Integrated Environmental Assessment and Management — Volume 00, Number 00—pp. 1–15
Received: 16 November 2023 | Revised: 1 February 2024 | Accepted: 1 February 2024


Environmental Management

# Fusing remote sensing data with spatiotemporal in situ samples for red tide (*Karenia brevis*) detection

*Ronald Fick,[1] Miles Medina,[1,2] Christine Angelini,[1] David Kaplan,[1] Paul Gader,[1,†] Wenchong He,[1,3] Zhe Jiang,[1,3] and Guangming Zheng[4,5]*

[1]*Center for Coastal Solutions, University of Florida, Gainesville, Florida, USA*
[2]*ECCO Scientific, LLC, St. Petersburg, Florida, USA*
[3]*Computer & Information Science & Engineering, University of Florida, Gainesville, Florida, USA*
[4]*NOAA/NESDIS Center for Satellite Applications and Research, College Park, Maryland, USA*
[5]*Cooperative Institute for Satellite Earth System Studies, Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA*
[†]*Deceased June 6, 2023.*

## Abstract

We present a novel method for detecting red tide (*Karenia brevis*) blooms off the west coast of Florida, driven by a neural network classifier that combines remote sensing data with spatiotemporally distributed in situ sample data. The network detects blooms over a 1-km grid, using seven ocean color features from the MODIS-Aqua satellite platform (2002–2021) and in situ sample data collected by the Florida Fish and Wildlife Conservation Commission and its partners. Model performance was demonstrably enhanced by two key innovations: depth normalization of satellite features and encoding of an in situ feature. The satellite features were normalized to adjust for depth-dependent bottom reflection effects in shallow coastal waters. The in situ data were used to engineer a feature that contextualizes recent nearby ground truth of *K. brevis* concentrations through a K-nearest neighbor spatiotemporal proximity weighting scheme. A rigorous experimental comparison revealed that our model outperforms existing remote detection methods presented in the literature and applied in practice. This classifier has strong potential to be operationalized to support more efficient monitoring and mitigation of future blooms, more accurate communication about their spatial extent and distribution, and a deeper scientific understanding of bloom dynamics, transport, drivers, and impacts in the region. This approach also has the potential to be adapted for the detection of other algal blooms in coastal waters. *Integr Environ Assess Manag* 2024;00:1–15. © 2024 SETAC

**KEYWORDS:** Machine learning; Neural networks; Remote sensing

## INTRODUCTION

Harmful algal blooms (HABs) in coastal and marine environments are a growing concern in the United States and around the world (Anderson et al., 2021; Sonak et al., 2018). In the eastern Gulf of Mexico, blooms of the marine dinoflagellate *Karenia brevis*—commonly referred to as red tides—occur nearly every year, with adverse effects on water quality, ecological communities, public health, and economic activity particularly along the southwestern coast of Florida (Bechard, 2021; Court et al., 2021; Sonak et al., 2018; Turley et al., 2022). Red tide bloom initiation, transport, intensification, and decline follow a complex cascade of atmospheric, oceanographic, and biogeochemical processes, making them difficult to understand, predict, and mitigate (Medina et al., 2022; Steidinger, 2009; Weisberg et al., 2019).

This article contains online-only Supporting Information.
Address correspondence to rfick@ufl.edu
Published on wileyonlinelibrary.com/journal/ieam.

Microscopic analysis of in situ samples is the conventional method for detecting and quantifying *K. brevis* blooms. Such field sampling has been instrumental in tracking blooms along Florida's Gulf Coast and in estimating associated environmental and economic impacts (Court et al., 2021; Stumpf et al., 2022). Indeed, monitoring in this area has been remarkably frequent and consistent over recent decades (Heil & Steidinger, 2009). However, the spatiotemporal resolution of in situ sampling is constrained by the large spatial domain and the cost of sample collection and processing. These sample data are therefore necessarily spatiotemporally coarse, limiting their utility in real-time decision-making and response efforts (Heil et al., 2014) and in advancing our fundamental scientific knowledge of bloom dynamics and drivers.

Remote sensing (RS)-based approaches for red tide detection, based on satellite imagery, hold promise for addressing this urgent need for reliable, spatially extensive, and highly resolved bloom detection in near-real time, with applications in the Gulf of Mexico and globally (Table 1). However, the majority of existing RS methods are based on simple

 

TABLE 1 Summary of methods published in the peer-reviewed literature for red tide detection

| Author | Sensor(s) | Approach | Spatial application |
|---|---|---|---|
| Stumpf et al. (2003) | SeaWiFS | Chlorophyll anomalies and local knowledge | Gulf of Mexico |
| Tomlinson et al. (2004) | SeaWiFS | Chlorophyll anomalies | Gulf of Mexico |
| Hu et al. (2005) | MODIS and SeaWiFS | Qualitative bloom detection | Gulf of Mexico |
| Ahn and Shanmugam (2006) | SeaWiFS | Red Tide Index | Northeast Asia |
| Cannizzaro et al. (2008) | MODIS and SeaWiFS | $Chl-a > 1.5\ mg/m^3$ and $b_{bp}(550)$ less than Morel (1988) relationship | Gulf of Mexico |
| Cannizzaro et al. (2009) | MODIS | $b_{bp}(550) \leq 0.0045\ m^2/mg$ and<br>$Chl-a > 1.5\ mg/m^3$ | Gulf of Mexico |
| Amin et al. (2009) | MODIS | Red band difference and *Karenia brevis* bloom index | Gulf of Mexico |
| Cheng et al. (2009) | MODIS | SVM and Random forest and KNN | Gulf of Mexico |
| Kim et al. (2009) | MODIS | Multistage algorithm | Korean Sea |
| Tomlinson et al. (2009) | MODIS and SeaWiFS | Chlorophyll anomalies and $b_{bp}(550)$ from Cannizzaro et al. (2008) and spectral shape at 490 nm | Gulf of Mexico |
| Carvalho et al. (2010) | MODIS | Bio-optical and empirical and hybrid approaches from de Araujo Carvalho (2008) | Gulf of Mexico |
| Gokaraju et al. (2011) | MODIS and SeaWiFS | Kernel SVM with KPCA and wavelet feature processing | Gulf of Mexico |
| Al Shehhi et al. (2013) | MODIS | Correlation of features with bloom conditions | Arabian Gulf |
| Lou and Hu (2014) | GOCI | Red Tide Index | East China Sea |
| Yunus et al. (2015) | Landsat OLI | Hot spot analysis | Tokyo Bay |
| Sakuno et al. (2019) | Sentinel-2 | Red tide index | Lake Koyama-ike, Japan |
| Lee et al. (2019) | Landsat-8 OLI | Neural networks | Korean Peninsula |
| Hill et al. (2020) | MODIS | CNN feature extraction, followed by classification with LSTM, MLP, SVM, or Random Forest | Gulf of Mexico and Arabian Gulf |
| Zhao et al. (2021) | HY-1D CZI | Modified U-Net (training labels from visual inspection) | East China Sea |
| Shin et al. (2021) | PlanetScope | U-Net (training labels from the red tide index) | Korean Peninsula |
| Soto et al. (2021) Bernard et al. (2021) | MODIS | $Chl-a > 1.5\,mg/m^3$<br>$nFLH > 0.01\,mW/cm^2\,\mu m\,sr$<br>$b_{bp}\ ratio < 1$ | Gulf of Mexico |

Abbreviations: CNN, convolutional neural network; KNN, K-nearest neighbor; KPCA, kernel principal component analysis; LSTM, long short-term memory; MLP, multilayer perceptron; SVM, support vector machine.

empirical or feature thresholding (e.g., comparing the chlorophyll-*a* level against a threshold or detecting anomalies in the chlorophyll-*a* level) (e.g., Ahn & Shanmugam, 2006; de Araujo Carvalho, 2008; Yunus et al., 2015). Further, these methods require manual tuning of the threshold through visual assessment and are thus difficult to automate.

The application of machine learning (ML) for remote detection of *K. brevis* blooms is a relatively recent development that presents the opportunity for automated learning of the relationship between RS data and bloom conditions. Existing approaches range from traditional ML methods such as K-nearest neighbors (KNNs), random forests, and support vector machines (Cheng et al., 2009) to more recent deep learning architectures such as convolutional neural networks (CNN), long short-term memory (LSTM), and U-Net. These ML-based approaches have used a variety of input features from a range of satellite sensors (e.g., MODIS, SeaWiFS, GOCI, Landsat, Sentinel2, PlanetScope), but they have rarely used in situ data for training or validation. The current state of the art for ML-based remote red tide detection (with the greatest reported accuracy) is a CNN classifier trained by a historical in situ record (Hill et al., 2020). Critically, however, comparison of the performance of ML methods presented in the literature is not

 

straightforward, since performance is most often reported solely in qualitative terms (e.g., visual comparison of pixel classifications with in situ data on a map). In rare cases, in which an individual study, such as Hill et al. (2020), reports quantitative performance metrics, the results are typically not directly comparable across studies due to differences in the input data (e.g., different input datasets, splits for training and testing data, or thresholds defining bloom and/or nonbloom classes).

We propose a novel ML-based remote red tide detection method that combines RS data with in situ sample data to predict *K. brevis* cell count concentration at a 1-km resolution over the West Florida Shelf and nearshore waters. The method uses a neural network to classify satellite image pixels into one of six classes of *K. brevis* cell concentration (0–1000 cells/L, 1000–10 000 cells/L, 10 000–50 000 cells/L, 50 000–100 000 cells/L, 100 000–1 000 000 cells/L, and 1 000 000+ cells/L) using MODIS ocean color features and in situ sample data. Our method introduces two key innovations: depth normalization of satellite features and engineering of an additional feature from in situ data. The depth normalization adjusts for the depth dependence of ocean color features (e.g., chlorophyll) in shallow coastal waters, where remotely detectable surface blooms often occur (Abbas et al., 2019). The engineered in situ feature incorporates ground-truth observations according to a KNN weighting scheme, representing a data assimilation approach (Deshmukh et al., 2016) that fully leverages ongoing in situ sampling efforts.

Finally, we provide a rigorous and comprehensive quantitative evaluation of our model's performance for comparison with existing methods proposed in the literature and applied in practice. Ultimately, there are significant opportunities to operationalize this classifier to advance scientific understanding of bloom dynamics, transport, evolution, and drivers; support efforts to mitigate blooms before they become too large to manage; guide coastal bloom monitoring programs; and forecast potential effects on wildlife, public health, and economic indicators. In addition, our novel approach to model development and rigorous evaluation of model performance has strong potential for application to the detection and monitoring of other HAB species.

## METHODS

### *Data*

This section outlines the datasets used in this work, including MODIS satellite imagery, in situ *K. brevis* cell count data from the Florida Fish and Wildlife Conservation Commission (FWC), and bedrock depth maps from NOAA's ETOPO1 model. These data sources were used to create model input features and the target variables to be predicted.

*MODIS-Aqua satellite imagery.* Satellite imagery was obtained from the MODIS-Aqua sensor (NASA Goddard Space Flight Center, 2021), chosen because it offers several desirable properties for this work. Specifically, MODIS-Aqua provides data every 1–2 days due to the revisit time of the satellite and the wide swath width of the sensor. It also has the requisite wavelength bands for computing fluorescence features that several authors have used for red tide detection (Al Shehhi et al., 2013; Cheng et al., 2009; Hu et al., 2005). Another benefit of the MODIS-Aqua satellite is that it launched in 2002, providing a long record of historical data for model training relative to other available sensors (e.g., Sentinel, Landsat-8). For this work, all MODIS-Aqua images from July 2002 to June 2021 covering Florida were obtained. The Florida region was defined between 23.98° N to 31.92° N and 79.15° W to 88.64° W. This region and time period corresponded to 12 291 satellite image granules.

MODIS-Aqua level 2 ocean color products were used because they have already been processed to compensate for atmospheric effects (Table 2). These products are provided with a spatial resolution of 1 km at nadir. Further details on the MODIS-Aqua level 2 products and the algorithms that are used to produce them can be found in the NASA documentation (NASA Goddard Space Flight Center, 2021).

*Red tide in situ samples.* In situ red tide sample data (*K. brevis* concentrations) were obtained from FWC's HAB Monitoring Database (Florida Fish and Wildlife Conservation Commission, 2021a). The FWC bins data into five different classification groups: 0–1000 cells/L, 1000–10 000 cells/L,

**TABLE 2** The top seven MODIS-Aqua level-2 ocean color features ranked according to recursive feature elimination (RFE)

| Abbreviated feature name | Full feature name | RFE ranking |
|---|---|---|
| *nflh* | Normalized fluorescence line height | 1 |
| $Rrs_{488}$ | Remote sensing reflectance at 488 nm | 2 |
| $Kd_{490}$ | Diffuse attenuation coefficient at 490 nm | 3 |
| *par* | Photosynthetically available radiation | 4 |
| $Rrs_{469}$ | Remote sensing reflectance at 469 nm | 5 |
| $chlor_a$ | Chlorophyll concentration, ocean color index (OCI) algorithm | 6 |
| $Rrs_{443}$ | Remote sensing reflectance at 443 nm | 7 |

 

10 000–100 000 cells/L, 100 000–1 000 000 cells/L, and 1 000 000+ cells/L. To train our models, we use the same classification groups, but we split the 10 000–100 000 cells/L group into two (10 000–50 000 cells/L and 50 000–100 000 cells/L). This allows for direct comparison with binary classification models that use a threshold of 50 000 cells/L.

As described in FWC's documentation, sampling intensity has varied over time depending on the availability of funding since the monitoring program began in 1954. Monitoring includes both routine and event-based sampling. The latter—sampling during red tide blooms in order to guide management decisions—introduces the potential for an imbalanced dataset due to bias toward high *K. brevis* concentrations. However, FWC has increased its routine sampling effort since 2000, and we account for the imbalanced-data problem in the Experimental Design and Evaluation Metrics Section.

From the FWC HAB Monitoring Database, we obtained a dataset of 89 608 red tide in situ samples from January 2000 through May 2020. These samples were spread along the southwest coast of Florida, with 9054 samples from Pinellas county, 2840 samples from Hillsborough county, 6999 samples from Manatee county, 30 105 samples from Sarasota county, 5666 samples from Charlotte county, 16 635 samples from Lee county, 9430 samples from Collier county, and 8879 samples from Monroe county. Figure 1 shows the locations of in situ samples that were successfully paired with MODIS-Aqua satellite data. Details about the pairing are described in the Dataset Preparation Section.

*ETOPO1 model for water depth.* The NOAA's ETOPO1 Global Relief Model (Amante & Eakins, 2009; NOAA National Centers for Environmental Information, 2022) provides an estimate of bedrock height to the nearest meter at a 1 arc-minute resolution (approximately 1.9 km). Each pixel in the MODIS satellite images was paired with the closest pixel from the ETOPO1 model as a proxy for water depth. Water depths in the study area range from sea level to 3562 m deep. However, the majority of FWC in situ samples have been collected in shallow water near the coastline; 90% of the in situ samples were collected at depths of 22 m or less.

*Dataset preparation.* This section describes dataset preparation for the training of several ML models. We followed a process similar to Hill et al. (2020) to extract relevant features from the MODIS-Aqua imagery, including RS reflectance, chlorophyll-*a*, and others, and the ETOPO1 Model.

Specifically, for each in situ sample, we constructed a 100 km × 100 km grid (101 × 101 pixels) centered on the sample location. MODIS-Aqua images from up to 10 days prior to the in situ sample date were used to fill the pixel grid by linearly interpolating each image to the grid and then averaging the gridded data over the 10-day window. The usage of level 2 data products means that all pixels containing significant cloud cover or sun glint were excluded from this analysis. This temporal averaging scheme was adopted to mitigate missing values in any one satellite image (e.g., due to cloud cover or noncoverage by the satellite). In addition, each grid pixel was matched with the nearest value from the NOAA ETOPO1 model for depth.

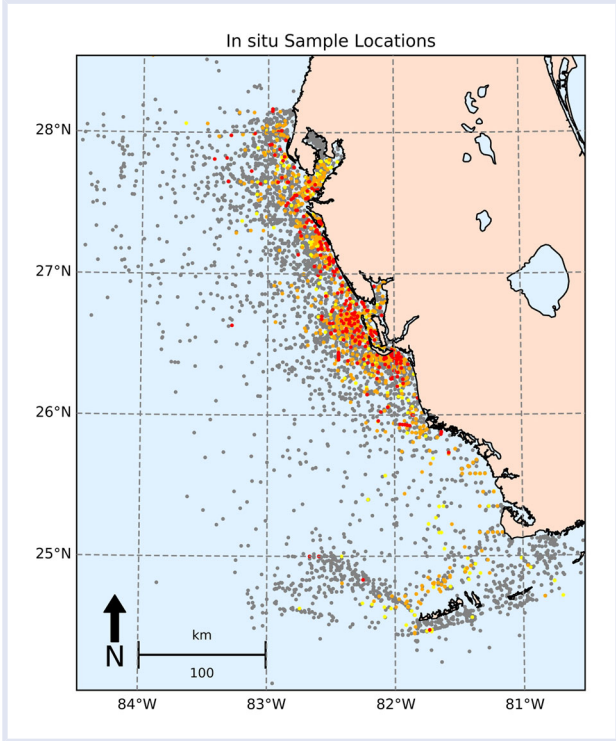


FIGURE 1 In situ sample locations with valid matching satellite data. This region forms the study area of this work and the samples shown are those used in the quantitative evaluation presented. The color of the dots indicates the in situ samples' cell concentration following Florida Fish and Wildlife Conservation Commission's categories: Gray is background (<1000 cells/L), white is very low (between 1000 and 10 000 cells/L), yellow is low (between 10 000 and 100 000 cells/L), orange is medium (between 100 000 and 1 000 000 cells/L), and red is high (>1 000 000 cells/L)

From the dataset containing 101 × 101 pixel grids, we also created a second dataset by extracting a single pixel (the central pixel) from each grid as well as the associated in situ sample. Most of the models that we consider utilized the central pixel only; the 101 × 101 grid of pixels was only used for comparison with the model from Hill et al. (2020).

### Satellite feature processing

*Feature selection.* An important step in ML development is feature selection, in which we select the features most relevant for prediction and discard the others. This step allows models to focus on the most useful information during training without getting confused by features that have little or no predictive value.

Recursive feature elimination (RFE) (Guyon et al., 2002) was applied to perform feature selection. Recursive feature elimination estimates feature importance according to some metric, removes the least important feature, and then recursively repeats the process until only one feature remains.

 

The metric used in this work to rank features was the Gini importance of a random forest classifier, using the RFE implementation provided in scikit-learn (Pedregosa et al., 2011). Among the MODIS-Aqua level 2 ocean color features, we developed models using the seven top-ranked features from the 20 candidate features: *nflh*, $Rrs_{488}$, $Kd_{490}$, *par*, $Rrs_{469}$, $chlor_a$, and $Rrs_{443}$ (Table 2).

*Feature processing.* Remotely sensed ocean color features are water depth dependent in optically shallow waters: Reflectance received by a remote sensor is influenced by backscatter and benthic effects in shallow and clear coastal waters (Abbas et al., 2019; Ha et al., 2014). Visualization of the aggregated satellite feature values as a function of water depth illustrates this dependence (Figure 2). The $chlor_a$, $Kd_{490}$, and *nflh* values increase at shallower depths, and the $Rrs_{443}$, $Rrs_{469}$, and $Rrs_{488}$ values decrease at shallower depths. Chlorophyll-*a* feature values are generally higher with greater variability at lower depths, presumably due to additional uncertainties introduced by bottom reflection. Such a spatially heterogeneous distribution poses a challenge to RS methods since the spectral signature of red tide pixels may differ across locations.

To account for depth dependence of the satellite features, we normalized the values of pixels in each feature channel based on the integer-rounded depth provided by the nearest ETOPO1 pixel. Specifically, we depth-normalized the raw color feature values based on the mean $\mu^{\tau}_{feature}$ and standard deviation $\sigma_{feature}{}^{\tau}$ of each discretized depth level $\tau$:

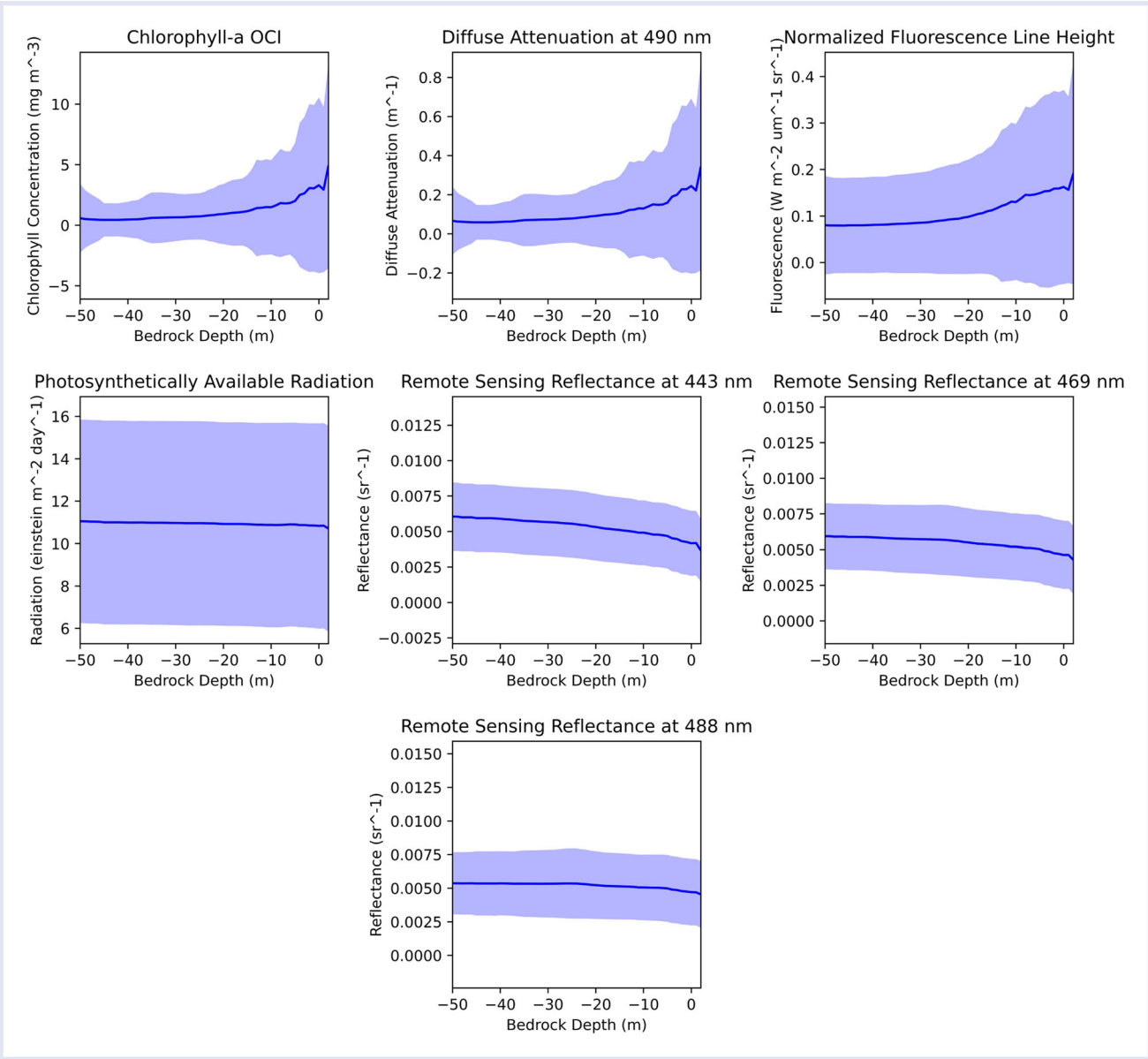


FIGURE 2 Means and standard deviations of MODIS satellite feature values as a function of ETOPO1 water depth; for visual clarity, standard deviations have been divided by 5. To compute these statistics, MODIS feature values from the full study period were paired with integer-rounded depths from the nearest ETOPO1 pixel

$$normalized\ feature = \frac{raw\ feature - \mu^{\tau}_{feature}}{\sigma^{\tau}_{feature}}. \quad (1)$$

An alternative method of using depth would be to include depth directly as an input feature to the neural network model, as was done in Hill et al. (2020). The approach taken here was chosen because the vast majority of the in situ training samples exist in shallow water, as noted in the ETOPO1 Model for Water Depth Section. This makes it challenging for the model to directly learn the relationship of depth with the satellite features.

### *Incorporating historical data with weighted KNNs*

Ongoing *K. brevis* sampling by FWC provides an opportunity to leverage recent in situ data to improve performance. In this section, we describe engineering of a new input feature based on recent in situ observations.

Because FWC red tide data are generally available three days after sample collection, we can use recent in situ sample data—from three to 10 days prior to the current date—to engineer a feature that will potentially improve the prediction of a pixel's bloom state by providing additional spatiotemporal context. We developed a weighted KNN regression model that weights all available in situ samples (from three to 10 days prior to the current date) according to their distance from the target location (the current pixel to be classified) in both space and time. This approach encodes our knowledge that recent, nearby samples should be more highly correlated than spatially and temporally distant samples. Equation (2) shows how the distance function is set up to use both spatial and temporal information, where $d(i,j)$ is the geodesic distance computed by the method from Karney (2013). For all experiments in this article, $\beta$ is set to 1.

$$spDistance(i, j) = d(i, j) + \beta \times (date_i - date_j), \quad (2)$$

$$d(i, j) = geodesic(latitude_i, longitude_i, latitude_j, longitude_j). \quad (3)$$

In order to facilitate training, we first rescaled the *K. brevis* concentration values, following Qin et al. (2017). Bloom magnitudes are conventionally expressed as log-transformed concentrations (from 0 to over $10^6$ cells/L), and we rescaled the concentration values on [0,1] (Equation 4).

$$\log Conc = \frac{\log_{10} conc}{\max \log_{10} conc}. \quad (4)$$

Distances were converted into inverse distances so that closer neighbors are more heavily weighted (Equation 5). The KNN weights were computed by dividing each sample's inverse distance by the sum of inverse distances across samples (Equation 6). The final KNN feature value is the sum of the products of the weights and the corresponding log-concentration values (Equation 4) across all in situ samples in the three- to 10-day window (Equation 7). Algorithm 1 details the full algorithm for the weighted KNN approach.

$$inverseDistance_i = \frac{1}{spDistance(i, j)}, \quad (5)$$

$$weight_i = \frac{inverseDistance_i}{\sum inverseDistance_i}, \quad (6)$$

$$estimatedConc = \sum weight_i \times \log Conc_i. \quad (7)$$

**Algorithm 1** *K-Nearest Neighbor Regression*

**Data**: Test point $i$: (latitude $lat_i$, longitude $lon_i$, $date_i$); Neighbors $j$: (latitudes $lat_j$, longitudes $lon_j$, dates $date_j$, and *K. brevis* concentrations $conc_j$ of water samples)

**Input**: Spatiotemporal weighting $\beta$

**Result**: Prediction for *K. brevis* concentration

1. Find subset of $j$ water samples whose dates are 3–10 days before the test point $i$
2. Compute $\log Conc_j = \frac{\log_{10} conc_j}{\max \log_{10} conc_j}$ for all $j$
3. Compute spatial distances $d_{i,j} = geodesic(lat_i, lon_i, lat_j, lon_j)$ for all $j$
4. Compute spatiotemporal distances $spDistance_{i,j} = d_{i,j} + \beta \times (date_i - date_j)$ for all $j$
5. Compute $inverseDistance_{i,j} = \frac{1}{spDistance_{i,j}}$ for all $j$
6. Compute $weight_{i,j} = \frac{inverseDistance_{i,j}}{\sum inverseDistance_{i,j}}$ for all $j$
7. Compute $estimatedConc_i = \sum weight_{i,j} \times \log Conc_j$
8. Return $estimatedConc_i$

### *Neural network architecture and training*

We developed neural networks using the seven satellite features that ranked highest in terms of Gini importance and the depth feature from the NOAA ETOPO1 Global Relief Model (Amante & Eakins, 2009; NOAA National Centers for Environmental Information, 2022), as described in the Feature Selection Section. In addition, we developed networks that accept an additional feature derived from recent in situ data weighted by the KNN algorithm (Incorporating Historical Data with Weighted K Nearest Neighbors Section). These models fuse the information from the RS data with the known ground truth from the recent spatiotemporal region, allowing the model to use both data sources simultaneously. Table 3 shows the full feature set used for neural network training.

Two different neural network architectures were explored in this work. The first is a multilayer perceptron that classifies individual pixels—the "central pixel neural network"—that is, for each image of 101 × 101 pixels, the model classifies the central pixel or the pixel closest to the in situ sample. A figure showing the model architecture is provided in the Supporting Information.

 

**TABLE 3** Full feature set used to train neural network models

| Abbreviated feature name | Full feature name |
|---|---|
| *nflh* | Normalized fluorescence line height |
| $Rrs_{488}$ | Remote sensing reflectance at 488 nm |
| $Kd_{490}$ | Diffuse attenuation coefficient at 490 nm |
| *par* | Photosynthetically available radiation |
| $Rrs_{469}$ | Remote sensing reflectance at 469 nm |
| $chlor_a$ | Chlorophyll concentration, OCI algorithm |
| $Rrs_{443}$ | Remote sensing reflectance at 443 nm |
| *depth* | Bathymetry from the NOAA ETOPO1 model |
| *knn* | (Optional) Spatiotemporal estimate of *K. brevis* concentration from the KNN algorithm |

Abbreviations: KNN, K-nearest neighbor; OCI, ocean color index.

This network outputs six values that indicate the network's prediction that the input pixel falls into one of six classes: 0 to 1000 cells/L, 1000 to 10 000 cells/L, 10 000 to 50 000 cells/L, 50 000 to 100 000 cells/L, 100 000 to 1 000 000 cells/L, and 1 000 000+ cells/L. These classes follow the abundance classifications used by FWC (2021b), with an additional threshold at 50 000 cells/L to allow comparison with other models that use that threshold.

The other model that we considered was a convolutional neural network (figure provided in the Supporting Information), whose structure closely follows Hill et al. (2020). Images of $101 \times 101$ pixels are given to the NASNet-Mobile network, which outputs a feature layer of size $1056 \times 7 \times 7$. Following the procedure in Hill et al. (2020), we extracted the central $3 \times 3$ region to obtain a $1056 \times 3 \times 3$ layer, and then flattened that layer to a set of 9504 features. Those features were then passed through fully connected layers to produce the classification output.

This model outputs two values indicating the model's prediction that the input pixel falls into one of two classes: 0 to 50 000 cells/L or 50 000+ cells/L. These classes follow the categorization used in Hill et al. (2020).

We experimented with the optional KNN feature in the convolutional model. This feature was added in the final layer of the model after the satellite and depth features had been aggregated by the convolutional layers. The KNN feature aggregates spatiotemporal information from the surroundings and is coupled with the convolutional network, which aggregates contextual information in the satellite data. Therefore, one of the questions that we explore in the experiment section is whether the KNN feature can add information that is not already present in the convolution.

### *Ablation study*

We performed an ablation study to examine the predictive value of the depth-normalization procedure (for satellite features) and the engineered KNN feature, in both the central pixel model and the convolutional model. Accordingly, we ran each of these models without the KNN feature or depth normalization, with the KNN feature only, and with both the KNN feature and the depth normalization. Each configuration was trained and tested 20 times (using a different year as the test set).

### *Implementation of baseline methods from the literature*

We implemented existing methods from the literature (Table 1) and provide the implementation details below for reproducibility. Source code is available on Github (https://github.com/CenterForCoastalSolutions/red-tide-conv). Most of these existing methods rely on traditional RS products, but the convolutional neural network in Hill et al. (2020) is the current state-of-the-art (described in the Neural Network Architecture and Training Section). The traditional RS products are described below. For any methods that use a wavelength band not present in the MODIS sensor, we used the closest available band.

From Amin et al. (2009), we compared against two methods. The first used the authors' proposed Red Band Difference (RBD), based on the normalized water-leaving radiance feature, *nLw*. We adopted the authors' suggested threshold $RBD > 0.15\,\mathrm{W/m^2/\mu m/sr}$.

$$RBD = nLw(678) - nLw(667). \tag{8}$$

The second method from Amin et al. (2009) used the RBD as well as the *K. brevis* bloom index (KBBI). We adopted the authors' suggested thresholds of $RBD > 0.15\,\mathrm{W/m^2/\mu m/sr}$ and $KBBI > 0.3 \times RBD$.

$$KBBI = \frac{nLw(678) - nLw(667)}{nLw(678) + nLw(667)}. \tag{9}$$

Cannizzaro et al. (2008) classified *K. brevis* blooms based on $chlorophyll > 1.5\,\mathrm{mg/m^3}$ and $b_{bp}(550)$ values less than the Morel (1988) relationship. Cannizzaro et al. (2008) calculated $b_{bp}(550)$ based on Equation (10) from Carder et al. (1999), and the Morel (1988) relationship refers to $b_{bp}(550)$ as computed by Equation (11).

$$b_{bp}(550) = 2.058 \times R_{rs}(551) - 0.00182, \tag{10}$$

$$b_{bp_{morel}}(550) = 0.30 \times Chl^{0.62} \times (0.002 + 0.02 \times (0.5 - 0.25 \times \log Chl)). \tag{11}$$

Similarly, Cannizzaro et al. (2009) detected *K. brevis* blooms based on $chlorophyll > 1.5\,\mathrm{mg/m^3}$ and $b_{bp}(550) \leq 0.0045\,\mathrm{m^2/mg}$.

Al Shehhi et al. (2013) proposed a method based on *nFLH*. The authors did not suggest a threshold value for $nFLH_{thresh}$, so where necessary, we adopted $nFLH_{thresh} = 0$.

Lou and Hu (2014) proposed a red tide index *RI* (Equation 12), modified from the original red tide index proposed by Ahn and Shanmugam (2006). We applied the authors' suggested threshold of $RI > 2.8$.

 

$$RI = \frac{R_{rs}(555) - R_{rs}(443)}{R_{rs}(490) - R_{rs}(443)}. \quad (12)$$

Stumpf et al. (2003) is based on chlorophyll anomalies, defined as the difference in chlorophyll value in a single image compared to the mean chlorophyll value in the period from two months prior to two weeks prior to the image date. We adopted a chlorophyll anomaly threshold of >0.5 μg/L as that which corresponds to a *K. brevis* concentration of 50 000 cells/L.

Tomlinson et al. (2009) proposed a spectral shape method as defined by Equation (13) and the threshold $SS(490) < 0$.

$$SS(490) = nLw(490) - nLw(443) - (nLw(510) - nLw(443)) \times \frac{490 - 443}{510 - 443}. \quad (13)$$

Soto et al. (2021) and Bernard et al. (2021) proposed a method applying thresholds across several features. The method detects *K. brevis* blooms based on the following criteria: $chlorophyll > 1.5\,\mathrm{mg/m^3}$, $nFLH > 0.01\,\mathrm{mW/cm^2/\mu m/sr}$, and $b_{bp}\ ratio < 1$. The $b_{bp}$ ratio refers to the ratio of $b_{bp}$ as computed by the Quasi-Analytical Algorithm from Lee et al. (2002) divided by $b_{bp}$ using the Morel (1988) relationship from Equation (11).

### *Experimental design and evaluation metrics*

We conducted experiments to test each model's ability to detect red tide bloom conditions. To set up the datasets, the data were split into training and test data according to year, rather than randomly, to prevent data in the training and test sets from being close in both space and time, which would potentially enable models to show unrealistically strong predictive performance. For instance, one test set may comprise all data from 2020, and the corresponding training set would comprise all data excluding 2020. As the basis for evaluating each model's performance, models were run 20 times using a different year for the test set.

Among the samples in the potential training set, we selected samples in order to balance the set across *K. brevis* cell concentrations. We randomly selected 350 samples in each of the following six cell count categories: 0–1000 cells/L, 1000–10 000 cells/L, 10 000 to 50 000 cells/L, 50 000–100 000 cells/L, 100 000 to 1 000 000 cells/L, and 1 000 000+ cells/L. These six classes follow the categories used in the central pixel model described in the Neural Network Architecture and Training Section, and as such serve as the labels for training.

The convolutional neural network model described in the Neural Network Architecture and Training Section used two classes: 0–50 000 cells/L and 50 000+ cells/L. To produce labels for this training, we combined the six categories described above to fit these two classes. Specifically, the samples from the 0 to 1000 cells/L category, the 1000 to 10 000 cells/L category, and the 10 000 to 50 000 cells/L category were combined into the 0 to 50 000 cells/L class, and the samples from the 50 000 to 100 000 cells/L category, the 100 000 to 1 000 000 cells/L category, and the 1 000 000+ cells/L category were combined into the 50 000+ cells/L class.

For quantitative evaluation, we used three metrics: classification accuracy, the F1 score, and Cohen's $\kappa$ score (Cohen, 1960). These metrics were computed based upon individual pixels corresponding to field samples. The classification accuracy is computed as the number of correct classifications divided by the total number of classifications:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (14)$$

where *TP* is the number of true positives, *TN* is the number of true negatives, *FP* is the number of false negatives, and *FN* is the number of false negatives.

We also computed the F1 score and Cohen's $\kappa$ score. The F1 score is the harmonic mean of precision and recall (Equations 15–17). Cohen's $\kappa$ score relates the observed classification accuracy $p_o$ to the hypothetical accuracy $p_e$ achievable by random chance (Equation 18). For instance, if 90% of test samples were from nonbloom conditions, a model could predict every test sample as a nonbloom and achieve 90% accuracy by artifact. The $\kappa$ score adjusts the score to account for such bias. Note that for all three of these evaluation metrics, higher scores indicate better performance.

$$F1 = 2 \times \frac{precision \times recall}{precision + recall}, \quad (15)$$

$$precision = \frac{TP}{TP + FP}, \quad (16)$$

$$recall = \frac{TP}{TP + FN}, \quad (17)$$

$$\kappa = \frac{p_0 - p_e}{1 - p_e}. \quad (18)$$

We also evaluated models using receiver operating characteristic (ROC) curves, which visualize performance as a function of the detection threshold (i.e., the minimum output value of the network required to classify a pixel as indicative of bloom conditions). For each threshold, the ROC curve displays the model's false-positive rate (the number of nonbloom pixels incorrectly classified as bloom pixels) and true-positive rate (the number of correctly classified bloom pixels). The top left corner of the ROC plot represents perfect performance (100% true-positive rate with 0% false-positive rate), and a dashed 1:1 line indicates performance equivalent to guessing (i.e., true-positive rate equals false-positive rate).

## RESULTS

To illustrate the proposed model's output in practice, we present predicted red tide maps over the course of a red tide bloom event that occurred off the southwest coast of

 

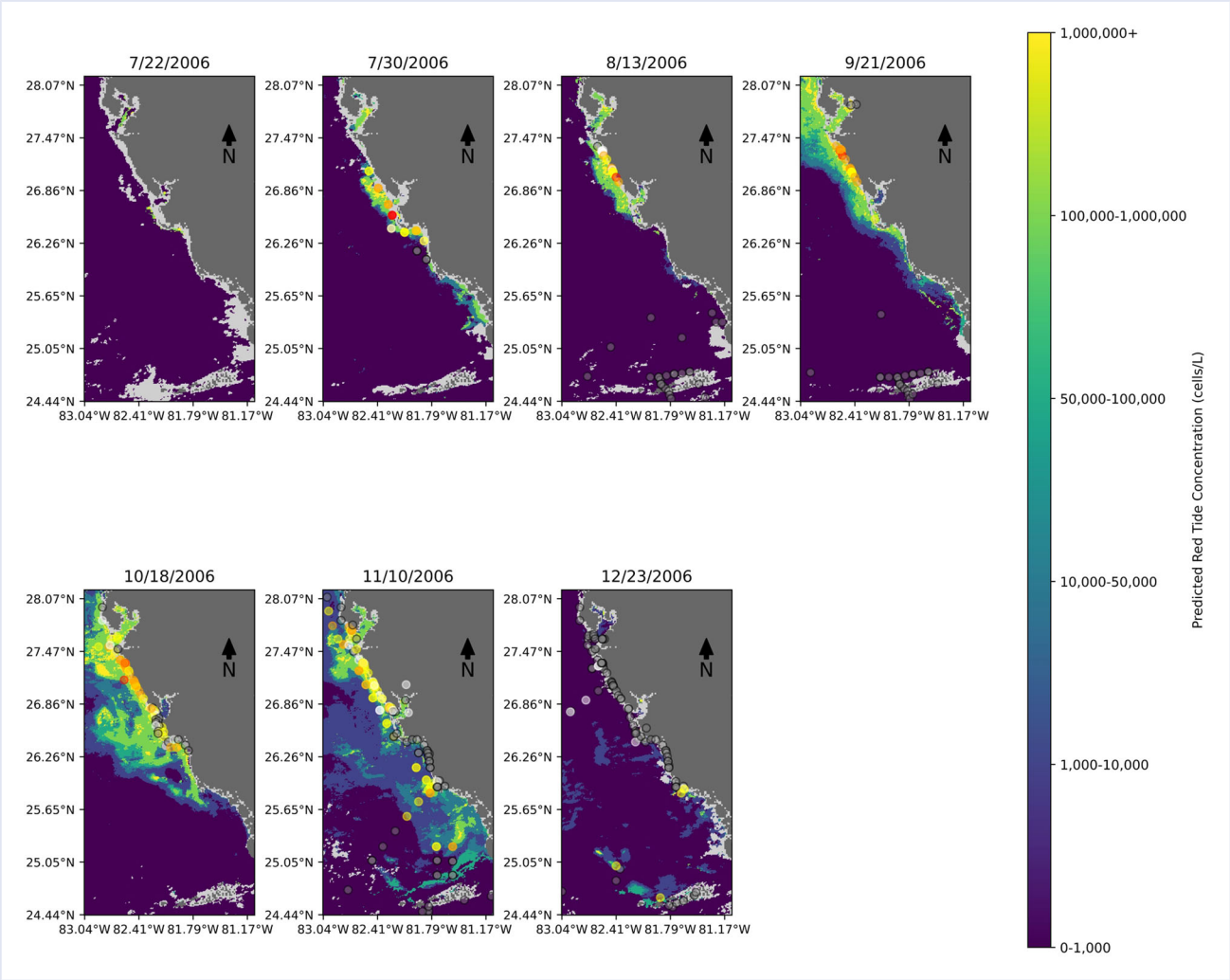


FIGURE 3 Example of model output during the 2006 *Karenia brevis* bloom event off the southwest coast of Florida. Maps visualize the predicted red tide concentration classification at each pixel with a blue-to-yellow color gradient. The dark gray pixels indicate land and light gray pixels indicate missing values due to persistent cloud cover or sun glint. Imagery was averaged over 10 days, matching the training of the model. Circles indicate in situ samples collected up to one week prior to the displayed date, and observed concentration values are indicated by color, following Florida Fish and Wildlife Conservation Commission's concentration categories: Gray is background (<1000 cells/L), white is very low (between 1000 and 10 000 cells/L), yellow is low (between 10 000 and 100 000 cells/L), orange is medium (between 100 000 and 1 000 000 cells/L), and red is high (>1 000 000 cells/L)

Florida in 2006 (Figure 3). The maps visualize the predicted *K. brevis* concentration classification at each pixel. For reference, the maps also display in situ data collected during the same period.

We conducted an ablation study to compare model architectures and specifications using a threshold of 50 000 cells/L for binary bloom and/or nonbloom classification. The ablation study indicated that the central pixel neural network outperformed the convolutional neural network whose architecture mirrors the network proposed by Hill et al. (2020). The central pixel model performed best with the inclusion of the KNN feature and the depth-normalization procedure, and this model achieved a mean accuracy of 0.70, an F1 score of 0.67, and $\kappa$ coefficient of 0.38 across replicate runs of the model based on 20 different train/test splits (Table 4).

TABLE 4 Ablation results: Comparison of models with and without the KNN feature and depth-normalized feature processing

| Model | Accuracy | F1 score | $\kappa$ score |
|---|---|---|---|
| Hill et al. (2020) | 0.60 (0.04) | 0.50 (0.08) | 0.19 (0.06) |
| Hill et al. (2020) + KNN | 0.62 (0.05) | 0.53 (0.09) | 0.22 (0.09) |
| Hill et al. (2020) + KNN + depth normalization | 0.62 (0.05) | 0.47 (0.11) | 0.20 (0.06) |
| Central pixel | 0.65 (0.03) | 0.62 (0.09) | 0.29 (0.08) |
| Central Pixel + KNN | 0.68 (0.04) | 0.65 (0.08) | 0.34 (0.08) |
| Central Pixel + KNN + Depth Normalization | 0.70 (0.02) | 0.67 (0.09) | 0.38 (0.05) |

*Note*: Performance metrics are expressed as mean (standard deviation) values across 20 replicate training and/or testing runs for each model.
Abbreviation: KNN, K-nearest neighbor.

TABLE 5 Performance metrics for the proposed model and existing models from the literature

| Model | Accuracy | F1 score | $\kappa$ score |
|---|---|---|---|
| Central Pixel + KNN + Depth Normalization (our method) | 0.70 (0.02) | 0.67 (0.09) | 0.38 (0.05) |
| Amin et al. (2009) RBD + KBBI | 0.65 (0.05) | 0.48 (0.09) | 0.26 (0.06) |
| Amin et al. (2009) RBD | 0.65 (0.05) | 0.48 (0.09) | 0.26 (0.06) |
| Soto et al. (2021), Bernard et al. (2021) | 0.64 (0.03) | 0.61 (0.07) | 0.26 (0.06) |
| Tomlinson et al. (2009) | 0.63 (0.05) | 0.65 (0.08) | 0.26 (0.08) |
| Lou and Hu (2014) | 0.62 (0.03) | 0.52 (0.08) | 0.22 (0.06) |
| Hill et al. (2020) | 0.60 (0.04) | 0.50 (0.08) | 0.19 (0.06) |
| Cannizzaro et al. (2008) | 0.59 (0.05) | 0.45 (0.10) | 0.16 (0.08) |
| Stumpf et al. (2003) | 0.56 (0.06) | 0.65 (0.08) | 0.13 (0.06) |
| Cannizzaro et al. (2009) | 0.55 (0.08) | 0.25 (0.09) | 0.06 (0.05) |
| Al Shehhi et al. (2013) | 0.49 (0.08) | 0.64 (0.09) | 0.02 (0.03) |

*Note*: Performance metrics are expressed as mean (standard deviation) values across 20 replicate training and/or testing runs for each model.
Abbreviations: KNN, K-nearest neighbor; RBD, Red Band Difference.

Our central pixel model also outperformed existing methods in the literature (Table 5). The methods of Soto et al. (2021), Bernard et al. (2021), and Tomlinson et al. (2009) achieved the next-highest performance when considering all three metrics, with mean accuracies of 0.64 and 0.63, mean F1 scores of 0.61 and 0.65, and mean $\kappa$ coefficients of 0.26 and 0.26, respectively. ROC curves show the mean performance of each model (across 20 runs) and indicate dominance of the central pixel model across all detection thresholds (Figure 4).

Additionally, we show a confusion matrix of the proposed method's performance against unseen text pixels in Figures 5 and 6. These two figures show that the confusion matrix first normalized by column and then by row. This allows insight into how the model's predicted labels correspond to true labels, and vice versa.

## DISCUSSION

Quantitative comparisons of model performance indicated that our central pixel neural network, including the KNN feature and depth-normalized satellite features, offers substantially improved detection of red tide blooms compared to existing methods (Tables 4 and 5). This work leveraged valuable, long-term datasets of in situ sampling of *K. brevis* and 20 years of MODIS-Aqua satellite imagery—datasets that have required significant and long-term state (FWC) and federal agency investment to maintain and curate. Further, we introduced two main innovations by normalizing satellite features to water depth and engineering a new KNN feature that makes use of the ongoing availability of in situ sample data. Importantly, this approach can be feasibly and efficiently operationalized to support the detection and tracking of red tide along Florida's Gulf Coast

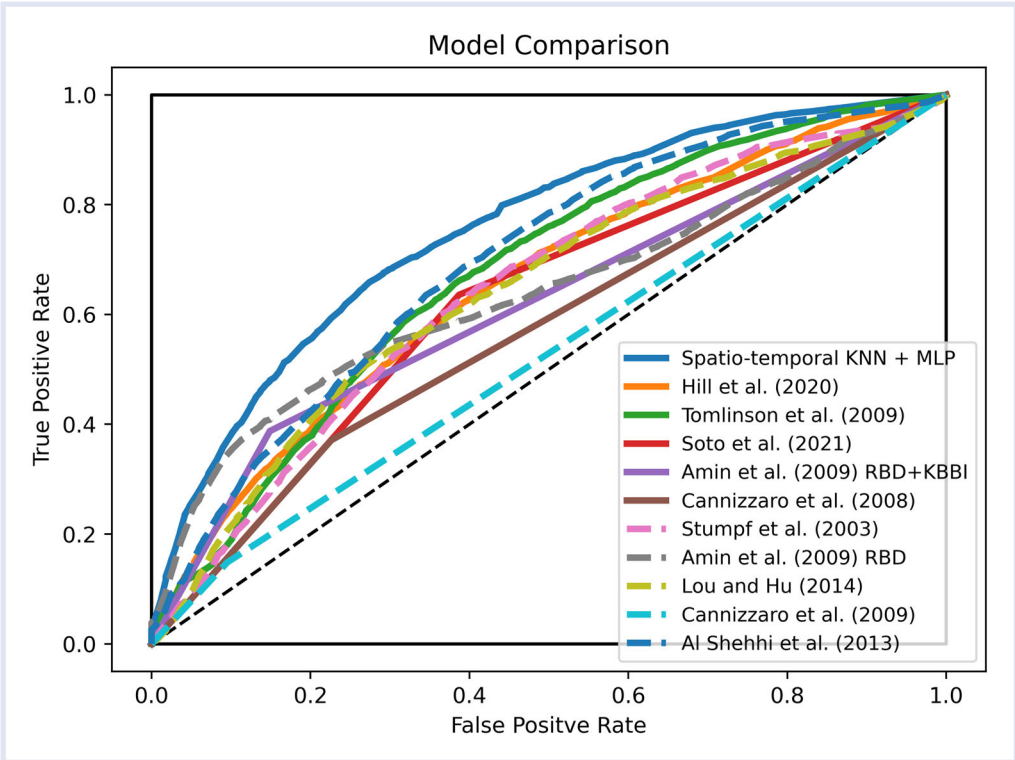


FIGURE 4 Receiver operating characteristic (ROC) curves for the proposed model and existing models from the literature. Each curve shows the specified model's mean ROC curve across 20 replicate training and/or testing runs. Vertical and horizontal axes indicate the true-positive rate and false-positive rate, respectively. KBBI, *Karenia brevis* bloom index; KNN, K-nearest neighbor; RBD, Red Band Difference

 

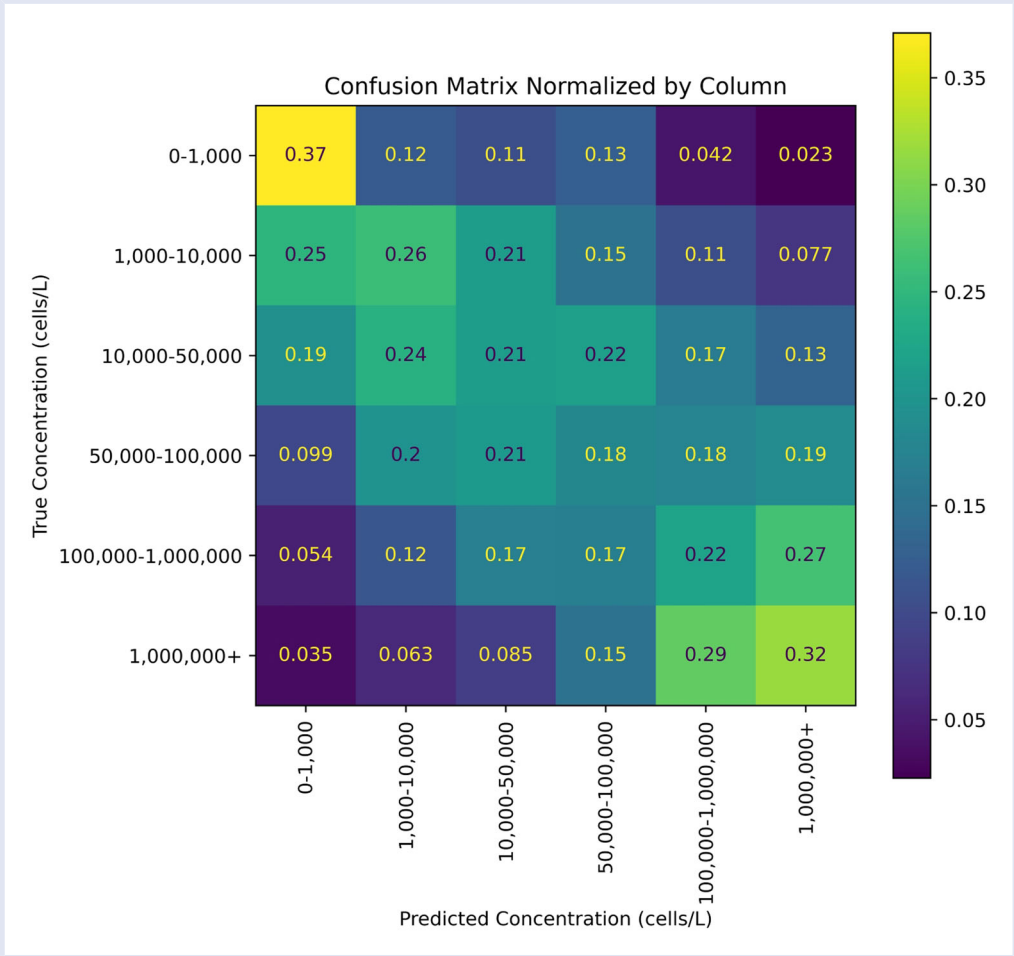


FIGURE 5 Confusion matrix normalized by column to show how each predicted cell count concentration class corresponds to the true classes. Values on the upper-left to lower-right diagonal line represent positive predictive values, or precision. Other values represent the false discovery rate associated with each individual class

and adapted to support the detection of algal blooms in other coastal areas characterized by heterogeneous water depths. Further, our rigorous evaluation of model performance relative to others presented in the literature and currently in applied in practice establishes a clear and consistent benchmark against which future studies may compare their accuracy.

### *Additional comments on improved accuracy with baseline methods*

Our review of ML-based red tide detection methods in the literature revealed that quantitative performance metrics are rarely reported. One exception is Hill et al. (2020), who reported better performance than we found based on our implementation of the method (Table 5). Hill et al. (2020) reported that their best-performing model achieved an accuracy of 0.91, an F1 score of 0.88, and a $\kappa$ score of 0.81. A possible explanation for the discrepancy is that these authors discarded samples with *K. brevis* cell concentrations between 1 and 50 000 cells/L. All of the samples in their nonbloom class thus had a labeled concentration of 0 cells/L, and all of the samples in the bloom class had a labeled concentration >50 000 cells/L. Ignoring concentrations between these two values makes the problem of detecting red tide easier, but it is not reflective of realistic conditions. These lower cell counts, discarded by Hill et al. (2020), may represent conditions when blooms are intensifying or abating—critical time periods from a detection and management perspective. Additionally, Hill et al. (2020) did not describe their procedure for splitting the data into training and test sets. If the splits were determined randomly, rather than by year (as we have done), the detection problem is artificially made easier, and the model may perform unrealistically well during testing. Regardless, our approach provides an objective "apples-to-apples" comparison, with substantive improvements realized by our best-performing model.

Inclusion of one of our key innovations—the KNN feature—improved performance of the central pixel model across all metrics (Table 4). This improvement was also observed in our implementation of the convolutional model proposed by Hill et al. (2020). The intent for the KNN feature was to aggregate information across a broader spatiotemporal context, but one might expect such context to be built into a convolutional network (without an in situ feature) due to aggregation across a large area. The KNN feature nonetheless provided additional useful information.

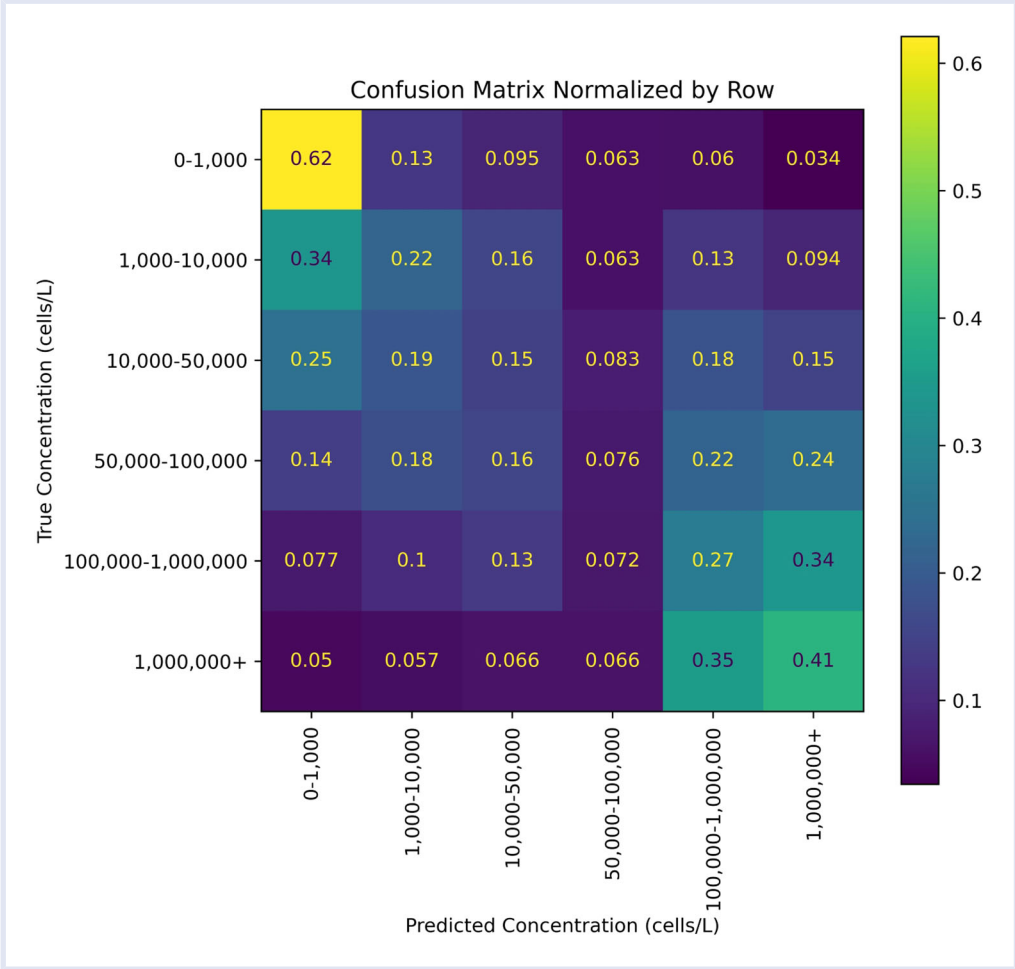


FIGURE 6 Confusion matrix normalized by row to show how each true cell count concentration class is predicted into different classes. Values on the upper-left to lower-right diagonal line represent true-positive rates, or sensitivity. Other values represent the false-negative rate associated with each individual class

Additionally, using CNN over a $101 \times 101$ window, as suggested by Hill et al. (2020), resulted in lower accuracy and a lower $\kappa$ coefficient than using the central pixel only (Table 4). One possible explanation is that using such a large geographic area to classify each satellite pixel can result in an oversmoothing effect due to high spatial correlation across windows.

### *Scope and limitations of our method*

The spatial footprint of red tide blooms is sometimes patchy, as discussed in Tomlinson et al. (2009). Since we fused in situ samples with low-resolution satellite pixels using KNNs, it is possible that our method might not adequately capture the subtle patterns of spatial and temporal variations (e.g., patchiness) of a bloom. The issue could be mitigated by denser in situ samples and satellite pixels with higher spatial and temporal resolutions. In addition, we focused on detecting the *K. brevis* concentration per pixel. Future work could evaluate the method for distinguishing red tide blooms from other nontoxic blooms (El-Habashi et al., 2016). In particular, work needs to be done to understand how microalgae may interact with *K. brevis* blooms to influence the resulting satellite signal, as data for such analysis become available. Finally, although ML approaches are adaptive to complex scenarios when sufficient amounts of observed data are available for training, a neural network model is a black box. As expressed in the literature, there is a growing need for explainability (Abdollahi & Pradhan, 2021; Kakogeorgiou & Karantzalos, 2021), and in particular, it would be advantageous to anticipate novel conditions under which our classifier would fail (Gawlikowski et al., 2022; Inkawhich et al., 2022).

### *Application for real-world red tide detection*

The 2006 case study demonstrates the utility of our approach in detecting and tracking a red tide bloom (Figure 3). Specifically, the image from 7/22/2006 shows the model prediction prior to bloom formation and correctly identifies a lack of bloom conditions (though we note that no in situ samples were collected during that week). The images from 7/30/2006 and 8/13/2006 show the early stages of bloom formation. Specifically, in the 7/30/2006 image, the model

 

identified the bloom appearing off the coast of Fort Myers, Florida. It also indicated the presence of a bloom further south along the coast. By 8/13/2006, samples from that southern area observed zero/background *K. brevis* cell counts, and the model predicted no *K. brevis* in that area. As one moves forward in time, peak bloom conditions become apparent in the images from 9/21/2006, 10/18/2006, and 11/10/2006, and visually, the model predictions appear to agree with the in situ data. The obvious advantage of using the satellite imagery is that it covers a larger spatial domain than the in situ samples alone. For instance, the model identified bloom conditions near Tampa Bay in the 9/21/2006 image, prior to in situ samples in that area, and later samples in the 10/18/2006 and 11/3/2006 images confirmed the presence of bloom conditions. Finally, the 12/14/2006 image shows the termination of bloom conditions in Fort Myers and the Tampa Bay regions, at least temporarily, as bloom conditions would reintensify in the spring of 2007. The model does identify small patches of potential blooms near the Florida Keys. Sampling in the area over the subsequent weeks would confirm bloom conditions.

Taken together, this 2006 case study demonstrates the utility of our approach in capturing the spatial and temporal evolution of a red tide bloom along Florida's heterogeneous and shallow coast. Indeed, the model was able to accurately detect the early appearance of the bloom at the surface, its distribution during peak bloom conditions, and its abatement in December. Access to such reliable, near real-time information about where blooms are located may accelerate our understanding of blooms' physical, biological, and chemical drivers and their effects on wildlife, water quality, economic activity, and public health. In addition, highly resolved maps may support targeted public health advisories and proactive management and mitigation measures.

The confusion matrices shown in Figures 5 and 6 have implications for the usefulness of this model in a management and response context. When faced with examples that are high concentration (full-blown bloom conditions) or very low concentration (nonbloom conditions), the model's predictions are very accurate. in intermediate bloom levels, confidence in the model's predictions should be lower (i.e., differentiating the 1000–10 000 cells/L group from the 10 000–50 000 cells/L group).

Future work can address why the model struggles to resolve the intermediate bloom levels. A possible hypothesis as to why these examples are so challenging is that they may reside on the margins of established blooms, and so the satellite data are aggregating information from either inside or outside of the bloom. Another possibility is that these intermediate pixels reside in blooms that are either emerging or dying off and the satellite information may not be capturing current conditions accurately.

## CONCLUSION

We developed a novel ML-based method for detecting red tide (*K. brevis*) blooms off the Florida Gulf Coast at a 1-km spatial resolution. The method offers improved performance over state-of-the-art red tide detection methods and consists of a neural network that leverages seven depth-normalized ocean color features (MODIS) as well as an engineered feature based on recent in situ sample data aggregated by a straightforward KNN weighting scheme. The depth-normalization procedure adjusted for backscatter and reflectance effects in shallow coastal waters, and the KNN feature encoded additional spatiotemporal context not provided by the satellite data. This model demonstrates potential for advancing scientific understanding of bloom dynamics and drivers, enhancing red tide management and mitigation efforts in near-real time, and adaptability for application to other types of algal blooms.

## AUTHOR CONTRIBUTION

**Ronald Fick**: Conceptualization; data curation; formal analysis; methodology; software; validation; visualization; writing—original draft; writing—review and editing. **Miles Medina**: Data curation; visualization; writing—review and editing. **Christine Angelini**: Funding acquisition; project administration; resources; supervision; writing—review and editing. **David Kaplan**: Writing—review and editing. **Paul Gader**: Data curation; funding acquisition; methodology; resources; supervision; writing—review and editing. **Wenchong He**: Methodology. **Zhe Jiang**: Methodology; project administration; supervision; writing—review and editing. **Guangming Zheng**: Methodology; writing—review and editing.

## ACKNOWLEDGMENT

The authors acknowledge the following funding sources: Florida Department of Environmental Protection (LPR0017), NSF Career Award (1652628) to CA, and philanthropic support to the University of Florida Center for Coastal Solutions. The author list covers all effort of this work.

## CONFLICT OF INTEREST

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

All data used in this work are publicly available and can be downloaded from its respective owner. The MODIS satellite data can be accessed at: https://oceancolor.gsfc.nasa.gov/data/find-data/. FWRI's in situ red tide database can be accessed at: https://myfwc.com/research/redtide/monitoring/database/. The ETOPO1 bathymetry data can be accessed at: https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ngdc.mgg.dem:316. All code used to process these data is available at our repository here: https://github.com/CenterForCoastalSolutions/red-tide-conv.

## ORCID

Ronald Fick http://orcid.org/0000-0002-4201-2304
Miles Medina http://orcid.org/0000-0002-4264-6307

 

## SUPPORTING INFORMATION

Graphic descriptions of the two neural network architectures that were used in the work.

## REFERENCES


Abbas, M. M., Melesse, A. M., Scinto, L. J., & Rehage, J. S. (2019). Satellite estimation of chlorophyll-a using moderate resolution imaging spectroradiometer (MODIS) sensor in shallow coastal water bodies: Validation and improvement. *Water*, *11*, 1621.

Abdollahi, A., & Pradhan, B. (2021). Urban vegetation mapping from aerial imagery using explainable AI (XAI). *Sensors*, *21*, 4738.

Ahn, Y. H., & Shanmugam, P. (2006). Detecting the red tide algal blooms from satellite ocean color observations in optically complex Northeast-Asia Coastal waters. *Remote Sensing of Environment*, *103*, 419–437.

Al Shehhi, M. R., Gherboudj, I., Zhao, J., Mezhoud, N., & Ghedira, H. (2013). Evaluating the performance of MODIS FLH ocean color algorithm in detecting the Harmful Algae Blooms in the Arabian Gulf and the Gulf of Oman. In *2013 OCEANS-San Diego* (pp. 1–7). IEEE. https://ieeexplore.ieee.org/abstract/document/6741163

Amante, C., & Eakins, B. W. (2009). *Etopo1 1 arc-minute global relief model: Procedures, data sources and analysis*. NOAA Technical Memorandum NESDIS NGDC-24.

Amin, R., Zhou, J., Gilerson, A., Gross, B., Moshary, F., & Ahmed, S. (2009). Novel optical techniques for detecting and classifying toxic dinoflagellate *Karenia brevis* blooms using satellite imagery. *Optics Express*, *17*, 9126–9144.

Anderson, D. M., Fensin, E., Gobler, C. J., Hoeglund, A. E., Hubbard, K. A., Kulis, D. M., Landsberg, J. H., Lefebvre, K. A., Provoost, P., Richlen, M. L., Smith, J. L., Solow, A. R., & Trainer, V. L. (2021). Marine harmful algal blooms (HABs) in the United States: History, current status and future trends. *Harmful Algae*, *102*, 101975.

Bechard, A. (2021). Gone with the wind: Declines in property values as harmful algal blooms are blown towards the shore. *The Journal of Real Estate Finance and Economics*, *62*, 242–257.

Bernard, S., Kudela, R., Robertson Lain, L., & Pitcher, G. (2021). *Observation of harmful algal blooms with ocean colour radiometry* (Technical Report 20). International Ocean Colour Coordinating Group.

Cannizzaro, J. P., Carder, K. L., Chen, F. R., Heil, C. A., & Vargo, G. A. (2008). A novel technique for detection of the toxic dinoflagellate, *Karenia brevis*, in the Gulf of Mexico from remotely sensed ocean color data. *Continental Shelf Research*, *28*, 137–158.

Cannizzaro, J. P., Hu, C., English, D. C., Carder, K. L., Heil, C. A., & Muller-Karger, F. E. (2009). Detection of *Karenia brevis* blooms on the west Florida shelf using in situ backscattering and fluorescence data. *Harmful Algae*, *8*, 898–909.

Carder, K. L., Chen, F., Lee, Z., Hawes, S., & Kamykowski, D. (1999). Semianalytic moderate-resolution imaging spectrometer algorithms for chlorophyll *a* and absorption with bio-optical domains based on nitrate-depletion temperatures. *Journal of Geophysical Research: Oceans*, *104*, 5403–5421.

Carvalho, G. A., Minnett, P. J., Fleming, L. E., Banzon, V. F., & Baringer, W. (2010). Satellite remote sensing of harmful algal blooms: A new multi-algorithm method for detecting the Florida Red Tide (*Karenia brevis*). *Harmful Algae*, *9*(5), 440–448. https://doi.org/10.1016/j.hal.2010.02.002

Cheng, W., Hall, L. O., Goldgof, D. B., Soto, I. M., & Hu, C. (2009). Automatic red tide detection from MODIS satellite images. In *2009 IEEE International Conference on Systems, Man and Cybernetics* (pp. 1864–1868). IEEE. https://ieeexplore.ieee.org/abstract/document/5346189

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Court, C., Ferreira, J., Ropicki, A., Qiao, X., & Saha, B. (2021). *Quantifying the socio-economic impacts of harmful algal blooms in Southwest Florida in 2018* (Technical report). Food and Resource Economics Department, University of Florida.

de Araujo Carvalho, G. (2008). *The use of satellite-based ocean color measurements for detecting the Florida Red Tide (Karenia brevis)* [Doctoral thesis, University of Miami].

Deshmukh, A. N., Deo, M., Bhaskaran, P. K., Nair, T. B., & Sandhya, K. (2016). Neural-network-based data assimilation to improve numerical ocean wave forecast. *IEEE Journal of Oceanic Engineering*, *41*, 944–953.

El-Habashi, A., Ioannou, I., Tomlinson, M. C., Stumpf, R. P., & Ahmed, S. (2016). Satellite retrievals of *Karenia brevis* harmful algal blooms in the West Florida shelf using neural networks and comparisons with other techniques. *Remote Sensing*, *8*, 377.

Florida Fish and Wildlife Conservation Commission. (2021a). *HAB monitoring database* [Dataset]. Retrieved August 17, 2021. https://myfwc.com/research/redtide/monitoring/database/

Florida Fish and Wildlife Conservation Commission. (2021b). *Red tide current status* [Dataset]. Retrieved August 17, 2021. https://myfwc.com/research/redtide/statewide/

Gawlikowski, J., Saha, S., Kruspe, A., & Zhu, X. X. (2022). An advanced Dirichlet prior network for out-of-distribution detection in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–19.

Gokaraju, B., Durbha, S. S., King, R. L., & Younan, N. H. (2011). A machine learning based spatio-temporal data mining approach for detection of harmful algal blooms in the gulf of Mexico. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *4*(3), 710–720. https://doi.org/10.1109/jstars.2010.2103927

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, *46*, 389–422.

Ha, N. T. T., Koike, K., & Nhuan, M. T. (2014). Improved accuracy of chlorophyll-a concentration estimates from MODIS imagery using a two-band ratio algorithm and geostatistics: As applied to the monitoring of eutrophication processes over Tien Yen Bay (Northern Vietnam). *Remote Sensing*, *6*, 421–442.

Heil, C. A., Dixon, L. K., Hall, E., Garrett, M., Lenes, J. M., O'Neil, J. M., Walsh, B. M., Bronk, D. A., Killberg-Thoreson, L., Hitchcock, G. L., Meyer, K. A., Mulholland, M. R., Procise, L., Kirkpatrick, G. J., Walsh, J. J., & 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 multi-year regional study. *Harmful Algae*, *38*, 127–140.

Heil, C. A., & Steidinger, K. A. (2009). Monitoring, management, and mitigation of *Karenia* blooms in the eastern Gulf of Mexico. *Harmful Algae*, *8*, 611–617.

Hill, P. R., Kumar, A., Temimi, M., & Bull, D. R. (2020). Habnet: Machine learning, remote sensing based detection of harmful algal blooms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 3229–3239.

Hu, C., Muller-Karger, F. E., Taylor, C. J., Carder, K. L., Kelble, C., Johns, E., & Heil, C. A. (2005). Red tide detection and tracing using MODIS fluorescence data: A regional example in SW Florida coastal waters. *Remote Sensing of Environment*, *97*, 311–321.

Inkawhich, N., Zhang, J., Davis, E. K., Luley, R., & Chen, Y. (2022). Improving out-of-distribution detection by learning from the deployment environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 2070–2086.

Kakogeorgiou, I., & Karantzalos, K. (2021). Evaluating explainable artificial intelligence methods for multi-label deep learning classification tasks in remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *103*, 102520.

Karney, C. F. (2013). Algorithms for geodesics. *Journal of Geodesy*, *87*, 43–55.

Kim, Y., Byun, Y., Kim, Y., & Eo, Y. (2009). Detection of Cochlodinium polykrikoides red tide based on two-stage filtering using MODIS data. *Desalination*, *249*(3), 1171–1179. https://doi.org/10.1016/j.desal.2009.05.009

Lee, M.-S., Park, K.-A., Chae, J., Park, J.-E., Lee, J.-S., & Lee, J.-H. (2019). Red tide detection using deep learning and high-spatial resolution optical satellite imagery. *International Journal of Remote Sensing*, *41*(15), 5838–5860. https://doi.org/10.1080/01431161.2019.1706011


 

Lee, Z., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, *41*, 5755–5772.

Lou, X., & Hu, C. (2014). Diurnal changes of a harmful algal bloom in the East China Sea: Observations from GOCI. *Remote Sensing of Environment*, *140*, 562–572.

Medina, M., Kaplan, D., Milbrandt, E. C., Tomasko, D., Huffaker, R., & Angelini, C. (2022). Nitrogen-enriched discharges from a highly managed watershed intensify red tide (*Karenia brevis*) blooms in southwest Florida. *Science of the Total Environment*, *827*, 154149.

Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). *Journal of Geophysical Research: Oceans*, *93*, 10749–10768.

NASA Goddard Space Flight Center. (2021). *Moderate-resolution imaging spectroradiometer (MODIS) aqua ocean level-2 data* [Dataset]. Retrieved August 17, 2021, from: https://doi.org/10.5067/AQUA/MODIS/L2/OC/2022

NOAA National Centers for Environmental Information. (2022). *Etopo1 1 arc-minute global relief model* [Dataset]. Retrieved August 17, 2021, from: https://doi.org/10.25921/fd45-gt74

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Qin, M., Li, Z., & Du, Z. (2017). Red tide time series forecasting by combining ARIMA and deep belief network. *Knowledge-Based Systems*, *125*, 39–52.

Sakuno, Y., Maeda, A., Mori, A., Ono, S., & Ito, A. (2019). A simple red tide monitoring method using sentinel-2 data for sustainable management of Brackish Lake Koyama-ike, Japan. *Water*, *11*(5), 1044. https://doi.org/10.3390/w11051044

Shin, J., Jo, Y.-H., Ryu, J.-H., Khim, B.-K., & Kim, S. M. (2021). High spatial-resolution red tide detection in the Southern Coast of Korea Using U-Net from PlanetScope Imagery. *Sensors*, *21*(13), 4447. https://doi.org/10.3390/s21134447

Sonak, S., Patil, K., Devi, P., & D'Souza, L. (2018). Causes, human health impacts and control of harmful algal blooms: A comprehensive review. *Environmental Pollution and Protection*, *3*, 40–55.

Soto, I. M., Hu, C., Cannizzaro, J., Wolny, J., & Muller-Karger, F. E. (2021). *Remote detection of neurotoxic dinoflagellate* Karenia brevis *blooms on the West Florida shelf* (Vol. *20*, pp. 61–71, Chapter 6). International Ocean-Colour Coordinating Group.

Steidinger, K. A. (2009). Historical perspective on *Karenia brevis* red tide research in the Gulf of Mexico. *Harmful Algae*, *8*, 549–561.

Stumpf, R., Culver, M., Tester, P., Tomlinson, M., Kirkpatrick, G., Pederson, B., Truby, E., Ransibrahmanakul, V., & Soracco, M. (2003). Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae*, *2*, 147–160.

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*, e0260755.

Tomlinson, M., Wynne, T., & Stumpf, R. (2009). An evaluation of remote sensing techniques for enhanced detection of the toxic dinoflagellate, *Karenia brevis*. *Remote Sensing of Environment*, *113*, 598–609.

Tomlinson, M. C., Stumpf, R. P., Ransibrahmanakul, V., Truby, E. W., Kirkpatrick, G. J., Pederson, B. A., Vargo, G. A., & Heil, C. A. (2004). Evaluation of the use of SeaWiFS imagery for detecting Karenia brevis harmful algal blooms in the eastern Gulf of Mexico. *Remote Sensing of Environment*, *91*(3–4), 293–303. https://doi.org/10.1016/j.rse.2004.02.014

Turley, B. D., Karnauskas, M., Campbell, M. D., Hanisko, D. S., & Kelble, C. R. (2022). Relationships between blooms of *Karenia brevis* and hypoxia across the West Florida shelf. *Harmful Algae*, *114*, 102223.

Weisberg, R. H., 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. *Journal of Geophysical Research: Oceans*, *124*, 2501–2512.

Yunus, A. P., Dou, J., & Sravanthi, N. (2015). Remote sensing of chlorophyll-a as a measure of red tide in Tokyo bay using hotspot analysis. *Remote Sensing Applications: Society and Environment*, *2*, 11–25.

Zhao, X., Liu, R., Ma, Y., Xiao, Y., Ding, J., Liu, J., & Wang, Q. (2021). Red tide detection method for HY−1D Coastal Zone imager based on U−Net convolutional neural network. *Remote Sensing*, *14*(1), 88. https://doi.org/10.3390/rs14010088

 