AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research®

Key Points:
- A new framework uses empirical orthogonal function, conditional simulation, and Bayesian kriging to estimate lake hypoxia extent
- In Lake Erie, the approach captures the dynamic nature of bottom hypoxia in offshore areas, but more sensors are needed in nearshore areas
- Open-source R code and a web application are provided for researchers to build upon the research and explore other hypoxia datasets

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

Correspondence to:
W. Xu,
xuwz.uiuc@gmail.com

Spatio-Temporal Analysis of Hypoxia in the Central Basin of Lake Erie of North America

Wenzhao Xu[1] [iD], **Paris D. Collingsworth[2]** [iD], **Richard Kraus[3]** [iD], and **Barbara Minsker[4]** [iD]

[1]Amazon A9, Palo Alto, CA, USA, [2]Department of Forestry and Natural Resources and Illinois-Indiana Sea Grant, Purdue University, West Lafayette, USA, [3]US Geological Survey, Sandusky, OH, USA, [4]Department of Civil and Environmental Engineering, Southern Methodist University, Dallas, TX, USA

Abstract We develop a spatio-temporal geostatistical interpolation framework to estimate hypoxia extent (dissolved oxygen [DO] concentrations below 2 mg/L) with data from a network of DO loggers. The framework uses empirical orthogonal functions and Bayesian kriging to identify the spatially varying temporal pattern and estimate the distribution of hypoxia, including estimation uncertainty. A prototype web application is also developed in R. The framework is applied to analyze spatio-temporal dynamics of DO in the central basin of Lake Erie in North America using data sampled from a logger network placed on the lake bottom during the summers of 2014, 2015, and 2016. Cross-validation results demonstrate that the framework is capable of capturing the dynamic nature of bottom hypoxia over offshore areas, but nearshore areas have poor interpolation performance due to the impacts of complex physical processes such as seiche events. The findings showed that in the central basin, hypoxia started to emerge in early August of 2014, while in 2015 and 2016 hypoxia began in July. The peak hypoxia extent occurred in late September 2014, mid-August 2015, and early September 2016. The prediction error of the overall spatial extent of hypoxia was as large as 25% of the interpolation area based on current logger deployment. Based on the cross-validation and interpolation error, we suggest placing more loggers in nearshore areas to reduce prediction error near the margins of the hypoxic zone.

1. Introduction

Hypoxia (dissolved oxygen [DO] concentrations lower than 2 mg/L) is a global water quality issue in productive lakes and estuaries (Breitburg et al., 2018; Diaz, 2001). In fresh water systems, hypoxia is caused by excessive inputs of phosphorus, a limiting nutrient, and excess algal production (Diaz, 2001). When algal cells die, they sink to the bottom of the lake where bacteria decompose the cells and consume oxygen. Then the bottom layers of water become hypoxic when oxygen depletion exceeds oxygen replenishment, causing fish and invertebrate stress or mortality, diet shifts, and changes in food web structure (Pihl et al., 1992; Rao et al., 2014; Scavia et al., 2014; Vanderploeg et al., 2009). Previous work has shown that 71 of 365 lakes worldwide have developed hypoxic conditions since the mid-19th century (Jenny et al., 2016) and many inland lakes suffer from seasonal hypoxia (Diaz, 2001), including Lake Erie, which is the focus of this study. Lake Erie is the shallowest and most productive of the Laurentian Great Lakes of North America and, therefore, has experienced many problems associated with cultural eutrophication (eutrophication intensified by human activity), including bottom hypoxia (Burns et al., 2005). Lake Erie has three distinct basins that vary in depth, temperature, and productivity and this unique bathymetry has implications for the formation of bottom hypoxia. While infrequent hypoxia can occur in the eastern and western basins, the most persistent and severe hypoxia occurs in the mesotrophic central basin of Lake Erie (Anderson et al., 1984; Scavia et al., 2014; Zhou et al., 2013), which tends to have a thin hypolimnion (ranging from 0.5 to 10m, depending on time and location [Rao et al., 2008 and Bouffard et al., 2013]) during stratified conditions, and oxygen depletion often exceeds replenishment rates.

Monitoring hypoxia in Lake Erie is important to achieve Great Lakes Water Quality Agreement (GLWQA, https://www.epa.gov/glwqa) objectives to reduce the severity and extent of hypoxia. Achieving these objectives is complicated by our inability to accurately predict hypoxia and thus evaluate progress toward the binational water quality goals. One way to monitor hypoxia development is through the deployment of sensor networks that measure DO continuously in different locations. Since a limited number of sensors can be deployed, interpolation that utilizes spatial and temporal correlations among loggers is needed to

estimate the values at unsampled locations and track progress in improving water quality as required under the GLWQA.

Previous research has modeled spatio-temporal data using spatio-temporal kriging, such as interpolating PM10 in the atmosphere (Gräler et al., 2016), soil water content (Snepvangers, et al., 2003), dissolved oxygen (Matli et al., 2018), harmful algal blooms (Fang et al., 2019) and surface water chloride concentrations (Jat & Serre, 2016). These models fit covariance functions of relevant variables, in which distances are defined in both space and time. Another discrete time approach is dynamic spatio-temporal models (DSTM) that model the process and relations at time based on state at time $t - 1$. One specific example of this approach is a first-order vector autoregression model, which incorporates residuals that have spatial correlations (Bakar & Sahu, 2015). However, both types of models assume a defined spatio-temporal trend so that the spatio-temporal residuals satisfy temporal and/or spatial stationarity. Therefore, when there is no such spatio-temporal trend that can be easily extracted, the above methods are not applicable. A more flexible approach from Lindström et al. (2014) is to model the data as a set of spatially varying temporal basis functions, or empirical orthogonal functions (EOF). The EOFs account for the temporal variability in data and are associated with spatially varying coefficients. Coefficients estimated by universal kriging then can be used to interpolate DO at target locations. EOF approach reduces the spatio-temporal problem into spatial problem only by decomposing temporal patterns into modes, which are modeled separately with different spatial structures. Yet the spatial nonstationarity problem may still exist. To estimate hypoxia uncertainty, Zhou et al. (2013) used conditional simulation with a spatial universal kriging model to simulate DO distribution given collected data from fixed stations and cruises. The simulation values are used to predict hypoxia values and their uncertainties.

In this study, we propose a novel framework that combines these two approaches, EOF kriging interpolation and conditional simulation, within a Bayesian framework that can better address kriging model uncertainty (Section 3). We apply and cross-validate the framework with lake bottom DO data in the central basin of Lake Erie sampled in 2014, 2015, and 2016. Using these methods, seasonal changes to hypoxia extent in Lake Erie are characterized (Section 4). We then discuss the model performance in nearshore and offshore areas and provide suggestions to optimize logger locations (Section 5).

2. Study Area and Data Descriptions

To estimate the spatial extent of hypoxia in Lake Erie, US Environmental Protection Agency–Great Lakes National Program Office (EPA-GLNPO) and U.S. Geological Survey (USGS) deployed a network of Onset HOBO U26 temperature and DO loggers that continuously sampled lake bottom DO in the central basin of Lake Erie during the summers of 2014, 2015, and 2016 (Figure 1). 10 loggers in the offshore areas remained in the same sampling position in all three years (stations starting with ER, where ER stands for Erie). The locations of nearshore loggers changed every year and were usually deployed at the intersection between the lake bed and the thermocline. The loggers were deployed in early summer and retrieved in early fall, although the deployment and retrieval dates are different for nearshore and offshore loggers. All loggers measured and recorded DO concentrations and temperatures every 10 min.

Each logger began measuring dissolved oxygen levels at different times due to vessel schedules. To standardize the data for analysis, we performed hourly averaging of the raw logger data. We also converted the longitude and latitude to Universal Transverse Mercator (UTM) coordinates using "WGS84" datum and zone number "17T." This projects longitude and latitude onto a 2-D plane with linear length units so that spatial Euclidean distance can be calculated for interpolation. After this preprocessing, some spikes were still evident in the data at the hourly scale. Nonetheless, no additional smoothing was undertaken because the spikes may indicate important lake processes such as the intrusion of surrounding hypoxic water or seiche events that mix the water column quickly.
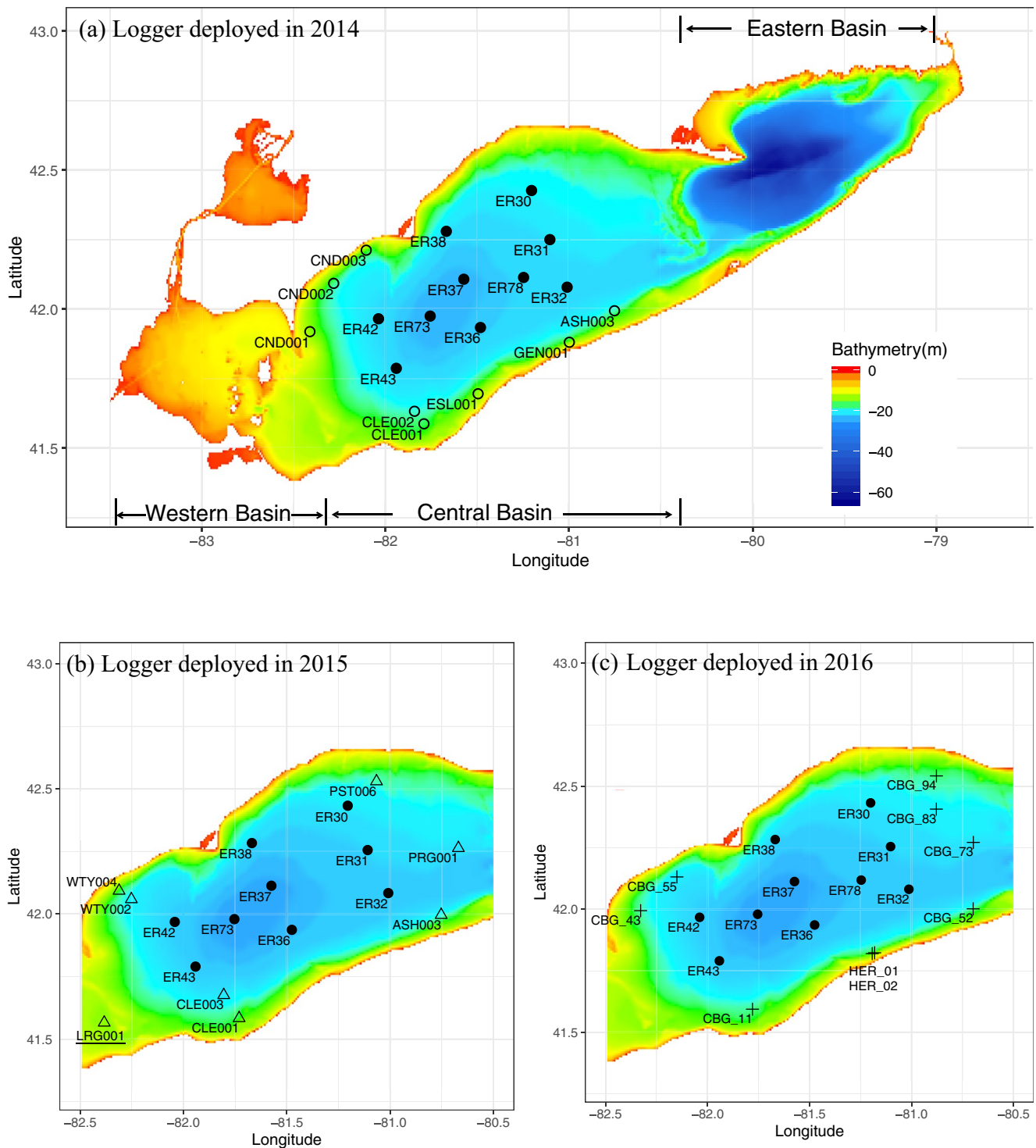
Figure 1. Lake Erie Bathymetry and logger locations in 2014–2016. Offshore loggers are deployed at stations whose names start with "ER." Other loggers are deployed to nearshore stations, which are different every year. Note ER78 is not presented in 2015.

3. Methodology

In the first step of the framework, we conduct two spatio-temporal interpolations, inverse distance weighting (IDW) as the baseline method (Section 3.1) and EOF kriging interpolations (Section 3.2), to estimate the spatial extent of hypoxia. The hourly averaged data are interpolated at unsampled locations during the time

period when all nearshore and offshore loggers have available data in each year (Section 3.3). Finally, a web application is built for interactively exploring the data set (Section 3.4)

3.1. IDW Interpolation at Each Time Step

Inverse distance weighting (IDW) is a commonly used deterministic spatial interpolation technique. The target location is interpolated using a weighted average of values from other sampled locations and the weights are computed based on the inverse of the distance between the target locations and sampled points. The mathematical equations are:

$$y(x) = \begin{cases} \dfrac{\sum_{i=1}^{N} \omega_i(x) y_i}{\sum_{i=1}^{N} \omega_i(x)} & (d(x, x_i) \neq 0) \\ y_i & (d(x, x_i) = 0) \end{cases} \tag{1}$$

where $y(x)$ is the interpolated value at target location x. y_i is the sampled value at location x_i. N is the total number of sampled points surrounding x that are considered.

The weights ω are calculated as:

$$\omega_i(x) = 1/d(x, x_i)^2 \tag{2}$$

where $d(x, x_i)$ is the spatial distance between the target location x and sampled location x_i. IDW interpolation is not able to incorporate covariate information such as bathymetry. The weights are solely depending on distances. N is a tuning parameter for IDW. We choose $N = 5$ for this study, which means that the five nearest sampling loggers are used to interpolate oxygen levels at a target location. This is a fitting parameter that is set to best fit the particular data set. A larger N increases the computational time and could bias the interpolation results due to different patterns at distant loggers. IDW interpolation is performed at every time step to generate interpolation time series at every spatial point.

3.2. Empirical Orthogonal Function (EOF) Kriging Interpolation

As an alternative to IDW, EOF kriging interpolation is also evaluated, following Lindström et al. (2014). DO data are modeled using the form of:

$$y(s, t) = \mu(s, t) + \epsilon(s, t) \tag{3}$$

where $y(s, t)$ denotes the spatio-temporal DO values, $\mu(s, t)$ is the mean trend, and $\epsilon(s, t)$ is the space-time residuals; s, t are the space and time indices, respectively.

The EOF kriging interpolation algorithm consists of the following steps:

(1) Decompose the mean trend term $\mu(s, t)$ using singular value decomposition (SVD) to obtain empirical orthogonal functions and corresponding basis coefficients;
(2) For coefficients in each basis function, conduct universal kriging on the target grid using maximum likelihood estimation and Bayesian framework to fit a variogram model;
(3) Reconstruct the interpolated trend term $\mu(s, t)$ at the target grid from the interpolated coefficients, calculated from Step 2;
(4) Conduct IDW interpolations on the residuals $\epsilon(s, t)$ for each time step t;
(5) Add the reconstructed trend and interpolated residuals as the final interpolated values.

More details on extracting the EOFs and conducting interpolations are given below.

3.2.1. Empirical Orthogonal Function

The SVD decomposition in Step 1 gives:

$$D = U\Sigma V^T \tag{4}$$

where D is the raw data matrix that is column-wise centered and scaled to mean zero and unit variance, with element $d_{i,j}$ representing the centered and scaled DO values at time from logger j; and j are the row and column indices. The column vectors in U and V are orthogonal eigenvectors of DD^T and $D^T D$. Σ is a diagonal matrix containing the eigenvalues of DD^T and $D^T D$ ordered by magnitudes. The first r columns of

Table 1
Prior Distributions of the Variogram Model

Parameter	Prior	Justification
Covariates coefficient (ω)	Flat	We have no prior knowledge of this parameter
Sill (σ)	Reciprocal distribution (p(σ^2) is proportional to $1/\sigma^2$).	No parameters of prior distributions are needed.
Nugget (τ)	Fixed as 0	We assume the measurement errors are zero with no small-scale variability for simplicity.
Range (ϕ)	Uniform distribution ranging from 20 to 70 km at 5 km intervals	20 km is the software's minimum recommended pairwise distance. Selected range is approximately 1/3 of the maximum pairwise distance (recommended default values in Golden Software and "gstat" R package). This range limitation is also enforced in the maximum likelihood estimation.

matrix U are defined as the temporal basis functions or empirical orthogonal functions (EOFs), which are kept and then smoothed using a 1-D spline. These basis functions are fundamental to building the logger data time series because they represent the basic DO patterns over time.

We then obtain the coefficients of the basis function using linear regression, so that:

$$y(s,t) = \sum_{k=1}^{r+1} \beta_k(s) f_k(t) + \epsilon(s,t) \tag{5}$$

where $y(s,t)$ are the DO values at sampling location s at time t, $f_k(t)$ is the kth EOF (i.e., the smoothed k^{th} column vector of matrix U), and $\beta_k(s)$ are the coefficients of $f_k(t)$, which vary across space. The $(r+1)^{th}$ EOF is the intercept term for which $f_{r+1}(t) \equiv 1$. $\epsilon(s,t)$ are the residuals. The function $f_k(t)$ is called the empirical orthogonal function (EOF) and has been used for analyzing the modes in other spatio-temporal datasets such as sea level pressure (Hannachi et al., 2007) and DO in Chesapeake Bay (Scully, 2016). The number of EOFs r in the SVD also needs to be pre-defined and cross-validation can be used to help choose the best r (following Lindström et al. (2014)).

3.2.2. Basis Coefficient Kriging Interpolation

After obtaining the basis functions $f_k(t)$, the coefficients $\beta_k(s)$ in Equation 5 are then interpolated using universal kriging. The detailed equations of universal kriging are summarized by Cressie (1993). The co-variates considered in the universal kriging interpolation include x and y in UTM system, bathymetry, and the square of the bathymetry. The square of the bathymetry is included because Zhou et al. (2013) found it important for kriging interpolation of DO based on Bayesian Information Criterion (BIC).

Note that in our framework, the covariates should be either consistent across time (so they can be included in the coefficient interpolation) or available across the entire spatial interpolation area at hourly resolution for detrending the original data. Therefore, covariates such as satellite-derived monthly surface temperatures and surface chlorophyll concentrations explored by Zhou et al. (2013) are not suitable for this work.

We use two approaches to estimate the variogram model: maximum likelihood (MLE) and Bayesian framework. MLE is used to numerically maximize the likelihood of the Gaussian process (i.e., kriging). Bayesian kriging assumes a prior distribution of the variogram model parameters and finds parameter posterior distributions, then the expected prediction values are calculated based on the parameter posterior distributions. In this study, we use an "exponential" variogram model, where the increase in covariance with distance is modeled by an exponential function, the same approach as Zhou et al. (2013).

To implement these approaches, we use "geoR" package in R (Ribeiro Jr and Diggle, 2001). For a given prior distributions, the geoR Bayesian inference framework derives analytical marginal and joint distributions of model parameters. Then parameters are simulated iteratively without Markov chain Monte Carlo (MCMC) approach. More details are provided by Diggle and Ribeiro (2002). The priors for the parameters of the exponential variogram model are summarized in Table 1. For simplicity, we assume that the measurement errors are zero with no small-scale variability, otherwise the nugget (τ) also needs a prior distribution that is unknown. The sensor accuracy is estimated to be ±0.2 mg/L up to 8 mg/L (https://www.onsetcomp.com/products/data-loggers/u26-001), but this variance cannot easily be converted to the variance of basis

coefficients. The reciprocal distribution for sill (σ) and uniform distribution for covariates coefficient (ω) are the default settings in the "krige.bayes" function in the "geoR" package. To examine the impacts of these priors, we conducted cross validation and found that the choice of priors did not significantly influence the interpolation results for this case study (Appendix A).

The number of basis functions determines how many spatial kriging interpolations are conducted. The trend predictions at a target cell from interpolated basis coefficients are reconstructed by calculating Equation 5, where $\beta_k(s)$ is now the interpolated coefficient at grid cell s and $\epsilon(s,t)$ is the interpolated residual, which is described in the next subsection. When the interpolated DO is less than 0, which can happen with a Gaussian distribution, we set it to zero.

Each kriging interpolation can generate expected prediction values given the prediction distribution. Using conditional simulations, following the approach of Zhou et al. (2013), this framework can also generate multiple predictions on the target grid that satisfy the spatial correlations defined by the variogram. We denote $\hat{\beta}_k^*(s)$ as the expected prediction value and $\hat{\beta}_k^i(s)$ as the ith prediction simulation.

3.2.3. Residual Interpolation

After interpolating the coefficients, we interpolate the residuals. First, the residuals $\epsilon(s,t)$ from Equation 5 need to be analyzed to find the appropriate method for interpolation because spikes or large values in the residuals may influence surrounding areas. Lindström et al. (2014) proposed fitting another spatial kriging model to the residuals. In order to reduce the computational cost and model complexity, we perform only a simple IDW interpolation at each time step (Section 3.1) on the residuals. Since the IDW is a linear combination of the values of surrounding points, the temporal correlations of logger data residuals are naturally preserved in the interpolated residuals.

3.3. Hypoxia Extent Estimation

Lastly, the spatial extent of hypoxia is estimated by interpolating DO values across the target area using both IDW and EOF kriging interpolation methods. In the central basin of Lake Erie, we interpolated DO on a grid with size 0.025° in longitude and latitude, thus, one grid cell occupies approximately 5.6 km². To prevent extrapolations, which are usually not reliable since there are no data points to constrain values, the grid is filtered by a convex hull defined by the sampling locations.

For IDW interpolation, we calculate $N_h(t)$, the number of grid cells $N_h(t)$ at time t where the interpolated DO was below the hypoxic threshold (in this case, 2 mg/L). The number of cells are then converted to areas to obtain a time series of the hypoxic areas. The IDW interpolation cannot give interpolation uncertainty so that the hypoxia extent estimation is considered deterministic in this case.

For kriging interpolation on basis coefficients, we perform conditional simulation for each basis kriging interpolation and aggregate simulation results to estimate the uncertainty of DO interpolations as well as the uncertainty of the estimated hypoxic areas. The steps are:

Step 1. Build kriging models on the coefficients of each EOF and generate M possible coefficient predictions using conditional simulations.

Step 2. For each basis function, independently and randomly select a prediction realization among the M simulations. Reconstruct the spatio-temporal DO interpolations (Section 3.2), obtaining the trend from the interpolated coefficients and then adding back the IDW residuals interpolation.

Step 3. Calculate the areas where interpolated DO values are smaller than the hypoxic threshold (2 mg/L) in this simulation through time.

Step 4. Repeat Steps 2 and 3 K times to generate K time series of the hypoxic area.

The simulated time series are then summarized to estimate the hypoxia extent with uncertainty. We choose $M = 100$, assuming that 100 data points can adequately capture the distribution of kriging predictions for the basis coefficients. We choose $K = 1000$, meaning that the final results are summarized from 1,000 simulations, same as Zhou et al. (2013).
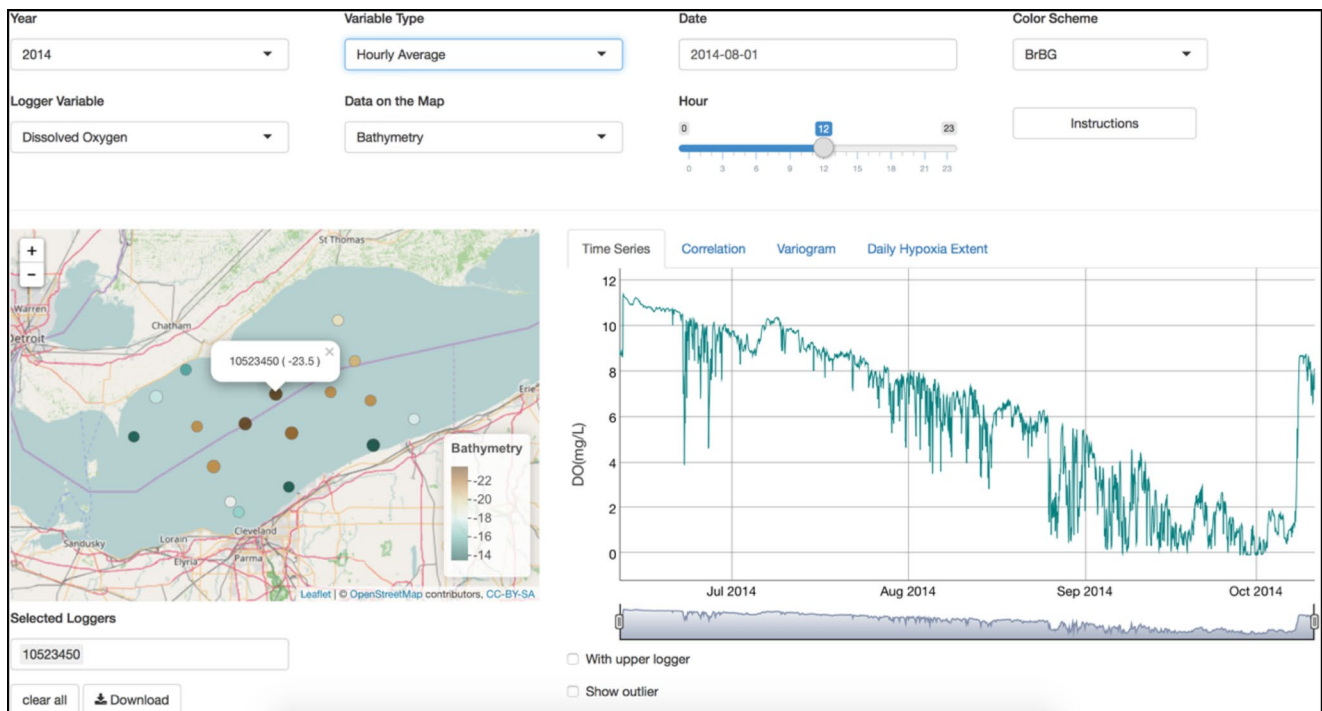
Figure 2. Interactive Web application for dissolved oxygen (DO) exploration, showing Lake Erie data.

3.4. Web Applications

We built an interactive web application called Hypoxia Interpolator using R "Shiny" packages (Chang et al., 2017). The software allows users to explore their DO data and calculate hypoxia extent with the framework developed in this paper, along with other statistics (Figure 2). In the left panel, users can see the spatial DO or temperature distribution at a certain time. In the right panel, users can visualize the time series of DO or temperature from multiple loggers using different time steps (hourly or daily), check correlations and variograms, or calculate daily hypoxia extent. More information and the source code can be found in the Supporting Information S1 and at https://eisalab.github.io/Hypoxia-Interpolator/. A live demo is at http://hypoxiamapping.smu.edu/hypoxia/.

4. Results

In this section, we apply the framework described in Section 3 to the Lake Erie data described in Section 2. The basis decompositions are first used to explore general hypoxia patterns (Section 4.1). Interpolation cross-validation results are then discussed in Section 4.2 to select the best method, which for this case study is the Bayesian method. This method is then used to examine spatio-temporal DO patterns and hypoxia extent in the central basin in Section 4.3.

4.1. Basis Decomposition and Residual Distributions

The basis functions can provide general insights on the data, particularly when they are ordered by the magnitude of eigenvalues. Figure 3 shows the first three EOFs (Equation 5) as examples from the Lake Erie DO data in 2014, 2015, and 2016.

The basis with the largest magnitude of eigenvalues, i.e., Basis_1 (leftmost basis graph in Figures 3a–3c), explains the most variance in the data. Different values of coefficients represent the extent to which such patterns are contained in the data series. Therefore, the magnitude of the coefficients of Basis_1 are the largest for most loggers (leftmost maps in Figures 3a–3c). With positive coefficients, Basis_1 reveals the most fundamental patterns: In this case, DO decreased during the sampling period in all three years, sharply
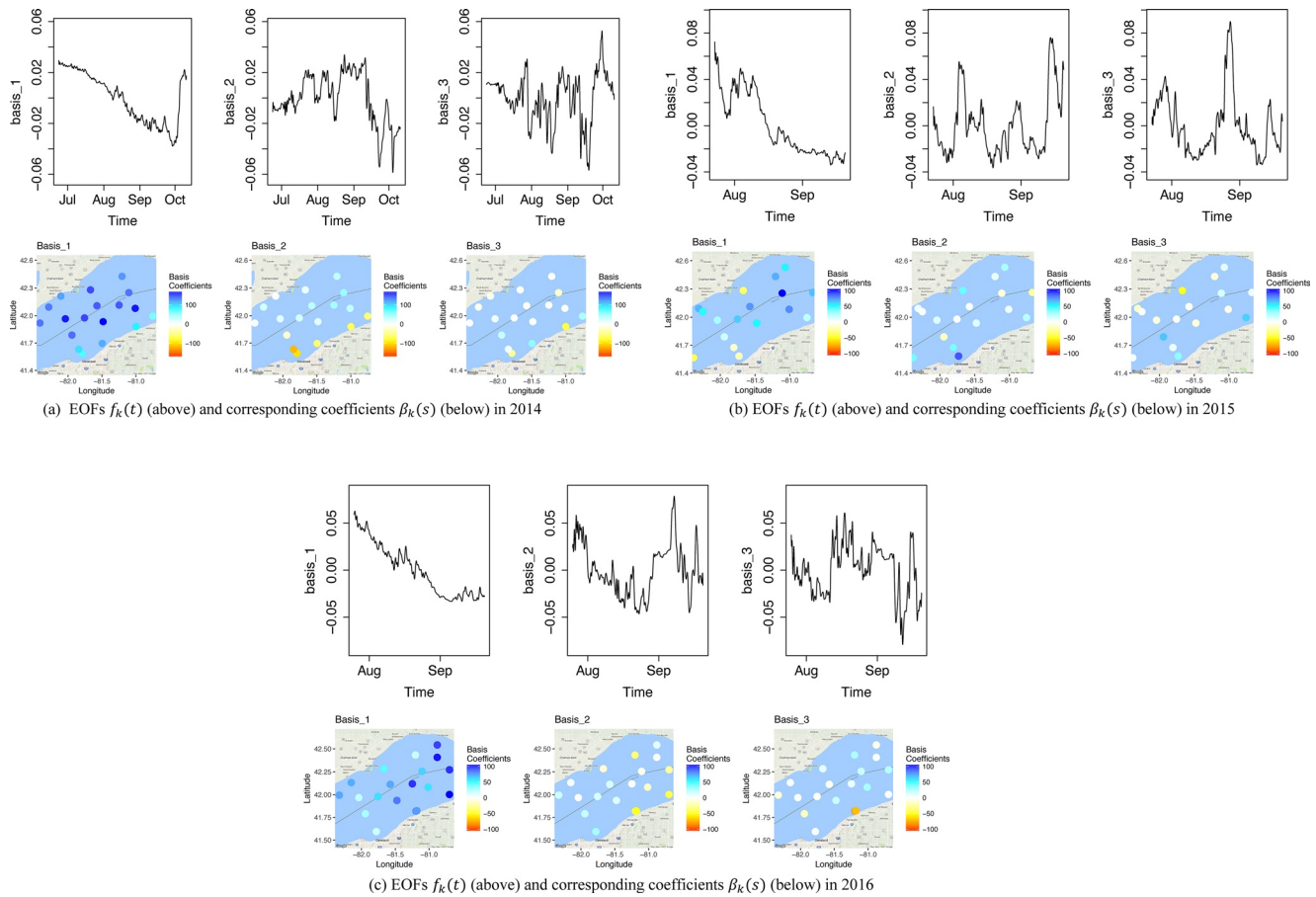
(a) EOFs $f_k(t)$ (above) and corresponding coefficients $\beta_k(s)$ (below) in 2014

(b) EOFs $f_k(t)$ (above) and corresponding coefficients $\beta_k(s)$ (below) in 2015

(c) EOFs $f_k(t)$ (above) and corresponding coefficients $\beta_k(s)$ (below) in 2016

Figure 3. Basis decompositions in 2014, 2015, and 2016.

increased in October 2014, and had a large fluctuation at the beginning of August 2015. Basis_1 also tends to be driven by the offshore logger patterns, which have relatively large positive coefficients (shown in the leftmost maps in Figure 3). Note that the values of the Basis_1 function across different years are not comparable since they have been normalized in each year.

Basis_2 and the remaining basis functions represent more detailed patterns, with coefficients generally becoming smaller and smaller, indicating that these patterns are usually less significant than Basis_1. However, some loggers have similar coefficient magnitudes as Basis_1, indicating that Basis_2 is also important for these loggers, where patterns may not be accurately approximated with only a few basis functions.

Loggers with similar basis coefficients indicate that the patterns are similar. In 2014, loggers can be roughly clustered into two large groups: south shore group and a group formed by the remaining loggers. Nearshore activities such as seiche events caused by wind could be the mechanistic cause of the two patterns, which will be discussed further in later sections. In 2015, there seems to be no clear patterns across groups of loggers, indicating that DO patterns could be influenced more by local effects such as local wind direction/strength and internal waves (Rao et al., 2008). In 2016, there is a rough trend showing that south west loggers are similar to each other; north east loggers are also similar. Factors causing east-west gradients could be responsible, such as sediment oxygen demand. Rowe et al. (2019) mechanistic model suggests there may be an east-west gradient of SOD due to a trophic gradient, which is related to early DO depletion in the western area (Mortimer, 1987)

The variogram cloud and fitted variograms for the basis coefficients of the first two basis in 2014 are shown below as examples (Figure 4). Data points are quite noisy as they contain mixed patterns of nearshore and offshore. With such small numbers of data points, posterior distribution is fat, causing large uncertainty in
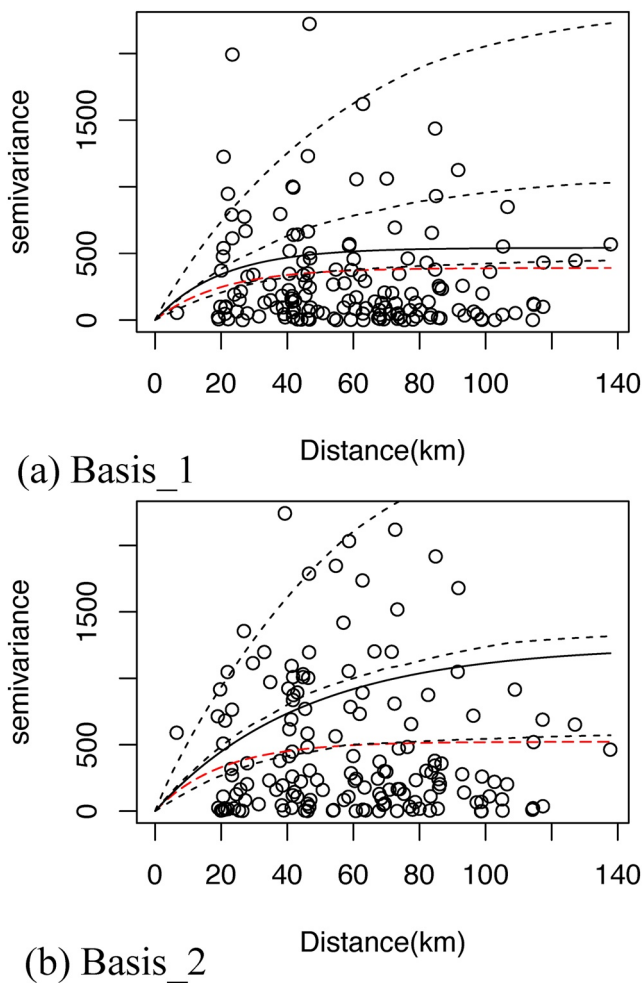
Figure 4. Variogram cloud and fitted variogram by Bayesian framework and maximum likelihood estimation for the first two bases in 2014. Semivariance is half the average squared difference between points separated at some distance. The dotted black lines are 0.95, 0.5, and 0.05 quantiles of variogram samples by the Bayesian method. The solid black line is the mode of the variogram for the Bayesian approach. The dashed red line is the variogram fit by maximum likelihood estimation.

the variogram model. Also, the maximum likelihood estimation will fit a small range (ϕ) value, indicating that the data are noisy and there may be no clear spatial correlations. Finally, the Sill (σ) by maximum likelihood estimation is less than the mode and median variogram by Bayesian estimation. All model parameters can be retrieved from intermediate modeling files provided in the Supporting Information S1.

For most time steps, residuals have insignificant effects on the interpolation results. Figure 5 shows the residuals after basis decomposition from all loggers when $r = 10$ (i.e., total of 11 basis functions, including the intercept basis) in 2014 as an example. Residuals have means close to 0 and low 25% and 75% quantile values (Figure 5b). Standard deviations of residuals range from 0.5 to 0.8 mg/L (detailed summary in Supporting Information S1). The spatio-temporal variogram (Figure 5a) reveals that the residuals still have some temporal correlation as the semivariogram increases with time at spatial distance = 0. However, no global spatial correlation patterns are obvious because the semivariance quickly rises to the plateau. These residuals are then to be interpolated by IDW interpolation at every time step.

4.2. Spatio-Temporal Interpolation Cross-Validation

For the spatio-temporal interpolations, we use cross-validation to evaluate IDW (Section 3.1) and EOF kriging interpolation (Section 3.2). We leave one logger out and interpolate its values based on all other loggers using our framework. One of the metrics is root-mean-square errors (RMSEs) between interpolation and the actual measurements across time. For N loggers, N RMSEs are calculated and summarized using boxplots (Figure 6).

In Figure 6 (left), cross-validation shows that IDW and EOF Kriging interpolation with MLE and Bayesian framework have similar RMSEs for different numbers of basis functions (r in Equation 5). The model also had generally larger RMSE in 2015 compared to 2014, indicating that in 2015, the patterns of each logger are harder to capture. Note that for MLE and Bayesian, RMSEs are calculated using the median values in all interpolation simulations. The model parameter shown in Figure 6 refers to the number of basis functions (r in Equation 5) for MLE and Bayesian method or the number of nearest sampling locations used in the IDW method.

Unlike IDW, EOF kriging models using MLE or Bayesian framework can provide estimation uncertainty. Figure 6 (right plot) shows another metric called CI coverage, which is the proportion of sampled data that fall into a given confidence interval (CI), with a value of zero indicating that no sampled data are within the confidence interval and one indicating that all of the sampled data are within the confidence interval. The figure shows the 90% CI (i.e., 5% and 95% sample quantiles as the lower and upper bounds, respectively). The results show that the Bayesian method has larger confidence intervals due to consideration of the variogram model uncertainty, so more observed data fall into the confidence interval. MLE underestimates the prediction uncertainty by not incorporating the variogram model uncertainty.

Overall, the Bayesian method with 10 basis functions performed best for this case study because: (a) compared to IDW interpolation, EOF kriging interpolation provides useful information on uncertainty of the interpolation and hypoxia extent, despite being more complicated; (b) compared to MLE method, the confidence interval from the Bayesian method contains more observed data and the CI metric has less variance, indicating that the model is less likely to underestimate uncertainty for unknown locations; and (c) $r = 10$ has better CI coverage performance compared to $r = 2$ and $r = 5$. The model with 15 basis functions has

(a) Spatio-temporal variogram

(b) Boxplot of the residuals
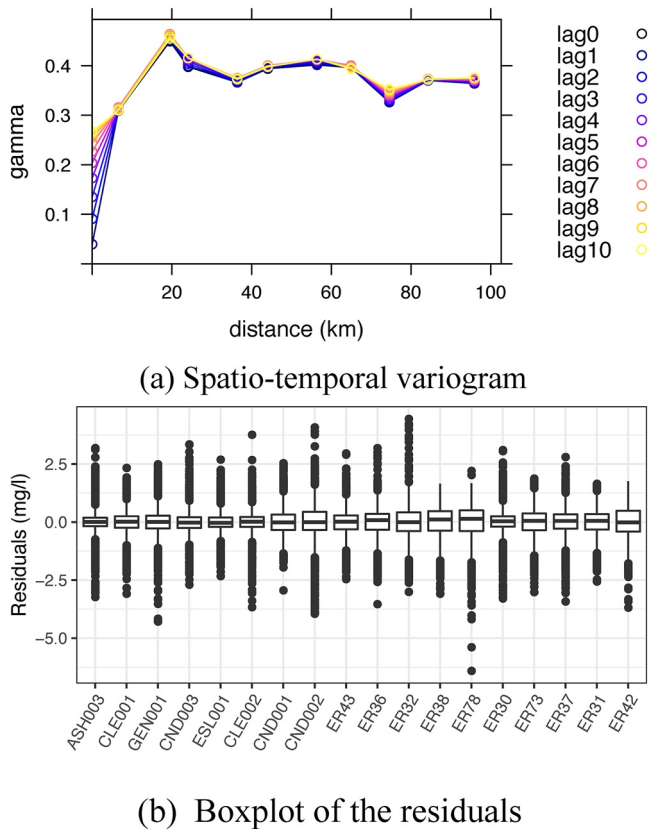
Figure 5. Residual statistics after detrending with 10 basis functions ($r = 10$) for hourly data in 2014. (a) Spatio-temporal semivariance (gamma, *y* axis) versus spatial distance (*x* axis) for different temporal lags (colors: lag0 means no difference in time dimension; lag1 means 1-h differences); (b) residuals left for each logger.

similar performance but is more complicated and the computational cost is heavier. Also, larger r incorporated more uncertainty through kriging approach, which could cause over estimating uncertainty.

It should be noted that the framework considers nearshore and offshore patterns all together, leading to high model uncertainty (wide range of fitted variogram by Bayesian method in Figure 4). Thus, uncertainty can be overestimated for some loggers, especially the offshore loggers where dynamics are relatively stable. There are also some loggers having low CI coverages, usually nearshore loggers where the patterns are completely different, thus violating the stationarity assumption (shown later in Figure 7). We discuss nearshore logger patterns in Section 5. In general, the similar RMSE across different model parameters in Figure 6 indicates that the final hypoxia extent estimation is not overly sensitive to the number of basis functions for this case study.

Visualizations of RMSE and CI coverage from cross validation reveals how model perform in different locations (Figure 7). The results show that in 2014, loggers near the shoreline have high RMSE errors and low CI coverage (top left in Figure 7). These results indicate that the DO patterns in these areas are fundamentally different from other locations and therefore the variances in the spatial random process and in the variogram model cannot completely account for them. Furthermore, Logger at ESL001 (bottom left) and GEN001 (top right) have quite different patterns from surrounding loggers. They are hard to predict because: (a) the predictions on nearshore loggers are extrapolated during cross-validation, which has worse performance in general; and (b) the DO dynamics are more complicated in nearshore areas since they are influenced by seiche events as discussed further below. The three loggers along the north-west shoreline have better performance than the south shore loggers because patterns are similar within the northwest areas, which may be from fewer seiche events due to weak wind activities. Advection of hypolimnion water by internal waves may also contribute to different patterns (Kraus et al., 2015). Additional current profiler data or mechanistic

models (Bouffard et al., 2013; Rowe et al., 2019) are needed to fully understand the mechanism. Similar results were found for 2015 and 2016 (shown in Appendix B), providing further indication that nearshore loggers have large cross-validation errors.

### 4.3. Spatio-Temporal DO Interpolation and Hypoxia Extent

Figure 8 shows the interpolated DO in 2014, 2015, and 2016, using EOF kriging interpolation with $r = 10$ and expected prediction coefficients (i.e., $\hat{\beta}_k^*(s)$, Section 2.3) within the Bayesian framework. We also provide animations of DO distribution in the Supporting Information S1. In 2014 (Figure 8a), DO depletion began in the southwestern portion of the central basin, then gradually expanded to the whole interpolation area. The hypoxic area began to contract around September 22 as the south shore zone had regained some DO, but then increased again. By early October, as the lake fall turnover started, the bottom water quickly mixed with DO-saturated surface water and the hypoxic area rapidly decreased. The phenomenon of hypoxia starting from the nearshore or shallower zones is consistent with previous numerical modeling studies (Bocaniov & Scavia, 2016). The development of hypoxia in nearshore areas, particularly along the south shore line, is caused by southwesterly winds that are common during the summer months which produce a thin hypolimnion along the south shore during most years and leads to the rapid development of hypoxia in this area (Rowe et al., 2019)

During 2015 and 2016, the sampling date ranges of nearshore and offshore loggers overlap for less than two months, starting from late July and ending in late September, thus hypoxia development and dissipation are not as completely captured as in 2014. In 2015 (Figure 8b), hypoxia was already present by late July along

(a) 2014

(b) 2015

(c) 2016
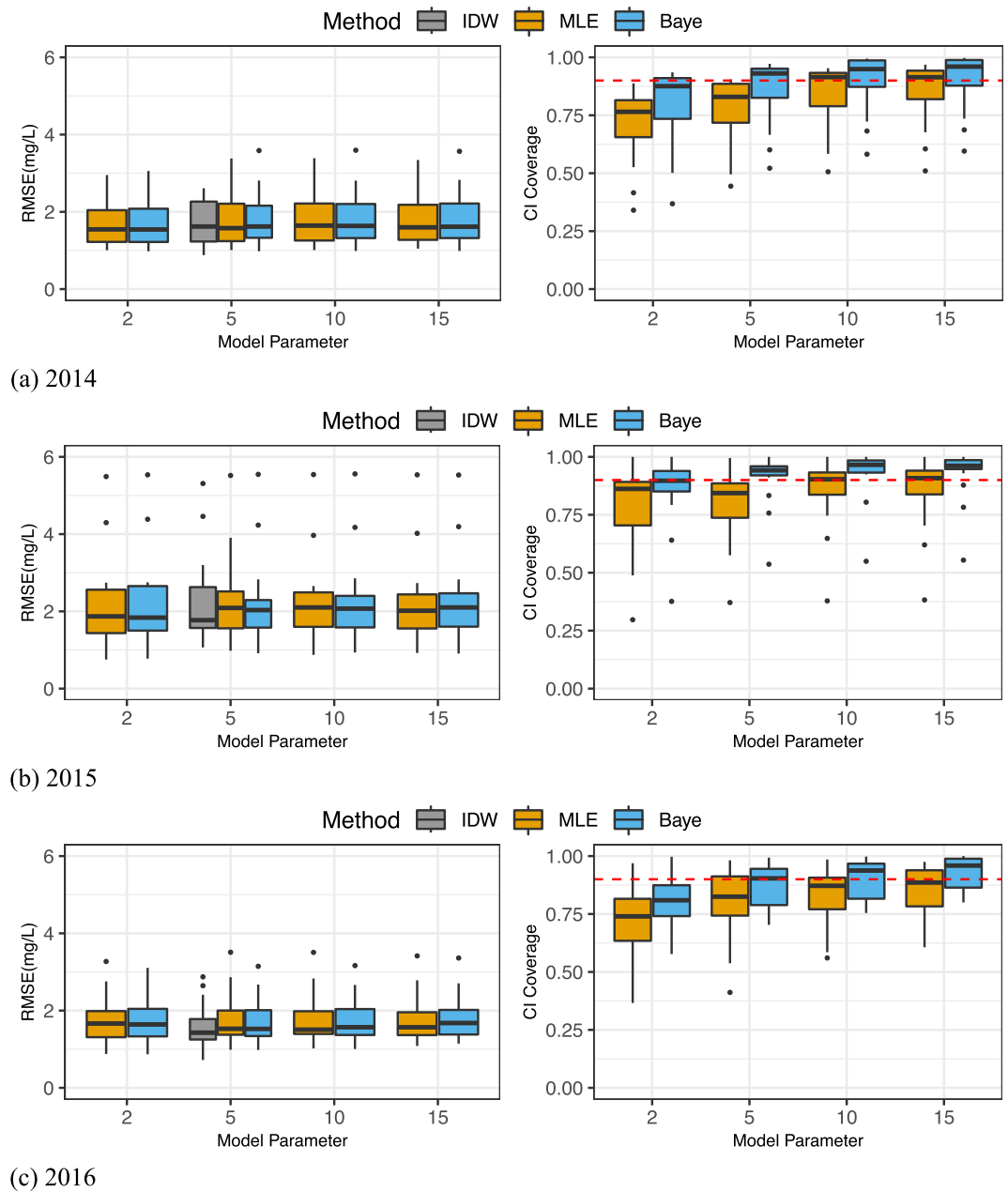
Figure 6. Box plots of cross-validation results for hourly aggregated data. Left: Root-mean-square error (RMSE) with different methods and model parameters in each year. Right: Coverage of 90% confidence intervals (CI) with MLE and Bayesian method. The box in the boxplot shows the first (Q1) and third quantile (Q3) and the black line shows the median value. Red dashed line is CI = 0.9. The outliers are the values beyond Q3+1.5IQR or Q1-1.5IQR, where the interquartile range IQR = Q3-Q1. X axis is the number of bases used if the method is MLE or Bayes or, for inverse distance weighting (IDW), the number of nearest points included in the interpolation.

the south shore and then expanded quickly by the end of the month. DO concentrations in the eastern zone increased briefly in early August but then declined. The maximum extent of the hypoxic zone occurred around mid-August, when almost the entire interpolated area was hypoxic. After mid-August, there were minor fluctuations in DO along the western edge of the interpolated area and along the northern shore, then the entire area began returning to normoxic (i.e., non-hypoxic) conditions by mid-September. DO first began to increase to normoxic DO levels in northern zones, which remained in that condition, indicating
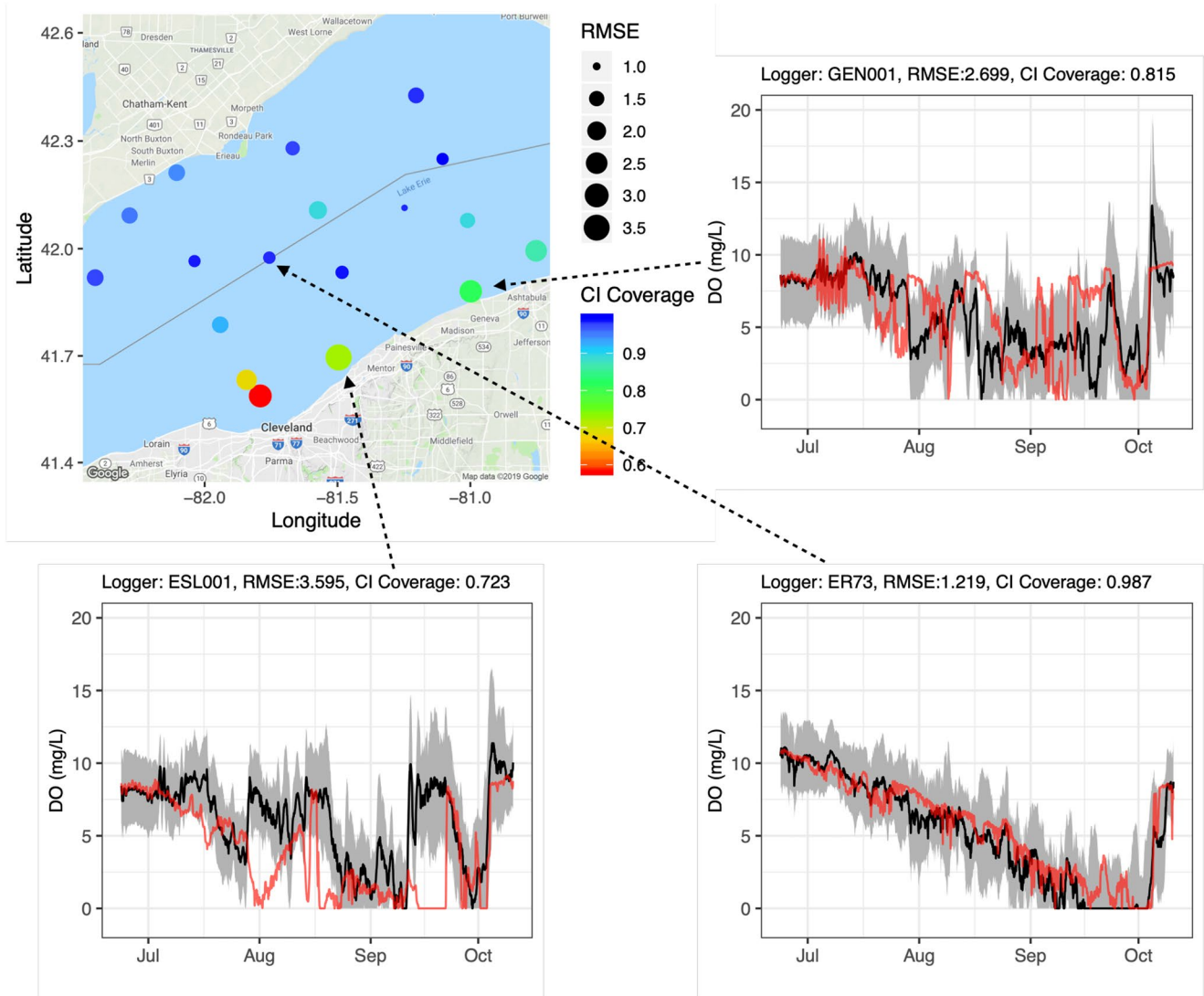
Figure 7. Cross-validation results from hourly aggregated data in 2014 with EOF kriging interpolation using the Bayesian framework and $r = 10$. In the top left, each dot indicates the sampling location, with the size representing root-mean-square error (RMSE) and the color representing confidence intervals (CI) coverage ratio. Other plots represent cross-validation predictions on specific loggers, where red is the observed data, black is the median values of all simulations, and gray areas are the 90% confidence intervals.

that the thermocline had dropped below the nearshore bathymetry and lake fall turnover would begin shortly.

During the summer of 2016 (Figure 8c), generally low DO conditions appeared along the north and western areas in late July, yet higher DO existed at the north western nearshore loggers. Hypoxia then quickly expanded across the whole central basin. The whole central basin remained hypoxic during most of September, when the loggers were recovered.

Zhou et al. (2013) and Burns et al. (2005) noted that in August, hypoxia is most common in the western and middle northern parts of the central basin, and then it expands east in September. Similar patterns exist in our data: The hypoxia gradually expands to the east, but with variations in locations. Rowe et al. (2019)'s mechanistic models suggest the spatial and temporal variance of hypoxia are largely due to physical processes, such as wind induced coastal upwelling (more detail in Section 5) and meteorological forcing that drives the thermocline depth. Another important factor to explain the retreat and progress of hypoxia, especially in offshore areas, is the internal current in the hypolimnion and strong wind induced vertical
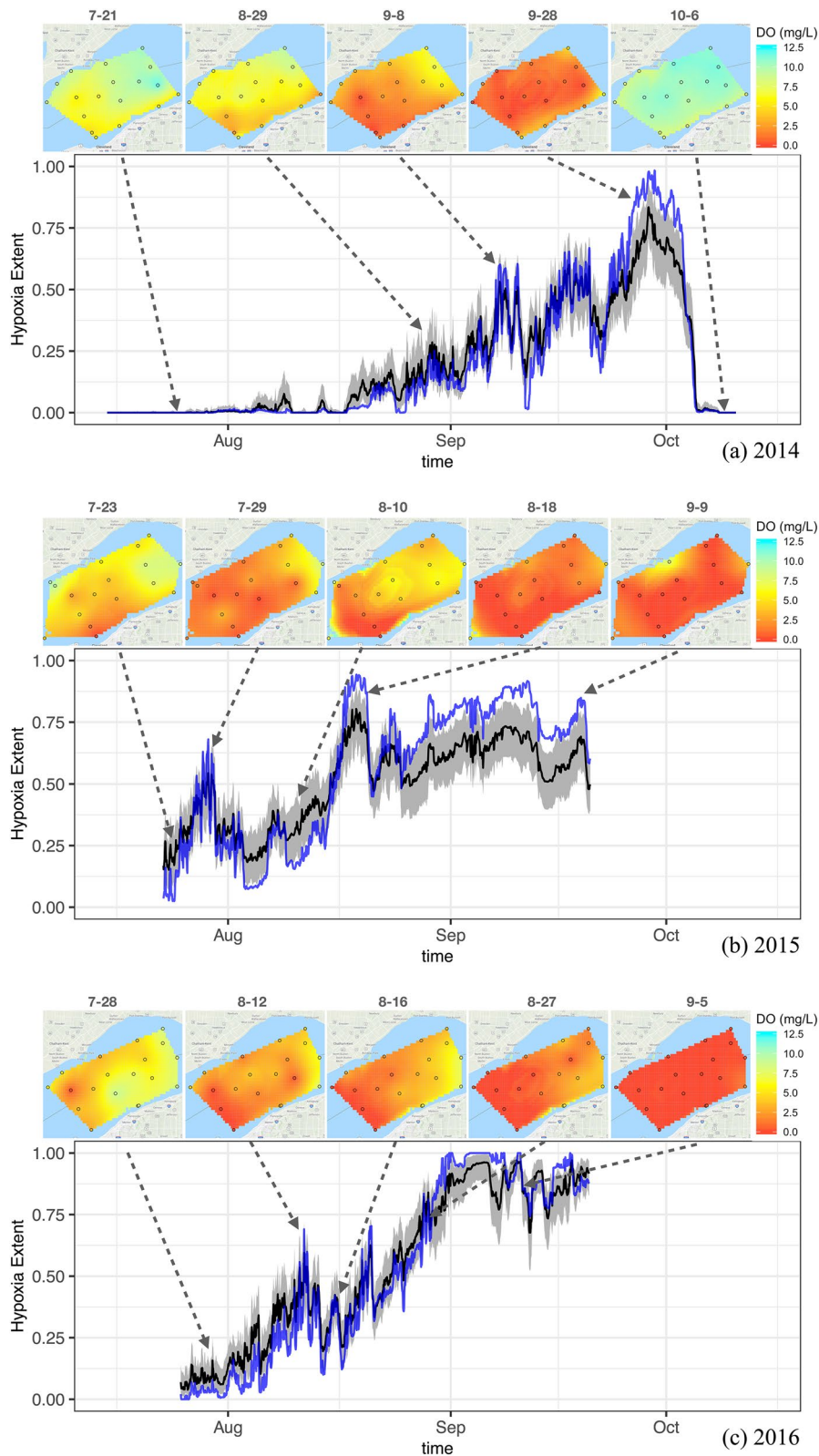
Figure 8.

mixing, which can cause hypolimnion DO changes in short time scales from hours to days (Rao et al., 2008). Varying sediment oxygen demand (SOD) may also contribute to differences in oxygen depletion at different locations (Lashaway & Carrick, 2010; Rowe et al., 2019).

During the periods with the largest hypoxia extents (late September 2014, mid-August 2015, and early September 2016), more than 75% of the interpolated areas had less than 2 mg/L DO. The uncertainty in hypoxia extent (range of confidence interval) was as large as 25% for both years (~1,700 km^2 in 2014, 2200 km^2 in 2015, and 1,900 km^2 in 2016).

Compared with the EOF Bayesian kriging approach, results from IDW interpolation at each time step have larger fluctuations that predict higher hypoxia extent during hypoxic events and lower extent when DO concentrations are recovering, generating a less smooth time series.

5. Discussion

The framework presented here provides a practical approach to hypoxia estimation that can quickly conduct spatio-temporal interpolations with estimation uncertainty. Building spatio-temporal kriging models requires stationarity in both spatial and temporal domain (Gneiting et al., 2006), which may not be easy to satisfy in such a highly dynamic system. Another possible approach is to interpolate at every time step, as the data set is spatially sparse but temporally dense (hourly resolution after aggregation). One can fit a single spatial variogram model and apply it at every time step, yet the spatial correlation structure may change seasonally. If we build different kriging models at every timestep, it is computationally burdensome.

EOF approach simplifies the spatio-termporal problem into spatial problems and temporal nonstationary patterns are decomposed into basis functions. Modeling temporal correlation is not needed. We use complex model structure (Bayesian kriging) on spatial interpolation of basis coefficient and only k kriging models are needed where k is the number of basis functions. Each kriging model is built separately so that k spatial correlation structure, each for one temporal pattern, are modeled and mixed in the DO interpolation. And we use the simple method of IDW to handle the non-important residuals. Overall, it considers different spatial structures along time dimension without modeling every time step, thus reducing the computational burden. One limitation of this framework is that we need to tune the number of basis functions. With more basis functions, more uncertainty will be added within the kriging framework.

Similar to other geostatistical approach, the performance of the presented interpolation methods depends on: (a) whether the processes are consistent across space, that is, the spatial stationarity assumption; and (b) how well the model captures the statistical correlations as the estimation of the variogram is critical for DO interpolation and hypoxia estimation, in both accuracy and variance. The results of this work show how the complexities of lake processes affect interpolation performance. For (a) DO patterns in nearshore areas are changing rapidly across space and time, but the statistical models still assume similarity to nearby loggers (i.e., stationarity). As nearshore and offshore patterns are quite different, a single variogram estimated using all data cannot consider such non-stationarity. Our approach is a compromise for simplicity. For (b), with limited data, the variogram model estimated have large uncertainty (Figure 4), which reflects variations in nearshore-nearshore, offshore-offshore, and nearshore-offshore patterns. The Bayesian structure accounts for such variations and provides additional uncertainty for interpolated DO, as shown by cross-validation results (i.e., high CI coverages). Notice that generally we should aim for accurate variance estimation, that is, neither overestimating nor underestimating. In our hypoxia cases, overestimation of variance estimates the hypoxia extent in worst cases, which is preferred.

In particular, nearshore patterns can be heavily influenced by seiche events. In Figure 9, wave heights in 2014 from one buoy on the south shore, north of Cleveland, OH are compared with DO data from a nearby logger. Spikes of DO are often correlated with higher wave heights (e.g., in late September and early October). A possible explanation is that when the southwest winds push warm surface water toward the

Figure 8. Dissolved oxygen (DO) interpolation results (above) reconstructed from expected prediction coefficients for multiple dates at 00:00 EDT, and hypoxia extent ratio (below, compared to the total interpolation area) with DO < 2 mg/L in 2014 (a), 2015 (b) and 2016 (c). The circles are the sampling locations. Shaded areas show the 90% confidence interval (from 5% to 95% quantiles) generated by conditional simulations. Black lines indicate median values from conditional simulations generated with the Bayesian approach. Blue lines give hypoxic area estimates calculated by inverse distance weighting (IDW). The total interpolation areas in 2014, 2015, and 2016 are approximately 6,985, 8,917, and 7,543 km^2, respectively.
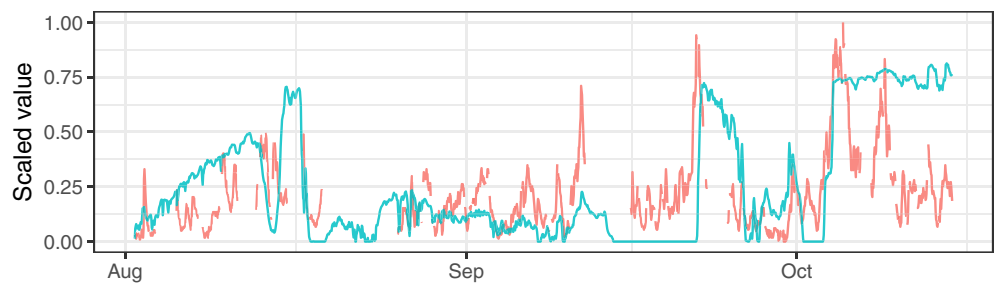
Figure 9. Comparison between nearshore dissolved oxygen (DO) patterns measured at station ESL001 (blue line) and wave heights at buoy 45,164 (red line) in 2014. Both are scaled to (0, 1) values. Buoy location and wave height data from the National Oceanic and Atmospheric Administration (https://www.ndbc.noaa.gov/station_page.php?station=45164).

shoreline due to Coriolis forces or Ekman transport (Rao & Schwab, 2007) and deepen the thermocline (i.e., downwelling events) below the sampling location, the whole water column mixes and bottom DO is replenished. On the contrary, when northeast winds drive surface water away from the southern shoreline, upwelling events may bring low DO water to nearshore areas, causing a sharp drop in DO concentrations.

To improve model performance, one could increase the number of loggers in the network so that more data points will contribute to a more confident estimation of the variogram parameters, but non-stationarity will still remain. With more data points, Bayesian estimation will just provide a maximum likelihood estimation that is a more accurate "average" of nearshore and offshore variogram patterns. However, increasing the density of loggers will make the interpolation more accurate, given that the distance between sampled locations will be smaller, which leads to higher correlations in sampled values and lower variances.

However, deploying more loggers increases both capital costs (e.g., one HOBO DO logger costs $1,250 as of September 2019 [https://www.onsetcomp.com/products/data-loggers/u26-001], plus a similar cost for initial deployment and operational and management costs, which are approximately $3,000 per year at this site). With a limited number of loggers, optimizing the sampling locations without compromising interpolation accuracy is very important to understanding the lake DO dynamics and maximizing monitoring yields. To this end, prediction variance is computed for each spatial point (the average across time of the standard deviation across 1,000 simulations at each time step, Appendix C). Both the cross-validation results and prediction variance maps suggest that more loggers should be placed in nearshore areas. Placing additional loggers at locations that are difficult to predict (large RMSE and low CI coverage) by other loggers in the cross-validation results, or at locations where model has the worst inferring confidence (Figure C1) should improve DO estimation and prediction in these areas. On the other hand, as our interpolation areas are covered by a convex hull defined by the outermost loggers for reliable interpolation rather than extrapolation, it is also possible that additional loggers could be valuable for extending the convex hull and estimating hypoxia in more of the entire central basin.

Besides deploying more loggers, approaches to mitigate stationarity could further improve the interpolation performance. Kriging within a moving neighborhood is one option so that variograms are built with a smaller neighborhood. However, this approach may not be suitable with limited loggers. Other external data and model outputs can be used as covariates to remove driving factors in the raw data (Matli et al., 2020; Murphy Rebecca et al., 2010). In our case, covariates can also be thermocline depth and whether the depth is below sensor depth or not, generated from mechanistic models. With thermocline predictions from the Great Lakes Coastal Forecasting Systems, one simple way to improve the model would be setting DO to saturated levels if the bathymetry of the interpolation location is above the thermocline in nearshore areas. With advanced mechanistic models that can predict bottom dissolved oxygen lake wide (Bouffard et al., 2013; Burns et al., 2005; Rao et al., 2008; Rowe et al., 2019), the framework proposed here may also be used to account for mechanistic model errors (residuals), which could be stationary in space and time.

It should be noted that the above discussion is a retrospective analysis based on three years in which the sampling data are already known. With more data in the future, better generalizations can be made as to which locations have consistently homogeneous or heterogeneous patterns that may need fewer or more loggers, respectively, for accurate interpolation. In addition, data from more years are required to study long-term trends and prevailing seasonal patterns of hypoxia extent.

6. Conclusions

In this study, we develop a spatio-temporal interpolation method and software for estimating DO levels from logger data, including conditional simulations to estimate hypoxia extent with uncertainty. DO sensor data from 2014 to 2016 in the central basin of Lake Erie are analyzed as a case study and a Web application is developed using R Shiny. Three different spatio-temporal DO interpolation models based on IDW, EOF MLE kriging, and EOF Bayesian kriging are compared via cross-validation. The results show that these methods have similar RMSE but Bayesian kriging estimates prediction uncertainty more accurately than the other methods.

Applying the models within the central basin of Lake Erie shows that hypoxia developed differently in each year studied. In 2014, hypoxia started in early August while in 2015 and 2016, hypoxia began in July. The peak hypoxia extent occurred in late September 2014, mid-August 2015, and early September 2016. With current logger placement and numbers, the uncertainty of the hypoxia extent can be as large as 25% of the interpolation areas. Further discussion of cross-validation results and interpolation variances suggests that some offshore loggers may be redundant. For example, removing one offshore logger with small RMSE in the cross validation (Figure 7) may not cause much differences in hypoxia estimations as this logger's DO measurements can be inferred from other loggers. However, given the variable dynamics of the lake from year to year, to avoid loss of information, researchers may want to check the cross-validation error of this logger over a longer period before considering removal and also consider other data availability at offshore sampling locations where EPA-GLNPO conducts Lake CTD (conductivity, temperature, depth) profiling every year. On the other hand, more loggers are clearly needed in the nearshore areas to improve estimations of hypoxia extent.

It should be noted that the interpolation framework and web application can also be modified for use with other water quality monitoring data, such as bottom temperature, as well as in other lake systems (more details in the Supporting Information S1). The DO interpolation values generated by our framework could also be incorporated into related studies about fish movements (Chamberlin et al., 2020; Kraus et al., 2015) and nutrient loadings (Rucinski et al., 2014) to better understand lake systems and guide lake management policies. Furthermore, the interpolated values can be compared against outputs from mechanistic models. The locations of the large discrepancies could indicate potential model inaccuracy or important physical mechanisms that are not captured by the statistical models.

Appendix A

The parameters that define the prior distributions in EOF Bayesian kriging interpolation were initially chosen based on the default parameters used in the "geoR" package (the "base" case, Table 1). To explore the effects of this selection, we also ran the models using different prior parameters. These alternative prior parameter sets are defined in Table A1. Parameter Sets 1 to 3 explore different priority distributions for the Range (ϕ). Parameter Sets 4 to 7 explore nugget effects. Parameter Set 8 changes the variogram model to another widely used "spherical model." Parameter Sets 9 and 10 change the priority distribution of the Partial Sill (σ^2).

We tested these new parameters with the same cross-validation approach described in Section 4.2 on the 2014 data. For logger i, we generate interpolation values and variance with base and alternative parameters, utilizing data from all other loggers. Then the RMSE and CI coverage for the logger i are calculated. The boxplots in Figure A1 summarize the differences of RMSE and CI coverage from the same logger between the base and alternative parameters in Table A1 at the same logger.

Given that the base RMSE is approximately 2 mg/L and CI coverage is about 0.9, the results show that the parameters for prior distributions do not have significant effects on the interpolations, with small differences

Table A1
*Alternative Prior Distributions for Bayesian Kriging*

| Parameter set #	Changed parameter	Parameter setting
1	Range (ϕ)	Reciprocal distribution from 20 to 70 km with 5 km intervals
2	Range (ϕ)	Uniform distribution with a sequence of 51 values from 0 to 2 times the maximum distance between the data locations
3	Range (ϕ)	Squared reciprocal distribution from 20 to 70 km at 5 km intervals

Table A1
Continued

Parameter set #	Changed parameter	Parameter setting
4	Nugget (τ^2)/Partial sill (σ^2)	Fixed at 0.1
5	Nugget (τ^2)/Partial sill (σ^2)	Fixed at 0.2
6	Nugget (τ^2)/Partial sill (σ^2)	Uniform distribution from 0 to 1 with an interval of 0.1
7	Nugget (τ^2)/Partial sill (σ^2)	Reciprocal distribution from 0 to 1 with an interval of 0.1
8	Variogram model	Spherical model
9	Partial sill (σ^2)	Scaled inverse chi-squared distribution with 1° of freedom
10	Partial Sill (σ^2)	Scaled inverse chi-squared distribution with 3° of freedom

in RMSE and CI coverage. Furthermore, no alternative parameters are able to consistently achieve better results on every logger in the cross-validation. Parameter sets 4 to 7 explore scenarios where measurement noise and small scale variation exists (nugget $[\tau^2] > 0$); these scenarios show inconsistent effects, with interpolation on some loggers performing better and on other loggers performing worse. Therefore, we conclude that the base set of parameters is sufficient for estimation.
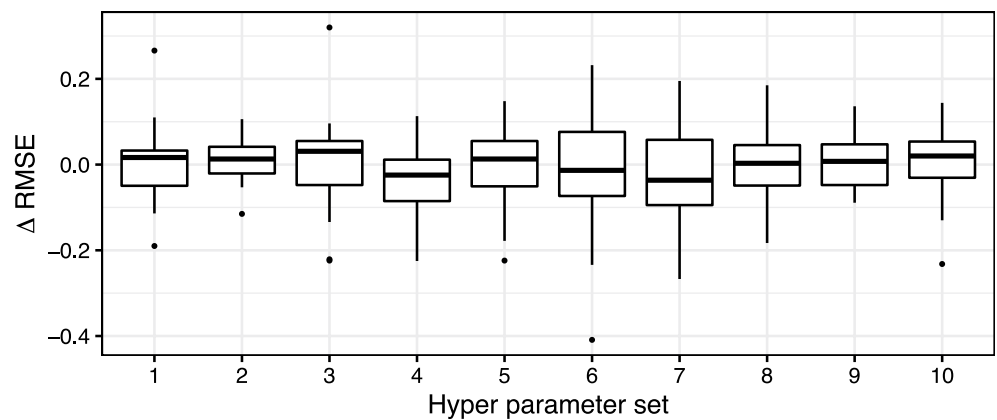
(a) RMSE differences (positive values indicate that the base parameter has lower RMSE and is better)
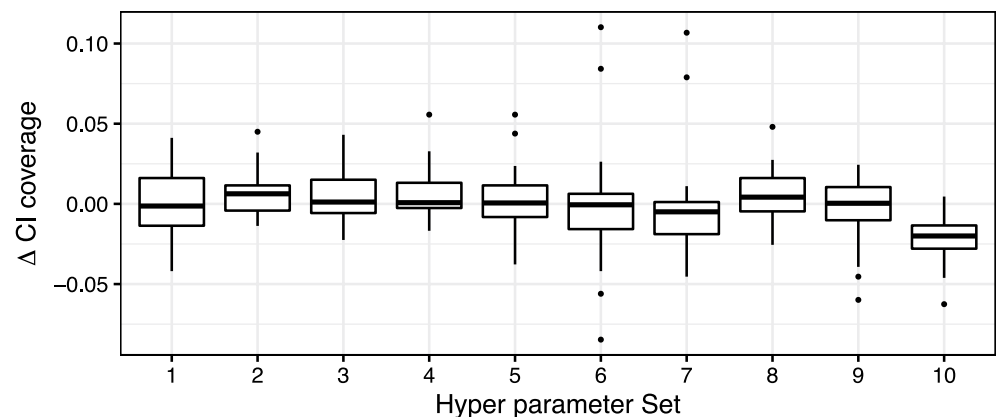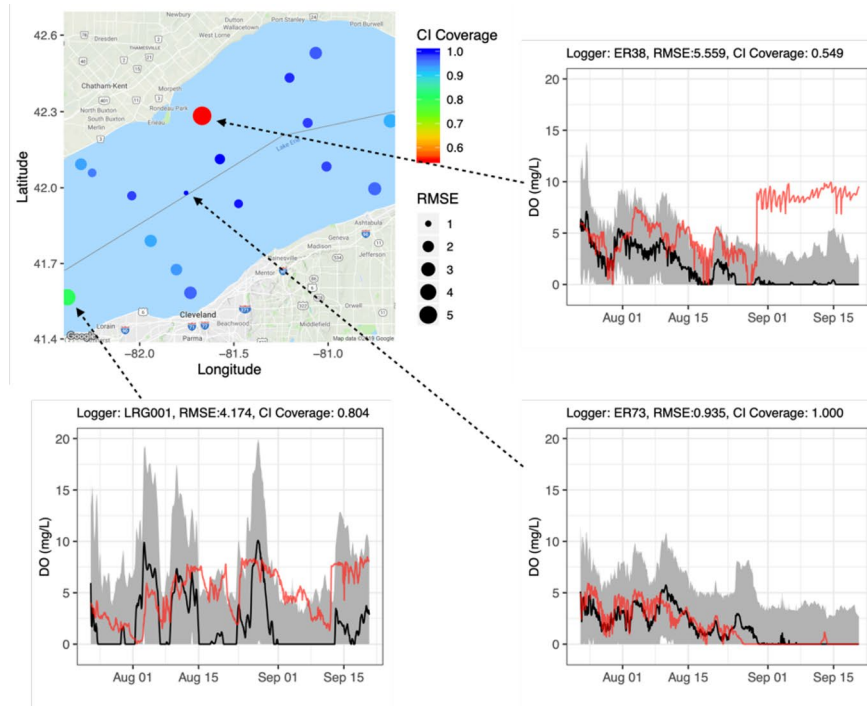

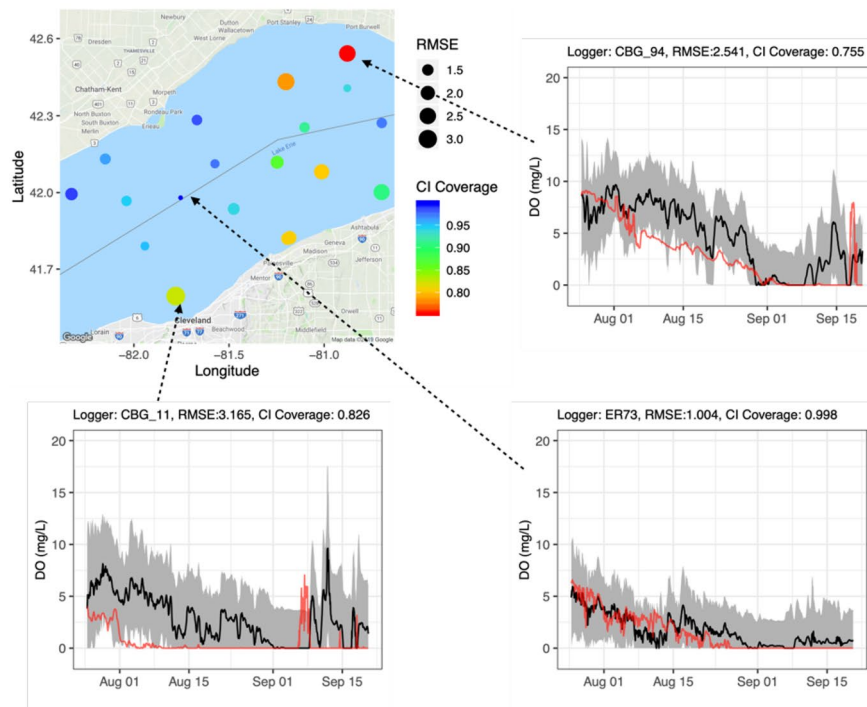(b) CI Coverage differences (positive values indicate that the alterative parameter has higher coverage and is better)

Figure A1. Root-mean-square error (RMSE) (a) and confidence intervals (CI) coverage (b) differences with alternative prior distributions.

Appendix B

Figure B1 provides CV results for 2015 and 2016.

(a) RMSE and CI for loggers in 2015

(b) RMSE and CI for loggers in 2016

Figure B1. Cross-validation results in hourly aggregated data in (a) 2015 and (b) 2016. In the top left, each dot indicates the sampling location, with size representing root-mean-square error (RMSE) and color representing confidence intervals (CI) coverage ratio. Other plots represent cross-validation predictions on specific loggers, where red is the observed data, black is the median values of all simulations, and gray areas are the 90% confidence intervals.

Appendix C

Figure C1 provides interpolation of temporally averaged standard deviations in 2014, 2015, and 2016.
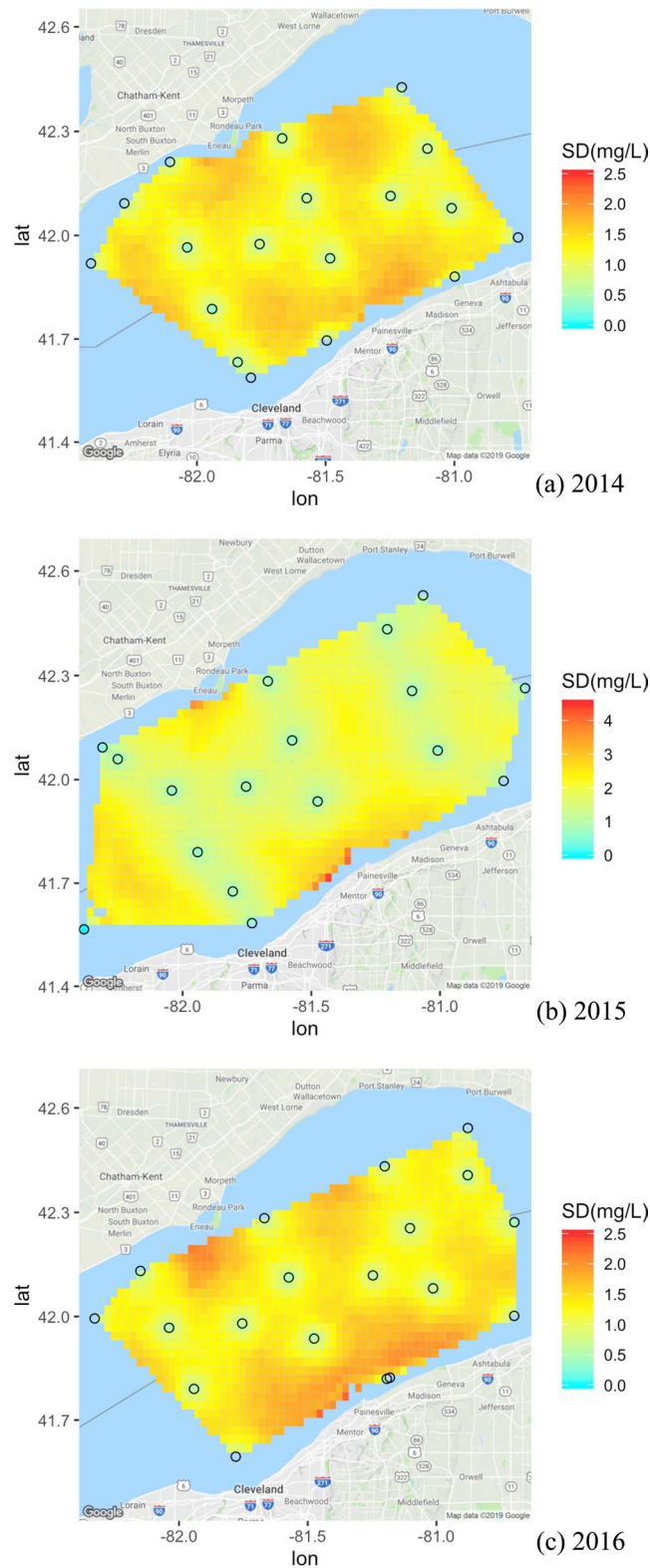

Figure C1. Interpolation of temporally averaged standard deviations in 2014, 2015, and 2016.

Data Availability Statement

We provide algorithm code and data in the Supporting Information S1. They are also publicly available at https://eisalab.github.io/Hypoxia-Interpolator/. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Anderson, J. E., El-Shaarawi, A. H., Esterby, S. R., & Unny, T. E. (1984). Dissolved Oxygen Concentrations in Lake Erie (U.S.A.-Canada): 1. Study of Spatial and Temporal Variability Using Cluster and Regression Analysis. *Journal of Hydrology*, *72*(3), 209–229. https://doi.org/10.1016/0022-1694(84)90081-7

Bakar, K. S., & Sahu, S. K. (2015). sptimer: Spatio-temporal bayesian modelling using R. *Journal of Statistical Software*, *63*(15), 1–32. https://doi.org/10.18637/jss.v063.i15

Bocaniov, S. A., & Scavia, D. (2016). Temporal and spatial dynamics of large lake hypoxia: Integrating statistical and three-dimensional dynamic models to enhance lake management criteria. *Water Resources Research*, *52*(6), 4247–4263. https://doi.org/10.1002/2015wr018170

Bouffard, D., Ackerman, J. D., & Boegman, L. (2013). Factors affecting the development and dynamics of hypoxia in a large shallow stratified lake: Hourly to seasonal patterns. *Water Resources Research*, *49*(5), 2380–2394. https://doi.org/10.1002/wrcr.20241

Breitburg, D., Levin, L. A., Oschlies, A., Gregoire, M., Chavez, F. P., Conley, D. J., et al. (2018). Declining Oxygen in the Global Ocean and Coastal Waters. *Science*, *359*(6371). https://doi.org/10.1126/science.aam7240

Burns, N. M., Rockwell, D. C. P. E. B. , D. M. D., JanCiborowski, J. H., Dolan, D. M., & Ciborowski, J. J. H. (2005). Trends in temperature, secchi depth, and dissolved oxygen depletion rates in the Central Basin of Lake Erie, 1983–2002. *Journal of Great Lakes Research*, *31*(January), 35–49. https://doi.org/10.1016/s0380-1330(05)70303-8

Chamberlin, D. W., Knight, C. T., Kraus, R. T., Gorman, A. M., Xu, W., & Collingsworth, P. D. (2020). Hypoxia augments edge effects of water column stratification on fish distribution. *Fisheries Research*, *231*(November), 105684. https://doi.org/10.1016/j.fishres.2020.105684

Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2017). *shiny: Web application framework for R*.

Cressie, N. A. C. (Ed.), (1993). *Spatial prediction and kriging. In: Statistics for spatial data, 105–209. Wiley Series in Probability and Statistics*. Hoboken: John Wiley & Sons, Inc.

Diaz, R. J. (2001). Overview of hypoxia around the world. *Journal of Environmental Quality*, *30*(2), 275–281. https://doi.org/10.2134/jeq2001.302275x

Diggle, P. J., & Ribeiro, P. J. (2002). Bayesian inference in gaussian model-based geostatistics. *Geographical and Environmental Modelling*, *6*(2), 129–146. https://doi.org/10.1080/1361593022000029467

Fang, S., Del Giudice, D., Scavia, D., Binding, C. E., Bridgeman, T. B., Chaffin, J. D., et al. (2019). A space-time geostatistical model for probabilistic estimation of harmful algal bloom biomass and areal extent. *The Science of the Total Environment*, *695*(December), 133776. https://doi.org/10.1016/j.scitotenv.2019.133776

Gneiting, T., Genton, M., & Guttorp, P. (2006). Geostatistical Space-Time Models, Stationarity, Separability, and Full Symmetry. In *C&H/CRC Monographs on Statistics & Applied Probability* (Vol. 151–75). Chapman and Hall/CRC. https://doi.org/10.1201/9781420011050.ch4

Gräler, B., Pebesma, E., & Heuvelink, G. (2016). Spatio-Temporal Interpolation Using Gstat. *The R Journal*. https://doi.org/10.32614/rj-2016-014. Retrieved from https://journal.r-project.org/archive/2016-1/na-pebesma-heuvelink.pdf

Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, *27*(9), 1119–1152. https://doi.org/10.1002/joc.1499

Jat, P., & Serre, M. L. (2016). Bayesian maximum entropy space/time estimation of surface water chloride in Maryland using river distances. *Environmental Pollution*, *219*(December), 1148–1155. https://doi.org/10.1016/j.envpol.2016.09.020

Jenny, J.-P., Francus, P., Alexandre, N., Lapointe, F., Perga, M.-E., Ojala, A., et al. (2016). Global spread of hypoxia in freshwater ecosystems during the last three centuries is caused by rising local human pressure. *Global Change Biology*, *22*(4), 1481–1489. https://doi.org/10.1111/gcb.13193

Kraus, R. T., Knight, C. T., Farmer, T. M., Gorman, A. M., Collingsworth, P. D., Warren, G. J., et al. (2015). Dynamic hypoxic zones in Lake Erie compress fish habitat, altering vulnerability to fishing gears. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(6), 797–806. https://doi.org/10.1139/cjfas-2014-0517

Lashaway, A. R., & Carrick, H. J. (2010). Effects of light, temperature and habitat quality on meroplanktonic diatom rejuvenation in Lake Erie: Implications for Seasonal Hypoxia. *Journal of Plankton Research*, *32*(4), 479–490. https://doi.org/10.1093/plankt/fbp147

Lindstrom, J., Szpiro, A. A., Sampson, P. D., Oron, A. P., Richards, M., Larson, T. V., & Sheppard, L. (2014). A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environmental and Ecological Statistics*, *21*(3), 411–433. https://doi.org/10.1007/s10651-013-0261-4

Mortimer, C. H. (1987). "Fifty years of physical investigations and related limnological studies on Lake Erie, 1928–1977. *Journal of Great Lakes Research*, *13*(4), 407–435. https://doi.org/10.1016/s0380-1330(87)71664-5

Murphy Rebecca, R., Curriero Frank, C., & Ball William, P. (2010). Comparison of spatial interpolation methods for water quality evaluation in the Chesapeake Bay. *Journal of Environmental Engineering*, *136*(2), 160–171. https://doi.org/10.1061/(asce)ee.1943-7870.0000121

Pihl, L., Baden, S. P., Diaz, R. J., & Schaffner, L. C. (1992). Hypoxia-induced structural changes in the diet of bottom-feeding fish and crustacea. *Marine Biology*, *112*(3), 349–361. https://doi.org/10.1007/bf00356279

Rao, Y. R., Hawley, N., Charlton, M. N., & Schertzer, W. M. (2008). Physical processes and hypoxia in the central basin of Lake Erie. *Limnology & Oceanography*, *53*(5), 2007–2020. https://doi.org/10.4319/lo.2008.53.5.2007

Rao, Y. R., Howell, T., Watson, S. B., & Abernethy, S. (2014). On hypoxia and fish kills along the north shore of Lake Erie. *Journal of Great Lakes Research*, *40*(1), 187–191. https://doi.org/10.1016/j.jglr.2013.11.007

Rao, Y. R., & Schwab, D. J. (2007). Transport and mixing between the coastal and offshore waters in the great lakes: A Review. *Journal of Great Lakes Research*, *33*(1), 202–218. https://doi.org/10.3394/0380-1330(2007)33[202:tambtc]2.0.co;2

Ribeiro, P. J., Jr., & Diggle, P. J. (2001). geoR: A Package for Geostatistical Analysis. *R News*, *1*(No 2), 1609–3631.

Rohith Reddy, M. V., Fang, S., Guinness, J., Rabalais, N. N., Kevin Craig, J., & Obenour, D. R. (2018). Space-time geostatistical assessment of hypoxia in the Northern Gulf of Mexico. *Environmental Science & Technology*, *52*(21), 12484–12493. https://doi.org/10.1021/acs.est.8b03474

Rohith Reddy, M. V., Laurent, A., Fennel, K., Craig, K., Krause, J., & Obenour, D. R. (2020). Fusion-based hypoxia estimates: Combining geostatistical and mechanistic models of dissolved oxygen variability. *Environmental Science & Technology*, *54*(20), 13016–13025. https://doi.org/10.1021/acs.est.0c03655

Rowe, M. D., Anderson, E. J., Beletsky, D., Stow, C. A., Moegling, S. D., Chaffin, J. D., et al. (2019). Coastal upwelling influences hypoxia spatial patterns and nearshore dynamics in Lake Erie. *Journal of Geophysical Research - C: Oceans*, *124*(8), 6154–6175. https://doi.org/10.1029/2019jc015192

Rucinski, D. K., DePinto, J. V., Scavia, D., & Beletsky, D. (2014). Modeling Lake Erie's hypoxia response to nutrient loads and physical variability. *Journal of Great Lakes Research*, *40*(Supplement 3), 151–161. https://doi.org/10.1016/j.jglr.2014.02.003

Scavia, D., David Allan, J., Arend, K. K., Bartell, S., Beletsky, D., Bosch, N. S., et al. (2014). Assessing and addressing the re-eutrophication of Lake Erie: Central basin hypoxia. *Journal of Great Lakes Research*, *40*(2), 226–246. https://doi.org/10.1016/j.jglr.2014.02.004

Scully, M. E. (2016). The contribution of physical processes to inter-annual variations of hypoxia in chesapeake bay: A 30-yr modeling study. *Limnology & Oceanography*, *61*(6), 2243–2260. https://doi.org/10.1002/lno.10372

Snepvangers, J. J. J. C., Heuvelink, G. B. M., & Huisman, J. A. (2003). Soil water content interpolation using spatio-temporal kriging with external drift. *Geoderma*, *112*(3), 253–271. https://doi.org/10.1016/s0016-7061(02)00310-5

Vanderploeg, H. A., Ludsin, S. A., Ruberg, S. A., Höök, T. O., Pothoven, S. A., Brandt, S. B., et al. (2009). Hypoxia affects spatial distributions and overlap of Pelagic Fish, Zooplankton, and Phytoplankton in Lake Erie. *Journal of Experimental Marine Biology and Ecology*, *381*(December), S92–S107. https://doi.org/10.1016/j.jembe.2009.07.027

Zhou, Y., Obenour, D. R., Scavia, D., Johengen, T. H., & Michalak, A. M. (2013). Spatial and temporal trends in Lake Erie hypoxia, 1987-2007. *Environmental Science & Technology*, *47*(2), 899–905. https://doi.org/10.1021/es303401b