Fusion-Based Hypoxia Estimates: Combining Geostatistical and Mechanistic Models of Dissolved Oxygen Variability

Venkata Rohith Reddy Matli, Arnaud Laurent, Katja Fennel, Kevin Craig, Jacob Krause, and Daniel R. Obenour

SEDAR74-RD86

September 2021

Environmental Science & Technology

pubs.acs.org/est

Article

Fusion-Based Hypoxia Estimates: Combining Geostatistical and Mechanistic Models of Dissolved Oxygen Variability

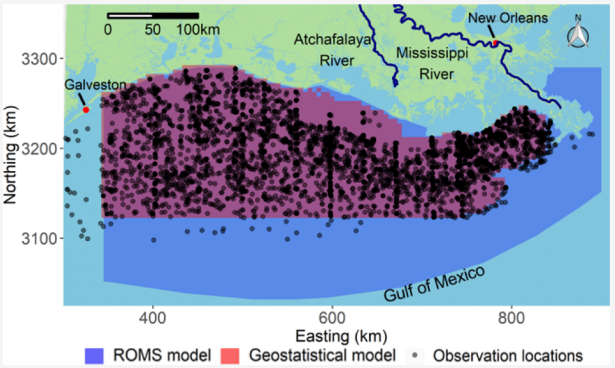
Venkata Rohith Reddy Matli,* Arnaud Laurent, Katja Fennel, Kevin Craig, Jacob Krause, and Daniel R. Obenour

Cite This: https://dx.doi.org/10.1021/acs.est.0c03655

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The need to characterize and track coastal hypoxia has led to the development of geostatistical models based on *in situ* observations of dissolved oxygen (DO) and mechanistic models based on a representation of biophysical processes. To integrate the benefits of these two distinct modeling approaches, we develop a space−time geostatistical framework for synthesizing DO observations with hydrodynamic−biogeochemical model simulations and meteorological time series (as covariates). This fusion-based approach is used to estimate hypoxia in the northern Gulf of Mexico across summers from 1985 to 2017. Deterministic trends with dynamic covariates explain over 35% of the variability in DO. Moreover, cross-validation results indicate that 58% of DO variability is explained when combining these trends with spatiotemporal interpolation, which is substantially better than mechanistic or conventional geostatistical hypoxia modeling alone. The fusion-based approach also reduces hypoxic area uncertainties by 11% on average and up to 40% in months with sparse sampling. Moreover, our new estimates of mean summer hypoxic area changed by >10% in a majority of years, relative to previous geostatistical estimates. These fusion-based estimates can be a valuable resource when assessing the influence of hypoxia on the coastal ecosystem.



1. INTRODUCTION

Hypoxia is a phenomenon in which dissolved oxygen (DO) concentrations are insufficient to support most marine life, and processes underlying the ecosystem function are altered or impaired.[1] Hypoxia is usually defined as DO below 2 mg/L, which is an important threshold for many marine species.[2] The main causes of anthropogenic hypoxia are excess nutrient loadings leading to increased organic matter production and decomposition (eutrophication), combined with water column stratification that inhibits reoxygenation of the bottom water column.[3,4] Such hypoxic regions are often referred to as "dead zones", as fish and shellfish are forced to move from these regions to more oxygen-rich waters, or else perish. In 2008, over 245,000 km^2 of the world's coastal waters were reported to experience hypoxia at varying levels of severity.[5] This number has further increased over the past decade, as eutrophication and algal blooms continue to increase.[6−8] Some of the largest hypoxic zones in the world are found in the Black Sea, Baltic Sea, East China Sea, and northern Gulf of Mexico.[5]

The magnitude of hypoxic zones and their effects on aquatic life have led to the inception of several governmental and academic programs that study the causes, extent, and consequences of hypoxia.[9−11] The hypoxic zone in the northern Gulf of Mexico is of particular interest because it is the second largest anthropogenic dead zone in the world and because of potential consequences for the region's fisheries.[10] Previous studies have explored correlations between annual estimates of the hypoxic area and fish and shrimp catch data.[12−14] Research in recent years, however, emphasizes the need for higher spatiotemporal resolution DO estimates in studying the growth, distribution, and behavior of fish and fisheries.[15−17] Also, an action plan was formulated in 2001 to implement watershed nutrient controls to reduce the hypoxic extent to 5000 km^2 (as a 5 year moving average) by 2015.[18] This plan has been revised over the years, and the current goal is to achieve the reduction by 2035.[10] Reliable estimates of hypoxia are crucial for evaluating progress toward this goal.

Hypoxia monitoring efforts often provide limited spatiotemporal coverage because of the time and resources required to conduct extensive cruises.[19−21] Sparse monitoring data can

Received: June 5, 2020
Revised: August 30, 2020
Accepted: September 3, 2020
Published: September 3, 2020

lead to uncertain estimates of hypoxic severity, as DO levels can vary dynamically across space and time. For example, high winds associated with low-pressure systems and tropical storms disrupt water-column stratification, promoting reoxygenation of the bottom waters and temporarily dissipating hypoxia.[22−24] Prior studies have shown that it takes around 2 weeks for hypoxia to re-form on the Louisiana shelf after large storm events.[25−27] In addition to meteorological changes, DO in coastal systems responds to physical and biogeochemical processes that are influenced by coastal circulation patterns.[28−30] To address these challenges, a number of statistical and process-based models,[27,31−33] each with distinct advantages and limitations,[34,35] have been developed to assess and predict hypoxia.[36]

Geostatistical models estimate spatial patterns in environmental variables, such as DO, using observations from monitoring programs and covariates, such as water depth.[37−39] The geostatistical framework allows for rigorous uncertainty quantification through preserving the covariance structure of the observations, and aggregated quantities (i.e., total hypoxic area) can be probabilistically estimated using spatial Monte Carlo simulations, often referred to as "conditional realizations".[40] While most geostatistical models focus on spatial variation in data,[37,38] space−time models rigorously account for temporal variability as well.[41,42] A recently developed space−time geostatistical model for the northern Gulf assimilates bottom water DO (BWDO) data from multiple research programs and allows for probabilistic estimation of hypoxia across each summer (May−September) from 1985 to 2016.[39] However, the model does not account for biophysical drivers of hypoxic variability (e.g., wind, nutrient load, and so forth), and for times and locations where observational data are sparse, the uncertainty in estimated values is high.

Mechanistic hypoxia models consider riverine and meteorological inputs using a series of process-based equations to simulate physical, chemical, and biological properties of aquatic systems.[43−45] For these models, development of DO predictions is less dependent on the availability of proximate observational data.[46] Also, these models can be used to evaluate the consequences of management strategies on hypoxic severity.[45,47] However, estimates obtained from mechanistic models are not always consistent with the available observational data.[44] Additionally, there is uncertainty associated with model structure, parameter estimates, and boundary and initial conditions,[27,48] which can be difficult to characterize, particularly for complex mechanistic models. Multiple hydrodynamic−biogeochemical models have been developed to study coastal hypoxia, including the regional ocean modeling system (ROMS).[43,44,49]

Integrating outputs from mechanistic models within a geostatistical framework can help account for variation in water quality resulting from biophysical processes. While such approaches are not common, Murphy et al. (2010) leveraged a mechanistic water quality model output in a geostatistical model for DO. In their study, the integrated approach was shown to substantially improve the predictive performance, though estimates were limited to spatial interpolations at the times of monitoring cruises.[38] Integrated models have also been implemented in the fields of geology and air quality. For example, Rühaak et al. (2014) developed a universal kriging model to estimate the subsurface temperature using limited borehole observations and output from a numerical temperature model.[50] Goovaerts et al. (2008) used output from a process-based atmospheric deposition model as a covariate in a geostatistical model to estimate the dioxin levels.[51] However, none of these studies modeled the temporal correlation structure to enable continuous estimation through time.

The primary objective of this study is to assess the efficacy of a spatial−temporal geostatistical framework that integrates in situ observations of BWDO with temporally dynamic covariates that are related to hypoxia formation. In addition to leveraging output from a mechanistic model, we also explore the utility of wind speed, precipitation, sea surface temperature, and solar irradiance data, which are related to biophysical processes controlling hypoxia (e.g., photosynthesis, vertical mixing, and so on).[44,52] These covariates are evaluated through statistical criterion-based variable selection. Improvements in predictive performance are described relative to a baseline geostatistical model without dynamic inputs.[39] The robustness of the improvements are assessed through cross-validation (CV), and the enhanced model is then used to estimate the hypoxic area (with quantified uncertainty) across each summer from 1985 to 2017 through conditional realization. Finally, DO and hypoxic area estimates from this new fusion-based approach are compared with previous estimates to help improve our understanding of historic hypoxic variability in the northern Gulf of Mexico.[39]

2. METHODS

2.1. Data. The DO sampling data used here were described by Matli et al. (2018).[39] Briefly, this dataset included 7000 observations collected from 150 monitoring cruises, as archived in National Oceanic and Atmospheric Administration (NOAA) National Center for Environmental Information (NCEI).[39] In this study, we also included observations from a 2017 monitoring cruise conducted by the Southeast Area Monitoring and Assessment Program (SEAMAP)[53] and 1992−1994 cruises conducted by Texas A&M University[54] (also available through NCEI). Observations collected from May to September of 1985−2017 were used in model development, as hypoxia occurs predominantly in summer. Distances were calculated using Universal Transverse Mercator Zone 15N, and bathymetry was extracted from a coastal relief model.[55] Based on the estimated maximum hypoxic zone extent determined by previous studies, the study area was limited to 94.605−89.512° W, 28.219−29.717° N, and 3−100 m depths.[37,39] This area extends from the outlet of Mississippi River to the Bolivar Peninsula, TX. Within this area, model estimates were resolved across a 5 km square estimation grid.

Meteorological data used in this study were obtained from NASA's Prediction Of Worldwide Energy Resources (POWER) project. These data include the surface water temperature, wind speed, precipitation, and solar irradiance from atmospheric models driven by satellite observations.[56] These coarse resolution (0.5° by 0.5°) data from POWER were resolved to the relatively fine-scale estimation grid of this study using inverse distance weighing (IDW). We note that POWER data correlated well with the wind data collected by NOAA buoys in the region (Supporting Information, SI-1).

Temperature, salinity, and DO predictions from ROMS, a hydrodynamic−biogeochemical model described in Laurent et al. (2017), were used in the current study (Supporting Information SI-2).[57] The ROMS grid varies from a fine-scale resolution (∼1 km) near the Mississippi and Atchafalaya River deltas to coarse resolution (∼20 km) in far-field deep ocean

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

regions. Water density was determined using temperature and salinity output.[58] The strength of density stratification was computed as the ratio of the density difference (surface to bottom) to the water-column depth. This measure of overall water-column stratification is comparable to the potential energy anomaly, which has been used to study Gulf hypoxia dynamics.[59,60] The irregular resolution ROMS output was mapped to the geostatistical estimation grid using IDW.

2.2. Geostatistical Modeling Framework. The space−time geostatistical model resolves the response variable (BWDO) as the combination of deterministic and stochastic components (eq 1).[39]

$$\text{BWDO} = x\beta + \eta + \varepsilon \tag{1}$$

The deterministic component ($x\beta$) is similar to multiple linear regression (MLR) with covariates, x, and trend coefficients, β. Matli et al. (2018)[39] found linear trends with respect to easting (E) and northing (N) and quadratic trends with respect to depth (D) and day-of-year (T) to be significant in explaining the BWDO variability. In addition to continuous trends, a year-specific (categorical) variable (i.e., "annual intercepts") was included to allow for shifts in mean annual BWDO. In this study, we considered additional linear and quadratic trends with respect to the wind speed, temperature, precipitation, solar irradiance, and ROMS DO′ and strength of stratification (S_{strat}). Meteorological variables were considered over multiple candidate averaging periods of 1, 2, 3, 7, 14, 21, and 28 days preceding the day of estimation. For these periods, the data were triangularly weighted (i.e., linearly declining weights from the prediction date backward in time), consistent with previous Gulf hypoxia modeling studies.[27,45]

The stochastic component ($\eta + \varepsilon$) of the geostatistical model resolves the remaining variability (not explained by the deterministic component) as the sum of correlated (η) and uncorrelated (ε) stochasticity (i.e., variance). Uncorrelated stochasticity is also known as "nugget", which accounts for random errors from sampling procedures and environmental microvariability.[61,62] The stochastic covariance was modeled as a function of spatial and temporal lags using a nonseparable exponential space−time covariance function (eq 2).[39]

$$Q(s_{i,j}, t_{i,j}) = \begin{cases} \sigma_\eta^2 + \sigma_\varepsilon^2, & s_{i,j} = t_{i,j} = 0 \\ \sigma_\eta^2 \times \exp\left(-\sqrt{\frac{s_{i,j}^2}{a^2} + \frac{t_{i,j}^2}{b^2}}\right), & s_{i,j} \text{ or } t_{i,j} > 0 \end{cases} \tag{2}$$

where Q is the covariance between two observations separated by a spatial lag $s_{i,j}$ and a temporal lag $t_{i,j}$. Parameter σ_η^2 is the partial sill (i.e., correlated stochasticity) and σ_ε^2 is the nugget variance. The spatial lag $s_{i,j}$ was determined using an anisotropy parameter, α, that accounts for greater correlation in the east−west dimension relative to the north−south dimension (Supporting Information SI-3).[63,64] Finally, a (km) and b (days) are scaling parameters that correspond to approximately one-third of the spatial and temporal correlation range, respectively. All covariance parameters were estimated using restricted maximum likelihood.[65]

2.3. Variable Selection. The current study considered a large pool of ROMS and meteorological trend variables (Section 2.1). To avoid overparameterizing the model, the Bayesian information criterion (BIC) was used to perform an exhaustive search for all possible combinations of the new variables. The set of variables with a minimum BIC, indicating an optimal balance between the model fit and parsimony, is preferred.[66] In this study, BIC was first applied ignoring spatiotemporal correlation for computational efficiency, as in MLR.[37,67] Residuals from the BIC-selected MLR were used to develop a preliminary estimate of the covariance function parameters. Then, variable selection was repeated using geostatistical BIC, considering spatiotemporal correlation.[68] Finally, residuals from the selected geostatistical model were used to develop a refined estimate of covariance parameters, which are then used in subsequent modeling.

The variable selection process was repeated multiple times, with different subsets of the candidate trend variables (covariates). Model version V1 explored the ability of meteorological variables from POWER to explain variations in DO. V2 evaluated the potential of ROMS output to explain variations in DO. Finally, V3 considered POWER and ROMS variables simultaneously. All model versions included the spatial and temporal trends selected in Matli et al. (2018) as a baseline, henceforth referred to as M18.[39]

2.4. Model Validation. Model performance and robustness were assessed through a k-fold CV[69] where the dataset was divided into k groups, also known as folds. Each fold comprised 100% (or 50%) of the observations from a particular monitoring cruise. Each fold was removed from the original dataset, in turn, and the remaining data were used to estimate model parameters and predict the response variable at the locations associated with the removed fold. The aggregated predictions were used to evaluate the goodness of fit (e.g., coefficient of determination, R^2, and root mean square error, RMSE).

We performed both "leave-one-cruise-out" (k = 75) and "leave-half-cruise-out" (k = 170) CVs to assess the benefits of having concurrent data available (as in the case of leave-half-cruise-out). For the leave-one-cruise-out validation, we did not remove data for years with only one (or zero) shelfwide cruise, as at least one major cruise is needed to robustly estimate the annual intercepts (Section 2.2). Shelfwide cruises were defined as cruises that collect samples across the study area, as opposed to specific transects or near-shore regions. For leave-half-cruise-out validation, we randomly removed half of the observations from cruises with at least two observations. We also assessed the advantage of accounting for both deterministic trends and spatiotemporal correlation ($x\beta + \eta$) relative to using deterministic trends only ($x\beta$). These metrics provided supplemental support in determining the best model version.

2.5. Conditional Realizations. Conditional realizations were used to quantify the uncertainty in hypoxic area estimates from the geostatistical model.[37,39,40] First, BIC-selected variables and covariance function parameters were used to determine a unique set of kriging weights (corresponding to the observations) for each estimation location and each day. Also, unconditional realizations (i.e., statistical simulations) of the response variable were determined using the procedure described in M18. These unconditional realizations were then conditioned to the observed data and trends using the kriging weights, thus producing conditional realizations.[40] Simulated grid locations with BWDO less than 2 mg/L were summed to develop realizations of the total hypoxic extent. This procedure was repeated 1000 times at a 3 day interval from May to September (1985 to 2017) to calculate the mean and

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Geostatistical Trend Coefficients (β) and Associated Standard Errors (σ_β) for BIC-Selected Variables, and Variance Explained by the Deterministic Component (R^2)[a]

		M18		V1		V2		V3	
variable	units	β	σ_β	β	σ_β	β	σ_β	β	σ_β
E	km	-4.4×10^{-3}	4.4×10^{-4}	-4.1×10^{-3}	4.2×10^{-4}	-3.6×10^{-3}	4.2×10^{-4}	-3.3×10^{-3}	4.1×10^{-4}
N	km	-8.2×10^{-3}	1.7×10^{-3}	-7.5×10^{-3}	1.7×10^{-3}	-7.9×10^{-3}	1.7×10^{-3}	-7.3×10^{-3}	1.6×10^{-3}
D	m	-2.3×10^{-1}	1.1×10^{-2}	-2.4×10^{-1}	1.1×10^{-2}	-2.3×10^{-1}	1.1×10^{-2}	-2.3×10^{-1}	1.1×10^{-2}
D^2	m^2	5.1×10^{-3}	2.2×10^{-4}	5.1×10^{-3}	2.1×10^{-4}	4.7×10^{-3}	2.2×10^{-4}	4.8×10^{-3}	2.2×10^{-4}
T	day	4.6×10^{-3}	3.2×10^{-3}	4.6×10^{-3}	3.0×10^{-3}	3.1×10^{-3}	2.9×10^{-3}	3.1×10^{-3}	2.8×10^{-3}
T^2	day^2	2.8×10^{-4}	4.9×10^{-5}	1.3×10^{-4}	5.2×10^{-5}	2.1×10^{-4}	4.6×10^{-5}	9.9×10^{-5}	5.0×10^{-5}
WS_{28}^2	$(m/s)^2$			8.8×10^{-2}	1.5×10^{-2}			7.1×10^{-2}	1.4×10^{-2}
DO′	mg/L					1.2×10^{-1}	1.4×10^{-2}	1.2×10^{-1}	1.4×10^{-2}
S_{strat}	kg/m^4					-7.1×10^{-1}	1.5×10^{-1}	-6.9×10^{-1}	1.5×10^{-1}
R^2		0.287		0.326		0.331		0.358	

[a]Results are for M18, V1 (POWER trends), V2 (ROMS trends), and V3 (POWER + ROMS trends). The units of β and σ_β are mg/L over units of the corresponding trend variable.

confidence intervals (CIs) of the hypoxic area. Daily estimates of hypoxia were also averaged over a 30 day and 90 day moving window using the procedure followed in M18.[39]

3. RESULTS AND DISCUSSION

3.1. Deterministic Component. The first model iteration (V1) considers meteorological variables (from POWER), and only the square of the mean wind speed weighted over 28 days (WS_{28}^2) is selected based on BIC (Table 1 and Supporting Information SI-4). A comparison of trend coefficients and the associated standard errors shows that, with the exception of the time of year (T), all selected trend variables are significant ($p < 0.05$). However, T is still included because the higher-order variable (i.e., T^2) is significant.[70] Overall, the trends explain 32.6% of the variance in BWDO, compared to 28.7% in M18 (Table 1).[39]

The coefficients for the selected deterministic trends indicate patterns and suggest drivers of BWDO variability (Table 1). Negative coefficients associated with easting (E) and northing (N) indicate decreasing BWDO closer to the Mississippi and Atchafalaya River outfalls. The model quantifies an increase in BWDO with an increase in the wind speed, which is to be expected as high wind events promote water column mixing. Including the trend with WS_{28}^2 reduces the coefficient (β) associated with T^2 by a notable 53.6% relative to the M18 model. This indicates that the temporal trends in M18 were partly accounting for the seasonality in wind speeds, with lower wind speeds in mid-summer, compared to May and October.[71] Including the trend with wind also shifts the low point in the quadratic trend with time (T and T^2) from July 24 (M18) to July 15 (Figure S6), indicating that other biophysical factors (e.g., nutrient loading, algal production, temperature, and so on) promote severe hypoxia in mid-July, independent of the effects of wind speed. The selection of WS_{28}^2 is consistent with the influence of wind stress (a function of the square of the wind speed) on stratification and reoxygenation.[27,28,72]

The second version of the geostatistical model (V2) considers linear and quadratic trends with ROMS BWDO (DO′) and strength of stratification (S_{strat}) instead of meteorological variables. Both DO′ and S_{strat} are selected in this model, increasing the variance explained to 33.1% (Table 1). The coefficient for DO′ indicates that every 1 mg/L (31.25 mmol m^{-3}) change in ROMS BWDO is predictive of a 0.12 mg/L change in observed BWDO. Furthermore, the negative coefficient for S_{strat} indicates that as stratification increases, observed BWDO decreases further, consistent with the importance of stratification in controlling hypoxic variability.[44,64] A 25.0% decrease in the β for T^2 is observed relative to M18, as the trends with DO′ and S_{strat} capture a portion of seasonal variation in BWDO. However, the quadratic trend with T now has a minimum on July 24, similar to M18.

The third version (V3) considers both the ROMS and POWER trend variables together. Model selection identifies W_{28}^2, DO′, and S_{strat} as optimal, explaining 35.8% of the variability in observed BWDO, outperforming V1 and V2, and representing a 24.7% relative increase compared to M18 (Table 1). Furthermore, there is a 64.6% reduction in the coefficient for T^2 compared to M18, indicating that substantial seasonal variability can now be attributed to the new trend variables (particularly wind, per V1). Day-of-year temporal trends (T and T^2) are the lowest on July 15, as in V1. However, no significant differences in other trend parameters are observed when compared to V1, V2, and M18 (Table 1). In addition, the annual intercepts are similar across all model versions for most years, though the uncertainties (i.e., standard errors) in the intercepts are typically lowest in V3 (Supporting Information SI-5).

For all model versions, ROMS and POWER variables are selected based on an exhaustive search. As there is significant correlation among some candidate variables, we note that minor changes in variables (i.e., different averaging periods for the same meteorological variable) yield a similar performance. For example, in V3, the wind speed weighed over 21 days (instead of 28 days) produces an R^2 of 0.355, similar to the R^2 for the best V3 model (Supporting Information SI-4). However, it is noteworthy that the precipitation, sea surface temperature, and solar irradiance are never selected in the top 5% of models (based on the BIC score), indicating that these factors are not strong drivers of intraseasonal hypoxic variability.

ROMS DO′ is selected in the V2 and V3 models. However, the β associated with DO′ (0.12) indicates limited agreement between ROMS predictions and BWDO observations. A β of 1.0 would indicate perfect coherence, but this is not a realistic expectation. BWDO observations are subject to environmental microvariability, as indicated by the substantial nugget (σ_ε^2) component of the model variance (Section 3.2). Furthermore, like all hypoxia models, ROMS is not expected to be a perfect predictor of BWDO, given the uncertainties in model forcing

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

and boundary conditions, among other factors, such that it will not capture small-scale spatiotemporal variability in observations.[48] At the same time, the persistence of significant trends with day-of-year (T^2), wind, and S_{strat} (in model V3) suggest that there are temporal patterns in BWDO that could potentially be captured by ROMS through further refinement. While ROMS DO′ is moderately correlated with BWDO across our study period (R^2 = 0.18), previous ROMS Gulf studies have indicated potential for a higher predictive performance (R^2 > 0.3).[44]

3.2. Stochastic Component. A unique set of covariance parameters is determined for each model version (Supporting Information SI-7). The improved deterministic trends of V1−V3 result in up to a 10% decrease in the total stochastic variance (i.e., $\sigma_\varepsilon^2 + \sigma_\eta^2$) relative to the M18 value of 3.1 $mg^2\ L^{-2}$. This improvement is manifested as a reduction in the partial sill (correlated stochasticity, σ_η^2), which accounts for spatially and temporally correlated stochasticity. Correlated stochasticity is reduced by up to 12% in V3 relative to M18, as deterministic trends now explain a larger portion of the spatial and temporal patterns in BWDO. Consistent with this improvement, the ranges of spatial and temporal correlations are also moderately reduced relative to M18 (Supporting Information SI-7). At the same time, the uncorrelated stochasticity (nugget σ_ε^2), representing environmental microvariability and measurement error, is essentially constant across all model versions (0.5 $mg^2\ L^{-2}$), indicating noise in the observational data that cannot be resolved through spatiotemporal trends.

3.3. Performance Assessment. The model performance is assessed through leave-half-cruise-out and leave-one-cruise-out CVs. Here, we compare the performance of the geostatistical model using trends only ($x\beta$) versus using trends plus spatiotemporal correlation with proximate observations ($x\beta + \eta$), also known as universal kriging.[40] Including the spatiotemporal correlation substantially improves the model performance (Table 2) demonstrating the benefit of geostatistical framework. Moreover, the model performance is substantially better for the leave-half-cruise-out validations (R^2 > 0.5) than the leave-one-cruise-out validations (R^2 < 0.3), demonstrating the importance of having spatially and temporally proximate observations for improving the universal kriging skill. Finally, when comparing the CV metrics across the different model versions (Table 2), we see that V3 has consistently superior performance, indicating the efficacy of incorporating multiple dynamic covariates into the geostatistical modeling framework.

Table 2. Performance Assessment Metrics (R^2 and RMSE) for the Leave-One-Cruise-Out and Leave-Half-Cruise-Out CV, across the Different Model Versions Considered in This Study[a]

	leave-one-cruise-out CV ($x\beta$)		leave-one-cruise-out CV ($x\beta + \eta$)		leave-half-cruise-out CV ($x\beta + \eta$)	
model	R^2	RMSE (mg/L)	R^2	RMSE (mg/L)	R^2	RMSE (mg/L)
M18	0.133	1.890	0.223	1.734	0.570	1.271
V1	0.173	1.837	0.246	1.707	0.576	1.261
V2	0.203	1.797	0.269	1.676	0.570	1.271
V3	0.227	1.766	0.282	1.663	0.582	1.253

[a]Results are shown using predictions from deterministic trends only ($x\beta$) and from universal kriging considering both trends and spatiotemporal correlation ($x\beta + \eta$).

3.4. Hypoxic Area Estimates and Associated Uncertainties. Comparisons of shelfwide hypoxic area estimates (1985−2017) across model versions show that results from V3 substantially vary from M18 in some years (Figures 1 and S8C). These variations appear to be largely a result of the effects of the wind speed on hypoxia dynamics (Supporting Information SI-2). For example, 20% higher than average wind speeds in mid-July of 2003 led to a 31% reduction of the estimated hypoxic extent compared to M18. In contrast, a 12% lower than average wind speed from May through mid-June of 2012 contributed to a 40% increase in hypoxic area estimates in these months. These wind effects are consistent with those determined by ROMS simulations, particularly for 2003 (Supporting Information SI-2).[46] On average, daily estimates of hypoxia increased by over 8% when comparing models V1 and V3 to M18 (Table 3). This suggests that including the wind speed (squared) as a covariate in the model has a relatively large influence on hypoxic area estimates. In contrast, V2, which only added ROMS variables, increased the average hypoxic area estimate by just 2% relative to M18.

The interannual variability in the hypoxic extent is assessed by averaging daily estimates of hypoxia over the entire summer (June−August) and the 30 day period of maximum hypoxia (this period varies from year to year). These aggregations (Figure 2) show that our new fusion-based (V3) estimates exhibit generally similar patterns to M18, though there are some notable differences. For example, summer-average hypoxic area estimates for the salient years of 2008 and 1997 saw 22.5 and 14.3% increases in magnitude, respectively, relative to M18 (Figure 2B). Moreover, the summer-average estimate for 2017 increased by 89%, with a similar increase in the 30 day maximum (Figure 2). The BWDO observations used to estimate hypoxia for 2017 were from a single cruise on the far west side of the study area in the last week of June. That cruise detected substantial hypoxia in the sampled region even though wind speeds were relatively high at the time of the cruise. However, later in the summer, wind speeds dropped substantially and the ROMS model results indicated a hypoxic extent of over 30,000 km^2 (Figures S2I and S8I). The fusion model, which considers the already low BWDO at the end of June (leading to a low intercept parameter value for V3, Supporting Information SI-5), coupled with the trends in the wind speed and ROMS output, produced the high fusion-based estimate of the hypoxic area (Figure 2). This estimate is relatively consistent with the record LUMCON midsummer cruise measurement of 22,720 km^2 in late July,[73] even though the data from this cruise are not yet available for inclusion in our models.

The fusion-based approach (V3) results in larger average and maximum hypoxic estimates in most years, relative to M18. This increase was greater than 10% in 15 years for summer-averages and in 8 years for 30 day maximum hypoxic area. However, there were 10% or greater reductions in 3 years and 8 years, for the summer-average and 30 day maxima, respectively. Interestingly, the greatest reductions generally occurred when hypoxia was mild (1988, 1992, and 2003). Thus, the fusion-based approach appears to capture extremes in a better way, making severe years larger and mild years milder. This is consistent with the fact that V3 has higher predictive skills (e.g., Table 2) and thus captures more spatiotemporal variability in BWDO. At the same time, linear

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

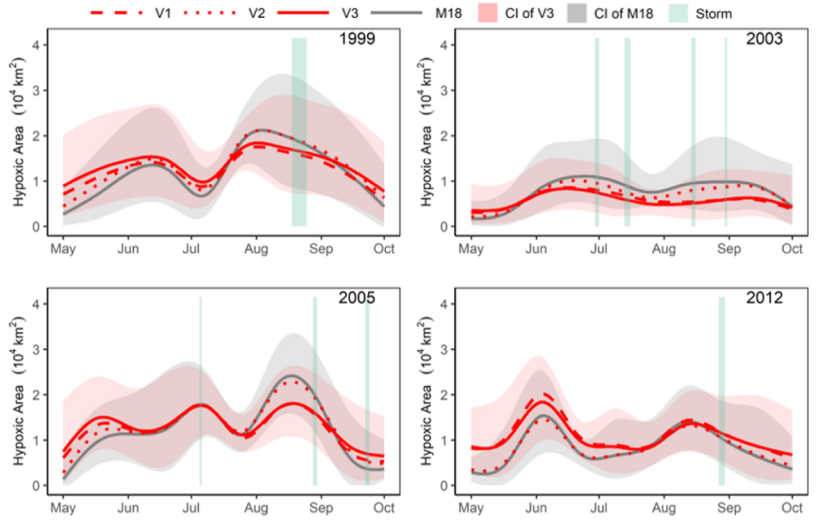


Figure 1. Daily estimates (at 3 day intervals) of the hypoxic extent across the summer for four example years. Estimates are from model versions V1, V2, V3, and M18. For clarity, 95% confidence intervals (CIs) are shown only for the preferred fusion model (V3) and the previous model (M18). Vertical bands represent times of major storms over the study area.[75]

Table 3. Mean of Estimated Hypoxic Extent (June−August), Mean 95% CI Width, and Mean Coefficient of Variation (C_v) for the Different Model Versions Considered in This Study

		daily estimates		summer averages	
model	hypoxic area (10^4 km^2)	95% CI (10^4 km^2)	C_v	95% CI (10^4 km^2)	C_v
M18	1.346	1.800	0.398	0.802	0.167
V1	1.456	1.819	0.370	0.782	0.152
V2	1.375	1.733	0.376	0.759	0.156
V3	1.467	1.756	0.356	0.739	0.143

trend analysis of the fusion-based estimates (performed similar to M18) identified no significant (p = 0.43) long-term trend in the hypoxic area over our study period, consistent with previous studies[37,45] and fairly stable levels of nutrient loading.[74]

Across all models, CIs indicate substantial uncertainty, reflecting temporal stochasticity in hypoxia dynamics that is not fully resolved by our available BWDO data and covariate trends (Figure 1, Table 3). In general, larger hypoxic area estimates lead to larger 95% CIs. However, when the uncertainty is scaled to the size of the hypoxic area, as in the coefficient of variation (C_v), our results show that model V3 has 11% less uncertainty than M18 (C_v of 0.356 vs 0.398). Furthermore, these reductions in C_v are found to vary based on the sampling frequency, with reductions of up to about 40% in months with sparse sampling (i.e., less than 30 samples per month, Supporting Information SI-9). In July, which is typically subject to intensive monitoring,[39] the average C_v reduction is just 6% (0.282 vs 0.299), compared to a 30% reduction in May (0.506 vs 0.725). These results demonstrate that the fusion approach is particularly effective in constraining uncertainties during periods of limited sampling. The fusion approach also reduces uncertainties in estimates of summer-average (Table 3, Figure 2) and 30 day maximum hypoxic area (Figure S10).

3.5. Summary and Implications. This study explores the use of meteorological variables and process-based model outputs as covariates within a geostatistical framework. While

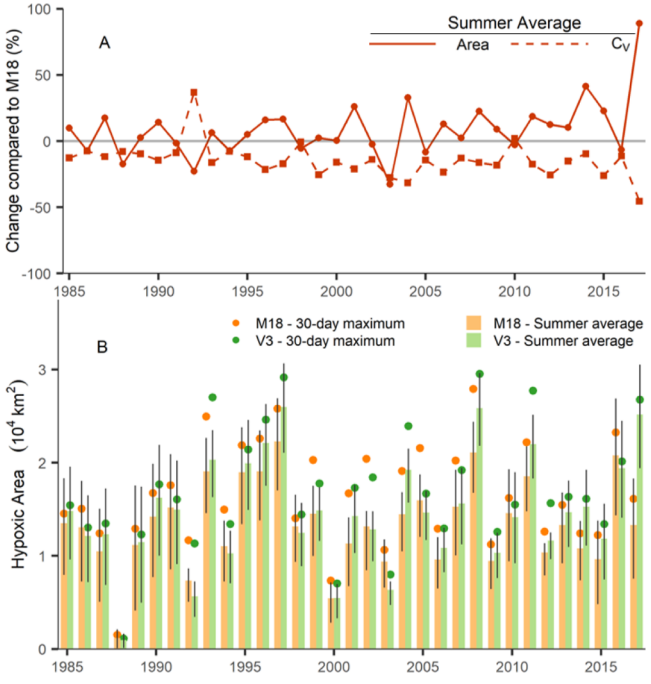


Figure 2. (A) Change in the estimated summer-average (June−August) hypoxic area and associated coefficient of variance (C_v) for the new fusion model relative to previous estimates (change = {V3 − M18}/M18) and (B) 30 day maximum and summer-average hypoxic extent from the new fusion model (V3) compared to M18. Error bars represent 95% CIs for the summer mean.

a limited number of previous studies considered similar covariates, they focused only on spatial correlation to provide snapshot estimates at times when data were collected.[38,50,51] This study considers the full spatiotemporal covariance structure of the response variable (BWDO) and allows for hypoxia estimation with rigorous uncertainty quantification across time. CV metrics indicate an improvement in the model performance because of the addition of these dynamic covariates, including the square of wind speed weighted over 28 days and predicted DO and strength of stratification from a hydrodynamic−biogeochemical model (ROMS).

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

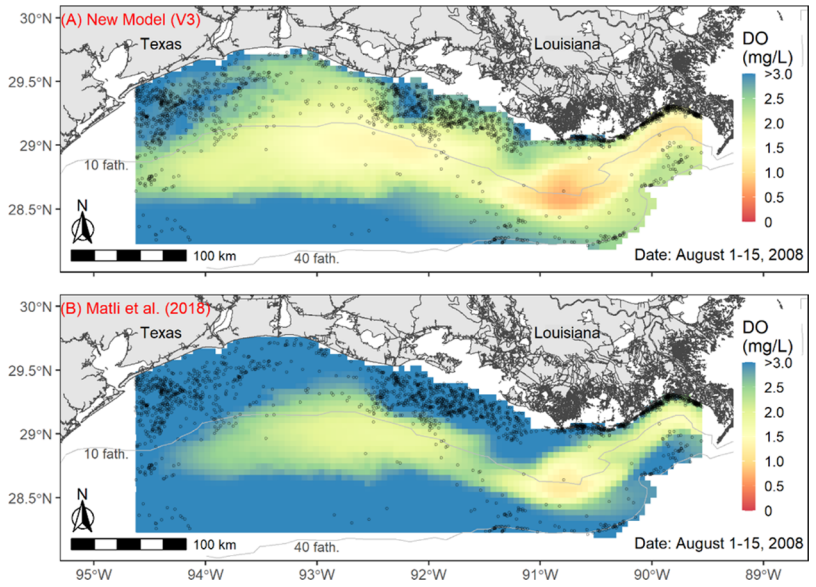


Figure 3. Two week average (August 1−15, 2008) bottom DO estimates from (A) the new model (V3) and (B) Matli et al. (2018) relative to the location of individual shrimp tows during the same period (black circles).[16,39] Note that this period was selected to illustrate a time when there is substantial variation between the two approaches.

We note that there are over 400 hypoxic zones worldwide, but few have been studied as intensively as the northern Gulf.[5] The proposed approach could be readily extended to other hypoxic systems where there are sufficient BWDO observations to define covariance parameters (>100 samples)[76] and wind speed data from local weather stations or climate models. Hydrodynamic−biogeochemical (e.g., ROMS) modeling output is a particularly valuable covariate but not explicitly required to improve the predictive performance (as indicated by model V1). Satellite remote sensing has also been linked to hypoxia[29] and could be considered as a potential covariate in future studies. Including dynamic covariates is expected to be more beneficial when assessing hypoxia in systems with limited monitoring,[77] as heavily monitored systems can be characterized more directly from the observational data.[78] This expectation is supported by our results showing larger uncertainty reductions in months with sparse sampling (Supporting Information SI-9). Also, in our study, the available covariates only explained about 23% of BWDO variability in CV. If covariates can be augmented to explain additional BWDO variability, then the benefits of the fusion-based approach could be greater than what is demonstrated here. In any case, by quantifying estimation uncertainty, the proposed approach can be used to assess the efficacy of different covariates and monitoring strategies.[42,79,80]

The new fusion-based estimates of hypoxic extent have substantially lower uncertainty (i.e., coefficient of variation) when compared to previous geostatistical estimates based on BWDO observations and static covariates only.[39] With increased confidence in estimated hypoxic areas, including 30 day maxima, this study provides enhanced information for tracking progress toward Gulf hypoxia reduction objectives (see Introduction).[10] The fusion-based estimates also better capture hypoxia dynamics related to wind mixing and coastal hydrodynamics, substantially revising our estimates of hypoxic severity in some years and months (e.g., Figure 3). These fusion-based estimates also go beyond the use of process-based models alone, by assimilating BWDO observations with ROMS output to help address model-data mismatch. At the same time, the study indicates covariate trends in hypoxic variability (with day-of-year, wind speed, and stratification) that could perhaps be resolved by ROMS through further model refinement.

The results from our study have implications for understanding the effects of ecosystem perturbations on economically important fisheries in the region. For example, the shrimp fishery is the highest valued single-species fishery in the northern Gulf and it peaks during summer months when hypoxia is most severe. Hypoxia affects the fishery in multiple ways, including by aggregating shrimp and fishing vessels near the edges of the hypoxic zone, suggesting locally high catch rates in these regions.[16,81] Initial comparisons suggest that the fusion model is generally consistent with these shifts in fleet distribution, with shrimping activity concentrated largely in moderately low oxygen bottom waters near the edges of the hypoxic zone (Figure 3A). In contrast, the previous (M18) geostatistical model does not always align as well with the spatial structure evident in the fishery (Figure 3B).[39] Because many harvested species are highly mobile and can move vertically in the water column,[82,83] future model developments that estimate the thickness and volume of hypoxia may also be beneficial. Overall, fusion-based hypoxia estimates, with enhanced accuracy and reduced uncertainty, have potential to improve the assessment and management of impacted coastal resources.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.0c03655.

Comparison of buoy and POWER wind speeds, time-series summary of POWER and ROMS data, spatial−temporal lag calculation, BIC scores, yearly intercepts, regression coefficients, parameter estimates, hypoxic area time−series plots, uncertainty across summer months, and comparison of new results with those of Matli et al.[39] (2018) (PDF)

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

AUTHOR INFORMATION

Corresponding Author

Venkata Rohith Reddy Matli − *Center for Geospatial Analytics, NC State University, Raleigh, North Carolina 27695, United States;* orcid.org/0000-0001-7788-0155; Phone: +1 919 749 8576; Email: vmatli@ncsu.edu

Authors

Arnaud Laurent − *Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada;* orcid.org/0000-0002-8545-9309

Katja Fennel − *Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada;* orcid.org/0000-0003-3170-2331

Kevin Craig − *National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Beaufort, North Carolina 28516, United States*

Jacob Krause − *National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Beaufort, North Carolina 28516, United States*

Daniel R. Obenour − *Center for Geospatial Analytics and Department of Civil, Construction and Environmental Engineering, NC State University, Raleigh, North Carolina 27695, United States;* orcid.org/0000-0002-7459-218X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.0c03655

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the NOAA grant NA16NOS4780203. This is an NGOMEX contribution 249.

REFERENCES

(1) Laffoley, D. D.; Baxter, J. M. *Ocean Deoxygenation: Everyone's Problem*; Union Internationale pour la Conservation de la Nature: Gland, Switzerland, 2019.

(2) Diaz, R. J.; Rosenberg, R. Marine benthic hypoxia: a review of its ecological effects and the behavioural responses of benthic macrofauna. *Oceanogr. Mar. Biol. Annu. Rev.* **1995**, *33*, 245−203.

(3) Rabalais, N. N.; Turner, R. E.; Scavia, D. Beyond Science into Policy: Gulf of Mexico Hypoxia and the Mississippi River. *BioScience* **2002**, *52*, 129−142.

(4) Meire, L.; Soetaert, K. E. R.; Meysman, F. J. R. Impact of global change on coastal oxygen dynamics and risk of hypoxia. *Biogeosciences* **2013**, *10*, 2633−2653.

(5) Diaz, R. J.; Rosenberg, R. Spreading dead zones and consequences for marine ecosystems. *Science* **2008**, *321*, 926−929.

(6) Glibert, P. M. Eutrophication, harmful algae and biodiversity—Challenging paradigms in a world of complex nutrient changes. *Mar. Pollut. Bull.* **2017**, *124*, 591−606.

(7) Glibert, P.; Burford, M. A.; Burford, M. Globally changing nutrient loads and harmful algal blooms: recent advances, new paradigms, and continuing challenges. *Oceanography* **2017**, *30*, 58−69.

(8) Rabalais, N. N.; Diaz, R. J.; Levin, L. A.; Turner, R. E.; Gilbert, D.; Zhang, J. Dynamics and distribution of natural and human-caused hypoxia. *Biogeosciences* **2010**, *7*, 585−619.

(9) *Harmful Algal Blooms and Hypoxia in the Great Lakes Research Plan and Action Strategy: An Interagency Report*; National Science and Technology Council: Washington D.C., 2017, https://cdn.coastalscience.noaa.gov/page-attachments/research/Harmful%20Algal%20Blooms%20Report%20FINAL%20August.2017.pdf (accessed May, 2020).

(10) *Mississippi River/Gulf of Mexico Watershed Nutrient Task Force2017 Report to Congress*; U.S. Environmental Protection Agency, Office of Wetlands, Oceans and Watersheds: Washington D.C., 2017, https://www.epa.gov/sites/production/files/2017-11/documents/hypoxia_task_force_report_to_congress_2017_final.pdf (accessed May, 2020).

(11) Texas A & M University. *Mechanisms Controlling Hypoxia*. 2020, https://sites.google.com/a/tamu.edu/hypoxia/home (accessed July, 2020).

(12) Craig, J.; Crowder, L. Hypoxia-induced habitat shifts and energetic consequences in Atlantic croaker and brown shrimp on the Gulf of Mexico shelf. *Mar. Ecol.: Prog. Ser.* **2005**, *294*, 79−94.

(13) O'Connor, T.; Whitall, D. Linking hypoxia to shrimp catch in the northern Gulf of Mexico. *Mar. Pollut. Bull.* **2007**, *54*, 460−463.

(14) Smith, M. D.; Asche, F.; Bennear, L. S.; Oglend, A. Spatial-dynamics of hypoxia and fisheries: The case of Gulf of Mexico brown shrimp. *Mar. Resour. Econ.* **2014**, *29*, 111−131.

(15) Briggs, K. B.; Craig, J. K.; Shivarudrappa, S.; Richards, T. M. Macrobenthos and megabenthos responses to long-term, large-scale hypoxia on the Louisiana continental shelf. *Mar. Environ. Res.* **2017**, *123*, 38−52.

(16) Purcell, K. M.; Craig, J. K.; Nance, J. M.; Smith, M. D.; Bennear, L. S. Fleet behavior is responsive to a large-scale environmental disturbance: Hypoxia effects on the spatial dynamics of the northern Gulf of Mexico shrimp fishery. *PloS One* **2017**, *12*, No. e0183032.

(17) Smith, M. D.; Oglend, A.; Kirkpatrick, A. J.; Asche, F.; Bennear, L. S.; Craig, J. K.; Nance, J. M. Seafood prices reveal impacts of a major ecological disturbance. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 1512−1517.

(18) *Action Plan for Reducing, Mitigating, and Controlling Hypoxia in the Northern Gulf of Mexico*; U.S. Environmental Protection Agency, Office of Wetlands, Oceans and Watersheds: Washington D.C., 2001, https://www.epa.gov/sites/production/files/2015-03/documents/2001_04_04_msbasin_actionplan2001.pdf (accessed July, 2020).

(19) *Gulf of Mexico Hypoxia Monitoring Implementation Plan*; Gulf of Mexico Hypoxia Monitoring Implementation Plan Steering Committee, 2009, https://coastalscience.noaa.gov/publications/handler.aspx?key=9218 (accessed May, 2020).

(20) Bever, A. J.; Friedrichs, M. A. M.; Friedrichs, C. T.; Scully, M. E.; Lanerolle, L. W. J. Combining observations and numerical model results to improve estimates of hypoxic volume within the Chesapeake Bay, USA. *J. Geophys. Res.: Oceans* **2013**, *118*, 4924−4944.

(21) Friedrich, J.; Janssen, F.; Aleynik, D.; Bange, H. W.; Boltacheva, N.; Çagatay, M. N.; Dale, A. W.; Etiope, G.; Erdem, Z.; Geraga, M.; Gilli, A.; Gomoiu, M. T.; Hall, P. O. J.; Hansson, D.; He, Y.; Holtappels, M.; Kirf, M. K.; Kononets, M.; Konovalov, S.; Lichtschlag, A.; Livingstone, D. M.; Marinaro, G.; Mazlumyan, S.; Naeher, S.; North, R. P.; Papatheodorou, G.; Pfannkuche, O.; Prien, R.; Rehder, G.; Schubert, C. J.; Soltwedel, T.; Sommer, S.; Stahl, H.; Stanev, E. V.; Teaca, A.; Tengberg, A.; Waldmann, C.; Wehrli, B.; Wenzhöfer, F. Investigating hypoxia in aquatic environments: diverse approaches to addressing a complex phenomenon. *Biogeosciences* **2014**, *11*, 1215−1259.

(22) Rabalais, N. N.; Turner, R. E.; Sen Gupta, B. K.; Boesch, D. F.; Chapman, P.; Murrell, M. C. Hypoxia in the northern Gulf of Mexico: Does the science support the plan to reduce, mitigate, and control hypoxia? *Estuaries Coasts* **2007**, *30*, 753−772.

(23) Forrest, D. R.; Hetland, R. D.; DiMarco, S. F. Multivariable statistical regression models of the areal extent of hypoxia over the Texas−Louisiana continental shelf. *Environ. Res. Lett.* **2011**, *6*, 045002.

(24) Feng, Y.; DiMarco, S. F.; Jackson, G. A. Relative role of wind forcing and riverine nutrient input on the extent of hypoxia in the northern Gulf of Mexico. *Geophys. Res. Lett.* **2012**, *39*, L09601.

(25) Walker, N. D.; Rabalais, N. N. Relationships among satellite chlorophylla, river inputs, and hypoxia on the Louisiana Continental shelf, Gulf of Mexico. *Estuaries Coasts* **2006**, *29*, 1081−1093.

(26) Scavia, D.; Evans, M. A.; Obenour, D. R. A scenario and forecast model for Gulf of Mexico hypoxic area and volume. *Environ. Sci. Technol.* **2013**, *47*, 10423−10428.
(27) Obenour, D. R.; Michalak, A. M.; Scavia, D. Assessing biophysical controls on Gulf of Mexico hypoxia through probabilistic modeling. *Ecol. Appl.* **2015**, *25*, 492−505.
(28) Yu, L.; Fennel, K.; Laurent, A. A modeling study of physical controls on hypoxia generation in the northern Gulf of Mexico. *J. Geophys. Res.: Oceans* **2015**, *120*, 5019−5039.
(29) Le, C.; Lehrter, J. C.; Hu, C.; Obenour, D. R. Satellite-based empirical models linking river plume dynamics with hypoxic area and volume. *Geophys. Res. Lett.* **2016**, *43*, 2693−2699.
(30) Fennel, K.; Testa, J. M. Biogeochemical controls on coastal hypoxia. *Annu. Rev. Mar. Sci.* **2019**, *11*, 105−130.
(31) Scavia, D.; Bertani, I.; Obenour, D. R.; Turner, R. E.; Forrest, D. R.; Katin, A. Ensemble modeling informs hypoxia management in the northern Gulf of Mexico. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 8823−8828.
(32) Justić, D.; Rabalais, N. N.; Turner, R. E. Simulated responses of the Gulf of Mexico hypoxia to variations in climate and anthropogenic nutrient loading. *J. Mar. Syst.* **2003**, *42*, 115−126.
(33) Greene, R. M.; Lehrter, J. C.; James, J. D. H., III Multiple regression models for hindcasting and forecasting midsummer hypoxia in the Gulf of Mexico. *Ecol. Appl.* **2009**, *19*, 1161−1175.
(34) Zirschky, J. Geostatistics for environmental monitoring and survey design. *Environ. Int.* **1985**, *11*, 515−524.
(35) Kearney, M.; Porter, W. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecol. Lett.* **2009**, *12*, 334−350.
(36) Scavia, D.; Justic, D.; Bierman, V. J. Reducing hypoxia in the Gulf of Mexico: advice from three models. *Estuaries* **2004**, *27*, 419−425.
(37) Obenour, D. R.; Scavia, D.; Rabalais, N. N.; Turner, R. E.; Michalak, A. M. Retrospective analysis of midsummer hypoxic area and volume in the northern Gulf of Mexico, 1985−2011. *Environ. Sci. Technol.* **2013**, *47*, 9808−9815.
(38) Murphy, R. R.; Curriero, F. C.; Ball, W. P. Comparison of spatial interpolation methods for water quality evaluation in the Chesapeake Bay. *J. Environ. Eng.* **2010**, *136*, 160−171.
(39) Matli, V. R. R.; Fang, S.; Guinness, J.; Rabalais, N. N.; Craig, J. K.; Obenour, D. R. Space-Time Geostatistical Assessment of Hypoxia in the Northern Gulf of Mexico. *Environ. Sci. Technol.* **2018**, *52*, 12484−12493.
(40) Chilès, J.; Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty*; John Wiley & Sons Inc.: New York, 1999.
(41) Holcomb, D. A.; Messier, K. P.; Serre, M. L.; Rowny, J. G.; Stewart, J. R. Geostatistical prediction of microbial water quality throughout a stream network using meteorology, land cover, and spatiotemporal autocorrelation. *Environ. Sci. Technol.* **2018**, *52*, 7775−7784.
(42) Fang, S.; Del Giudice, D.; Scavia, D.; Binding, C. E.; Bridgeman, T. B.; Chaffin, J. D.; Evans, M. A.; Guinness, J.; Johengen, T. H.; Obenour, D. R. A space-time geostatistical model for probabilistic estimation of harmful algal bloom biomass and areal extent. *Sci. Total Environ.* **2019**, *695*, 133776.
(43) Wang, L.; Justić, D. A modeling study of the physical processes affecting the development of seasonal hypoxia over the inner Louisiana-Texas shelf: Circulation and stratification. *Cont. Shelf Res.* **2009**, *29*, 1464−1476.
(44) Fennel, K.; Laurent, A.; Hetland, R.; Justić, D.; Ko, D. S.; Lehrter, J.; Murrell, M.; Wang, L.; Yu, L.; Zhang, W. Effects of model physics on hypoxia simulations for the northern Gulf of Mexico: A model intercomparison. *J. Geophys. Res.: Oceans* **2016**, *121*, 5731−5750.
(45) Del Giudice, D.; Matli, V. R. R.; Obenour, D. R. Bayesian mechanistic modeling characterizes Gulf of Mexico hypoxia: 1968−2016 and future scenarios. *Ecol. Appl.* **2019**, *30*, No. e02032.
(46) Laurent, A.; Fennel, K. Time-evolving, spatially explicit forecasts of the northern Gulf of Mexico hypoxic zone. *Environ. Sci. Technol.* **2019**, *53*, 14449−14458.
(47) Fennel, K.; Laurent, A. N and P as ultimate and proximate limiting nutrients in the northern Gulf of Mexico: implications for hypoxia reduction strategies. *Biogeosciences* **2018**, *15*, 3121−3131.
(48) Mattern, J. P.; Fennel, K.; Dowd, M. Sensitivity and uncertainty analysis of model hypoxia estimates for the Texas-Louisiana shelf. *J. Geophys. Res.: Oceans* **2013**, *118*, 1316−1332.
(49) Fennel, K.; Hetland, R.; Feng, Y.; DiMarco, S. A coupled physical-biological model of the Northern Gulf of Mexico shelf: model description, validation and analysis of phytoplankton variability. *Biogeosciences* **2011**, *8*, 1881−1899.
(50) Rühaak, W.; Bär, K.; Sass, I. Combining numerical modeling with geostatistical interpolation for an improved reservoir exploration. *Energy Procedia* **2014**, *59*, 315−322.
(51) Goovaerts, P.; Trinh, H. T.; Demond, A.; Franzblau, A.; Garabrant, D.; Gillespie, B.; Lepkowski, J.; Adriaens, P. Geostatistical modeling of the spatial distribution of soil dioxins in the vicinity of an incinerator. 1. Theory and application to Midland, Michigan. *Environ. Sci. Technol.* **2008**, *42*, 3648−3654.
(52) Peña, M. A.; Katsev, S.; Oguz, T.; Gilbert, D. Modeling dissolved oxygen dynamics and hypoxia. *Biogeosciences* **2010**, *7*, 933.
(53) Water temperature, salinity, dissolved oxygen, and other measurements from CTD taken from NOAA Ship Oregon II in the Gulf of Mexico from 2017-06-09 to 2017-07-19 as part of the Southeast Area Monitoring and Assessment Program (SEAMAP) (NCEI Accession 0166524). https://accession.nodc.noaa.gov/0166524 (accessed 2020-05).
(54) NODC Standard Product: Texas-Louisiana Shelf Circulation and Transport Processes Study: Hydrographic, Drifters, ADCP, and miscellaneous sensors (NCEI Accession 9800141). https://accession.nodc.noaa.gov/9800141 (accessed 2020-05).
(55) ETOPO1 Global Relief model. https://www.ngdc.noaa.gov/mgg/global/global.html (accessed 2020-05).
(56) Stackhouse, P. W., Jr.; Zhang, T.; Westberg, D.; Barnett, A. J.; Bristow, T.; Macpherson, B.; Hamilton, B. A. *POWER Release 8.0. 1 (With GIS Applications) Methodology (Data Parameters, Sources, & Validation) Documentation Date December 12, 2018 (All Previous Versions Are Obsolete)*, (Data Version 8.0. 1; Web Site Version 1.1. 0), 2018.
(57) Laurent, A.; Fennel, K.; Cai, W. J.; Huang, W. J.; Barbero, L.; Wanninkhof, R. Eutrophication-induced acidification of coastal waters in the northern Gulf of Mexico: Insights into origin and processes from a coupled physical-biogeochemical model. *Geophys. Res. Lett.* **2017**, *44*, 946−956.
(58) Millero, F. J.; Poisson, A. International one-atmosphere equation of state of seawater. *Deep-Sea Res., Part A* **1981**, *28*, 625−629.
(59) Simpson, J. H. The shelf-sea fronts: implications of their existence and behaviour. *Phil. Trans. Roy. Soc. Lond. Math. Phys. Sci.* **1981**, *302*, 531−546.
(60) Fennel, K.; Hu, J.; Laurent, A.; Marta-Almeida, M.; Hetland, R. Sensitivity of hypoxia predictions for the northern Gulf of Mexico to sediment oxygen consumption and model nesting. *J. Geophys. Res.: Oceans* **2013**, *118*, 990−1002.
(61) Clark, I. Statistics or geostatistics? Sampling error or nugget effect? *J. South. Afr. Inst. Min. Metall.* **2010**, *110*, 307−312.
(62) Kitanidis, P. K. *Introduction to Geostatistics: Applications in Hydrogeology*; Cambridge University Press, 1997.
(63) Li, Y.; Nowlin, W. D.; Reid, R. O. Mean hydrographic fields and their interannual variability over the Texas-Louisiana continental shelf in spring, summer, and fall. *J. Geophys. Res.: Oceans* **1997**, *102*, 1027−1049.
(64) Obenour, D. R.; Michalak, A. M.; Zhou, Y.; Scavia, D. Quantifying the impacts of stratification and nutrient loading on hypoxia in the northern Gulf of Mexico. *Environ. Sci. Technol.* **2012**, *46*, 5489−5496.

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

(65) Diggle, P. J.; Ribeiro, P. J., Jr. Bayesian inference in Gaussian model-based geostatistics. *Geogr. Environ. Model.* **2002**, *6*, 129−146.
(66) Schwarz, G. Estimating the dimension of a model. *Ann. Math. Stat.* **1978**, *6*, 461−464.
(67) Cavanaugh, J. E. *Model Selection: Bayesian Information Criterion*; Wiley Online Library, 2009.
(68) Mueller, K. L.; Yadav, V.; Curtis, P. S.; Vogel, C.; Michalak, A. M. Attributing the variability of eddy-covariance CO2 flux measurements across temporal scales using geostatistical regression for a mixed northern hardwood forest. *Global Biogeochem. Cycles* **2010**, *24*, GB3023.
(69) Roberts, D. R.; Bahn, V.; Ciuti, S.; Boyce, M. S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J. J.; Schröder, B.; Thuiller, W.; Warton, D. I.; Wintle, B. A.; Hartig, F.; Dormann, C. F. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913−929.
(70) Faraway, J. J. *Linear Models with R*; CRC Press, 2014.
(71) Muller-Karger, F. E.; Smith, J. P.; Werner, S.; Chen, R.; Roffer, M.; Liu, Y.; Muhling, B.; Lindo-Atichati, D.; Lamkin, J.; Cerdeira-Estrada, S.; Enfield, D. B. Natural variability of surface oceanographic conditions in the offshore Gulf of Mexico. *Prog. Oceanogr.* **2015**, *134*, 54−76.
(72) Walker, N. D.; Wiseman, W. J., Jr.; Rouse, L. J., Jr.; Babin, A. Effects of river discharge, wind stress, and slope eddies on circulation and the satellite-observed structure of the Mississippi River plume. *J. Coast Res.* **2005**, *216*, 1228−1244.
(73) Louisiana Universities Marine Consortium. *Shelfwide Cruise*, 2017, July 24−July 21, Press Release; Gulf Hypoxia, 2017, https://gulfhypoxia.net/research/shelfwide-cruise/?y=2017&p=press_release (accessed July, 2020).
(74) *USGS Nutrient Flux for the Mississippi River Basin and Subbasins*; U.S. Geological Survey, 2017, https://toxics.usgs.gov/hypoxia/mississippi/nutrient_flux_yield_est.html (accessed July, 2020).
(75) *Atlantic Hurricane Season*; National Hurricane Center, 2020, https://www.nhc.noaa.gov/data/tcr/(accessed May, 2020).
(76) Oliver, M. A.; Webster, R. *Basic Steps in Geostatistics: The Variogram and Kriging*; Springer, 2015.
(77) Zhou, Y.; Obenour, D. R.; Scavia, D.; Johengen, T. H.; Michalak, A. M. Spatial and temporal trends in Lake Erie hypoxia, 1987−2007. *Environ. Sci. Technol.* **2013**, *47*, 899−905.
(78) Lee, Y. J.; Lwiza, K. M. M. Characteristics of bottom dissolved oxygen in Long Island sound, New York. *Estuar. Coast Shelf Sci.* **2008**, *76*, 187−200.
(79) Dowdall, M.; Gerland, S.; Karcher, M.; Gwynn, J. P.; Rudjord, A. L.; Kolstad, A. K. Optimisation of sampling for the temporal monitoring of technetium-99 in the Arctic marine environment. *J. Environ. Radioact.* **2005**, *84*, 111−130.
(80) Golnick, P. C.; Chaffin, J. D.; Bridgeman, T. B.; Zellner, B. C.; Simons, V. E. A comparison of water sampling and analytical methods in western Lake Erie. *J. Great Lake. Res.* **2016**, *42*, 965−971.
(81) Craig, J. Aggregation on the edge: effects of hypoxia avoidance on the spatial distribution of brown shrimp and demersal fishes in the Northern Gulf of Mexico. *Mar. Ecol.: Prog. Ser.* **2012**, *445*, 75−95.
(82) Zhang, H.; Ludsin, S. A.; Mason, D. M.; Adamack, A. T.; Brandt, S. B.; Zhang, X.; Kimmel, D. G.; Roman, M. R.; Boicourt, W. C. Hypoxia-driven changes in the behavior and spatial distribution of pelagic fish and mesozooplankton in the northern Gulf of Mexico. *J. Exp. Mar. Biol. Ecol.* **2009**, *381*, S80−S91.
(83) Langseth, B. J.; Schueller, A. M.; Shertzer, K. W.; Craig, J. K.; Smith, J. W. Management implications of temporally and spatially varying catchability for the Gulf of Mexico menhaden fishery. *Fish. Res.* **2016**, *181*, 186−197.

https://dx.doi.org/10.1021/acs.est.0c03655
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Supporting Information

for

"Fusion-based hypoxia estimates: combining geostatistical and mechanistic models of dissolved oxygen variability"

by

Venkata Rohith Reddy Matli,* Arnaud Laurent, Katja Fennel, Kevin Craig, Jacob Krause, and Daniel R Obenour

consisting of

10 sections with 22 figures and 5 tables in 27 pages.

*corresponding author, vmatli@ncsu.edu

Table of Contents

SI-1: Comparison of windspeeds from Buoys and POWER

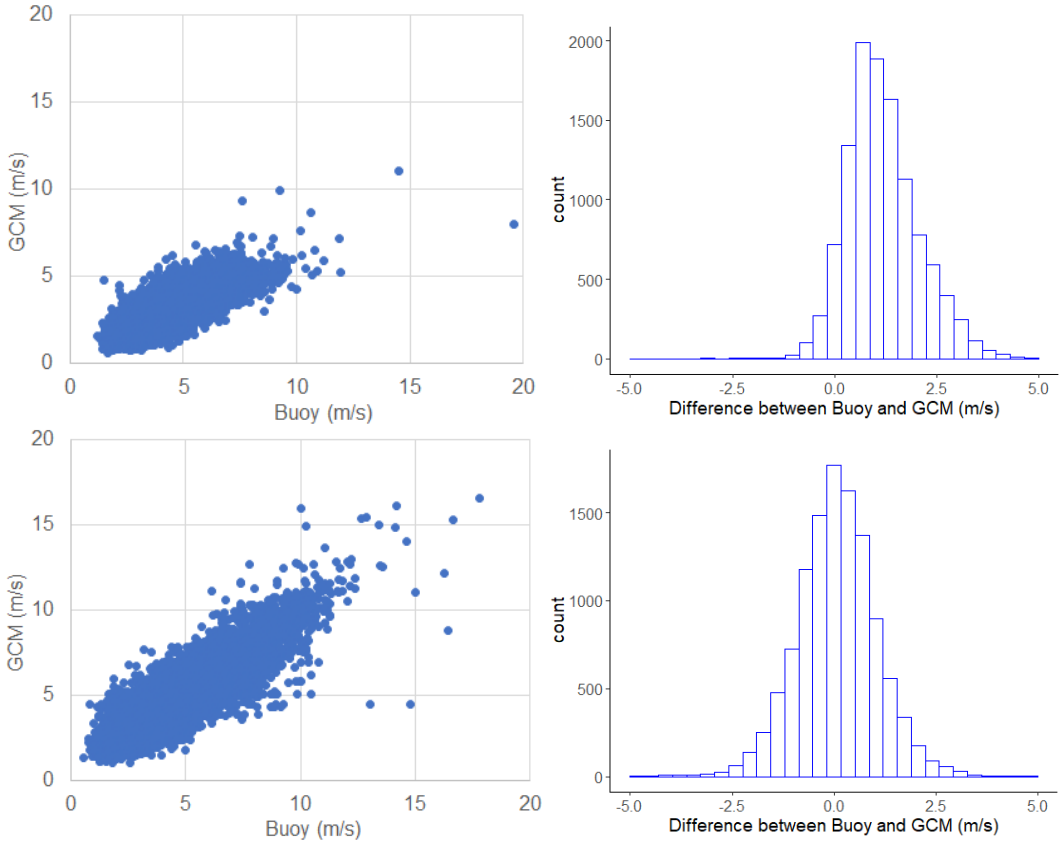


Figure S1. Comparison of wind speeds from POWER (GCM) and Buoy data (left) and histogram of difference between Buoy and GCM data (right). Figures on top are data collected from Buoy on the west end of the study area (Sabine Pass, Texas-SRST2). Figures on bottom are data collected from Buoy on the east end of the study area (Southwest Pass, Louisiana-BURL1). The Buoy data are retrieved from National Data Buoy Center (NDBC) archives.

SI-2: Shelf wide wind speeds from POWER and hypoxia extent from ROMS model

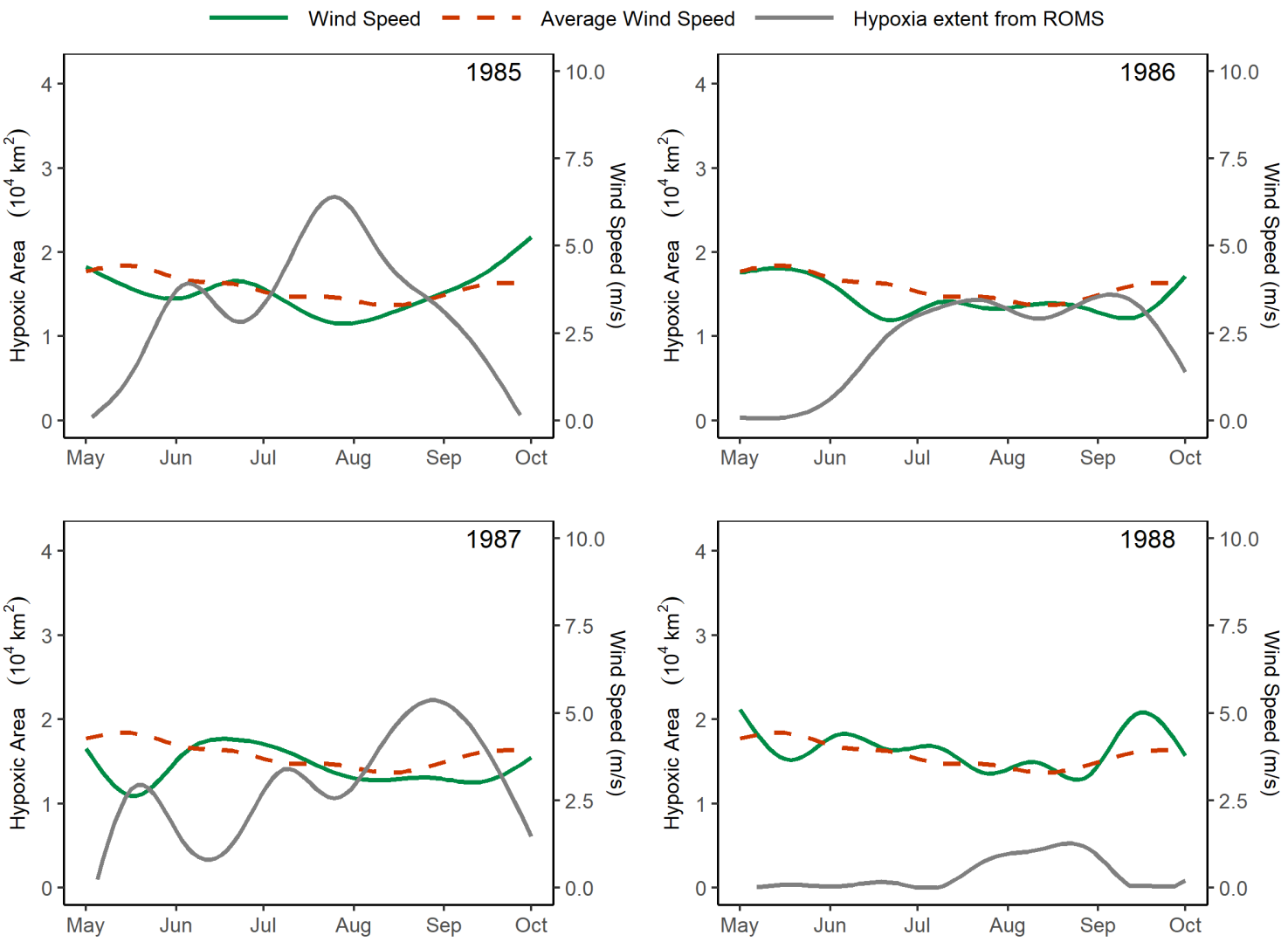


Figure S2A. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 1985-1988

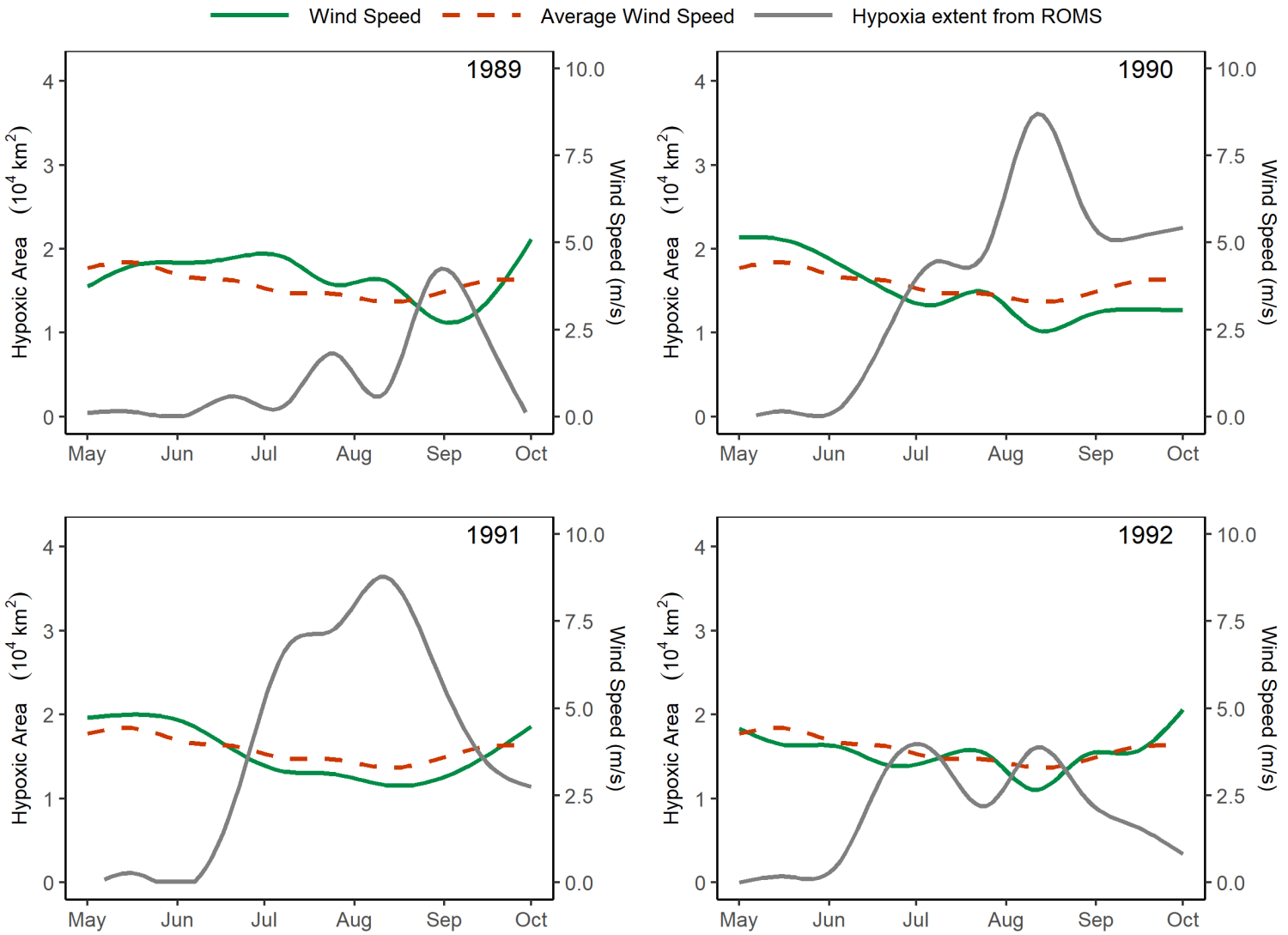


Figure S2B. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 1989-1992

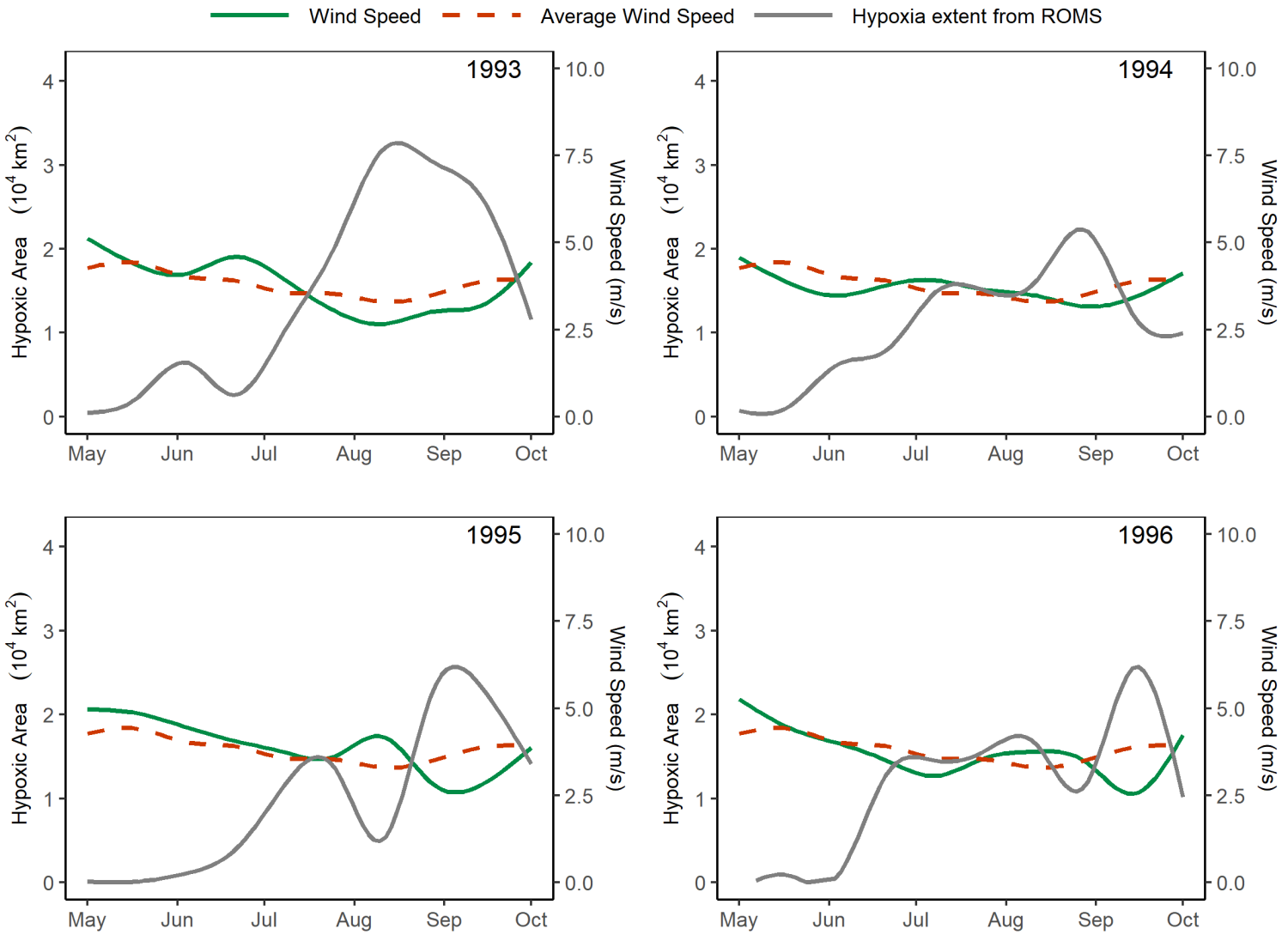


Figure S2C. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 1993-1996

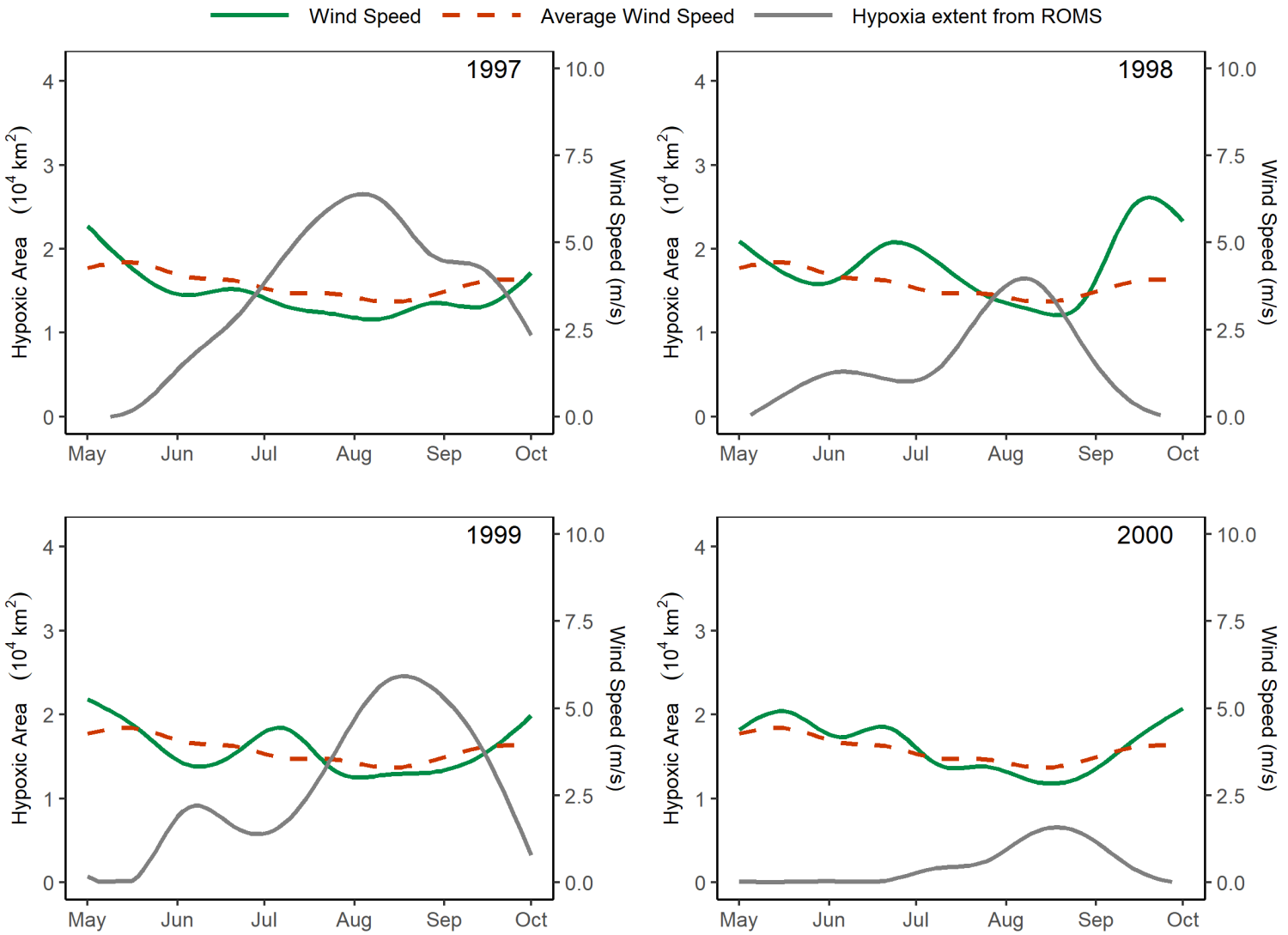


Figure S2D. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 1997-2000

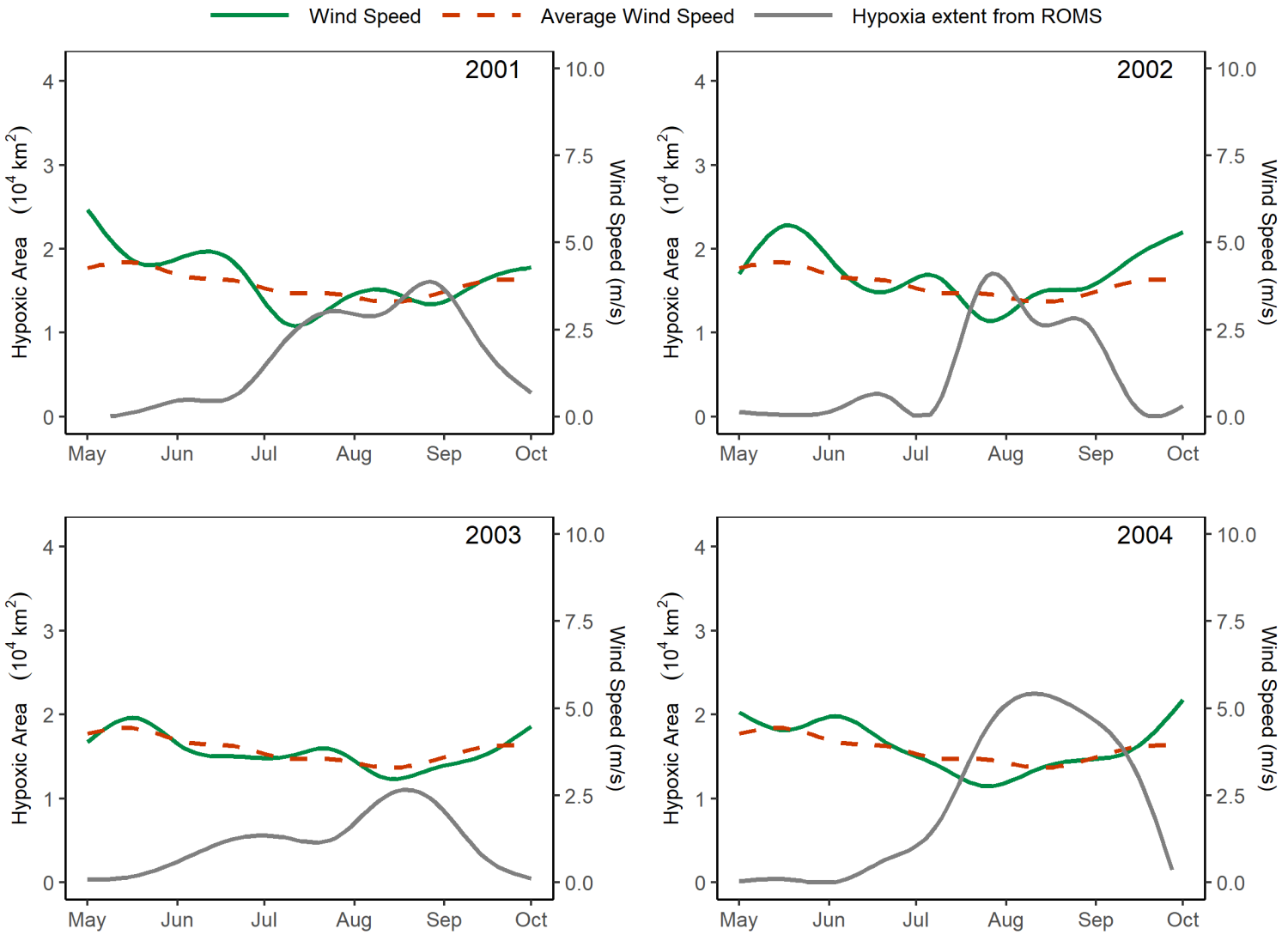


Figure S2E. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 2001-2004

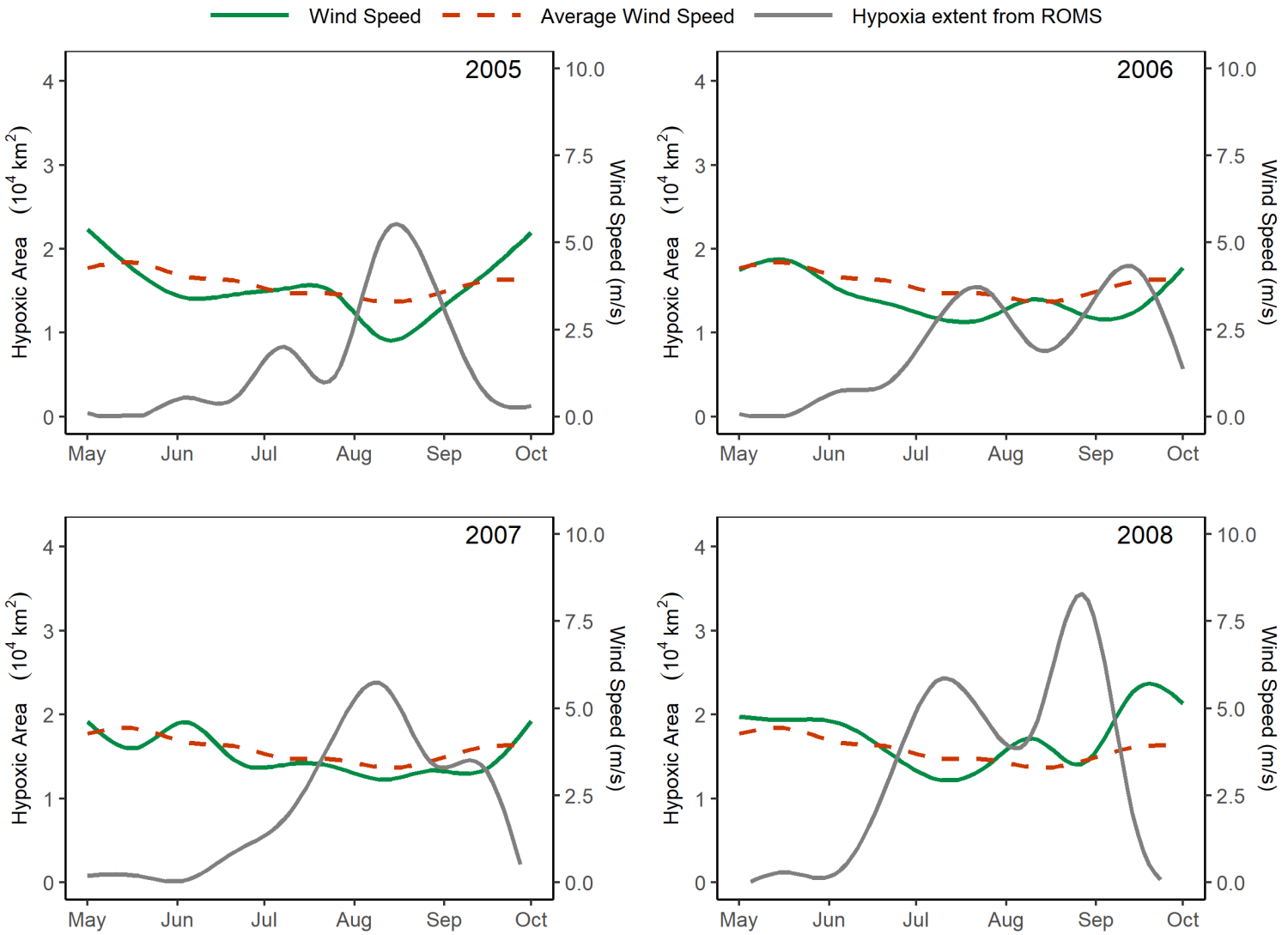


Figure S2F. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 2005-2008

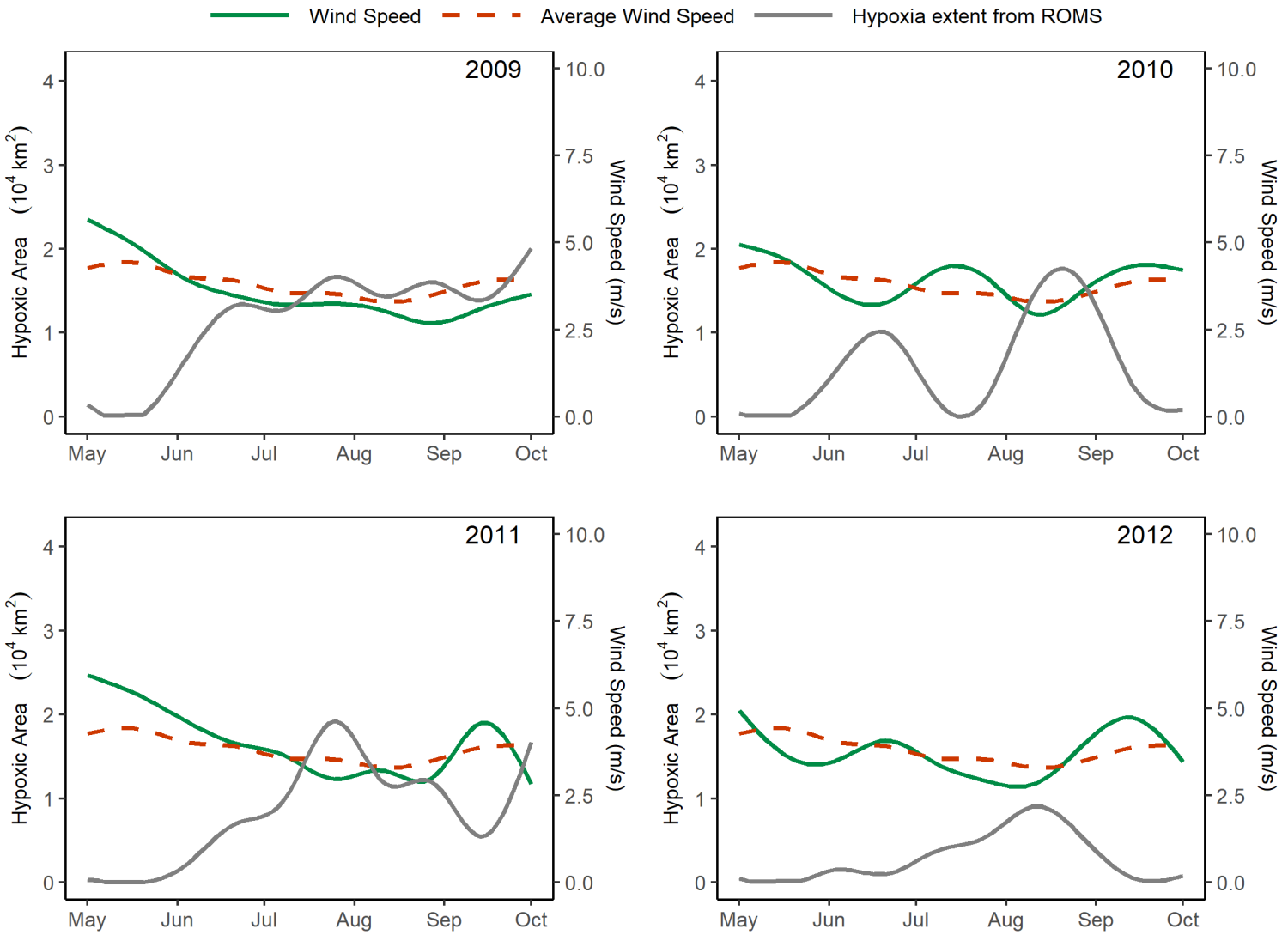


Figure S2G. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 2009-2012

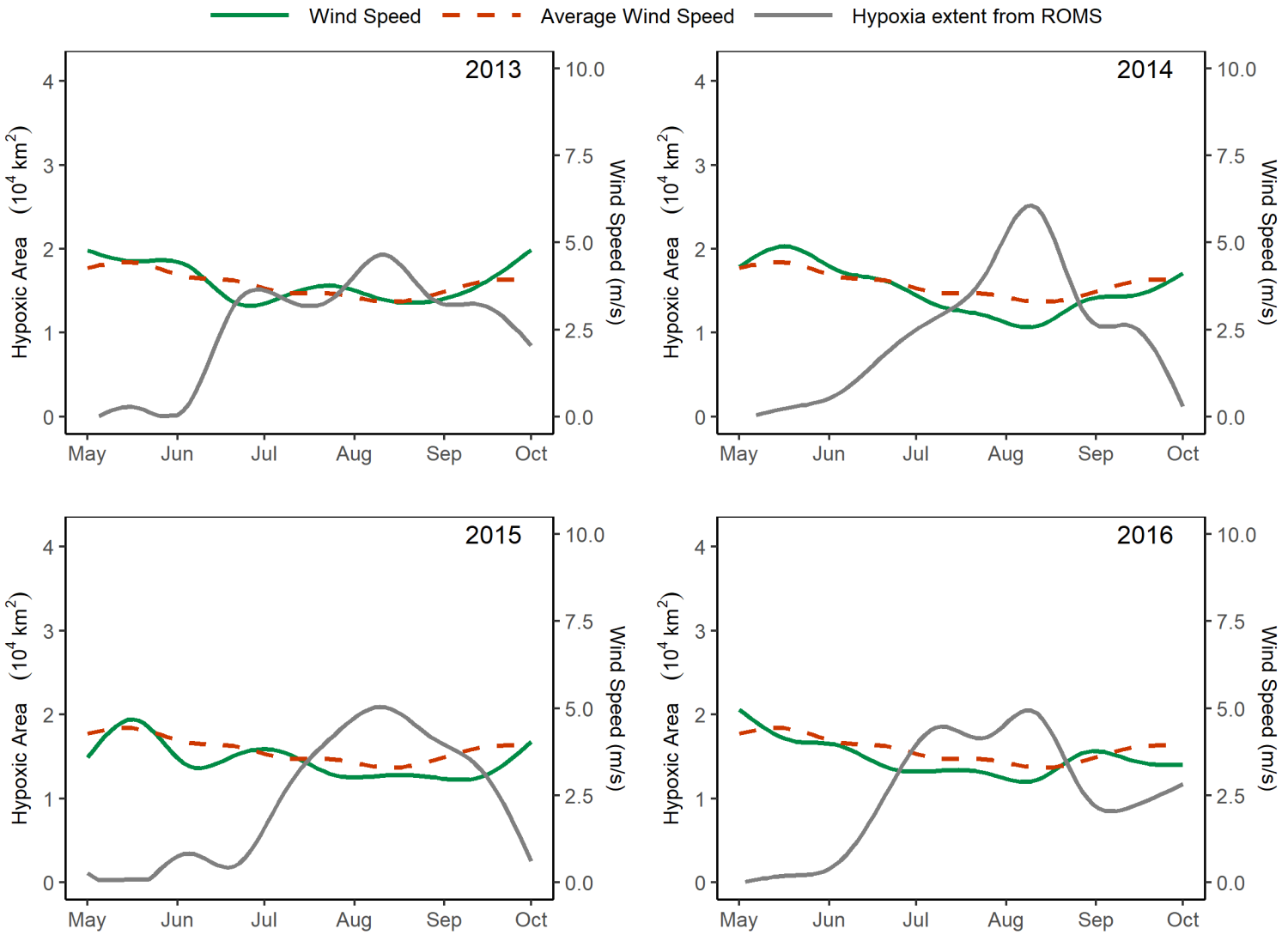


Figure S2H. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 2013-2016

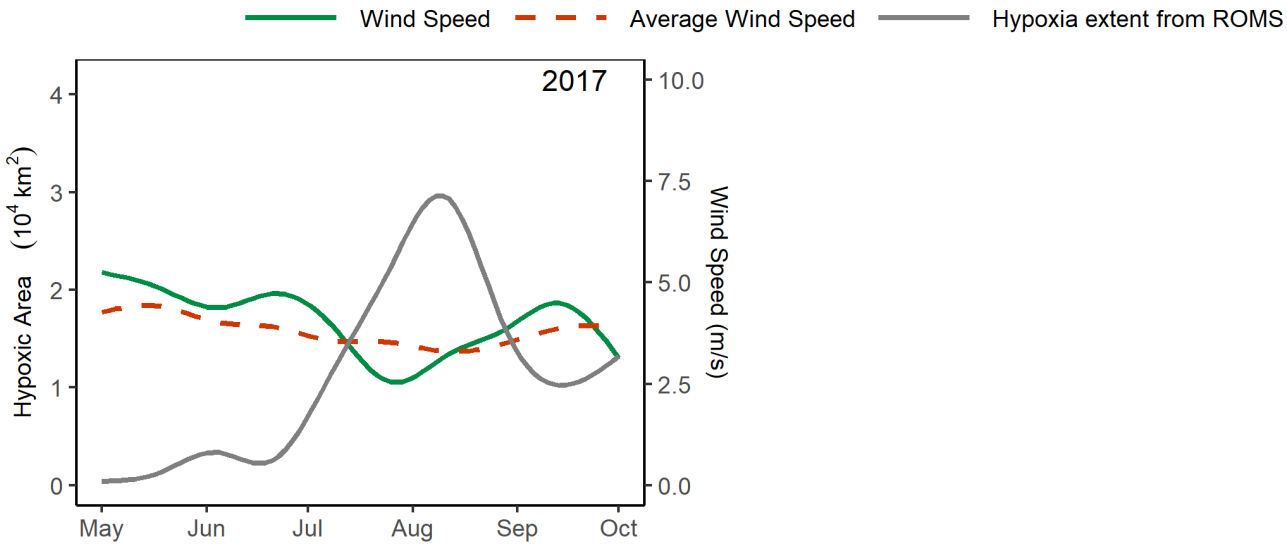


Figure S2I. Daily shelfwide wind speeds compared to multi-year averages and hypoxic area obtained from ROMS model from 2017

SI-3: Spatial-temporal lag calculation

Spatial lag ($s_{i,j}$, km) $= \sqrt{(\alpha * s_{yy})^2 + s_{xx}^2}$

Where α is spatial anisotropy and s_{yy} is the distance between two observations (i and j) along North-South direction and s_{xx} is the distance between two observations along East-West direction.

Temporal lag ($t_{i,j}$, days) = t_i - t_j

Where t_i and t_j are the dates on which observations i and j were collected.

SI-4: BIC scores and R^2 for example variable combinations tested

Variables Used											BIC score	R^2
N	E	D	D^2	T	T^2	W_{28}^2	DO'	S_{strat}	--	--	7872	0.358
N	E	D	D^2	T	T^2	W_{21}^2	DO'	S_{strat}	--	--	7873	0.355
N	E	D	D^2	T	T^2	W_{28}^2	DO'	DO'^{2}	S_{strat}	--	7874	0.359
N	E	D	D^2	T	T^2	W_{28}	W_{28}^2	DO'	S_{strat}	--	7876	0.356
N	E	D	D^2	T	T^2	W_{21}^2	DO'	DO'^{2}	S_{strat}	--	7876	0.357
N	E	D	D^2	T	T^2	W_{28}^2	DO'	--	--	--	7885	0.357
N	E	D	D^2	T	T^2	W_{21}^2	DO'	--	--	--	7886	0.354
N	E	D	D^2	T	T^2	DO'	S_{strat}	--	--	--	7888	0.333
N	E	D	D^2	T	T^2	W_{14}^2	DO'	--	--	--	7889	0.351
N	E	D	D^2	T	T^2	DO'	--	--	--	--	7903	0.331
N	E	D	D^2	T	T^2	W_{28}^2	--	--	--	--	7991	0.326
N	E	D	D^2	T	T^2	W_{21}^2	--	--	--	--	7992	0.325
N	E	D	D^2	T	T^2	--	--	--	--	--	8095	0.287

SI-5: Yearly intercepts (β, mgL^{-1}) and associated standard errors (σ_β, mgL^{-1}) from models M18, V1, V2 and V3.

year	V3		V1		V2		M18	
	β	σ_β	β	σ_β	β	σ_β	β	σ_β
1985	4.25	0.53	4.86	0.54	4.06	0.56	4.84	0.58
1986	4.13	0.59	4.90	0.59	4.00	0.61	4.97	0.63
1987	4.12	0.63	5.20	0.61	3.99	0.66	5.35	0.66
1988	6.14	0.70	7.24	0.69	6.30	0.74	7.69	0.75
1989	4.35	0.76	5.49	0.76	4.28	0.80	5.72	0.81
1990	4.32	0.60	5.29	0.59	4.08	0.63	5.30	0.63
1991	4.60	0.59	5.29	0.60	4.28	0.62	5.16	0.64
1992	4.74	0.35	5.34	0.34	4.67	0.35	5.42	0.35
1993	3.64	0.39	4.52	0.37	3.45	0.40	4.55	0.38
1994	4.33	0.42	5.28	0.39	4.23	0.43	5.41	0.42
1995	3.56	0.48	4.54	0.46	3.36	0.50	4.60	0.50
1996	3.83	0.45	4.56	0.45	3.69	0.47	4.62	0.48
1997	3.50	0.44	4.18	0.44	3.31	0.46	4.15	0.47
1998	3.78	0.43	5.00	0.37	3.63	0.44	5.17	0.38
1999	3.97	0.39	4.84	0.36	3.84	0.40	4.94	0.38
2000	4.96	0.38	5.76	0.36	5.06	0.39	6.08	0.38
2001	4.48	0.38	5.35	0.35	4.47	0.39	5.57	0.37
2002	4.26	0.38	5.09	0.36	4.31	0.39	5.37	0.37
2003	4.52	0.35	5.33	0.33	4.44	0.36	5.47	0.34
2004	4.02	0.36	4.80	0.34	3.97	0.36	4.95	0.35
2005	4.04	0.35	4.82	0.33	4.00	0.36	5.00	0.34
2006	4.82	0.32	5.32	0.31	4.79	0.32	5.41	0.32
2007	4.36	0.33	5.01	0.32	4.29	0.34	5.11	0.33
2008	3.51	0.39	4.33	0.37	3.40	0.40	4.44	0.39
2009	4.67	0.37	5.38	0.36	4.61	0.38	5.51	0.38
2010	3.70	0.43	4.79	0.39	3.64	0.44	5.01	0.41
2011	3.64	0.38	4.52	0.35	3.55	0.39	4.67	0.37
2012	4.39	0.35	5.09	0.34	4.44	0.36	5.32	0.35
2013	4.28	0.41	5.06	0.40	4.17	0.43	5.16	0.42
2014	4.43	0.40	5.19	0.39	4.34	0.42	5.30	0.41
2015	4.53	0.51	5.39	0.51	4.48	0.53	5.56	0.54
2016	3.77	0.45	4.36	0.45	3.64	0.47	4.38	0.48
2017	3.47	0.59	4.90	0.54	3.25	0.62	5.04	0.58

SI-6: Trend coefficients (β) and associated standard errors (σ_β) of normalized BIC-selected trend variables from all model versions.

Variable	M18		V1		V2		V3	
	β	σ_β	β	σ_β	β	σ_β	β	σ_β
E	-0.605	0.060	-0.559	0.057	-0.487	0.057	-0.453	0.055
N	-0.270	0.057	-0.245	0.055	-0.259	0.054	-0.239	0.053
D	-2.519	0.114	-2.535	0.114	-2.468	0.117	-2.485	0.116
D^2	2.639	0.112	2.657	0.112	2.480	0.113	2.503	0.113
T	0.131	0.090	0.131	0.084	0.088	0.084	0.089	0.080
T^2	0.482	0.086	0.233	0.091	0.368	0.080	0.172	0.086
WS_{28}^2	-	-	0.412	0.069	-	-	0.331	0.066
DO'	-	-	-	-	0.227	0.027	0.218	0.027
S_{strat}	-	-	-	-	-0.142	0.030	-0.139	0.030

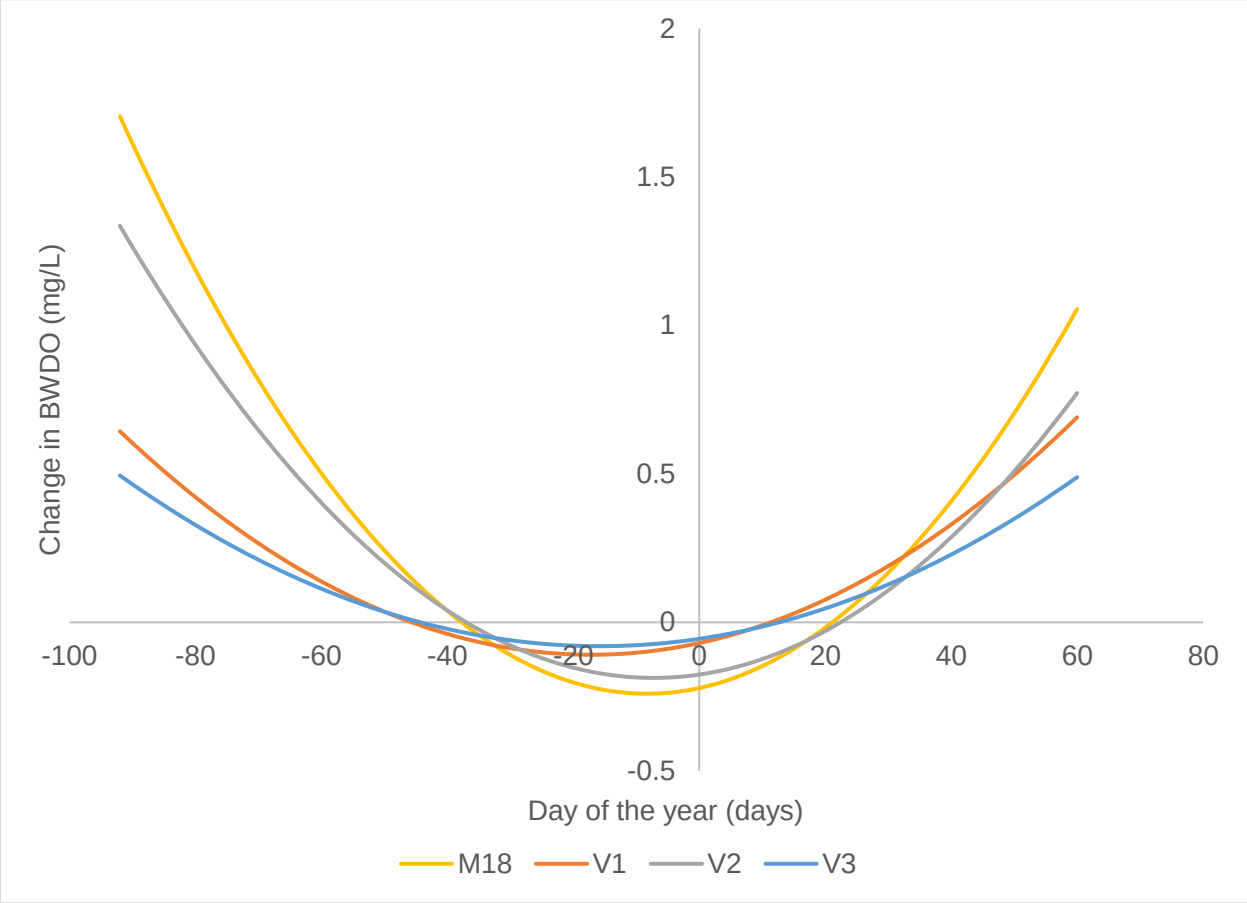


Fig S6: Temporal trend between BWDO and day of the year (centered on August 1st, day 0). Note that this is just one component of the deterministic trend function (xβ) and geostatistical model (eq 1); other factors such as wind speed and ROMS DO also inform temporal variability.

SI-7: Geostatistical covariance function parameters σ_ε^2 (nugget), σ_η^2 (partial sill), $\sigma_\varepsilon^2 + \sigma_\eta^2$ (total sill), α (spatial anisotropy ratio, east-west to north-south), a (spatial range parameter), b and (temporal range parameter) of all model versions.

Parameters	M18	V1	V2	V3
σ_ε^2 ($mg^2 L^{-2}$)	0.49	0.49	0.50	0.49
σ_η^2 ($mg^2 L^{-2}$)	2.57	2.44	2.34	2.26
$\sigma_\varepsilon^2 + \sigma_\eta^2$ ($mg^2 L^{-2}$)	3.06	2.93	2.84	2.76
a (km)	63.54	58.63	57.92	54.89
b (days)	8.92	8.37	8.34	7.91
α	2.26	2.26	2.21	2.21

SI-8: Daily estimates of hypoxia from models M18, V1, V2, and V3 for years 1985-2017

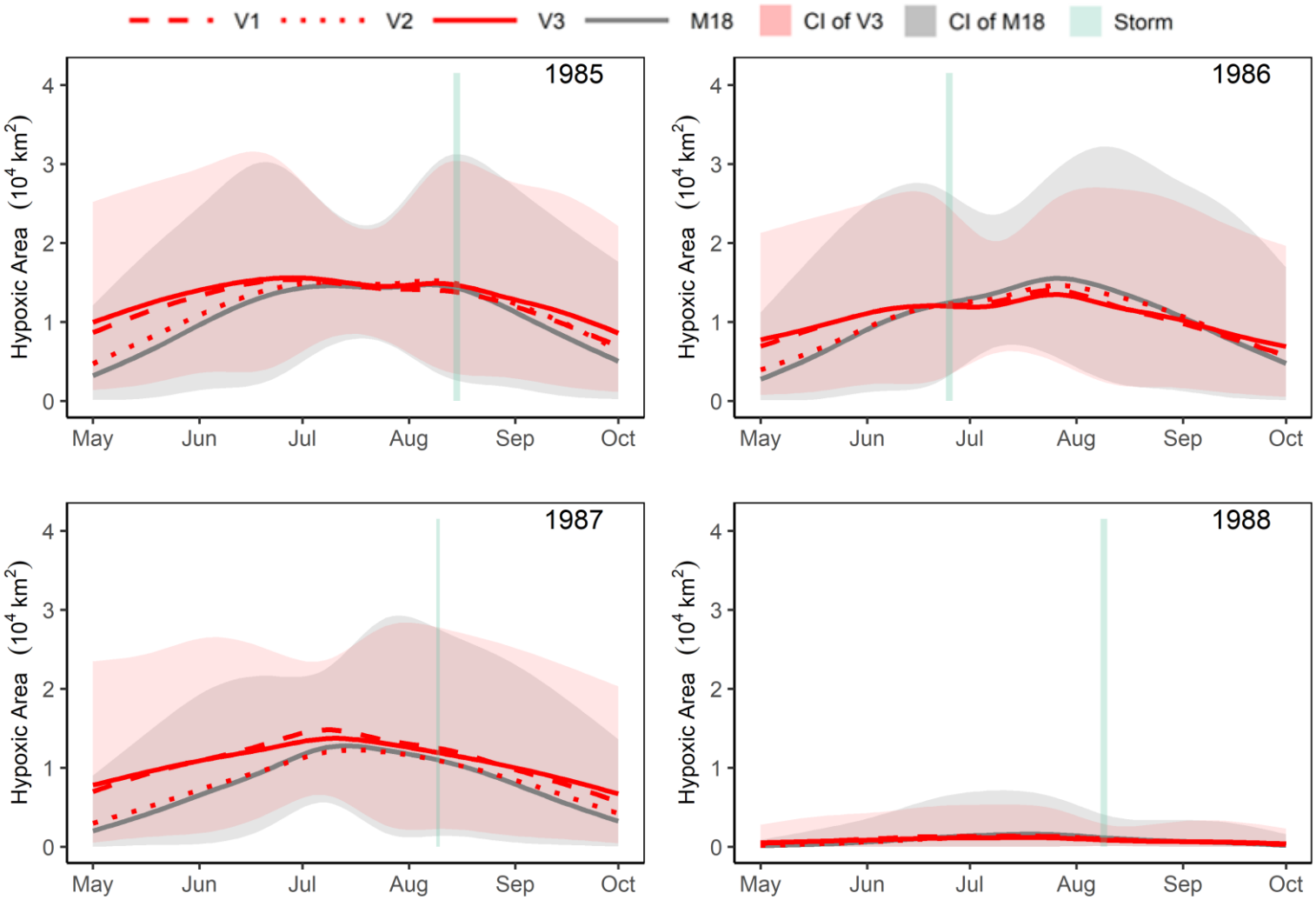


Figure S8A. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 1985-1988

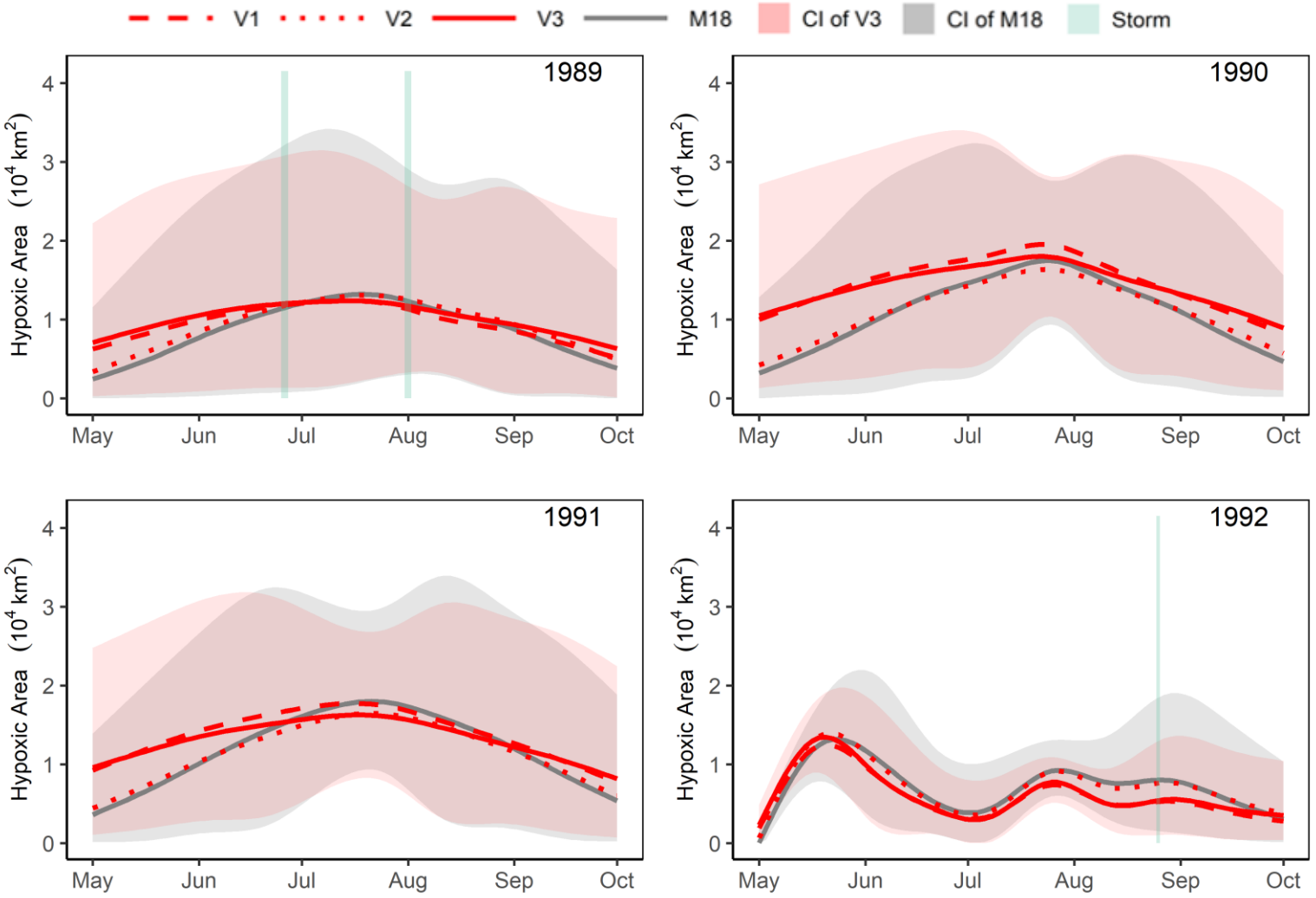


Figure S8B. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 1989-1992

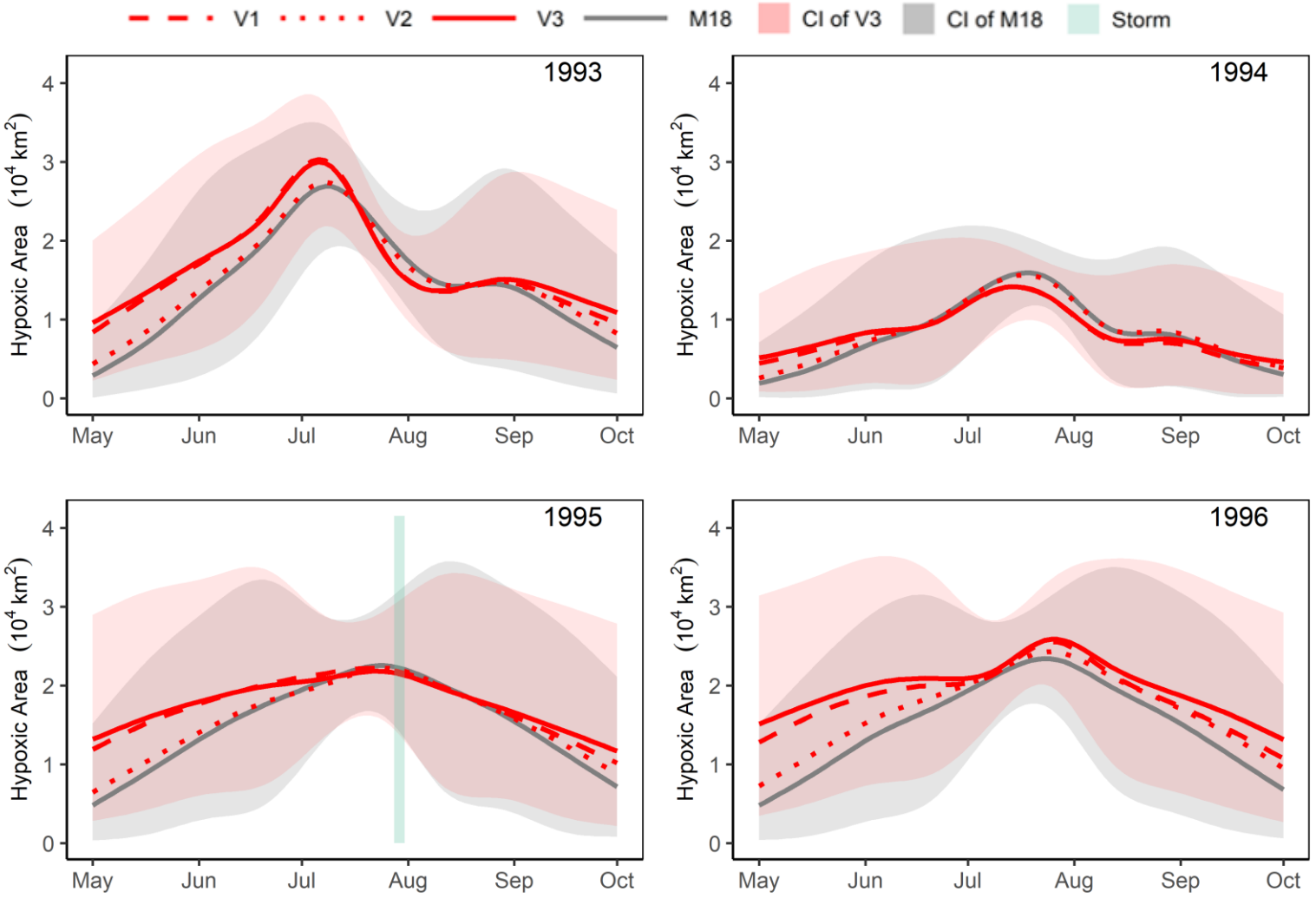


Figure S8C. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 1993-1996

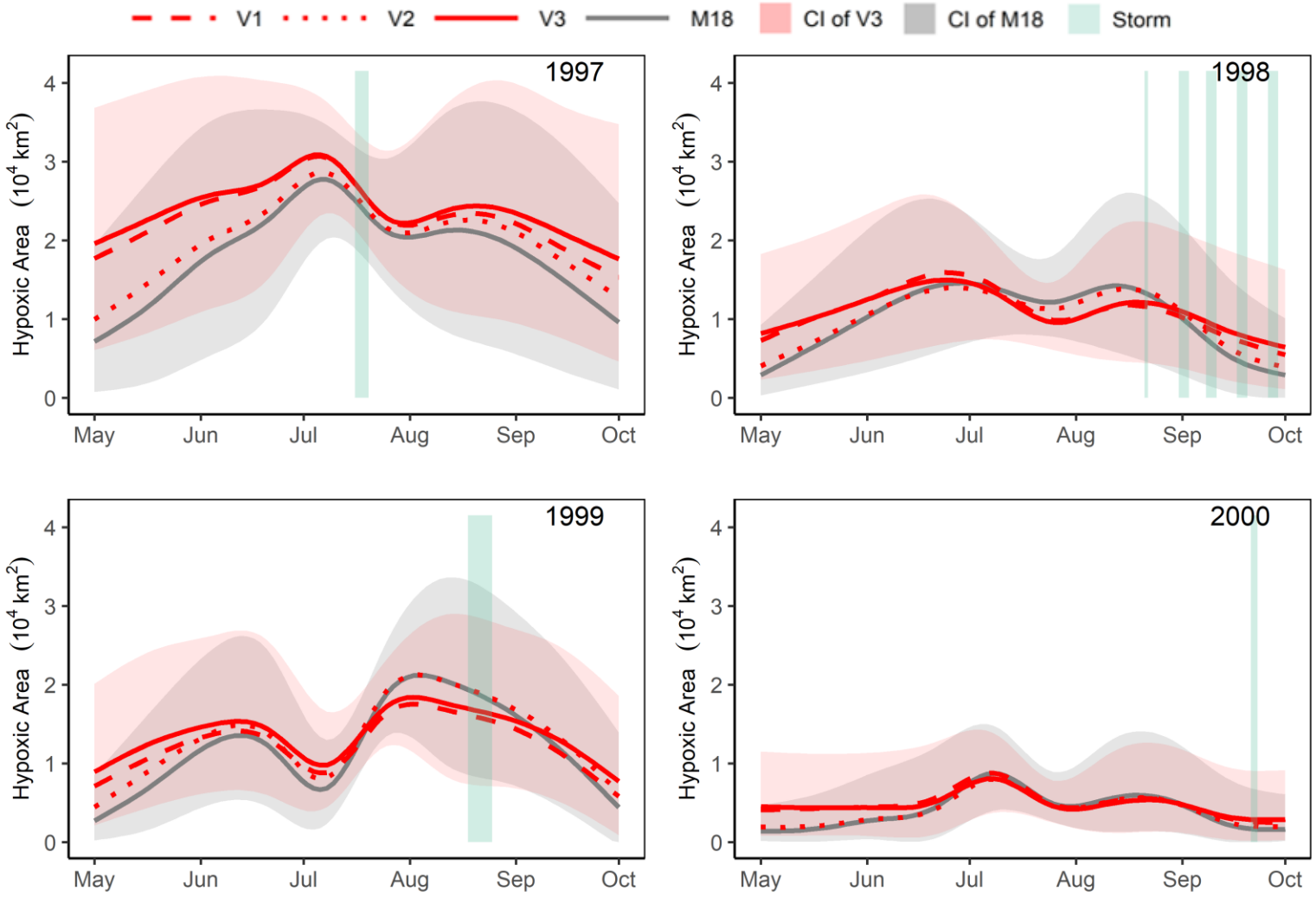


Figure S8D. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 1997-2000

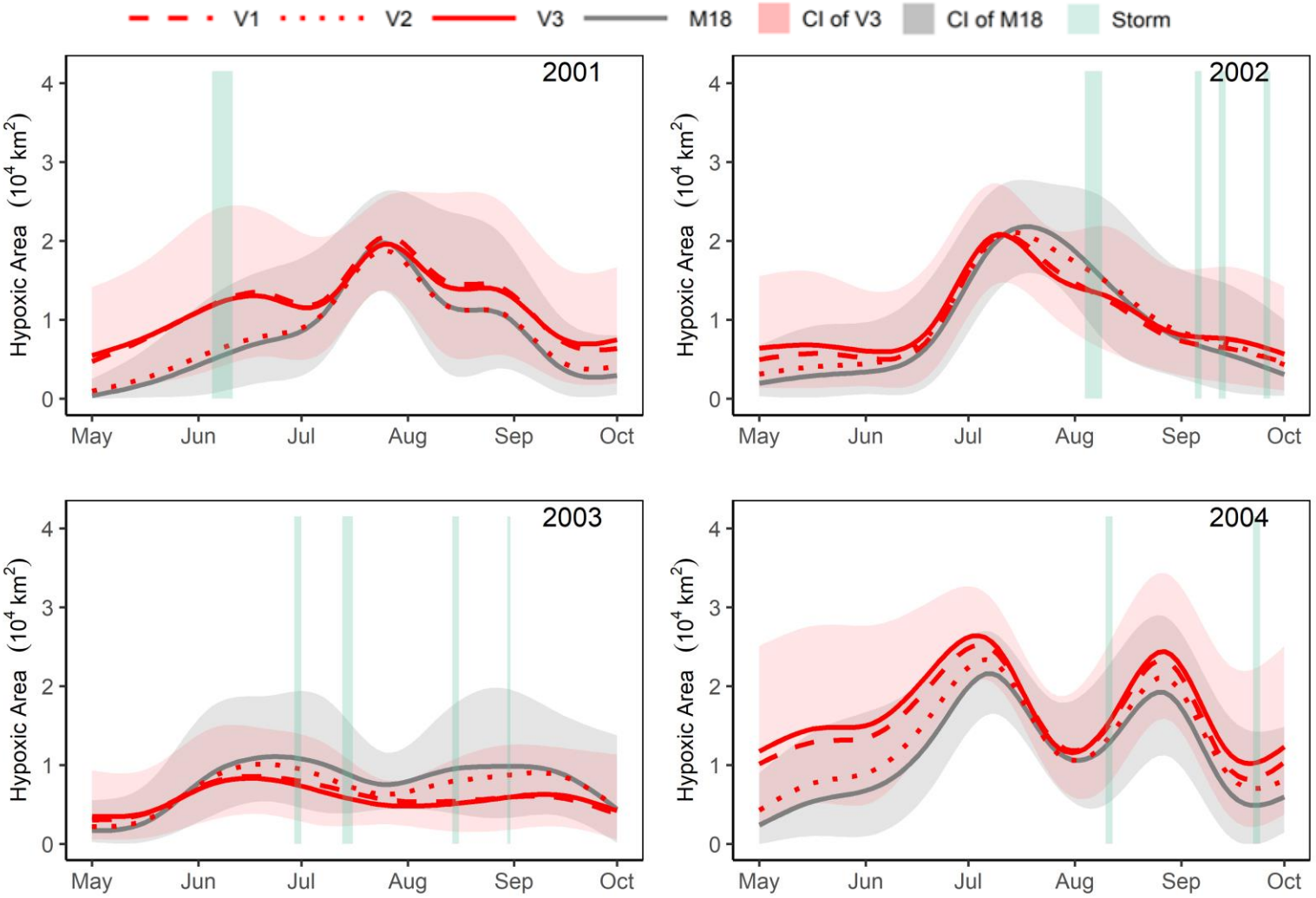


Figure S8E. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 2001-2004

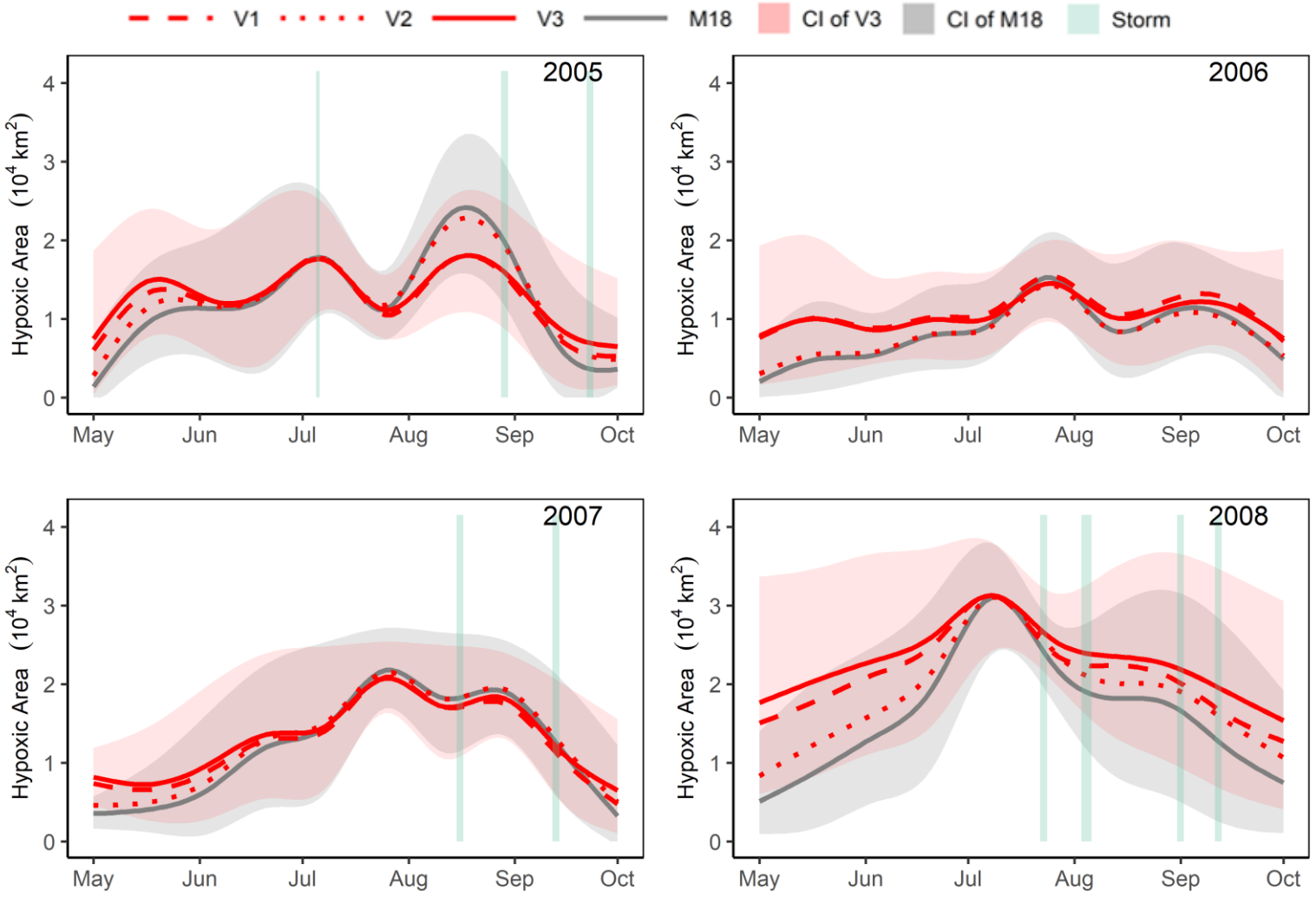


Figure S8F. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 2005-2008

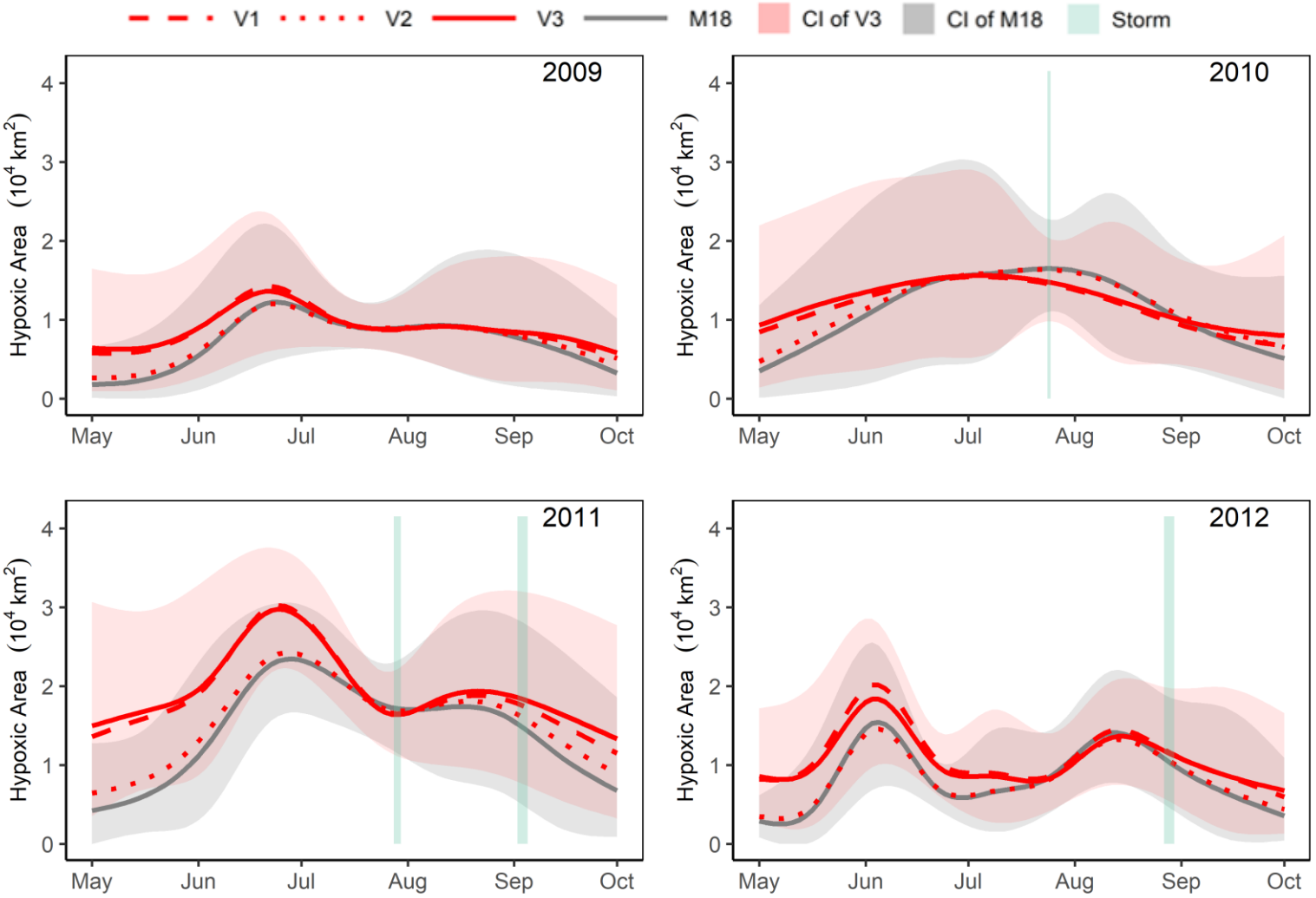


Figure S8G. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 2009-2012

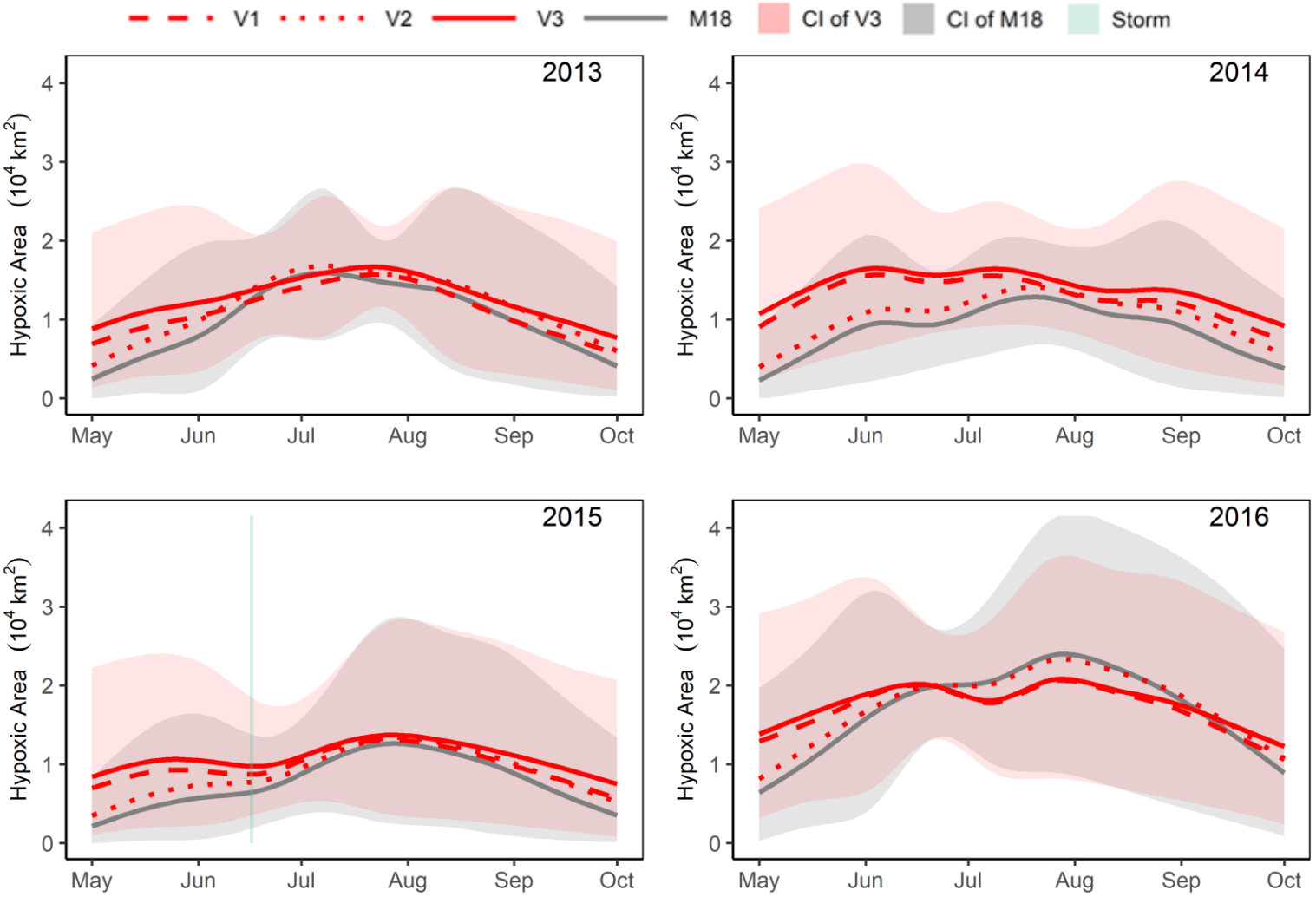


Figure S8H. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 2013-2016

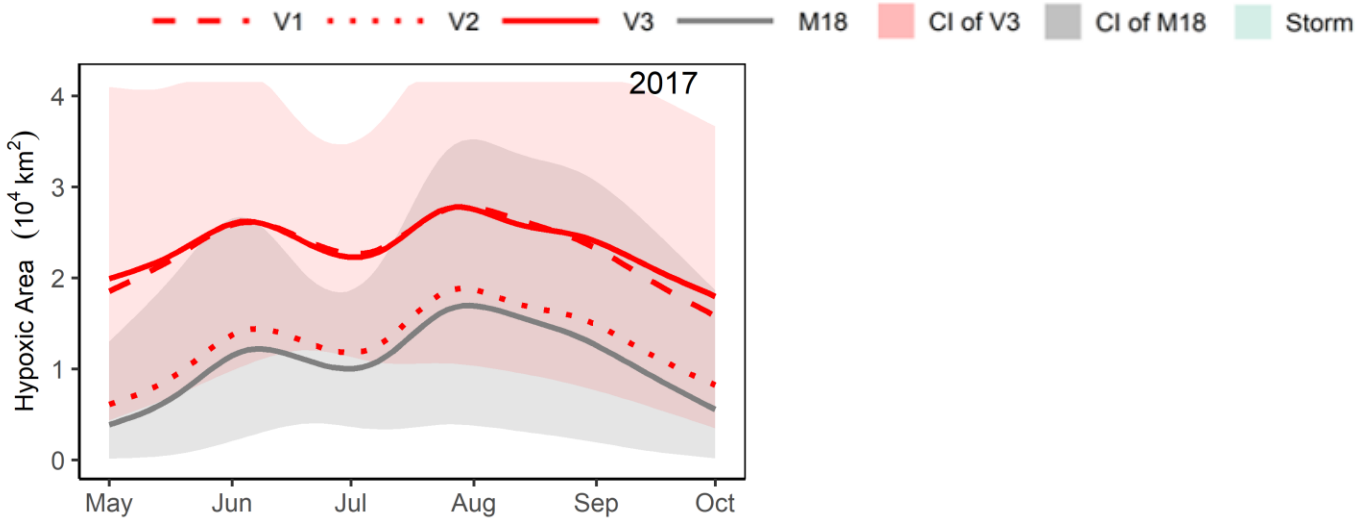


Figure S8I. Daily estimates (at 3-day intervals) of hypoxic extent and their 95% CI (confidence intervals) across the summer for 2017

SI-9: Coefficient of variations (C_v) of daily hypoxic area from each month of all models

Month	M18	V1	V2	V3
May	0.725	0.547	0.632	0.506
June	0.470	0.414	0.437	0.396
July	0.299	0.288	0.289	0.282
August	0.423	0.409	0.401	0.391
September	0.649	0.577	0.582	0.530

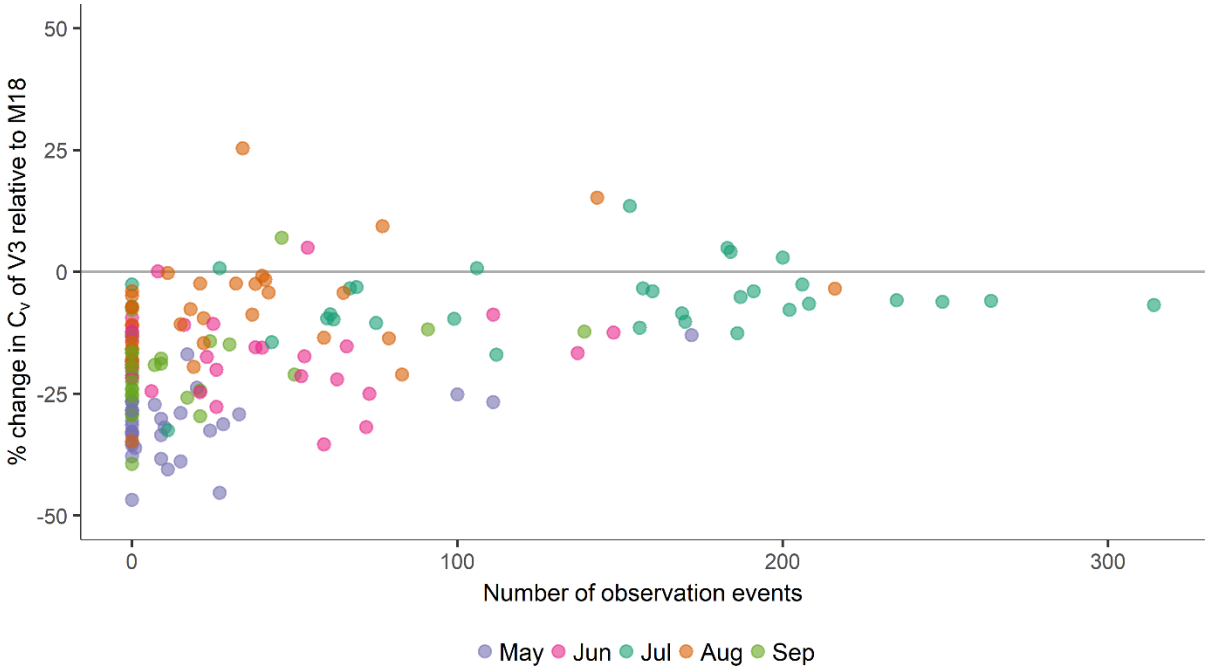


Figure S9. Change in C_v of V3 relative to M18 compared to number of observation events categorized by month (below zero indicates a reduction in uncertainty).

(change = {C_v of V3 – C_v of M18}/ C_v of M18)

SI-10: Comparisons of estimates and coefficients of variations of summer aggregates from V3 to M18

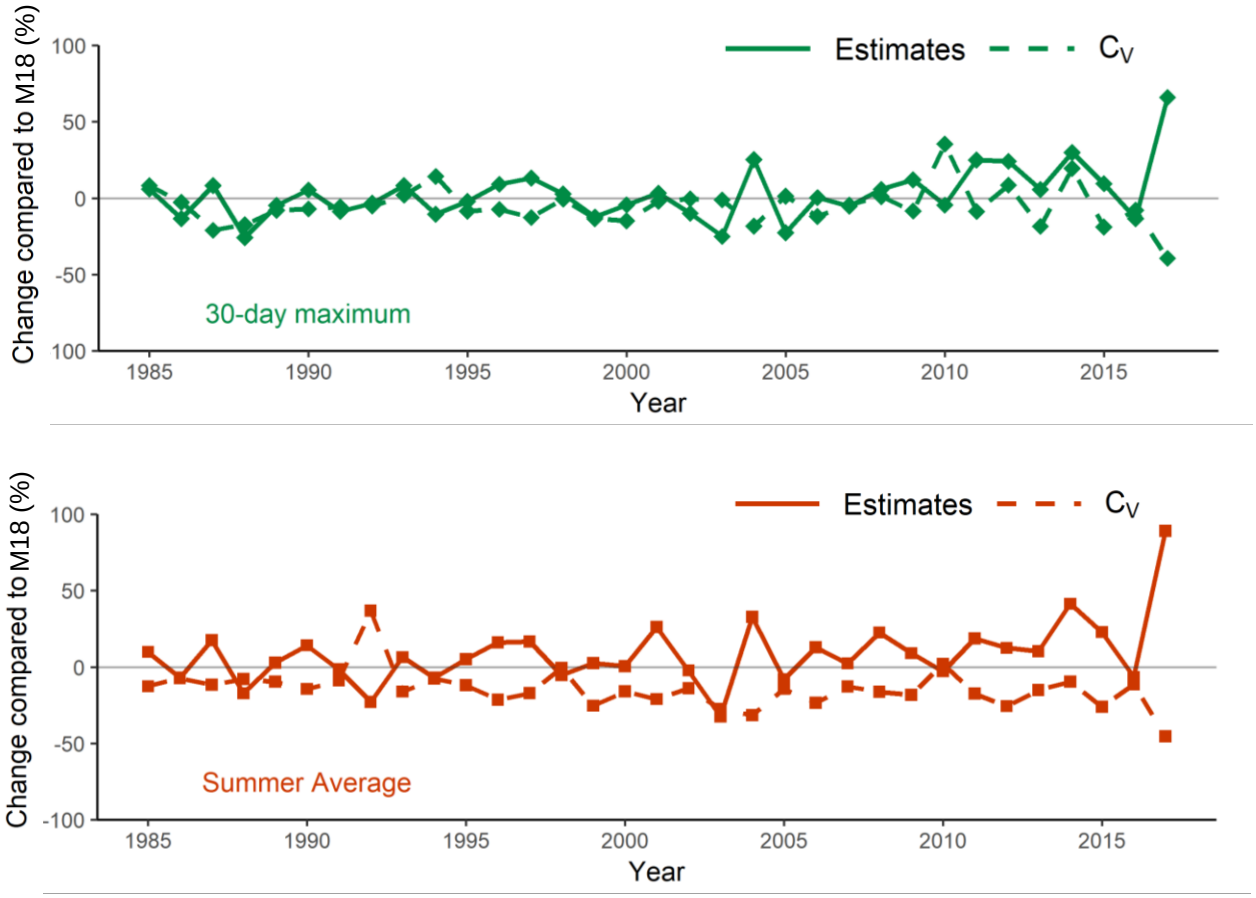


Figure S10. Changes in estimated values and uncertainty of 30-day maximum and summer average hypoxic extent from V3 compared to estimates from M18
(Change = {V3 – M18}/M18)