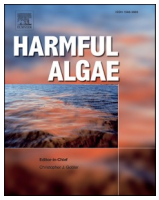
Harmful Algae 139 (2024) 102729

Contents lists available at ScienceDirect

Harmful Algae

journal homepage: www.elsevier.com/locate/hal



An early-warning forecast model for red tide (*Karenia brevis*) blooms on the southwest coast of Florida

Miles Medina [a,*], Paul Julian II [b], Nicholas Chin [c], Stephen E. Davis [b]

[a] *ECCO Scientific, LLC, St. Petersburg, Florida, USA*
[b] *The Everglades Foundation, Palmetto Bay, Florida, USA*
[c] *Engineering School of Sustainable Infrastructure & Environment, University of Florida, Gainesville, FL, USA*

ARTICLE INFO

Editor: Edited by Chennai Guest Editor

Keywords:
Harmful algal blooms (HABs)
Machine learning
Random forest
Lake Okeechobee
Caloosahatchee River
Charlotte Harbor

ABSTRACT

Karenia brevis blooms occur nearly annually along the southwest coast of Florida, and effective mitigation of ecological, public health, and economic impacts requires reliable real-time forecasting. We present two boosted random forest models that predict the weekly maximum *K. brevis* abundance category across the Greater Charlotte Harbor estuaries over one-week and four-week forecast horizons. The feature set was restricted to data available in near-real time, consistent with adoption of the models as decision-support tools. Features include current and lagged *K. brevis* abundance statistics, Loop Current position, sea surface temperature, sea level, and riverine discharges and nitrogen concentrations. During cross-validation, the one-week and four-week forecasts exhibited 73 % and 84 % accuracy, respectively, during the 2010–2023 study period. In addition, we assessed the models' reliability in forecasting the onset of 10 bloom events on time or in advance; the one-week and four-week models anticipated the onset eight times and five times, respectively.

1. Introduction

The marine dinoflagellate *Karenia brevis* forms harmful blooms over the West Florida Shelf (eastern Gulf of Mexico) nearly annually (Liu et al., 2016; Medina et al., 2022). These blooms—known colloquially as red tides—initiate offshore following a cascade of physical, chemical, and biological processes and may be transported to the southwest coast of Florida via favorable circulation patterns (Steidinger, 2009; Vargo, 2009; Heil et al., 2014; Weisberg et al., 2019). Once a *K. brevis* bloom appears on the coast, a variety of nutrient sources may perpetuate its growth and maintenance (Vargo, 2009; Heil et al., 2014), including anthropogenic nutrient loads from developed watersheds (Medina et al., 2020, 2022; Phlips et al., 2023; Tomasko et al., 2024). The organism produces a neurotoxin (brevetoxin) that causes wildlife mortality including massive fish kills that liberate nutrients that can further fuel blooms and disrupt ecosystem function (Walsh et al., 2006; Heil et al., 2014). In addition, brevetoxin results in hospitalizations from ingestion of contaminated shellfish or due to respiratory illness following aerosolization and transport of the neurotoxin up to several kilometers inland (Kirkpatrick et al., 2004; Fleming et al., 2005; Watkins et al., 2008; Stumpf et al., 2022). These ecological and public health impacts translate into significant regional economic losses to the tourism, fishing, and real estate sectors (Bechard, 2021; Ferreira et al., 2023).

Reliable forecasting of harmful algal blooms (HABs) enables timely mobilization of resources to mitigate adverse impacts via public advisories, shellfish harvest closures, and cleanup operations such as the harvest and disposal of dead fish (Liu et al., 2023). In addition, reliable forecasts may enable mitigation of the severity of the blooms themselves, to the extent that engineered riverine discharges fuel the growth and maintenance of downstream HABs. For instance, in southwest Florida, managed discharges from the Caloosahatchee River and Lake Okeechobee convey nitrogen loads that intensify and prolong *K. brevis* blooms along the coast (Medina et al., 2022; Tomasko et al., 2024). Forecasts may indicate when to withhold such discharges in anticipation of bloom conditions.

Nonetheless, HAB forecasting is challenging due to the complex interplay of factors influencing bloom initiation, growth, maintenance, transport, and termination and how these processes translate into negative impacts on human and ecological communities in space and time (McGowan et al., 2017; Medina et al., 2022). Approaches to forecasting HABs include numerical hydrodynamic models that simulate particle advection and transport (e.g., Liu et al., 2023); coupled models that simulate transport as well as chemical mass balance and biological growth (e.g., Sherwood et al., 2016); machine learning and statistical

* Corresponding author.
E-mail address: research@eccoscientific.com (M. Medina).

https://doi.org/10.1016/j.hal.2024.102729
Received 20 May 2024; Received in revised form 22 August 2024; Accepted 19 September 2024
Available online 24 September 2024

models (e.g., Nelson et al., 2018; Izadi et al., 2021); and equation-free modeling based on low-dimensional reconstructions of system dynamics in phase space (e.g., McGowan et al., 2017).

For *K. brevis* along the Gulf coast of Florida, the current state-of-the-art model generates 3.5-day forecasts of bloom trajectories based on Lagrangian particle tracking in a numerical ocean model that uses the locations of grab samples exhibiting elevated cell counts as initial positions (Liu et al., 2023). This study introduces a machine learning model that substantially extends the forecast horizon by generating weekly 'early warning' forecasts for *K. brevis* conditions one week and four weeks in the future, representing a valuable complement to the existing forecast paradigm. The model is a random-forest classifier that predicts the weekly maximum *K. brevis* abundance category across the Greater Charlotte Harbor estuary system, in terms of five order-of-magnitude abundance categories used by the Florida Fish and Wildlife Conservation Commission (FWC): Background (0–10^3 cells/L), Very Low (10^3–10^4 cells/L), Low (10^4–10^5 cells/L), Medium (10^5–10^6 cells/L), and High (>10^6 cells/L). Inputs to the model are restricted to 'live data' (data that are available in near-real time), consistent with operational use of the model on a weekly basis. These input data include current and lagged *K. brevis* cell count statistics and environmental data for sea level, ocean temperature, Loop Current position, and riverine discharge and nitrogen concentrations.

2. Methods

2.1. Study area

The southwest Florida study area includes the forecast target domain—extending from Lemon Bay south to Estero Bay, including Charlotte Harbor, the Caloosahatchee River estuary, and San Carlos Bay—and a broader feature domain that extends north to Sarasota Bay (Fig. 1). The three major rivers that discharge to the target domain include the Myakka River, the Peace River, and the Caloosahatchee River. Discharges from the Caloosahatchee River are released at the S-79 flow control structure operated by the U.S. Army Corps of Engineers, and this river conveys surface flows originating within its own drainage basin (Caloosahatchee watershed) as well as Lake Okeechobee (discharged via the S-77 structure). Discharges to the Caloosahatchee River estuary and Charlotte Harbor are highly seasonal with wet (May to October) and dry (November to April) seasons.

A recent study of *K. brevis* blooms between 2007 and 2023 indicated that maximum cell counts typically exceed 100,000 cells/L starting around October (following the summer wet season) and persist for an average of more than five months, although a bloom that first appeared on the coast in late 2017 persisted for approximately 15 months (Tomasko et al., 2024).

2.2. Data and feature engineering

The study period for the models is May 2010 through April 2023. In developing these models, we restricted the feature set to 'live data' that are available in near-real time, consistent with adoption as an operational forecast tool. As a result, we chose to forego several potentially informative data sources that report data following a substantial delay. For instance, grab sample water quality data that undergo several months of verification and quality control would not be available to inform real-time forecasting.

The forecast target for each model is the maximum *K. brevis* abundance category observed over a one-week or four-week forecast horizon, within the spatial domain bounded by latitudes 26.4 and 27.1 and longitudes –82.5 and –81.5. The *K. brevis* abundance categories include Background (0–10^3 cells/L), Very Low (10^3–10^4 cells/L), Low (10^4–10^5 cells/L), Medium (10^5–10^6 cells/L), and High (>10^6 cells/L). Respiratory irritation can occur at Very Low cell counts, and shellfish harvesting is closed when abundance exceeds 5000 cells/L (Stumpf et al., 2022).

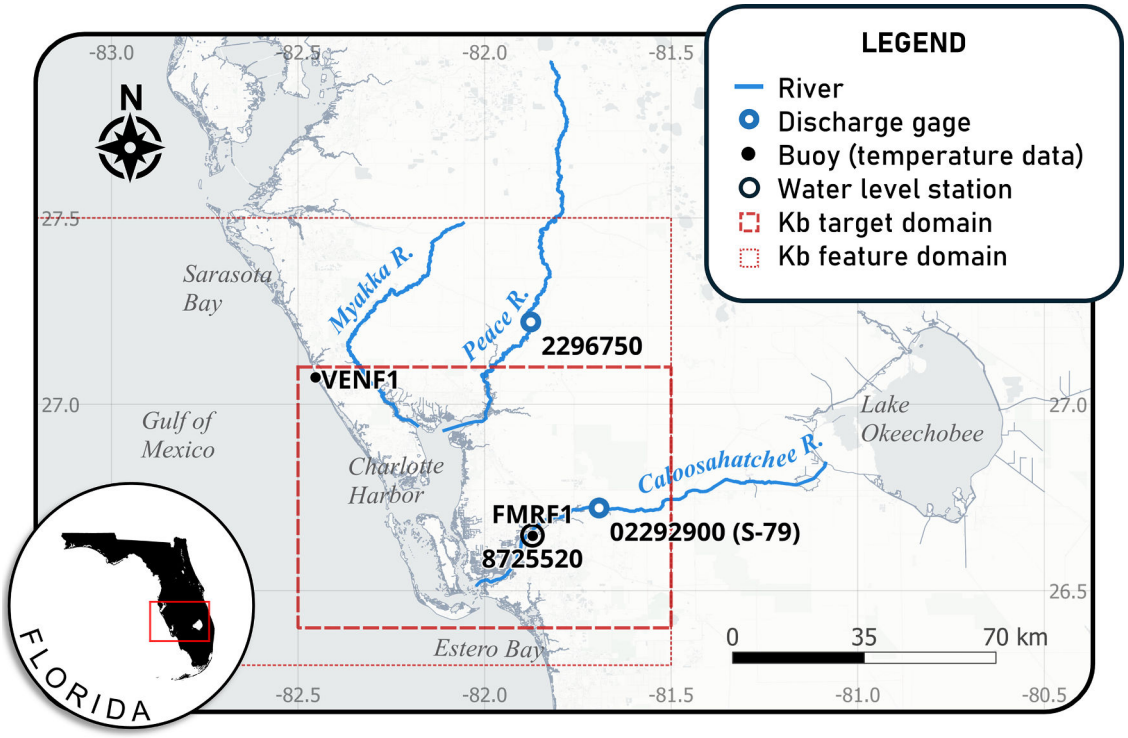


Fig. 1. The southwest Florida study area. The forecast target is the weekly maximum *K. brevis* (Kb) abundance category within the target domain (red dashed-line box). Features include *K. brevis* statistics aggregated over the feature domain (red dotted-line box), discharges at stream gages (blue circles) along the Peace and Caloosahatchee Rivers, NOx concentration at the S-79 structure co-located with the Caloosahatchee gage, mean water temperature at Venice and Fort Myers (black points), and the mean low low water level near Fort Myers (black circle). The Loop Current feature domain is not displayed.

Fish kills are considered possible at Low abundances and probable at Medium or greater abundances (Heil et al., 2014). The *K. brevis* data are collected *in situ* by the FWC and its partners and were obtained from the NOAA HABSOS database (https://habsos.noaa.gov/).

Both the one-week and the four-week forecast models rely on an identical feature set. Features include *K. brevis* cell count statistics and environmental variables including sea level, sea temperature, Loop Current position, and riverine discharges and nitrogen concentrations (Table 1). The *K. brevis*, sea temperature, sea level, and riverine features were built by aggregating available data on weekly and/or 4-week ('monthly') time frames—e.g., the weekly or monthly mean value—and the feature set includes several lags of these features. Feature name suffixes in Table 1 indicate the data aggregation and lag periods: The 'L' and 'M' suffixes indicate weekly and 4-week aggregation periods, respectively, and the subsequent digit indicates the lag period relative to the forecast target. For instance, 'L1' indicates a weekly-aggregated feature that is delayed one week (i.e., the week prior to the target period), and 'L2' indicates a weekly-aggregated feature that is delayed two weeks. An 'M1' feature is aggregated over the four-week period prior to the target, and an 'M2' feature is aggregated over the period five to eight weeks prior to the target.

Table 1
Model targets and features.

Model target or feature	Abbreviation	Lags*	Data source
Weekly maximum *K. brevis* cell count category (target)	Kb.category	–	NOAA HABSOS
4-week maximum *K. brevis* cell count category (target)	Kb. category.4wk	–	NOAA HABSOS
Weekly or 4-week maximum *K. brevis* cell count	Kb.max	L1, L2, M1–M3	NOAA HABSOS
Weekly or 4-week 95th percentile of *K. brevis* cell count	Kb.95p	L1, L2, M1–M3	NOAA HABSOS
Weekly or 4-week median *K. brevis* cell count	Kb.median	L1, L2, M1–M3	NOAA HABSOS
Weekly or 4-week mean *K. brevis* cell count	Kb.mean	L1, L2, M1–M3	NOAA HABSOS
Weekly or 4-week standard deviation of *K. brevis* cell count	Kb.sd	L1, L2, M1–M3	NOAA HABSOS
Weekly or 4-week proportion of *K. brevis* samples exceeding 10^4 cells/L	Kb. exceed_1e4	L1, L2, M1–M3	NOAA HABSOS
Weekly or 4-week proportion of *K. brevis* samples exceeding 10^5 cells/L	Kb. exceed_1e5	L1, L2, M1–M3	NOAA HABSOS
Weekly or 4-week proportion of *K. brevis* samples exceeding 10^6 cells/L	Kb. exceed_1e6	L1, L2, M1–M3	NOAA HABSOS
Weekly minimum mean low low water	MLLW.min	L1	NOAA Tides & currents
Weekly mean of mean low low water	MLLW.mean	L1	NOAA Tides & currents
Weekly maximum mean low low water	MLLW.max	L1	NOAA Tides & currents
Weekly mean sea temperature	NDBC.temp	L1–L6	NOAA NDBC
Weekly Loop Current center of mass position (latitude)	LC.lat	L0, L1, L5, L9, L13, L17, L21, L25, L29	Copernicus
Weekly Loop Current center of mass position (longitude)	LC.lon	L0, L1, L5, L9, L13, L17, L21, L25, L29	Copernicus
Weekly sum of Peace River discharges at Arcadia	Peace.Q. log10	L1–L6	USGS NWIS
Weekly sum of Caloosahatchee River discharges at S-79	Caloo.Q. log10	L1–L6	USGS NWIS
Weekly or 4-week maximum NOx concentration at S-79	NOx.max	L1, L2, M1–M6	USGS NWIS and SFWMD DBHYDRO

* One-week lags are denoted by "L" and the number of weeks prior to the target week. For instance, an "L1" lag corresponds to the week prior to the target. Four-week lags are denoted by "M" and the number of four-week periods prior to the target. For instance, an "M1" lag corresponds to the period one to four weeks prior to the target, and an "M2" lag corresponds to the period five to eight weeks prior to the target.

The *K. brevis* features include descriptive statistics for cell counts (maximum, 95th percentile, median, mean, and standard deviation) as well as the proportions of *K. brevis* samples exceeding thresholds of 10^4, 10^5, and 10^6 cells/L. The spatial domain of the *K. brevis* features is bounded by latitudes 26.3 and 27.5 and longitudes –87 and –81.5. The *K. brevis* feature domain is therefore a superset of the *K. brevis* target domain.

The sea level features were developed using mean low low water (MLLW) data from NOAA station 8725520 at Fort Myers, Florida. The sea temperature features were developed by averaging data from NOAA NDBC buoys VENF1 and FMRF1 at Venice, Florida and Fort Myers, Florida, respectively, collected at 1 m below MLLW. The riverine discharge features were developed using data from USGS gages at the Caloosahatchee River and Peace River (gages 02292900 and 02296750, respectively). Nitrogen concentration features were developed using nitrite+nitrate (NOx) concentration data collected at the S-79 flow control structure on the lower segment of the Caloosahatchee River. These concentration data were obtained from the USGS gage (02292900) and co-located South Florida Water Management District (SFWMD) monitoring stations S-79 (currently active) and CES01 (inactive).

The geographic coordinates of the Loop Current position were estimated as the center of mass of the sea surface height within the eastern Gulf of Mexico, bounded by latitudes 23.1 and 28.9 and longitudes –88 and –83. The sea surface height data were obtained from Copernicus ocean physics reanalysis and analysis/forecast products for the date ranges January 2000 through October 2020 and November 2020 through October 2023, respectively. Because Copernicus provides forecasted values, the Loop Current features include an unlagged 'L0' value that coincides with the target week. Initially, we intended to rely solely on the analysis/forecast product, since reanalysis data are not available in real time, but the analysis/forecast products are removed from the Copernicus repository once reanalysis products are published, and our attempts to obtain the earlier analysis/forecast values (i.e., prior to November 2020) were unsuccessful.

We elected not to include salinity as a feature, because we did not expect it to provide predictive value on the spatial scale of the target domain (Fig. 1). It is true that *K. brevis* is a marine species with a salinity range of about 20 to 45 PSU (Steidinger, 2009) and that the target domain includes mesohaline estuarine environments where salinity conditions may at times be too fresh to support population growth. However, one would not expect salinity to be predictive of the forecast target—the *maximum* abundance category for *K. brevis*—within the target spatial domain, which extends west into the Gulf of Mexico.

2.3. Random forest model development

A random forest classifier comprises an ensemble of classification *trees* that map input *features* to output *classes* (Fig. 2). Each tree partitions the feature space using a random subset of features and observations to predict the target with minimal misclassification error (i.e., by minimizing some loss function) while avoiding overfitting with cross-validation (Hastie et al., 2023; James et al., 2013). Whereas any individual tree is a weak classifier, the ensemble prediction of the forest typically achieves markedly greater accuracy (Cutler et al., 2007; James et al., 2013). *Boosting* further improves performance by constructing

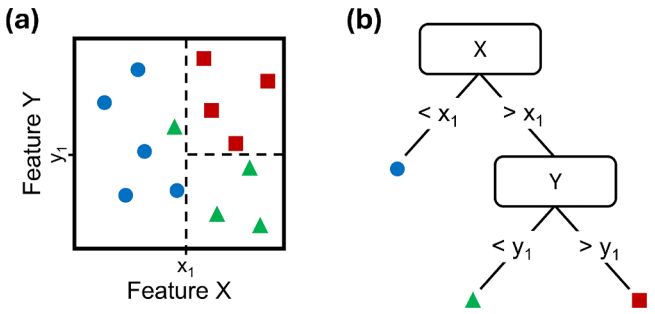


Fig. 2. A random forest classifier comprises a large set of classification trees. In this simplified example, a single classification tree predicts the output class (blue circle, green triangle, or red square) based on two input features, X and Y. (a) During model training, the feature space is partitioned using the values x_1 and y_1 to minimize impurity within each region. Any new observation can then be classified according to the region it occupies in the feature space. (b) The decision tree represents the set of sequential 'rules' for classifying new observations.

classification trees in sequence, rather than in parallel, with each subsequent tree placing more weight on observations that were misclassified by the previous tree (Hastie et al., 2023; Friedman 2001). Key advantages of the random forest approach include high classification accuracy, the ability to model linear and nonlinear interactions, and interpretability of the model in terms of ranking the importance of features (Cutler et al., 2007). Other machine learning approaches that have been used to forecast HABs include neural networks and support vector machines (e.g., Izadi et al., 2021).

We developed two boosted random forest models to predict the weekly maximum *K. brevis* abundance category, with $k = 5$ output classes modeled as unordered factors (multinomial), observed within the target domain over (1) one-week and (2) four-week forecast horizons, using the *gbm* package (Greenwell et al., 2022) in R version 3.4.1 (R Core Team, 2023). Hyperparameters that control the algorithm include the number of trees, interaction depth (i.e., the maximum number of feature interactions allowed within each tree), and the learning rate. Following experimentation with several combinations of hyperparameter values, we adopted values of 5000, 10, and 10^{-9}, respectively, for each model. The models output probabilities for each of the $k = 5$ classes, and the predicted class is the class with the largest probability. Feature importance is ranked according to features' relative influence in reducing the multinomial loss function.

2.4. Model evaluation

We evaluated model performance by measuring accuracy and Cohen's kappa score during k-fold cross-validation, using each water year as a fold. Model accuracy is computed as the number of correct classifications divided by the total number of classifications:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FP} \quad (1)$$

where TP, TN, FP, and FN are the number of true positives, true negatives, false positives, and false negatives, respectively. The kappa score κ accounts for bias that may arise from imbalanced classes by relating observed model accuracy (p_o) to the hypothetical probability of correct classification by random chance (p_e):

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (2)$$

Our feature set includes statistics describing past *K. brevis* abundances (e.g., weekly maximum cell count) that one should expect to be highly predictive of the target (maximum *K. brevis* abundance category), introducing the potential for an autoregressive model that generates lagged forecasts. Such a model may achieve high accuracy and kappa metrics during cross-validation but provide little value as an operational management tool, because forecasts that systematically lag reality will miss the onset of *K. brevis* blooms and require one or more time steps to 'catch up.' Therefore, in addition to computing accuracies and kappa scores, we visualized the observed and forecasted classes over time to assess whether the models generated systematically lagged forecasts that missed the onset of blooms.

Finally, we aggregated the model forecasts into three *K. brevis* abundance superclasses (Background/Very low, Low, and Medium/High), re-computed accuracies and kappa scores, and visualized the aggregated forecasts over time. Although this post-hoc aggregation reduces the resolution of the forecast, the three superclasses are nonetheless relevant to *K. brevis* bloom management, mitigation, and ecological effects.

3. Results

We designed two classifiers that forecasted the weekly maximum *K. brevis* abundance category over one-week and four-week horizons with accuracies of 0.503 and 0.590, respectively, during cross-validation (Table 2). When the forecasts were aggregated into three superclasses—Background/Very Low, Low, and Medium/High—the accuracies increased to 0.730 and 0.844, respectively. In each case, the models classified the extreme classes well but exhibited confusion among intermediate classes (Fig. 3). Both the one-week and the four-week models exhibited confusion between the two upper classes—Medium and High—with a greater bias toward overestimation (i.e., classifying Medium observations as High) than underestimation (Fig. 3a, 3c). The improved accuracy achieved by aggregating the forecasts into three superclasses arose largely from combining these two classes (Figs. 3b, 3d).

Because accurately forecasting the onset of a *K. brevis* bloom is a particularly important goal for real-time management and mitigation efforts, we plotted the forecasted classes over time to assess the models' ability to predict the onset of a bloom on time or in advance (Fig. 4). Here, we define the onset of a bloom event in terms of the first observed Medium or High class following a non-bloom period, and we consider the forecast of the onset to be correct if the Medium or High class is predicted on time or in advance of the observation. By this definition, the record contains 10 onsets, occurring in the latter parts of years 2011 through 2017, 2019, 2020, and 2022. The one-week model forecasted eight of ten bloom onsets on time or in advance, missing the onsets of the 2020 and 2022 blooms by one time step (Fig. 4a–b). The four-week model forecasted five of ten bloom onsets on time, missing the onsets of blooms in 2011, 2012, 2014, 2020, and 2022, often by multiple time steps (Fig. 4c–d).

Both the one-week and the four-week models ranked several "current" *K. brevis* features (i.e., one-week prior to the target) highest for feature importance (Fig. 5). Other features that ranked relatively highly included the Loop Current latitude, Peace River discharge, and the maximum NOx concentration at the Caloosahatchee River. These rankings comport with current mechanistic understanding based on earlier studies (e.g., Liu et al., 2023; Medina et al., 2022) and therefore

Table 2
Model evaluation.

Model	Training		Testing	
	accuracy	Kappa	accuracy	Kappa
1-week forecast				
5 abundance classes	0.753	0.680	0.503	0.351
3 abundance classes			0.730	0.509
4-week forecast				
5 abundance classes	0.836	0.711	0.590	0.418
3 abundance classes			0.844	0.702

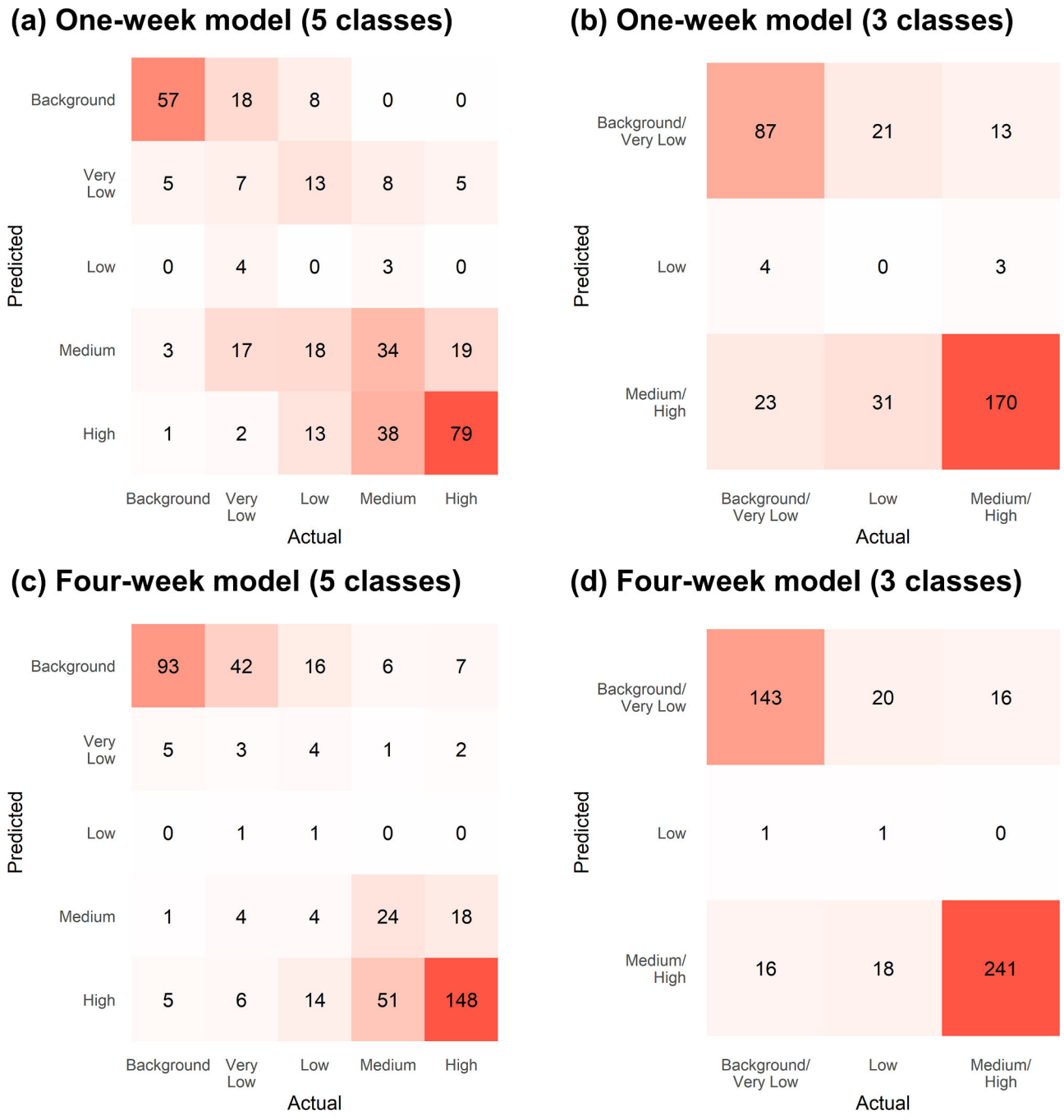


Fig. 3. Confusion matrices display *K. brevis* abundance classes predicted by the one-week (*a, b*) and four-week (*c, d*) forecast models against actual abundance classes. Panels *b* and *d* display results following post-hoc aggregation of the five original classes into three superclasses.

serve as a valuable 'sanity check' indicating that the models likely encoded some expected real-world relationships. However, we note that feature importance should not be interpreted as a ranking of causal mechanisms in the real world, particularly among correlated features such as Peace and Caloosahatchee River discharges. Feature importance simply indicates which features were most strongly associated with minimizing the loss function.

4. Discussion

We developed two classifiers that generate weekly forecasts of *K. brevis* conditions along the southwest Florida coast near Charlotte Harbor, over one-week and four-week forecast horizons. The feature set was intentionally limited to data that are available in real time (Table 1). As such, our results demonstrate the feasibility of short-term HAB forecasting under realistic management conditions (i.e., based only on 'live' data), and the models may serve as an operational decision-support tool to inform efforts to mitigate the severity of *K. brevis* blooms along the southwest coast of Florida, as well as their ecological, public health, and economic impacts. In particular, we recommend adoption of both forecast models simultaneously, considering a key performance disparity between them: The four-week forecasts exhibited greater classification accuracy than the one-week forecasts (84 % and 73 %, respectively), whereas the one-week forecasts correctly anticipated the onset of bloom events more often than did the four-week forecasts (80 % and 50 %, respectively). Finally, our modeling approach may serve as a framework for developing data-driven forecast models in other coastal areas.

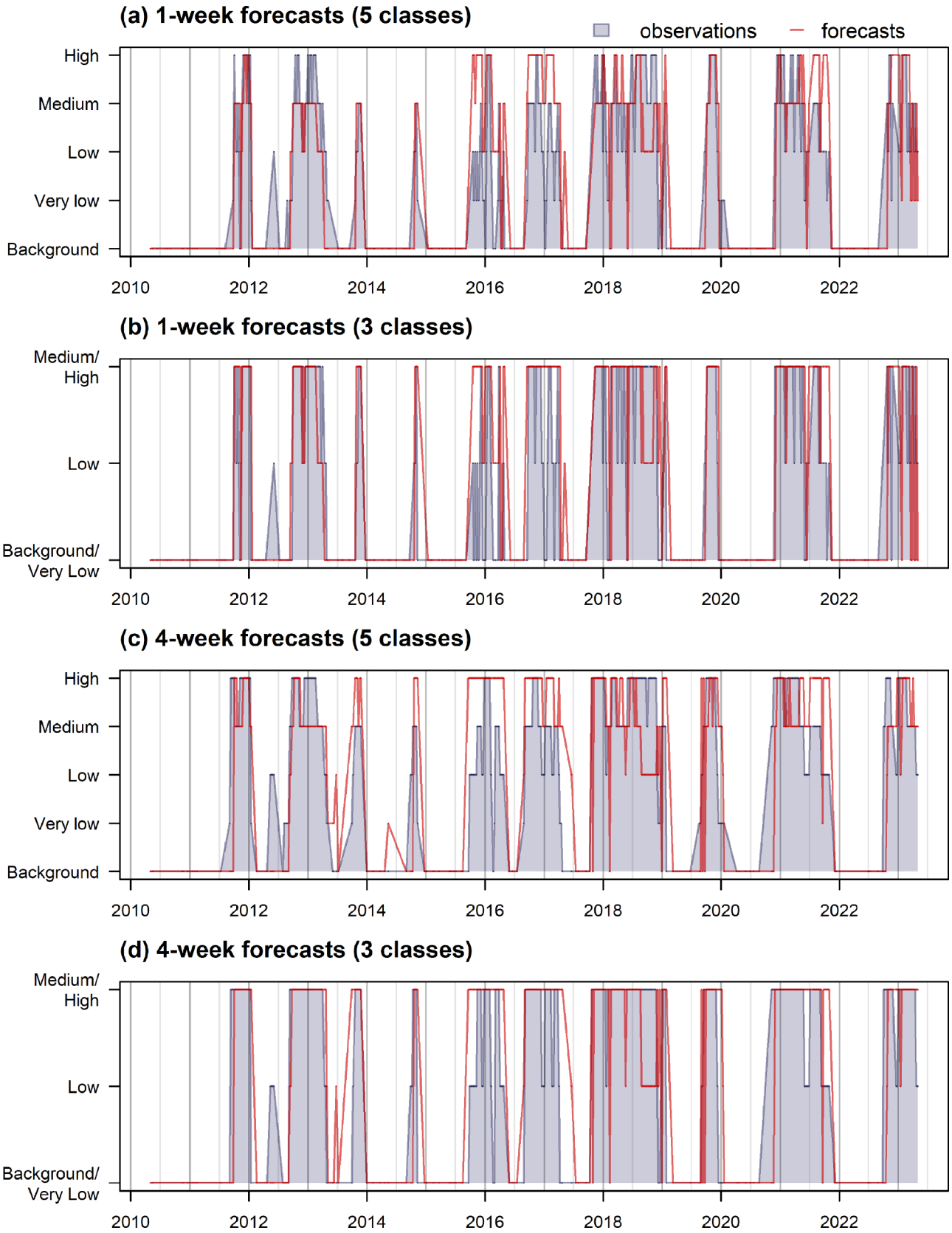


Fig. 4. Observed *K. brevis* abundance classes (gray) and out-of-sample forecasts (red) over time, for the one-week (*a, b*) and four-week (*c, d*) models. Panels *b* and *d* display results following post-hoc aggregation of the five original classes into three superclasses.

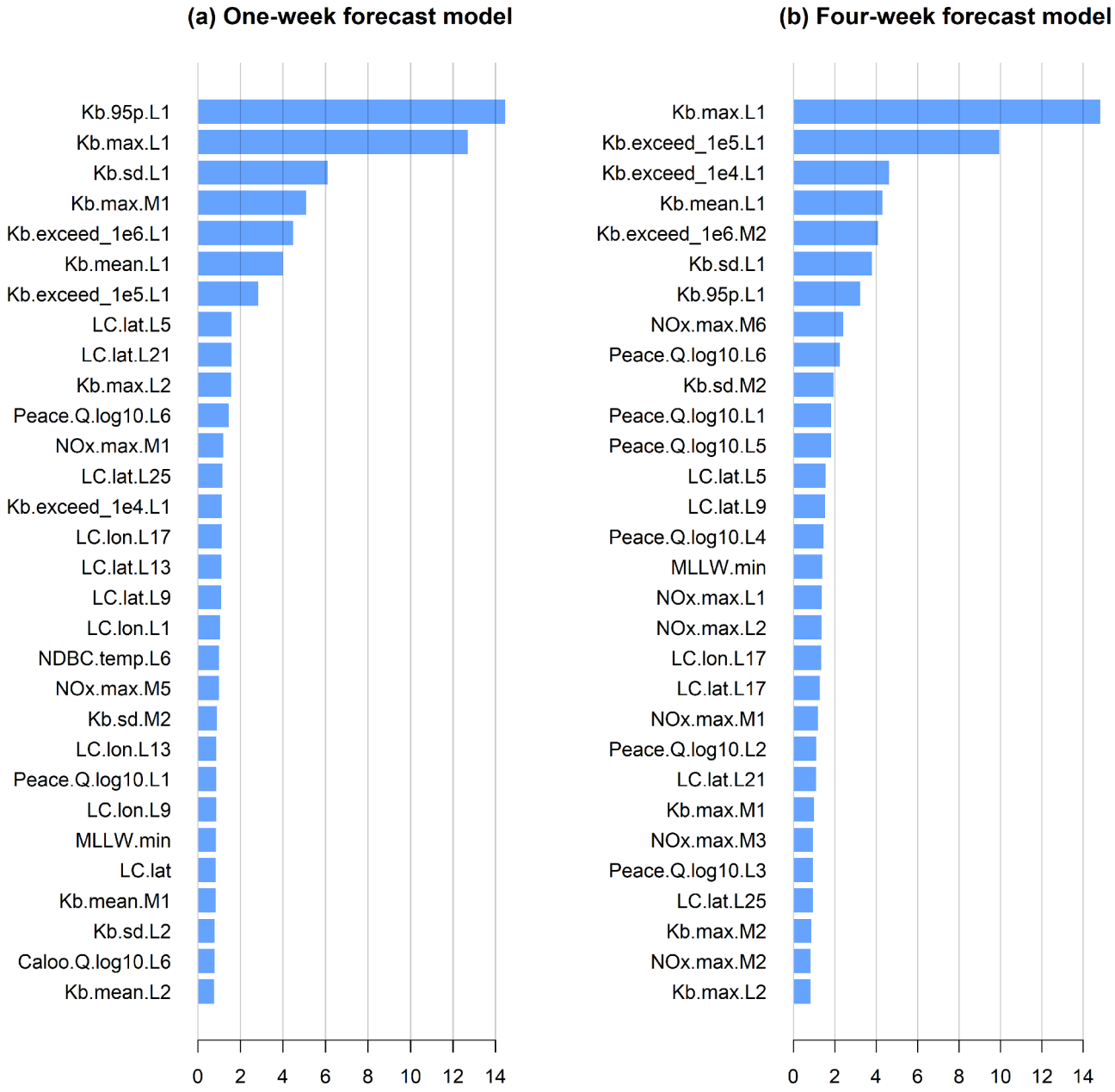


Fig. 5. Relative influence plots display the 30 top-ranked features in the one-week (*a*) and four-week (*b*) models.

The current state-of-the-art model for forecasting Florida red tides produces 3.5-day forecasts of nearshore *K. brevis* bloom trajectories using output from two ocean physics models that simulate advection and transport processes (Lagrangian particle tracking) with initial positions provided by *in situ* cell count data (Liu et al., 2023). This forecasting tool, developed and maintained at the University of South Florida (USF), is executed twice weekly on the university's high-performance computing cluster. Although our model is relatively reductive in its conceptualization of the problem—reducing a spatially heterogeneous HAB phenomenon to a single categorical target—our forecasts may nonetheless serve as a valuable complement to the existing USF forecasts by providing an 'early warning' of bloom conditions, up to four weeks in advance, at relatively low computational cost.

We visualized the models' out-of-sample forecasts over time (Fig. 4) in order to assess their ability to reliably predict the onset of blooms, after recognizing that forecasts may lag reality and yet exhibit high accuracy during cross-validation. This potential forecasting problem often arises when forecast models rely on lagged values of the target as features (i.e., $y_t \sim y_{t-1} + \ldots$)—an appropriate and often necessary practice that nonetheless carries a risk of generating lagged forecasts due to autocorrelation in the target. That is, during training, a model may learn to minimize error by repeating the most recent observation as its forecast: Given low cell counts 'this week,' the model predicts low cell counts 'next week.' In the extreme, this strategy generates reasonable albeit trivial forecasts when the target changes slowly, and the forecast lags behind reality whenever abrupt changes occur, requiring one or more time steps to 'catch up.' In the context of HABs, a systematically lagged forecast means missing the onset of a bloom, thus providing little operational value to end users relying upon the forecast to inform consequential decisions such as public health advisories and environmental management actions in real time. Because such a model may perform well during cross-validation despite systematically lagged forecasts, reporting cross-validation metrics alone will typically fail to reveal the problem. We therefore recommend that researchers proposing machine-learning models for HAB forecasting present plots of observed and predicted values over time, in addition to accuracy metrics. The

absence of such plots may, whether intentionally or unintentionally, give a false impression that the model is a suitable candidate for operational use.

We initially approached *K. brevis* forecasting as a regression problem and built models to predict, for instance, the weekly maximum cell count or the proportions of samples exceeding certain abundance thresholds. A few of these models performed reasonably well during cross-validation, but they invariably generated lagged forecasts due to autocorrelation, as described above. More often, our attempts generated poor forecasts when confronted with out-of-sample data. We suspect that their poor performance is ultimately attributable to our use of a single highly aggregated quantity as the forecast target in each regression model, considering the spatiotemporal complexity of HAB phenomena. Although an alternative model architecture such as a neural network might improve performance in both the regression and classification settings, we suspect that any substantial improvement will require spatial representations of the target and at least some of the features. For instance, the use of remote sensing (satellite) features or raster maps of *K. brevis* abundance (Izadi et al., 2021; Fick et al., 2024) as model inputs would encode the spatial distribution and temporal evolution of blooms and perhaps also implicitly encode the complex hydrodynamics of the study area (Dye et al., 2020; Hewageegana et al., 2023; Shi et al., 2023).

CRediT authorship contribution statement

Miles Medina: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Paul Julian:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Nicholas Chin:** Writing – review & editing, Methodology, Conceptualization. **Stephen E. Davis:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank the editor and reviewers for their efforts and constructive comments on this manuscript.

References

Bechard, A., 2021. Gone with the wind: declines in property values as harmful algal blooms are blown towards the shore. J. Real Estate Financ. Econ. 62, 242–257. https://doi.org/10.1007/s11146-020-09749-6.

Cutler, R.D., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88 (11), 2783–2792. https://doi.org/10.1890/07-0539.1.

Ferreira, J.P., Saha, B.B., Carrero, G.C., Kim, J., Court, C., 2023. Impacts of red tide in peer-to-peer accommodations: a multi-regional input-output model. Tour. Econ. 29 (3), 812–834. https://doi.org/10.1177/1354816621106827 6.

Fick, R., Medina, M., Angelini, C., Kaplan, D., Gader, P., He, W., 2024. Fusing remote sensing data with spatiotemporal in-situ samples for red tide (*Karenia brevis*) detection. Integr. Environ. Asses. Manag. https://doi.org/10.1002/ieam.4908.

Fleming, L., Backer, L., Baden, D., 2005. Overview of aerosolized Florida red tide toxins: exposures and effects. Environ. Health Perspect. 113, 618–620. https://doi.org/10.1289/ehp.7501.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232.

Greenwell, B., Boehmke, B., Cunningham, J., G.B.M. Developers, 2022. gbm: generalized boosted regression models. R package version, 2.1.8.1. https://CRAN.R-project.org/package=gbm.

Hastie, T., Tibshirani, R., Friedman, J., 2023. The Elements of Statistical learning: Data mining, inference, and Prediction, 2nd ed. Springer.

Heil, C.A., Dixon, L.K., Hall, E., Garrett, M., Lenes, J.M., O'Neil, J.M., Weisberg, R.W., 2014. Blooms of Karenia brevis (Davis) G. Hansen & Ø. Moestrup on the West Florida Shelf: nutrient sources and potential management strategies based on a multiyear regional study. Harmful Algae 38, 127–140. https://doi.org/10.1016/j.hal.2014.07.016.

Izadi, M., Sultan, M., Kadiri, R.E., Ghannadi, A., Abdelmohsen, K., 2021. A remote sensing and machine learning-based approach to forecast the onset of harmful algal bloom. Remote Sens. (Basel) 13 (19), 3863. https://doi.org/10.3390/rs13193863.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning With Applications in R, 2nd ed. Springer.

Kirkpatrick, B., Fleming, L., Squicciarini, D., Backer, L., Clark, R., Abraham, W., 2004. Literature review of Florida red tide: implications for human health. Harmful Algae 3, 99–115.

Liu, Y., Weisberg, R.H., Lenes, J.M., Zheng, L., Hubbard, K., Walsh, J.J., 2016. Offshore forcing on the "pressure point" of the West Florida Shelf: anomalous upwelling and its influence on harmful algal blooms. J. Geophys. Res. Oceans 121, 5501–5515. https://doi.org/10.1002/2016JC011938.

Liu, Y., Weisberg, R.H., Zheng, L., Hubbard, K.A., Muhlbach, E.G.0., Garrett, M.J., Hu, C., Cannizzaro, J.P., Xie, Y., Chen, J., John, S., Liu, L.Y., 2023. Short-term forecast of Karenia brevis trajectory on the West Florida Shelf. Deep Sea Res. Part II Top. Stud. Oceanogr. 212, 105335. https://doi.org/10.1016/j.dsr2.2023.105335.

McGowan, J.A., Deyle, E.R., Ye, H., Carter, M.L., Perretti, C.T., Seger, K.D., Sugihara, G., 2017. Predicting coastal algal blooms in southern California. Ecology 98 (5), 1419–1433. https://doi.org/10.1002/ecy.1804.

Medina, M., Huffaker, R., Jawitz, J.W., Muñoz-Carpena, R., 2020. Seasonal dynamics of terrestrially sourced nitrogen influence Karenia brevis blooms of Florida's southern Gulf Coast. Harmful Algae 98, 101900. https://doi.org/10.1016/j.hal.2020.101900.

Medina, M., Kaplan, D., Milbrandt, E., Tomasko, D., Huffaker, R., Angelini, C., 2022. Nitrogen-enriched discharges from a highly managed watershed intensity red tide (Karenia brevis) blooms in southwest Florida. Sci. Total Environ. 827, 154149. https://doi.org/10.1016/j.scitotenv.2022.154149.

Nelson, N.G., Muñoz-Carpena, R., Phlips, E.J., Kaplan, D., Sucsy, P., Hendrickson, J., 2018. Revealing biotic and abiotic controls of harmful algal blooms in a shallow subtropical lake through statistical machine learning. Environ. Sci. Technol. 52, 3527–3535. https://doi.org/10.1021/acs.est.7b05884.

Phlips, E.J., Badylak, S., Mathews, A.L., Milbrandt, E.C., Montefiore, L.R., Morrison, E.S., Nelson, N., Stelling, B., 2023. Algal blooms in a river-dominated estuary and nearshore region of Florida, USA: the influence of regulated discharges from water control structures on hydrologic and nutrient conditions. Hydrobiologia 850, 4385–4411. https://doi.org/10.1007/s10750-022-05135-w.

R Core Team, 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Sherwood, E.T., Greening, H., Garciz, L., Kaufman, K., Janicki, T., Pribble, R., Cunningham, B., Penne, S., Fitzpatrick, J., Dixon, K., Wessel, M., 2016. Development of an integrated ecosystem model to determine effectiveness of potential watershed management projects on improving Old Tampa Bay. Assessments for Watershed Management. In: Stringer, C.E., Krauss, K.W., Latimer, J.S. (Eds.), Headwaters to estuaries: advances in watershed science and management. Proceedings of the Fifth Interagency Conference on Research in the Watersheds, North Charleston, South Carolina. General Technical Report SRS-211. Asheville, NC: U.S. Department of Agriculture Forest Service. Southern Research Station.

Steidinger, K.A., 2009. Historical perspective on Karenia brevis red tide research in the Gulf of Mexico. Harmful Algae 8, 549–561. https://doi.org/10.1016/j.hal.2008.11.009.

Stumpf, R.P., Li, Y., Kirkpatrick, B., Litaker, R.W., Hubbard, K.A., Currier, R.D., Harrison, K.K., Tomlinson, M.C., 2022. Quantifying Karenia brevis bloom severity and respiratory irritation impact along the shoreline of Southwest Florida. PLoS ONE 17 (1), e0260755. https://doi.org/10.1371/journal.pone.0260755.

Tomasko, D., Landau, L., Suau, S., Medina, M., Hecker, J., 2024. An evaluation of the relationships between the duration of red tide (*Karenia brevis*) blooms and watershed nitrogen loads in southwest Florida (USA). Florida Scientist 87 (2), 61–72.

Vargo, G.A., 2009. A brief summary of the physiology and ecology of Karenia brevis Davis (G. Hansen & Moestrup comb. nov.) red tides on the West Florida Shelf and hypotheses posed for their initiation, growth, maintenance, and termination. Harmful Algae 8, 573–584. https://doi.org/10.1016/j.hal.2008.11.002.

Walsh, J.J., Jolliff, J.K., Darrow, B.P., Lenes, J.M., Milroy, S.P., Remsen, A., Dieterle, D. A., Carder, K.L., Chen, F.R., Vargo, G.A., Weisberg, R.H., Fanning, K.A., Muller-Karger, F.E., Shinn, E., Steidinger, K.A., Heil, C.A., Tomas, C.R., Prospero, J.S., Lee, T.N., Kirkpatrick, G.J., Whitledge, T.E., Stockwell, D.A., Villareal, T.A., Jochens, A. E., & P.S. Bontempi. (2006). Red tides in the Gulf of Mexico: where, when, and why? J. Geophys. Res. 111 (C11), C11003. 10.1029%2F2004JC002813.

Watkins, S.M., Reich, A., Fleming, L.E., Hammond, R., 2008. Neurotoxic shellfish poisoning. Mar. Drugs 6, 431–455. https://doi.org/10.3390/md20080021.

Weisberg, R., Liu, Y., Lembke, C., Hu, C., Hubbard, K., Garrett, M., 2019. The coastal ocean circulation influence on the 2018 West Florida Shelf K. brevis red tide bloom. J. Geophys. Res. Oceans 124, 2501–2512. https://doi.org/10.1029/2018JC014887.