





Article

Forecasting Seasonal *Vibrio parahaemolyticus* Concentrations in New England Shellfish

Meghan A. Hartwick^{1,2}, Erin A. Urquhart^{1,3}, Cheryl A. Whistler^{1,2}, Vaughn S. Cooper⁴ ,
Elena N. Naumova⁵  and Stephen H. Jones^{1,2,3,*}

¹ Northeast Center for Vibrio Disease and Ecology, University of New Hampshire, Durham, NH 03824, USA; mah2002@wildcats.unh.edu (M.A.H.); Urquhart.erin@epa.gov (E.A.U.); Cheryl.Whistler@unh.edu (C.A.W.)

² Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA

³ Department of Natural Resources and the Environment, University of New Hampshire, Durham, NH 03824, USA

⁴ Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA; vaughn.cooper@pitt.edu

⁵ Division of Nutrition Data Sciences, Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA 02111, USA; elena.naumova@tufts.edu

* Correspondence: stephen.jones@unh.edu; Tel.: +1-(603)-862-5124

Received: 18 September 2019; Accepted: 30 October 2019; Published: 7 November 2019



Abstract: Seafood-borne *Vibrio parahaemolyticus* illness is a global public health issue facing resource managers and the seafood industry. The recent increase in shellfish-borne illnesses in the Northeast United States has resulted in the application of intensive management practices based on a limited understanding of when and where risks are present. We aim to determine the contribution of factors that affect *V. parahaemolyticus* concentrations in oysters (*Crassostrea virginica*) using ten years of surveillance data for environmental and climate conditions in the Great Bay Estuary of New Hampshire from 2007 to 2016. A time series analysis was applied to analyze *V. parahaemolyticus* concentrations and local environmental predictors and develop predictive models. Whereas many environmental variables correlated with *V. parahaemolyticus* concentrations, only a few retained significance in capturing trends, seasonality and data variability. The optimal predictive model contained water temperature and pH, photoperiod, and the calendar day of study. The model enabled relatively accurate seasonality-based prediction of *V. parahaemolyticus* concentrations for 2014–2016 based on the 2007–2013 dataset and captured the increasing trend in extreme values of *V. parahaemolyticus* concentrations. The developed method enables the informative tracking of *V. parahaemolyticus* concentrations in coastal ecosystems and presents a useful platform for developing area-specific risk forecasting models.

Keywords: *Vibrio parahaemolyticus*; seasonality; seafood illness; forecasting; climate change

1. Introduction

Vibrio parahaemolyticus is the leading cause of seafood-borne gastroenteritis in the US and worldwide [1–3]. Most strains are believed to be non-pathogenic and the strains that do cause gastroenteritis and septicemia in humans have been historically associated with warm water environments [4–6]. Over the past decade, however, illnesses caused by *V. parahaemolyticus* have become more frequent in some cold and temperate water environments where illnesses were previously rare [7–14]. This new pattern of *V. parahaemolyticus* disease likely stems from a combination of observed trends, such as introduced and ecosystem establishment of pathogenic strains, increased summertime production and consumption of raw shellfish, and climate related changes causing warmer sea surface

temperatures and more variable salinities [7,8,13–19]. In the Northeast United States (US) where pathogenic *V. parahaemolyticus* is now established, foodborne illness is most frequently acquired from the consumption of raw or undercooked shellfish [3]. Post-harvest management has effectively reduced the incidence of *V. parahaemolyticus* disease outbreaks in this region. However, illness still occurs and achieving effective post-harvest control is both resource and time intensive. Effective pre-harvest *V. parahaemolyticus* forecasting tools would be valuable to shellfish growers and managers alike to make informed decisions about the *V. parahaemolyticus* risk conditions at the time of harvest and potentially reduce the risk and cost of *V. parahaemolyticus* management.

V. parahaemolyticus is a naturally occurring bacterial species that persists in a wide range of conditions in most marine and estuarine environments [5,20–30]. In multiple studies, temperature and salinity correlate most strongly with *V. parahaemolyticus*, but the strength of this relationship varies by region and season [31]. Similarly, nutrients, chlorophyll *a*, pH and turbidity were inconsistent and depended on the region and the variability of these factors. Therefore, region and even harvest area-specific studies are necessary to provide an accurate description of the influence of environmental conditions on *V. parahaemolyticus* concentration [32].

Long-term monitoring has been established in the Great Bay Estuary (GBE) by the Northeast Center for *Vibrio* Disease and Ecology at the University of New Hampshire (UNH) since 2007 [33–36]. The GBE is located on the border of New Hampshire and Maine (Figure 1) and has a long history of studies on pathogenic *Vibrio* spp. [37–39]. It is a regionally significant estuary that experiences wide-ranging environmental, climatic, and biological conditions [10], and thus serves as a useful model representative of regional estuaries. It is unique in that *V. parahaemolyticus* illnesses are still rare [40], although the *V. parahaemolyticus* population in the Northeast is evolving [13,14] and commercial shellfish harvests are rapidly increasing. The ongoing surveillance enables the development of pre-harvest risk-forecasting models.



Figure 1. Study area and sites for oyster and water sampling in the Great Bay Estuary, New Hampshire, USA. OR = Oyster River; NI = Nannie Island.

The goal of this study was to develop an integrated modeling approach to predict *V. parahaemolyticus* concentrations in shellfish at a pre-harvesting stage as a tool for managing this significant public health issue. We used data from 2007 to 2016 to capture long-term trends, seasonal fluctuations in a broad range of environmental and climatic predictors of *V. parahaemolyticus* dynamics, aiming to create a model development approach that could be transferable to other estuaries.

2. Materials and Methods

2.1. Study Sites, Environmental Sampling and Bacterial Analysis

The study area was the Great Bay estuary in New Hampshire. The two sampling locations (Figure 1) were near Nannie Island (NI) in Great Bay, where shellfish harvest classification is approved, and a site in the tidal portion of the Oyster River (OR), where harvesting is prohibited because of proximity to the Durham NH wastewater treatment facility. Both are locations of significant oyster (*Crassostrea virginica*) beds and long-term monitoring locations [36,38] and have different ecosystem and environmental conditions. The estuary has been monitored for over 30 consecutive years (March–December) through efforts by multiple agencies, including the Great Bay National Estuarine Research Reserve (GBNERR) and their System Wide Monitoring Program [41]. The average temperature, salinity, dissolved oxygen (DO), pH, and turbidity data were calculated from continuous (Q15) measurements obtained from the SWMP database for 2007–2016 for times simultaneous with and preceding oyster sampling in 12 h periods to account for ecological lag times and capture a more complete assessment of the potential environmental conditions that may have contributed *V. parahaemolyticus* concentrations observed at the time of collection. Monthly SWMP samples provided nutrient (total dissolved nitrogen (TDN)) and chlorophyll *a* (concentration by fluorescence; CHL) data for monitoring sites in close proximity to the NI and OR sampling locations. Meteorological data were acquired from several weather stations (SWMP; UNH) in the Great Bay region. Water quality parameters were measured in situ at the time of sampling using calibrated YSI 6600 and EXO multiprobe datasondes (Yellow Springs Instruments, Yellow Springs, OH, USA).

2.2. Oyster Sample Collection and Processing

Oyster samples were collected from the two oyster beds at NI and OR except during the period January–March from June 2007 through December 2016. For each sampling date, 10–12 oysters were cleaned and aseptically shucked into a sterile beaker (liquor and meat), weighed and diluted 1:1 with alkaline peptone water (APW (pH 8.6, 1% NaCl), and homogenized. A volume of 20 mL homogenate was further diluted in 80 mL APW for a starting dilution of 1:10. A volume of 1 mL of 1:10 solution was added to three tubes and then serially diluted with 1 mL aliquots into three serial dilutions containing 9 mL APW (pH 8.6, 1% NaCl). Each tube was incubated at 37 °C overnight (18–20 h) following the U.S. Food and Drug Administration Bacteriological Analytical Manual (BAM) [42].

Following incubation, turbid APW tubes were scored positive for growth. From 2007 to 2010, turbid tubes were streaked to Thioglycollate–Citrate–Bile–Salt (TCBS) agar (Beckton Dickson (BD), Franklin Lakes, NJ, USA) and incubated at 37 °C for 18–20 h. From 2011 to 2016, turbid tubes were streaked onto *Vibrio* CHROMAgar (CHROMagar, Paris, France) and incubated at 37 °C for 18–20 h. Sucrose negative (green) colonies from TCBS or purple colonies from CHROMagar were streaked onto tryptic soy agar (TSA; BD) and incubated at room temperature for 18–20 h. TSA isolates were inoculated in Heart Infusion (HI) broth for 18–20 h. Then, 1 mL HI aliquots were pelleted for 5 min at 8000 rpm, re-suspended in 1 mL molecular biology grade water (Phenix Research Products, Candler, NC, USA), boiled at 100 °C for 10 min and debris removed by centrifugation. Species identity of isolates was determined by polymerase chain reaction (PCR) performed using 2 µL cleared supernatant in 13 µL Mastermix, iQSupermix (BioRad, Hercules, CA, USA) using a BIO RAD T100 thermocycler and published primers and conditions [43] for 2007–2014, with slight modifications for 2015–2016 [19]. The PCR amplicons were visualized on 1.2% agarose gel with addition of Gel Red (Phenix Research

Products, Candler, NC, USA) under UV light. The concentration (Most Probable Number) was calculated from *V. parahaemolyticus* species-specific gene (tlh)-confirmed isolates from enrichment tubes and the BAM Most Probable Number (MPN) tables.

2.3. Statistical Analysis

All statistical computations were performed in the R Statistical Program and Environment, version 3.5.1 [44] with add-on packages MGCV [45]. Graphics were produced with ggplot2 [46]. Multiple steps of data analysis were performed to evaluate the relationship between environmental determinants, seasonality and *V. parahaemolyticus* concentrations in the GBE. MPN values for *V. parahaemolyticus* concentrations were log-transformed for analysis and model development to approximate normality and reduce skewness. Sampling events with missing environmental measurements ($n = 29$) when *V. parahaemolyticus* was not detected ($n = 71$) and one sample that exceeded >2 standard deviations were excluded from concentration model development. Statistical significance for all analysis in this study was determined using an alpha level of $p < 0.05$.

2.3.1. Model Development Strategy

All measurements were arranged in chronological order based on the date of measurement and multiple time series were compiled for the entire study period. The relationships between the time series for water quality variables, including water temperature, salinity, pH, DO, turbidity, CHL, TDN, rainfall and *V. parahaemolyticus* concentrations in oysters were evaluated using correlation and regression analysis. We used log-transformed values of *V. parahaemolyticus* concentrations (Y_t) and applied a Gaussian family distribution with an identity link function relating the expected value of response variable Y_t to selected predictors [47,48]. The transformation of water temperature, salinity, pH, DO, turbidity, CHL, TDN and rainfall was also explored as response variables in seasonality analysis and as predictor variables for *V. parahaemolyticus* in regression analysis with log or log + 1. We assessed the shape of relationships (linear and non-linear) between *V. parahaemolyticus* concentrations in oysters and environmental predictors. Variables that were significant in univariate regression were used to develop multiple regression models. We also assessed seasonality and trends over time and explored alternative variables representing seasonality with respect to their ability to improve the stability of forecasting. Assumptions of inter-correlation among predictors were evaluated using Spearman correlation analysis. Below we provide the detailed description of model building.

2.3.2. Seasonality and Trend Analysis

To explore the seasonality and the general trend throughout the whole study period (2007–2016) in all variables—*V. parahaemolyticus* concentrations, temperature, DO, salinity, pH, turbidity, CHL, TDN and rainfall—we developed two models with different ways of presenting the periodicity of seasonal oscillations. Model 1 contains variables for a linear trend and photoperiod. Model 2 uses terms for a linear trend and harmonic regression terms for the calendar day in the study as follows:

$$\text{Model1} : E(Y_t) = \beta_0 + \beta_1 t + \beta_p \text{Photoperiod}, \quad (1)$$

$$\text{Model2} : E(Y_t) = \beta_0 + \beta_1 t + \beta_s \sin(2\pi\omega t) + \beta_c \cos(2\pi\omega t). \quad (2)$$

In both models, Y_t is the daily time series for the outcome of interest, β_0 is the intercept, t is the daily time series, β_1 indicated a general trend in the outcome of interest, β_s and β_c are the coefficients of the harmonic terms and ω is the term representing the annual cycle (365.25 days, $\omega = 1 / 365.25$). The harmonic terms in Model 2 are expected to depict the periodic oscillation that can also be captured by the β_p in Model 1. The phase shift of periodic oscillations identified by Model 2 was determined as follows:

$$\psi = \arctan\left(\frac{\hat{\beta}_s}{\hat{\beta}_c}\right) + k. \quad (3)$$

When estimates of $\hat{\beta}_S$ and $\hat{\beta}_c$ were positive, $k = 0$. If $\hat{\beta}_S < 0$ and $\hat{\beta}_c > 0$, then $k = 2\pi$. If $\hat{\beta}_S$ and $\hat{\beta}_c$ were negative, or if $\hat{\beta}_S > 0$ and $\hat{\beta}_c < 0$, then $k = \pi$. The phase shift (ψ) was multiplied by 365.25 days in order to calculate peak timing. Covariance of β_s and β_c ($\sigma_{\beta_c\beta_s}$) and variance of β_s and β_c ($\hat{\beta}_s^2$ and $\hat{\beta}_c^2$) estimated the variance of the phase shift (ψ) as:

$$var(\psi) = \frac{(\sigma_{\beta_s\beta_c})^2 + (\sigma_{\beta_c\beta_s})^2 - (2\sigma_{\beta_c\beta_s}\beta_s\beta_c)}{(\hat{\beta}_s^2 + \hat{\beta}_c^2)(\hat{\beta}_s^2 + \hat{\beta}_c^2)}, \tag{4}$$

and confidence intervals of the peak timing were determined as: $1.96 \times \sqrt{Var(\psi) \times 365.25/2\pi}$. Secular trends were assessed using nine default thin-plate splines (f) from the MGCV package in R [46] in Model 3 and Model 4 as shown

$$\text{Model3 : } E(Y_t) = \beta_0 + \beta_1 f(t) + \beta_p \text{Photoperiod}, \tag{5}$$

$$\text{Model4 : } E(Y_t) = \beta_0 + \beta_1 f(t) + \beta_s \sin(2\pi\omega t) + \beta_c \cos(2\pi\omega t). \tag{6}$$

The models' performance was determined by the deviance explained, residual variation, Akaike's Information Criterion (AIC), and coefficient of determination (r^2) value. The trend term was determined to be non-linear based on visual assessment, positive ΔAIC and positive Δr^2 and $\Delta \text{Deviance} > 0.1$.

2.3.3. Extreme Value Trend Analysis

In addition to a general trend and Mann–Kendall trend analysis, we explored potential trends in high values of *V. parahaemolyticus* concentration as well as TDN, pH and salinity based on their importance in multiple regression models to estimate *V. parahaemolyticus* concentration by determining the number of events when the observations were above its 75th percentile. For other variables, trends were evaluated using the number of observations within the 25th and 75th percentile.

2.3.4. Variable Selection and Non-Linearity Assessment

To explore the relationship between the response variable, *V. parahaemolyticus* concentrations, and predictor variables, we incorporated each environmental parameter individually into linear (Model 5) and non-linear (Model 6) regression models. These two models were applied to the log-transformed values of *V. parahaemolyticus* concentrations (Y_t):

$$\text{Model5 : } E(Y_t) = \beta_0 + \beta_1 X_t, \tag{7}$$

where Y_t is the daily time series for the log-transformed *V. parahaemolyticus* concentrations in oyster, X_t is the daily time series for an environmental predictor and β_1 reflects the degree of captured linear relation in the daily time series of response and predictor variables.

Non-linear relationships were initially assessed using nine default thin-plate splines (f) from the MGCV package in R [46] as shown in

$$\text{Model6 : } E(Y_t) = \beta_0 + \beta_1 f(X_t). \tag{8}$$

The relationships between the environmental conditions and *V. parahaemolyticus* concentrations were overlaid with loess curves to visualize the relationship. Non-linear relationships were evaluated by the differences between the significance of the coefficient, residual variation, AIC, and coefficient of determination (r^2) value. Positive values indicate that the measure improved in Model 6 compared to Model 5 and negative values indicate a decrease in the model evaluation measurement. Variables were determined to be non-linear based on visual assessment, positive ΔAIC and positive Δr^2 and $\Delta \text{Deviance} > 0.1$. When strong non-linear non-monotonic relationships were detected, we re-parametrized the predictor by centering the variable around its *V. parahaemolyticus* concentration maximum and created

a new variable to provide biological interpretability to the model [49]. For example, a new variable for pH was created by squaring the difference between the observed pH values and the value of 7.8 selected for the centering. Re-parametrized variables are indicated as C-variable name (e.g., C-pH).

2.3.5. Model Building

The environmental parameters determined to be significant in univariate models (Models 5 and 6) were incorporated into a multivariate general linear regression model using Gaussian (GLM-G) and negative binomial (GLM-NB) distributional assumptions. For GLM-NB, the dispersion was determined by the index of dispersion: $\emptyset = \text{variance}/\text{mean} = 1$, where $\emptyset < 1$ refers to under-dispersion and $\emptyset > 1$ refers to over-dispersion. We started with the sequential model building (Model 7):

$$\text{Model7} : E(Y_t) = \beta_0 + \beta_1 X_{1,t} + \beta_k X_{k,t}, \quad (9)$$

where Y_t is the daily time series for the outcome of interest, β_0 is the intercept and t is the daily time series; $X_{1,t} \dots X_{k,t}$ are the daily time series for environmental predictors, including the reparametrized centered variables and interaction terms; $\beta_1 \dots \beta_k$ are the corresponding coefficients.

We then added variables to reflect the trend and seasonal oscillations and fine-tuned the model by using the photoperiod variable (Model 8), or harmonic terms (Model 9). In both models:

$$\text{Model8} : E(Y_t) = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \dots \beta_l t + \beta_p \text{Photoperiod}, \quad (10)$$

$$\text{Model9} : E(Y_t) = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \dots \beta_l t + \beta_s \sin(2\pi\omega t) + \beta_c \cos(2\pi\omega t), \quad (11)$$

where Y_t is the daily time series for the outcome of interest, β_0 is the intercept and t is the daily time series; $X_{1,t} \dots X_{k,t}$ are the daily time series for the selected environmental predictors, including the reparametrized centered variables and interaction terms; $\beta_1 \dots \beta_l$ are the corresponding coefficients. In Model 8 and 9, β_p is the coefficient of the photoperiod variable. In Model 10, β_s and β_c are the coefficients of the harmonic terms and ω is the term representing the annual cycle (365.25 days), as in Model 2.

For these hybrid models, we employed sequential model building using both Gaussian and negative binomial distributional assumptions in parallel and explored the contribution of interaction terms to the model's fit. Overall performance of GLMs was evaluated by evaluation of Akaike's Information Criterion (AIC) [50], residual variation, and deviance explained to determine the number and combination of variables that provided the strongest fit for the full time period of 2007–2016. Model fit was evaluated by the differences between the significance of the coefficient, residual variation, AIC, and coefficient of determination (r^2) value. Model selection was based on AIC value and improvement of r^2 and deviance explained >0.1 .

Using the parameters of the harmonic terms, e.g., the estimates of β_s and β_c regression coefficients and their error from Model 9, we applied the δ -method [51,52] to estimate seasonal peak timing along with its error term, expressed in days.

2.4. Assessment of Model Forecasting Ability

The predictive skill or forecasting ability of the selected versions of Models 7, 8 and 9 were evaluated by splitting the whole dataset into two datasets representing two periods: a training dataset from 2007 to 2013, and a test dataset from 2014 to 2016. Correlations between environmental variables and *V. parahaemolyticus* concentrations were compared for the full, training, and testing intervals. The forecasting ability and model performance were determined by coefficient of determination (r^2), and overall residual deviance. Forecasting error was evaluated by root mean square error (RMSE).

3. Results

3.1. *V. parahaemolyticus* Concentrations in the GBE, 2007–2016

V. parahaemolyticus was detected in 144 oyster samples during May through December from 27 June 2007 to 5 December 2016 for both the NI ($n = 77$) and OR ($n = 67$) study sites that included complete sets of data for environmental variables. There were no significant between-site differences for *V. parahaemolyticus* concentrations or measured condition parameters (data not shown) at the two sites, so all the following analyses use a combined-site database. Our analysis of the samples from 2007 to 2016 included the detection of *tdh* and *trh*, the traditional indicator markers for presence of ‘pathogenic’ *V. parahaemolyticus*. However, these markers were only detected in two samples during 2009 and were not detected again until 2015. Thus, our study focused on total *V. parahaemolyticus* concentrations because it is important to understand the ecosystem dynamics of the population of this species in shellfish harvest areas [36] as a proxy for risk assessment, and, in part because not all *V. parahaemolyticus* strains in clinical cases in the Northeast US and elsewhere contain either *tdh* or *trh*. Three main aspects of the full ten-year database are the marked seasonality, upward trend in high concentration values, and the wide variability/dispersion of *V. parahaemolyticus* concentrations between years and within each year (Figure 2). The observed *V. parahaemolyticus* concentrations were highly seasonal, ranging from 0.036 MPN/g oyster tissue during cold seasonal conditions to 4600 MPN/g during warm summertime conditions. The highest annual *V. parahaemolyticus* concentrations were higher during the later years than in early years. The detailed analysis of the trends and seasonality is presented below.

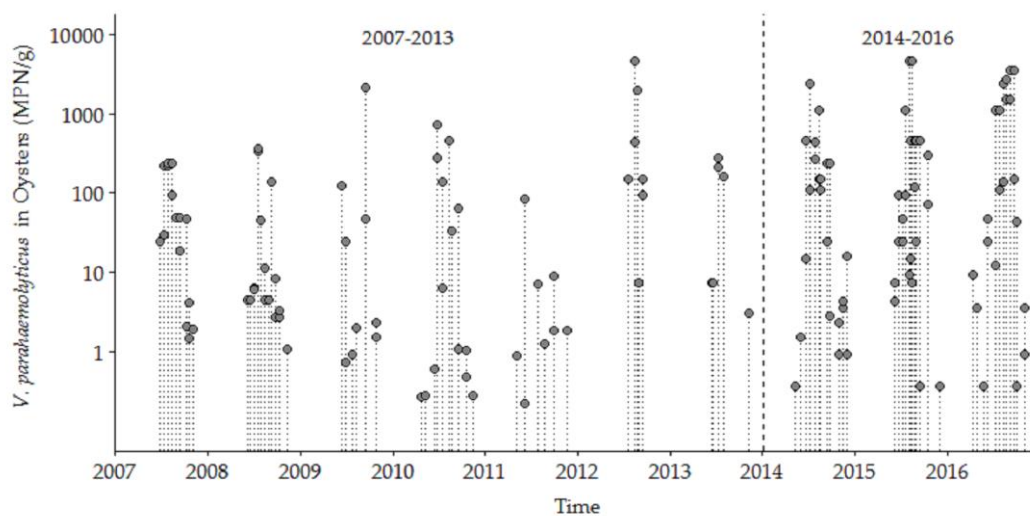


Figure 2. *Vibrio parahaemolyticus* concentrations in oysters from NI and OR at low tide in the Great Bay Estuary (GBE) in 2007–2016.

3.1.1. Trends and Seasonality

V. parahaemolyticus concentrations in oysters and environmental variables in the GBE were formally assessed for seasonality by using a photoperiod (Model 1) and a harmonic regression model (Model 2). These models allowed us to determine whether the study variables displayed re-occurring periodicity and a linear trend using calendar day of study to assess change over time. We examined trends, peak timing and seasonal oscillations in water temperature, DO, salinity, pH, turbidity, CHL, TDN and rainfall. The patterns of data had various shapes, including an extended period of *V. parahaemolyticus* detection during fall compared to spring. The variability in *V. parahaemolyticus* concentration in oysters, water temperature, DO and salinity were highly seasonal and well detected by both photoperiod and harmonic regression models, though the harmonic regression model provided a better fit in all instances (Table 1).

Table 1. Trend and seasonality estimates detected by Model 1 and Model 2 for *V. parahaemolyticus* concentrations and environmental variables (Model 1, top and Model 2, bottom).

Variable ^a	Coefficients ^b		Standard Error		<i>r</i> ²	Deviance	AIC	Peak Timing ^c
	Trend	Seasonality	Trend	Seasonality				
<i>Vp</i> (MPN/g)	0.0005 ***	0.57 ***	0.0001	0.11	0.19	0.21	673.4	222 ± 5
	0.0006 ***	-2.87 *** -3.66 ***	0.0001	0.34 0.33	0.50	0.51	597.4	
Water Temperature (°C)	<0.001	2.01 ***	<0.001	0.15	0.53	0.54	774.1	213 ± 2
	0.002 *	-5.81 *** -10.22 ***	<0.001	0.24 0.23	0.93	0.93	497.9	
Dissolved Oxygen (mg/L)	<0.001	-0.31 ***	<0.001	0.05	0.22	0.23	441.5	220 ± 6
	<0.001	1.45 *** 1.91 ***	<0.001	0.15 0.14	0.58	0.59	352.0	
Salinity (ppt)	0.001 ***	-0.19	0.0003	0.20	0.12	0.13	849.4	251 ± 18
	0.002 ***	-4.06 *** -1.77 **	0.0003	0.76 0.72	0.26	0.28	825.5	
pH	<0.001 ***	-0.02 *	<0.001	0.01	0.08	0.10	19.9	298 ± 98
	<0.001 ***	-0.06 0.03	0.006	0.05 0.05	0.09	0.11	20.9	
Turbidity (NTU)	-0.02 ***	3.93	0.007	4.10	0.06	0.09	1723.6	135 ± 111
	-0.02 ***	-6.34 -9.83	0.007	16.77 15.87	0.06	0.08	1716.5	
Chlorophyll- <i>a</i> (µg/L)	-0.0002	0.62 ***	0.005	0.0002	0.09	0.10	775.3	180 ± 37
	<0.001	0.11 -2.02 ***	<0.001	0.65 0.61	0.09	0.10	778.2	
Total Dissolved Nitrogen (mg/L)	<0.001 ***	-0.008 *	<0.001	0.005	0.15	0.16	-229.0	206 ± 45
	<0.001 ***	0.02 0.04 *	<0.001	0.02 0.02	0.15	0.17	-228.2	
Rainfall (mm)	<0.001	0.01 *	<0.001	<0.001	0.01	0.02	-76.7	209 ± 38
	<0.001	-0.03 -0.07 **	<0.001	0.001 <0.001	0.01	0.04	-74.6	

^a Variables are shown for Model 1, top row and Model 2, two bottom rows for sine and cosine terms; ^b the significance of coefficients is indicated as *** 0.001, ** 0.01, and * 0.1; ^c peak timing estimates are represented by the mean and standard error values; for two parameters, dissolved oxygen (DO) and total dissolved nitrogen (TDN), the estimates reflect the seasonal nadir. AIC, Akaike's Information Criterion.

The peak timing of *V. parahaemolyticus* (day 222 ± 5) was determined to be approximately 10 days after the peak timing of water temperature that occurred at day 213 ± 2. The peak timing of salinity and pH were within 25 days of the peak timing of *V. parahaemolyticus*, though as the strength of the seasonality of the variable decreased the confidence interval around the corresponding the peak timing was observed to increase. Neither model offered a fit to the variability observed in other variables. For instance, less than 4% of the variability in rainfall was attributed to seasonality. Rainfall and turbidity measurements above zero were episodic and model fit did not improve above 1% variance explained with log or log + 1 transformation. Figure 3 provides an explanation for the model fit by superimposing daily values for each year and depicting seasonal patterns for each variable.

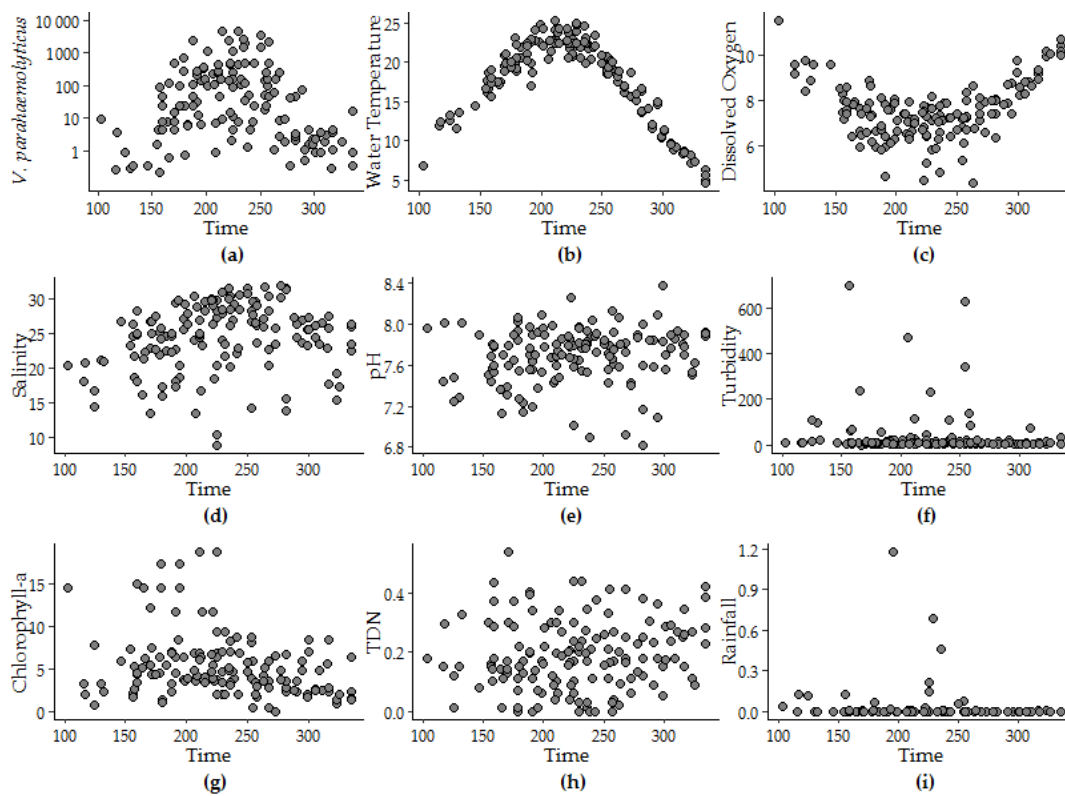


Figure 3. Patterns in (a) *V. parahaemolyticus* concentration, (b) water temperature, (c) dissolved oxygen, (d) salinity, (e) pH, (f) turbidity, (g) CHL, (h) TDN, and (i) rainfall versus the calendar day of the year superimposed from 2007 to 2016.

Over the ten-year period of surveillance there were significant increases in *V. parahaemolyticus* concentrations, salinity, pH and TDN. Only turbidity decreased during this same period (Table 1). To further explore these findings, we used thin-plate splines to assess secular trends (Table S1) and examined the trend in extreme values for salinity, pH, TDN, and *V. parahaemolyticus* concentrations that were above the 75% percentile (Table 2; Figure 4). The improvement in fit from the non-linear trend term in the photoperiod model (Model 1 and Model 3) was not seen in the harmonic regression model (Model 2 and Model 4). The change in fit can be largely attributed to interannual variation that was accounted for in the harmonic regression model and so a linear trend term was applied moving forward. For pH, the range of observed pH values decreased over time with more observations occurring within the pH range of 7.56 (25th percentile) to 7.88 (75th percentile). For salinity and pH, more than 55% and 62.1% of days observed in 2015 were above 27.0 ppt for salinity and within the 7.56–7.88 pH range, respectively. TDN above 0.27 mg/L was observed in six of ten years of the study. In later years, at least 45.0% percent of measured TDN was above the 75% percentile. Yearly *V. parahaemolyticus* concentrations above 240 MPN/g oyster tissue increased from 11.8% of samples in 2007 to 38.1% in 2016. Kendall-Mann trend analysis identified significant upward trends in extreme values for *V. parahaemolyticus*, salinity, pH, and TDN ($p < 0.05$).

Table 2. Trends of the frequency of days when *V. parahaemolyticus* concentrations, water temperature and salinity exceeded the 75th percentile of data and pH data were within the 25th to 75th percentile range in GBE during the period 2007 to 2016.

Year	<i>V. Parahaemolyticus</i>		Salinity		TDN		pH	
	75th Percentile				25th and 75th Percentile			
	220 MPN/g		27 ppt		0.27 mg/L		7.56–7.88	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
2007	2/17	11.8%	196/488	40.2%	6/17	35.3%	215/488	44.1%
2008	2/18	11.1%	10/465	2.2%	0/18	0.0%	148/465	31.8%
2009	1/11	9.1%	18/463	3.9%	1/11	9.0%	173/449	38.5%
2010	3/14	21.4%	58/451	12.9%	0/14	0.0%	157/451	34.8%
2011	0/9	0.0%	46/377	12.2%	0/9	0.0%	102/430	23.7%
2012	3/7	42.9%	135/475	28.4%	0/7	0.0%	217/447	48.5%
2013	1/6	16.7%	65/438	14.8%	3/6	50.0%	231/438	52.7%
2014	7/22	31.8%	135/432	31.3%	13/22	59.1%	277/432	64.1%
2015	8/24	33.3%	205/443	46.3%	10/22	45.5%	230/408	56.3%
2016	8/21	38.1%	266/479	55.5%	4/18	22.2%	289/465	62.1%

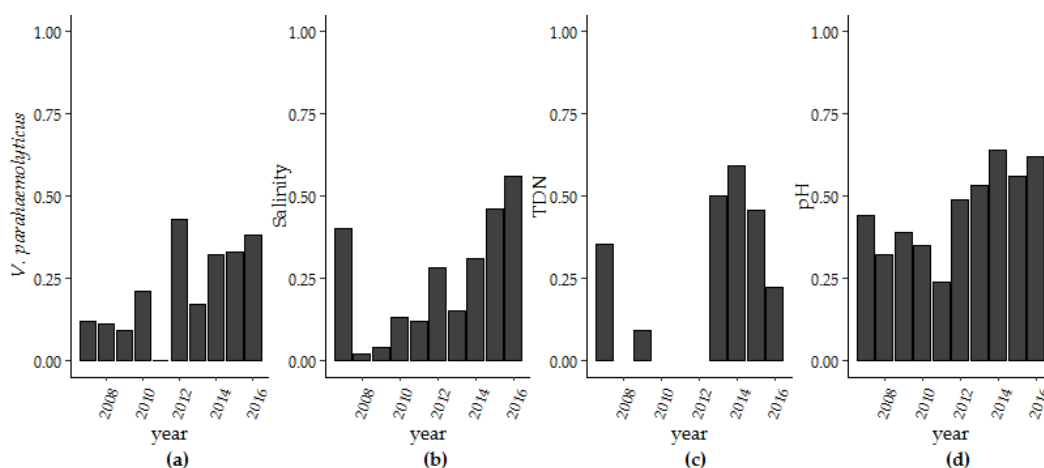


Figure 4. The number of observations per year above the 75th percentile for (a) *V. parahaemolyticus* concentrations, (b) salinity, (c) TDN and between the 25th and 75th percentile for (d) pH.

3.1.2. Univariate Regression

Individual linear and non-linear regression analyses conducted between *V. parahaemolyticus* concentration in oysters and eight measured variables from 2007 to 2016 identified water temperature, salinity, DO, pH, CHL and rainfall as significant model parameters in linear or non-linear regression. Model fit improved by less than 0.1 with log or log + 1 transformation of the independent variables. Water temperature accounted for the largest degree of *V. parahaemolyticus* variation (48.1%), DO accounted for 32.1%, followed by salinity (11.0%), pH (4.8%), CHL (2.8%) and rainfall (2.3%). The significance of pH increased in non-linear versus linear regression (Table 3), and the variability explained by pH also increased from 4.8% to 13.4%.

Table 3. The relationship between *V. parahaemolyticus* concentrations and environmental variables and fit improvement based on linear (Model 5) and non-linear (Model 6) regression models in GBE in 2007–2017. Positive values indicate that the measure improved in Model 6 compared to Model 5 and negative values indicate a decrease in the model evaluation measurement.

Variable	Model 5	Model 6	ΔModel 6–Model 5		
	<i>p</i> -Value	<i>p</i> -Value	Δ <i>r</i> ²	ΔDeviance	ΔAIC
Water Temperature (°C)	<0.001	<0.001	0.03	0.03	8.27
Dissolved Oxygen (mg/L)	<0.001	<0.001	0.04	0.05	7.28
Salinity (ppt)	<0.001	<0.001	−0.01	0.0	0.0
pH	0.009	0.002	0.14	0.08	8.48
Chlorophyll <i>a</i> (μg/L)	0.05	0.09	0.01	0.29	0.11
Rainfall (mm)	0.03	0.02	0.04	0.04	−6.31
Turbidity (NTU)	0.27	0.48	0.01	0.25	0.43
Total Dissolved Nitrogen (mg/L)	0.38	0.31	0.02	0.03	3.20

The form of the relationship between the environmental conditions and *V. parahaemolyticus* concentrations was further explored using loess smoothing to determine the parameters for each variable (Figure 5). The strength and significance of the linear response (Model 5) can be observed between *V. parahaemolyticus* and temperature, salinity and DO. Likewise, the non-linear relationship between pH and *V. parahaemolyticus*, identified by Model 6, is also highlighted by the loess smoothing. Though the fit between *V. parahaemolyticus* and CHL improved in Model 4 compared to Model 3, visual inspection of this relationship shows that this improvement can be attributed to rare events in the extremes of the observations.

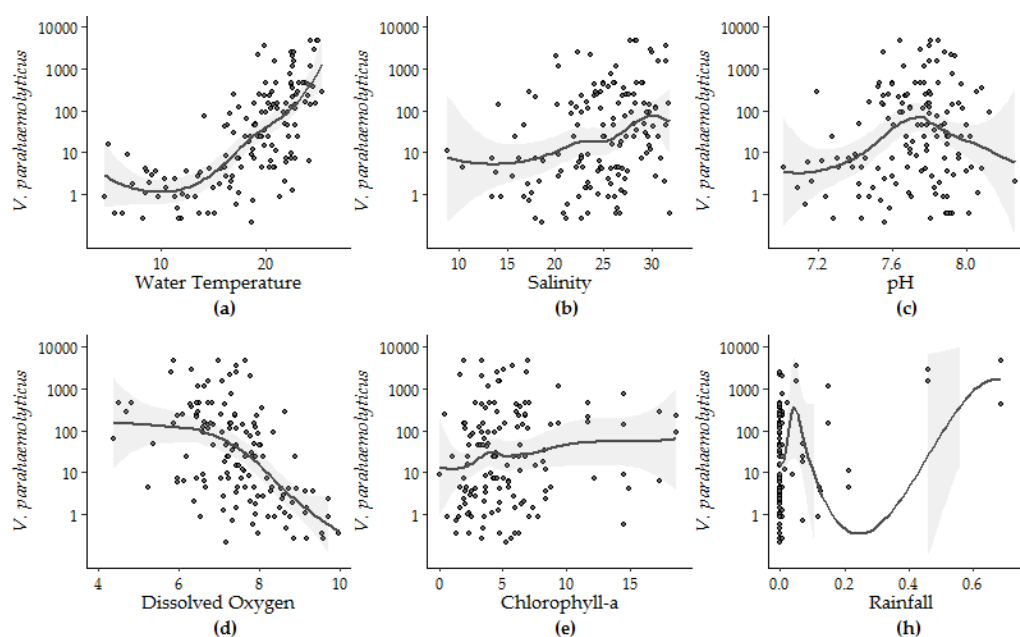


Figure 5. Loess smoothing applied to *V. parahaemolyticus* concentrations and (a) water temperature, (b) salinity, (c) pH, (d) DO—dissolved oxygen, (e) CHL—chlorophyll-*a*, and (h) rainfall.

The nonlinear regression between pH and *V. parahaemolyticus* was first improved with the addition of thin-plate splines. Based on the application of loess smoothing, pH was then re-parametrized as the square of the difference between the observed pH and 7.8, an apparent ecological optimum relative to observed *V. parahaemolyticus* concentrations in the study area. Re-parameterization of pH improved the percent variability explained, r^2 and p values from 4.8%, 0.04 and 0.008 (for the unmodified pH data) to 8.6%, 0.1 and 0.0003, respectively.

3.2. Sequential Model Building

A multiple regression model was next developed to determine a set of environmental variables that predict *V. parahaemolyticus* concentrations in oysters between 2007 and 2016 (Table 4). Water temperature was a foundational model variable for multiple regression model development and thus used in all multiple variable regression models. Single and multi-parameter models excluding water temperature explained less than the 48.1% of *V. parahaemolyticus* concentration variation explained by water temperature alone (data not shown). The addition of the trend term, photoperiod and harmonic regression variables to the environmental variables in negative binomial regression optimized model estimations (Figure 6, Table S2).

Table 4. The sequential building of multiple regression models for *V. parahaemolyticus* concentrations in oysters using Gaussian (GLM-G) and negative binomial (GLM-NB) models (Models 7, 8, 9).

Model Composition ^a	Coefficients	St. Error	Deviance	AIC	Coefficients	St. Error	Deviance	AIC
Model 7 GLM-G					GLM-NB			
1. Temperature	0.34 ***	0.03	0.54	586.9	0.34 ***	0.03	0.48	1533.4
Salinity	0.12 **	0.03			0.13 ***	0.03		
2. Temperature	0.37 ***	0.03	0.57	583.1	0.41 ***	0.03	0.51	1521.6
C-pH ^b	-4.73 ***	0.94			-5.52 ***	0.91		
3. Temperature	0.35 ***	0.03	0.59	572.5	0.34 ***	0.02	0.53	1518.3
C-pH	-3.93 ***	0.99			-4.38 ***	0.93		
Salinity	0.07 **	0.03			0.07 **	0.02		
4. Temperature	0.35 ***	0.02	0.61	567.8	0.34 ***	0.02	0.57	1507.1
C-pH	4.61	3.35			5.52 *	0.03		
Salinity	0.11 ***	0.04			0.10 **	2.97		
C-pH*Salinity	-0.41 ***	0.16			-0.53 **	0.14		
Model 8 GLM-G					GLM-NB			
1. Trend	0.0003 **	0.0001			0.0003 ***	0.0001		
Photoperiod	-0.35 **	0.11	0.62	564.4	-0.32 ***	0.09	0.58	1501.9
Temperature	0.46 ***	0.04			0.43 ***	0.03		
C-pH	-3.77 ***	0.95			-4.52 ***	0.86		
2. Trend	0.0002 **	0.001			0.0003 ***	0.0001		
Photoperiod	-0.32 **	0.11	0.62	565.9	-0.32 ***	0.13	0.58	1503.9
Temperature	0.44 ***	0.04			0.43 ***	0.04		
C-pH	-3.77 ***	0.99			-4.48 ***	0.89		
Salinity	0.02	0.06			-0.004	0.03		
Model 9 GLM-G					GLM-NB			
1. Trend	0.0003 **	0.0001			0.0003 ***	0.0001		
Sin(.)	0.07	0.69			-0.28	0.56		
Cos(.)	1.47	1.12	0.62	566.3	0.79	0.91	0.58	1504.2
Temperature	0.50 ***	0.11			0.41 ***	0.09		
C-pH	-3.78 ***	0.96			-4.49 ***	0.87		
2. Trend	0.0003 **	0.0001			0.0003 ***	0.0001		
Sin(.)	0.15	0.69			-0.31	0.57		
Cos(.)	1.47	1.12	0.62	567.7	0.77	0.91	0.58	1506.2
Temp	0.49 ***	0.11			0.41 ***	0.09		
C-pH	-3.60 ***	0.99			-4.61 ***	0.89		
Salinity	0.03	0.04			-0.006	0.03		

^a The significance of coefficients is indicated as *** 0.001, ** 0.01, and * 0.1; ^b C-pH data were treated as reparametrized C-pH variables.

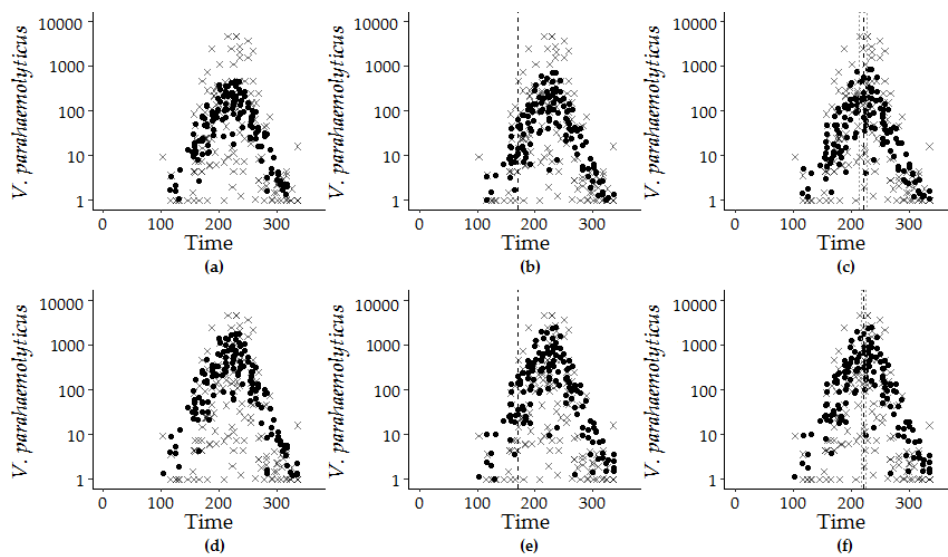


Figure 6. Model estimations (filled circle) and observed *V. parahaemolyticus* concentrations (x) are superimposed by the calendar day of the year from 2007 to 2016: GLM-G for (a) Model 7.4, (b) Model 8.1, and (c) Model 9.1 and GLM-NB for (d) Model 7.4, (e) Model 8.1, and (f) Model 9.1. The dashed vertical line at day 170 for the hybrid model (b,e) marks the longest day of the year, and the dashed lines at day 222 ± 5 days and at day 221 ± 7 days indicate the calculated peak timing of *V. parahaemolyticus* concentration for Model 7.1 for (c) GLM-G and (f) GLM-NB versions.

Spearman rank correlation analysis of the individual intervals indicates that photoperiod, water temperature, DO, pH and salinity were significantly correlated with *V. parahaemolyticus* concentrations in all time intervals, though the correlation between pH and *V. parahaemolyticus* varied between intervals (Figure 7). Inter-variable correlations were observed between water temperature and DO ($R = -0.69, p < 0.0001$), salinity ($R = 0.20, p = 0.014$), CHL ($R = 0.21, p \leq 0.0001$) and rainfall ($R = 0.21, p = 0.015$). pH and DO ($R = 0.43, p < 0.0001$) and pH and salinity ($R = 0.37, p < 0.0001$) were also correlated. Rainfall was only significant in the test dataset, while CHL was significant in the entire and training dataset (2007–2013) intervals but not significant in the test dataset (2014–2016) interval. Significant associations were observed between photoperiod and *V. parahaemolyticus* concentrations, water temperature, DO, salinity, CHL and pH ($R = 0.28, p \leq 0.001$; $R = 0.5, p \leq 0.001$; $R = -0.36, p \leq 0.001, R = 0.18, p = 0.04$; $R = 0.33, p = 0.001$; $R = -0.20, p = 0.02$ respectively) (Figure 7). Of the nine variables considered, only turbidity, TDN and rainfall were not correlated to *V. parahaemolyticus* concentrations or photoperiod.



Figure 7. Spearman correlation analysis of *V. parahaemolyticus* concentrations and environmental variables for three intervals: (a) 2007–2016, (b) 2007–2013 and (c) 2014–2016. Red indicates positive and blue negative correlations and the degree of significance is highlighted by color intensity.

3.3. Model Performance Prediction

The hybrid model (Model 9.1) provided the best overall fit for each dataset time interval, with consistently lower RMSE and higher r^2 values compared to the harmonic regression (Model 10.1) and environmental model (Model 7.4) (Table 5). The fits for all three models were relatively consistent even though the significance of some variables changed between time intervals. Although the estimations of precision for the harmonic regression model across training/test datasets were slightly lower than for other models, it is advantageous because important attributes of the data can be identified. For example, the *V. parahaemolyticus* concentrations peaked on 222 ± 5 day of the 365.25-day period for all three intervals. Similarly, the peak timing of water temperature and salinity were stable between the overall, training and test datasets (212 ± 2 day and 251 ± 18 day, respectively).

Table 5. The performance of three selected models: environmental model (Model 7.4), hybrid model (Model 8.1), and harmonic regression model (Model 9.1) for three time periods: full (P1), training (P2), and testing (P3) intervals.

Model	Variable ^a	Time Interval		
		P1	P2	P3
Model 7.4	Coefficient: Temperature	0.34 ***	0.37 ***	0.31 ***
	Salinity	0.10 ***	0.08 **	0.24 **
	C-pH	5.51 *	5.12	266.01 ***
	Salinity*C-pH	−0.53 ***	−0.53 ***	−11.01 ***
	r^2	0.54	0.58	0.57
	Deviance	0.57	0.58	0.54
	RMSE	1.91	1.79	1.96
Model 8.1	Coefficient: Trend	0.0003 ***	0.0003	0.0007
	Photoperiod	−0.31 ***	−0.28 **	−0.48 **
	Temperature	0.43 ***	0.45 ***	0.44 ***
	C-pH	−4.51 ***	−4.32 ***	−5.10
	r^2	0.61	0.57	0.61
	Deviance	0.58	0.59	0.53
	RMSE	1.85	1.81	1.92
Model 9.1	Coefficient: Trend	0.0004 ***	0.0004 *	0.0008
	Sin(.)	−0.41	−1.88 *	1.72 *
	Cos(.)	0.63	−1.54	4.66 **
	Temperature	0.40 ***	0.29 **	0.74 ***
	C-pH	−4.30 ***	−4.20 ***	1.60
	r^2	0.61	0.55	0.63
	Deviance	0.58	0.60	0.54
RMSE	1.81	1.82	1.83	

^a The significance of coefficients is indicated as *** 0.001, ** 0.01, and * 0.1.

The environmental (Model 7.4), hybrid (Model 8.1), and harmonic regression (Model 9.1) models developed with the observations from the training dataset accurately predict the overall trend, seasonality, and dispersion of the test dataset (Figure 8). The overall fits of all models were high, and the RMSE values increased in the short test time period. The hybrid model performed equally well in describing *V. parahaemolyticus* concentrations in the training dataset and predicting *V. parahaemolyticus* concentrations in the test dataset. This model contains a minimum number of environmental variables, photoperiod and calendar day of the study, and provides a good fit for capturing and predicting the seasonality, trend and dispersion of *V. parahaemolyticus* concentration during the study period.

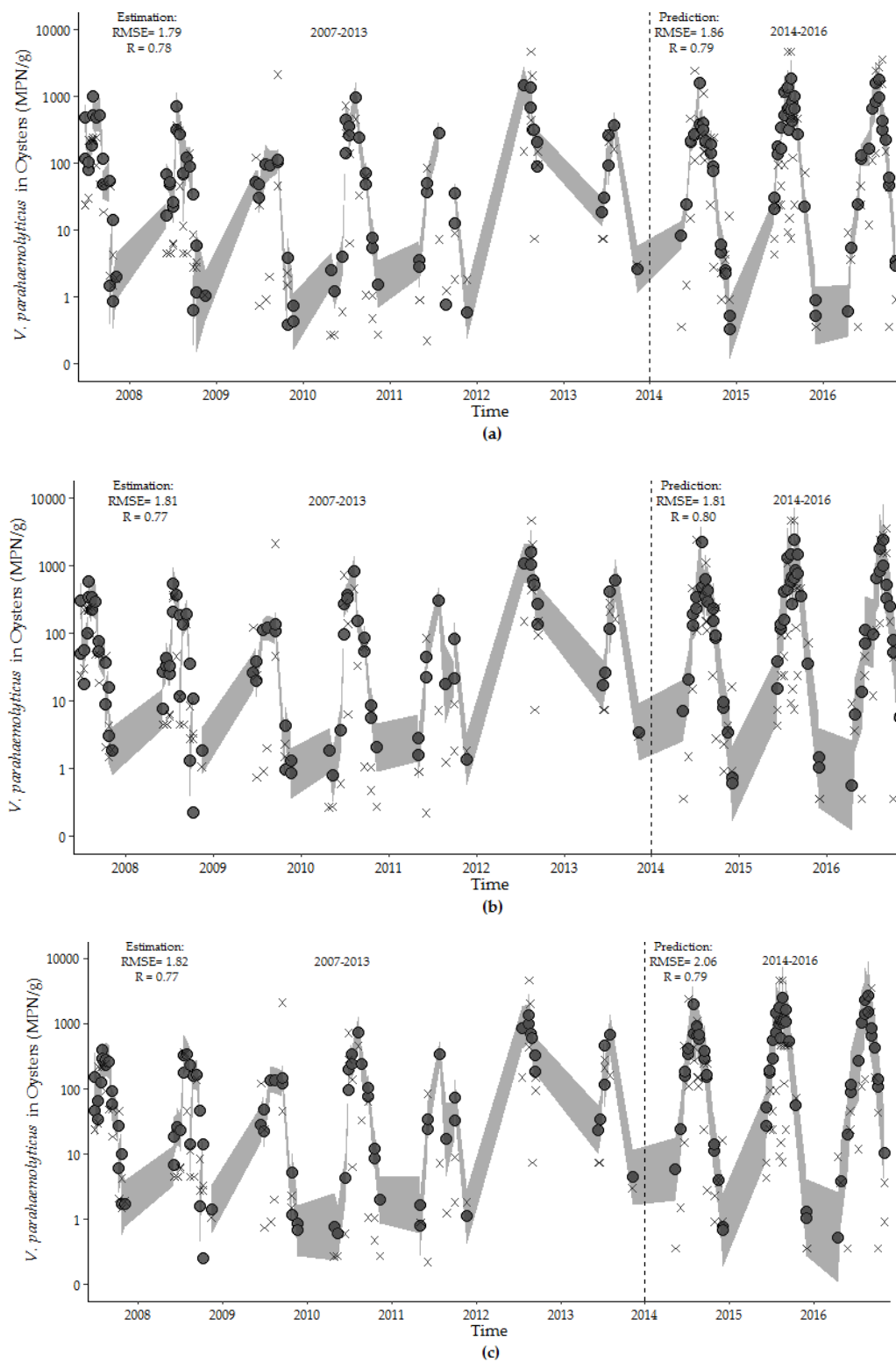


Figure 8. Estimates of *V. parahaemolyticus* concentrations (closed circle) with observed *V. parahaemolyticus* concentrations for: (a) environmental model, (b) hybrid model, and (c) harmonic regression model for the training (2007–2013) and test (2014–2016) periods. The 95th percentile prediction interval is represented by the gray shading. Model fit values are shown in the upper left corner of each figure.

4. Discussion

The intrinsic link that *V. parahaemolyticus* has with coastal ecosystems has been well studied and characterized. Previous studies have provided many useful site- and time-specific descriptive models for describing *V. parahaemolyticus* concentration dynamics. However, few of them have been evaluated for their ability to forecast *V. parahaemolyticus* dynamics, or to be generalizable and transferable to other geographic areas or time periods. A wide range of environmental conditions and ecological interactions have been reported to influence, or at least correlate with, *V. parahaemolyticus* concentrations including water temperature, salinity, inorganic and organic nutrients, suspended solids-turbidity, chlorophyll-*a* and plankton levels, light availability, and meteorological conditions [4,5,16,17,27,29,36,38,52–62]. The temporal and spatial data analysis methods vary greatly in these studies, from simple correlation to more complex models [31]. These have often included the application of multiple regression analysis to characterize and model the interactions between multiple environmental parameters and *V. parahaemolyticus* levels [5,18,28,36,63,64], even though they have not been useful for forecasting *V. parahaemolyticus* dynamics and risk conditions. Based on clearly observable aspects of the *V. parahaemolyticus* concentration data for this study and some initial analyses, the combination of models applied here incorporate seasonality, trend and dispersion concepts to characterize *V. parahaemolyticus* dynamics and accurately predict *V. parahaemolyticus* concentrations. Model accuracy is in part a function of using variables that are known and consistent such as photoperiod or day of the year that are ecologically interpretable, but stable for effective *V. parahaemolyticus* forecasting. This approach of seasonality and trend analysis has the potential to be transferable for developing similar forecasting models patterns of *V. parahaemolyticus* dynamics in other locations.

V. parahaemolyticus concentrations in the GBE during this study followed the same pattern each year as concentrations increased rapidly each springtime as water temperatures increased, and after peak concentrations during the warmest summer conditions, decreased as water temperatures decreased in the fall each year. Such seasonality, where regular and predictable changes in environmental and climatic conditions re-occur every calendar year, tends to become more pronounced with increasing distance from the equator and is largely due to extreme temperature variation driven by variable photoperiod [65]. Water temperature accounted for approximately 48.1% of the variation observed in *V. parahaemolyticus* concentrations in this study, similar to what has been observed globally and especially in highly seasonal, temperate water regions [27,28,60]. Thus, seasonality is a significant aspect of *V. parahaemolyticus* concentration dynamics in temperate coastal areas like New Hampshire and the Northeast US.

Photoperiod and harmonic regression models along with correlation analysis showed that *V. parahaemolyticus* concentration, water temperature, dissolved oxygen, pH, salinity and chlorophyll-*a* are significantly related to variables that mirror seasonal patterns in the GBE. Likewise, these variables accurately estimate *V. parahaemolyticus* concentrations in oysters. The synchronized seasonal periodic oscillation is one probable explanation for why regression modeling favors water temperature as the most significant model parameter. A complex combination of biological and physical environmental variables certainly drives *V. parahaemolyticus* population dynamics. However, many of these environmental variables are, in turn, driven mainly by seasonal temperature. Therefore, the variability they contribute to *V. parahaemolyticus* concentrations is not significantly different than what is provided by water temperature. For example, dissolved oxygen was negatively correlated with *V. parahaemolyticus* concentrations, similar to what has been previously reported [20] and was the second strongest variable, estimating over 32% of the variability in *V. parahaemolyticus* over the course of the study in the GBE. Since *V. parahaemolyticus* is a facultative anaerobe, this finding has the potential to elucidate important biological dimensions of the ecology of *V. parahaemolyticus*. Water temperature is a dominant driver of dissolved oxygen concentrations so collinearity between these variables is likely. In addition, because of the constraints of mathematical modeling, well-fit models are not necessarily mechanistically or ecologically descriptive [66] and, in this case, dissolved oxygen

was omitted from model development to avoid multicollinearity in favor of water temperature as a stronger model variable.

Salinity and water temperature are both seasonally variable parameters that, together are the most commonly cited environmental drivers of *V. parahaemolyticus* concentration variation [16,31,67]. Salinity was a significant predictive parameter for *V. parahaemolyticus* concentration in this study, though the significance of salinity was dependent on the time interval (2007–2013 versus 2014–2016) of the data and the trend adjustment in the model (Table S2). This type of variability has also been observed in risk assessment [68,69] and in previous studies where salinity sometimes shows a strong positive correlation with *V. parahaemolyticus* [5,6,25,63], whereas for others [28,53,60,70], salinity and *V. parahaemolyticus* dynamics do not correlate. Thus, the finding that salinity and other variables reported to be significant in other *V. parahaemolyticus* concentration models were not included in this study's final model may be, at least in part, a function of both the specific conditions at this study site and time period and a function of the in-depth statistical approach used.

Though most studies find little to no correlation between pH and *V. parahaemolyticus* concentration [20,21,28], non-linear regression and correlation analysis identified pH as an important parameter for the predictive models in the GBE. Loess smoothing highlighted the marked non-linearity of the relationship between pH and *V. parahaemolyticus* concentrations and suggested a biological optimum/optimal range for pH where *V. parahaemolyticus* concentrations decreased as pH increased or decreased relative to pH 7.8. For the purposes of optimal model development, a new pH variable was constructed by reparametrizing the measurements to create a linear response in *V. parahaemolyticus* as pH measurements moved from the optima of 7.8. An optimal pH of 7.8 is near the pH (8.5) of alkaline peptone water medium used to optimally enrich for *Vibrio* species [42] and has also been suggested as an optimal pH by laboratory-based observations [71]. Wong et al., (2015) [72] found that exposure to more acidic environments tended to reduce cell density and cause stress responses in *V. parahaemolyticus*. In this study, we observe that pH measurements in the GBE appeared to become less variable and more basic in recent years, which was also reported by Lopez-Hernandez et al., (2015) [5]. Thus, going beyond simple linear regression and including the use of non-linear analysis reveals pH as an important and ecologically linked variable to explain *V. parahaemolyticus* population dynamics.

In other studies [36,60,63], variables other than salinity and pH were significant for estimating *V. parahaemolyticus* concentrations in univariate regression. However, in this study, they provided an insignificant amount of improvement to a multiple regression model that included water temperature. For example, chlorophyll-*a*, considered a proxy measurement for phytoplankton abundance [21,31], was significantly related to *V. parahaemolyticus* concentrations in correlation and univariate regression, but it was not significant in a multiple regression model that included water temperature. Chlorophyll-*a* was thus omitted from further model development because it did not contribute additional information in describing *V. parahaemolyticus* variation. Many studies have suggested an important ecological interaction between *V. parahaemolyticus* and plankton [16,27,57,63,69,73,74], and though chlorophyll-*a* was not included in the multiple regression models, we have also conducted a parallel study to explore the relationship between *V. parahaemolyticus* and plankton species across several years in the GBE [75,76] to determine covarying plankton species. These have included phytoplankton that have been reported to be significantly associated with *V. parahaemolyticus* elsewhere [77,78] that could provide more in-depth insight into the importance of phytoplankton and the proxy chlorophyll-*a* to the *V. parahaemolyticus* concentration dynamics observed in the GBE.

Approximately half of the variability of *V. parahaemolyticus* in the GBE could be predicted using the contribution of photoperiod (in hours), sine and cosine of the day of the study in harmonic regression, and the day of the study. Even though the model consisting solely of environmental variables was potentially more ecologically informative, the trend and seasonality variables of calendar day of the study, photoperiod, sine and cosine were more stable to estimate and predict the patterns of seasonality and trend of increasingly high concentrations over time in *V. parahaemolyticus* than salinity and to a lesser degree pH and do not require in situ measurements. Additionally, evaluation of the environmental

model for its forecasting ability highlighted that some evaluation measures were discordant, while the harmonic regression and photoperiod model goodness-of-fit and forecasting error were in agreement. This highlights that though multiple evaluation measures can cause complexities in model selection, in this study the model with conflicting evaluation measures may indicate underlying issues, while the models where evaluation measurements were in agreement provided stronger prediction accuracy. Harmonic regression analyses also lead to identification of the day of year for peak *V. parahaemolyticus* concentration that occurs in mid-August (day 222 ± 5 days) that followed the peak timing of water temperature (213 ± 2), while the longest day of the year is 21 June (day 170). This highlights a loading, or hysteresis in the system and provides the basis for understanding the 'fall shoulder' of elevated concentrations of *V. parahaemolyticus* that extend into the late September.

Peak timing was used to assess each environmental variable individually to detect how environmental variables may contribute to the development of ideal conditions for *V. parahaemolyticus*. Data in this study were collected either monthly or biweekly, while *V. parahaemolyticus* replicates every eight minutes under ideal conditions. In this instance, the accurate detection of lagged effects on *V. parahaemolyticus* would require more frequent sampling and fine-scale temporal resolution. Due to this level of biological complexity and the irregular temporal intervals of the data in our study, the mean from 12 h preceding collection was used for regression with environmental variables and peak timing was used to assess temporally how each environmental variable may contribute to the development of ideal conditions for *V. parahaemolyticus*. Using this approach, we determined that significant predictive variables peak in advance of *V. parahaemolyticus* potentially contributing to a hysteresis or loading of the systems, setting up conditions that are optimal for *V. parahaemolyticus*. Davis et al., 2019 [79], recently reported that environmental variables approximately one month preceding collection were significant to predicting *V. parahaemolyticus* concentrations in the Chesapeake Bay suggesting they might also be observing this type of lagged effect from a loading of the system. The application of harmonic regression and peak timing here demonstrates how biological complexities and limitations of sampling frequency can be overcome while also providing the resolution to detect temporal patterns between dependent and independent variables. The determination of peak timing is also a potentially important tool for forecasting the commonly observed mid-summer peaks in illnesses in the Northeast US [80].

A major characteristic of the *V. parahaemolyticus* concentration data is their wide dispersion. The comparison between Gaussian and negative binomial GLMs determined that the dispersion of *V. parahaemolyticus* concentrations, especially the extreme high concentrations, was best fit by the negative binomial model, as it can better account for the wide range of *V. parahaemolyticus* concentrations (0.3 to 4600 MPN/g) observed annually in the GBE. Effective risk models, with negative binomial regression as an essential model attribute, developed to predict the increasing and more dispersed *V. parahaemolyticus* concentrations will become more important as global warming and other climate and ecosystem changes will probably cause increased concentrations and persistence of *V. parahaemolyticus* in temperate coastal areas [8,81–83] with a likely increase in public health risks.

Model evaluation, estimations, and predictions illustrate how each model provides fit and prediction ability of the variability in *V. parahaemolyticus* concentration observed over the course of the study. Though a forecasting model consisting of environmental variables could be more appealing because of its ecological interpretability, there are potential limitations to models that rely solely on environmental predictors. For example, it is unlikely that a well-fit model can contain all the environmental variables that effect *V. parahaemolyticus* given its ecological complexity and the collinearity between seasonal-driven variables that relate to *V. parahaemolyticus* dynamics. Further, the strength of environmental variables to predict *V. parahaemolyticus* over time can change, as was observed in the interaction between pH and salinity between time intervals. Additionally, salinity became insignificant when the model was adjusted for a linear trend. The negative binomial harmonic regression and hybrid models fit the seasonality and trend features, and account well for the dispersion of *V. parahaemolyticus*. All models demonstrated good forecasting ability. Importantly, these models

also enabled the determination of key characteristics of *V. parahaemolyticus* in the GBE including peak timing and a seasonal loading contributing to prolonged elevated concentrations that last into fall months. The hybrid model provides the optimal level of ecological interpretability a reasonable ability to capture the dynamics of *V. parahaemolyticus* concentrations in oysters in the GBE, and a stable platform for forecasting *V. parahaemolyticus* concentrations in coming seasons. Thus, the use of both significant environmental variables and stable parameters in the hybrid negative binomial regression model lead to successful forecasting model development that captures seasonality, temporal trends, and the high degree of data variability and dispersion.

The increased incidence of illnesses caused by *V. parahaemolyticus* infections in the Northeast US has co-occurred with increases in regional surface water temperatures and other environmental parameters, as shown in this study, suggesting an increase in the presence of pathogenic *V. parahaemolyticus* strains and/or population evolution [13,14]. The model approach developed in this study illustrates how characteristics of *V. parahaemolyticus* dynamics can be captured as environmental conditions continue to become more favorable for the pathogen to enable accurate prediction of public health risk to shellfish consumers and recreational users of coastal waters. This information, coupled with recent advances [13,14,19] that improve detection methods for endemic and invasive pathogenic *V. parahaemolyticus* sequence types (ST) in the Northeast, could be useful for shellfish harvest management in the Northeast US based on this new improved and integrated capacity to forecast concentration dynamics of both total and pathogenic *V. parahaemolyticus* populations and potential disease outbreak risks. The developed modeling approach also has the potential to inform more in-depth mechanistic studies in order to gain a better understanding of the ecology of *V. parahaemolyticus* and other water-borne pathogens.

5. Conclusions

This study suggests that transferable models can be developed for forecasting public health risks related to *V. parahaemolyticus* concentrations in shellfish. Ecological monitoring data and statistical modeling are necessary to effectively characterize relationships between ecological variables and *V. parahaemolyticus* concentrations. From among many ecological variables, easy to measure water temperature and pH were all that was required when combined with seasonality and trend variables within hybrid statistical models to capture both long-term increasing trends for *V. parahaemolyticus* concentrations and to provide capacity for forecasting *V. parahaemolyticus* concentrations. The determination of peak timing is useful for assessing how each environmental variable may contribute to the development of optimal conditions for *V. parahaemolyticus*. This approach may be best applied in temperate, seasonally driven regions like the Northeast, US, as it relies on characteristics of *V. parahaemolyticus* ecology that are shared by most temperate regions.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1660-4601/16/22/4341/s1>: Table S1: Trend and seasonality estimates detected by Model 3 and Model 4. Table S2: Stepwise Development of Models 6 and 7.

Author Contributions: Conceptualization, M.A.H., E.A.U., C.A.W., V.S.C., E.N.N. and S.H.J.; methodology, M.A.H., E.A.U., C.A.W., V.S.C., E.N.N. and S.H.J.; software, M.A.H. and E.A.U.; validation, C.A.W., V.S.C., E.N.N. and S.H.J.; formal analysis, M.A.H. and E.A.U.; resources, C.A.W., V.S.C. and S.H.J.; data curation, M.A.H. and E.A.U.; writing—original M.A.H. and E.A.U.; writing—review and editing, C.A.W., V.S.C., E.N.N. and S.H.J.; visualization, M.A.H. and E.A.U.; supervision, S.J.H.; funding acquisition, C.A.W., V.S.C. and S.H.J.

Funding: The authors gratefully acknowledge partial funding support from the National Science Foundation EPSCoR IIA-1330641, USDA National Institute of Food and Agriculture Hatch NH00574, NH00609 (accession 233555), and NH00625 (accession 1004199), and the National Oceanic and Atmospheric Administration College Sea Grant program and New Hampshire Sea Grant program grants R/CE-137, R/SSS-2, R/HCE-3, and from NSF IRES Track I: Collaborative Research: U.S.-Indonesian Research Experience for Students on Sustainable Adaptation of Coastal Areas to Environmental Change (award #1826939).

Acknowledgments: The authors would like to thank Jennifer Mahoney, Meg Striplin, Brian Schuster, Crystal Ellis, Jong Yu, Eliot Jones, Michael Taylor, Ashley Marcinkiewicz, Feng Xu, Tom Gregory, Chris Peters, Jackie Lemaire, Audrey Berenson, Sarah Richards, Emily Schulz, and Elizabeth Deyett for their help with sampling, sample

processing, detection analysis, and database management. Also, we thank Iago Hale, Alexandra Kulinkina and Tania M. Alarcon Falconi for support with implementing harmonic regression analysis, peak timing calculations in the R environment, and other statistical approaches.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. CDC. *Foodborne Diseases Active Surveillance Network (FoodNet): FoodNet 2015 Surveillance Report (Final Data)*; U.S. Department of Health and Human Services, CDC: Atlanta, GA, USA, 2017.
2. Nilsson, W.B.; Paranjpye, R.N.; Hamel, O.S.; Hard, C.; Strom, M.S. *Vibrio parahaemolyticus* risk assessment in the Pacific Northwest: It's not what's in the water. *FEMS Microbiol. Ecol.* **2019**, *95*, fiz027. [[CrossRef](#)] [[PubMed](#)]
3. Newton, A.; Garrett, N.; Stroika, S.; Halpin, J.; Turnsek, M.; Mody, R. Notes from the field: Increase in *Vibrio parahaemolyticus* infections associated with consumption of Atlantic coast shellfish—2013. *Morb. Mortal. Wkly. Rep. (MMWR)* **2014**, *63*, 335–336. [[PubMed](#)]
4. Ceccarelli, D.; Hasan, N.A.; Huq, A.; Colwell, R.R. Distribution and dynamics of epidemic and pandemic *Vibrio parahaemolyticus* virulence factors. *Front. Cell. Infect. Microbiol.* **2013**, *3*, 97. [[CrossRef](#)] [[PubMed](#)]
5. López-Hernández, K.M.; Pardío-Sedas, V.T.; Lizárraga-Partida, L.; Williams, J.D.J.; Martínez-Herrera, D.; Flores-Primo, A.; Uscanga-Serrano, R.; Rendón-Castro, K. Environmental parameters influence on the dynamics of total and pathogenic *Vibrio parahaemolyticus* densities in *Crassostrea virginica* harvested from Mexico's Gulf coast. *Mar. Pollut. Bull.* **2015**, *91*, 317–329. [[CrossRef](#)] [[PubMed](#)]
6. Vezzulli, L.; Colwell, R.R.; Pruzzo, C. Ocean warming and spread of pathogenic Vibrios in the aquatic environment. *Microb. Ecol.* **2013**, *65*, 817–825. [[CrossRef](#)]
7. Baker-Austin, C.; Trinanés, J.A.; Taylor, N.G.; Hartnell, R.; Siitonen, A.; Martínez-Urtaza, J. Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat. Clim. Chang.* **2013**, *3*, 73. [[CrossRef](#)]
8. Baker-Austin, C.; Trinanés, J.; González-Escalona, N.; Martínez-Urtaza, J. Non-cholera vibrios: The microbial barometer of climate change. *Trends Microbiol.* **2017**, *25*, 76–84. [[CrossRef](#)]
9. Centers for Disease Control and Prevention (CDC). Increase in *Vibrio parahaemolyticus* Illnesses Associated with Consumption of Shellfish from Several Atlantic Coast Harvest Areas, United States, 2013. Available online: <https://www.cdc.gov/vibrio/investigations/> (accessed on 1 November 2019).
10. Jones, S.H. Microbial Pathogens and Biotoxins: State of the Gulf of Maine Report. Gulf of Maine Council on the Marine Environment. Available online: <http://www.gulfofmaine.org/2/sogom-homepage/> (accessed on 1 November 2019).
11. Martínez-Urtaza, J.; Baker-Austin, C.; Jones, J.; Newton, A.; DePaola, A. Spread of Pacific Northwest *Vibrio parahaemolyticus* strain. *N. Engl. J. Med.* **2013**, *369*, 1573–1574. [[CrossRef](#)]
12. 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. Outbreak of *Vibrio parahaemolyticus* gastroenteritis associated with Alaskan oysters. *N. Engl. J. Med.* **2005**, *353*, 1463–1470. [[CrossRef](#)]
13. Xu, F.; Ilyas, S.; Hall, J.A.; Jones, S.H.; Cooper, V.S.; Whistler, C.A. Genetic characterization of clinical and environmental *Vibrio parahaemolyticus* from the Northeast USA reveals emerging resident and non-indigenous pathogen lineages. *Front. Microbiol.* **2015**, *6*, 272. [[CrossRef](#)]
14. Xu, F.; González-Escalona, N.; Drees, K.P.; Sebra, R.P.; Cooper, V.S.; Jones, S.H.; Whistler, C.A. Parallel evolution of two clades of an Atlantic-endemic pathogenic lineage of *Vibrio parahaemolyticus* by independent acquisition of related pathogenicity islands. *Appl. Environ. Microbiol.* **2017**, *83*, e01168-17. [[CrossRef](#)]
15. Lipp, E.K.; Huq, A.; Colwell, R.R. Effects of global climate on infectious disease: The cholera model. *Clin. Microbiol. Rev.* **2002**, *15*, 757–770. [[CrossRef](#)] [[PubMed](#)]
16. Lovell, C.R. Ecological fitness and virulence features of *Vibrio parahaemolyticus* in estuarine environments. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 1781–1794. [[CrossRef](#)] [[PubMed](#)]
17. Martínez-Urtaza, J.; Blanco-Abad, V.; Rodríguez-Castro, A.; Ansedo-Bermejo, J.; Miranda, A.; Rodríguez-Alvarez, M.X. Ecological determinants of the occurrence and dynamics of *Vibrio parahaemolyticus* in offshore areas. *ISME J.* **2012**, *6*, 994–1006. [[CrossRef](#)] [[PubMed](#)]

18. Urquhart, E.A.; Zaitchik, B.F.; Guikema, S.D.; Haley, B.J.; Taviani, E.; Chen, A.; Brown, M.E.; Huq, A.; Colwell, R.R. Use of Environmental Parameters to Model Pathogenic Vibrios in Chesapeake Bay. *J. Environ. Inf.* **2016**, *26*, 1–13. [[CrossRef](#)]
19. Whistler, C.A.; Hall, J.A.; Xu, F.; Ilyas, S.; Siwakoti, P.; Cooper, V.S.; Jones, S.H. Use of whole-genome phylogeny and comparisons for development of a multiplex PCR assay to identify sequence type 36 *Vibrio parahaemolyticus*. *J. Clin. Microbiol.* **2015**, *53*, 1864–1872. [[CrossRef](#)]
20. Blackwell, K.D.; Oliver, J.D. The ecology of *Vibrio vulnificus*, *Vibrio cholerae*, and *Vibrio parahaemolyticus* in North Carolina estuaries. *J. Microbiol.* **2008**, *46*, 146–153. [[CrossRef](#)]
21. Caburlotto, G.; Haley, B.J.; Lleò, M.M.; Huq, A.; Colwell, R.R. Serodiversity and ecological distribution of *Vibrio parahaemolyticus* in the Venetian Lagoon, Northeast Italy. *Environ. Microbiol. Rep.* **2010**, *2*, 151–157. [[CrossRef](#)]
22. Cook, D.W.; Bowers, J.C.; DePaola, A. Density of total and pathogenic (tdh+) *Vibrio parahaemolyticus* in Atlantic and Gulf Coast molluscan shellfish at harvest. *J. Food Prot.* **2002**, *65*, 1873–1880. [[CrossRef](#)]
23. Froelich, B.A.; Ayrapetyan, M.; Fowler, P.; Oliver, J.D.; Noble, R.T. Development of a matrix tool for the prediction of *Vibrio* species in oysters harvested from North Carolina. *Appl. Environ. Microbiol.* **2015**, *81*, 1111–1119. [[CrossRef](#)]
24. Johnson, C.N.; Flowers, A.R.; Noriega, N.F.; Zimmerman, A.M.; Bowers, J.C.; DePaola, A.; Grimes, D.J. Relationships between environmental factors and pathogenic vibrios in the northern Gulf of Mexico. *Appl. Environ. Microbiol.* **2010**, *76*, 7076–7084. [[CrossRef](#)] [[PubMed](#)]
25. Johnson, C.N.; Bowers, J.C.; Griffith, K.J.; Molina, V.; Clostio, R.W.; Pei, S.; Laws, E.; Paranjpye, R.N.; Strom, M.S.; Chen, A.; et al. Ecology of *Vibrio parahaemolyticus* and *Vibrio vulnificus* in the coastal and estuarine waters of Louisiana, Maryland, Mississippi, and Washington (United States). *Appl. Environ. Microbiol.* **2012**, *78*, 7249–7257. [[CrossRef](#)] [[PubMed](#)]
26. Julie, D.; Solen, L.; Antoine, V. Ecology of pathogenic and non-pathogenic *Vibrio parahaemolyticus* on the French Atlantic coast. Effects of temperature, salinity, turbidity and chlorophyll a. *Environ. Microbiol.* **2010**, *12*, 929–937. [[CrossRef](#)] [[PubMed](#)]
27. Kaneko, T.; Colwell, R.R. Ecology of *Vibrio parahaemolyticus* in Chesapeake Bay. *J. Bacteriol.* **1973**, *113*, 24–32.
28. Paranjpye, R.N.; Nilsson, W.B.; Liermann, M.; Hilborn, E.D.; George, B.J.; Li, Q.; Bill, B.D.; Trainer, V.L.; Strom, M.S.; Sandifer, P.A. Environmental influences on the seasonal distribution of *Vibrio parahaemolyticus* in the Pacific Northwest of the USA. *FEMS Microbiol. Ecol.* **2015**, *91*, 121. [[CrossRef](#)]
29. Parveen, S.; Hettiarachchi, K.A.; Bowers, J.C.; Jones, J.L.; Tamplin, M.L.; McKay, R.; Beatty, W.; Brohawn, K.; DaSilva, L.V.; DePaola, A. Seasonal distribution of total and pathogenic *Vibrio parahaemolyticus* in Chesapeake Bay oysters and waters. *Int. J. Food Microbiol.* **2008**, *128*, 354–361. [[CrossRef](#)]
30. Parveen, S.; DaSilva, L.; DePaola, A.; Bowers, J.; White, C.; Munasinghe, K.A.; Brohawn, K.; Mudoh, M.; Tamplin, M.L. Development and validation of a predictive model for the growth of *Vibrio parahaemolyticus* in post-harvest shellstock oysters. *Int. J. Microbiol.* **2013**, *161*, 1–6. [[CrossRef](#)]
31. Takemura, A.; Chien, D.; Polz, M. Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Front. Microbiol.* **2014**, *5*. [[CrossRef](#)]
32. Froelich, B.A.; Noble, R.T. *Vibrio* bacteria in raw oysters: Managing risks to human health. *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371*, 20150209. [[CrossRef](#)]
33. Crystal, E.N.; Schuster, B.M.; Striplin, M.J.; Jones, S.H.; Whistler, C.A.; Cooper, V.S. Influence of seasonality on the genetic diversity of *Vibrio parahaemolyticus* in New Hampshire shellfish waters as determined by multilocus sequence analysis. *Appl. Environ. Microbiol.* **2012**, *78*, 3778–3782.
34. Mahoney, J.C.; Gerding, M.J.; Jones, S.H.; Whistler, C.A. Comparison of the pathogenic potentials of environmental and clinical *Vibrio parahaemolyticus* strains indicates a role for temperature regulation in virulence. *Appl. Environ. Microbiol.* **2010**, *76*, 7459–7465. [[CrossRef](#)] [[PubMed](#)]
35. Schuster, B.M.; Tyzik, A.L.; Donner, R.A.; Striplin, M.J.; Almagro-Moreno, S.; Jones, S.H.; Cooper, V.S.; Whistler, C.A. Ecology and genetic structure of a northern temperate *Vibrio cholerae* population related to toxigenic isolates. *Appl. Environ. Microbiol.* **2011**, *77*, 7568–7575. [[CrossRef](#)] [[PubMed](#)]
36. Urquhart, E.A.; Jones, S.H.; Jong, W.Y.; Schuster, B.M.; Marcinkiewicz, A.L.; Whistler, C.A.; Cooper, V.S. Environmental conditions associated with elevated *Vibrio parahaemolyticus* concentrations in Great Bay Estuary, New Hampshire. *PLoS ONE* **2016**, *11*, e0155018. [[CrossRef](#)] [[PubMed](#)]

37. Bartley, C.H.; Slanetz, L.W. Occurrence of *Vibrio parahaemolyticus* in Estuarine Waters and Oysters of New Hampshire. *Appl. Environ. Microbiol.* **1971**, *21*, 965–966.
38. Jones, S.; Summer-Brason, B. Incidence and detection of pathogenic *Vibrio* spp. in a northern New England estuary, USA. *J. Shellfish Res.* **1998**, *17*, 1665–1669.
39. O'Neill, K.; Jones, S.; Grimes, D. Incidence of *Vibrio vulnificus* in northern New England water and shellfish. *FEMS Microbiol. Lett.* **1990**, *72*, 163–167. [[CrossRef](#)]
40. Nash, C.; Dejadon, B. *New Hampshire Vibrio Parahaemolyticus Risk Evaluation*; New Hampshire Department of Environmental Services, Shellfish Program: Portsmouth, NH, USA, 2019.
41. National Estuarine Research Reserve System-Centralized Data Management Office. Available online: <https://cdmo.baruch.sc.edu/dges/>. (accessed on 2 October 2019).
42. Kaysner, C.; DePaola, A. Chapter 9: *Vibrio*. In *Bacteriological Analytical Manual Online*; US Food and Drug Administration: Silver Spring, MD, USA, 2004. Available online: <https://www.fda.gov/food/laboratory-methods-food/bam-vibrio> (accessed on 17 June 2019).
43. Panicker, G.; Call, D.R.; Krug, M.J.; Bej, A.K. Detection of pathogenic *Vibrio* spp. in shellfish by using multiplex PCR and DNA microarrays. *Appl. Environ. Microbiol.* **2004**, *70*, 7436–7444. [[CrossRef](#)]
44. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria; Available online: <https://www.R-project.org/> (accessed on 1 November 2019).
45. Faraway, J.J. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2016.
46. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, NY, USA, 2016.
47. Hastie, T.; Tibshirani, R. *Generalized Additive Models*; CRC Press: Boca Raton, FL, USA, 1990; Volume 43.
48. Wood, S.N. *Generalized Additive Models: An Introduction with R*, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2017.
49. Ruckstuhl, A. Introduction to Nonlinear Regression. In IDP Institut für Datenanalyse und Prozessdesign, Zürcher Hochschule für Angewandte Wissenschaften. 2010, p. 365. Available online: <https://stat.ethz.ch/~stahel/courses/cheming/nlreg10E.pdf> (accessed on 14 July 2019).
50. Akaike, H. A New Look at the Statistical Model Identification. In *Selected Papers of Hirotugu Akaike*; Springer: New York, NY, USA, 1974; pp. 215–222.
51. Naumova, E.N.; MacNeill, I.B. Seasonality Assessment for Biosurveillance. In *Advances in Statistical Methods for the Health Sciences*; Auger, J.L., Balakrishnan, N., Mesbah, M., Molenberghs, G., Eds.; Birkhäuser: Basel, Switzerland, 2007; pp. 437–450.
52. Alarcon Falconi, T.M.; Cruz, M.; Naumova, E.N. The shift in seasonality of legionellosis in the U.S. *Epidemiol. Infect.* **2018**, *146*, 1824–1833. [[CrossRef](#)]
53. Cox, A.; Gomez-Chiarri, M. *Vibrio parahaemolyticus* in Rhode Island coastal ponds and the estuarine environment of Narragansett Bay. *Appl. Environ. Microbiol.* **2012**, *78*, 2996–2999. [[CrossRef](#)]
54. DePaola, A.; Hopkins, L.H.; Peeler, J.T.; Wentz, B.; McPhearson, R.M. Incidence of *Vibrio parahaemolyticus* in US coastal waters and oysters. *Appl. Environ. Microbiol.* **1990**, *56*, 2299–2302.
55. Deepanjali, A.; Kumar, H.S.; Karunasagar, I. Seasonal variation in abundance of total and pathogenic *Vibrio parahaemolyticus* bacteria in oysters along the southwest coast of India. *Appl. Environ. Microbiol.* **2005**, *71*, 3575–3580. [[CrossRef](#)] [[PubMed](#)]
56. Kaneko, T.; Colwell, R.R. The annual cycle of *Vibrio parahaemolyticus* in Chesapeake Bay. *Microb. Ecol.* **1997**, *4*, 135–155. [[CrossRef](#)] [[PubMed](#)]
57. Oberbeckmann, S.; Wichels, A.; Wiltshire, K.H.; Gerdt, G. Occurrence of *Vibrio parahaemolyticus* and *Vibrio alginolyticus* in the German Bight over a seasonal cycle. *Antonie Van Leeuwenhoek* **2011**, *100*, 291–307. [[CrossRef](#)] [[PubMed](#)]
58. Su, Y.C.; Liu, C. *Vibrio parahaemolyticus*: A concern of seafood safety. *Food Microbiol.* **2007**, *24*, 549–558. [[CrossRef](#)] [[PubMed](#)]
59. Turner, J.W.; Malayil, L.; Guadagnoli, D.; Cole, D.; Lipp, E.K. Detection of *Vibrio parahaemolyticus*, *Vibrio vulnificus* and *Vibrio cholerae* with respect to seasonal fluctuations in temperature and plankton abundance. *Environ. Microbiol.* **2014**, *16*, 1019–1028. [[CrossRef](#)]
60. Oberbeckmann, S.; Fuchs, B.M.; Meiners, M.; Wichels, A.; Wiltshire, K.H.; Gerdt, G. Seasonal dynamics and modeling of a *Vibrio* community in coastal waters of the North Sea. *Microb. Ecol.* **2012**, *63*, 543–551. [[CrossRef](#)]

61. Davis, B.J.; Jacobs, J.M.; Davis, M.F.; Schwab, K.J.; DePaola, A.; Curriero, F.C. Environmental determinants of *Vibrio parahaemolyticus* in the Chesapeake Bay. *Appl. Environ. Microbiol.* **2017**, *83*, e01147-17. [[CrossRef](#)]
62. Vezzulli, L.; Pezzati, E.; Moreno, M.; Fabiano, M.; Pane, L.; Pruzzo, C.; VibrioSea Consortium. Benthic ecology of *Vibrio* spp. and pathogenic *Vibrio* species in a coastal Mediterranean environment (La Spezia Gulf, Italy). *Microb. Ecol.* **2009**, *58*, 808–818. [[CrossRef](#)]
63. Froelich, B.; Bowen, J.; Gonzalez, R.; Snedeker, A.; Noble, R. Mechanistic and statistical models of total *Vibrio* abundance in the Neuse River Estuary. *Water Res.* **2013**, *47*, 5783–5793. [[CrossRef](#)]
64. Hsieh, J.; Fries, J.; Noble, R. Dynamics and predictive modelling of *Vibrio* spp. in the Neuse River Estuary, North Carolina, USA. *Environ. Microbiol.* **2008**, *10*, 57–64. [[CrossRef](#)]
65. Tonkin, J.D.; Bogan, M.T.; Bonada, N.; Rios-Touma, B.; Lytle, D.A. Seasonality and predictability shape temporal species diversity. *Ecology* **2017**, *98*, 1201–1216. [[CrossRef](#)] [[PubMed](#)]
66. Shmueli, G. To explain or to predict? *Stat. Sci.* **2010**, *25*, 289–310. [[CrossRef](#)]
67. DePaola, A.; Nordstrom, J.L.; Bowers, J.C.; Wells, J.G.; Cook, D.W. Seasonal abundance of total and pathogenic *Vibrio parahaemolyticus* in Alabama oysters. *Appl. Environ. Microbiol.* **2003**, *69*, 1521–1526. [[CrossRef](#)] [[PubMed](#)]
68. Young, I.; Gropp, K.; Fazil, A.; Smith, B.A. Knowledge synthesis to support risk assessment of climate change impacts on food and water safety: A case study of the effects of water temperature and salinity on *Vibrio parahaemolyticus* in raw oysters and harvest waters. *Food Res. Int.* **2015**, *68*, 86–93. [[CrossRef](#)]
69. Parveen, S.; Jahncke, M.; Elmahdi, S.; Crocker, H.; Bowers, J.; White, C.; Gray, S.; Morris, A.C.; Brohawn, K. High salinity relaying to reduce *Vibrio parahaemolyticus* and *Vibrio vulnificus* in Chesapeake Bay oysters (*Crassostrea virginica*). *J. Food Sci.* **2017**, *82*, 484–491. [[CrossRef](#)] [[PubMed](#)]
70. Turner, J.W.; Good, B.; Cole, D.; Lipp, E.K. Plankton composition and environmental factors contribute to *Vibrio* seasonality. *ISME J.* **2009**, *3*, 1082. [[CrossRef](#)]
71. Wong, H.C.; Wang, P. Induction of viable but nonculturable state in *Vibrio parahaemolyticus* and its susceptibility to environmental stresses. *J. Appl. Microbiol.* **2004**, *96*, 359–366. [[CrossRef](#)] [[PubMed](#)]
72. Wong, H.C.; Peng, P.Y.; Han, J.M.; Chang, C.Y.; Lan, S.L. Effect of mild acid treatment on the survival, enteropathogenicity, and protein production in *Vibrio parahaemolyticus*. *Infect. Immun.* **2015**, *66*, 3066–3071.
73. Main, C.R.; Salvitti, L.R.; Whereat, E.B.; Coyne, K.J. Community-level and species-specific associations between phytoplankton and particle-associated *Vibrio* species in Delaware’s inland bays. *Appl. Environ. Microbiol.* **2015**, *81*, 5703–5713. [[CrossRef](#)]
74. Grimes, D.J.; Ford, T.E.; Colwell, R.R.; Baker-Austin, C.; Martinez-Urtaza, J.; Subramaniam, A.; Capone, D.G. Viewing marine bacteria, their activity and response to environmental drivers from orbit: Satellite remote sensing of bacteria. *Microb. Ecol.* **2014**, *67*, 489–500. [[CrossRef](#)]
75. Hartwick, M.; Jones, S.H. *Vibrio parahaemolyticus* Population Dynamics Within an Estuarine Microbial Community. *Int. Symp. Microb. Ecol.* **2016**. Poster presentation.
76. Hartwick, M.; Berenson, A.; Whistler, C.A.; Cooper, V.S.; Jones, S.H. Seasonal Plankton Communities and Plankton-Associated *Vibrio parahaemolyticus* Improve the Estimation of *Vibrio parahaemolyticus* Concentration in Oysters in the Great Bay Estuary. In preparation.
77. Gilbert, J.A.; Steele, J.A.; Caporaso, J.G.; Steinbrück, L.; Reeder, J.; Temperton, B.; Huse, S.; McHardy, A.C.; Knight, R.; Joint, I.; et al. Defining seasonal marine microbial community dynamics. *ISME J.* **2012**, *6*, 298. [[CrossRef](#)] [[PubMed](#)]
78. Rehnstam-Holm, A.; Godhe, A.; Härnström, K.; Raghunath, P.; Saravanan, V.; Collin, B.; Karunasagar, I.; Karunasagar, I. Association between phytoplankton and *Vibrio* spp. along the southwest coast of India: A mesocosm experiment. *Aquat. Microb. Ecol.* **2010**, *58*, 127–139.
79. Davis, B.J.; Jacobs, J.M.; Zaitchik, B.; DePaola, A.; Curriero, F.C. *Vibrio parahaemolyticus* in the Chesapeake Bay: Operational in situ prediction and forecast models can benefit from inclusion of lagged water quality measurements. *Appl. Environ. Microbiol.* **2019**, *85*, e01007-19. [[CrossRef](#)] [[PubMed](#)]
80. Schillaci, C. *Vibrio parahaemolyticus* (Vp) Illness in Massachusetts and Impact on the Cape and Island’s Shellfish Industry. In Proceedings of the 4th Annual Cape Coastal Conference, Hyannis, MA, USA, 6–7 December 2016; Available online: http://www.waquoitbayreserve.org/wp-content/uploads/CC4_Schillaci_Vibrio.pdf (accessed on 2 October 2019).

81. Semenza, J.C.; Trinanes, J.; Lohr, W.; Sudre, B.; Löfdahl, M.; Martinez-Urtaza, J.; Nichols, G.L.; Rocklöv, J. Environmental suitability of *Vibrio* infections in a warming climate: An early warning system. *Environ. Health Perspect.* **2017**, *125*, 107004. [[CrossRef](#)] [[PubMed](#)]
82. Vezzulli, L.; Grande, C.; Reid, P.C.; Hélaouët, P.; Edwards, M.; Höfle, M.G.; Brettar, I.; Colwell, R.R.; Pruzzo, C. Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E5062–E5071. [[CrossRef](#)]
83. Vezzulli, L.; Pezzati, E.; Brettar, I.; Höfle, M.; Pruzzo, C. Effects of global warming on vibrio ecology. *Microbiol. Spectr.* **2015**, *3*. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).