Water Quality Predictors for Harmful Algal Blooms Occurrence in the Mississippi Sound, USA

Vladimir J. Alarcon[1(✉)], Paul F. Mickle[1], and Ruben A. Alvarado[2]

[1] Northern Gulf Institute, 1021 Balch Blvd., Stennis Space Center, Mississippi 39529, USA
{valarcon,pmickle}@ngi.msstate.edu
[2] Universidad Católica de Salta, Salta, CP 4400, República Argentina

Abstract. From July 2019 to September 2019, harmful algal blooms (HABs) were observed in the State of Mississippi coast at the Northern Gulf of America (Mississippi Sound). In this paper, several multiple linear regression models (MLR models) based on observed water quality data and simulated water quality data are presented. Fourteen environmental and water quality observed variables and four model-simulated water quality variables were included in the study. Several combinations of those variables were used to generate MLR models which could help identify the main water quality and/or environmental predictors for HAB occurrence in the Mississippi Sound, during 2019.

Keywords: Mississippi Sound · Harmful Algal Blooms · Water Quality · Multiple Linear Regression · Microcystis

1 Introduction

The Mississippi Sound (MS Sound) is an estuarial water body located between the Mississippi Gulf Coast and the barrier islands extending from Lake Borgne (Louisiana) to Mobile Bay (Alabama) (Fig. 1). It is connected to Lake Pontchartrain (Louisiana) through a narrow natural channel [1]. The Bonnet Carré Spillway (BCS) is a flood control structure that diverts Mississippi River water to Lake Pontchartrain (in Louisiana) allowing floodwaters from the Mississippi River to flow into Lake Pontchartrain (Louisiana). The opening of the Bonne Carre Spillway has a strong effect on the water quality of the MS Sound.

Nutrient loadings from urban areas [2] affect water quality and biological species in the MS Sound, both of which support the economy of the State of Mississippi coastal area [3]. Nitrogen and phosphorus loading and their effects to the MS coastal waters should be studied to support sound management strategies [4].

In addition to contaminant inputs from urban areas, the MS Sound receives pollutant loads from Lake Pontchartrain, especially when the Bonnet Carré Spillway (BCS) is open. The BCS is a flood control structure which diverts waters from the Mississippi River to Lake Pontchartrain (Louisiana) when water stage at the river increases flood risk

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16761, pp. 542–556, 2027.
https://doi.org/10.1007/978-3-032-30527-5_35

in the city of New Orleans. Since the Lake is a shallow water body connected to the Gulf of America through the Rigolets strait (which allows bidirectional flow depending on the tidal regime) flushing of lake waters to the MS Sound can occur relatively fast depending on the magnitude of the inflow rate through the BCS. Alarcon et al. [1] report flushing times between 25 to 40 days. This means that pollutants present in Lake Pontchartrain before BCS openings may also be flushed out during BCS openings and could reach the MS Sound in 30 to 50 days. Over the past few decades, the impacts of BCS openings to adjacent estuaries have been scientifically investigated [5]. Nevertheless, observed water quality conditions in the MS Sound during HAB events may also be a factor for HAB presence.

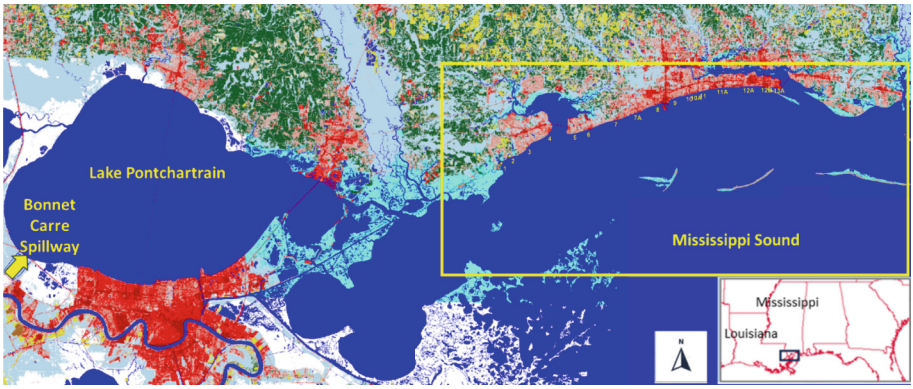

Fig. 1. Mississippi Sound (MS Sound). The MS Sound is located at the coast of the State of Mississippi (USA) (Northern Gulf of America).

In this research, a multiple linear regression model (MLR model) is developed based on observed and simulated water quality data. Fourteen environmental and water quality observed variables and five model-simulated water quality variables were included in the study. Several combinations of those variables were explored to identify the main water quality and/or environmental predictors for HAB occurrence in the Mississippi Sound.

2 Methods

2.1 Study Area

Figure 2 shows the study area including Mississippi Department of Environmental Quality sampling locations along the State of Mississippi coast. The figure also shows the predominant land use and land cover in the inland urban or undeveloped watersheds that drain to the MS Sound. As shown, dense urban areas located along the coast are potential sources of pollution. Rivers and minor streams in the inland area may also collect non-point source pollutants from agricultural areas (cultivated crops, pasture/hay) located outside in undeveloped areas, while stormwater outfalls from urban areas transport sediment loads and other contaminants to the MS Sound.

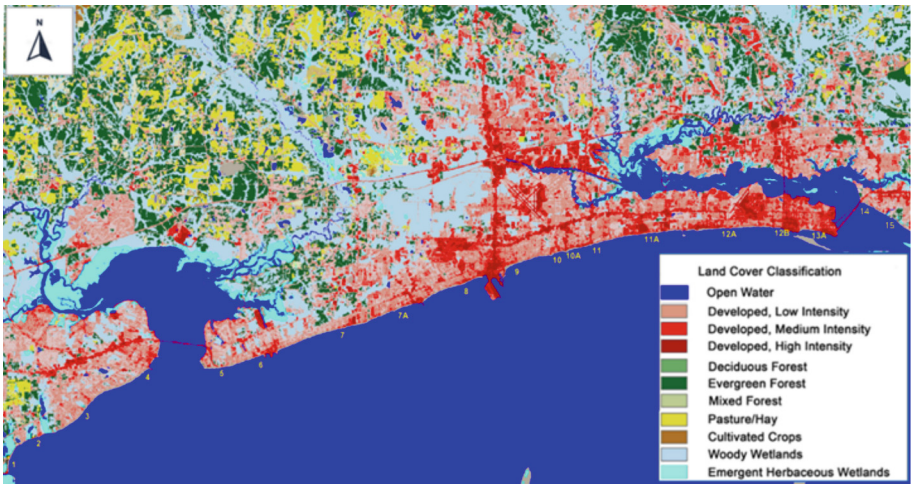

Fig. 2. Study Area. Land use and land cover in the Mississippi Sound. Land use/land cover in urban areas draining to the MS Sound are shown. MDEQ water quality sampling locations along the Mississippi coast are shown.

2.2 Observed and Simulated Water Quality Data

The Mississippi Department of Environmental Quality (MDEQ) performs routine water quality analyses in the Mississippi Sound. Under the MDEQ Mississippi Beach Monitoring Program, coastal water (at locations shown in Fig. 2) is monthly tested for the presence of nutrients, totals suspended solids (TSS), chlorophyl-a, fecal bacteria, and other water quality constituents. During the months of July, August, and September of 2019, MDEQ additionally analyzed for HAB. The species for which presence was mostly observed were: Microcystis spp. and Chattonella spp. Since Microcystis blooms were significantly more present than Chattonella during 2019, in this research Microcystis blooms were chosen as harmful algal blooms (HABs) indicator.

Observed Water Quality Data Used in the Study
Monthly observed Total Kjeldahl Nitrogen (TKN), Total Phosphorus (TP), Total Suspended Solids (TSS), salinity, Total Suspended Solids (TSS), Total Organic Carbon (TOC), chlorophyll-a, turbidity, TKN/TP ratio, and Temperature are used in the model development. The number of Microcystis bloom (BLOOM) sightings were monthly summed. Data from stations 1 through 11 (Fig. 2) were used for the development of the MLR models. Data from stations 12B, 14, and 15, located at the eastern-most border of the study area (Fig. 2) were used for assessing the predictive ability of the MLR models.

Simulated Water Quality Data used in the Study
A water quality model (described in [6]) was used to simulate Total Kjeldahl Nitrogen (TKNsim), Total Phosphorus (TPsim), and Total Nitrogen (TNsim). The ratios TKNsim/TPsim and TNsim/TPsim were also computed. Since the water quality model outputs are at hourly temporal resolution, the simulated data correspond exactly to the time of day in which HAB blooms were reported by MDEQ.

In addition to water quality data, tidal data at three locations along the MS sound were also used as input data: Biloxi, Gulfport, and Saint Joseph Island. Also, since the opening of the Bonnet Carre Spillway is a key factor in MS Sound HAB events, the inflow rate (Q in m3/s) was also input as predictor variable.

2.3 Multiple Linear Regression Model (MLR Model)

A multiple linear regression model can be represented by the following equation [7].

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \ldots + \beta_k x_{i,k} + \varepsilon_i.$$

In this research, y_i represents predicted number of HAB sightings per month, x_i are the predictor water quality variables included in the model, k is the number of predictor water quality variables, and the β_k coefficients are the values that minimize the sum of squared errors for the sample. ε_i are the residuals which are assumed to have a normal distribution with mean 0 and constant variance σ^2. In this research, several combinations of variables were performed to achieve the best statistical fit when comparing observed HAB sightings versus predicted HAB sightings. The statistical fit was measured calculating the r^2 coefficient of determination for each combination of variables included in the analysis.

Potential predictor water quality variables (observed and simulated) were organized in 3 batches: a) observed and simulated water quality data, b) only observed water quality data, c) only simulated water quality data. Each of these batches were used for generating MLR models with 2, 3, 4, and 5, minimum number of predictor variables. Therefore, 12 linear models were developed and assessed for statistical fit.

3 Results

3.1 Microsystis sp. and Chattonella sp. in the MS Sound

Figure 3 shows observed monthly blooms sightings for Microcystis sp. and Chattonella sp. The figure includes number of monthly bloom sightings for July, August, and September 2019. For brevity, the most representative locations are shown: Lakeshore (station 1), Bay Saint Louis (station 4), Jourdan River (station 4A), Long Beach (station 8), and Edgewater (station 11A). Microcistys sp. is shown to have a more pervasive presence than Chattonella sp. during all months, therefore this research focuses primarily on Microcystis blooms. Figure 3 shows that Bay Saint Louis (station 4) was the location where the largest number of blooms were recorded during July and August, 2019. The second 2019-opening of the Bonnet Carre Spillway took place from May 10 to July 27th, 2019. Nevertheless, peak inflow (4559 m^3/s) took place on May 22. According to Alarcon et al. [1], for an inflow of this magnitude Lake Pontchartrain flushing time is 33 days. Therefore, HAB present in Lake Pontchartrain may have reached the Northern Gulf of America by late-June 2019. Since Lakeshore and Bay Saint Louis are the locations closest to the Rigolets (Lake Pontchartrain's outfall to the gulf), if HAB would have been transported from the lake, HAB should have been observed in July 2019. This is clearly illustrated in Fig. 3: 13 bloom sightings in Lakeshore, and 15 sightings

in Bay Saint Louis were recorded during July 2019. Since Bay Saint Louis has more urban areas than Lakeshore, nutrients from those areas may have enhanced blooms in the area. This is consistent with the number of blooms observed in Edgewater and Long beach during July 2019; those locations also have large coastal urban areas. In station 4A (Jourdan River outfall) 10 bloom sightings during July were also reported. Nutrients from agricultural areas in the river's watershed together with hydrodynamically transported Microcystis may explain those sightings. Interestingly, during August 2019, Bay Saint Louis and Jourdan River are the locations with the greatest number of bloom sightings. The rest of the stations still recorded at least 7 HAB sightings in August. During September, the closest location to Lake Pontchartrain (Lakeshore) showed the greatest number of HAB sightings, which indicates potential HAB migration or remanent HAB present in the lake.

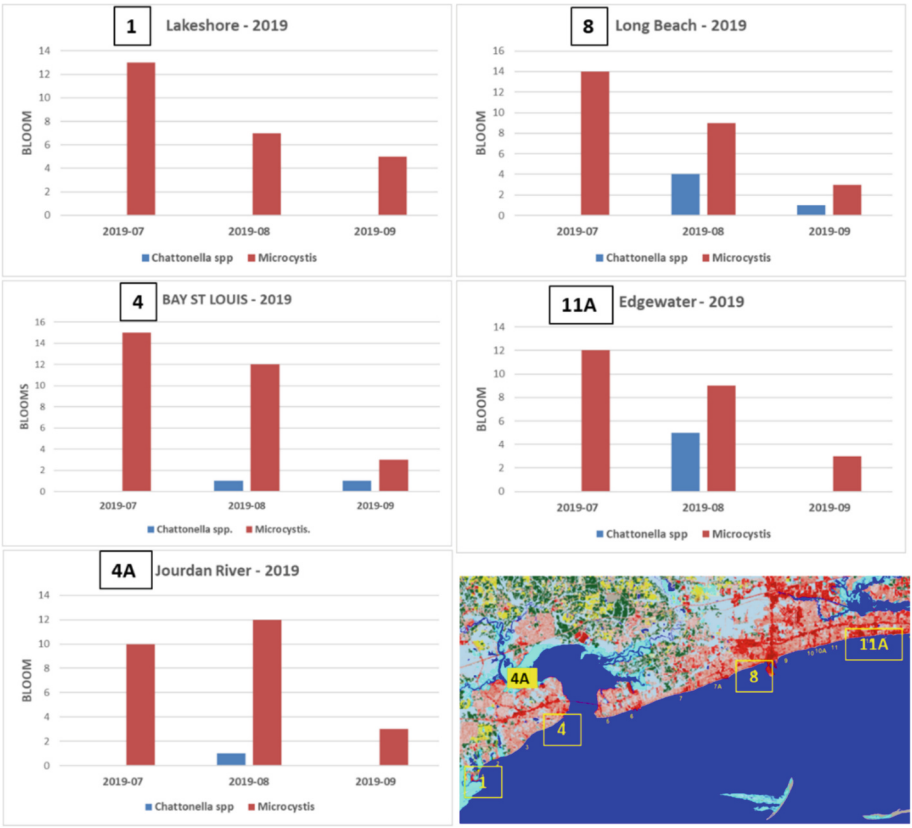

Fig. 3. Harmful algal bloom sightings during July, August, and September 2019.

Usually, HAB presence is related to nutrient enrichment and chlorophyll-a is a standard HAB indicator. To explore the relationship between nutrient-related observed water quality constituents, total nitrogen (TN), total phosphorus (TP), chlorophyll-a (Chl-a),

and total organic carbon (TOC) were scatter-plotted. As shown in Fig. 4, Chl-a strongly correlates with TN, TP, and TOC.

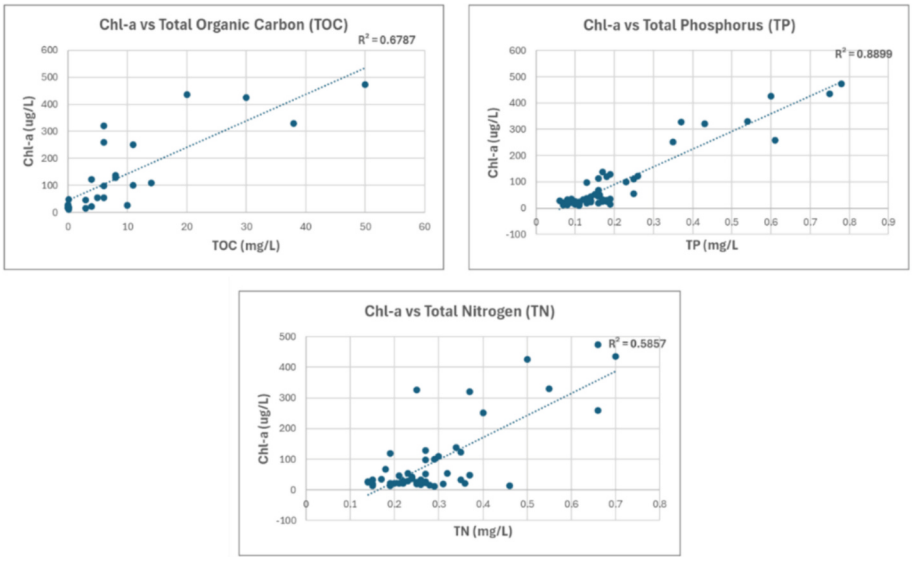

Fig. 4. Relationship between total nitrogen (TN), total phosphorus (TP), and chlorophyll-a (Chl-a) in the MS Sound.

Nevertheless, none of the nutrient constituents (TN or TP) and Total Organic Carbon (TOC) showed a strong correlation with HAB presence (Fig. 5).

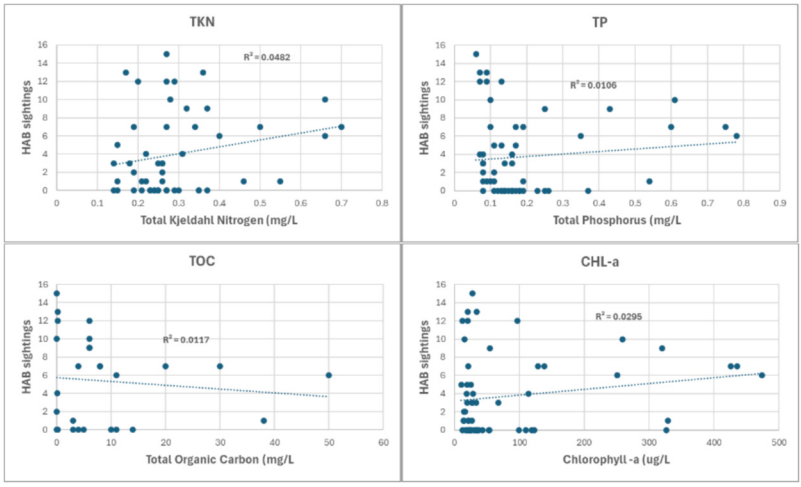

Fig. 5. HAB sightings versus nutrients (TN, TP), and Chlorophyll-a (CHL-a) concentrations.

As shown in Fig. 6, salinity was the only water quality variable that showed moderate correlation ($R^2 = 0.51$). While nutrients or chlorophyll-a correlated poorly with HAB presence, the ratio TKN/TP was shown to have a moderate correlation ($R^2 = 0.55$).

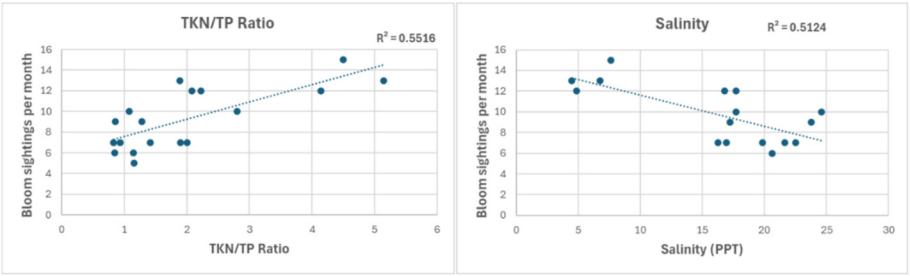

Fig. 6. Relationship between salinity and TKN/TP ratio with HAB presence.in the MS Sound.

Since one-to-one relationships are shown to be not evident, combining water quality variables to predict HAB seems to be necessary. These results reinforce the objective of this research.

3.2 Multiple Linear Regression Models' Results

As described earlier, for the development of MLR models the data was organized in three groups, each of which was tested for 2, 3, 4, and 5 predictor variables: a) observed and simulated water quality data, b) only observed water quality data, c) only simulated water quality data. Nevertheless, the MLR models for which the R^2 value was smaller than 0.5 were considered unsatisfactory, and the results are not reported below. The results for acceptable MLR models are as follows.

MLR Models with Observed and Simulated Predictor Variables
Figure 7 shows that MLR models that mix input data (observed and simulated) as predictor variables to estimate HAB presence have good statistical fit between observed HAB data and predicted HAB data. All R^2 indicators are above 0.57 and the Root Mean Square errors (RMS) range from 2.21. to 2.75 (Table 1).

MLR Models with only Observed Predictor Variables
Figure 8 shows that using only observed data to predict HAB presence slightly decreases the quality of the statistical fit between predicted HAB and observed HAB. Nevertheless, all R^2 values range between 0.53 and 0.66, and RMS values ranged from 2.59 to 2.94.

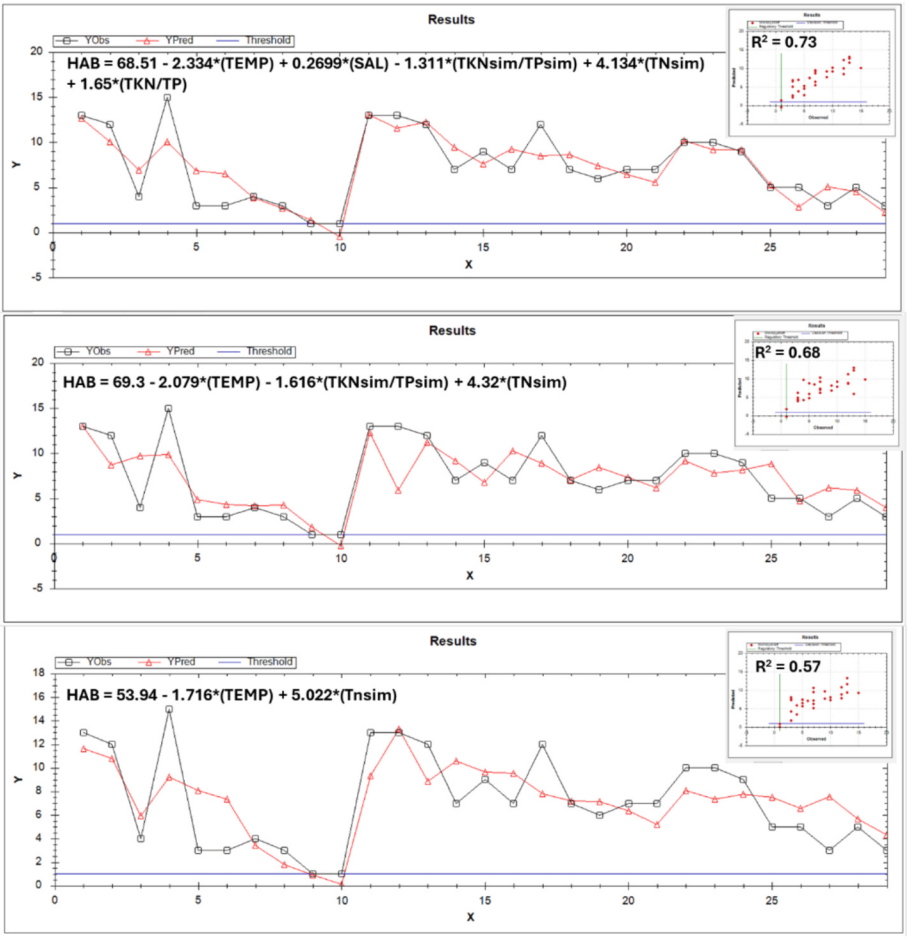

Fig. 7. MLR models generated with observed and simulated water quality data. The resulting models include 5, 4, and 3 predictor variables. Resulting linear model equations are shown in each chart, including the response variable (HAB) and the predictor variables.

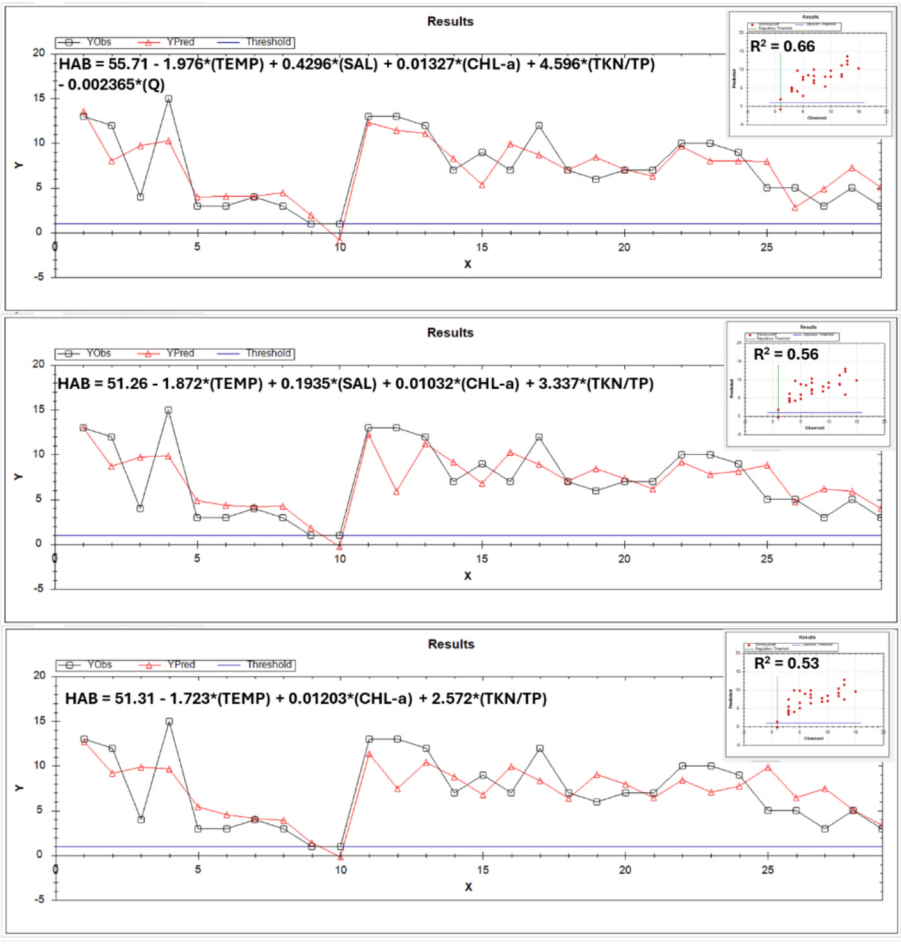

Fig. 8. MLR models generated with only observed water quality data. The resulting models include 5, 4, and 3 predictor variables. Resulting linear model equations are shown in each chart, including the response variable (HAB) and the predictor variables.

Adjusted R^2 values (Table 1) suggest that models 1 through 4 provide better quality fit between observed and simulated response variable (Adj. $R^2 > 0.63$). The adjusted Akaike criterion indicates that models 1, 2, and 4 seem to be better candidates for HAB prediction. The Bayesian criterion suggests that among the models that used simulated and observed variables, models 1 and 2 are more parsimonious than model 3. Among the models that used only observed variables, model 4 is shown to be the most parsimonious.

Table 1. Summary of indicators or fit for MLR models.

Model	Number of predictor variables	R^2	Adj. R^2	RMS	Adj. Akaike	Bayes. C.
1	2 simulated, 3 observed	0.73	0.73	2.21	80.33	56.16
2	2 simulated, 1 observed	0.68	0.65	2.42	84.05	57.16
3	1 simulated, 1 observed	0.57	0.54	2.75	90.49	62.22
4	5 observed	0.66	0.63	2.59	89.65	65.49
5	4 observed	0.56	0.52	2.89	95.11	69.59
6	3 observed	0.53	0.49	2.94	95.21	68.31

Table 1 shows that mixing simulated and observed predictor variables may provide MLR models which generate better quality HAB predictions than using only observed predictor variables. Nevertheless, recent research on HAB prediction shows similar statistical fit values as those shown in Table 1. Wen et al. [8] used machine learning to predict Alexandrium sp. algal blooms in the Gulf of Maine (USA). In their study, the following predictor variables were selected: PO4, NH4, Si(OH)4, NO2 + NO3, and Chl-a. The R^2 values reported in [8] range between 0.53 to 0.83. Yu et al. [9] proposed a method based on machine learning to predict the occurrence of Lingulodinium polyedrum algal blooms in the Southern California coastal ocean, and Sishili bay (Shandong province, China). Yu et al. [9] found that for California coastal ocean the main predictor variables were ammonia, chlorophyll-a and temperature. For Sishili bay (China) Yu et al. [9] selected DO, Ammonia, chlorophyll-a, and nitrite. The resulting R^2 in the study values range between 0.501 and 0.956 [9]. Therefore, the statistical fit results achieved in this research and shown in Table 1 are within ranges reported in the literature.

Table 2 shows additional statistics for the resulting MLR models. All models include water temperature (TEM)P as primary predictor variable as shown by the corresponding t-statistic and P-value. Salinity (SAL) is a predictor variable in models 1, 4 and 5. Nevertheless, Although Model 5 is shown to have a decent R2 value (see Table1), salinity's t-statistic is smaller than 2.0, and its P-value is 0.18 (the highest for all models). Also, in Model 5, predictor variable chlorophyll-a (CHL-a) has P-value = 0.0379 and t-statistic = 2.1976, which makes Model 5 the weakest of all MLR models. On the other hand, all the other models have |t-statistic| > 2.47 and P-value < 0.022, which, together with the R2 and RMS values shown in Table 1, makes them valid representation of the observed data.

Multiple Linear Regression (MLR) models have been used for HAB prediction in lakes and coastal waters, as well as water quality and/or biological species prediction. MLR models (as opposed to other prediction methods), aim at predicting continuous outcomes, based on independent variables which may include various water quality parameters, hydrological parameters, and meteorological parameters [10].

Table 2. Resulting MLR models.

	Predictor variables	Coeff.	Std. Err.	t-Stat.	P-Value	Std.Coeff.
Model 1	Intercept	68.509	9.081	7.5443	1.16E-07	
	TEMP	−2.3337	0.3019	−7.7288	7.71E−08	−1.0039
	SAL	0.2699	0.1091	2.4727	0.0212	0.4342
	TKNsim/TPsim	−1.3111	0.517	−2.5362	0.0185	−0.3194
	TNsim	4.1341	1.1285	3.6634	0.0013	0.5546
	TKN/TP	1.6496	0.6599	2.4998	0.02	0.5011
Model 2	Intercept	69.299	9.7248	7.1261	1.81E−07	
	TEMP	−2.0788	0.3126	−6.6509	5.72E−07	−0.8942
	TKNsim/TPsim	−1.6158	0.5457	−2.9608	0.0066	−0.3936
	TNsim	4.3199	0.952	4.5378	0.0001	0.5795
Model 3	Intercept	53.936	9.3331	5.779	4.36E−06	
	TEMP	−1.7159	0.3263	−5.2592	1.70E−05	−0.7382
	TNsim	5.0223	1.0462	4.8006	5.69E−05	0.6738
Model 4	Intercept	55.709	8.9929	6.1948	2.55E−06	
	TEMP	−1.9759	0.3325	−5.9425	4.66E−06	−0.85
	SAL	0.4296	0.155	2.7716	0.0109	0.6912
	CHL-a	0.0133	0.0044	3.0456	0.0057	0.4805
	Q	−0.0024	0.0009	−2.654	0.0142	−0.4177
	TKN/TP	4.5964	0.8755	5.2499	2.52E−05	1.3963
Model 5	Intercept	51.263	9.8458	5.2067	2.46E−05	
	TEMP	−1.8717	0.3679	−5.0871	3.33E−05	−0.8052
	SAL	0.1935	0.1414	1.3679	0.184	0.3113
	CHL-a	0.0103	0.0047	2.1976	0.0379	0.3735
	TKN/TP	3.337	0.8199	4.0699	0.0004	1.0137
Model 6	Intercept	51.306	10.0093	5.1258	2.69E−05	
	TEMP	−1.723	0.3574	−4.8216	5.92E−05	−0.7412
	CHL-a	0.012	0.0046	2.616	0.0149	0.4356
	TKN/TP	2.5717	0.6093	4.2204	0.0003	0.7812

Table 3 summarizes the MRL models equations generated in this research. From a water quality management perspective, MLR models are a straight-forward and simple method which could be used by managers. Nevertheless, a model with many predictor variables may be costly to operate, especially if those predictor variables depend on field campaigns and laboratory analysis. An MLR model could also be computationally and/or financially expensive if the model relies on simulated data from other models (e.g.,

[11–13]. For example, MLR models 1 and 2 seem to be the best models because of their statistical performance (see Tables 1, 2 and 3). However, Model 1, besides depending on observed TEMP and SAL (which are relatively inexpensive to measure), also depends on observed TKN and TP, as well as simulated TKN, TP and TN (TKNsim, TPsim, TNsim), which makes Model 1 a costly alternative. Models 2 and 3, on the other hand, seem to be more implementable, although their dependance on simulated data might is a drawback.

Table 3. Summarized MLR models' equations

Model	MLR models using observed and simulated predictor variables
1	HAB = 68.51 – 2.334*(TEMP) + 0.2699*(SAL) - 1.311*(TKNsim/TPsim) + 4.134*(TNsim) + 1.65*(TKN/TP)
2	HAB = 69.3 – 2.079*(TEMP) – 1.616*(TKNsim/TPsim) + 4.32*(TNsim)
3	HAB = 53.94 – 1.716*(TEMP) + 5.022*(TNsim)
	MLR models using observed and simulated predictor variables
4	HAB = 55.71 – 1.976*(TEMP) + 0.4296*(SAL) + 0.01327*(CHL-a) + 4.596*(TKN/TP) – 0.002365*(Q)
5	HAB = 51.26 – 1.872*(TEMP) + 0.1935*(SAL) + 0.01032*(CHL-a) + 3.337*(TKN/TP)
6	HAB = 51.31 – 1.723*(TEMP) + 0.01203*(CHL-a) + 2.572*(TKN/TP)

From the MLR models based on observed data, Model 4 has the greatest R^2 (Table 1) and the t-statistic and P-value for the predictor variables are optimal. Variables TEMP, SAL, TKN and TP, are routinely measured by MDEQ and USGS. CHL-a could be acquired from public-domain NASA imagery [14]. Bonnet Carre inflow rate (Q) can be provided by USACE during spillway openings. Nevertheless, its actual predictive ability must be assessed.

3.3 Using the Developed MLR Models to Predict HAB Sightings

As described earlier, water quality data from stations 1 through 11 were used to generate the MLR models. To assess the predictive ability of the models, water quality data from stations 12B (Biloxi East Central), 14 (Frontbeach), and 15 (Shearwater) were used. Figure 9 shows HAB sighting predictions corresponding to models 4, 5, and 6.

Model 4 is shown to provide limited HAB prediction capability. Although the decreasing trend shown by observed data is replicated, Model 4 overestimates HAB sightings for August and July 2019. Nevertheless, HAB sightings equal or below 10 are predicted acceptably. Models 5 and 6 successfully capture the trend and magnitude of HAB sightings throughout the observation period. The R^2 for these models is also better than the R^2 corresponding to Model 4. Model 6 captures better the magnitude of the HAB sightings which is reflected on its R^2 (0.54). The three models underpredict one HAB sightings for September 2019, being one of those predictions negative. Nevertheless, the occurrence of this false negative can be avoided in future research using

shrinkage methods that apply a penalty term to the loss function embedded in the linear regression to avoid over- and underfitting [15, 16].

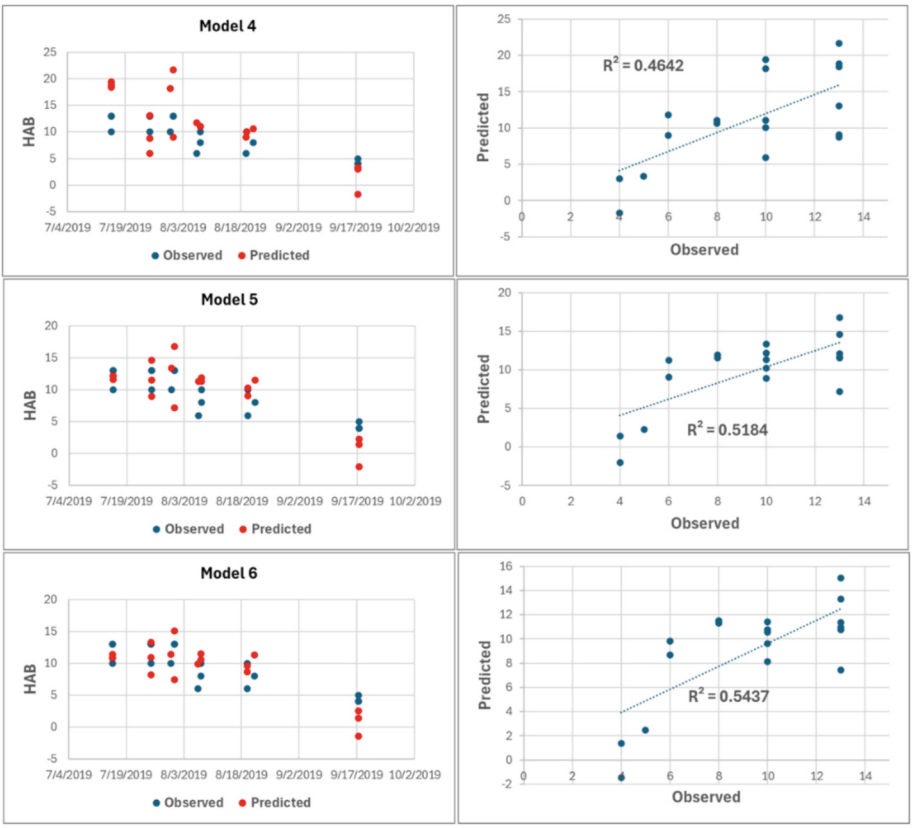

Fig. 9. Prediction of HAB sightings. Model 6 is shown to be the best predictor of HAB sightings.

Between the models that were developed using only observed data, Model 4 showed the best statistical fit (Fig. 8). However, its predictive ability is shown to be limited when compared to models 5 and 6 (Fig. 9). This seems to be a typical case of overfitting. Overfitting may lead to poor generalization of new data [17] because increasing input dimensionality does not necessarily guarantee improved predicting performance [18][19], as shown in Fig. 9.

There have been several modeling approaches to predict HAB which have relied on conventional water quality indicators, such as temperature, and nutrient concentrations but they use chlorophyll-a as the primary response variable [15, 20, 21]. The models presented in this paper predict HAB sightings, which are a closer surrogate to actual HAB concentrations. All the equations shown in Table 3 identify temperature, chlorophyl-a, and nutrients as predictor variables which coincides with previous research on this topic [22–26].

4 Conclusions

Six multiple linear regression (MRL) models based on observed water quality data and simulated water quality data are presented. Fourteen environmental and water quality observed variables (salinity, temperature, chl-a, turbidity, TP, TN, TKN, TOC, TSS, TKN/TP, Biloxi tides, Gulfport tides, St. Joseph tides, Q at Bonnet Carre) and four model-simulated water quality variables (TPsim, TNsim, TKNsim, TKNsim/TPsim) were included in the study. Several combinations of those variables were used to develop the MLR models presented in this paper. The models using only observed data were assessed in their predictive ability of HAB sightings since models dependent on simulated data are costly to be used for water quality management. Model 6, which predictor variables are water temperature, chlorophyll-a, and TKN/TP ratio, is shown to have the best predictive ability.

Acknowledgements. This research was funded by grant: MS GOMESA 2022 MS-363507.

References

1. Alarcon, V.J., Mickle, P.F., Alvarado, R.A.: A Hydrodynamic model of the Mississippi sound. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15890, pp. 281–294. Springer, Cham. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-97606-3_19
2. 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 **15**(10), 3121–3131 (2018). https://doi.org/10.5194/bg-15-3121-2018
3. Sankar, M.S., et al.: Seasonal changes of trace elements, nutrients, dissolved organic matter, and coastal acidification over the largest oyster reef in the western Mississippi sound, USA. Environ. Monitor. Assess. **195**(1) (2023). https://doi.org/10.1007/s10661-022-10719-z
4. Camacho, R.A., Martin, J.L., Watson, B., Paul, M.J., Zheng, L., J.B.: Stribling, modeling the factors controlling phytoplankton in the St. Louis Bay estuary, Mississippi and evaluating estuarine responses to nutrient load modifications. J. Environ. Eng. **141**(3) (2015). https://doi.org/10.1061/(ASCE)EE.1943-7870.0000892
5. Beck, R., Xu, M., Zhan, S., et al.: Comparison of satellite reflectance algorithms for estimating turbidity and cyanobacterial concentrations in productive freshwaters using hyperspectral aircraft imagery and dense coincident surface observations. J. Great Lakes Res. **45**(3), 413–433 (2019). https://doi.org/10.1016/j.jglr.2018.09.001
6. Alarcon and Mickle 2025: A Conceptual Model for Water Quality Simulation in the Mississippi Sound, USA. State of the Coast Conference, New Orleans (2025). https://www.stateofthecoast.org/wp-content/uploads/2025/05/2025_-Session-24_-Vladimir-Alarcon.pdf
7. Pardoe, I.: Applied Regression Modeling. Wiley, New York (2012). https://doi.org/10.1002/9781118345054
8. Wen, J., Yang, J., Li, Y., Gao, L.: Harmful algal bloom warning based on machine learning in maritime site monitoring. Knowl.-Based Syst. **245**, 108569 (2022). https://doi.org/10.1016/j.scitotenv.2024.173546
9. Yu, P., Gao, R., Zhang, D., Liu, Z.-P.: Predicting coastal algal blooms with environmental factors by machine learning methods. Ecol. Indic. **123**, 107334 (2021). https://doi.org/10.1016/j.ecolind.2020.107334

10. Park, J., Patel, K., Hyong Lee, W.: Recent advances in algal bloom detection and prediction technology using machine learning. Sci. Tot. Environ. **938**, 173546 (2024). https://doi.org/10.1016/j.scitotenv.2024.173546
11. Alarcon, V.J., Johnson, D., McAnally, W.H., et al.: Nested hydrodynamic modeling of a coastal river applying dynamic-coupling. Water Resour. Manag. **28**, 3227–3240 (2014). https://doi.org/10.1007/s11269-014-0671-6
12. Alarcon, V.J., et al.: Coastal inundation under concurrent mean and extreme sea-level rise in Coral Gables. Nat. Hazards **2022**, 1 (2022). https://doi.org/10.1007/s11069-021-05163-0
13. Alarcon, V.J., Linhoss, A., Kelble, C., Mickle, P.: Potential challenges for the restoration of Biscayne Bay (Florida, USA) in the face of climate change effects revealed with predictive models. In: Elsevier (ed.) Ocean & Coastal Management, vol. 247, pp. 25–33. Elsevier, Amsterdam (2024). https://doi.org/10.1016/j.ocecoaman.2023.106929
14. Alarcon, V.J., van der Zwaag, J., Moorhead, R.: Estimation of estuary phytoplankton using a web-based tool for visualization of hyper-spectral images. In: 35th IEEE Applied Imagery and Pattern Recognition Workshop (AIPR'06), Washington, DC, USA, pp. 25–25 (2006). https://doi.org/10.1109/AIPR.2006.22
15. Izadi, M., Sultan, M., Kadiri, R.E., Ghannadi, A., Abdelmohsen, K.: A remote sensing and machine learning-based approach to forecast the onset of harmful algal bloom. Remote Sens. **13**(19), 3863 (2021). https://doi.org/10.3390/rs13193863
16. James, G., Witten, D., Hastie, T., Tibshirani, R.: An Introduction to Statistical Learning. Springer, New York (2000). ISBN 978-1-4614-7137-0
17. Yan, Z., Kamanmalek, S., Alamdari, N.: Predicting coastal harmful algal blooms using integrated data-driven analysis of environmental factors. Sci. Total Environ. **912**(2024), 169253 (2023). https://doi.org/10.1016/j.scitotenv.2023.169253
18. Lee, Y.K., Oh, H., Leeb, B.-M., Hur, J.: Deciphering and predicting algal bloom variability using size-fractionated organic matter and machine learning in a complex watershed. Ecol. Indicat. **178**, 114067 (2025). https://doi.org/10.1016/j.ecolind.2025.114067
19. Acuña-Alonso, C., Barba-Barragáns, D., Martínez-Sánchez, J., Álvarez, X.: Predictive mapping of harmful algal blooms with Sentinel-2 imagery and random forest modelling. Water Secur. **27**(2026), 100194 (2026). https://doi.org/10.1016/j.wasec.2025.100194
20. Lee, D., Kim, M., Lee, B., Chae, S., Kwon, S., Kang, S.: Integrated explainable deep learning prediction of harmful algal blooms Technol. Forecast. Soc. Chang. **185**, 122046 (2022)
21. Mermer, O., Zhang, E., Demir, I.: A comparative study of ensemble machine learning and explainable AI for predicting harmful algal blooms. Big Data Cogn. Comput. **9**(2025), 138 (2025)
22. Goldman, J.C., Carpenter, E.J.: A kinetic approach to the effect of temperature on algal growth. Limnol. Oceanogr. **19**, 756–766 (1974)
23. Hallegraeff, G.M.: Ocean climate change, phytoplankton community responses, and harmful algal blooms: a formidable predictive challenge. J. Phycol. **46**, 220–235 (2010)
24. Bricaud, A., Bosc, E., Antoine, D.: Algal biomass and sea surface temperature in the Mediterranean Basin: intercomparison of data from various satellite sensors, and implications for primary production estimates. Remote Sens. Environ. **81**, 163–178 (2002)
25. Errera, R.M., Yvon-Lewis, S., Kessler, J.D., Campbell, L.: Reponses of the dinoflagellate Karenia brevis to climate change: PCO2 and sea surface temperatures. Harmful Algae **37**, 110–116 (2014)
26. Sarma, Y.V.B., Al-Hashmi, K., Smith, S.L.: Sea surface warming and its implications for harmful algal blooms off Oman. Int. J. Mar. Sci. **3**, 65–71 (2013)