Bayesian mechanistic modeling characterizes Gulf of Mexico hypoxia: 1968–2016 and future scenarios

Dario Del Giudice [iD],[1,3] V. R. R. Matli,[2] and Daniel R. Obenour[1,2]

[1]*Department of Civil, Construction & Environmental Engineering, NC State University, Raleigh, North Carolina 27695 USA*
[2]*Center for Geospatial Analytics, NC State University, Raleigh, North Carolina 27695 USA*

Abstract. The hypoxic zone in the northern Gulf of Mexico is among the most dramatic examples of impairments to aquatic ecosystems. Despite having attracted substantial attention, management of this environmental crisis remains challenging, partially due to limited monitoring to support model development and long-term assessments. Here, we leverage new geostatistical estimates of hypoxia derived from nearly 150 monitoring cruises and a process-based model to improve characterization of controlling mechanisms, historic trends, and future responses of hypoxia while rigorously quantifying uncertainty in a Bayesian framework. We find that November–March nitrogen loads are important controls of sediment oxygen demand, which appears to be the major oxygen sink. In comparison, only ~23% of oxygen in the near-bottom region appears to be consumed by net water column respiration, which is driven by spring and summer loads. Hypoxia typically exceeds 15,600 km^2 in June, peaks in July, and declines below 10,000 km^2 in September. In contrast to some previous Gulf hindcasting studies, our simulations demonstrate that hypoxia was both severe and worsening prior to 1985, and has remained relatively stable since that time. Scenario analysis shows that halving nutrient loadings will reduce hypoxia by 37% with respect to 13,900 km^2 (1985–2016 median), while a +2°C change in water temperature will cause a 26% hypoxic area increase due to enhanced sediment respiration and reduced oxygen solubility. These new results highlight the challenges of achieving hypoxia reduction targets, particularly under warming conditions, and should be considered in ecosystem management.

Key words: Bayesian inference; climate change; dead hypoxic zones; eutrophication; Gulf of Mexico; hindcasts and projections; process-based modeling; riverine nitrogen; uncertainty quantification.

Introduction

The northern Gulf of Mexico harbors one of the largest hypoxic or "dead" zones in the world, often exceeding 20,000 km^2 in mid-summer (Obenour et al. 2013). Hypoxic zones are areas with low bottom water oxygen concentration (<2 mg/L), which can have a number of negative effects on aquatic ecosystems and coastal economies (Diaz and Rosenberg 2008, Smith et al. 2017). Even though an Action Plan to reduce the spatial extent of Gulf hypoxia to 5,000 km^2 has been in place since 2001 (HTF 2001), the area of bottom-water hypoxia does not appear to be decreasing (Forrest et al. 2011, Obenour et al. 2015). In this context, mathematical models are essential tools to understand how hypoxia will respond to future nutrient management measures and hydroclimatic conditions that influence the development and severity of hypoxia (Scavia et al. 2017).

Models of intermediate (Justić et al. 2002, Scavia et al. 2013, Obenour et al. 2015) and high (Justić and Wang 2014, Feist et al. 2016, Fennel et al. 2016) mechanistic complexity have linked Gulf hypoxia to both natural and anthropogenic causes. Stratification, which inhibits reoxygenation of bottom waters, is principally a natural phenomenon influenced predominantly by wind and freshwater discharge (Wiseman et al. 1997, Obenour et al. 2012). Anthropogenic impacts include nutrients mostly leached from agricultural watersheds and transported through the Mississippi and Atchafalaya Rivers (MARs; McIsaac et al. 2001). Once in the Gulf, these nutrients promote the production of organic matter (i.e., phytoplankton), which is ultimately subject to microbial decomposition that depletes oxygen in the bottom waters (Justić et al. 2002, Turner et al. 2008).

While the broad causes of Gulf hypoxia are understood, there is uncertainty related to how oxygen depletion responds to anthropogenic and environmental factors over different temporal scales, from daily to decadal (Obenour et al. 2015, Yu et al. 2015). Past studies have explored some of these issues using coupled hydrodynamic and water quality modeling (Hetland and

DiMarco 2008, Laurent et al. 2018). However, there has been limited uncertainty quantification, which is essential to understand how much confidence should be placed in model applications to future scenarios (Reichert and Borsuk 2005, Robson 2014). Two notable exceptions include Obenour et al. (2015), which focused on probabilistic inference of biophysical parameters controlling interannual variability in mid-summer hypoxia, and Mattern et al. (2013), which tested hypoxia model sensitivity to various uncertain model inputs. However, there remains a need for studies that rigorously characterize parameter uncertainties and allow for probabilistic predictions throughout the hypoxic season. Additionally, most models of hypoxia formulate sediment oxygen demand (SOD) either as a constant process disconnected from incoming nutrients (Hetland and DiMarco 2008, Obenour et al. 2015) or, at the opposite end of the spectrum, as instantaneous remineralization of the settled organic matter (Fennel et al. 2013). These parameterizations, which are not well-suited to study the delayed impacts of nutrient loads over multiple seasons, may limit our ability to accurately predict hypoxia, especially considering indications that SOD is the predominant oxygen sink for Gulf bottom waters (Dortch et al. 1994, Quiñones-Rivera et al. 2007, Yu et al. 2015). Finally, while studies have investigated the effects of either reducing nutrient loads (Scavia et al. 2013, Feist et al. 2016, Fennel and Laurent 2018) or climate warming (Justić et al. 1996, Laurent et al. 2018), there is currently no probabilistic assessment of the potentially countervailing effects of increased temperature and nutrient reductions on Gulf hypoxia.

To address these issues, we enhance the mechanistically parsimonious model of hypoxia from Obenour et al. (2015) hereafter referred to as OMS15. Parsimonious models fall within the modeling spectrum ranging from simple regressions (Greene et al. 2009, Le et al. 2016, Del Giudice et al. 2018b) to high-resolution hydrodynamic-biogeochemical models (Meier et al. 2011, Fennel et al. 2016), and have been successfully employed to understand and predict eutrophication-related phenomena in the Gulf (Justić et al. 1996, Scavia and Donnelly 2007, Scavia et al. 2013, Obenour et al. 2015) and other aquatic systems (Rucinski et al. 2010, Chapra et al. 2016). Also, we leverage new, geostatistically derived estimates of Gulf dissolved oxygen from Matli et al. (2018) hereafter referred to as MFO18. Compared to previous Gulf studies focusing only on mid-summer hypoxic conditions (Greene et al. 2009, Forrest et al. 2011, Scavia et al. 2013, Obenour et al. 2015), these new estimates allow for calibration and prediction of hypoxia over the entire summer. The main objectives of this study are to (1) provide new insights into oxygen demand by incorporating autumn-spring loads in a new formulation of sediment respiration; (2) generate for the first time, daily probabilistic (Bayesian) predictions of hypoxic area (HA) over June–September; (3) hindcast oxygen demands and HA from 1968 to 1984 (before regular monitoring began) and assess long-term

trends in Gulf hypoxia through 2016; and (4) conduct scenario analyses of the effects of changing nitrogen loads, hydrology, and water temperatures on future HA and seasonal duration.

Materials and Methods

Study area and data

We consider the Louisiana–Texas (LaTex) shelf from Galveston Bay to the Mississippi River Delta, where hypoxia is most common (Wiseman et al. 1997, Matli et al. 2018). The shelf is divided into east and west sections relative to the Atchafalaya River outflow (Appendix S1: Fig. S1), as these regions can have different bottom-water dissolved oxygen (BWDO) dynamics (Obenour et al. 2013, Fennel et al. 2016). We take advantage of a large sample of cruises from multiple monitoring programs ($n = 149$; Matli et al. 2018) to characterize the variability in oxygen conditions on the LaTex shelf. For each cruise date, scattered measurements of BWDO are converted through geostatistical modeling to "observations" of section-specific average BWDO (Appendix S1: Table S1) and shelf-wide HA (Matli et al. 2018). Specifically, we use MFO18's space-time geostatistical approach, which associates uncertainty bounds to estimates of BWDO and HA.

To model hypoxia, we use multiple environmental inputs. Riverine data include monthly discharge and nitrogen (N) loads from the U.S. Geological Survey (data *available online*).[4] As in OMS15, we use total bioavailable N load, which is composed of nitrate, nitrite, ammonia, and 12% organic nitrogen. We also use daily discharge data from the U.S. Army Corps of Engineers at Simmesport and Tarbert Landing to estimate daily loads for the Atchafalaya and the Mississippi River, respectively (data *available online*).[5] As in OMS15, we use wind inputs (speed and zonal velocity) derived from National Data Buoy Center (NDBC) stations (Appendix S1: Fig. S1; data *available online*).[6]

To reconstruct hypoxia from 1968 to 1984, data sources described above required some imputation. Similar to Scavia et al. (2003) we impute missing MARs loadings (Appendix S1: Section S1) using a regression with USGS Mississippi NO_2+NO_3 loads and river flows, which started being recorded in 1967. For wind, we extend NDBC data backward from 1985 using NCAR Reanalysis velocities (Kalnay et al. 1996) extracted from two 2.5° square grid cells covering a large part of the modeling domain and being centered at 27.5° and 30° latitude and −92.5° longitude (data *available online*).[7]

---

[4] https://toxics.usgs.gov/hypoxia/mississippi/nutrient_flux_yield_est.html
[5] https://www.mvn.usace.army.mil/Missions/Engineering/Stage-and-Hydrologic-Data/Historical-Discharges/
[6] https://www.ndbc.noaa.gov/
[7] https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.pressure.html

NCAR wind inputs for 1968–1984 are bias-corrected through comparison with NDBC inputs during their overlapping period (1985–2016).

Mechanistic model

We model oxygen dynamics using a process-based formulation of intermediate biochemical complexity. This mass balance model, developed and validated by OMS15, represents each shelf section as two compartments separated by the pycnocline (Fig. 1). Freshwater and N from rivers are assumed to be advected above the pycnocline by wind-driven coastal currents. Late spring and summer N loads are transformed into algal production, represented as organic carbon (C), which settles to the lower compartments. Here, dissolved oxygen is lost due to the microbial decomposition of C, represented in the model as water column oxygen demand (WCOD). Oxygen levels are further depleted through sediment oxygen demand (SOD), which, unlike OMS15, now varies based on empirical relationships with long-term N loading through March. We note that while much of the WCOD consumption is expected to occur in the water column below the pycnocline (Murrell and Lehrter 2011), it is likely that some of this near-term production settles to the seafloor over the course of the summer (Rowe et al. 2002). Thus, SOD and WCOD primarily reflect the effects of N loadings over different time scales, similar to the representation of oxygen demands in Katin et al. (2019), rather than represent the exact depths at which oxygen consumption occurs. Finally, oxygen is replenished through exchanges with the surface layer, which is approximately saturated.

OMS15 model is based on a steady-state solution to mass balance differential equations. Using time-varying inputs, we predict daily BWDO (O_b) concentration (mg/L) from 1 June to 30 September of each year:

$$O_b = \frac{(k_a O_s - D_w)}{k_a - D_s / O_f} - 1 \qquad (1)$$

where k_a is the reaeration rate (m/d), D_w is the WCOD, and D_s is the SOD at O_f, a reference oxygen concentration set to 3 mg/L. The solution is adjusted based on a reanalysis of cruise profiles, which indicates that BWDO is approximately 1 mg/L lower than the average oxygen concentration below the pycnocline. As we model oxygen dynamics over four months, we set the saturation concentration, O_s, to 6.54, 6.52, 6.48, 6.53 mg/L for June, July, August, September, respectively. These concentrations are derived (Greenberg et al. 1992) from average surface layer temperature and salinity measurements and suggest little temporal variability.

Consistent with OMS15, the net WCOD ($g \cdot m^{-2} \cdot d^{-1}$) of each lower compartment is represented as

$$D_w = J\gamma\omega = \left(\lambda \frac{L_r + L_u}{(Q_r + Q_u + Q_g)/\nu + A}\right)\gamma\omega. \quad (2)$$

here, J is the downward carbon flux ($g \cdot m^{-2} \cdot d^{-1}$), γ is the mass ratio of oxygen demand to organic carbon set to 3.5 (Justić et al. 1996, Chapra 1997), λ is the ratio of organic carbon to nitrogen set to 5.68 (Redfield et al. 1963), A is the area of the shelf section (Gm^2), ω is an oxygen demand adjustment factor, and ν is the effective settling velocity (m/d), which incorporates both the production and sinking of organic matter. These latter two parameters are calibrated through Bayesian inference. The variables Q and L represent the near-term flows (Gm^3/d) and N loads (Gg/d) entering the upper compartments, respectively. Subscripts r, u, and g denote the origin of these fluxes: MARs, upstream (i.e., eastern) shelf section, and the greater Gulf of Mexico. Q_g is approximated as a dilution factor (3.2, derived from surface salinity data) multiplied by mean Mississippi River discharge (1.6 Gm^3/d).
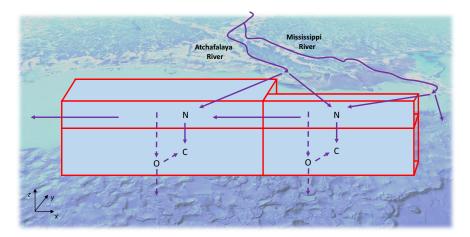

Fig. 1. Schematic representation of the LaTex shelf, its conceptual compartmentalization, and main fluxes considered by the model. N, O, and C stand for nitrogen, oxygen, and carbon, respectively. Oxygen sources and sinks are represented by dashed lines.

The reaeration rate for each section is determined as a function of wind stress (representing shear-induced turbulence) and freshwater flow (representing buoyancy):

$$k_a = \beta_0 + \beta_1 \frac{U^2}{Q_s} \frac{A}{10{,}000} \qquad (3)$$

where U is the 14-d weighted mean wind speed for the shelf section (m/s), Q_s is the river discharge entering the section [Gm^3/d], and β_0 and β_1 are calibration parameters.

The model uses near-term inputs 30–90 and 0–60 d before the date of prediction. The 30–90 d inputs regulate WCOD, taking into account the additional time required for algal growth and subsequent trophic processes (e.g., grazing, excretion, decomposition). For the mid-summer cruises taking place around the end of July, this period approximately covers May–June, which have been shown to be the months whose nutrient loads are highly predictive of mid-summer hypoxic area (Scavia et al. 2003, Forrest et al. 2011). Shorter-term (0–60 d) inputs instead regulate water stability and thus reaeration (Obenour et al. 2015). For both inputs, partitioning is computed through

$$F_W = 0.5 - \beta_e v_e \qquad (4)$$

where F_W is the fraction of abovementioned flows and loads transported westward over the shelf, v_e is the mean eastward wind velocity (m/s), and β_e is a calibration parameter. The 0.5 indicates that, in absence of wind, inputs from both rivers would equally partition westward and eastward (Obenour et al. 2015).

The main mechanistic enhancement with respect to OMS15 is a more realistic SOD formulation that is also unique relative to other Gulf hypoxia models (Fennel et al. 2013, Feist et al. 2016). While the original model only included a constant respiratory flux B, SOD is now represented as

$$D_s = B \sqrt{\frac{L}{\bar{L}}} \theta^{T - \bar{T}} \qquad (5)$$

where L (Mg/month) is the combined nutrient loading from the MARs, averaged until March, which is the month when near-term nutrient inputs start affecting WCDO and thus BWDO in early summer. We normalize these pre-spring loads relative to their long-term average ($\bar{L}$) for the study period. The load dependency is square-root transformed, reflecting the fact that sediment-related oxygen demands tend to saturate at high levels of organic matter (Di Toro 2001). Also, SOD is temperature dependent because higher temperatures increase the metabolism of microbes decomposing organic matter in the sediments (Thamdrup et al. 1998, Hetland and DiMarco 2008). Temperature dependence is based on the Arrhenius model with $\theta = 1.07$, a formulation often used to correct respiration rates (Chapra 1997, Gujer 2008). Rates are corrected when temperatures deviate

from $\bar{T}$, the summertime average. Here, T is the monthly mean temperature determined from cruise measurements at a water depth of approximately 30 m (25.6, 26.2, 25.7, and 25.4 °C for June, July, August, and September, respectively).

In the model, SOD represents the remineralization of less labile organic matter that is produced and settled before summer. This formulation reflects the expectation that production over the long-term can lead to accumulation of organic carbon in coastal sediments (Justić et al. 2002), which influences the amount of detritus available for remineralization in later summer months (Turner et al. 2008). Multiple accumulation windows are tested, beginning in the preceding autumn or multiple years back (starting in October, the beginning of the water year). This wide range of windows is motivated by previous studies indicating that hypoxia may be influenced by loadings occurring over the previous year (Matli et al. 2018) or over multiple years (Turner et al. 2006, Del Giudice et al. 2018b). We test these alternative loading windows based on their ability to improve the model's predictive skill for BWDO.

Inference and predictions

Our approach combines the advantages of mechanistic modeling with those of Bayesian inference, enabling us to make probabilistic statements on parameters and predictions of water quality (Arhonditsis et al. 2008, Sikorska et al. 2015). We describe the BWDO response (y_e and y_w for east and the west shelf sections, respectively) as a combination of mechanistic model output (O_b from Eq. 1) and a Gaussian white noise error $\in$ largely representing model structural uncertainty:

$$y = O_b(\vartheta) + \in (\vartheta). \qquad (6)$$

Both the mechanistic and stochastic components of the model have parameters ϑ to be estimated. In a Bayesian framework, parameter calibration is equivalent to estimating the posterior probability distribution given the BWDO concentration observations (y_e, y_w):

$$p(\vartheta | y_e, y_w) \propto p(\vartheta) p(\bar{J}_e(\vartheta)) p(y_e | O_b(\vartheta), \sigma_e(\vartheta))$$
$$p(y_w | O_b(\vartheta), \sigma_w(\vartheta)). \qquad (7)$$

here, $p(\vartheta)$ represents the joint prior distribution that formalizes our knowledge of parameters (Appendix S1: Table S2) before introducing the cruise-based measurements of BWDO (Appendix S1: Table S1). Model parameters include both process coefficients (e.g., v, the effective settling velocity) and error metrics. In addition, $p(\bar{J}_e(\vartheta))$ represents a prior constraint for the average eastern-section carbon flux of $\mathcal{N}$ (0.29,0.05) g·m^{-2}·d^{-1} (Redalje and Fahnenstiel 1994, Obenour et al. 2015). Finally, $p(y_e | O_b(\vartheta), \sigma_e(\vartheta)) p(y_w | O_b(\vartheta), \sigma_w(\vartheta))$ is the likelihood function, which represents the probability that the measured BWDO is generated by the biophysical-

Bayesian model. Here, $\sigma_e(\vartheta)$ and $\sigma_w(\vartheta)$ are standard deviations for the two shelf sections, which combine $\sigma_{m,e}$ and $\sigma_{m,w}$, model structural uncertainty from $\in(\vartheta)$, and measurement uncertainty specified by the geostatistical model (Matli et al. 2018).

After inference based on oxygen concentrations, we predict HA (y_H) at the bottom of each section by using a transformation, $g()$, of BWDO with error π (Appendix S1: Section S2):

$$y_H = g(O_b(\vartheta) + \in(\vartheta)) + \pi. \qquad (8)$$

The posterior predictive distribution of HA is computed via

$$p(y_H|y_e, y_w) \propto \int p(y_H|\vartheta)p(\vartheta|y_e, y_w)d\vartheta. \qquad (9)$$

This marginalization over the posterior distribution is performed via Monte Carlo simulations by propagating 1,000 posterior parameter samples through Eq. 8, enabling us to generate time series of HA that account for total predictive uncertainty. Inference and prediction are computed in R version 3.5.3 (R Core Team 2019). Sampling of the posterior uses an adaptive Metropolis algorithm (Haario et al. 2001, Del Giudice et al. 2018a).

Scenario projections

The model is used to predict the effects of reduced nutrient loading and increased temperature and freshwater discharge. We combine the strengths of ceteris paribus analysis, in which sensitivities to individual input types are quantified (Obenour et al. 2015, Scavia et al. 2017, Del Giudice et al. 2018b), and classical scenario analysis, in which multiple input types are plausibly varied at once (Justić et al. 1996, Meier et al. 2011, Laurent et al. 2018). Specifically, we consider four main types of changes: reduced MAR nutrient concentrations relative to historical levels, increased water temperatures, and reduced nutrients combined with increased temperature with or without increased discharges. For nutrient management scenarios, we select a range of reduction percentages comparable to previous assessments (HTF 2017, Scavia et al. 2017). For warming scenarios, we consider temperature increases consistent with previous studies of climate change in the Gulf (Justić et al. 1996, Laurent et al. 2018). For scenarios combining nutrient decreases and higher water temperatures, we select +2°C, congruous with other studies of hypoxia under climate change projections (Justić et al. 1996, Del Giudice et al. 2018b, Irby et al. 2018, Laurent et al. 2018). Scenarios consider the effect of temperature on both oxygen saturation (O_s) and benthic respiration (D_s) through the mechanisms described above. Consistent with Irby et al. (2018), we consider a homogeneous increment in surface and bottom water temperatures as both are likely to increase over the long term. This approach also represents a parsimonious solution to the discordant findings on warming of Gulf waters: according to data analysis

bottom temperatures are warming more than surface temperatures (Turner et al. 2017), yet some modeling results indicate the opposite (Laurent et al. 2018). While temperature increases are the most likely effect of climate change in the Gulf (Biasutti et al. 2012), we additionally consider a scenario group in which changes in nutrients and temperature are accompanied by a 10% increase in river discharge owing to climate change (Laurent et al. 2018). Finally, we acknowledge that other driving factors, such as winds, may also be affected by climate change. However, we focus exclusively on warming and increased freshwater inputs, as projections for other factors are less certain in terms of both magnitude and direction (Biasutti et al. 2012, Feng et al. 2012, Laurent et al. 2018).

For each scenario, we recalculate the distribution of average hypoxic extent and duration over the 32-yr study period, under the proposed changes in riverine nutrient concentration, flow, and/or water temperature. In practice, the extent distribution is approximated by propagating a large sample of the posterior $p(\vartheta|y_e, y_w)$ through the model (Eq. 8) and averaging each model realization through time. The duration of hypoxia is calculated by counting, for each realization, the fraction of days having HA exceeding a given threshold.

RESULTS AND DISCUSSION

Impact of variable SOD rate on model performances

In this study, we explore how SOD varies based on long-term N loading, which represents an important modeling enhancement relative to OMS15 and other hypoxia modeling studies. The variable SOD formulation is compared to a null model utilizing a constant SOD. Calibrating this null model to the BWDO data ($n = 149$, per shelf) leads to $R^2 = 0.576$ (fraction of variance explained). As we hypothesized, including pre-spring loading (L) in the model's SOD formulation (Eq. 5) generally enhances predictive performances. Specifically a loading window beginning in September, October, or November results in the best model performances ($R^2 = 0.638$ for all three), whereas considering longer or shorter loading periods leads to a gradual performance deterioration (e.g., $R^2 = 0.632$ for August, $R^2 = 0.636$ for December). Multi-annual loads lead to little or no improvement with respect to the null model. For instance, starting loading accumulation in October of two or seven years earlier only leads to a R^2 of 0.579 or 0.571, respectively. The finding that autumn–winter loads increase the (absolute) variance explained by the model by more than 6% without the need for additional calibration parameters is remarkable, as previous models of Gulf hypoxia have focused mostly on May (Forrest et al. 2011, Obenour et al. 2012, Scavia et al. 2013) or May–June (Scavia et al. 2003, Obenour et al. 2015) loads. Thus, these results emphasize the role of sediments in accumulating oxygen demand over longer

time periods, consistent with what has been found for other hypoxic systems such as the Baltic Sea (Pitkänen et al. 2001) and Lake Erie (Del Giudice et al. 2018a). At the same time, the time scales of nutrient loading most influencing Gulf SOD and hypoxia (<1 yr) appear to be shorter than the multi-year time scales of those more enclosed systems. This outcome is consistent with the Gulf of Mexico being a relatively dynamic system for which seasonal accumulation of organic matter is more important than multi-annual storage (Justić et al. 2002). Interestingly, November–April riverine inputs have been shown to influence SOD in another hypoxic coastal system (Katin et al. 2019). Overall, considering nutrient loadings before November does not improve model performances further, which is also consistent with the regression results of MFO18. Therefore, the SOD formulation of the final model uses average November–March N loading from MARs.

Bayesian parameter estimates

Calibration of the final model generates posterior parameter distributions $p(\vartheta|y_e, y_w)$ that are more precise than the priors (Fig. 2). This high posterior identifiability suggests that the model structure is appropriate and the amount of calibration data is sufficient (Omlin and Reichert 1999, Del Giudice et al. 2015, McElreath 2018). The average correlation between parameters is low ($r = 0.10$), yet is moderate for β_1 and B ($r = 0.85$), which implies that the relative importance of reaeration and average SOD is well constrained yet their absolute value can be somewhat uncertain. This latter correlation is not surprising as the two parameters have similar yet opposing effects on BWDO: an increase in β_1 enhances mean reaeration while an increase in B increases mean SOD. Of these two parameters, B appears to be more robustly resolved than β_1, considering its lower coefficient of variation (16.9% vs. 19.4%) and smaller shift with respect to the original OMS15 posteriors (2.3% vs. 87.3%). Compared to OMS15, parameter uncertainties are lower, on average, which largely reflects the availability of almost six times the previous number of geostatistical BWDO observations.

Posterior estimates of effective settling velocity v and average sediment respiration rate B are similar to OMS15's estimates. These consistencies indicate that the mid-summer BWDO data used in OMS15 are sufficient to estimate v and B, and that these parameters are robust to moderate changes in model formulation (e.g., introduction of Eq. 5) and a substantial increase in calibration data (27 vs. 149 cruises). The consistency of the sediment respiration parameter is notable, as other models have shown high sensitivity of hypoxia to different SOD representations (Fennel et al. 2013). At the same time, the richer calibration data set enables us to refine estimation of the net water column demand as controlled by the WCOD adjustment (ω), reaeration coefficients (β_0 and β_1), and the east-west advection

coefficient (β_e). The lower ω suggests that off-shelf losses of primary production, horizontal influx of oxygenated water, and/or photosynthetic oxygen production might be more important than previously thought. The latter interpretation is consistent with Yu et al. (2015) who found that primary productivity can offset up to 72% of the total respiration below the pycnocline. Our reaeration rates for the east and west sections range from 0.18–0.72 and 0.21–0.66 m/d (95% intervals), respectively, which is comparable to the range of 0.17–0.86 m/d determined by Justić et al. (2002). Additionally, low ω could also represent a portion of the seasonally-produced organic matter settling and being subsumed into SOD. Model residual (error) standard deviations ($\sigma_{m,w}$, $\sigma_{m,e}$) have posteriors very similar to those in OMS15, at about 0.35 mg/L. This demonstrates that a model initially developed to predict mid-summer oxygen depletion is sufficiently adaptable to predict the evolution of hypoxia over the entire summer with only a small increase in residual uncertainty. These error parameters also show that, when using the same (current) BWDO data, structural uncertainty of the current model is lower than that of OMS15 model (Appendix S1: Fig. S3), further supporting adoption of the updated SOD formulation.

Oxygen demand apportionment and multi-annual trends

The model distinguishes between two types of oxygen consumption. The first is associated with the decay of less labile organic matter accumulated in the sediments (i.e., SOD), which is influenced by November–March nutrient loads. The second is associated with organic matter from spring and summer N loads, which is expected to manifest largely as WCOD. Both oxygen sinks vary over time. For the period of recorded hypoxia (1985–2016), the two oxygen sinks have different magnitudes ($g \cdot m^{-2} \cdot d^{-1}$), with SOD (mean = 0.33, SD = 0.05) being three times larger than net WCOD (mean = 0.10, SD = 0.05). The fraction of overall oxygen demand accounted for by SOD ranges from 70% in 1990 to 88% in 1988 (Fig. 3). The finding that sediments are the predominant sink of BWDO is in contrast with previous studies that have attributed preponderant importance to WCOD (Scavia et al. 2003, 2013, Murrell and Lehrter 2011). However, our partitioning of WCOD and SOD is consistent with findings that 68% of respiration in the bottom 5 m (Yu et al. 2015) and ~74% of respiration in the bottom ~1 m (Dortch et al. 1994, Quiñones-Rivera et al. 2007) are derived from the sediments. Additionally, calculated SOD is within the approximate range of 5–30 mmol $O_2 \cdot m^{-2} \cdot d^{-1}$ (0.16–0.96 g $O_2 \cdot m^{-2} \cdot d^{-1}$) reported by other Gulf modeling (Yu et al. 2015) and field studies (Dortch et al. 1994, Rowe et al. 2002, Murrell and Lehrter 2011). While our SOD results are largely consistent with the studies discussed above, we note that determinations of water and sediment respiration can depend
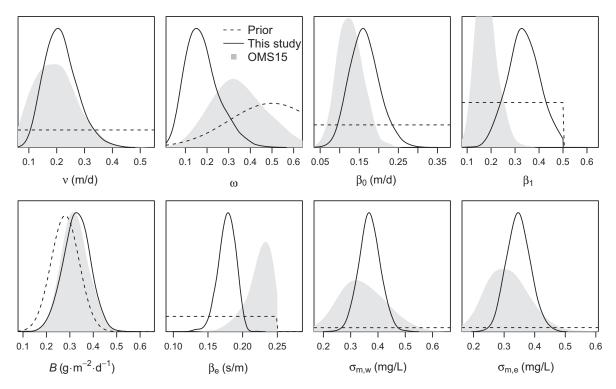
FIG. 2. Prior and posterior distributions of six biophysical process parameters and two error parameters (Appendix S1: Table S2). Parameters estimated in this study are also compared to previous estimates using the same priors (Obenour et al. 2015). Posteriors are depicted as smoothed kernel density estimates based on 1,000 Markov Chain Monte Carlo samples. The y-axis represents probability density.

on context-specific definitions of the bottom water layer (Yu et al. 2015). As discussed in the introduction to the mechanistic model (in *Methods*), WCOD and SOD are defined conceptually, rather than at precise vertical locations, in order to distinguish the impact of near-term vs. longer-term nutrient inputs. Interestingly, annual SOD and WCOD are only weakly correlated ($r = 0.29$), implying that years with lower than average spring and summer loads can still experience intense hypoxia caused by high November–March loads. An example is 2005, which had high SOD (81st percentile) and low WCOD (19th percentile) but experienced HA 12% larger than average. In comparison, for 2005, models that did not directly account for winter loads tended to underpredict HA (Fennel et al. 2013, Yu et al. 2015). Overall, HA is strongly correlated to total oxygen demand ($r = 0.86$), and moderately to WCOD ($r = 0.80$) and SOD ($r = 0.64$; Fig. 2). Interestingly, WCOD is more strongly correlated with HA than SOD, despite the latter being a more important oxygen sink. A plausible reason is that annual WCOD, but not SOD, is also somewhat anticorrelated with mean reaeration ($r = -0.43$). Seasonal nutrient loading and stratification (enhancing water stability and thus impeding reaeration) tend to be somewhat correlated based on their common dependence on river flow (Obenour et al. 2012).

Analysis of years with unusual hypoxia shows that they are driven either by unusual nutrient loading (and thus oxygen demand) or hydrometeorologic conditions (and thus reaeration). Specifically, model results indicate that 1993 experienced the highest oxygen demand and HA, while 2000 experienced the lowest demand and HA. The year 1988 experienced similarly low HA, yet with a total oxygen demand that was only 11% lower than the 1985–2016 mean. However, the summer of 1988 experienced above-average wind stress (Scavia et al. 2013) and the lowest summer discharge, leading to high reaeration of bottom waters (Eq. 3).

Our hindcasts, covering five decades, shed light on whether hypoxia in the Gulf has been worsening. Prior studies have argued that BWDO has been decreasing (Justić et al. 2002) and that the spatial extent of Gulf hypoxia has been increasing since the 1980s (Turner et al. 2006, 2008) and even doubled since 1985 (Sylvan et al. 2006). However, more recently OMS15, analyzing model residuals, found no significant decrease in BWDO for 1985–2011. Our new estimates (Fig. 3) also do not exhibit a significant multi-annual trend for 1985–2016, yet show that the propensity of the system to become hypoxic has increased over the period leading up to the beginning of regular hypoxia monitoring (1985). In particular, between 1968 and 1985, we estimate that oxygen demand increased at a rate of
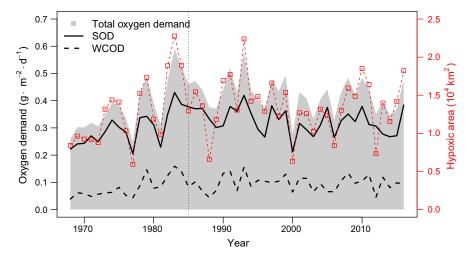
FIG. 3. Modeled summertime (June–September) mean bottom-water dissolved oxygen (BWDO) consumption and hypoxic extent for the LaTex Shelf (1968–2016). Model output is obtained using the mean of the posterior parameter distributions. While water column oxygen demand (WCOD; here an area-weighted average of net water column respiration of east and west shelf sections) is regulated by spring and summer nutrient loads, sediment oxygen demand (SOD) is controlled by November–March loads. The vertical dotted line marks the year 1985 when hypoxia monitoring cruises started.

14 mg·m^{-2}·d^{-1} per year ($p = 3 \cdot 10^{-4}$) and HA at a rate of 530 km^2 per year ($p = 5 \cdot 10^{-3}$). These results are generally consistent with studies that indicate there was a substantial increase in hypoxia in the late 1970s and early 1980s due to increasing nitrogen loads (Justić et al. 2002, Scavia et al. 2003, Scavia and Donnelly 2007, Turner et al. 2008). However, our results conflict with the conclusion that, before 1978 or 1981 (depending on the particular study; Justić et al. 2002, Scavia et al. 2003, Turner et al. 2006, 2008), loads were insufficient to generate systematic hypoxia. Instead, we find that, from 1968 to 1979, summertime HA ranged from 5,900 (lowest HA) to 17,400 km^2 (88th percentile of HA, 1985–2016), which is comparable to the estimates of Scavia and Donnelly (2007) and Greene et al. (2009). When compared to all previous hindcasting studies, the model applied here benefits from a longer and more accurate calibration data set (Matli et al. 2018). Longer calibration spanning a variety of conditions is likely to generate more representative model parameters (Del Giudice et al. 2018a). Finally, the finding that hypoxia was both substantial and increasing prior to 1985 is corroborated by paleoindicators in Gulf sediments (Rabalais et al. 2007). Specifically, sediment cores show that the shelf has experienced substantial oxygen stress since the 1950s and that conditions have been deteriorating at least until the 1980s.

Daily predictions and intraseasonal variability of hypoxic area

For every summer day, our Bayesian mechanistic model can predict HA and simultaneously quantify three major types of uncertainties, specifically associated with calibration parameters, system representation, and transformation from BWDO to HA. Additionally, the method explicitly accounts for a fourth source of uncertainty, namely that associated with BWDO observations. This approach is different from most Gulf modeling studies that only showed the most likely conditions yet did not quantify the impact of model limitations and imperfection of calibration data (Justić et al. 2002, Turner et al. 2006, Justić and Wang 2014). However, rigorous uncertainty analysis can provide essential information for environmental planning and management (Reichert and Borsuk 2005, Arhonditsis et al. 2008), for instance answering questions about the probability of exceedance of critical thresholds over various time scales.

Daily predictions of HA show that total uncertainty of the model is comparable and often lower than the uncertainty around the geostatistical HA observations (Fig. 4), particularly for observations corresponding to cruises with relatively sparse sampling. Model predictions and observations generally agree when distinguishing between months with mild and severe hypoxia (Appendix S1: Figs. S4–S11), with the model predicting hypoxia to be maximum in July in 56% of the years, June in 38% of years, and August in 6% of years. Consistent with these results, July has the largest average hypoxic area (15,800 km^2) and September has the lowest (9,900 km^2) of the summer season. When comparing our intraseasonal predictions with those from other mechanistic models, there are notable differences. For example, predictions from Yu et al. (2015) suggest average June HA is mostly below 5,000 km^2, and HA peaks in August or September in 75% of years. A similar apparent underestimation of HA in June and
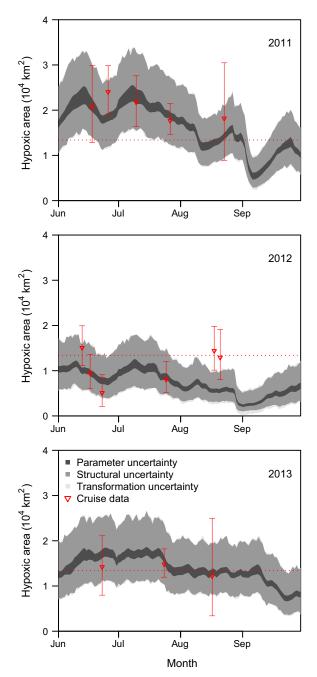
FIG. 4. Daily model predictions of hypoxia for the whole LaTex Shelf for years with severe (2011), mild (2012), and intermediate (2013) hypoxia. The gray areas indicate 95% interquantile ranges of the predictive distribution (see Eq. 9 for total uncertainty) and highlight the effect of parameter uncertainty, model structural error, and uncertainty in transformation from BWDO to hypoxic area (often negligible). The mean and 95% interquantile ranges of the geostatistical estimates from cruise data are also displayed (Matli et al. 2018). The dotted horizontal lines represent the average of all cruise-based estimates of hypoxic areas and help distinguish between years with mild or severe hypoxia. Appendix S1: Figs. S4–S11 show all years.

overestimation in September is also observed in other mechanistic model formulations (Fennel et al. 2016). A likely reason for the more accurate representation of intraseasonal hypoxic patterns is our extensive and robust data set of shelf-wide BWDO that spans hypoxic conditions from late spring to early autumn, thus enabling a more realistic tuning of model parameters.

Scenario projections

Projection results highlight the sensitivity of average hypoxic area to specific changes in watershed nutrient management and climate (Fig. 5). While these scenarios are not intended to represent HA at any specific time, they seek to quantify how historic hypoxia could be affected by realistic nutrient (HTF 2017) and hydroclimatic (Laurent et al. 2018) conditions for the mid-to-late 21st century. The reference case for the projections is the June-September average over 1985–2016, which has a median of approximately 13,900 km^2. Given the long period over which averages are taken, uncertainty in these projections primarily reflects parameter uncertainties. Our projections also indicate the probability that hypoxia exceeds a specific threshold of extent (5, 10, or 15·10^3 km^2) during a summer day (1 June to 30 September).

Results of nutrient reduction scenarios (Fig. 5A and D) show that, under constant climatic conditions, reducing riverine nutrient loads (while keeping discharges constant) is expected to reduce hypoxia through its effect on WCOD and SOD. Within the range of scenarios analyzed, the average response of the system is almost linear with a 50% nutrient reduction translating into a 37% reduction in total HA (35% reduction in the west section and 41% in the east). This quasilinear response of HA averaged over multiple years is consistent with results of more complex biogeochemical modeling (Feist et al. 2016). Besides calculating average hypoxia, our probabilistic predictions naturally lend themselves to quantifying the days with hypoxia larger than a given areal threshold, which provides a more complete picture of the occurrence of severe oxygen depletion. Taking a 50% nutrient reduction as an example, the curves show that <10% of summer days would witness HA > 15,000 km^2 while HA > 5,000 km^2 would still occur more than 80% of the summer.

With no nutrient reductions, increased temperature is likely to exacerbate oxygen depletion by enhancing sediment respiration rates and reducing oxygen saturation concentrations and thus reoxygenation (Fig. 5B and E). In the case of increasing temperatures, the model shows ~12% increase in HA per each degree of warming. While with historic temperatures (1985–2016), only 39% of days are expected to have HA > 15,000 km^2; with +2°C, this number is predicted to increase to 62%. This temperature increase is similar to the warming to be expected at the end of the century in a "business as usual" scenario (Laurent et al. 2018). While an increase
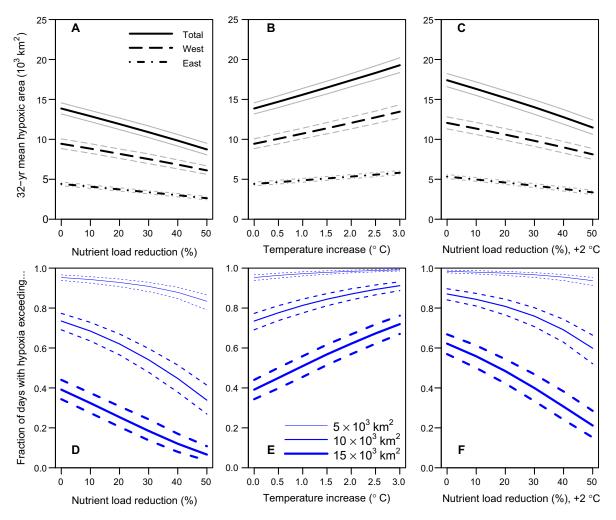
FIG. 5. Ceteris paribus scenario analysis of impacts of nitrogen reduction and/or water temperature increase on Gulf hypoxia. On the top (A, B, C), the average summertime (June–September) hypoxic extents with median (thick black lines) and 95% credible intervals (thin gray lines) are shown. On the bottom (D, E, F), the fraction of summer days exceeding different thresholds of total hypoxic area is shown.

of 2°C represents a slight extrapolation with respect to the range of summer temperatures considered, our projections are likely to be more robust than previous ones based on models of similar or lower complexity calibrated over shorter periods and using more extreme temperature changes (Justić et al. 1996, Del Giudice et al. 2018b). Panels C and F show the countervailing effects of temperature increase on nutrient reduction. With +2°C, a load reduction of ~30% will be required to maintain HA at historic (1985–2016) levels. Additionally, even with a 50% nutrient reduction, 21% of summer days would witness HA > 15,000 km^2 while HA > 5,000 km^2 would still persist during more than 90% of the summer. Compounding this temperature increase with a plausible 10% increase in riverine discharge (Appendix S1: Fig. S12) will further countervail the efforts to reduce HA, largely due to enhanced stratification and thus reduced oxygen replenishment (Eq. 3). In this scenario

group, a load reduction ~35% will be required to maintain HA at historic levels, while with no nutrient reduction, hypoxia would be on average 30% larger than historic levels. In general, these results corroborate that climate change will exacerbate hypoxia in the Gulf (Justić et al. 1996, Laurent et al. 2018), consistent with studies of other large aquatic ecosystems (Meier et al. 2011, Del Giudice et al. 2018b, Irby et al. 2018).

The current strategy to achieve the target HA of 5,000 km^2 is based on a 45% nutrient reduction (HTF 2017). Our results suggest that even in the conservative case of no climate change, this 45% reduction recommendation will result in an average summer hypoxia of ~9,300 km^2, almost twice the target (Fig. 5). This finding is close to recent multimodel projections (Scavia et al. 2017) computing that a 45% load reduction will reduce average HA to ~9,000 km^2. However, our results suggest the interim 20% nutrient reduction goal recently

recommended by HTF (2017) may have some benefit by reducing the time hypoxia exceeds 15,000 km^2 by almost 40%, at least under current climate (Fig. 5D). We note that this study does not consider all possible changes in environmental factors such as alterations in wind patterns. As explained in *Methods*, climate change impacts beyond effects on temperatures are substantially less certain (Biasutti et al. 2012, Feng et al. 2012, Laurent et al. 2018), which motivated us to focus on more predictable and management-relevant changes.

Modeling approach and future directions

This study takes advantage of mechanistic (deductive) and Bayesian (inductive) modeling to advance our understanding of oxygen depletion processes and drivers and of the temporal variability in Gulf of Mexico hypoxia, from daily to multidecadal scales. Our approach mitigates the risk of overfitting, which can lead to parameter estimates with reduced interpretability and predictive power (McElreath 2018). Overfitting is prevented by making use of prior information on interpretable model parameters and process-based relationships that constrain model outputs. Model robustness is further increased by assimilating a large data set and estimating only six mechanistic model parameters. This hybrid deductive-inductive approach seeks to extract the maximum amount of information from field observations and prior knowledge of the system (Robson 2014) and has advantages over purely statistical models based only on patterns in the data but with no embedded physical or biogeochemical mechanism (Forrest et al. 2011, Feng et al. 2012, Le et al. 2016, Del Giudice et al. 2018b). These inductive models usually have parameters without a clear biophysical meaning, which makes it difficult to incorporate prior information and interpret their parameter values relative to ecosystem processes.

The approach we present also has advantages and disadvantages relative to highly mechanistic three-dimensional hydrodynamic-biogeochemical models (Feist et al. 2016, Fennel et al. 2016). First, simulation with our model is almost instantaneous and only requires easy-to-obtain hydrometeorological and nutrient loading data. Computational efficiency facilitates a long study period and rigorous calibration through Bayesian parameter estimation. Parameters are represented by probability distributions and uncertainties by stochastic processes. This makes our model predictions probabilistic, which is particularly important to assess future impacts of climate change and nutrient management strategies (Reichert and Borsuk 2005). Finally, our SOD formulation taking winter nutrients into account represents a potentially more realistic formulation than the instantaneous remineralization of settled organic matter used in other modeling studies (Fennel et al. 2006, 2013, Feist et al. 2016). A limitation of the current approach is its coarse spatial resolution, which only distinguishes between eastern and

western regions of the shelf. Also, our simple hydrodynamic formulation considers westward transport driven by wind and thus does not represent more complex current patterns. Further, plankton dynamics are currently represented only implicitly and a parsimonious approach is employed to convert nutrients into oxygen demand. Finally, while our model formulation appears to capture the primary drivers of hypoxia, we recognize the value of developing ensemble predictions across a range of formulations, especially when extrapolating over substantial changes in environmental forcings (Meier et al. 2011, Scavia et al. 2017).

Regardless of the complexity of other models, the current study has the advantage of assimilating and comparing results against a large number of cruises ($n = 149$), which is substantially more than previously used for calibration of statistical models or verification of complex mechanistic models. For example, some statistical modeling studies have used as few as 9 or 12 cruise observations (Wiseman et al. 1997, Le et al. 2016), while complex mechanistic models have typically been compared to even fewer HA estimates (Feist et al. 2016, Fennel et al. 2016). Our multidecadal calibration period with multiple cruise observations in most summers helps ensure the robustness of both intraseasonal and interannual modeling results. Additionally, model and data are merged in a Bayesian framework, which probabilistically considers different types of uncertainty.

In summary, combining Bayesian mechanistic modeling and an abundant data set of hypoxia observations, this work generates important insights into controls on oxygen depletion in the Gulf of Mexico. Unique from some previous studies, we find that hypoxia is more severe in June than in September and that a significant increasing trend in hypoxia likely ended in the mid-1980s. Moreover, our results strongly suggest that autumn-winter nutrient loads are a major driver of SOD and hypoxia in the following summer. We further demonstrate how the Bayesian framework allows for daily oxygen predictions and long-term climate and nutrient management scenario forecasts with quantified uncertainties. We envision future advances in deductive-inductive hypoxia modeling to focus on more explicitly representing phytoplankton production. While complex phytoplankton dynamics do not seem necessary to accurately model the variability in Gulf hypoxia (Justić et al. 2002, Scavia et al. 2013), satellite measurements of chlorophyll might help better constrain this additional state variable and thus model parameters (Le et al. 2016). Additionally, we envision the inclusion of phosphorus (P) inputs in the organic matter generation. While N seems more important than P in driving overall hypoxic extent (Turner et al. 2006, Rabalais et al. 2007, Scavia and Donnelly 2007), there is evidence that P represents a proximate limiting nutrient near the Mississippi Delta during May–July (Sylvan et al. 2006, Fennel and Laurent 2018). Finally, as geostatistical estimates of hypoxic volume become more available, the model could

be enhanced to probabilistically predict volume, in addition to hypoxic area. Mid-summer cruise data suggest that hypoxic layer thickness varies substantially over time (Obenour et al. 2013), and volume may respond more strongly than area to nutrient loading reductions (Scavia et al. 2018).

LITERATURE CITED

Arhonditsis, G. B., D. Papantou, W. Zhang, G. Perhar, E. Massos, and M. Shi. 2008. Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management. Journal of Marine Systems 73:8–30.

Biasutti, M., A. H. Sobel, S. J. Camargo, and T. T. Creyts. 2012. Projected changes in the physical climate of the Gulf Coast and Caribbean. Climatic Change 112:819–845.

Chapra, S. C. 1997. Surface water-quality modeling. McGraw-Hill, New York, New York, USA.

Chapra, S. C., D. M. Dolan, and A. Dove. 2016. Mass-balance modeling framework for simulating and managing long-term water quality for the lower Great Lakes. Journal of Great Lakes Research 42:1166–1173.

Del Giudice, D., P. Reichert, V. Bareš, C. Albert, and J. Rieckermann. 2015. Model bias and complexity—Understanding the effects of structural deficits and input errors on runoff predictions. Environmental Modelling & Software 64:205–214.

Del Giudice, D., R. L. Muenich, M. McCahon Kalcic, N. S. Bosch, D. Scavia, and A. M. Michalak. 2018a. On the practical usefulness of least squares for assessing uncertainty in hydrologic and water quality predictions. Environmental Modelling & Software 105:286–295.

Del Giudice, D., Y. Zhou, E. Sinha, and A. M. Michalak. 2018b. Long-term phosphorus loading and springtime temperatures explain interannual variability of hypoxia in a large temperate lake. Environmental Science & Technology 52:2046–2054.

Di Toro, D. M. 2001. Sediment flux modeling. Wiley-Interscience, New York, New York, USA.

Diaz, R. J., and R. Rosenberg. 2008. Spreading dead zones and consequences for marine ecosystems. Science 321:926–929.

Dortch, Q., N. N. Rabalais, R. E. Turner, and G. T. Rowe. 1994. Respiration rates and hypoxia on the Louisiana shelf. Estuaries 17:862–872.

Feist, T. J., J. J. Pauer, W. Melendez, J. C. Lehrter, P. A. DePetro, K. R. Rygwelski, D. S. Ko, and R. G. Kreis. 2016. Modeling the relative importance of nutrient and carbon loads, boundary fluxes, and sediment fluxes on Gulf of Mexico hypoxia. Environmental Science & Technology 50:8713–8721.

Feng, Y., S. F. DiMarco, and G. A. Jackson. 2012. Relative role of wind forcing and riverine nutrient input on the extent of hypoxia in the northern Gulf of Mexico. Geophysical Research Letters 39:L09601.

Fennel, K., and A. Laurent. 2018. N and P as ultimate and proximate limiting nutrients in the northern Gulf of Mexico: implications for hypoxia reduction strategies. Biogeosciences 15:3121–3131.

Fennel, K., J. Wilkin, J. Levin, J. Moisan, J. O'Reilly, and D. Haidvogel. 2006. Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget. Global Biogeochemical Cycles 20:GB3007.

Fennel, K., J. Hu, A. Laurent, M. Marta-Almeida, and R. Hetland. 2013. Sensitivity of hypoxia predictions for the northern Gulf of Mexico to sediment oxygen consumption and model nesting. Journal of Geophysical Research: Oceans 118:990–1002.

Fennel, K., A. Laurent, R. Hetland, D. Justić, D. S. Ko, J. Lehrter, M. Murrell, L. Wang, L. Yu, and W. Zhang. 2016. Effects of model physics on hypoxia simulations for the northern Gulf of Mexico: a model intercomparison. Journal of Geophysical Research: Oceans 121:5731–5750.

Forrest, D. R., R. D. Hetland, and S. F. DiMarco. 2011. Multivariable statistical regression models of the areal extent of hypoxia over the Texas-Louisiana continental shelf. Environmental Research Letters 6:045002.

Greenberg, A. E., L. S. Clesceri, and A. D. Eaton. 1992. Standard methods for the examination of water and wastewater. 18th edition. American Public Health Association, Washington, D.C., USA.

Greene, R. M., J. C. Lehrter, and J. D. Hagy. 2009. Multiple regression models for hindcasting and forecasting midsummer hypoxia in the Gulf of Mexico. Ecological Applications 19:1161–1175.

Gujer, W. 2008. Systems analysis for water technology. Springer, Berlin, Germany.

Haario, H., E. Saksman, and J. Tamminen. 2001. An adaptive metropolis algorithm. Bernoulli 7:223–242.

Hetland, R. D., and S. F. DiMarco. 2008. How does the character of oxygen demand control the structure of hypoxia on the Texas-Louisiana continental shelf? Journal of Marine Systems 70:49–62.

HTF. 2001. Hypoxia Task Force, action plan for reducing, mitigating, and controlling hypoxia in the Northern Gulf of Mexico. https://www.epa.gov/ms-htf/hypoxia-task-force-2001-action-plan

HTF. 2017. Hypoxia Task Force, report to Congress. https://www.epa.gov/ms-htf/hypoxia-task-force-reports-congress

Irby, I. D., M. A. M. Friedrichs, F. Da, and K. E. Hinson. 2018. The competing impacts of climate change and nutrient reductions on dissolved oxygen in Chesapeake Bay. Biogeosciences 15:2649–2668.

Justić, D., and L. Wang. 2014. 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. Continental Shelf Research 72:163–179.

Justić, D., N. N. Rabalais, and R. E. Turner. 1996. Effects of climate change on hypoxia in coastal waters: A doubled CO2 scenario for the northern Gulf of Mexico. Limnology and Oceanography 41:992–1003.

Justić, D., N. N. Rabalais, and R. E. Turner. 2002. Modeling the impacts of decadal changes in riverine nutrient fluxes on coastal eutrophication near the Mississippi River Delta. Ecological Modelling 152:33–46.

Kalnay, E., et al. 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society 77:437–472.

Katin, A., D. Del Giudice, and D. R. Obenour. 2019. Modeling biophysical controls on hypoxia in a shallow estuary using a

Bayesian mechanistic approach. Environmental Modelling & Software 120:104491.

Laurent, A., K. Fennel, D. S. Ko, and J. Lehrter. 2018. Climate change projected to exacerbate impacts of coastal eutrophication in the northern Gulf of Mexico. Journal of Geophysical Research: Oceans 123:3408–3426.

Le, C., J. C. Lehrter, C. Hu, and D. R. Obenour. 2016. Satellite-based empirical models linking river plume dynamics with hypoxic area and volume. Geophysical Research Letters 43:2693–2699.

Matli, V. R. R., S. Fang, J. Guinness, N. N. Rabalais, J. K. Craig, and D. R. Obenour. 2018. A space-time geostatistical assessment of hypoxia in the northern Gulf of Mexico. Environmental Science & Technology 52:12484–12493.

Mattern, J. P., K. Fennel, and M. Dowd. 2013. Sensitivity and uncertainty analysis of model hypoxia estimates for the Texas-Louisiana shelf. Journal of Geophysical Research: Oceans 118:1316–1332.

McElreath, R. 2018. Statistical rethinking: a Bayesian course with examples in R and Stan. Chapman & Hall/CRC Press, Boca Raton, FL.

McIsaac, G. F., M. B. David, G. Z. Gertner, and D. A. Goolsby. 2001. Eutrophication: nitrate flux in the Mississippi River. Nature 414:166–167.

Meier, H. E. M., H. C. Andersson, K. Eilola, B. G. Gustafsson, I. Kuznetsov, B. Müller-Karulis, T. Neumann, and O. P. Savchuk. 2011. Hypoxia in future climates: a model ensemble study for the Baltic Sea. Geophysical Research Letters 38: L24608.

Murrell, M. C., and J. C. Lehrter. 2011. Sediment and lower water column oxygen consumption in the seasonally hypoxic region of the Louisiana continental shelf. Estuaries and Coasts 34:912–924.

Obenour, D. R., A. M. Michalak, Y. Zhou, and D. Scavia. 2012. Quantifying the impacts of stratification and nutrient loading on hypoxia in the northern Gulf of Mexico. Environmental Science & Technology 46:5489–5496.

Obenour, D. R., D. Scavia, N. N. Rabalais, R. E. Turner, and A. M. Michalak. 2013. Retrospective analysis of midsummer hypoxic area and volume in the northern Gulf of Mexico, 1985–2011. Environmental Science & Technology 47:9808–9815.

Obenour, D. R., A. Michalak, and D. Scavia. 2015. Assessing biophysical controls on Gulf of Mexico hypoxia through probabilistic modeling. Ecological Applications 25:492–505.

Omlin, M., and P. Reichert. 1999. A comparison of techniques for the estimation of model prediction uncertainty. Ecological Modelling 115:45–59.

Pitkänen, H., J. Lehtoranta, and A. Räike. 2001. Internal nutrient fluxes counteract decreases in external load: the case of the estuarial eastern Gulf of Finland, Baltic Sea. AMBIO: A Journal of the Human Environment 30:195–201.

Quiñones-Rivera, Z. J., B. Wissel, D. Justić, and B. Fry. 2007. Partitioning oxygen sources and sinks in a stratified, eutrophic coastal ecosystem using stable oxygen isotopes. Marine Ecology Progress Series 342:69–83.

Rabalais, N. N., R. E. Turner, B. K. S. Gupta, E. Platon, and M. L. Parsons. 2007. Sediments tell the history of eutrophication and hypoxia in the northern Gulf of Mexico. Ecological Applications 17:S129–S143.

R Core Team. 2019. R version 3.5.3. R Project for Statistical Computing, Vienna, Austria. www.R-project.org.

Redalje, D. G., and G. L. Fahnenstiel. 1994. The relationship between primary production and the vertical export of particulate organic matter in a river-impacted coastal ecosystem. Estuaries 17:829–838.

Redfield, A. C., B. H. Ketchum, and F. A. Richards. 1963. The influence of organisms on the chemical composition of seawater. Pages 26–77 in M. N. Hill, editor. The sea. Interscience, New York, New York, USA.

Reichert, P., and M. E. Borsuk. 2005. Does high forecast uncertainty preclude effective decision support? Environmental Modelling & Software 20:991–1001.

Robson, B. J. 2014. When do aquatic systems models provide useful predictions, what is changing, and what is next? Environmental Modelling & Software 61:287–296.

Rowe, G. T., M. E. C. Kaegi, J. W. Morse, G. S. Boland, and E. G. E. Briones. 2002. Sediment community metabolism associated with continental shelf hypoxia, Northern Gulf of Mexico. Estuaries 25:1097–1106.

Rucinski, D. K., D. Beletsky, J. V. DePinto, D. J. Schwab, and D. Scavia. 2010. A simple 1-dimensional, climate based dissolved oxygen model for the central basin of Lake Erie. Journal of Great Lakes Research 36:465–476.

Scavia, D., and K. A. Donnelly. 2007. Reassessing hypoxia forecasts for the Gulf of Mexico. Environmental Science & Technology 41:8111–8117.

Scavia, D., N. N. Rabalais, R. E. Turner, D. Justić, and W. J. Wiseman. 2003. Predicting the response of Gulf of Mexico hypoxia to variations in Mississippi River nitrogen load. Limnology and Oceanography 48:951–956.

Scavia, D., M. A. Evans, and D. R. Obenour. 2013. A scenario and forecast model for Gulf of Mexico hypoxic area and volume. Environmental Science & Technology 47:10423–10428.

Scavia, D., I. Bertani, D. R. Obenour, R. E. Turner, D. R. Forrest, and A. Katin. 2017. Ensemble modeling informs hypoxia management in the northern Gulf of Mexico. Proceedings of the National Academy of Sciences USA 114:8823–8828.

Scavia, D., D. Justic, D. R. Obenour, K. Craig, and L. Wang. 2018. Hypoxic volume is more responsive than hypoxic area to nutrient load reductions in the northern Gulf of Mexico – and it matters to fish and fisheries. Environmental Research Letters 14:024012.

Sikorska, A. E., D. Del Giudice, K. Banasik, and J. Rieckermann. 2015. The value of streamflow data in improving TSS predictions—Bayesian multi-objective calibration. Journal of Hydrology 530:241–254.

Smith, M. D., A. Oglend, A. J. Kirkpatrick, F. Asche, L. S. Bennear, J. K. Craig, and J. M. Nance. 2017. Seafood prices reveal impacts of a major ecological disturbance. Proceedings of the National Academy of Sciences USA 114:1512–1517.

Sylvan, J. B., Q. Dortch, D. M. Nelson, A. F. Maier Brown, W. Morrison, and J. W. Ammerman. 2006. Phosphorus limits phytoplankton growth on the Louisiana shelf during the period of hypoxia formation. Environmental Science & Technology 40:7548–7553.

Thamdrup, B., J. W. Hansen, and B. B. Jørgensen. 1998. Temperature dependence of aerobic respiration in a coastal sediment. FEMS Microbiology Ecology 25:189–200.

Turner, R. E., N. N. Rabalais, and D. Justic. 2006. Predicting summer hypoxia in the northern Gulf of Mexico: riverine N, P, and Si loading. Marine Pollution Bulletin 52:139–148.

Turner, R. E., N. N. Rabalais, and D. Justic. 2008. Gulf of Mexico hypoxia: alternate states and a legacy. Environmental Science & Technology 42:2323–2327.

Turner, R. E., N. N. Rabalais, and D. Justić. 2017. Trends in summer bottom-water temperatures on the northern Gulf of Mexico continental shelf from 1985 to 2015. PLoS ONE 12: e0184350.

Wiseman, Wm. J., N. N. Rabalais, R. E. Turner, S. P. Dinnel, and A. MacNaughton. 1997. Seasonal and interannual variability within the Louisiana coastal current: stratification and hypoxia. Journal of Marine Systems 12:237–248.

Yu, L., K. Fennel, A. Laurent, M. C. Murrell, and J. C. Lehrter. 2015. Numerical analysis of the primary processes controlling oxygen dynamics on the Louisiana shelf. Biogeosciences 12:2063–2076.

SUPPORTING INFORMATION

Additional supporting information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/eap.2032/full