https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

Biogeosciences Discussions

Hydrodynamic and Biochemical Impacts on the Development of Hypoxia in the Louisiana–Texas Shelf Part II: Statistical Modeling and Hypoxia Prediction

Yanda Ou[1] , Bin Li[2], Z. George Xue[1,3,4]

[1]Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, LA, 70803, USA.
[2]Department of Experimental Statistics, Louisiana State University, Baton Rouge, LA, 70803, USA
[3]Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, 70803, USA.
[4]Coastal Studies Institute, Louisiana State University, Baton Rouge, LA, 70803, USA

Correspondence to: Z. George Xue (zxue@lsu.edu)

Abstract. In this study, a novel ensemble regression model was developed for hypoxic area (HA) forecast in the Louisiana–Texas (LaTex) Shelf. The ensemble model combines a zero-inflated Poisson generalized linear model (GLM) and a quasi-Poisson generalized additive model (GAM) and considers predictors with hydrodynamic and biochemical features. Both models were trained and calibrated using the daily hindcast (2007–2020) by a three-dimensional coupled hydrodynamic–biogeochemical model embedded in the Reginal Ocean Modeling System (ROMS). A promising HA forecast is provided by the ensemble model with a low RMSE (3,204 km^2), a high R^2 (0.8005), and a precise performance in capturing hypoxic area peaks in the summers. To test its robustness, the model was further applied to a global forecast model and produces HA prediction from 2019 to 2020 with the adjusted predictors from the HYbrid Coordinate Ocean Model (HYCOM). Predicted HA shows a high agreement with the ROMS hindcast time series (RMSE=4,571 km^2, R^2=0.8178). Our model can also predict the magnitude and onsets of summer HA peaks in both 2019 and 2020 with high accuracy. To the best of our knowledge, this ensemble model is by far the first one providing fast and accurate daily HA predictions for the LaTex Shelf while considering both hydrodynamic and biochemical effects. This study demonstrates that it is feasible to perform regional ocean HA prediction using global ocean forecast.

1 Introduction

The Louisiana–Texas (LaTex) Shelf has become a center of hypoxia (bottom dissolved oxygen, DO<2 mg L-1) study since the 1980s (Rabalais et al., 2002; Rabalais et al., 2007a; Justić and Wang, 2014). Regular mid-summer Shelfwide cruises documented that the area and volume of hypoxic bottom water could reach up to 23,000 km^2 and 140 km^3, respectively (Rabalais and Turner, 2019; Rabalais and Baustian, 2020). The aquatic environments, fisheries, and coastal economies are under threat of recurring hypoxia in summer (Chesney and Baltz, 2001; Craig and Bosman, 2013; De Mutsert et al., 2016; LaBone et al., 2020; Rabalais and Turner, 2019; Rabotyagov et al., 2014; Smith et al., 2014). Water column stratification and sediment oxygen consumption (SOC) are two main factors regulating the formation, evolution, and deconstruction of bottom

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

hypoxia from mid-May through mid-September (Bianchi et al., 2010; Conley et al., 2009; Fennel et al., 2011, 2013, 2016; Feng et al., 2014; Hetland and DiMarco, 2008; Justić and Wang, 2014; Laurent et al., 2018; McCarthy et al., 2013; Murrell and Lehrter, 2011; Rabalais et al., 2007b; Wang and Justić, 2009; Yu et al., 2015). However, prevailing prediction models for the hypoxic area (HA) rely most on nutrient-induced mechanisms rather than the hydrodynamic features. Turner et al. (2006) built a multiple linear regression model for summer HA prediction using the annual and May nitrogen flux (nitrate+nitrite) of the Mississippi River as the predictors. The model provides a robust annual prediction when no strong wind was present but underestimates the HA in windy years. Obenour et al. (2015) modeled HA using the empirical relationship between HA and bottom DO concentration derived by a Bayesian biophysical model. Their model accounted for primary biophysical processes solved for steady-state conditions, water transport, May total nitrogen loads by rivers, and parameterized water reaeration. Katin et al. (2021) further adjusted the Bayesian model by taking into account river flows, riverine bioavailable nitrogen loadings, and wind velocity in both summer (June–September) and non-summer (November–May) months. Summer riverine inputs are projected using non-summer riverine variables, river basin precipitation, and river basin temperature, while, however, summer wind velocity is resampled from historical records from 1985 to 2016. Therefore, the model is known as a pseudo-forecast model since predictors in future stages only include riverine inputs. The pseudo-forecast model explains 71 % and 41 %–48 % of the variability in hindcast (Del Giudice et al., 2020) and geostatistically estimated HA (Matli et al., 2018), respectively. Another Bayesian model was proposed for summer bottom DO concentration prediction taking account of May total nitrogen loads, distance from the Mississippi River mouth, and downstream velocity (Scavia et al., 2013). The summer HA is determined by hypoxic length (HA=57.8 hypoxic length) derived from summer bottom DO concentration. The model explains 69 % of the variability in observed HA by the mid-summer Shelfwide cruises. Different from linear regression and Bayesian analysis, Laurent and Fennel (2019) developed a weighted mean forecast method calibrated on the May nitrate loads and three-dimensional hindcast simulations (1985–2018). Once calibrated, the model requires the May nitrate loads for the forecast year as the only input to produce the seasonal forecast. The model can explain up to 76 % of the year-to-year variability of the HA observation. However, the model is not favorable for years with strong wind events during summer.

These above-mentioned models share some similar shortages: (1) The effects of water column stratification are not included or only partially considered even though stratification is documented as a crucial factor in regulating HA variability. (2) The information of future conditions is limited although some models are built upon multiple predictors, thus these forecast models are indeed “pseudo-forecast” ones. (3) Most models only capture year-to-year HA variability and fail whenever winds are strong in summers. According to the hindcast results by our three-dimensional coupled hydrodynamic–biogeochemical model described in the accompanying paper (Part I), monthly and daily variabilities of HA cannot be neglected before and after strong wind events. In this study, we aimed to provide a new technique in HA prediction considering both stratification and biochemical effects and executing daily forecasts based on the forecasts of selected predictors. An important assumption is that the future conditions of predictors are accessible. Indeed, it can be fulfilled by using global forecast products such as the HYbrid Coordinate Ocean Model (HYCOM), which provides operational hydrodynamics forecasts for up to one week (eight

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

65 days). In the accompanying paper (Part I), we demonstrated that the hypoxic volume and the bottom HA over the LaTex Shelf
66 are highly correlated. The former can be reproduced by the latter alone with a quadratic relationship. Thus, in this study, we
67 focused on bottom HA predictions. The rest of the paper is organized as follows: a detailed description of methods and data is
68 given in section 2. The employment of generalized linear models (GLMs) and generalized additive models (GAMs) is given
69 in section 3. The ensemble HA prediction and its application using the global HYCOM is discussed in section 4.

70 **2 Methods**

71 **2.1 Data descriptions**

72 We adapted a three-dimensional coupled hydrodynamic–biogeochemical model embedded in the framework of the Regional
73 Ocean Modeling System (ROMS) on the platform of Coupled Ocean–Atmosphere–Wave–Sediment Transport modeling
74 system (COAWST, Warner et al., 2010) to the GoM (Gulf–COAWST, for detailed descriptions, validations, and results of the
75 numerical model see Part I). Numerical hindcasts (hereafter denoted as ROMS hindcasts or ROMS simulations) are output
76 daily from 1 January 2007 to 26 August 2020 and spatially averaged over the LaTex Shelf. In this study, we aim to produce a
77 fast and accurate daily forecast of the shelf HA using models trained from the ROMS outputs.

78 **2.1.1 Hydrodynamic-related predictors**

79 Both water stratification and bottom biochemical processes modulate the variability of bottom DO concentration in the LaTex
80 Shelf. Potential energy anomaly (PEA, in J m^{-3}) is introduced as an estimate of water column stratification according to:
81
82 $$PEA = \frac{1}{H}\int_{-h}^{\eta}(\bar{\rho} - \rho)gzdz, \quad (1)$$
83
84 where ρ is water density profile (estimated by water temperature and salinity profiles) over water column of depth $H = h + \eta$,
85 h is the location of the bed, η is water surface elevation, g is the gravitational acceleration (9.8 m s^{-2}), z is the vertical axis, $\bar{\rho}$
86 is the depth-integrated water density given by $\bar{\rho} = \frac{1}{H}\int_{-h}^{\eta}\rho dz$ (Simpson and Hunter, 1974; Simpson et al., 1978; Simpson,
87 1981; Simpson and Bowers, 1981). The PEA represents the amount of energy per volume to homogenize the entire water
88 column (Simpson and Hunter, 1974). Thus, a greater PEA value represents a more stratified water column. As a river-
89 dominated area, water stratification in the LaTex Shelf is highly affected by freshwater-induced buoyancy from the Mississippi
90 and Atchafalaya Rivers. Sea surface salinity (SSS) is a good proxy in representing the distribution and variability of river
91 freshwater across the shelf. Indeed, the correlation of regionally averaged PEA and SSS is significantly high up to -0.88
92 (p<0.001; Figure 1a) which emphasizes the importance of freshwater-induced stratification. Therefore, we considered SSS as
93 another candidate predictor besides PEA.
94

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022


Biogeosciences Discussions Open Access EGU

In the meantime, surface heating and wind mixing are other two factors influencing water stratification (Simpson and Hunter, 1974; Simpson et al., 1978) and can be quantified as follows:

$$\frac{d(PEA)}{dt} = \frac{\alpha g h}{2c} Q - \delta k_a \rho_a W^3, \quad (2)$$

where Q is the rate of surface heat input, α is the volume expansion coefficient, c is water specific heat capacity, δ is coefficient of wind mixing, k_a is drag coefficient, ρ_a is humid air density near the sea surface, and W is the wind speed near sea surface. The first term on the right-hand side of Eq. (2) represents the rate of change of water stratification due to surface heating, while the second term is the rate of working by wind stress contributing negatively to water stratification. Therefore, the heat-induced change of PEA is proportional to the product of heat input and water depth, which is,

$$\mathrm{d}(PEA)_{heat} \propto Qh, \quad (3)$$

The total net heat flux, a sum of net shortwave and net longwave radiation flux, is derived from the National Centers for Environmental Prediction Climate Forecast System (NCEP) Reanalysis (CFSR) 6-hourly products (Saha et al., 2010; 2011) in this study. The term (Qh) is added to the candidate list of predictors and is denoted as $\mathrm{PEA_{heat}}$ (heat-induced PEA changes) for simplification.

Daily variability of term ($\delta k_a \rho_a W^3$) is dominated by that of W^3, since the ρ_a fluctuates much less than the W^3 in a daily scale (Figure A1). We obtained the ρ_a according to (Picard et al., 2008) :

$$\rho_a = \frac{pM_d}{ZRT}\left[1 - x_v\left(1 - \frac{M_v}{M_d}\right)\right], \quad (4)$$

where p represents the absolute air pressure, M_d (=28.96546 g mol^{-1}) is the molar mass of dry air, M_v (=18.01528 g mol^{-1}) is the molar mass of water vapor, Z indicates compressibility, R (=8.314472 J mol^{-1} K^{-1}) is the molar gas constant, T is thermodynamic temperature, x_v is the mole fraction of water vapor. We assumed that air parcels at the sea surface are ideal gases ($Z = 1$) and are always saturated with water vapor. Thus, x_v is a function of absolute air pressure (p) and saturation vapor pressure of water (p_{sat}) and can be calculated as follows:

$$x_v = \frac{p_{sat}}{p}, \quad (5)$$

According to the Tetens equation (Monteith and Unsworth, 2014), p_{sat} (in Pa) can be estimated for the following:

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

$$p_{sat} = 610.78e^{\frac{17.27(T-237.3)}{T}}, \quad (6)$$

Substitute Eqs. (5)–(6) to Eq. (4) with the assumption of $Z = 1$, we obtained air density as a function of both air pressure and air temperature in the following:

$$\rho_a = \rho_a(T,p) = \frac{pM_d}{ZRT}\left[1 - \frac{1}{p}\left(1 - \frac{M_v}{M_d}\right)e^{\frac{17.27(T-237.3)}{T}}\right], \quad (7)$$

The ρ_a is then estimated using sea surface air pressure and air temperature 2 meters above the sea surface provided by NCEP CFSR 6-hourly products. Correlation of daily $\rho_a W^3$ and W^3 (provided by NCEP CFSR 6-hourly products) is significantly high as 0.9989 (p<0.001, Figure A1) emphasizing the importance of term W^3 in controlling the daily variability of wind-induced *PEA* changes over the shelf. We, thus, approximated the relationship as:

$$\mathrm{d}(PEA)_{wind} \propto W^3, \quad (8)$$

The term W^3 is introduced as another candidate predictor and is denoted as $\mathrm{PEA_{wind}}$ (wind-induced PEA changes) for simplification.

2.1.2 Biochemical-related predictors

Sedimentary biochemical processes directly influence the bottom DO consumption rate. However, by far, global forecast model systems like HYCOM does not include biochemical fields. Therefore, the biochemical-related term SOC needs to be replaced by an alternative term (denoted as SOCalt) that does not rely on biochemical simulations. According to the SOC scheme stated in Eq. (8) and Eq. (10) in Part I, the biochemical features are attributed to the sedimentary particulate organic nitrogen (PONsed, derived from ROMS hindcasts) concentration. The total nitrate and nitrite load by the Mississippi River are used to represent the PONsed variability, because inorganic nitrogen is the primary nutrient resource for plankton bloom. Daily updates of measured riverine nitrate+nitrite loads are accessible from U.S. Geological Survey (USGS) National Water Information System (NWIS). Due to lateral transports and vertical settling of particulate organic matter, a leading period should be introduced to the time series of riverine nutrient loads. The optimal length of leading days is obtained by examining the highest linear correlation of regionally averaged ROMS-hindcast SOC and SOCalt following Eq. (9) and is calculated as 19 days (Figure A2a). The exponential term in Eq. (9) estimates the temperature-dependent decomposition rate of organic matter. A significant correlation coefficient between daily SOCalt and ROMS-hindcast SOC is found as 0.8157 (p<0.001, Figure A2).

$$SOCalt = \text{Mississippi River inorganic nitrogen loads (led by 19 days)} \cdot \mathrm{e}^{0.0693T_b}, \quad (9)$$

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

where T_b indicates bottom water temperature (in ˚C). Along with SOCalt, the temperature-dependent decomposition rate $e^{0.0693 \cdot T_b}$ is also considered as a candidate predictor in statistical models and is denoted as DCP_{Temp} for simplification.

2.1.3 HA estimation

As listed in Table 1, there are six candidate predictors considered in the statistical models including four stratification-related variables (PEA, SSS, (Qh), and W^3) and two bottom biochemical variables (SOCalt and $e^{0.0693 \cdot T_b}$). For simplification, we denoted this variable as (Qh), W^3, and $e^{0.0693 \cdot T_b}$ as PEA_{heat}, PEA_{wind}, and DCP_{Temp}, respectively. Correlation coefficients matrix (Figure 1a) indicates that multicollinearity may become a problem in regression models since linear correlations among some predictors are significantly high, e.g., 0.76 (p<0.001) between PEA and SOCalt, and -0.88 (p<0.001) between PEA and SSS. The frequency distribution of HA (Figure 1b) illustrates that the response variable is highly right-skewed with ~51 % of samples (2,506 out of 4,968) being exactly zero. The HA is estimated by the number of hypoxia cells times a constant value (area of the computational cell). Thus, the HA can be estimated by the number of grid cells when the Poisson and negative binomial regression models are applied. However, the great portion of zero samples leads to overdispersion (magnitude of variance ≫ magnitude of mean, i.e., 52,161,613 ≫ 4,378) and zero-inflated problems (Lambert, 1992). The overdispersion issue violates the mean-variance equality assumption employed in regular Poisson regression models, while zero-inflated problems can weaken the model performances.

Table 1. Description of daily response variable and candidate predictors. The data cover a time range from 1 January 2007 to 26 August 2020. Prescribed min and max are used for min–max normalization.

Variables [units]	Description	Min	Median	Mean	Max	Prescribed (Min:Max)
HA [km^2]	Area of extremely low dissolved oxygen concentration (< 2 mg L^{-1})	0	0	4,378	40,561	Non-normalized
PEA [J m^{-3}]	Potential energy anomaly measuring the water stratification	3.1	36.9	49.2	190.4	(0:200)
SSS [non-dim]	Sea surface salinity	20.7	31.8	31.4	34.4	(0:40)
PEA_{heat} [W m^{-3}]	=Qh, an approximation of surface heat-induced water stratification	-1,472.9	3,986.3	3,717.2	6,829.7	(-2,000:7,000)

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

PEA_{wind} [$m^3\ s^{-3}$]	$=W^3$, an approximation of water stratification changes due to wind mixing	0.8	175.1	305.4	6,415.8	(0:6,500)
SOCalt [$mmol\ m^{-3}\ s^{-1}$]	An alternative term for sediment oxygen consumption.	874,870	10,103,864	12,604,970	41,530,153	(800,000:42,000,000)
DCP_{Temp}[non-dim]	$= e^{0.0693 \cdot T_b}$, temperature-dependent decomposition rate of organic matter	2.6	5.1	5.2	8.0	(0:10)

177

178 **2.2 Data pre-processes**

179 We applied the spatially averaged daily ROMS-derived predictors over the LaTex Shelf, then applied the min–max
180 normalization (Eq. (10)) to the one-dimensional time series. Predictive models can be beneficial from the min–max
181 normalization when applying to a new dataset since the method guarantees that the normalized predictors from different
182 datasets range from 0 to 1 as the minimum and maximum values are prescribed. Note that the response is not normalized.
183

184 $$X_{nor} = \frac{X_{org} - Min_{prescribed}}{(Max_{prescribed} - Min_{prescribed})}, \quad (10)$$

185

186 where X_{nor}, X_{org}, $Min_{prescribed}$, and $Max_{prescribed}$ represent normalized value, original value, prescribed minimum, and
187 prescribed maximum, respectively. The daily samples are then split into a training set (for model construction) accounting for
188 80 % of the total samples and a test set (for assessment of model performances) accounting for the rest 20 %. To maintain the
189 HA distribution in both sets, a random resampling method is applied in different HA intervals individually. For example, 80
190 % of samples with HA=0 is chosen randomly for the training set out of all daily samples with HA=0, while the rest of samples
191 with HA=0 is grouped into the test set. The HA=0 is the first interval to which the resampling process is applied, while the rest
192 of samples are split every 5,000 km^2. However, the distribution of HA from each year is similar with a right-skewed structure
193 and numerous zero values. Thus, even though through random processes, both the training and test sets contain samples from
194 each year including samples with non-peak and peak HA. Samples shown in Figure 4 are listed sequentially in the time
195 dimension from 2007 to 2020 but are not equally distributed along time, which means that the listed samples should not be
196 regarded as time series. This splitting method increases the model applicability and provides a comprehensive assessment of
197 prediction performances on both non-peak and peak HA.

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

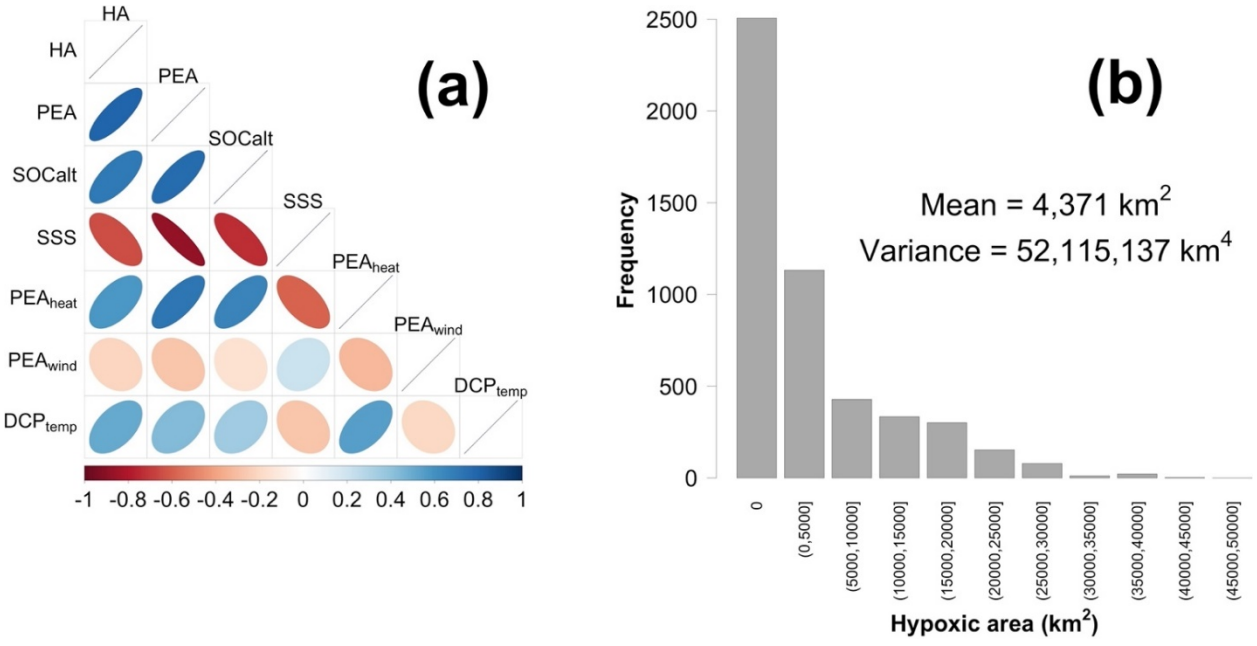


Figure 1. (a) Correlation coefficient matrix of the response variable and candidate predictors, and (b) frequency distribution of HA.

3 Model construction

3.1 Model built-up process

Several regression models are explored using the statistical programming language R. To find the "best" model balancing both model interpretability and prediction performance, a procedure is conducted for model selection (Figure 2) and is summarized below. (1) Choose a regression model. (2) Apply an exhaustive best-subset searching approach to the chosen model. Models with possible combinations of candidate predictors from the ROMS training set are built. A 10-fold cross-validation (CV) method is applied to each model yielding 10 root-mean-square errors (RMSEs) and 1 corresponding mean. The candidate predictors of PEA and SOCalt are forced into each subset. Thus, the number of fitted models with a subset size of k is $C(6-2, k-2) = \frac{4!}{(6-k)!(k-2)!}$, $2 \le k \le 6$ (the total number of candidate predictors is 6). The optimal subset of this size is found as the one with the lowest mean CV RMSE among these models. The best subset is then obtained by comparing mean CV RMSEs of the optimal subsets of different sizes. (3) Steps (1)–(2) are repeated for the selected M candidate regression models. (4) Prediction performances of different models with the corresponding best subsets are assessed by the 10-fold CV RMSEs and Bootstrap (1,000 iterations) aggregating (i.e., Bagging) ensemble algorithms. The Bagging method builds the given model N (=1,000) times during each of which the given model is trained using different samples chosen randomly and repeatedly from the ROMS training set and is executed for HA prediction using samples in the ROMS test set. The ensemble means and ensemble 95 % prediction intervals (PIs) of forecast HA are given according to the prediction results in the 1,000

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

iterations. The best model (Model X in Figure 2) is chosen according to the comparisons of the 10-fold CV RMSEs and the Bagging results.

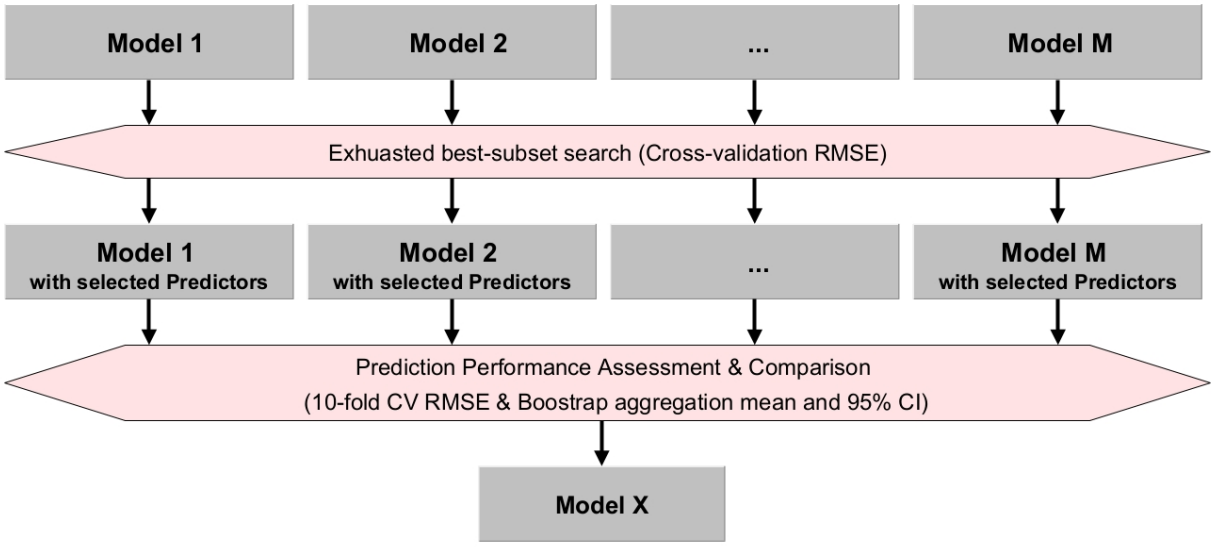


Figure 2. A flow chart of building up regression models.

3.2 Generalized linear models (GLMs)

3.2.1 Regular GLMs and zero-inflated GLMs

The response variable can be treated as count data. Regular Poisson (function glm in R package "stats" version 3.6.2), quasi-Poisson (function glm in R package "stats" version 3.6.2), and negative binomial (function glm.nb in R package "MASS" version 7.3-54; Venables and Ripley, 2002) GLMs are explored in this section. The latter two GLMs are known for solving overdispersion problems by relaxing the mean-variance equality assumption. These GLMs make use of a natural log link function. Thus, a natural logarithm of the area of a single ROMS cell (~ 25.56 km^2) is added to the models as an offset term (an additional intercept term).

In addition, the overdispersion issue can result from the great percentage (~51 %) of zero values in the response variable (Figure 1b). Zero-inflated GLMs (using function zeroinfl in R package "pscl" version 1.5.5; Jackman, 2020; Zeileis et al., 2008) are developed for dealing with response variables of this kind. Rather than resetting dispersion parameters, a zero-inflated count model is a two-component mixture model blending a count model and a zero-excess model. The count model is usually a Poisson or negative binomial GLM (with log link), while the zero-excess model is a binomial GLM (with logit link in this study) estimating the probability of zero inflation. An offset term of log (25.56) is also introduced into the count model. Instead of applying the best-subset searching to the count and zero-excess models simultaneously, in this study, the searching

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

is conducted respectively for these two models to reduce demands of computational resources. The best subset of the zero-excess model (binomial GLM) is given first. The best subset of the count model (Poisson or negative binomial GLMs) is then provided blending the zero-excess model with the corresponding selected best subset fixed.

However, it is hard to determine whether a given zero value of HA is excessive, instead, it is relatively easy to model hypoxia occurrence assuming that all the zero values are excessive. A new binary response, hypoxia, stated in Eq. (11) is introduced for modeled hypoxia occurrence using regular binomial GLMs (function glm in R package "stats" version 3.6.2). The hypoxia is equal to 0 when HA is 0 (no hypoxia), otherwise, is equal to 1. The optimal model selected three predictors: PEA, SOCalt, and DCP_{Temp} (Figure 3b).

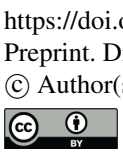
$$hypoxia = \begin{cases} 0, & no\ hypoxia \\ 1, & hypoxia\ occurs \end{cases}, \tag{11}$$

3.2.2 Performance of GLMs

The zero-inflated Poisson GLM serves as the best GLM in terms of prediction performances since it has the lowest mean CV RMSE (Figure 3a) among the five candidate GLMs. The relaxation of the mean-variance equality assumption by the negative binomial GLM and the quasi-Poisson GLM does not guarantee salient improvement of performances when comparing their CV RMSEs to those of regular Poisson GLM. The zero-inflated negative binomial GLM yields poorest perfomance with the largest mean CV RMSE. The mean CV RMSEs of zero-inflated Poisson GLM hit the trough (3,621 km^2) at the size of five. However, the greatest drop of RMSEs (3,671 km^2) occurs at the size of three beyond which the RMSEs remain stable. It is worth considering a model with fewer predictors satisfying model interpretability. Thus, the best zero-inflated Poisson GLM accounts for three predictors (PEA, SOCalt, and DCP_{Temp}) in the count model and three predictors (PEA, SOCalt, and DCP_{Temp}) in the zero-excess model. As indicated in the correlation matrix (Figure 1a), the robustness of a model can be imparied by multicollinearity which can be estimated by variance inflation factors (VIFs). VIFs among the selected predictors are 2.60, 2.43, and 1.23 for PEA, SOCalt, and DCP_{Temp}, respectively. The VIFs are all less than 5 suggesting that predictors subsets involved in both the count and the zero-excess models are merely violated by multicollinearity. For simplicity, the best zero-inflated Poisson GLM is symbolized as GLMzip3.

The Bagging ensemble method is implemented to estimate the prediction performance of GLMzip3 (Figure 4a). The Bagging means of predicted HA provides an RMSE of 3,610 km^2 and an R^2 of 0.7511 against the ROMS hindcasts. The Bagging 95 % PIs of the predicted values are restricted within a narrow range with a slight increase at the predicted peaks. We want to address that the comparisons are not between the time series. The training set and test set are resampled according to different HA intervals, while the distributions of HA in each year are similar. Thus, HA in both the training set and test set contains

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022



observations of peak and non-peak values in each year. The results suggest that GLMzip3 is capable of providing not only accurate but also stable HA forecasts. Nevertheless, we noted salient overestimations (e.g., around the 30th and 920th samples) and underestimations (e.g., around the 540th and 830th samples) at some peaks. Instead of the prediction performance at non-peak HA, here we are more focused on forecasts at HA peaks which impose more threatens to the shelf ecosystem. In section 3.3, GAMs are investigated with an expectation of further improvements in peak predictions by considering non-parametric or non-linear effects of the predictors.

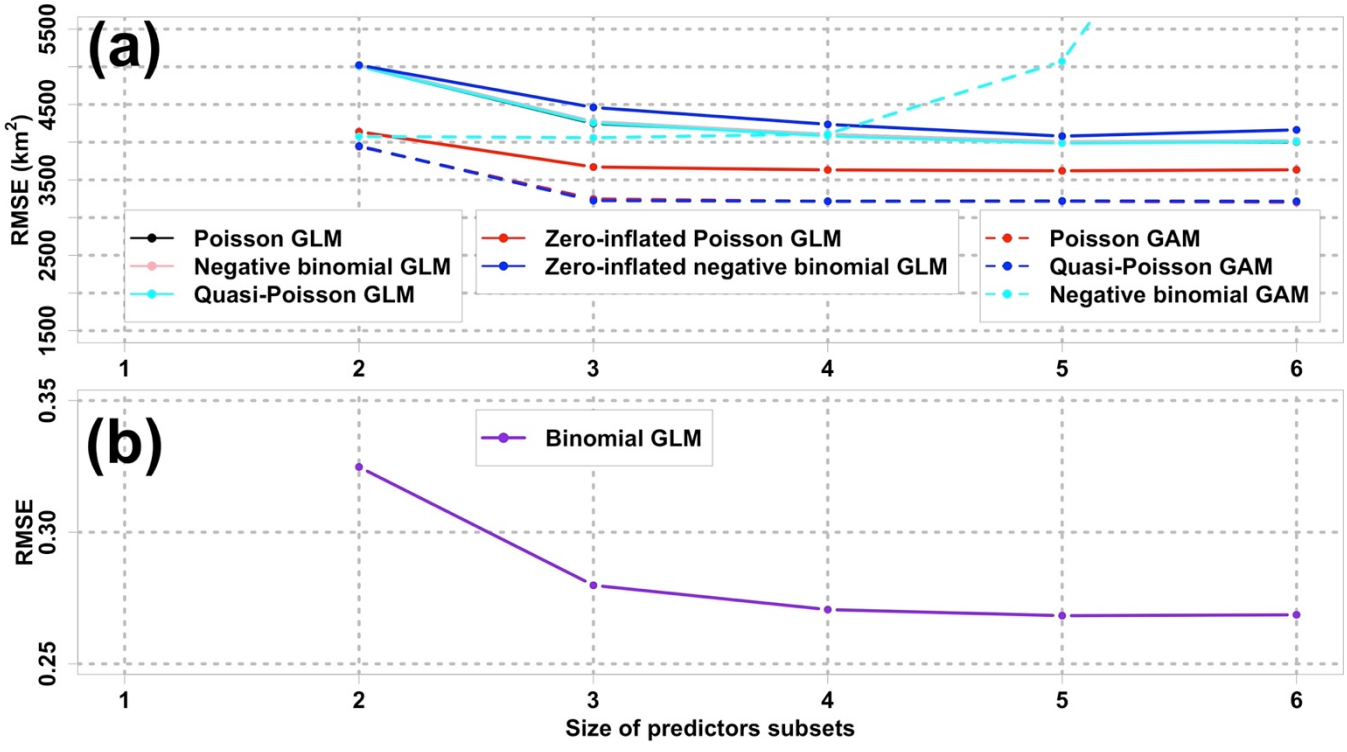


Figure 3. Comparisons of mean 10-fold CV RMSEs among different regression models with various sizes of predictors subsets. The response variable in (b) binomial GLM and (a) other models is hypoxia occurrence (hypoxia) and hypoxic area (HA), respectively. Note that the CV RMSE of negative binomial GAM with the size of six is out of the range shown. CV RMSE curves of the Poisson GLM, negative binomial GLM, and quasi-Poisson GLM overlap, while those of Poisson GAM and quasi-Poisson GAM overlap. The minimum size of predictor subsets is two since PEA and SOCalt are forced into every subset.

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

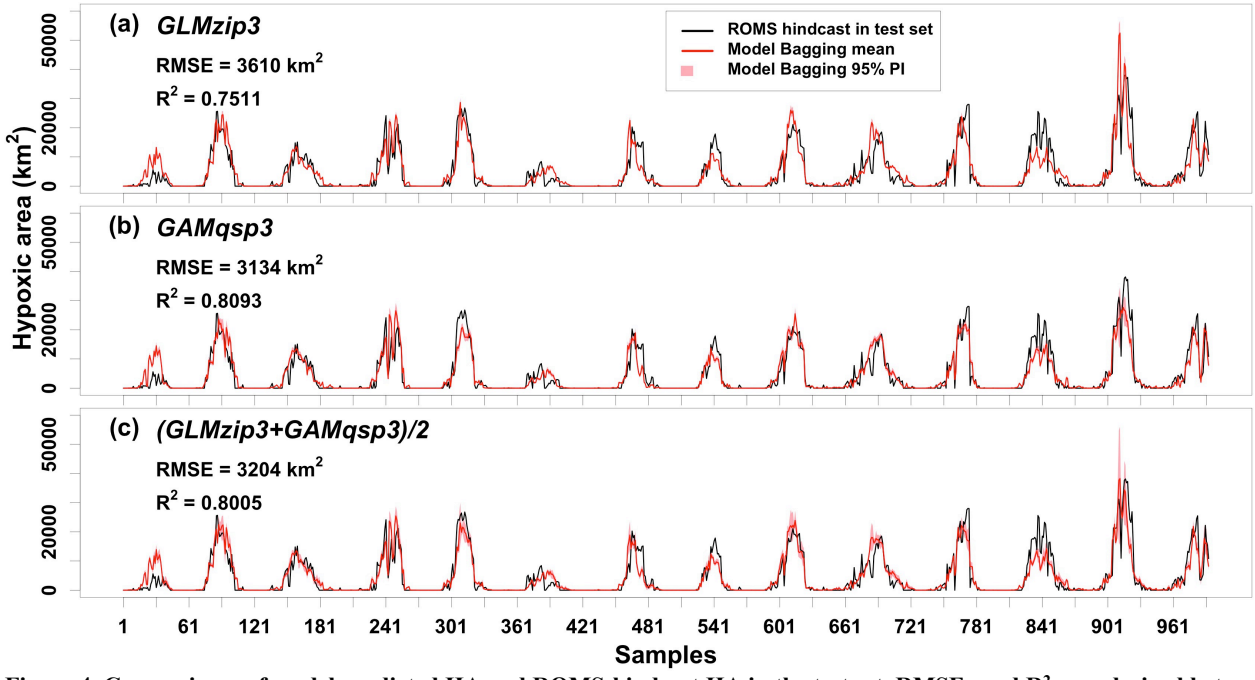


Figure 4. Comparisons of model predicted HA and ROMS-hindcast HA in the test set. RMSEs and R^2s are derived between model Bagging mean and ROMS-hindcast HA.

3.2.3 Model interpretation for GLMzip3

We applied the complete ROMS training set to the model construction of GLMzip3 and found the coefficients for PEA, SOCalt, and DCP_{Temp} (Table 2) are all significantly positive (p<0.001) in the count model, while coefficients for these predictors are significantly negative (p<0.001) in the zero-excess model. The count model simulates the HA while the zero-excess model estimates the probability of HA being zero. Higher PEA is consistent with stronger water stratification, while higher SOCalt and DCP_{Temp} are both corresponding to higher sediment oxygen consumption. Therefore, there is no surprise that higher PEA, SOCalt, and DCO_{Temp} are related to greater HA and higher hypoxia occurrence or lower probability of HA being zero. Results indicate that the GLMzip3 essentially builds up reasonable relationships between the response and predictors variables with a high agreement with physical and biochemical mechanisms.

Table 2. Regression coefficients of GLMzip3.

Count model coefficients (Poisson with log link):					Zero-excess model coefficients (binomial with logit link):				
	Estimate	Std. Error	z value	Pr (> |z|)		Estimate	Std. Error	z value	Pr (> |z|)
Intercept	1.9897	0.0021	948.2	<2E-16***	Intercept	9.1993	0.3181	28.9	<2E-16***
PEA	2.6763	0.0016	1681.4	<2E-16***	PEA	-10.0945	0.5986	-16.9	<2E-16***
SOCalt	0.9228	0.0014	663.6	<2E-16***	SOCalt	-8.7784	0.5508	-15.9	<2E-16***
DCP_{Temp}	3.5940	0.0031	1168.2	<2E-16***	DCP_{Temp}	-9.4939	0.4346	-21.9	<2E-16***

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022


Biogeosciences Discussions Open Access EGU

Significance codes:	0 (***)	0.001 (**)	0.01 (*)
Log-likelihood: -2.4E6 on 8 degrees of freedom			

294

295 **3.3 Generalized additive models (GAMs)**

296 GAMs are explored with an expectation of improving prediction performance in HA peaks by introducing non-parametric
297 effects of predictors. Using function "gam" in R package “mgcv” (version 1.8-36; Wood, 2011) with smooth functions as pure
298 thin plate regression splines (degree of freedom=9; Wood, 2003), three GAMs are studied and compared, i.e., Poisson GAM,
299 quasi-Poisson GAM, and negative binomial GAM. Following the same procdurel in GLM exploration, the best subset
300 searching approach is applied to the GAMs first. The mean 10-fold CV RMSEs for the Poisson and quasi-Poisson GAMs
301 (Figure 3a) exhibit insignificant differences and are the lowest among those for all GLMs and GAMs studied. Although the
302 mean CV RMSEs for these two types of GAMs both reach the lowest at the size of five, the best size is considered as three
303 (considering PEA, SOCalt, and DCP_{Temp}) at which CV RMSEs exhibits most saline decline, and beyond which mean CV
304 RMSEs stabilize around 3,200 km^2. The negative binomial GAM has the greatest mean CV RMSEs among the GAMs studied
305 and has an extremely high mean CV RMSE at the size of six. It is, therefore, dropped out of the list of candidate models. The
306 quasi-Poisson GAM with three predictors involved (symbolized as GAMqsp3) is chosen as the best GAM since it relaxes the
307 mean-variance equality assumption which should not be applied to the HA dataset due to the overdispersion issue.
308
309 Component plots of model GAMqsp3 (Figure 5) imply that HA generally increases as the chosen predictors increase. The
310 smooth functions of PEA and DCP_{Temp} are considerably greater than the smooth function of SOCalt indicating that the
311 contributions of the former two predictors are greater than the effect of SOCalt on the daily variability of HA. Note that the
312 fitted HA equals the summation of all smooth function terms. Such results agree with those found by model GLMzip3.
313 However, the component plots provide more detailed information about the rate of changes of HA. The effective degrees of
314 freedom range from 8 to 8.54 indicating strong non-linear effects of the predictors on the changes of HA. The HA is more
315 sensitive to the predictors in the low-value ranges but becomes nearly stable in the medium- and high-value ranges of
316 predictors. It implies that bottom hypoxia develops rapidly in early summer when water stratification and sediment oxygen
317 demand start to increase. The bottom hypoxic water further extends with a much lower expansion speed as the stratification
318 and SOC further intensify. Nevertheless, the smooth function of PEA is slightly greater also with a more acute slope than those
319 found for SOCalt and DCP_{Temp} in the medium- and high-value regimes of the predictors. It indicates that the HA variability is
320 more related to the hydrodynamic changes in the shelf than the biochemical effects. The result is consistent with the findings
321 by previous studies of the shelf hypoxia (Yu et al., 2015; Mattern et al., 2013) emphasizing that the physical impacts are
322 stronger than the biological impacts on HA estimates. A short conclusion is made that the GAMqsp3 model provides reasonable
323 interpretations on the hypoxic area mechanisms.

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

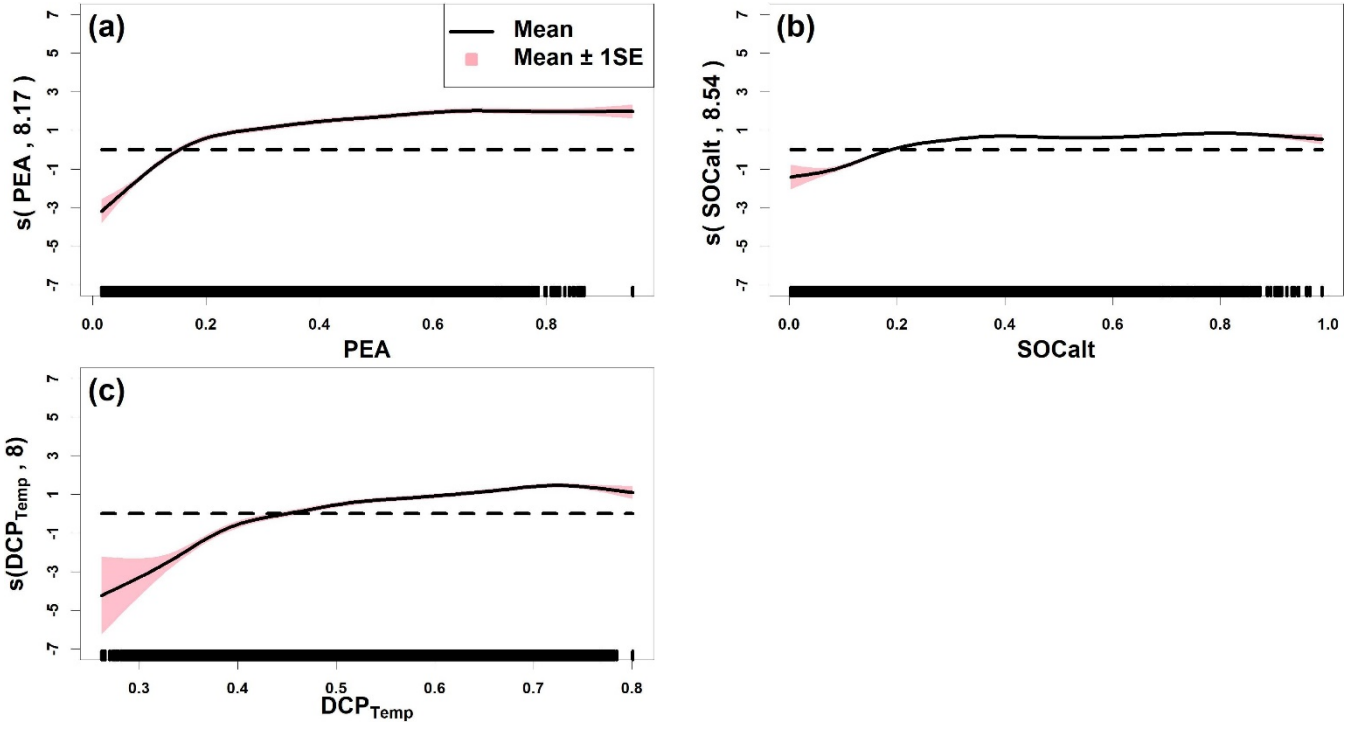


Figure 5. Component plots of model GAMqsp3. Solid black lines represent the mean of the smooth function, while the pink area denotes the range of mean ± 1SE. Numbers in brackets represent effective degrees of freedom for the corresponding smooth terms. Black bars at the x axis indicate the density of corresponding predictors. Dashed black lines are straight lines of zero along the predictor domains. Note that the predictors shown have been normalized.

The prediction performance of GAMqsp3 is estimated using the Bagging ensemble method (Figure 4b). The RMSE and R^2 between the Bagging mean and ROMS-hindcast HA is 3,134 km^2 and 0.8093, respectively. They are 13 % lower and 8 % higher than the corresponding statistics found for the GLMzip3, respectively, suggesting that GAMqsp3 outcompetes GLMzip3 in terms of overall performance. However, GAMqsp3 tends to produce underestimated predictions at HA peaks (like peaks around the 310th and 920th samples) some of which are overestimated by the GLMzip3. Therefore, instead of determining the best model out of the two, ensemble HA predictions blending efforts of both GLMzip3 and GAMqsp3 are carried out with an expectation to improve model performance in the peak forecast. We assumed that the contributions of GLMzip3 and GAMqsp3 are equally weighted since there is no clue showing the apparent superiority of either model in HA peak predictions. We thus averaged the predicted HA by GLMzip3 and GAMqsp3 and calculated the 95 % PIs given the Bagging results of these models (Figure 4c). As expected, the overall performance of the ensemble forecast is somewhere between the performance of GLMzip3 and GAMqsp3 with an RMSE of 3,204 km^2 and an R^2 of 0.8005. However, some HA peak events (like peaks around the 310th and 920th samples) which are overestimated by GLMzip3 but are underestimated by GAMqsp3 are accurately predicted by the ensemble approach. The ensemble model provides higher accuracy in peak forecast given minor sacrifices in overall performance.

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022


Biogeosciences Discussions

4 Discussion

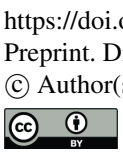
A promising HA forecast is provided by the ensemble model with a low RMSE (3,204 km^2), a high R^2 (0.8005), and a precise performance in capturing hypoxic area peaks in the summers. The power of the prediction model relies on the availability of the forecast of predictors. In this section, we discuss the model's transferability using an independent global ocean product.

The Global Ocean Forecasting System (GOFS) 3.1 provides global daily analysis products and an eight-day forecast in a daily interval with a horizontal resolution of 1/12 °. The products (hereafter referred to HYCOM-derived products) are derived by a 41-layer HYCOM global model with data assimilated via the Navy Coupled Ocean Data Assimilation (NCODA) system (Cummings, 2005; Cummings and Smedstad, 2013). Daily data from 1 January 2007 to 26 August 2020 are retrieved and studied. Predictors of PEA, SOCalt, and DCP_{Temp} are reconstructed using HYCOM-derived variables and Mississippi River daily total nitrate and nitrite loadings downloaded from the USGS NWIS. Relationships of ROMS-derived and HYCOM-derived predictors are examined in Figure 6. The magnitudes of HYCOM-derived SOCalt and DCP_{Temp} match up with the corresponding ROMS-derived predictors, respectively, although HYCOM-derived predictors are found slightly greater. Simple linear regression for these predictors illustrates that the linear relationships between the ROMS and HYCOM products are significant with the R^2 ranging from 0.93 to 0.95. The intercept terms are at least one-order smaller than the magnitudes of corresponding predictors. Therefore, the HYCOM global products are deemed to agree with the ROMS hindcasts for SOCalt and DCP_{Temp}. Nevertheless, the magnitude of HYCOM-derived PEA is found much lower than the ROMS-derived PEA (Figure 6a). Simple linear regression indicates a significant linear relationship between the natural log transformation of PEA from the two datasets (R^2=0.69).

At land-sea interfaces, the HYCOM global model is forced by monthly riverine discharges, which weakens the model performance in coastal regions. The hydrodynamics in the LaTex Shelf is highly affected by the freshwater and momentum from the Mississippi and the Atchafalaya Rivers. Monthly river forcings in HYCOM are essentially weaker than daily forcings used in our ROMS set up and can result in a less stratified water column (i.e., lower PEA). Therefore, it is necessary to scale the magnitude of HYCOM-derived PEA to that of the ROMS hindcast. It can be achieved by using the natural log transformation and simple linear regression as discussed. We then adjusted HYCOM-derived PEA but kept the HYCOM-derived SOCalt and DCP_{Temp} unchanged before the application of the ensemble model.

The Bagging approach is implemented again to assess the performances of the ensemble model. During each iteration (N=1,000), the GLMzip3 and GAMqsp3 are trained using the ROMS training set and then applied to the adjusted HYCOM-derived predictors for HA prediction from 1 January 2019 to 26 August 2020. The ensemble method provides averages and 95 % PIs of predicted HA blending Bagging results by GLMzip3 and GAMqsp3. Compared to the ROMS-hindcast HA, the ensemble model performs an overall accurate HA forecast with an RMSE and an R^2 of 4,571 km^2 and 0.8178, respectively

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022



(Figure 7). The HA peaks in both 2019 and 2020 summers are well captured by the model with slight underestimates at the first peak and slight overestimates at the second. The width of 95 % PI is larger during high HA periods suggesting less stability in the HA peak forecast. Discharges measurements for the Mississippi and the Atchafalaya Rivers are provided and updated daily by USGS NWIS, assuring a possible improvement of HYCOM model performance in the LaTex Shelf. Once the performance of hydrodynamics predictions in the LaTex Shelf is guaranteed, predictions performance of the ensemble model on HA would be further improved.

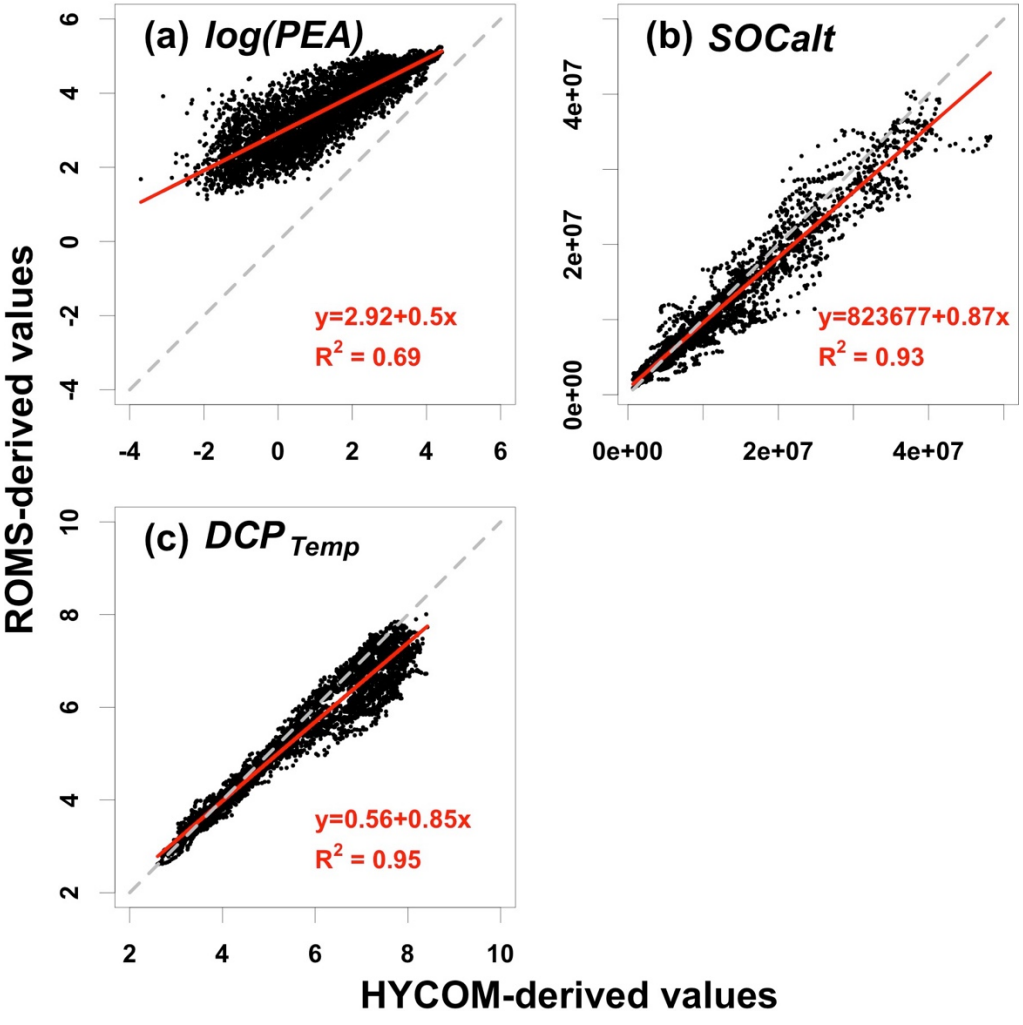


Figure 6. Scatter plots of (a) log(PEA) (unit: log (J m^{-3})), (b) SOCalt (unit: mmol m^{-3} s^{-1}), and (c) DCP_{Temp} (unit: 1) between ROMS and HYCOM simulations. Note that the solid red lines represent linear regression lines, while the dashed grey lines are diagonals with a slope of 1 and an intercept of 0. Daily data compared are from 2007 to 2020.

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

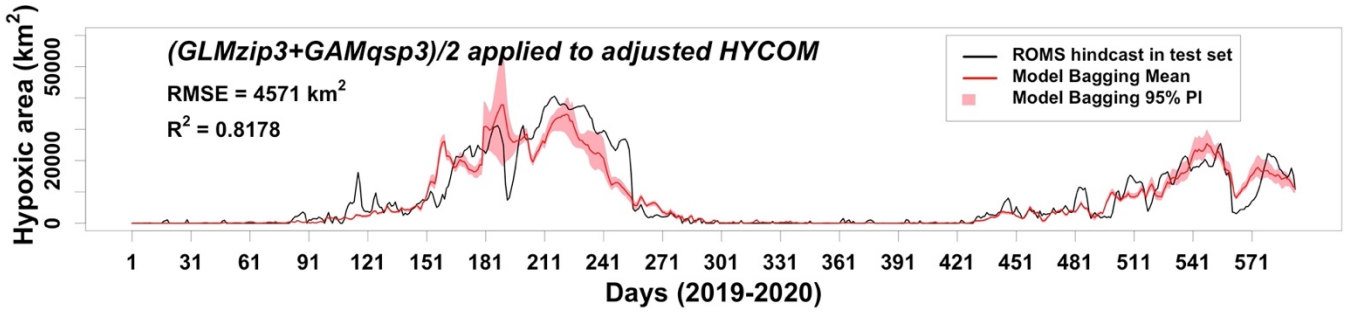


Figure 7. Daily time series of predicted HA by ensemble model ((GLMzip3+GAMqsp3)/2) when applied to adjusted HYCOM products and ROMS-hindcast HA from 2019 to 2020.

5 Conclusion

In this study, an ensemble HA forecast model for the LaTex Shelf is developed using the state-of-the-art statistic programming language R. The model is trained using numeric simulations from 1 January 2007 to 26 August 2020 generated by a coupled hydrodynamic–biogeochemical model. Before splitting data into a training set and a test set, we applied regional average over the LaTex Shelf and min-max normalization to the hindcast data.

Multiple GLMs (regular Poisson GLMs, quasi-Poisson GLMs, negative binomial GLMs, zero-inflated Poisson GLMs, and zero-inflated negative binomial GLMs) and GAMs (regular Poisson GAMs, quasi-Poisson GAMs, and regular negative binomial GAMs) are assessed for HA prediction. Comparisons of model prediction performance illustrate that an ensemble model combing the prediction efforts of a zero-inflated Poisson GLM and a quasi-Poisson GAM provides the most accurate HA forecast with a variability explanation high up to 80 %, a low overall RMSE of 3,204 km^2, and a high precision in forecasting peak HA when compared to the hindcasts by the coupled model. Predictors PEA, SOCalt, and DCP_{Temp} are involved in the GLM and GAM. Statistically significant coefficients for the predictors (for the GLMzip3) and component plots (for the GAMqsp3) agree well with the physical and biochemical mechanisms.

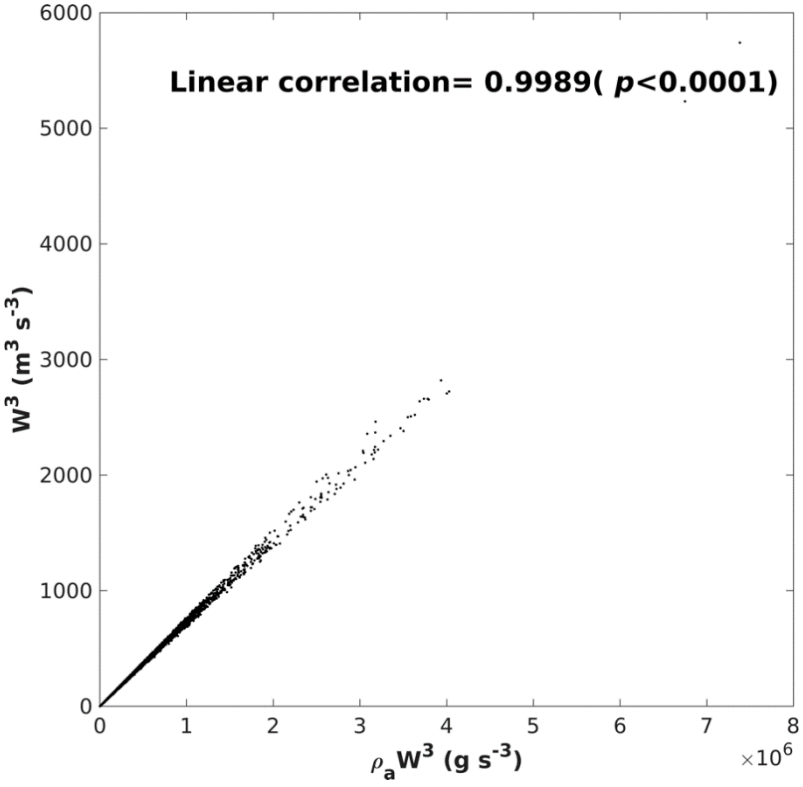
The ensemble model is then migrated to the GOFS 3.1 products based on HYCOM,s which provides eight-day forecast of global hydrodynamics. The ensemble model is trained using the ROMS training set and then is used for the HA prediction covering the period from January 1st, 2019 to August 26th, 2020. The prediction is robust when compared to the ROMS simulations (2019–2020), with a low overall RMSE (4,571 km^2) and a high R^2 (0.8178). The model can also accurately predict the magnitude and onset of summer HA peaks in 2019 and 2020, respectively. To our best knowledge, this ensemble model is the first model providing efficient yet accurate daily HA forecast for the LaTex Shelf while considering both hydrodynamic and biochemical effects. This model is also the first model successfully applying global hydrodynamic forecast in regional HA predictions.

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

413 **Author contribution:** Bin Li and Z. George Xue designed the experiments and Yanda Ou carried them out. Yanda Ou
414 developed the model code and performed the simulations. Yanda Ou, Bin Li, and Z. George Xue prepared the manuscript.
415
416 **Competing interests:** The authors declare that they have no conflict of interest.
417
418 **Acknowledgment:** Research support was provided through the Bureau of Ocean Energy Management (M17AC00019,
419 M20AC10001). We thank Dr. Jerome Fiechter at UC Santa Cruz for sharing his NEMURO model codes. Computational
420 support was provided by the High-Performance Computing Facility (clusters SuperMIC and QueenBee3) at Louisiana State
421 University.
422
423

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

424 **Appendix A:**



425

426 **Figure A1. A scatter plot of $\rho_a W^3$ and W^3 and their linear correlation.**

https://doi.org/10.5194/bg-2022-4
Preprint. Discussion started: 9 February 2022

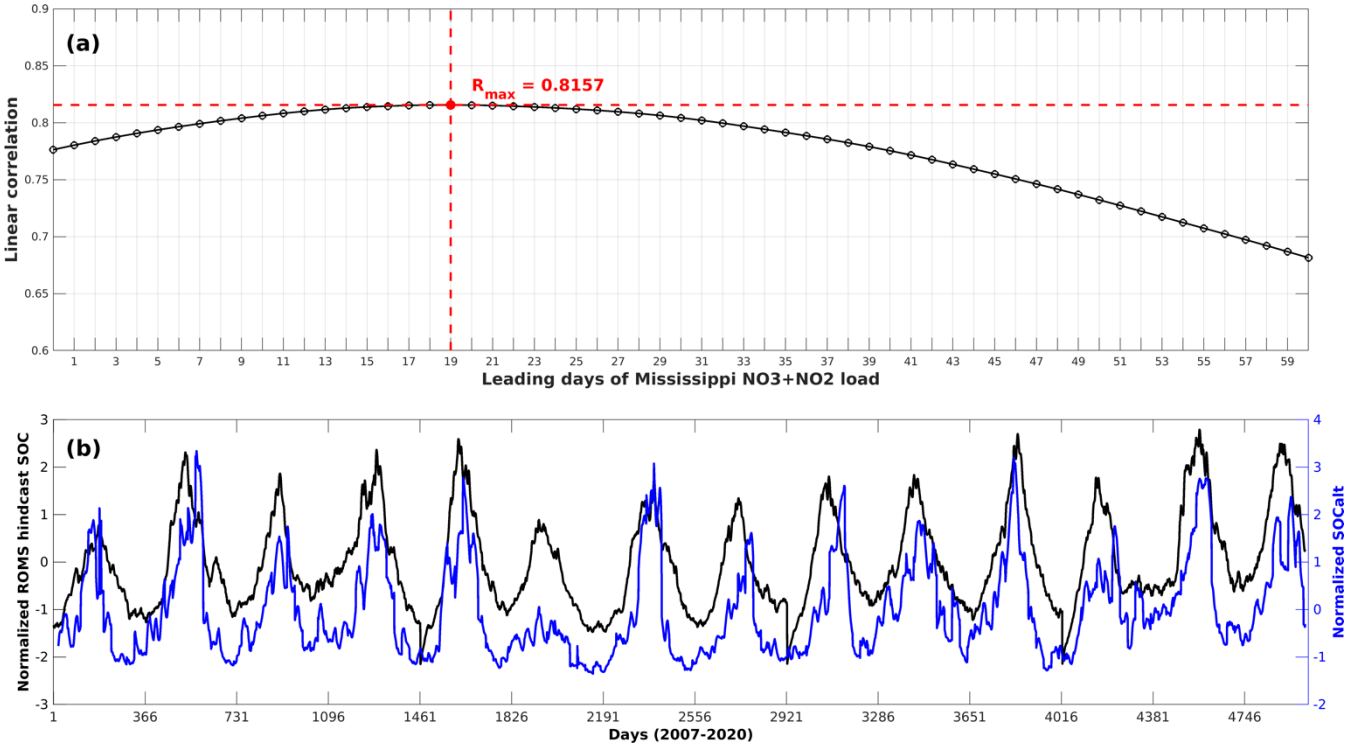


Figure A2. (a) Lead/lag correlation coefficients between ROMS hindcast daily SOC and SOCalt (= Mississippi River inorganic nitrogen loads · $e^{0.0693T_b}$) with the Mississippi nitrogen loads led by different days; (b) daily time series of ROMS hindcast SOC and SOCalt when the Mississippi nitrogen loads led by 19 days. Time series of compared is averaged over the LaTex Shelf and is normalized.

Reference

Bianchi, T. S., DiMarco, S. F., Cowan, J. H., Hetland, R. D., Chapman, P., Day, J. W., and Allison, M. A.: The science of hypoxia in the northern Gulf of Mexico: A review, Sci. Total Environ., 408, 1471–1484, https://doi.org/10.1016/j.scitotenv.2009.11.047, 2010.

Chesney, E. J. and Baltz, D. M.: The effects of hypoxia on the northern Gulf of Mexico Coastal Ecosystem: A fisheries perspective, in: Coastal Hypoxia: Consequences for Living Resources and Ecosystems, American Geophysical Union, 321–354, https://doi.org/10.1029/CE058p0321, 2001.

Conley, D. J., Paerl, H. W., Howarth, R. W., Boesch, D. F., Seitzinger, S. P., Havens, K. E., Lancelot, C., and Likens, G. E.: Controlling Eutrophication: Nitrogen and Phosphorus, Science, 323, 1014–1015, https://doi.org/10.1126/science.1167755, 2009.

Craig, J. K. and Bosman, S. H.: Small Spatial Scale Variation in Fish Assemblage Structure in the Vicinity of the Northwestern Gulf of Mexico Hypoxic Zone, Estuaries and Coasts, 36, 268–285, https://doi.org/10.1007/s12237-012-9577-9, 2013.

Cummings, J. A.: Operational multivariate ocean data assimilation, Q. J. R. Meteorol. Soc., 131, 3583–3604, https://doi.org/10.1256/qj.05.105, 2005.

Cummings, J. A. and Smedstad, O. M.: Variational Data Assimilation for the Global Ocean, in: Data Assimilation for

Atmospheric, Oceanic and Hydrologic Applications, vol. II, edited by: Park, S. K. and Xu, L., Springer Berlin Heidelberg, 303–343, https://doi.org/10.1007/978-3-642-35088-7_13, 2013.

Feng, Y., Fennel, K., Jackson, G. A., DiMarco, S. F., and Hetland, R. D.: A model study of the response of hypoxia to upwelling-favorable wind on the northern Gulf of Mexico shelf, J. Mar. Syst., 131, 63–73, https://doi.org/10.1016/j.jmarsys.2013.11.009, 2014.

Fennel, K., Hetland, R., Feng, Y., and Dimarco, S.: A coupled physical-biological model of the Northern Gulf of Mexico shelf: Model description, validation and analysis of phytoplankton variability, Biogeosciences, 8, 1881–1899, https://doi.org/10.5194/bg-8-1881-2011, 2011.

Fennel, K., Hu, J., Laurent, A., Marta-Almeida, M., and Hetland, R.: Sensitivity of hypoxia predictions for the northern Gulf of Mexico to sediment oxygen consumption and model nesting, J. Geophys. Res. Ocean., 118, 990–1002, https://doi.org/10.1002/jgrc.20077, 2013.

Fennel, K., Laurent, A., Hetland, R., Justic, D., Ko, D. S., Lehrter, J., Murrell, M., Wang, L., Yu, L., and Zhang, W.: Effects of model physics on hypoxia simulations for the northern Gulf of Mexico: A model intercomparison, J. Geophys. Res. Ocean., 121, 5731–5750, https://doi.org/10.1002/2015JC011516, 2016.

Del Giudice, D., Matli, V. R. R., and Obenour, D. R.: Bayesian mechanistic modeling characterizes Gulf of Mexico hypoxia: 1968–2016 and future scenarios, Ecol. Appl., 30, 1–14, https://doi.org/10.1002/eap.2032, 2020.

de Mutsert, K., Steenbeek, J., Lewis, K., Buszowski, J., Cowan, J. H., and Christensen, V.: Exploring effects of hypoxia on fish and fisheries in the northern Gulf of Mexico using a dynamic spatially explicit ecosystem model, Ecol. Modell., 331, 142–150, https://doi.org/10.1016/j.ecolmodel.2015.10.013, 2016.

Hetland, R. D. and DiMarco, S. F.: How does the character of oxygen demand control the structure of hypoxia on the Texas-Louisiana continental shelf?, J. Mar. Syst., 70, 49–62, https://doi.org/10.1016/j.jmarsys.2007.03.002, 2008.

Jackman, S.: pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, https://github.com/atahk/pscl/, 2020.

Justić, D. and Wang, L.: Assessing temporal and spatial variability of hypoxia over the inner Louisiana-upper Texas shelf: Application of an unstructured-grid three-dimensional coupled hydrodynamic-water quality model, Cont. Shelf Res., 72, 163–179, https://doi.org/10.1016/j.csr.2013.08.006, 2014.

Katin, A., Del Giudice, D., and Obenour, D.: Daily hypoxia forecasting and uncertainty assessment via Bayesian mechanistic model for the Northern Gulf of Mexico, Hydrol. Earth Syst. Sci. Discuss., 35, 1–17, https://doi.org/10.5194/hess-2021-207, 2021.

LaBone, E., Rose, K., Justic, D., Huang, H., and Wang, L.: Effects of spatial variability on the exposure of fish to hypoxia: a modeling analysis for the Gulf of Mexico, Biogeosciences Discuss., 1–35, https://doi.org/10.5194/bg-2020-51, 2020.

Lambert, D.: Zero-inflated poisson regression, with an application to defects in manufacturing, Technometrics, 34, 1–14, https://doi.org/10.1080/00401706.1992.10485228, 1992.

Laurent, A. and Fennel, K.: Time-Evolving, Spatially Explicit Forecasts of the Northern Gulf of Mexico Hypoxic Zone, Environ. Sci. Technol., 53, 14449–14458, https://doi.org/10.1021/acs.est.9b05790, 2019.

Laurent, A., Fennel, K., Ko, D. S., and Lehrter, J.: Climate change projected to exacerbate impacts of coastal Eutrophication in the Northern Gulf of Mexico, J. Geophys. Res. Ocean., 123, 3408–3426, https://doi.org/10.1002/2017JC013583, 2018.

484 Matli, V. R. R., Fang, S., Guinness, J., Rabalais, N. N., Craig, J. K., and Obenour, D. R.: Space-Time Geostatistical Assessment
485 of Hypoxia in the Northern Gulf of Mexico, Environ. Sci. Technol., 52, 12484–12493, https://doi.org/10.1021/acs.est.8b03474,
486 2018.

487 Mattern, J. P., Fennel, K., and Dowd, M.: Sensitivity and uncertainty analysis of model hypoxia estimates for the Texas-
488 Louisiana shelf, J. Geophys. Res. Ocean., 118, 1316–1332, https://doi.org/10.1002/jgrc.20130, 2013.

489 McCarthy, M. J., Carini, S. A., Liu, Z., Ostrom, N. E., and Gardner, W. S.: Oxygen consumption in the water column and
490 sediments of the northern Gulf of Mexico hypoxic zone, Estuar. Coast. Shelf Sci., 123, 46–53,
491 https://doi.org/10.1016/j.ecss.2013.02.019, 2013.

492 Monteith, J. and Unsworth, M.: Principles of environmental physics: plants, animals, and the atmosphere, 4th ed., Academic
493 Press, https://doi.org/https://doi.org/10.1016/C2010-0-66393-0, 2014.

494 Murrell, M. C. and Lehrter, J. C.: Sediment and Lower Water Column Oxygen Consumption in the Seasonally Hypoxic Region
495 of the Louisiana Continental Shelf, Estuaries and Coasts, 34, 912–924, https://doi.org/10.1007/s12237-010-9351-9, 2011.

496 Obenour, D. R., Michalak, A. M., and Scavia, D.: Assessing biophysical controls on Gulf of Mexico hypoxia through
497 probabilistic modeling, Ecol. Appl., 25, 492–505, https://doi.org/10.1890/13-2257.1, 2015.

498 Picard, A., Davis, R. S., Gläser, M., and Fujii, K.: Revised formula for the density of moist air (CIPM-2007), Metrologia, 45,
499 149–155, https://doi.org/10.1088/0026-1394/45/2/004, 2008.

500 Rabalais, N. N. and Baustian, M. M.: Historical Shifts in Benthic Infaunal Diversity in the Northern Gulf of Mexico since the
501 Appearance of Seasonally Severe Hypoxia, Diversity, 12, https://doi.org/10.3390/d12020049, 2020.

502 Rabalais, N. N. and Turner, R. E.: Gulf of Mexico Hypoxia: Past, Present, and Future, Limnol. Oceanogr. Bull., 28, 117–124,
503 https://doi.org/10.1002/lob.10351, 2019.

504 Rabalais, N. N., Turner, R. E., and Wiseman, W. J.: Gulf of Mexico hypoxia, a.k.a. "The dead zone," Annu. Rev. Ecol. Syst.,
505 33, 235–263, https://doi.org/10.1146/annurev.ecolsys.33.010802.150513, 2002.

506 Rabalais, N. N., Turner, R. E., Sen Gupta, B. K., Boesch, D. F., Chapman, P., and Murrell, M. C.: Hypoxia in the northern
507 Gulf of Mexico: Does the science support the plan to reduce, mitigate, and control hypoxia?, Estuaries and Coasts, 30, 753–
508 772, https://doi.org/10.1007/BF02841332, 2007a.

509 Rabalais, N. N., Turner, R. E., Gupta, B. K. S., Platon, E., and Parsons, M. L.: Sediments tell the history of eutrophication and
510 hypoxia in the northern Gulf of Mexico, Ecol. Appl., 17, 129–143, https://doi.org/10.1890/06-0644.1, 2007b.

511 Rabotyagov, S. S., Kling, C. L., Gassman, P. W., Rabalais, N. N., and Turner, R. E.: The economics of dead zones: Causes,
512 impacts, policy challenges, and a model of the gulf of Mexico Hypoxic Zone, Rev. Environ. Econ. Policy, 8, 58–79,
513 https://doi.org/10.1093/reep/ret024, 2014.

514 Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H.,
515 Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R.,
516 Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A.,
517 Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S.,
518 Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: NCEP
519 Climate Forecast System Reanalysis (CFSR) 6-hourly Products, January 1979 to December 2010,
520 https://doi.org/10.5065/D69K487J, 2010.

521 Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H., Iredell, M., Ek, M.,

522 Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E.: NCEP Climate Forecast
523 System Version 2 (CFSv2) 6-hourly Products, https://doi.org/10.5065/D61C1TXF, 2011.

524 Scavia, D., Evans, M. A., and Obenour, D. R.: A scenario and forecast model for gulf of mexico hypoxic area and volume,
525 Environ. Sci. Technol., 47, 10423–10428, https://doi.org/10.1021/es4025035, 2013.

526 Simpson, J. H.: The shelf-sea fronts: implications of their existence and behaviour, Philos. Trans. R. Soc. London. Ser. A,
527 Math. Phys. Sci., 302, 531–546, https://doi.org/10.1098/rsta.1981.0181, 1981.

528 Simpson, J. H. and Bowers, D.: Models of stratification and frontal movement in shelf seas, Deep Sea Res. Part A, Oceanogr.
529 Res. Pap., 28, 727–738, https://doi.org/10.1016/0198-0149(81)90132-1, 1981.

530 Simpson, J. H. and Hunter, J. R.: Fronts in the Irish Sea, Nature, 250, 404–406,
531 https://doi.org/https://doi.org/10.1038/250404a0, 1974.

532 Simpson, J. H., Allen, C. M., and Morris, N. C. G.: Fronts on the Continental Shelf, J. Geophys. Res., 83, 4607–4614,
533 https://doi.org/https://doi.org/10.1029/JC083iC09p04607, 1978.

534 Smith, M. D., Asche, F., Bennear, L. S., and Oglend, A.: Spatial-dynamics of hypoxia and fisheries: The case of Gulf of
535 Mexico brown shrimp, Mar. Resour. Econ., 29, 111–131, https://doi.org/10.1086/676826, 2014.

536 Turner, R. E., Rabalais, N. N., and Justic, D.: Predicting summer hypoxia in the northern Gulf of Mexico: Riverine N, P, and
537 Si loading, Mar. Pollut. Bull., 52, 139–148, https://doi.org/10.1016/j.marpolbul.2005.08.012, 2006.

538 Venables, W. N. and Ripley, B. D.: Modern Applied Statistics with S, Fourth., Springer, New York,
539 https://doi.org/10.1007/978-0-387-21706-2, 2002.

540 Wang, L. and Justić, D.: A modeling study of the physical processes affecting the development of seasonal hypoxia over the
541 inner Louisiana-Texas shelf: Circulation and stratification, Cont. Shelf Res., 29, 1464–1476,
542 https://doi.org/10.1016/j.csr.2009.03.014, 2009.

543 Warner, J. C., Armstrong, B., He, R., and Zambon, J. B.: Development of a Coupled Ocean-Atmosphere-Wave-Sediment
544 Transport (COAWST) Modeling System, Ocean Model., 35, 230–244, https://doi.org/10.1016/j.ocemod.2010.07.010, 2010.

545 Wood, S. N.: Thin plate regression splines, J. R. Stat. Soc. Ser. B Stat. Methodol., 65, 95–114, https://doi.org/10.1111/1467-
546 9868.00374, 2003.

547 Wood, S. N.: Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized
548 linear models, J. R. Stat. Soc. Ser. B Stat. Methodol., 73, 3–36, https://doi.org/10.1111/j.1467-9868.2010.00749.x, 2011.

549 Yu, L., Fennel, K., and Laurent, A.: A modeling study of physical controls on hypoxia generation in the northern Gulf of
550 Mexico, J. Geophys. Res. Ocean., 120, 5019–5039, https://doi.org/10.1002/ 2014JC010634, 2015.

551 Zeileis, A., Kleiber, C., and Jackman, S.: Regression Models for Count Data in R, J. Stat. Softw., 27,
552 https://doi.org/10.18637/jss.v027.i08, 2008.