Time-Evolving, Spatially Explicit Forecasts of the Northern Gulf of Mexico Hypoxic Zone

Arnaud Laurent* and Katja Fennel

Department of Oceanography, Dalhousie University, Halifax B3H 4R2, Nova Scotia, Canada

ABSTRACT: The Mississippi–Atchafalaya River Basin delivers large amounts of freshwater and nutrients to the northern Gulf of Mexico promoting the development of a large hypoxic zone every summer. Statistical and semiempirical models have long been used to provide seasonal forecasts of the mid-summer hypoxic extent using historic time series of spring nutrient load and mid-summer hypoxic extent. These forecasts consist of a scalar estimate of the hypoxic area with uncertainty but do not include spatial distributions or temporal evolution of hypoxic conditions. Three-dimensional (3D) circulation-biogeochemical models of the coastal ocean simulate the temporal evolution of hypoxia in a spatially explicit manner but have not yet been used for seasonal hypoxia forecasting. Here, we present a hybrid method for seasonal, spatially explicit, time-evolving forecasts of the hypoxic zone that combines statistical forecasting with information from a 3D biogeochemical model. The hybrid method uses spring nitrate load and a multiyear (1985–2018) 3D hindcast simulation to produce a seasonal forecast. Validation shows that the method explains up to 76% of the observed year-to-year variability in the hypoxic area. The forecasts suggest that the maximum seasonal extent of hypoxia is reached, on average, on August 13, 2 weeks after the completion of the annual cruise. An analysis of month-to-month variations in hypoxia forecasts due to variability in wind speed and freshwater discharge allows estimates of weather-related uncertainties in the forecast.

INTRODUCTION

An increase of nutrient loading from rivers to the coastal ocean in the past decades1 has fueled coastal eutrophication and led to the expansion of coastal hypoxic zones, now a growing concern worldwide.2,−5 Bottom-water hypoxia (O<sub>2</sub> < 62.5 mmol m<sup>−3</sup> or 2 mg L<sup>−1</sup>) is detrimental to benthic communities6 and has negative repercussions for the coastal foodweb.7 In the northern Gulf of Mexico, one of the world’s largest seasonal hypoxic zones (15 000 km<sup>2</sup> on average in 1985–2018) has been observed every summer since the 1980s.8 Hypoxia in the region is a consequence mainly of increased nitrogen (N) loads from the Mississippi and Atchafalaya rivers due to the intensification of agricultural practices and increased use of fertilizers since the 1960s.9,−12 Hypoxic conditions significantly affect the whole foodweb from bacterial communities to fish13−15 including the northern Gulf’s fisheries.16,17

The prerequisite for occurrence of hypoxia is that an O<sub>2</sub> sink acts on a water body for long enough to deplete O<sub>2</sub> while its resupply is inhibited; in other words, the time scale to hypoxia development is shorter than the time scale of O<sub>2</sub> resupply.3 During the hypoxic season (June–Sept) in the northern Gulf of Mexico, the O<sub>2</sub> consumption rate is high (the time scale to hypoxia, thus, short), and vertical density stratification prevents reoxygenation of bottom waters from the surface long enough for hypoxia to develop. Spring nitrate load (i.e., the combination of riverine nitrate concentration and discharge) fuels respiration in the water column and sediments (the oxygen sink) and is highly correlated with the size of the mid-summer hypoxic zone.16−20 Sediment O<sub>2</sub> consumption (SOC) is the main driver of hypoxia in this region because it prevails in the bottom boundary layer (BBL) where hypoxia occurs (55−76% of total respiration in the BBL is from SOC).21,22 The distribution and strength of vertical density stratification are controlled by freshwater discharge, buoyancy- and wind-driven circulation, and vertical mixing.24,25 Hypoxic conditions are, thus, highly sensitive to the combination of strong stratification and high SOC, their spatial and temporal variability determining the location and evolution of the hypoxic zone.21,26

Since 1985, an annual mid-summer survey has been carried out to map the extent of the hypoxic zone (http://www.gulfhypoxia.net). In 2001, the Mississippi River/Gulf of Mexico Watershed Nutrient Task Force (henceforth Hypoxia Task Force) set the goal to reduce the mid-summer hypoxic
zone to a 5-year average size below 5000 km², which, according to a recent amendment to the plan, should be reached by 2035. To date, no significant reductions in nutrient load and hypoxic area have occurred. The averages of annual nitrate loads and mid-summer hypoxic areas are similar in the first and last decades of the historical record (1985–1994: 106 × 10⁶ kg N and 13.2 × 10⁴ km² and 2008–2017: 107 × 10⁶ kg N and 15.3 × 10⁴ km²) although there is large year-to-year variability.

Regression models have been developed to relate the observed mid-summer hypoxic extent to year-to-year variations in nutrient load, surface chlorophyll, and physical factors. These models have provided further information, such as hypoxic volume and O₂ consumption rate. These models, referred to hereafter as "simple biophysical models", are used to provide seasonal forecasts of the size of the hypoxic zone during the mid-summer survey (e.g.,) and to assess the effects of nutrient load changes and reductions needed to reach the long-term goal of 5000 km². The forecasts are generally good at reproducing the interannual variability in the mid-summer hypoxic zone, but their forecast skill drops for the years where episodic meteorological events (e.g., storms or persistent atypical wind directions) resulted in vertical mixing or unusual circulation patterns before or during the mid-summer hypoxia survey.

Three-dimensional (3D) circulation models with biogeochemistry (hereafter biogeochemical models) of the northern Gulf have been used extensively to underpin our mechanistic understanding of hypoxia formation, dynamics, and variability. These models have also been used for retrospective hypoxia assessments and to simulate the effects of nutrient reduction scenarios. For example, recent studies with simple biophysical and biogeochemical models agree that a 45% reduction in nutrient load is necessary to meet the long-term 5000 km² goal. But up to now, the biogeochemical models have not been used in seasonal forecasting.

Thus far, seasonal forecasts have yielded the hypoxic extent during the mid-summer survey (i.e., a scalar estimate with uncertainty) but not provided information on the location, severity, or temporal evolution of the hypoxic condition (e.g., ). Given the large spatio-temporal variability in hypoxic conditions in this region and indications that hypoxia will become more intense (i.e., lower O₂ concentrations) with global warming, an ability to provide spatially explicit, time-evolving forecasts of hypoxic extent and severity is desirable. The biogeochemical models provide highly resolved, spatially and temporally explicit information but are typically used in hindcasting mode or retrospectively because they require atmospheric forcing data. Thus, their forecasting window is limited to the weather band of 1–2 weeks.

Here, we propose a hybrid method that bridges the gap between simple biophysical and 3D biogeochemical models and provides temporally and spatially resolved seasonal forecasts of the hypoxic zone. Such resolution has not been achieved before. This work is consistent with NOAA's ecological forecasting strategy and directly addresses the expressed goal of developing dynamical, highly resolved, regional ecological forecasts on seasonal time scales. The hybrid method is similar to the simple biophysical models in that it only requires spring N load as the input but differs in using a multiyear (1985–2018) retrospective analysis from an extensively validated 3D biogeochemical model.

We first describe the methodology underlying the hybrid method and then present historical hypoxic area forecasts for validation. To assess the forecast's uncertainty related to meteorological forcing, we carry out an analysis of wind- and freshwater-driven variations around the mean forecast, which provides insights into the drivers of variability of the hypoxic zone. The implications for hypoxia mitigation are also discussed.

### METHODS

**Input Data.** Prior to producing a forecast, the hybrid method requires only one observation type as the input to produce the seasonal forecast: May nitrate load from the Mississippi–Atchafalaya River Basin. The estimate becomes available every year in early June as a part of monthly nutrient flux estimates provided by the US Geological Survey; we use the Load Estimator value (LOADEST, methods and monitoring network are described in ). For validation, bottom O₂ data were retrieved from the National Centers for Environmental Information ( ). Years 1988, 1989, and >2014 were not available.

**Coupled Hydrodynamic-Biogeochemical Model.** The biogeochemical model is a high-resolution, regional configuration of the Regional Ocean Modeling System (ROMS) that covers the northern Gulf of Mexico shelf (see Supporting Information Figure S2a). A detailed description and validation of the circulation model is available in . The biogeochemistry is based on the pelagic N-cycle model of Fennel expanded to include phosphate, dissolved oxygen, and river-derived dissolved organic matter (DOM). The model has 10 state variables: phytoplankton, chlorophyll, zooplankton, nitrate, ammonium, phosphate, oxygen, and three detrital organic matter pools (small and large detritus, and river-derived DOM) and has been extensively validated with respect to its ability to reproduce nutrient and chlorophyll distributions, process rates, and bottom-water oxygen distributions. A full description and the model equations are available in the Supporting Information of . A schematic of the model is available in Supporting Information Figure S2b.

**Sediment–water interactions** are simulated with an empirical relationship between sediment oxygen consumption (SOC) and bottom-water temperature and oxygen concentration. The original parameterization of was modified here by adding two scale factors that account for the spatial distribution and dependence on May nitrate load of...
particulate organic matter (POM) deposition to the sediment, such that

\[
SOC(x, y, t) = 6 \times 2^{T(x,y,t)/10} \left(1 - \exp\left(-\frac{O_2(x, y, t)}{30}\right)\right) \times I_{POM}(x, y) \times I_{ML}(t)
\]

(1)

where \(T\) and \(O_2\) are bottom-water temperature and oxygen, respectively, at time \(t\) and location \((x, y)\). \(I_{POM}(x,y)\) is a spatially varying factor accounting for the average spatial pattern of POM deposition (\(I_{POM} \leq 1\), see Supporting Information Figure S1). \(I_{ML}(t)\) is an annually varying factor that scales POM deposition with total nitrogen load in May. \(I_{POM}\) reaches its maximum of 1 close to the Mississippi Delta and near Terrebonne and Atchafalaya Bays (see Supporting Information Figure S1), where POM deposition is high and decreases southward and westward of these (regions, see Supporting Information Figure S1). \(I_{ML}\) is calculated as the ratio between the current May load (\(N\)) and the long-term average \((\bar{N}), 1982–2018\) such that \(I_{ML} = (N/\bar{N})\). The exponent \(\gamma = 0.6\) was determined by calibration over the period 2000–2017 (it mimics the differences between simulated and observed mid-summer hypoxic areas). For a year with above-average May load, \(I_{ML} > 1\) so that for years with similar density stratification (i.e., similar freshwater discharge and coastal circulation), \(I_{ML}\) amplifies hypoxia when nitrate load is higher than average. Further details on the determination of \(I_{POM}\) and \(I_{ML}\) are available in the Supporting Information (Text S1). Denitrification in the sediment (representing all microbial pathways of \(N_2\) gas production) and the efflux of \(NH_4^+\) to the water column \((F_{NH_4}\)) are parameterized using an empirically derived relationship between SOC and denitrification in a range of aquatic environments.\(^{59,60}\) such that \(F_{NH_4}/SOC = 0.036 \text{ mol N (mol O}_2^{-1}.^{21}\) We assume that the efflux of phosphate \((F_{PO_4}^{-})\) is proportional to the efflux of dissolved inorganic carbon (DIC), such that \(F_{PO_4}^{-} = F_{DIC}/106\). This parameterization simulates large phosphate efflux at high respiration rates but does not represent increased phosphate efflux as bottom water approaches anoxic conditions as observed by Ghasas\(^{61}\) at station C6 in the central Louisiana Shelf. There is significant uncertainty about \(P\) release from the sediment in the region. For instance, in a recent study that includes 5 stations along the Louisiana Shelf, Berelson\(^{62}\) found large \(P\) release near station C6 but no clear relationship between \(O_2\) levels and \(P\) release overall. Given this uncertainty and the fact that \(O_2\) concentrations rarely reach anoxic levels in the study region, we believe that the lack of \(O_2\) sensitivity on \(P\) release in the parameterization is not a major limitation of the model.

Atmospheric forcing is specified using 3-hourly near-surface winds from NARR\(^{63}\) and climatological surface heat and freshwater fluxes.\(^{65,65}\) Freshwater discharge from the Mississippi and Atchafalaya Rivers is prescribed using daily freshwater transport estimates by the US Army Corps of Engineers at Tarbert Landing and Simmesport, respectively (see the location in Supporting Information Figure S3). Nutrient, POM, and DOM river concentrations are calculated using monthly flux estimates from the U.S. Geological Survey.\(^{54}\) Open boundary conditions are prescribed from climatology. The model was run from January 1, 1984 to December 31, 2018, using the first year as spin up.

### Weighted Mean Forecast

The forecast consists of a times series of daily hypoxic area values, \(H_i^f\), and a corresponding hypoxic area map, \(M_f(x,y,t)\), for each day. Both are calculated as weighted means of times series and maps from the multiyear 3D biogeochemical model hindcast. Years with sustained anomalous wind patterns in July (e.g., strong or upwelling favorable wind) were excluded because they are outliers from the relationship between May nitrate load (hereafter May load) and summer hypoxic area. More specifically, 1998 and 2009 were excluded because they had strong upwelling favorable wind, as was done in Scavia\(^{37}\), 5 years (1989, 1995, 2003, 2010 and 2013) were excluded because of unusually high average wind stress in July (\(>0.034 \text{ Pa, i.e., 85th percentile}\)) and 2018 because a storm event occurred at the time of the annual mapping cruise. For the 26 remaining years, there is a strong relationship between May load and the mid-summer hypoxic area (\(r = 0.85, p < 0.001\)). Model output from these years was smoothed to remove the high-frequency variability associated with meteorological events, bias corrected with respect to observations, and interpolated over May loads to a uniform load sampling frequency (\(S \times 10^5 \text{ kg N}$^-1\) over the range \(0–21.4 \times 10^7\) kg N (see details in Supporting Information Text S2). The dataset used in the forecast is then comprised of 108 time series of the hypoxic area \(H_i(t)\) and 108 \times 184 daily binary maps \(M_f(x,y, t)\) of hypoxia occurrence \(M_f(x,y,t)\), where \(i = 1,..,108\).

The assumption underlying this method is that the hindcast simulation provides an ensemble of spatio-temporally explicit hypoxia seasons over a range of May loads that are representative of the conditions in the northern Gulf of Mexico. The weight attributed to each ensemble member is based on the similarity of its May load to the May load for which the forecast is to be made.

The weighted mean forecast is calculated as

\[
\chi^f = \sum_{i=1}^{n} w_i X_i / \sum_{i=1}^{n} w_i
\]

(2)

where \(X\) refers to the time series \(H\) or map \(M\) and subscript \(i\) refers to the ensemble member. The normally distributed weights \(w_i\) are assigned to individual ensemble members \(i\) based on the difference between their respective May load \(N_i\) and the May load for which the forecast is to be made, \(N_f\) (kg N), as

\[
w_i = \frac{1}{\alpha \sqrt{2\pi}} e^{-[(N_i - N_f)^2]/2\alpha^2}
\]

(3)

The spread \(\alpha\) is set to \(1.8 \times 10^7\) kg N aiming for a narrow enough distribution to avoid a sampling bias at high load but ensuring enough ensemble members for a robust forecast. Increasing \(\alpha\) above this value did not improve the forecast skill significantly (maximum \(r^2 = 0.77\) versus 0.76), whereas the sampling bias increased. An example of weight distribution is provided in Figure S4. The 95% confidence interval around \(H^f\), which provides information about the uncertainty around the weighted mean, is calculated from the weighted standard error\(^{66}\) such that

\[
\text{CI}_{n,a=1}^{f} = H^f \pm t_{n,a=1} \sqrt{\frac{n}{(n-1)w_k^2} \sum_{i=1}^{n} w_i^2 (H_i - H^f)^2}
\]

(4)

where \(t_{n,a=1}\) is the \(t\)-score and \(\alpha = 0.05\) and \(w_k = \sum_{i=1}^{n} w_i\).
Since the maps are binary with values of only zero and one, the forecast map $M_f$ that results from the weighted mean of the 108 $M_i$ maps is a probability distribution of hypoxia occurrence ($0 \leq M_f(x,y) \leq 1$). At any given time, the hypoxic zone is assumed to occur at the locations with highest probability such that its area is equal to $H_f(t)$.

The script and input files required to run the forecast are provided in the Supporting Information and available as a Matlab toolbox on the MathWorks File Exchange.

## RESULTS AND DISCUSSION

### Mid-Summer Forecast

The hybrid method was used to issue forecasts in 2018\(^{68,69}\) and 2019\(^{36,70}\) (on June 7, 2018 and June 10, 2019). To assess the forecast skill of the hybrid method on a larger dataset, it was also applied to the historical May loads from 1985 onward, but each time excluding the year for which the prediction was generated from the weighted mean calculation.

The 2018 and 2019 forecasts for the time of the hypoxia cruise were 15.9 and $21.6 \times 10^3$ km$^2$, respectively, similar to the ensemble of forecasts from the simple biophysical models of 15.0 and 20.3 for 2018 and 2019, respectively. The observed hypoxic areas in both years were 7.0 and $18.0 \times 10^3$ km$^2$, respectively. In 2018, the disagreement between forecast and observed hypoxic area is due to a storm event that occurred during the cruise and temporarily ventilated bottom waters. The 2019 forecast was slightly larger than the observed hypoxic area, partly due to Tropical Storm Barry that occurred a week before the mid-summer cruise\(^{71}\).

The retrospective forecasts from 1985 to 2017 agree well with the observation-based hypoxia estimate from the annual hypoxia mapping cruises (Figure 1), indicating that the hybrid method explains 76% of the variability for the 26 years retained in the analysis. The forecast skill degrades in the low and high hypoxic area regions ($H_f < 8$ and $H_f > 21$) where input information is limited (Figure 1b). The forecast skill is in the middle of the range of simple biophysical models described in Scavia\(^{37}\), which explain between 64 and 92% of the variability in mid-summer hypoxic extent. Inclusion of the 7 years with anomalous wind patterns lowers the predictive skill to 49% of...
the variability, demonstrating that the forecast works best when typical weather conditions prevail in summer. The forecast method is not capable of accounting for high-frequency variability associated with meteorological events, similar to the other simple biophysical models. 72

The hybrid method relies on the observed positive relationship between May load and the mid-summer hypoxic area (Figure 1a). This relationship is relatively linear for May loads under $16 \times 10^7$ kg N but levels off at high loads, similar to the curvilinear relationship used by Turner to predict mid-summer hypoxic extent. Since the hybrid method is independent of time, the same sensitivity of hypoxia to nitrate load (i.e., hypoxic area per unit load) is assumed since 1985. The mid-summer observations do not show a change in sensitivity over time for the 26 years analyzed ($r = 0.35, p = 0.09$), which is consistent with Obenour. 35,73 The small increase in hypoxic area per unit load over time is driven by the year 1988 that had unusually low loads and a very small hypoxic area early on in the time record ($r = 0.25, p = 0.24$ when excluding 1988).

Overall, the hybrid method yields mid-summer hypoxic area forecasts with similar skills as the simple biophysical models.

**Temporally and Spatially Resolved Forecast.** For a given May load ($N^M$), the hybrid method provides continuous predictions of the size ($H^f$) and location ($M^f$) of the hypoxic zone (Figures 1a, 2 and Supporting Information Movie S1), significantly expanding the forecast capability of the current simple biophysical models. For instance, the hybrid method predicts the year-integrated hypoxic area ($H^f; \text{km}^2 \text{yr}$) as a function of May load, which is a potentially useful integrated measure of hypoxia exposure or severity. $H^f$ increases linearly with loads, similar to the mid-summer forecast relationship, for loads less than $13 \times 10^7$ kg N but levels off for N loads between 13 and $17 \times 10^7$ kg N. Unlike the mid-summer forecast, $H^f$ keeps increasing at larger loads (Figure 1a) but since only one year with N load larger than $20 \times 10^7$ kg N is available in the dataset, the increase in $H^f$ for large loads is uncertain.

The spatial skill of the mid-summer forecast is measured for all years with available bottom O$_2$ data, i.e., all analyzed years except 1988, 2015, and 2017, using a point-to-point comparison between hypoxic conditions in the observations and their corresponding location on the forecast map. On average, the spatial skill during the mid-summer cruises is 64% ($\pm 7.3\%$ SD, Figure S6). This is similar to the spatial skill of biogeochemical models. 21 As illustration, Figure 2 compares the historical forecast with the mid-summer observations from the mapping cruises for years with contrasting nutrient loads, specifically 8.4, 13.0, and $19.1 \times 10^7$ kg N (2005, 1994, and 2017, respectively). The variation in size and location of the hypoxic zone in the forecast is consistent with the mid-summer observations. The evolution of the hypoxic area has been described as initially developing in June, peaking in July–August, and receding in the early fall. 74–76 At small May loads, the hypoxic zone develops later and terminates earlier than at larger loads, and the hypoxic season is shorter by about a month. At intermediate and large May loads, there is no difference in timing for the onset and termination of hypoxia but the size of the hypoxic area varies significantly in July–August (Figure 2). In terms of location, $M^f$ is most persistent along the coast on the eastern shelf and within the 10–30 m depth range in the mid-shelf (south of Atchafalaya Bay) and the western shelf (west of Atchafalaya Bay, Figures 2, 3 and Supporting Information Figure S7), consistent with observations. 74

With increasing May load, hypoxic waters tend to cover a wider bathymetric range and extend further west, although there is considerable variability in the location of the hypoxic zone (Figure 2). This variability is inherent to the system and, thus, to the biogeochemical model results. 16 May load is correlated with discharge ($r = 0.87, p < 0.001$). Therefore, the stratification envelope that prevents the ventilation of bottom waters expands westward with increasing loads. 25,42 The largest discrepancies between $M^f$ and observations occur on the western shelf (Figure 2), farthest downstream from the Mississippi freshwater source, where SOC weakens and physical factors have more control on hypoxia formation. 35,73,77 In this region, stratification weakens and is more spatially variable due to the influence of wind-driven currents on the location of the Atchafalaya River plume. 24,42 Sediment O$_2$ demand also weakens due to lower organic matter deposition (see Supporting Information Figure S1). 78

The shelf surveys are typically carried out at the end of July and may not coincide with the maximum hypoxic extent. The size of the hypoxic zone can vary significantly in July–August. Matli 16 recently showed, based on 1985–2016 observations, that the size of the mid-summer hypoxic area is not always representative of the maximum annual extent of hypoxia. In our forecast, the maximum extent of the hypoxic zone occurs
from early to mid-August, i.e., after the shelf-wide cruise. To assess if the forecast is realistic at the time of maximum hypoxia, we compare historical forecasts with observations for the 4 years when extensive sampling occurred in August, i.e., 2004, 2005, 2007, and 2012 (Figure 3), with May loads of 9.41, 8.41, 14.80, and 6.18 × 10^7 kg N, respectively (Table S2). Overall, the size and location of the hypoxic zone forecasts are in good agreement with the observations available (Figure 3). Point-by-point comparisons between the observations and the historical forecasts indicate skills of 82% in 2005, 71% in 2007, and 82% in 2012. The skill drops to 65% in 2004 when hypoxia is underestimated in the Mississippi Bight. Despite the good skill in August 2012, hypoxia is overestimated on the western shelf (Figure 3). May load in 2012 is the second lowest of the time record, and mid-summer hypoxia was underestimated for that year (Table S2).

Few values are available to constrain the forecast at very low N loads and, thus, the uncertainty is larger. In 2007, when May load is above average, hypoxic water covers a wider area on the eastern shelf and is located farther offshore, consistent with observations. The hypoxic area is similar in August 2004 and 2005, years with similar May load levels. The westward extent of the hypoxic zone is well represented in the 2005 historical forecast. The survey did not cover the westernmost part of the shelf in 2004, thus the occurrence of hypoxia in this region cannot be validated.

Given the intrinsic hydrodynamic variability of the northern Gulf, the skill of the hybrid method is promising for seasonal hypoxia predictions throughout the summer. The forecast should be suitable for informing managers of the risk of hypoxia and providing time-resolved estimates of its spatial distribution. High-frequency variations cannot be forecast with the hybrid method and likely can only be assessed in real-time or retrospectively. However, month-to-month variability associated with physical processes can be assessed independently, as will be discussed next.

Seasonal Variability Associated with Wind and Discharge. Physical factors such as wind and freshwater input are important contributors to density stratification on the shelf and, thus, partly control the variability in the extent of the hypoxic zone. The hybrid method relies only on May loads and does not account for high-frequency variability in wind speed and freshwater discharge. The same limitation applies to simple biophysical models and affects their forecast skill under unusual meteorological conditions.

Retrospective biogeochemical model simulation with high-resolution physical forcing provides a full dynamical picture of seasonal hypoxia evolution after the fact and can complement the forecasts. Figure 4 illustrates the difference between the forecast and the retrospective analysis for 2003 and 2018, two of the years that were excluded from the weighted mean because of anomalous conditions (with May loads of 10.6 and 11.5 × 10^7 kg N, respectively). It is obvious that although the retrospective analysis follows the general shape of the forecast, significant discrepancies with the forecast exist (Figure 4e,f). These discrepancies can be explained by the magnitude and changes in water column stratification, measured by the potential energy anomaly (PEA, Figure 4a–d). On average, freshwater discharge was 29% lower (March–October, Figure S8) and wind stress was 29% higher (Figure 4a,b) in 2003 compared to 2018, which explains the lower hypoxic area and the overestimation of the hypoxic area by the 2003 forecast (Figure 4e,f). A series of strong wind-induced mixing events occurred in July and August 2003 (Figure 4a,c) and resulted in a temporary oxygenation of bottom waters. Similar events but of different intensities occurred in 2018. In both years, mixing events preceded the annual mapping cruise and the observed extent of the hypoxic zone was much smaller than expected based on the May load only.

Variations in the discharge also explain some of the differences between the forecast and retrospective analysis in 2018. March–April discharge was 46 200 m^3 s^{-1} that year.
Figure S8), compared to the long-term mean of \(31,000 \text{ m}^3 \text{s}^{-1}\) (1985–2018). The large discharge resulted in strong stratification and the development of an unusually large hypoxic area in late spring during a period of weak winds (Figure 4b,d,f). Forrest\(^3\) and Greene\(^2\) also found the improved skill in their regression models of mid-summer hypoxic area when including wind and discharge as predictors, respectively. The two variables are associated with a high level of uncertainty in the size of the hypoxic area.\(^4\)

Wind and discharge have opposite effects on stratification and, thus, can compensate for each other. The magnitudes of these effects can be estimated from the forcing and physical outputs of the original biogeochemical model simulation (1985–2018) using a two-step method. First, stratification strength can be characterized by the potential energy anomaly (PEA, \(\text{J m}^{-3}\)), which represents the work per unit volume required to homogenize the water column through mixing,\(^7\) and can be calculated from simulated temperature and salinity distributions. We calculated a multiple linear regression between the monthly (July–September), shelf-averaged PEA from the biogeochemical model dataset, monthly averaged wind speed (\(\text{WS, m s}^{-1}\)), and the previous month’s freshwater discharge (\(\text{FW, m}^3 \text{s}^{-1}\)) from the forcing dataset (\(\text{PEA} = 132.53 - 20.60 \times \text{WS} + 2.73 \times \text{FW}\)). This relationship explains 79% of the month-to-month variability in PEA (Figure 5a,b). The previous month’s discharge (i.e., June–August) was used to account for the travel time of freshwater over the shelf. Then, we calculated a quadratic relationship between the monthly averaged hypoxic area (\(H_M \times 10^3 \text{ km}^2\)) and PEA (\(\text{PEA m}^{-3}\)) (\(H_M = 4.1330 + 0.2510 \times \text{PEA} - 0.0007 \times \text{PEA}^2\)). The relationship explains 42% of the month-to-month variability in the hypoxic area (Figure 5c) and can be used to retrospectively assess the effect of a deviation from typical (median, 1985–2018) stratification conditions (\(\text{PEA} = 104.53 \text{ J m}^{-3}\)) on the hypoxic area (\(\bar{H}_M = 13.86 \times 10^3 \text{ km}^2\), Figure 5d). Increased stratification somewhat enhances the size of the hypoxic zone (up to 32% increase), as seen when comparing 2013 (Figure 4c,e) and 2018 (Figure 4d,f), but the effect of destratification on hypoxia is much larger (up to 87% decrease in the hypoxic area) and nearly linear (Figure 5d). By combining the two steps above, we can assess the combined effect of a deviation from typical (median, 1985–2018) wind (\(\text{WS} = 3.95 \text{ m s}^{-1}\)) and/or discharge (\(\text{FW} = 18.36 \text{ m}^3 \text{s}^{-1}\)) on the hypoxic area (Figure 6). This relationship is informative and complementary to the forecast in providing an estimate of the variability around the forecast that can be expected for specific forcing conditions. Within the observed range, month-to-month variations in freshwater discharge or wind speed result in about ±40% and −85 to +15% change in the hypoxic area, respectively. For a ±25% variation from FW or WS, the relationship indicates a variation in the hypoxic area of −11/+9 and −17/+14%, respectively, whereas the dual (opposite) effect is −30/+22%. These values are smaller than the uncertainties estimated by Mattern.\(^4\)

Our analysis only estimates month-to-month variations. Larger variability is expected at shorter time scales, as shown in Figure 4. Wind direction can also have a significant influence...
on the size and location of the hypoxic zone (e.g., 39,42), but this effect is not included in our analysis.

**Implications for Hypoxia Mitigation.** The Hypoxia Task Force has set the goal of reducing the five-year running average areal extent of the hypoxic zone to less than 5000 km2 by the year 2035, with an intermediate target of a 20% reduction in N and P load by 2025 relative to the 1980–1996 average.17 The forecast model can be used to estimate the effect of nutrient reductions on the hypoxic area. Since P loads are highly correlated with N loads in May (1982−2018, r = 0.85, p < 0.001), nutrient reductions here should be interpreted as dual reductions. The hybrid method predicts that the hypoxic area reduction goal of 5000 km2 will be reached for a May load of 6.3 × 10^7 kg N (Figure 1a and Supporting Information Figure S9). This corresponds to a 52% (45−57) reduction from the 1980−1996 average of 13.0 × 10^7 kg N load. This May load reduction level is similar to other recent estimates, which indicate that a >45% reduction will be necessary to meet the long-term goal.37,41,63,45

The intermediate 20% N and P load reduction relative to the 1980−1996 average would result in a 12.9 × 10^3 km^2 (11.4−14.2) mid-summer hypoxic area, corresponding to a 21% reduction in size. This mid-summer estimate is within the model ensemble mean of Scavia37 ranging from 11.1 to 16.4 × 10^3 km^2. The ratio of the hypoxic area to May load reduction is 1.06 for the 20% intermediate load reduction and 1.34 for the 5000 km2 goal, following the slope of the mid-summer relationship (Figure 1a). This indicates that the efficiency of nutrient load reduction measures will increase toward the long-term goal, which is consistent with the results from biogeochemical model scenario simulations.43

Despite the low efficiency for an initial 20% load reduction, this intermediate load reduction is an important step in hypoxia mitigation.

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.9b05790.

Parameterization of sediment—water fluxes80 (Text S1, Table S1); preprocessing steps (Text S2); Supporting Table S2, Figures S1−S9 (PDF)

Hypoxic area time series (left panel) and mid-summer hypoxic zone (right panels) for 3 years with contrasting May loads (MP4)

LaurentFennel2019_EST_toolbox (ZIP)

**AUTHOR INFORMATION**

**ORCID**

Arnaud Laurent: 0000-0002-8545-9309

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This paper is the result of research funded by the National Oceanic and Atmospheric Administration’s National Centers for Coastal Ocean Science Competitive Research Program under award NA16OAR4320199 to Dalhousie University. This is NGOMEX Contribution 244.

**REFERENCES**


Environmental Science & Technology


(70) Fennel, K.; Laurent, A. Low-oxygen “dead zone” in the northern Gulf of Mexico predicted to grow to 26,000 km² this summer http://memg.ocean.dal.ca/news/2019/forecast.html.


(80) Rowe, G. T.; Cruz Kaei, M. E.; Morse, J. W.; Boland, G. S.; Escobar Briones, E. G. Sediment Community Metabolism Associated with Continental Shelf Hypoxia, Northern Gulf of Mexico. Estuaries 2002, 25, 1107–1106.