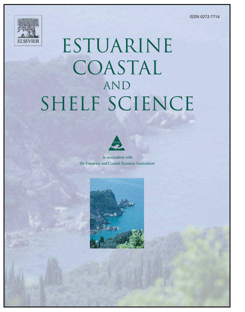

Impact of the future coastal water temperature scenarios on the risk of potential growth of pathogenic *Vibrio* marine bacteria

Habiba Ferchichi, André St-Hilaire, Taha B.M.J. Ouarda, Benoît Levesque

PII: S0272-7714(20)30825-8

DOI: https://doi.org/10.1016/j.ecss.2020.107094

Reference: YECSS 107094

To appear in: *Estuarine, Coastal and Shelf Science*

Received Date: 14 February 2020

Revised Date: 9 September 2020

Accepted Date: 2 November 2020

Authors' contributions

H.Ferchichi, A.St-Hilaire, T.B.M.J.Ouarda, and B.Lévesque participated in the design of the study. H.Ferchichi had completed the analysis under the supervision of A.St-Hilaire, T.B.M.J.Ouarda, and B.Lévesque.

H.Ferchichi had completed the bulk of the writing and prepared the Figures 1-3, with assistance from A.St-Hilaire, T.B.M.J.Ouarda, and B.Lévesque. All authors reviewed the manuscript.

Impact of the future coastal water temperature scenarios on the risk of potential growth of pathogenic *Vibrio* marine Bacteria

Habiba Ferchichi[1,*], André St-Hilaire[1, 2], Taha B.M.J Ouarda[3], Benoît Levesque[4, 5]

Vibrio (V), a genus of marine bacteria, are common inhabitants of warm coastal waters and estuaries. _Vibrio_ includes _V. parahaemolyticus_ and _V. vulnificus_ species that can cause human infections through the consumption of contaminated shellfish (as bivalve molluscs). The growth of pathogenic _Vibrio_ is related to ambient water temperature and seems to increase at 15 °C and over. The expansion of _Vibrio_ infection outbreak is increasing worldwide due to the increase of the sea surface temperature as a result of ocean warming. Canada's coast is not an exception to this worldwide _Vibrio_ spread. Faced with this issue, this study focuses on modelling the future potential _Vibrio_ growth risk along the coasts of the St. Lawrence Gulf and Estuary, where the shellfish industry is well developed. This is done using the adequate machine learning model with explanatory variables that include air temperature and wind speed for predicting future water temperatures. Based on the predicted future water temperature scenarios and a threshold of 15 °C to determine the conditions favorable to the growth of _Vibrio_ bacteria, we modelled the _Vibrio_ growth risk indicator, i.e. the number of days exceeding the minimum temperature for _Vibrio_ pathogenic growth (15 °C), in the horizon 2040-2100. Simulations show that the number of days, where the minimum temperature (15 °C) will be reached ,will increase spatially and even seasonally and all the shellfish beds would meet the temperature condition for _Vibrio_ growth regardless of the climate scenario (optimistic and pessimistic).

Key words: Coastal Water temperature, Vibrio bacteria, machine learning models, modelling, prediction, climate change.

Abbreviations: ANN, Artificial Neural Networks; Bagging, Bootstrap Aggregating; CANOPA, CANadian Océan PArallélisé; CART, Classification and Regression Tree; CDC, Centers for Disease Control and Prevention; GHG, Greenhouse Gas GSL; Gulf of St. Lawrence; IDW, Inverse Distance Weighted; MA, Moving Average; MK, Mann-Kendall; MLP, Multilayer Perceptron; MMK, Modified Mann-Kendall; MSE, Mean Square Error; Nash, Nash-Sutcliffe coefficient; Ntry, The number of bootstrap input variables at each split of a tree; OOB, Out-Of-Bag; PEI, Prince Edward Island; ppt, parts per thousand; rBias, Relative mean bias; RCM, Regional Climate Model; RCP, Representative Concentration Pathways; RF, Random Forests; RFE, Recursive feature Elimination; RMSE, Root Mean Square Error; SST, Sea Surface Temperature; *Vibrio*, *V*; *Vibrio parahaem*olyticus, *Vp*; *Vibrio vulnificus, Vv*.

[1] Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, 490 de la Couronne, Québec, QC, G1K 9A9, Canada.

[2] Canadian River Institute, University of New Brunswick, Fredericton, NB, Canada

[3] Canada Research Chair in Statistical Hydro-climatology, Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, 490 de la Couronne, Québec, QC, G1K 9A9, Canada.

[4] Direction de la santé environnementale et de la toxicologie, Institut national de santé publique du Québec, 945 avenue Wolfe, Québec, Québec, G1V 5B3, Canada

[5] Département de médecine sociale et préventive, Faculté de médecine, Université Laval, Québec.

* Corresponding author: Habiba Ferchichi (ferchichihabiba1993gmail.com)

1. Introduction

*Vibrio (V.) parahaem*olyticus and *Vibrio vulnificus* belong to the family of *Vibrionacea*, a group of aquatic microorganisms that includes other human pathogens such as *V. cholera.* They are natural inhabitants of warm coastal waters (>15 °C) and estuaries with low salinity (<25 ppt) (Baker-Austin et al., 2010; Heng et al., 2017; Kaspar and Tamplin, 1993; Motes et al., 1998; Vezzulli et al., 2013).

*V. parahaem*olyticus (*Vp*) is recognised as a leading cause of gastroenteritis associated with seafood consumption worldwide (Martinez-Urtaza et al., 2010) and it was the cause of significant outbreaks of infections in North America. For example, the largest outbreak of *Vp* in Canadian history, associated with the consumption of raw oysters, occurred in summer 2015 in British Colombia and resulted in the highest reported *Vp* cases (82 cases) (Taylor et al., 2018) since the 1997 outbreak (Fyfe et al., 1997).

V. vulnificus (*Vv*) infections are less frequent. However, *Vv* is a lethal opportunistic human pathogen responsible for the majority of deaths related to seafood consumption worldwide. For instance, in the USA, more than 95% of seafood-related deaths are caused by this bacterium (Oliver, 2013). Consumption of raw or undercooked bivalve shellfish (oysters, mussels, clams, etc.) contaminated with *Vv* can lead to major infections such as septicemia, with subsequent highest mortality (sometimes exceeding 50%) than any foodborne pathogen (Dechet et al., 2008; Feldhusen, 2000; Oliver, 2005). In the USA, the CDC estimates an average of 100 foodborne infections associated to *Vv* annually, resulting in 50 fatalities per year (Mead et al., 1999).

The growth of pathogenic *Vibrio* species causing human illness is directly related to the exceedance of a threshold of water temperature (about 15 °C) (Baker-Austin et al., 2013; Jacobs et al., 2015; Martinez-Urtaza et al., 2010; McLaughlin et al., 2005). Because most of bivalves are filter feeders, *Vibrio* bacteria may concentrate in their tissues. When the water temperature exceeds a certain threshold, shellfish are more likely to be contaminated with *Vibrio*. These contaminated shellfish transmit, in turn, the *Vibrio* bacteria to humans through consumption of raw or undercooked shellfish (Baker-Austin et al., 2017; Davis et al., 2017; McLaughlin et al., 2005; Motes and DePaola, 1996; Zimmerman et al., 2007).

Several reports and scientific researches show that the incidence of *Vibrio* infections has increased significantly worldwide (Centers for Disease and Prevention, 2013; Martinez-Urtaza et al., 2010; Newton et al., 2012). For instance, during the Canadian *Vp* outbreak, the number of reported cases was 2.5 times the number of expected cases and the outbreak unusually occurs earlier than expected (June-July) (Taylor et al., 2018). This unusual outbreak emergence was associated with abnormally high sea surface temperatures (SST>15°C) and the human *Vp* incidence decreased when the SST decreased below 15°C (Taylor et al., 2018). Numerous studies on this Canadian *Vp* outbreak show that the sea surface temperature is the most significant environmental predictor of the *Vp* proliferation in oysters and the *Vp* illness incidence is strongly associated with the increase of SST and the exceedance of the temperature threshold (Galanis et al., 2020; Konrad et al., 2017; Taylor et al., 2018).

In addition to the increase of spatial spread worldwide, sudden *Vibrio* outbreaks had emerged in new temperate and even cold regions including Peru (Martinez-Urtaza et al., 2008), Europe (Baker-Austin et al., 2010), Chile (Narjol et al., 2005) and Alaska (McLaughlin et al., 2005). This unusual outbreak emergence of *Vibrio* infections coincides with water temperatures anomalies (SST>15°C) (Baker-Austin et al., 2017; Martinez-Urtaza et al., 2010). Many microbiologists agree that climate change may explain this increase of *Vibrio* infections spread worldwide as well as the likelihood of its geographical expansion in new areas (Baker-Austin et al., 2012; Deeb et al., 2018; Martinez-Urtaza et al., 2010; McLaughlin, 2005; Vezzulli et al., 2013; Vezzulli et al., 2016). They have even considered *Vibrio* pathogens as microbial barometer of climate change (Baker-Austin et al., 2017).

The global average land-ocean temperature has risen by approximately 0.85 °C since the late nineteenth century (IPCC, 2013). This increase in SST, caused by atmospheric warming, heavily affects the coastal ecosystems (Baker-Austin et al., 2017; Burge et al., 2014; Halpern et al., 2008), resulting in significant warming of 70% of the world's coastline (Baker-Austin et al., 2017; Lima and Wethey, 2012).

In order to evaluate and manage the *Vibrio* infection risk, various models have been developed. Among these models, some are related to the prediction of *Vibrio* concentration, which is based either on only sea surface temperature (SST) (Chu et al., 2011) or both SST and salinity (Jacobs et al., 2014; United States Food and Drug

Administration (FDA), 2005). Another category includes models developed to explain the relation of *Vibrio* infections exposure in response to SST threshold exceedances (Semenza et al., 2017). Given the paucity of *Vibrio* concentration and *Vibrio* infection data, the aim of this study is to evaluate the *Vibrio* growth risk through its relation with SST threshold exceedances.

The harvesting of molluscs is an important part of the Canadian economy. It is well developed in the provinces of Quebec, and Prince Edward Island (PEI). PEI is Canada's top shellfish producer with about 49434 tons in 2018 between wild shellfish and aquaculture, while Quebec produces about 1840 tons (Statistics Canada, 2019). The shellfish beds are distributed over coastal zones of the Estuary and Gulf of St. Lawrence (GSL), located in the eastern part of Canada, including Rimouski, Gaspe, Baie des Chaleurs, the Quebec North Shore, Magdalen Islands and PEI (Fig 1).

As the rest of worldwide marine ecosystems affected by ocean warming, the SST of GSL has increased by 1 to 1.5 °C during 1982-2011 by calculating the annual average of temperatures from May to November (Galbraith et al., 2012). The predicted SST in Eastern Canada, through climate scenarios projections, indicate a possible rise by more than one degree Celsius during the next century (Galbraith et al., 2012). Therefore, this increase of water temperature could lead to the proliferation of *Vibrio* pathogens as well as shellfish contamination and human infections.

In fact, a recent surveillance study on the diversity and dynamics of the *Vibrio* communities in Canada's coasts (British Columbia [Pacific Coast], Nova Scotia [Atlantic Coast] and Gaspe) highlights the emergence of *V. cholerae* in temperate Canadian estuaries and the detection of pathogenic strains of *V. parahaemolyticus* in bivalve molluscs harvested in Canada (clams, mussels and oysters) with increasing trend during the warmest months of 2006-2016 (Banerjee et al., 2018).

In order to protect the shellfish industry as well as human health, modelling the future scenarios of SST in the Estuary and the GSL, with the aim of mapping future potential risk areas, is primordial. Predicting SST in GSL has been generally realized through deterministic model, based on physical and mathematical representation of the climatic and ocean processes, such as the three-dimensional coastal ice-ocean model called CANOPA (CANadian Océan PArallélisé) (Long et al., 2015; Saucier, 2003). Recently, we used machine learning models (Artificial Neuron Networks-ANN, and Random Forest-RF) in predicting daily SST in the GSL by entering a combination of predictors (also known as features) explaining most of SST variation: 3-day trailing Moving Average (MA) of daily mean air temperature (i.e. average of daily mean air temperature of the present and two previous days), the 30 day-MA of daily mean wind speed, the 30 day-MA of maximum daily tidal range, 120 day-MA mean St. Lawrence freshwater runoff and 60 day-MA of North Atlantic Oscillation (Ferchichi et al., 2019). The MAs are used as filters for smoothing predictors' data and detecting a better association between the dependent (SST) and independent variables. The results showed that Random Forests provided the best SST prediction accuracy of historical SST in the GSL (Ferchichi et al., 2019). In the same study, it has been demonstrated that both of the air temperature and wind speed are the most relevant predictors by explaining more than 70% of SST variation for most of the stations (Ferchichi et al., 2019). A recent study, focusing on coastal water temperature prediction, shows the impact of daily maximum and the average air temperature of previous 1 and 2 days on the daily water temperature variation. Considering this lag time factor and entering these variables as predictors improved significantly the daily coastal temperature prediction (Trinh et al., 2019).

In this paper, we present the future scenarios of *Vibrio* growth risk in the GSL by modelling the future water temperatures under different climate scenarios (optimistic and pessimistic). By entering the most relevant and readily available predictors (air temperature and wind speed) to the machine learning models (ANN and RF), we test their performance predictions and select the best inputs for each model using the backward selection method (Recursive Feature Elimination-RFE). After choosing the best model and entering the climate projections of the selected predictors, we produce the future water temperature in both of optimist and pessimist climatic scenarios. Finally, we map the future potential *Vibrio* growth risk area in the Estuary and GSL by interpolating the calculated risk indicator, in relation to the theoretical proliferation of pathogenic *Vibrio*, over our study area.

2. Study Area

The St. Lawrence River is the second largest river in North America (El-Sabh and Murty, 1990), with an average flow of approximately 12100 m^3/s (Galbraith et al., 2017). Originating from the Great Lakes, it reaches a vast estuary, where the fresh water of the river and salt water from the Atlantic Ocean mix. It flows over approximately 250 km to Pointe-des-Monts where it becomes the Gulf of St-Lawrence, opened to the Atlantic Ocean through the straits of Cabot and Belle-Isle. The GSL is one of the largest and most diverse marine ecosystems in the world covering an area of 225000 km^2.

The study region, as shown in (Fig 1), covers the coastal areas of the Estuary (downstream limit near Rimouski) and the Gulf of St-Lawrence (the Quebec North Shore, Gaspe, Baie-des-Chaleurs, PEI and Magdalen Islands).

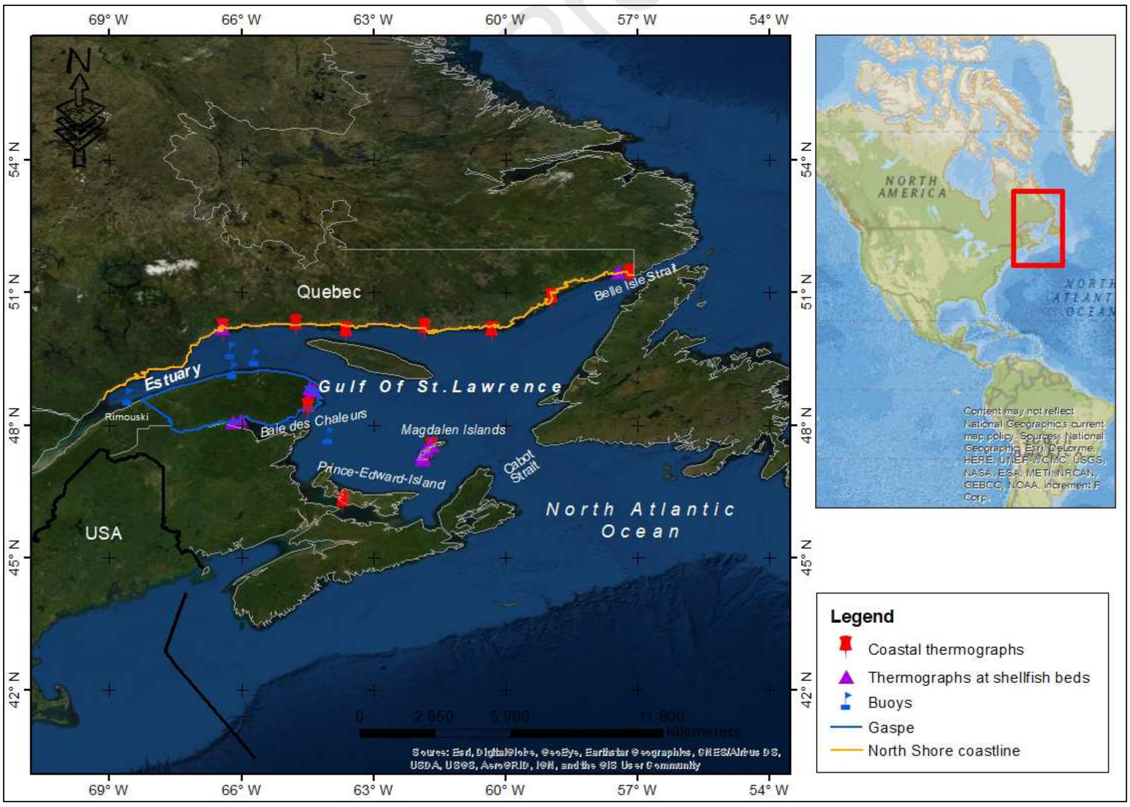

Figure 1. Geographic location of the thermographs and buoys in the Estuary and Gulf of Saint Lawrence

The most abundantly harvested shellfish in the GSL are oysters, mussels, clams and scallops (Statistics Canada, 2019). The major shellfish aquaculture techniques are the intertidal, subtidal and suspended cultures. The shellfish produced using the intertidal method are growing directly in the substrate. Since the clams and the oysters may be farmed intertidally, they are more exposed to higher water temperature at low tide, which increases the risk of *Vibrio proliferation*. In order to provide assurance of bivalve molluscs' safety for human consumption, a Canadian Shellfish Sanitation Programme (CSSP) was established (Sauvé, 2010). The aim of this program is to monitor the shellfish growing areas, classify them with regard to environmental conditions and water quality to determine the safety of shellfish consumption, to monitor the marine biotoxins and to control the shellfish harvesting and processing in these areas (Sauvé, 2010).

3. Material and methods

3.1 Data collection

For modelling daily mean water temperature, we use the 3 day-MA mean air temperature and 30 day-MA wind speed as they present the most relevant explanatory variables of SST variation in the most of stations. The daily air temperature and wind speed data are available online from the Government of Canada through this site: http://climate.weather.gc.ca/historical_data/search_historic_data_e.html.

The daily sea surface temperatures of the buoys and costal thermographs were supplied by Fisheries and Oceans Canada and the Maurice-Lamontagne Institute.

The daily sea surface temperatures of coastal thermographs, located at shellfish beds, were provided by MERINOV-Québec Centre for Innovation in Aquaculture and Fisheries.

The predictor projections (daily air temperature and the 30-day MA wind speed) come from eight climate simulations obtained from Ouranos-a climate-science consortium based in Quebec (Martynov et al., 2013; Šeparović et al., 2013), and the CORDEX program (Giorgi et al., 2009). These simulations are generated through regional climate models driven by global climate model under one of the two Representative Concentration Pathways (RCPs; RCP4.5 (Knutti and Sedláček, 2013) or RCP8.5 (Meinshausen et al., 2011)).

The scenarios used in this study were the average of the regional model outputs mentioned in the Table 1.

Table 1. List of the Regional Climate Models (RCMs) used in simulations.

Sources of RCMs	Modelling groups	Regional Climate Model (RCM)
Ouranos	_	CRCM5
CORDEX	DMI (Danish Meteorological Institute)	HIRHAM5
	UQUAM (L'Université du Québec à Montréal)	CRCM5

CCCma (Canadian Centre for Climate Modelling and Analysis)	CanRCM4

3.2 Modelling water temperature

We model the target variable (the daily water temperature of each buoy and coastal thermograph) by entering the selected predictors (3 day-MA air temperature and 30-day MA wind speed) into tested models (RF and ANN). The RFE was selected as feature selection method in order to choose the best subset of the predictors. In this case, only two combinations of the predictors were likely to be selected, either the air temperature as the sole input variable, or both air temperature and wind speed. 80% of original data are used for training and the remaining data serve as test data to evaluate the model predictive power. We use the k-fold cross validation (10-fold cross validation) as model validation technique.

3.2.1 Artificial Neural Network Model

In this study, we use a Multilayer Perceptron (MLP), a feedforward Artificial Neural Network, trained by using the supervised learning based on the error gradient back-propagation algorithm. This class of model is composed of three layers: the input layer includes the predictors which are standardized by subtracting each variable by its means and dividing by its standard deviations, the output layer, composed of single node, produces the response variable (water temperature) and the hidden layer connects both of the input and output layers. At the state of hidden layer, the ANN attributes weights to the set of inputs (x_i) and applies an activation function (f_1) on the weighted sum of inputs. Then, a linear function (f_2) is applied on the output of the hidden layer to produce the desired output (O), given by:

$$O = f_2\left[\sum_{j=1}^{n} w_{jk}\left[f_1\left(\sum_{j=1}^{n} w_{ij}x_i + b_j\right)\right] + b_0\right] \tag{1}$$

Where w_{ij} is the weight between the input x_i and hidden neuron j, b_j is the bias associated to each hidden neuron j, w_{jk} is the weight between the hidden neuron j and the output neuron k, and b_0 is the bias associated to the output neuron .

The activation function (f_1) used in this study is the sigmoid, given by:

$$f_1(x) = \frac{1}{1 + e^{-\mathrm{x}}} \tag{2}$$

The used MLP is composed by one hidden layer holding only one neuron. A single hidden layer is sufficient to approximate any continuous function, but there is no general rule for selecting the appropriate number of hidden nodes in the hidden layer (Haykin, 1994; Piotrowski et al., 2015). The ANN architecture selection is done via the trial-error method. It tests the model performance by using a grid search of hidden neurons number (1 to 10 hidden nodes). Adding more than one neuron does not improve significantly the prediction accuracy.

3.2.2 Random Forests Model

Random Forests (RF) is a recent machine learning algorithms (2000) developed by Braiman (Breiman, 2001). The RF is a tree-based ensemble method that randomly selects a subset of predictors to build a binary tree based on bootstrap samples of the training data (Breiman, 2001). The overall prediction is the average of the predictions from all the generated decision trees (Aggregation).

The RF generalization error is estimated by averaging the prediction error of each tree using the Out-Of-Bag (OOB) samples, i.e. samples that are not included in the bootstrap training sets (1/3 bootstrap samples of the training sets). This OOB error is computed with a Mean Square Error (MSE) as shown below:

$$MSE^{OOB} = \frac{1}{n} \times \sum_{i=1}^{n} [\hat{Y}(X_i) - Y_i]^2 \tag{3}$$

Where n is the size of the OOB sample, $\hat{Y}(X_i)$ corresponds to the RF output given the input sample X_i, and Y_i represents the actual output.

The parameters set are: the number of trees, the number of bootstrap input variables at each split of a tree (Ntry) and the minimum node size of each tree. The minimum node size of each tree, recommended by RF creators, and used in many studies is 5. The smaller the minimum node size is, the deeper that the tree is. The Ntry, recommended by the RF developers, is the number of input variables divided by three (Breiman, 2001) (Ntry=1). The number of trees would be experimentally set through plotting the OOB error plot in function of tree numbers (number of trees=50).

3.2.3 Feature Selection

The Recursive feature Elimination (RFE) presents a backward selection technique used in selecting the best subset of input variables (features) that contribute the most in model accuracy during the training process. The RFE consists in training the model by entering all the features then removing the variables with the lowest contribution in model accuracy, i.e., based on variable importance. Using the new reduced feature subset, it retrains the model (Guyon et al., 2002). The best selected subset is the one that optimizes the most of the chosen performance criteria. In this study, we chose the RMSE as performance criterion for best subset features selection.

We use the k-fold cross validation (k=10) in performance prediction assessment of the possible feature subsets. In fact, the goodness of models fit, according to each feature subset, is assessed by computing the 10-fold cross validated RMSE ($RMSE_{CV}$), i.e. the average of the ten RMSE computed over the 10 validation sets:

The ($RMSE_{CV}$) is given by:

$$RMSE_{CV} = \frac{1}{k}\sum_{j=1}^{k} RMSE_j \quad (4)$$

Where the RMSE calculated for each fold j is given by:

$$RMSE_j = \sqrt{\sum_{i=1}^{N} \frac{(y_{i,p} - y_{i,obs})^2}{N}} \qquad (5)$$

Where the k is the number of folds (10 in this example), j is one of the k folds, N is the sample size of the fold j, $y_{i,p}$ are predicted data and $y_{i,obs}$ are the observed data.

3.2.4 Performance evaluation of models

In this paper, we compare the prediction accuracy of tested models (ANN and RF) according to three performance criteria: the Root Mean Square Error (RMSE), Nash-Sutcliffe coefficient (Nash) and Relative mean bias (rBias).

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_{i,obs} - y_{i,p})^2}{n}} \qquad (6)$$

- Nash-Sutcliffe coefficient (Nash)

$$Nash = 1 - \frac{\sum_{i=1}^{n}(y_{i,obs} - y_{i,p})^2}{\sum_{i=1}^{n}(y_{i,obs} - \bar{y}_{obs})^2} \qquad (7)$$

- Relative mean bias (rBias)

$$rBias = \frac{100}{n} \times \sum_{i=1}^{n} \frac{y_{i,obs} - y_{i,p}}{y_{i,obs}} \qquad (8)$$

Where $y_{i,obs}$ is the observed data, $y_{i,p}$ is the predicted data, $\bar{y}_{obs}$ corresponds to the means of observed data and n is the size of observed data. Generally, a computed value of Nash greater than 0.5 indicates a relative satisfactory model performance, with a value of 1 corresponding to an ideal model (N. Moriasi et al., 2007). Low values of RMSE and rBias indicate better performing models.

3.3 Trend Analysis

After selecting the best model for all the stations, we generate the projections of the daily water temperature for the horizon 2040-2100. In order to cover the range of plausible water temperature scenarios, we select two different climatic scenarios: a relatively optimistic scenario (RCP4.5) and a pessimistic scenario (RCP8.5), known as “Business as usual” (i.e. continuous rise in GHG emissions). These scenarios lead to a warming of average air temperatures of 2.5 °C to 5 ° C around 2100. We perform a trend analysis of the predicted water temperature for each station under pessimistic scenario during the horizon 2040-2100 by using the Modified Mann Kendall (MMK) test, which takes into account the serial correlation. Then, we compute the trend slope of each station using the Theil-Sen’s slope estimator.

3.3.1 Modified Mann Kendall test (MMK)

The non-parametric Mann-Kendall (MK) test is commonly used for detecting monotonic trends in time series (Kendall, 1975; Mann, 1945). The null hypothesis H_0 of this test is that there is no trend in the series.

The MK test statistic S is calculated as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_i) \tag{9}$$

Where x_i and x_j denote the values of observations for the respective periods i and j (j> i), n is the length of the times series and $sgn(x_i\text{-}x_j)$ presents the sign function given by:

$$sgn(x_j - x_i) \begin{cases} 1 \; if \; x_j - x_i > 1 \\ 0 \; if \; x_j - x_i = 0 \\ -1 \; if \; x_j - x_i < 1 \end{cases} \tag{10}$$

Mann (1945) and Kendall (1975) have noted that for large values of n (n≥8), the distribution of the S statistic is approximately normal (Kendall, 1975; Mann, 1945), with the mean E and variance V of the statistic S, are defined as follows:

$$E(S) = 0 \tag{11}$$

$$V(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{m} t_i(i-1)(2i+5)}{18} \tag{12}$$

Where m represents the number of tied groups in the data set and the t_i represents the number of values in the ith tied group.

The standardized statistic Z_S is calculated by:

$$Z_S = \begin{cases} \dfrac{S-1}{\sqrt{V(s)}} & if\ S > 0 \\ 0 & if\ S = 0 \\ \dfrac{S+1}{\sqrt{V(s)}} & if\ S < 0 \end{cases} \tag{13}$$

The sign of the statistic Z_S indicates the direction of the trend whether it is upward (positive Z_S) or downward (negative Z_S). The standardized statistic Z_S follows the standard normal distribution with a mean of 0 and variance of 1. The null hypothesis would be rejected, implying the presence of a significant trend, when Z_S is higher than a critical value $Z_{1-\alpha/2}$, where α represents the chosen significance level (5% in this study) and $Z_{1-\alpha/2}$ could be deduced from the standard normal cumulative distribution tables.

In order to account for the autocorrelation that may exist in the time series, Hamed and Rao proposed to modify the variance of the MK test (Hamed and Rao, 1998). The variance is corrected through multiplying by the factor n/n^*, where n^* presents the effective sample size. Yue and Yang have demonstrated that incorporating the effective sample size in variance correction limits effectively the effect of serial correlation on the MK test (Yue and Wang, 2004).

$$\frac{n}{n^*} = 1 + \frac{2}{n(n-1)(n-2)} \sum_{i=1}^{p} (n-i)(n-i-1)(n-i-2)\ \rho_s\,(i) \tag{14}$$

Where n is the actual sample size, n^* is the effective sample size to account for autocorrelations in the data, ρ_s presents the autocorrelation function of the ranks of the observations for lag i and p is the maximum of time lags taking into account.

In this paper, the variance is corrected through considering complete autocorrelations (all lags) in the effective sample size computation, proposed by Yue and Wang (2004), and applied by using the mmky R package.

3.3.2 Theil-Sen's slope estimator

Theil-Sen's slope estimator, proposed by Theil (1950) (Theil, 1950) and Sen (1968) (Sen, 1968), allows to capture the direction and the strength of significant trend slope. It has been considered as robust estimate of the magnitude of trend's slope (Yue and Wang, 2004).

It is given by the following equation:

$$b = median\left(\frac{x_j - x_i}{j - i}\right) \; \forall \; i < j \tag{15}$$

3.4 Mapping future potential risk growth *Vibrio*

In order to evaluate the potential risk of *Vibrio* growth, we chose as thermal metric: the number of days above the minimum known temperature for *Vibrio* growth (15 °C). This thermal metric, selected as *Vibrio* growth risk indicator, is computed from the produced daily future water temperatures for both of optimistic and pessimistic scenarios.

We calculate the monthly average of this risk indicator, during the study period (June-October), averaged over twenty years during the study horizon 2040-2100. Subsequently, we interpolate the risk indicator computed for the available stations over the study area using the Inverse Distance Weighting (IDW) method, using ARCGIS. As a result of a comparison between IDW and kriging, we select IDW interpolation as it is the simplest method, given similar interpolation errors to kriging. We produce maps for two selected future horizons 2040-2060 and 2080-2100 under both climate change scenarios (optimistic-RCP4.5 and pessimistic-RCP8.5) in order to compare the level of potential *Vibrio* expansion risk over the study area.

We compute the root mean squared error (RMSE) of the interpolation using a leave-one-out procedure. The relative RMSE is calculated by dividing the RMSE of the IDW interpolation, produced for every month i of study period, by the areal average of the risk indicator for the same month. Sometimes the relative error is quite strong, which may be due to the large spatial variation of the selected variable. For cases that are too uncertain (relative error>50%), interpolation is useless.

3.4.1 Inverse Weighted Distance (IWD) interpolation method

The IDW method, a deterministic spatial interpolation approach, allows to compute an average of a selected variable in ungauged sites using values from nearby weighted sites.

The weights, accorded to gauged locations, are proportional to the distance between the gauged and ungauged sites and determined by the IDW power coefficient. The larger the power coefficient is, the stronger the weights are attributed to the closest locations. The estimated variable at ungauged location (z_j*) is defined by the following equation:

$$z_j^* = \frac{\sum_{i=1}^{n} w_{ij} x_i}{\sum_{i=1}^{n} w_{ij}} \tag{16}$$

Where:

$$w_{ij} = \frac{1}{d_{ij}^p} \tag{17}$$

Where x_i is the variable value of a neighboring gauged site, w_{ij} is the weight assigned to the gauged sites (i), d_{ij} is the distance between the gauged (i) and ungauged sites (j), n is the number of gauged sites and p is the exponent of the distance. In this study, the interpolation was performed using the ArcGIS software and a default value of p = 2 was chosen.

4 Results

After applying the Recursive Feature Elimination (RFE) on the selected potential predictors, air temperature and wind speed, we found that the ANN uses only the air temperature as input for most of the stations while the RF uses both predictors. Table 2 presents the results of performance criteria for ANN and RF in SST prediction of the tested dataset.

Table 2. Performance criteria results of tested models (RF and ANN)

Models	**Stations**	**RMSE (°C)**	**Nash**	**rBiais(%)**
	Grande Rivière	1.7	0.708	-2.05
	Borden	1.991	0.585	-1.529
	Blanc Sablon	2.437	0.331	-13.774
	Courant de Gaspé	1.368	0.84	-1.235
	Havre St Pierre	2.435	0.29	-12.67
	Ile Shag	2.097	0.651	-4.008
ANN	Montlouis	1.323	0.854	-1.647
(Artificial Neural Network)	Natashquan	2.028	0.709	-6.409
	Rimouski	1.139	0.827	-0.779
	Rivière aux Tonnerre	2.015	0.627	-7.809
	Romaine	1.345	0.816	-2.929
	Tabatière	2.342	0.391	-9.577
	Sept-Îles	1.957	0.702	-3.175
	Shediac Valley	1.056	0.893	-0.568
	Mean	**1.802**	**0.659**	**-4.868**
	Grande Rivière	1.654	0.708	-1.877
	Borden	1.837	0.585	-0.66
	Blanc Sablon	2.29	0.331	-10.614
	Courant de Gaspé	1.37	0.84	-1.018
	Havre St Pierre	2.519	0.29	-10.844
	Ile Shag	1.958	0.685	-3.478
RF	Montlouis	1.418	0.687	-1.407
(Random Forests)	Natashquan	2.082	0.693	-4.756
	Rimouski	1.157	0.827	-0.794
	Rivière aux Tonnerre	2.08	0.627	-6.377
	Romaine	1.404	0.816	-2.496
	Tabatière	2.323	0.391	-7.193
	Sept-Îles	1.982	0.702	-3.676
	Shediac Valley	1.02	0.893	-0.43
	Mean	**1.792**	**0.648**	**-3.973**

By using the paired t-test (DF=13, t-value=0.436) on RMSE values at all of our sites for both models, we note that there is no significant difference between the performances of models in terms of RMSE (p-value=0.67) at a significance level of 5%. The average RMSE performance for all the stations is approximately 1.8 °C for both of RF and ANN. Both models present good performing results in terms of Nash-Sutcliff criterion, i.e. higher than 0.5, and low relative mean bias (<5%). Given that there is no significant difference between RF and ANN, we choose the ANN as it is the most parsimonious model for modelling future water temperatures. Then, we proceed to generate future daily mean water temperature for each station for the horizon 2040-2100 for a pessimistic climate scenario, RCP8.5, and a more optimistic one (RCP4.5), by using the projections of air temperature.

We proceed with a trend analysis of the predicted water temperature for each station under pessimistic scenario during the horizon 2040-2100 by using the Modified Mann Kendall (MMK) test, which takes into account the serial correlation. Significant positive trends (p-value<1%) in future daily mean SST, for the period from June to October in the horizon 2040-2100, were revealed in all the tested stations at a significance level of 5%. Table 3 presents the results of Theil Sen's slope computed for each station after applying the MMK test. By averaging the trend slopes of all the stations over the horizon (2040-2100), we found that the water temperatures are likely to increase by 0.4 °C per decade, for a total of 2.5 °C up to 2100.

426 *Table 3. Theil-Sen's slope for projected daily mean water temperature times series*
427 *during the horizon (2040-2100)*

Stations	**Theil-Sen's slope**	**[a]Slope-10 years**	**[b]Slope-60years**
Natashquan	4.59E-04	0.711	4.266
Baie Cascapedia	1.69E-04	0.262	1.57
Baie Trascapedia	1.73E-04	0.269	1.613
Sept-Îles	2.42E-04	0.376	2.255
Baie Plaisance	1.34E-04	0.208	1.245
Lagune Havre	3.31E-04	0.513	3.079
Bassin Havre	3.71E-04	0.575	3.447
Lagune Grande	3.16E-04	0.489	2.935
Belles Amours	2.81E-04	0.436	2.618
Blanc Sablon	2.64E-04	0.41	2.458
Borden	1.49E-04	0.232	1.39
Grande Rivière	1.66E-04	0.257	1.54
Iles Shag	1.50E-04	0.233	1.4
Havre St Pierre	1.75E-04	0.271	1.624
Rivière aux Tonnerre	4.10E-04	0.635	3.813
Romaine	4.49E-04	0.696	4.176
Tarbatière	3.30E-04	0.512	3.072
Rimouski	3.77E-04	0.584	3.506
Courant Gaspé	2.57E-04	0.399	2.391
Gyre Anticosti	2.02E-04	0.312	1.875

Montlouis	2.63E-04	0.407	2.443
Shediac Valley	2.03E-04	0.315	1.889
Mean	**2.67E-04**	**0.414**	**2.482**

[a] **Slope-10 years:** the average of Theil-Sen's slope for 10 years for the time period between June and October

[b] **Slope-60years:** the average of Theil-Sen's slope for 60 years for the time period between June and October

After generating the future daily means of SST, we calculate the number of days exceeding the threshold of 15 °C, the minimum temperature for *Vibrio* growth, as a risk indicator. Then, we interpolate the values of this risk indicator, averaged for each month over 20 years. We present the results of August and September as examples of spatial interpolation of the risk indicator over the study area in Figures 2 and 3 respectively.

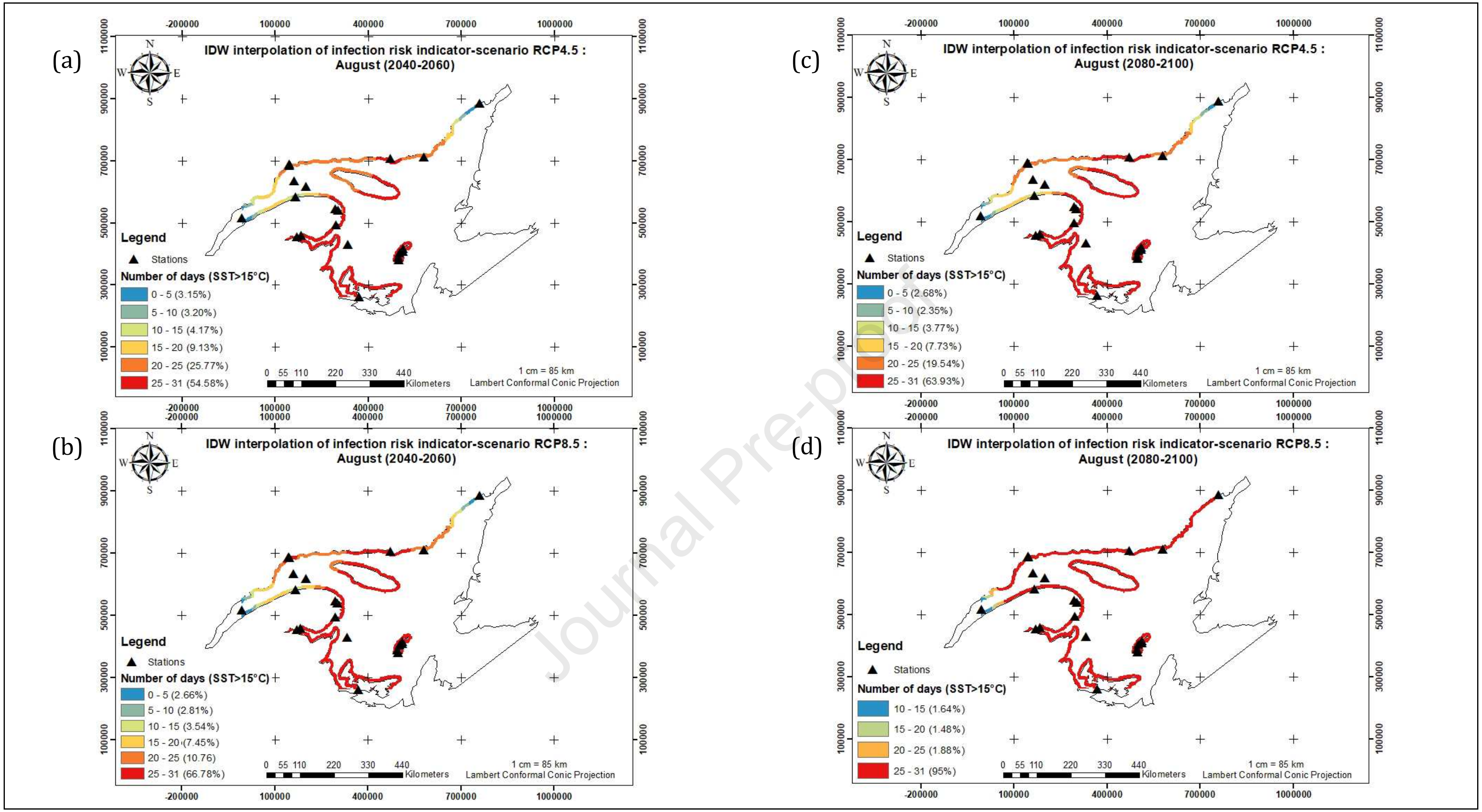

Figure 2. Inverse Distance Weighting (IDW) interpolation of risk indicators (number of days exceeding the threshold (15 °C)) over the Estuary and Gulf of St. Lawrence under pessimistic and optimistic climatic scenario for the horizons (2040-2060) and (2080-2100) in August. ***(a)*** *IDW interpolation of the risk indicator in August during the horizon (2040-2060) under optimistic scenario (RCP4.5).* ***(b)*** *IDW interpolation of the risk indicator in August during the horizon (2040-2060) under pessimistic scenario (RCP8.5).* ***(c)*** *IDW interpolation of the risk indicator in August during the horizon (2080-2100) under optimistic scenario (RCP4.5).* ***(d)*** *IDW interpolation of the risk indicator in August during the horizon (2080-2100) under pessimistic scenario (RCP8.5).*

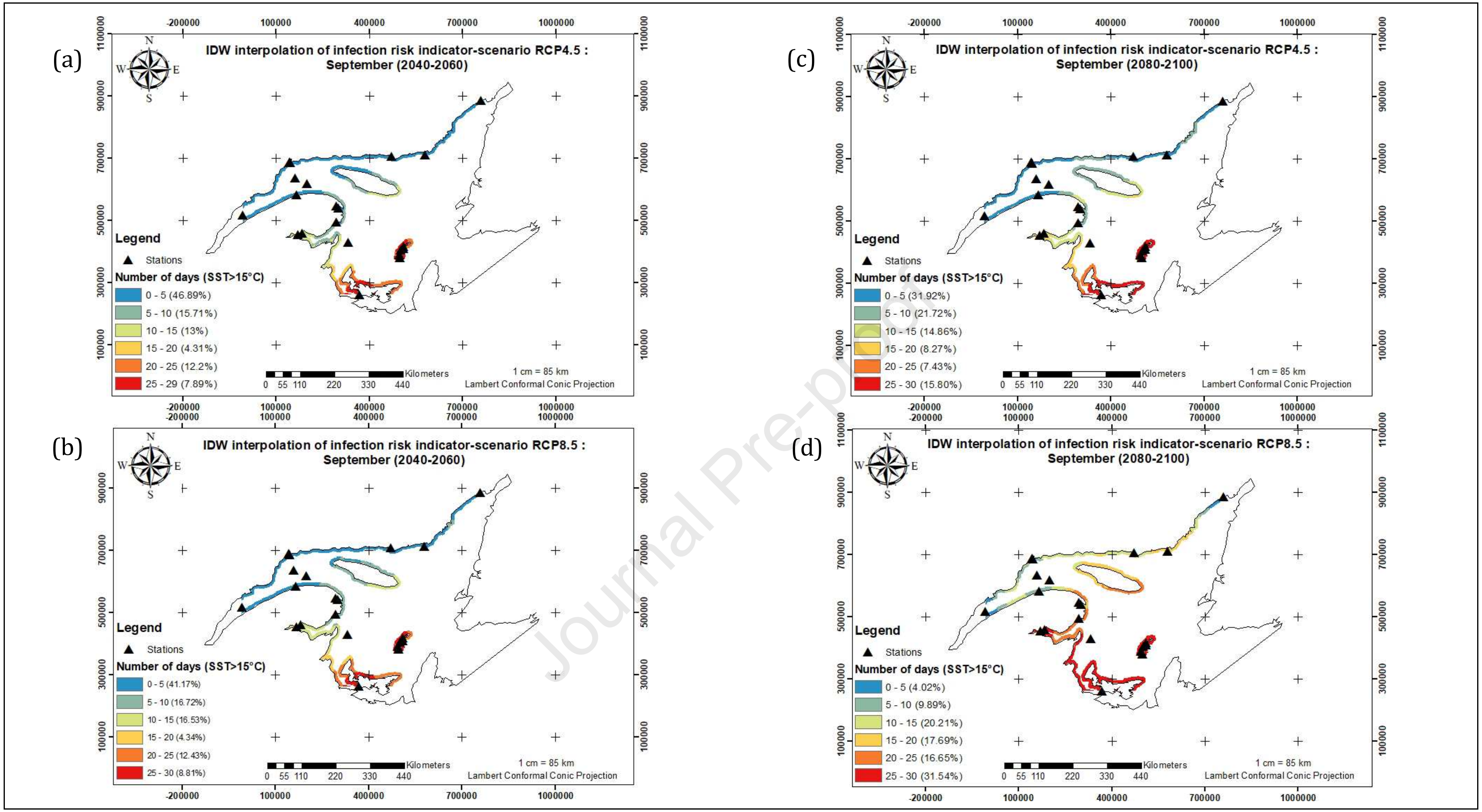

Figure 3. Inverse Distance Weighting (IDW) interpolation of risk indicators (number of days exceeding the threshold (15 °C)) over the Estuary and Gulf of St. Lawrence under pessimistic and optimistic climatic scenario for the horizons (2040-2060) and (2080-2100) in September. ***(a)*** *IDW interpolation of the risk indicator in September during the horizon (2040-2060) under optimistic scenario (RCP4.5).* ***(b)*** *IDW interpolation of the risk indicator in September during the horizon (2040-2060) under pessimistic scenario (RCP8.5).* ***(c)*** *IDW interpolation of the risk indicator in September during the horizon (2080-2100) under optimistic scenario (RCP4.5).* ***(d)*** *IDW interpolation of the risk indicator in September during the horizon (2080-2100) under pessimistic scenario (RCP8.5).*

Figure 2(a) shows that, during the horizon 2040-2060, under an optimistic scenario (RCP4.5), the waters in the shellfish beds of Magdalen Islands, PEI, the Gaspe Peninsula and Baie des Chaleurs are likely to be at high risk of infection as the risk indicator (number of days above the 15 °C) exceeds 25 days. The risk indicator of some stations in the Quebec North Shore coast along the GSL varies between 20 and 25 days, so the shellfish beds in this area might also be under high risk of *Vibrio* growth. For the same horizon, but under a pessimistic scenario as shown in Figure 2(b), most shellfish beds would probably be at high risk of *Vibrio* growth since the calculated risk indicator exceeds 25 days for approximately 67% of coastal areas.

By comparing the interpolation maps of August (Fig 2(c) and Fig 2(d)) during the horizon 2080-2100 for both scenarios, we note that the risk indicator, exceeding 25 days, covers between 64% (scenario RCP4.5) and 95% (scenario RCP8.5) of the total coastal area. Therefore, by 2100, most of the stations located in shellfish beds, where harvesting occurs, are likely to be at high risk of *Vibrio* growth whatever the considered scenario.

During September (Fig 3), the risk indicator at the shellfish beds of the Magdalen Islands and PEI exceeds 20 days for both climate scenarios and horizons. So, they might be at a higher risk of *Vibrio* growth. During the horizon (2080-2100) under a pessimistic scenario (Fig 3(d)), we note that in addition to the Magdalen Islands and PEI, the shellfish beds of Gaspe Peninsula and Baie des Chaleurs would probably be at high risk of *Vibrio* growth as the number of days above the 15 °C threshold exceeds 20 days.

Table 4 presents the results of the risk indicator calculated for the month of October for some stations under a pessimistic scenario during the horizon 2080-2100. It was not possible to perform a spatial interpolation because too few stations had non-zero values. Table 4 shows that the risk of *Vibrio* growth may also occur during October on the coasts of PEI and Magdalen Islands. Therefore, the risk of *Vibrio* growth would probably expand both spatially and temporally (i.e. into the fall for some regions).

Table 4. The number of days exceeding the minimum temperature threshold (15 °C) for the growth of pathogenic Vibrio during October in the horizon (2080-2100) under pessimistic scenario (RCP-8.5)

Localisation	**stations**	**October**
PEI	Borden	19
Magdalen Islands	Baie Plaisance	6
	Bassin Havre	20
	Lagune Havre	16
	Lagune Grande	14
	Ile Shag	4

The risk assessment was not just limited to one risk indicator. We also compute the number of days exceeding a threshold of 20 °C. In fact, the increasing water temperature trend under the pessimistic scenario results in exceedance of the higher temperatures thresholds (20 °C) associated with higher abundance of *Vibrio*, in contrast with the optimistic scenario where the water temperature does not exceed the 20°C threshold. Table 5 presents the stations that would be at high risk of pathogenic *Vibrio* growth. For the rest of the stations, the number of days exceeding 20 °C is zero so they have not been included in the table. Table 5 shows that during 2080-2100 under a pessimistic scenario, the threshold of 20 °C would be exceeded along the coast of the Magdalen Islands , where the blue mussels, clams and scallops are harvested, the threshold of 20 °C would be exceeded for about 31 days in August against an average of 18 days in September".

The shellfish beds of mussels and scallops in the Gaspe Peninsula could be at the same degree of risk with about 30 days in August while the shellfish beds of PEI (mussels, oysters and clams) and North Shore would be under a lower risk with an average of 16 days.

Table 5. The number of days exceeding the threshold (20 °C) in the horizon (2080-2100) under pessimistic scenario (RCP 8.5)

Localisation	**stations**	**August**	**September**
North Shore	Natashquan	17	0
Magdalen Islands	Bassin Havre	31	22
	Lagune Havre	31	16
	Lagune Grande	31	16
Gaspé	Havre Gaspé	30	4
PEI	Borden	16	0

5 Discussion

By using one or both of the most relevant predictors, air temperature and wind speed, the results show that the SST prediction performance of ANN and RF were similar. However, RF requires more predictors than ANN to achieve similar prediction performance. Thus, we select the ANN for SST prediction and used the mean of climate models projections of air temperature, as input for ANN model, without exploring the variability between the projections.

Modelling future water temperature through ANN constitutes a useful tool to predict the plausible future water temperatures in the coastline of the Estuary and GSL where shellfish beds occur. The trend analysis results for daily mean water temperature, using the MMK test, indicates that our study area exhibits a significant increasing trend by 2.5 °C up to 2100 under a pessimistic scenario. This positive trend in water temperature implies a rise in risk indicator of *Vibrio* growth. This result is demonstrated in the interpolation maps between the horizons 2040-2060 and 2080-2100. In fact, by comparing the interpolation maps in August for both of the horizons under the pessimistic scenario (RCP8.5), we note an expansion of *Vibrio* growth risk from 64% to 95% of the total coastal area of the Estuary and the Gulf of Saint Lawrence. The risk indicator distribution during July was similar to August so the shellfish beds would be exposed to a similar risk as in August under both scenarios. Whereas, in June the risk of *Vibrio* growth would be less severe than August and July except over the horizon (2080-2100) under pessimistic scenario, where all the shellfish beds on the coasts of North Shore, Gaspe Peninsula, Baie des Chaleurs, Magdalen Islands and PEI would be under high risk of *Vibrio* growth. Their risk indicator could exceed an average of 20 days.

The results of risk indicator interpolation in August, suggest that *Vibrio* growth risk may increase under both of pessimistic or optimistic scenario so all the shellfish beds practically (on the coasts of North Shore, Gaspe Peninsula, Baie des Chaleurs, Magdalen Islands and PEI) would be at risk of *Vibrio* growth regardless the scenario. In addition to this spatial spread, the *Vibrio* growth risk would extend seasonally by occurring out of the summer time during September and even October, especially on the coasts of Magdalen Islands and PEI.

The lowest temperature threshold for the *Vibrio* growth, based on the literature, is 15 °C. Computing the number of days exceeding a higher temperature threshold (20 °C) allowed to locate the shellfish beds that would be at higher risk of *Vibrio* growth, like the Magdalen Islands and PEI.

This study focused on surface temperatures (average depth of 1.5 m). Wild molluscs can be found or harvested at this depth, e.g. on the foreshore or near the islands. However, in some cases, molluscs harvesting occurs in deeper water (e.g. oysters). The risk maps produced may be biased, i.e. the risk of *Vibrio* growth may be overestimated, as there may be significant thermal stratification in some coastal sites. Molluscs that are found under the thermocline may not be as much as risks as those in shallow, well mixed areas.

In this study, we focused on water temperature as the main factor affecting *Vibrio* growth. However, it should be noted that many studies confirm the importance of salinity in *Vibrio* growth. So, future work on *Vibrio* risk management should concentrate of the combined effects of water temperature and salinity on the proliferation of pathogenic *Vibrio* and use the projections of both variables to better locate the potential risk areas followed by sampling of water and shellfish to confirm the presence of *Vibrio*.

Acknowledgements

The authors would like to thank Ouranos for funding this research.

We also thank Ouranos for generating and supplying the Canadian Regional Climate Model (CRCM5) data.

The authors would like to thank Pr. Yves Gratton for helping to find data sources.

The authors would like to thank MERINOV for providing us with the data in the shellfish areas.

We would also like to thank Estelle Pedneault and Céline Campagna for their contribution and help.

The authors would like to thank Gabriel Rondeau-Genesse for providing us with the required variable projections.

The Canadian Regional Climate Model (CRCM5) was developed by the ESCER Centre at UQAM (Université du Québec à Montréal) with the collaboration of Environment and Climate Change Canada.

CRCM5 computations were made on the supercomputer guillimin from McGill University, managed by Calcul Québec and Compute Canada. The operation of this supercomputer is funded by the Canada Foundation For Innovation (CFI), the ministère de l'Économie, de la science et de l'innovation du Québec (MESI) and the Fonds de recherché du Québec- Nature et technologies (FRQ-NT).

We also thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and The Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) for the funding of the development of the CRCM5.

We acknowledge the World Climate Research Programme's Working Group on Regional Climate, and the Working Group on Coupled Modelling, former coordinating body of CORDEX and responsible panel for CMIP5. We also thank the climate modelling groups (listed in Table 1 on this paper) for producing and making available their model output. We also acknowledge the U.S. Department of Defense ESTCP for its support of the NA-CORDEX data archive.

The authors are grateful to the Editor, Dr. Steve Mitchell, and to anonymous reviewers for their comments which helped improve the quality of the manuscript.

References

Baker-Austin, C., Stockley, L., Rangdale, R., Martinez-Urtaza, J., 2010. Environmental occurrence and clinical impact of Vibrio vulnificus and Vibrio parahaemolyticus: a European perspective. Environmental microbiology reports 2, 7-18.

Baker-Austin, C., Trinanes, J., Gonzalez-Escalona, N., Martinez-Urtaza, J., 2017. Non-Cholera Vibrios: The Microbial Barometer of Climate Change. Trends in microbiology 25, 76-84.

Baker-Austin, C., Trinanes, J.A., Taylor, N.G., Hartnell, R., Siitonen, A., Martinez-Urtaza, J., 2013. Emerging Vibrio risk at high latitudes in response to ocean warming. Nature Climate Change 3, 73.

Baker-Austin, C., Trinanes, J.A., Taylor, N.G.H., Hartnell, R., Siitonen, A., Martinez-Urtaza, J., 2012. Emerging Vibrio risk at high latitudes in response to ocean warming. Nature Climate Change 3, 73-77.

Banerjee, S.K., Rutley, R., Bussey, J., 2018. Diversity and dynamics of the Canadian coastal Vibrio community: an emerging trend detected in the temperate regions. Journal of bacteriology 200, e00787-00717.

Breiman, L., 2001. Random forests. Machine learning 45, 5-32.

Burge, C.A., Mark Eakin, C., Friedman, C.S., Froelich, B., Hershberger, P.K., Hofmann, E.E., Petes, L.E., Prager, K.C., Weil, E., Willis, B.L., Ford, S.E., Harvell, C.D., 2014. Climate change influences on marine infectious diseases: implications for management and society. Annual review of marine science 6, 249-277.

Centers for Disease, C., Prevention, 2013. Incidence and trends of infection with pathogens transmitted commonly through food - foodborne diseases active surveillance network, 10 U.S. sites, 1996-2012. MMWR. Morbidity and mortality weekly report 62, 283-287.

Chu, C., Do, Y., Kim, Y., Saito, Y., Lee, S.-D., Park, H., Lee, J.-K., 2011. Mathematical modeling of Vibrio vulnificus infection in Korea and the influence of global warming. Osong public health and research perspectives 2, 51-58.

Davis, B.J.K., Jacobs, J.M., Davis, M.F., Schwab, K.J., DePaola, A., Curriero, F.C., 2017. Environmental determinants of Vibrio parahaemolyticus in the Chesapeake Bay. Applied and environmental microbiology.

Dechet, A.M., Yu, P.A., Koram, N., Painter, J., 2008. Nonfoodborne Vibrio infections: an important cause of morbidity and mortality in the United States, 1997-2006. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 46, 970-976.

Deeb, R., Tufford, D., Scott, G.I., Moore, J.G., Dow, K., 2018. Impact of Climate Change on Vibrio vulnificus Abundance and Exposure Risk. Estuaries and Coasts 41, 2289-2303.

El-Sabh, M.I., Murty, T.S., 1990. Mathematical modelling of tides in the St. Lawrence Estuary, Oceanography of a Large-Scale Estuarine System. Springer, pp. 10-50.

Feldhusen, F., 2000. The role of seafood in bacterialfoodborne diseases. Microbes and infection 2, 1651-1660.

Ferchichi, H., St-Hilaire, A., Ouarda, T.B.M.J., Lévesque, B., 2019. Modélisation des scénarios futurs de température de l’eau en milieu côtier et implications sur les infections potentielles par Vibrio parahaemolyticus et Vibrio vulnificus : application aux bancs coquillers de l’estuaire et du golfe du st-laurent (Master's thesis), Centre Eau Terre Environnement. Institut national de la recherche scientifique, Quebec, Canada, p. 132.

Fyfe, M., Yeung, S.T., Daly, P., Schallie, K., Kelly, M.T., Buchanan, S., 1997. Outbreak of Vibrio parahaemolyticus related to raw oysters in British Columbia. Can Commun Dis Rep 23, 145-148.

Galanis, E., Otterstatter, M., Taylor, M., 2020. Measuring the impact of sea surface temperature on the human incidence of Vibrio sp. infection in British Columbia, Canada, 1992–2017. Environmental Health 19, 58.

Galbraith, P.S., Chassé, J., Gilbert, D., Larouche, P., Brickman, D., Pettigrew, B., Devine, L., Gosselin, A., Pettipas, R., Lafleur, C., 2017. Physical oceanographic conditions in the Gulf of St. Lawrence in 2016. Canadian Science Advisory Secretariat.

Galbraith, P.S., Larouche, P., Chassé, J., Petrie, B., 2012. Sea-surface temperature in relation to air temperature in the Gulf of St. Lawrence: Interdecadal variability and long term trends. Deep Sea Research Part II: Topical Studies in Oceanography 77-80, 10-20.

Giorgi, F., Jones, C., Asrar, G.R., 2009. Addressing climate information needs at the regional level: the CORDEX framework. World Meteorological Organization (WMO) Bulletin 58, 175.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning 46, 389-422.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R., Watson, R., 2008. A global map of human impact on marine ecosystems. Science 319, 948-952.

Hamed, K.H., Rao, A.R., 1998. A modified Mann-Kendall trend test for autocorrelated data. Journal of hydrology 204, 182-196.

Haykin, S., 1994. Neural networks: a comprehensive foundation. Prentice Hall PTR.

Heng, S.-P., Letchumanan, V., Deng, C.-Y., Ab Mutalib, N.-S., Khan, T.M., Chuah, L.-H., Chan, K.-G., Goh, B.-H., Pusparajah, P., Lee, L.-H., 2017. Vibrio vulnificus: An Environmental and Clinical Burden. Frontiers in microbiology 8, 997-997.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jacobs, J., Moore, S.K., Kunkel, K.E., Sun, L., 2015. A framework for examining climate-driven changes to the seasonality and geographical range of coastal pathogens and harmful algae. Climate Risk Management 8, 16-27.

Jacobs, J.M., Rhodes, M., Brown, C.W., Hood, R.R., Leight, A., Long, W., Wood, R., 2014. Modeling and forecasting the distribution of Vibrio vulnificus in Chesapeake Bay. Journal of applied microbiology 117, 1312-1327.

Kaspar, C.W., Tamplin, M.L., 1993. Effects of temperature and salinity on the survival of Vibrio vulnificus in seawater and shellfish. Applied and environmental microbiology 59, 2425-2429.

Kendall, M., 1975. Rank Correlation Methods. Griffin, London.

Knutti, R., Sedláček, J., 2013. Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change 3, 369.

Konrad, S., Paduraru, P., Romero-Barrios, P., Henderson, S.B., Galanis, E., 2017. Remote sensing measurements of sea surface temperature as an indicator of Vibrio parahaemolyticus in oyster meat and human illnesses. Environmental Health 16, 92.

Lima, F.P., Wethey, D.S., 2012. Three decades of high-resolution coastal sea surface temperatures reveal more than warming. Nature Communications 3, 704.

Long, Z., Perrie, W., Chassé, J., Brickman, D., Guo, L., Drozdowski, A., Hu, H., 2015. Impacts of Climate Change in the Gulf of St. Lawrence. Atmosphere-Ocean 54, 337-351.

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica: Journal of the Econometric Society, 245-259.

Martinez-Urtaza, J., Bowers, J.C., Trinanes, J., DePaola, A., 2010. Climate anomalies and the increasing risk of Vibrio parahaemolyticus and Vibrio vulnificus illnesses. Food Research International 43, 1780-1790.

Martinez-Urtaza, J., Huapaya, B., Gavilan, R.G., Blanco-Abad, V., Ansede-Bermejo, J., Cadarso-Suarez, C., Figueiras, A., Trinanes, J., 2008. Emergence of asiatic vibrio diseases in south america in phase with El Niño. Epidemiology 19, 829-837.

Martynov, A., Laprise, R., Sushama, L., Winger, K., Šeparović, L., Dugas, B., 2013. Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian Regional Climate Model, version 5: model performance evaluation. Climate Dynamics 41, 2973-3005.

McLaughlin, 2005. climate anomalies and the increasing risk of Vibrio.

McLaughlin, J.B., DePaola, A., Bopp, C.A., Martinek, K.A., Napolilli, N.P., Allison, C.G., Murray, S.L., Thompson, E.C., Bird, M.M., Middaugh, J.P., 2005. Outbreak of Vibrio parahaemolyticus gastroenteritis associated with Alaskan oysters. New England Journal of Medicine 353, 1463-1470.

Mead, P.S., Slutsker, L., Dietz, V., McCaig, L.F., Bresee, J.S., Shapiro, C., Griffin, P.M., Tauxe, R.V., 1999. Food-related illness and death in the United States. Emerging infectious diseases 5, 607.

Meinshausen, M., Smith, S.J., Calvin, K., Daniel, J.S., Kainuma, M.L.T., Lamarque, J.-F., Matsumoto, K., Montzka, S.A., Raper, S.C.B., Riahi, K., Thomson, A., Velders, G.J.M., van Vuuren, D.P.P., 2011. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Climatic Change 109, 213.

Motes, M.L., DePaola, A., 1996. Offshore suspension relaying to reduce levels of Vibrio vulnificus in oysters (Crassostrea virginica). Applied and environmental microbiology 62, 3875-3877.

Motes, M.L., DePaola, A., Cook, D.W., Veazey, J.E., Hunsucker, J.C., Garthright, W.E., Blodgett, R.J., Chirtel, S.J., 1998. Influence of Water Temperature and Salinity on Vibrio vulnificus in Northern Gulf and Atlantic Coast Oysters (Crassostrea virginica). Applied and environmental microbiology 64, 1459-1465.

N. Moriasi, D., G. Arnold, J., W. Van Liew, M., L. Bingner, R., D. Harmel, R., L. Veith, T., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE 50, 885-900.

Narjol, G.-E., Viviana, C., Claudia, A., María, L.R., Juan, A.V., Felipe, C., Jaime, R., Romilio, T.E., 2005. Vibrio parahaemolyticus Diarrhea, Chile, 1998 and 2004. Emerging Infectious Disease journal 11, 129.

Newton, A., Kendall, M., Vugia, D.J., Henao, O.L., Mahon, B.E., 2012. Increasing rates of vibriosis in the United States, 1996-2010: review of surveillance data from 2 systems. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 54 Suppl 5, S391-395.

Oliver, J.D., 2005. Wound infections caused by Vibrio vulnificus and other marine bacteria. Epidemiology and Infection 133, 383-391.

Oliver, J.D., 2013. Vibrio vulnificus: Death on the Half Shell. A Personal Journey with the Pathogen and its Ecology. Microbial ecology 65, 793-799.

Piotrowski, A.P., Napiorkowski, M.J., Napiorkowski, J.J., Osuch, M., 2015. Comparing various artificial neural network types for water temperature prediction in rivers. Journal of Hydrology 529, 302-315.

Saucier, F.J., 2003. Modeling the formation and circulation processes of water masses and sea ice in the Gulf of St. Lawrence, Canada. Journal of Geophysical Research 108.

Sauvé, G., 2010. Official control monitoring programmes for live bivalve molluscs–legislative and regulatory approaches: Canada. In: Rees G, Pond K, Kay D, Bartram J, Domingo JS (eds) Safe Management of Shellfish and Harvest Waters, pp.217–232. World Health Organization (WHO), London.

Semenza, J.C., Trinanes, J., Lohr, W., Sudre, B., Lofdahl, M., Martinez-Urtaza, J., Nichols, G.L., Rocklov, J., 2017. Environmental Suitability of Vibrio Infections in a Warming Climate: An Early Warning System. Environmental health perspectives 125, 107004.

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. Journal of the American statistical association 63, 1379-1389.

Šeparović, L., Alexandru, A., Laprise, R., Martynov, A., Sushama, L., Winger, K., Tete, K., Valin, M., 2013. Present climate and climate change over North America as simulated by the fifth-generation Canadian regional climate model. Climate Dynamics 41, 3167-3201.

Statistics Canada, 2019. Table 32-10-0107-01 Aquaculture, production and value.

Taylor, M., Cheng, J., Sharma, D., Bitzikos, O., Gustafson, R., Fyfe, M., Greve, R., Murti, M., Stone, J., Honish, L., 2018. Outbreak of Vibrio parahaemolyticus associated with consumption of raw oysters in Canada, 2015. Foodborne pathogens and disease 15, 554-559.

Theil, H., 1950. A rank-invariant method of linear and polynominal regression analysis (parts 1-3), Ned. Akad. Wetensch. Proc. Ser. A, pp. 1397-1412.

Trinh, N.X., Trinh, T.Q., Phan, T.P., Thanh, T.N., Thanh, B.N., 2019. Water Temperature Prediction Models in Northern Coastal Area, Vietnam.

United States Food and Drug Administration (FDA), 2005. Quantitative Risk Assessment on the Public Health Impact of Pathogenic Vibrio parahaemolyticus.

Vezzulli, L., Colwell, R.R., Pruzzo, C., 2013. Ocean warming and spread of pathogenic vibrios in the aquatic environment. Microbial ecology 65, 817-825.

Vezzulli, L., Grande, C., Reid, P.C., Helaouet, P., Edwards, M., Hofle, M.G., Brettar, I., Colwell, R.R., Pruzzo, C., 2016. Climate influence on Vibrio and associated human diseases during the past half-century in the coastal North Atlantic. Proceedings of the National Academy of Sciences of the United States of America 113, E5062-5071.

Yue, S., Wang, C., 2004. The Mann-Kendall Test Modified by Effective Sample Size to Detect Trend in Serially Correlated Hydrological Series. Water Resources Management 18, 201-218.

Zimmerman, A.M., DePaola, A., Bowers, J.C., Krantz, J.A., Nordstrom, J.L., Johnson, C.N., Grimes, D.J., 2007. Variability of total and pathogenic Vibrio parahaemolyticus densities in northern Gulf of Mexico water and oysters. Applied and environmental microbiology 73, 7589-7596.

The highlights are:

- The statistical modelling (Random forest and Artificial Neural Network) of the coastal water temperature from air temperature and wind speed.
- Future coastal water temperature scenarios were produced under optimistic and pessimistic climate scenarios, using Artificial Neural Network model.
- Maps of the future Vibrio growth risk indicator were produced, from the future water temperature scenarios.
- Maps show that the Vibrio risk will increase spatially and seasonally regardless the climate change scenario.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: