal would provide information on the plausible routes of infection to guide control strategies. Boat movement, baitfish commercialization and fish stocking, and connection of rivers, among other factors, may be relevant for anticipating how VHSV could increase its dispersal capacity in the Great Lakes region (Phelps *et al.* 2014). For example, Brockmann & Helbing (2013) compared air traffic vs. geographic distance to explore which of these variables explain human pathogens spread at coarse scales. Authors found that passenger flow is more important than the spatial distance between infected and disease-free areas. In this context, human-driven virus transport, more than geographic distance between water bodies, may be a key component in VHSV epidemics, but such phenomena have not been explored in detail.

Finally, our results showed extended environmental suitability to areas where VHSV is currently absent, including water bodies in Minnesota, Lake Winnipeg and northern areas of Lake Superior. This suitability suggests areas where VHSV could be able to survive. The truncation of model predictions due to limitation of environmental values available in the study area represents an important challenge on the macroecology of VHSV. We found that (i) more areas could be predicted suitable if the study area increases (Fig. 4a) and (ii) suitable areas were found beyond the current distribution of the

virus. Under this scenario, VHSV appears to be in ecological non-equilibrium (Guisan & Thuiller 2005). Invasive species tend to expand their distribution to all the environmentally suitable areas if abiotic interaction and movement barriers do not limit their invasion dispersal. An invasive species reaches ecological equilibrium when the species occupies all the invadable environments. For VHSV, the equilibrium state is uncertain considering (i) the early stage of its invasion (i.e. ~2003; Elsayed *et al.* 2006), (ii) the discovery of vast number of hosts species affected (Williams & Jackson 2007); (iii) the broad environmental range; and (iv) the truncation of the model due to environments not available in the study area (e.g. Fig. 4). In fact, one of our assumptions was violated: we assumed that the environmental space was larger with respect to the VHSV niche. Our exploration suggests that VHSV may find suitable environmental conditions that are not present in our study area. Fine-scale patterns of environmental suitability associated with population growth (infection rates) still need theoretical and empirical consideration (Drake 2015). Thus, continued proactive management efforts and early identification of VHSV outbreaks are warranted. Other fine-scale factors with potential importance in VHSV outbreaks deserve exploration including toxic chemicals, presence of dreissenids and composition of the fish community at a given site, among others. This will allow understanding the factors triggering disease emergence at local scale.

Conclusion

Viral haemorrhagic septicaemia virus will likely continue to be a threat for fish in the Great Lakes region. While we did not consider the presence of susceptible fish species in our model, the expansion of the virus to water bodies currently free of infection would occur in the future if susceptible hosts exist with appropriate densities, environmental conditions are suitable and the virus breaches geographic barriers to arrive to the VHSV-free area. Seasonality and connectivity of water bodies should be a topic of high importance to anticipate and prevent VHSV spread to virus-free areas. VHSV invasion process is currently static, but more areas could be reported infected considering the suitable conditions found beyond areas already reported VHSV positive (Escobar *et al.* 2014). Thus, VHSV has the potential to interact with novel geographic ranges, environmental

conditions and host species to yield more ecological surprises.

Acknowledgements

LEE was supported by the Minnesota Environment and Natural Resources Trust Fund, the Minnesota Aquatic Invasive Species Research Center and the Clean Water Land and Legacy. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the US Government.

References

- Bain M.B., Cornwell E.R., Hope K.M., Eckerlin G.E., Casey R.N., Groocock G.H., Getchell R.G., Bowser P.R., Winton J.R., Batts W.N., Cangelosi A. & Casey J.W. (2010) Distribution of an invasive aquatic pathogen (Viral Hemorrhagic Septicemia virus) in the Great Lakes and its relationship to shipping. *PLoS ONE* **5**, 13–15.
- Barve N., Barve V., Jiménez-Valverde A., Lira-Noriega A., Maher S.P., Peterson A.T., Soberón J. & Villalobos F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* **222**, 1810–1819.
- Blehert D.S., Hicks A.C., Behr M., Meteyer C.U., Berlowski-Zier B.M., Buckles E.L., Coleman J.T.H., Darling S.R., Gargas A., Niver R., Okoniewski J.C., Rudd R.J. & Stone W.B. (2009) Bat white-nose syndrome: an emerging fungal pathogen? *Science* **323**, 227–227.
- Brockmann D. & Helbing D. (2013) The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342.
- Burnham K.P., Anderson D.R. & Huyvaert K.P. (2011) AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* **65**, 23–35.
- Carroll M.L., DiMiceli C.M., Sohlberg R.A. & Townshend J.R.G. (2005) 250m MODIS Normalized Difference Vegetation Index, 250ndvi28920033435, Collection 4. University of Maryland. Available at <http://glcf.umd.edu/data/ndvi/> (Accessed 2 June 2015).
- Cornwell E.R., Eckerlin G.E., Getchell R.G., Groocock G.H., Thompson T.M., Batts W.N., Casey R.N., Kurath G., Winton J.R., Bowser P.R., Bain M.B. & Casey J.W. (2011) Detection of Viral Hemorrhagic Septicemia virus by quantitative reverse transcription polymerase chain reaction from two fish species at two sites in Lake Superior. *Journal of Aquatic Animal Health* **23**, 207–217.
- Cornwell E.R., Eckerlin G.E., Thompson T.M., Batts W.N., Getchell R.G., Groocock G.H., Kurath G., Winton J.R., Casey R.N., Casey J.W., Bain M.B. & Bowser P.R. (2012a) Predictive factors and viral genetic diversity for Viral Hemorrhagic Septicemia virus infection in Lake Ontario and the St. Lawrence River. *Journal of Great Lakes Research* **38**, 278–288.
- Cornwell E.R., Getchell R.G., Groocock G.H., Walsh M.G. & Bowser P.R. (2012b) Low prevalence of VHSV detected in round goby collected in offshore regions of Lake Ontario. *Journal of Great Lakes Research* **38**, 575–579.
- Cornwell E.R., Primus A., Wong P.T., Anderson G.B., Thompson T.M., Kurath G., Groocock G.H., Bain M.B., Bowser P.R. & Getchell R.G. (2014) Round gobies are an important part of VHSV genotype IVb ecology in the St. Lawrence River and eastern Lake Ontario. *Journal of Great Lakes Research* **40**, 1002–1009.
- Dale O.B., Ørpetveit I., Lyngstad T.M., Kahns S., Skall H.F., Olesen N.J. & Dannevig B.H. (2009) Outbreak of Viral Hemorrhagic Septicemia (VHS) in seawater-farmed rainbow trout in Norway caused by VHS virus Genotype III. *Diseases of Aquatic Organisms* **85**, 93–103.
- Do Amaral K.B., Alvares D.J., Heinzmann L., Borges-Martins M., Siciliano S. & Moreno I.B. (2015) Ecological niche modeling of *Stenella dolphins* (Cetartiodactyla: Delphinidae) in the southwestern Atlantic Ocean. *Journal of Experimental Marine Biology and Ecology* **472**, 166–179.
- Drake J.M. (2015) Range bagging: a new method for ecological niche modelling from presence-only data. *Journal of the Royal Society Interface* **12**, 20150086.
- Dudík M., Phillips S.J. & Schapire R.E. (2007) Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. *Journal of Machine Learning Research* **8**, 1217–1260.
- Eckerlin G.E., Farrell J.M., Casey R.N., Hope K.M., Groocock G.H., Bowser P.R. & Casey J. (2011) Temporal variation in prevalence of Viral Hemorrhagic Septicemia virus Type IVb among Upper St. Lawrence River smallmouth bass. *Transactions of the American Fisheries Society* **140**, 529–536.
- Einer-Jensen K., Ahrens P., Forsberg R. & Lorenzen N. (2004) Evolution of the fish rhabdovirus viral haemorrhagic septicaemia virus. *Journal of General Virology* **85**, 1167–1179.
- Elith J., Phillips S.J., Hastie T., Dudík M., Chee Y.E. & Yates C.J. (2011) A statistical explanation of Maxent for ecologists. *Diversity and Distributions* **17**, 43–57.
- Elsayed E., Faisal M., Thomas M., Whelan G., Batts W. & Winton J.R. (2006) Isolation of Viral Hemorrhagic Septicemia virus from muskellunge, *Esox masquinongy* (Mitchill), in Lake St Clair, Michigan, USA reveals a new sublineage of the North American genotype. *Journal of Fish Diseases* **29**, 611–619.
- Escobar L.E. & Peterson A.T. (2013) Spatial epidemiology of bat-borne rabies in Colombia. *Pan American Journal of Public Health* **34**, 135–136.
- Escobar L.E., Lira-Noriega A., Medina-Vogel G. & Peterson A.T. (2014) Potential for spread of White-nose fungus (*Pseudogymnoascus destructans*) in the Americas: using Maxent and NicheA to assure strict model transference. *Geospatial Health* **11**, 221–229.
- Escobar L.E., Peterson A.T., Papeş M., Favi M., Yung V., Restif O., Qiao H. & Medina-Vogel G. (2015a) Ecological approaches in veterinary epidemiology: mapping the risk of bat-borne rabies using vegetation indices and night-time light satellite imagery. *Veterinary Research* **46**, 92.

- Escobar L.E., Ryan S.J., Stewart-Ibarra A.M., Finkelstein J.L., King C.A., Qiao H. & Polhemus M.E. (2015b) A global map of suitability for coastal *Vibrio cholerae* under current and future climate conditions. *Acta Tropica* **149**, 202–211.
- ESRI (2015) *ArcGIS Desktop: Release 10.2*. Environmental Systems Research Institute, Redlands, CA.
- Estrada-Peña A., Ostfeld R.S., Peterson A.T., Poulin R. & de la Fuente J. (2014) Effects of environmental change on zoonotic disease risk: an ecological primer. *Trends in Parasitology* **30**, 205–214.
- Faisal M. & Winters A.D. (2011) Detection of Viral Hemorrhagic Septicemia virus (VHSV) from *Diporeia* spp. (Pontoporeiidae, Amphipoda) in the Laurentian Great Lakes, USA. *Parasites & Vectors* **4**, 2.
- Faisal M., Shavaliar M., Kim R.K., Millard E.V., Gunn M.R., Winters A.D., Schulz C.A., Eissa A., Thomas M.V., Wolgamood M., Whelan G.E. & Winton J.R. (2012) Spread of the emerging Viral Hemorrhagic Septicemia virus strain, genotype IVb, in Michigan, USA. *Viruses* **4**, 734–760.
- Fisher M.C., Garner T.W.J. & Walker S.F. (2009) Global emergence of *Batrachochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host. *Annual Review of Microbiology* **63**, 291–310.
- Foreman M.G.G., Guo M., Garver K.A., Stucchi D., Chandler P., Wan D., Morrison J. & Tuele D. (2015) Modelling Infectious Hematopoietic Necrosis virus dispersion from marine salmon farms in the Discovery Islands, British Columbia, Canada. *PLoS ONE* **10**, e0130951.
- Fratini S.A., Grocock G.H., Getchell R.G., Wooster G.A., Casey R.N., Casey J.W. & Bowser P.R. (2011) A 2006 survey of Viral Hemorrhagic Septicemia (VHSV) virus type IVb in New York state waters. *Journal of Great Lakes Research* **37**, 194–198.
- Gallardo B., Zu Ermgassen P.S.E. & Aldridge D.C. (2013) Invasion ratcheting in the zebra mussel (*Dreissena polymorpha*) and the ability of native and invaded ranges to predict its global distribution. *Journal of Biogeography* **40**, 2274–2284.
- Garver K.A., Traxler G.S., Hawley L.M., Richard J., Ross J.P. & Lovy J. (2013) Molecular epidemiology of viral haemorrhagic septicaemia virus (VHSV) in British Columbia, Canada, reveals transmission from wild to farmed fish. *Diseases of Aquatic Organisms* **104**, 93–104.
- GEBCO (2014) GEBCO_2014 Grid — A global 30 arc-second interval grid. *General Bathymetric Chart of the Oceans*. Available at <http://www.gebco.net/> (Accessed 01 January 2016).
- Goodwin A.E. & Merry G.E. (2011) Mortality and carrier status of bluegills exposed to Viral Hemorrhagic Septicemia virus genotype IVb at different temperatures. *Journal of Aquatic Animal Health* **23**, 85–91.
- Grocock G.H., Getchell R.G., Wooster G.A., Britt K.L., Batts W.N., Winton J.R., Casey R.N., Casey J.W. & Bowser P.R. (2007) Detection of viral hemorrhagic septicemia in round gobies in New York State (USA) waters of Lake Ontario and the St. Lawrence River. *Diseases of Aquatic Organisms* **76**, 187–192.
- Guisan A. & Thuiller W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**, 993–1009.
- Hawley L.M. & Garver K.A. (2008) Stability of Viral Hemorrhagic Septicemia virus (VHSV) in freshwater and seawater at various temperatures. *Diseases of Aquatic Organisms* **82**, 171–178.
- Hengl T., Kilibarda M., Carvalho-Ribeiro E.D. & Reuter H.I. (2015) Worldgrids — a public repository and a WPS for global environmental layers. *WorldGrids*. Available at <http://worldgrids.org/doku.php?id=about&rev=1427534899> (Accessed 14 September 2015).
- Herdendorf C.E. (1990) Great Lakes Estuaries. *Estuaries* **13**, 493–503.
- Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G. & Jarvis A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965–1978.
- Holt R.D. (2009) Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences of the United States of America* **106**(Suppl), 19659–19665.
- Hondzo M. & Stefan H.G. (1993) Lake water temperature simulation model. *Journal of Hydraulic Engineering Division of the American Society of Civil Engineers* **119**, 1251–1273.
- Horlyck V., Møllergaard S., Dalsgaard I. & Jørgensen P.E.V. (1984) Occurrence of VHS in Danish Maricultured Rainbow Trout. *Bulletin of the European Association of Fish Pathologists* **4**, 11–13.
- Horning N., Robinson J., Sterling E., Turner W. & Spector S. (2010) *Remote Sensing for Ecology and Conservation*. Oxford University Press, New York.
- Hurlbert S.H. (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* **54**, 187–211.
- Kane-Sutton M., Kinter B., Dennis P.M. & Koonce J.F. (2010) Viral Hemorrhagic Septicemia virus infection in Yellow perch, *Perca flavescens*, in Lake Erie. *Journal of Great Lakes Research* **36**, 37–43.
- Kennedy C.R. (2009) The ecology of parasites of freshwater fishes: the search for patterns. *Parasitology* **136**, 1653–1662.
- Kim R. & Faisal M. (2011) Emergence and resurgence of the Viral Hemorrhagic Septicemia virus (Novirhabdovirus, Rhabdoviridae, Mononegavirales). *Journal of Advanced Research* **2**, 9–23.
- Kumar S., Spaulding S.A., Stohlgren T.J., Hermann K.A., Schmidt T.S. & Bahls L.L. (2009) Potential habitat distribution for the freshwater diatom *Didymosphenia geminata* in the continental US. *Frontiers in Ecology and the Environment* **7**, 415–420.
- LP DAAC (2009) Land Water Mask Derived from MODIS and SRTM L3 Global 250 m SIN Grid. *MOD44W*. Available at https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod44w (Accessed 16 November 2015).

- Lumsden J.S., Morrison B., Yason C., Russell S., Young K., Yazdanpanah A., Huber P., Al-Hussiney L., Stone D.M. & Way K. (2007) Mortality event in freshwater drum *Aplodinotus grunniens* from Lake Ontario, Canada, associated with Viral Hemorrhagic Septicemia virus, type IV. *Diseases of Aquatic Organisms* **76**, 99–111.
- MDNR (2010) Special notice: VHS found in Lake Superior. *Minnesota Department of Natural Resources*. Available at http://www.dnr.state.mn.us/fish_diseases/vhs.html (Accessed 06 September 2015).
- Merow C., Smith M.J. & Silander J.A. (2013) A practical guide to Maxent for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* **36**, 1–12.
- Meyers T.R., Short S. & Lipson K. (1999) Isolation of the North American strain of viral hemorrhagic septicemia virus (VHSV) associated with epizootic mortality in two new host species of Alaskan marine fish. *Diseases of Aquatic Organisms* **38**, 81–86.
- Millard E.V. & Faisal M. (2012) Heterogeneity in levels of serum neutralizing antibodies against Viral Hemorrhagic Septicemia virus genotype IVb among fish species in Lake St. Clair, Michigan, USA. *Journal of Wildlife Diseases* **48**, 405–415.
- Mills E.L., Leach J.H., Carlton J.T. & Secor C.L. (1993) Exotic species in the Great Lakes: a history of biotic crises and anthropogenic introductions. *Journal of Great Lakes Research* **19**, 1–54.
- Munro E.S., McIntosh R.E., Weir S.J., Noguera P.A., Sandilands J.M., Matejusova I., Mayes A.S. & Smith R. (2015) A mortality event in wrasse species (Labridae) associated with the presence of Viral Hemorrhagic Septicemia virus. *Journal of Fish Diseases* **38**, 335–341.
- Nekouee N., Hamidi S.A., Roberts P.J.W. & Schwab D.J. (2015) Assessment of a 3D hydrostatic model (POM) in the near field of a Buoyant River plume in Lake Michigan. *Water, Air, & Soil Pollution* **226**, 210.
- NOAA (2000) National Oceanic and Atmospheric Administration. About our Great Lakes: Lake by lake profiles. *Great Lakes Environmental Research Laboratory*. Available at <http://www.glerl.noaa.gov/pr/ourlakes/lakes.html> (Accessed 07 November 2016).
- Oliveira M.D., Hamilton S.K. & Jacobi C.M. (2010) Forecasting the expansion of the invasive golden mussel *Limnoperna fortunei* in Brazilian and North American rivers based on its occurrence in the Paraguay River and Pantanal wetland of Brazil. *Aquatic Invasions* **5**, 59–73.
- Peterson A.T. (2006) Ecologic niche modeling and spatial patterns of disease transmission. *Emerging Infectious Diseases* **12**, 1822–1826.
- Peterson A.T. (2014) *Mapping Disease Transmission Risk: Enriching Models using Biology and Ecology*. Johns Hopkins University Press, Baltimore.
- Peterson A.T., Papes M. & Kluza D.A. (2003) Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science* **51**, 863–868.
- Peterson A.T., Lash R.R., Carroll D.S. & Johnson K.M. (2006) Geographic potential for outbreaks of Marburg hemorrhagic fever. *The American Journal of Tropical Medicine and Hygiene* **75**, 9–15.
- Peterson A.T., Soberón J., Pearson R.G., Anderson R.P., Martínez-Meyer E., Nakamura M. & Bastos M. (2011) *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton, NJ.
- Phelps N.B.D., Goodwin A.E., Marecaux E. & Goyal S.M. (2013) Comparison of treatments to inactivate Viral Hemorrhagic Septicemia virus (VHSV-IVb) in frozen baitfish. *Diseases of Aquatic Organisms* **102**, 211–216.
- Phelps N.B.D., Craft M.E., Travis D., Pelican K. & Goyal S.M. (2014) Risk-based management of Viral Hemorrhagic Septicemia virus in Minnesota. *North American Journal of Fisheries Management* **34**, 373–379.
- Phillips S.J., Dudik M. & Schapire R.E. (2004) A maximum entropy approach to species distribution modeling. Twenty-first international conference on Machine learning - ICML '04 21, 83.
- Phillips S.J., Anderson R.P. & Schapire R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–259.
- Qiao H., Soberón J. & Peterson A.T. (2015) No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and Evolution* **6**, 1126–1136.
- Qiao H., Peterson A.T., Campbell L.P., Soberón J., Ji L. & Escobar L.E. (2016) NicheA: creating virtual species and ecological niches in multivariate environmental scenarios. *Ecography*. doi:10.1111/ecog.01961 [Epub ahead of print].
- R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raja-Halli M., Vehmas T.K., Rimaila-Pärnänen E., Sainmaa S., Skall H.F., Olesen N.J. & Tapiovaara H. (2006) Viral Hemorrhagic Septicemia (VHS) outbreaks in Finnish rainbow trout farms. *Diseases of Aquatic Organisms* **72**, 201–211.
- Res C., Pettorelli N., Ryan S., Mueller T., Bunnefeld N., Jedrzejewska B., Lima M. & Kausrud K. (2011) The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. *Climate Research* **46**, 15–27.
- Ruby E.G., Greenberg E.P. & Hastings J.W. (1980) Planktonic marine luminous bacteria: species distribution in the water column. *Applied and Environmental Microbiology* **39**, 302–306.
- Schindler D.W. & Nighswander J.E. (1970) Nutrient supply and primary production in Clear Lake, eastern Ontario. *Journal of the Fisheries Research Board of Canada* **27**, 2009–2036.
- Skall H.F., Slierendrecht W.J., King J.A. & Olesen N.J. (2004) Experimental infection of rainbow trout *Oncorhynchus mykiss* with viral haemorrhagic septicaemia virus isolates from European marine and farmed fishes. *Diseases of Aquatic Organisms* **58**, 99–110.
- Soberón J. & Peterson A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* **2**, 1–10.

- Thompson T.M., Batts W.N., Faisal M., Bowser P., Casey J.W., Phillips K., Garver K.A., Winton J.R. & Kurath G. (2011) Emergence of Viral Hemorrhagic Septicemia virus in the North American Great Lakes region is associated with low viral genetic diversity. *Diseases of Aquatic Organisms* **96**, 29–43.
- Throckmorton E., Peters A., Brenden T. & Faisal M. (2015) Direct and indirect evidence suggests continuous presence of Viral Hemorrhagic Septicemia virus (Genotype IVb) in Budd Lake, Michigan: management implications. *North American Journal of Fisheries Management* **35**, 503–511.
- Tyberghein L., Verbruggen H., Pauly K., Troupin C., Mineur F. & De Clerck O. (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* **21**, 272–281.
- U.S. Geological Survey (2013) Molecular epidemiology of aquatic pathogens -viral hemorrhagic septicemia virus (MEAP-VHSV) database. Available at <http://gis.nacse.org/vhsv/> (Accessed 1 June 2015).
- VHSV Expert Panel and Working Group (2010) Viral Hemorrhagic Septicemia virus (VHSV-IVb) risk factors and association measures derived by expert panel. *Preventative Veterinary Medicine* **94**, 28–139.
- Vo N.T.K., Bender A.W., Lee L.E.J., Lumsden J.S., Lorenzen N., Dixon B. & Bols N.C. (2015) Development of a walleye cell line and use to study the effects of temperature on infection by Viral Hemorrhagic Septicemia virus group IVb. *Journal of Fish Diseases* **38**, 121–136.
- Warren D.L. & Seifert S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications* **21**, 335–342.
- Warren D.L., Glor R.E. & Turelli M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607–611.
- Williams J.W. & Jackson S.T. (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment* **5**, 475–482.
- Winton J.R., Kurath G. & Batts W. (2007) Detection of Viral Hemorrhagic Septicemia Virus. *USGS*. Available at <http://pubs.usgs.gov/fs/2007/3055/fs20073055.pdf> (Accessed 10 September 2015).
- Zambrano L., Martínez-Meyer E., Menezes N. & Peterson A.T. (2006) Invasive potential of common carp (*Cyprinus carpio*) and Nile tilapia (*Oreochromis niloticus*) in American freshwater systems. *Canadian Journal of Fisheries and Aquatic Sciences* **63**, 1903–1910.
- Zengeya T.A., Robertson M.P., Booth A.J. & Chimimba C.T. (2013) Ecological niche modeling of the invasive potential of Nile tilapia *Oreochromis niloticus* in African river systems: concerns and implications for the conservation of indigenous congeners. *Biological Invasions* **15**, 1507–1521.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Material S1 Correlation matrix of comparisons between the environmental variables used to model the geographic distribution of the Viral Hemorrhagic Septicemia virus in the Great Lakes region.

Material S2 Description of environmental conditions in the Great Lakes.

Material S3 Occurrences of Viral Hemorrhagic Septicemia virus according to location, host species of isolation, year of isolation, and data source.

Material S4 Binary ecological niche model of Viral Hemorrhagic Septicemia virus (VHSV) in the geographic space.

Material S5 Table of water bodies hosting suitable conditions for Viral Hemorrhagic Septicemia virus (VHS) occurrence in the Great Lakes region.

Received: 21 January 2016

Revision received: 19 March 2016

Accepted: 21 March 2016