Marine Pollution Bulletin 188 (2023) 114715

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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

Spatial-temporal dynamics of decaying stages of pelagic *Sargassum* spp. along shorelines in Puerto Rico using Google Earth Engine

Mariana C. León-Pérez *, Anthony S. Reisinger, James C. Gibeaut

Harte Research Institute for Gulf of Mexico Studies, Texas A&M University-Corpus Christi, 6300 Ocean Drive, Unit 5869, Corpus Christi, TX 78412, United States of America

ARTICLE INFO

Keywords:
Pelagic sargassum influxes
sargassum monitoring
Sargassum-brown-tide
Google Earth Engine
Sentinel-2

ABSTRACT

Coastal social-ecological systems in the Caribbean are affected by pelagic *Sargassum* spp. influxes and decomposition, but most satellite monitoring efforts focus on offshore waters. We developed a method to detect and spatial-temporally assess sargassum accumulations and their decaying stages along the shoreline and nearshore waters. A multi-predictor Random Forest model combining Sentinel-2 MultiSpectral Instrument reflectance bands and several vegetation, seaweed, water, and water quality indices was developed within the online Google Earth Engine platform. The model achieved 97 % overall accuracy and identified both fresh and decomposing sargassum, as well as the Sargassum-brown-tide generated from decomposing sargassum. We identified three hotspots of sargassum accumulation in La Parguera, Puerto Rico and found that sargassum was present every month in at least one of its forms during the entire time series (September 2015–January 2022). This research provides information to understand sargassum impacts and areas where mitigation efforts need to focus.

1. Introduction

Wider Caribbean and West African coastlines have been dealing with the arrival, accumulation, and decomposition of pelagic *Sargassum* spp. (*Sargassum natans* and *Sargassum fluitans* referred to as sargassum from hereon), as a result of a new sargassum source region in the Equatorial Atlantic (Gower et al., 2013). For more than a decade, these annual events have been detrimental to coastal and marine ecosystems (Azanza Ricardo and Pérez Martín, 2016; Cabanillas-Terán et al., 2019; Hernández et al., 2022; Rodríguez-Martínez et al., 2019; van Tussenbroek et al., 2017). Floating living sargassum accumulated along the shoreline is a threat to benthic organisms that depend on sunlight for photosynthesis (CRFM, 2016), and is a nuisance for sea turtles that nest on the beach (Azanza Ricardo and Pérez Martín, 2016; Gavio and Santos-Martinez, 2018). However, most of the impacts of sargassum on the natural and social systems are related to the decaying process sargassum undergoes when it is trapped along the shoreline. After 48 h, noxious hydrogen sulfide emissions of decomposing sargassum can trigger neurological, digestive, and respiratory disorders in nearby human populations (Resiere et al., 2020, 2019). Further decomposition of sargassum organic material results in the formation of Sargassum-brown-tide (Sbt), murky brown water with reduced light, oxygen, and pH, and increased nutrient loads (van Tussenbroek et al., 2017). Negative effects of Sbt expand further than the sargassum accumulation and include seagrass die offs, partial or total mortality of nearshore corals, and fauna mortality (Rodríguez-Martínez et al., 2019; van Tussenbroek et al., 2017).

Efforts to detect sargassum accumulations on and near the coastline using remote sensing are limited (Zhang et al., 2022) mainly due to challenges associated with spectral, spatial, and temporal resolutions (Arellano-Verdejo et al., 2019; Wang and Hu, 2021; Zhang et al., 2022). Several studies, for instance, Arellano-Verdejo et al. (2019), Cuevas et al. (2018), Dierssen et al. (2015), and Wang and Hu (2021, 2016), have demonstrated the usefulness of threshold-based methods, deep neural networks, and Random Forest models to detect floating sargassum in nearshore and offshore open waters. Very recently, Zhang et al. (2022) used a deep learning model to extract sargassum features from Planet Dove satellite imagery along Miami and Cancun shorelines. However, resource managers are still in need of national and site-level monitoring of stranded sargassum (Oxenford et al., 2021), which involves not only the detection of living fresh sargassum, but decaying sargassum and Sbt. So far, the available literature only provides a method to differentiate sargassum on the water from sargassum on the beach (Zhang et al., 2022), but not specifically for the sargassum

* Corresponding author.
E-mail addresses: mleonperez@islander.tamucc.edu (M.C. León-Pérez), anthony.reisinger@gmail.com (A.S. Reisinger), james.gibeaut@tamucc.edu (J.C. Gibeaut).

https://doi.org/10.1016/j.marpolbul.2023.114715
Received 8 September 2022; Received in revised form 23 January 2023; Accepted 5 February 2023
Available online 11 February 2023

decaying process. Here we show a preliminary assessment of Sentinel-2 Multispectral Instrument (MSI) data highlighting the differences in spectral signature of sargassum at different decomposition stages on the southern coast of Puerto Rico (Fig. 1). Given the differences in spectral signatures of decomposing sargassum and Sbt from fresh sargassum, a different methodological approach from traditional threshold-based methods for sargassum detection (Dierssen et al., 2015; Wang and Hu, 2016) is needed.

While fresh sargassum detection is well documented in the literature, no research has addressed the detection of sargassum decomposing stages. To fill this data gap, we developed a method to detect sargassum and its decaying stages on the shoreline and nearshore waters using satellite data and to characterize these occurrences temporally and spatially. We accessed and processed Sentinel-2 MSI data in Google Earth Engine (GEE) and developed a multi-predictor Random Forest model to detect fresh and decomposing sargassum, as well as Sbt. Spatial and temporal dynamics of sargassum accumulation hotspots in La Parguera, Puerto Rico were identified and assessed.

2. Materials and methods

2.1. Study area

La Parguera Nature Reserve (La Parguera hereinafter) is located along the southwest coast of the main island of Puerto Rico in the Caribbean Sea at 17.9592°N, 67.0442°W (Fig. 2). Puerto Rico's climate is characterized by a wet season from May to November and a dry season from December to April (Glynn, 1973). The main island topography and the dominant easterly trade winds create high spatial variability in annual precipitation (PRCCC Working Group 1, 2013), creating a rain shadow from east to west and from north to south (Miller and Lugo, 2009). Because of the absence of perennial rivers and low rainfall (Pittman et al., 2010), coastal waters in La Parguera tend to be clear. The main ocean water circulation is westward (Hernández-Guerra, 2000) and tides are a mix of diurnal and semidiurnal (Kjerfve, 1981), with a

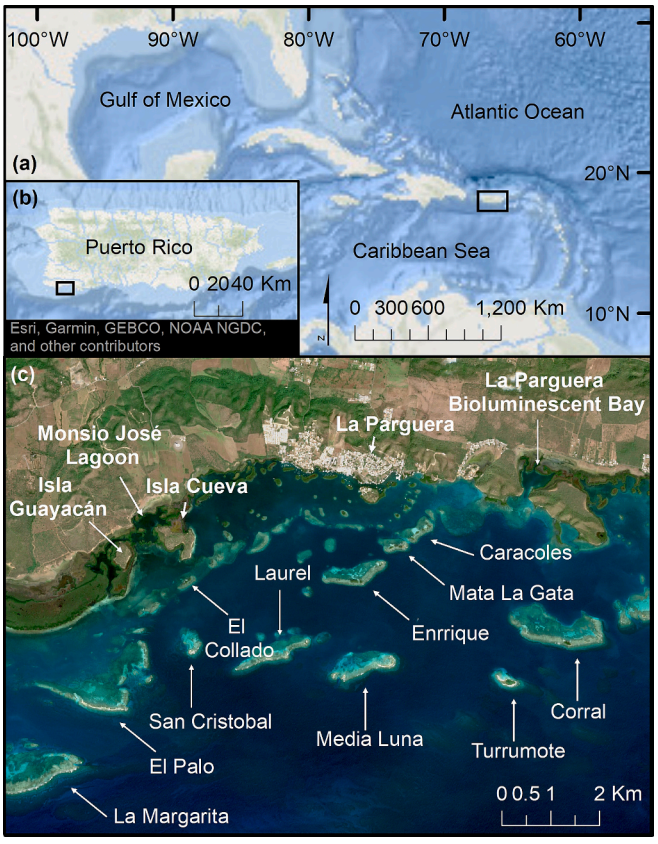


Fig. 2. Study area in Puerto Rico. a) Caribbean Sea, b) the archipelago of Puerto Rico, c) Sentinel-2 MSI image from La Parguera Nature Reserve showing main geographic locations and offshore cays.

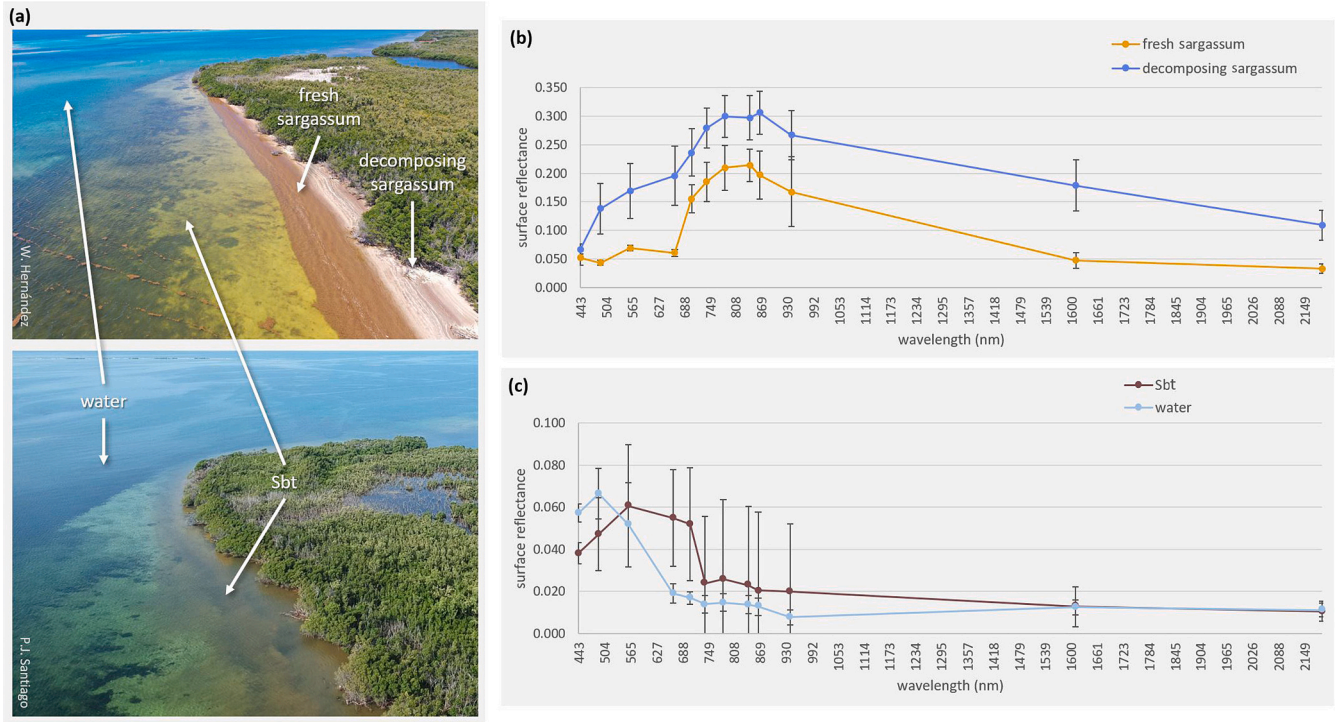


Fig. 1. Spectral signatures derived from Sentinel-2 MSI pixels. (a) Drone photographs showing sargassum decaying stages at La Parguera, Puerto Rico. (b) Fresh sargassum and decomposing sargassum spectral signatures from a May 27, 2021, image. (c) Spectral signatures of Sbt and water from pixels from a February 12, 2019, image.

diurnal tide range of <0.6 m (NOAA, 2019). The dominant wind direction in La Parguera is south-southeast (WindFinder.com GmbH and Co. KG, 2022).

La Parguera encompasses a variety of coastal and marine ecosystems, such as mangroves, seagrass beds, coral reefs, emergent cays, and bioluminescent bays (Valdés-Pizzini and Schärer-Umpierre, 2014). Its broad shelf and coastal embayment provide a sheltered shallow-water environment for these diverse and productive communities (Pittman et al., 2010). La Parguera is an important tourist destination in Puerto Rico, where most of the uses are related to nautical activities such as boating, snorkeling, SCUBA diving, fishing, kayaking, and wind and kite surfing (Valdés-Pizzini and Schärer-Umpierre, 2014).

2.2. Data sources

2.2.1. Sentinel-2 MSI data

The Sentinel-2 mission consists of two polar-orbiting satellites that provide a revisit time of 5 days at the equator and 2–3 days at mid-latitudes under cloud-free conditions (ESA, 2019). Both satellites are equipped with the MSI sensor that offers 13 spectral bands ranging from the visible spectrum to the shortwave infrared (Table 1). We used the level-1C product, a top-of-the-atmosphere (TOA) reflectance product, that has been radiometrically and geometrically corrected imagery.

2.2.2. Training and validation data

Field data was collected from 2019 to 2022. While navigating a boat, photographs of the shoreline and some cays were taken. We also had access to photographs taken from drones for some of the field visits when collaborators from the University of Puerto Rico at Mayaguez participated. Although the goal was to collect field data the same dates as the Sentinel-2 satellite passed, on some occasions it was not possible. We only considered field data collected no more than two days before or after a Sentinel-2 pass.

Each field datum was assigned to a class based on a scheme, which consisted of three sargassum classes (fresh, decomposing, and Sbt) and three non-sargassum classes (water, mangroves, and clouds) (Table 2). Sargassum accumulated above the water line in cays was not considered.

Table 1
Sentinel-2 MSI reflectance bands (ESA, 2019).

Spectral bands		Central wavelength (nm)	Spatial resolution (m)
B1	Coastal aerosol	443	60
B2	Blue	490	10
B3	Green	560	10
B4	Red	665	10
B5	Vegetation Red Edge	705	20
B6	Vegetation Red Edge	740	20
B7	Vegetation Red Edge	783	20
B8	Near Infrared (NIR)	842	10
B8A	NIR narrow	865	20
B9	Water vapour	940	60
B10	Short Wave Infrared (SWIR)-Cirrus	1375	60
B11	SWIR	1610	20
B12	SWIR	2190	20

Areas represented by each class were delineated using the geometry tool in GEE and converted to feature collections for training and validation (León-Pérez et al., in press).

To provide an independent set of data for training and validation of the model, the 16 Sentinel-2 scenes that corresponded with field data were grouped by month and randomly assigned to either training or validation. The scenes provide a balanced representation of the three sargassum classes and drone data for both training and validation. Approximately two-thirds of the imagery corresponding to field data were assigned to training (11 scenes) and one-third for validating the model (5 scenes).

Using a priori knowledge of the scenes (Chen and Stow, 2002), an additional set of visually interpreted data was also included to increase the representation of classes that had few pixels represented in the field data and therefore strengthen the model. We selected different scenarios of atmospheric conditions (e.g., presence of clouds and Saharan dust) and sea conditions (e.g., calm seas and turbulent seas). The visual interpretation was conducted using both red-green-blue and infrared (bands: 8, 3 and 2) band combinations. Visually interpreted imagery was also randomly assigned to training (6 scenes) and validation (8 scenes).

2.3. Imagery pre-processing and analysis

All image preprocessing, processing and analysis was conducted in GEE platform, specifically in the Earth Engine Code Editor, a web-based IDE for the Earth Engine JavaScript API, which is designed to develop complex geospatial workflows (Gorelick et al., 2017).

2.3.1. Imagery pre-processing

An initial image collection was built for La Parguera (tile: T19QGV) with acquisition dates from September 16, 2015 to March 31, 2022. However, scenes from January 25, 2022 onwards were later removed from the analysis (see Section 4.1). For each of the initial 443 Sentinel-2 TOA scenes, a cloud mask was applied using a predefined Sentinel-2 cloud probability mask function available in GEE (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_CLOUD_PROBABILITY#description).

The NIR band was not used for masking land areas (Su and Huang, 2019) because we observed that reflectance of accumulated sargassum on the shoreline was similar to that of mangroves. We also decided not to manually delineate a land mask from an image without accumulated sargassum because the mangrove/water boundary changed during the time series due to hurricanes and sargassum impacts. Thus, land areas were removed using a manually delineated land mask that excluded fringing mangroves, which allowed the classification method to determine which pixels represented mangroves and which sargassum or other classes. Although, there are no light-colored sandy beaches along the shoreline that could be misclassified as decomposing sargassum, there are some light-colored above-water coral rubbles along offshore cays that were masked out to prevent confusion. Lastly, the region of interest was defined with a polygon extending to −67.1107 west, −67.0049 east, 17.9800 north and 17.9150 south.

2.3.2. Supervised classification

A Random Forest (RF) supervised learning algorithm (built in GEE)

Table 2
Classification scheme used in this study.

	Class name	Description
Sargassum	Fresh sargassum	Pelagic sargassum mats drifting on open water or accumulated along the shoreline but not yet decomposed.
	Decomposing sargassum	Sargassum that is in a decaying process in water and turned whitish in color.
	Sbt	Colored dissolved organic matter created by sargassum decay (van Tussenbroek et al., 2017).
Non-sargassum	Water	Any water areas not representing Sbt.
	Mangroves	Fringing mangrove (mostly *Rhizophora mangle*) not included in the land mask (see Section 2.3.1).
	Clouds	Clouds not removed by the cloud masking process (see Section 2.3.1).

was used to classify the image collection. This algorithm creates multiple decision trees on data samples and merges them to finally select the best solution (Donges, 2022). We used a similar method to the multi-index approach developed by Cuevas et al. (2018), but we incorporated additional predictors to the model to be able to identify not only fresh sargassum but decomposing sargassum and Sbt (Table 2). Since our method is not a threshold-based approach attempting to derive a universal threshold for sargassum detection (Wang and Hu, 2016), we computed the band ratios and indices using TOA reflectance, which provided the longest temporal dataset. The ".importance()" function in GEE was used to measure the relative importance of a predictor to the model performance and to decide which inputs are important and which could be removed, reducing the chances of overfitting (Donges, 2022). We express the predictors importance as a percent, calculated as the importance score for each predictor divided by the total sum of the predictors' importance scores multiplied by 100.

The RF classifier was trained using an independent set of data that consisted of a feature collection for each of the 17 MSI scenes assigned for training. These feature collections consisted of manually digitized polygons representing the 6 classes of the classification scheme (Table 2) and were digitized based on field data. Finally, we selected 150 as the optimal number of decision trees for the RF classifier after trying different values and observing the effects on the accuracy assessment.

2.3.3. Accuracy assessment

Independent sets of validation data, from 5 images with field data collected and 8 images with interpreted data, were used to assess the accuracy of the classification model. Two error matrixes were created to determine the classification accuracy of each class, the overall accuracy, and to calculate the producer's and user's accuracies (Schowengerdt, 1983). The producer's accuracy, or omission error, indicates the probability of a reference pixel being correctly classified in the map, while the user's accuracy, or commission error, indicates the probability that a pixel classified on the map indeed represents that category on the ground (Congalton, 1991).

2.3.4. Spatial-temporal analysis of sargassum accumulation hotspots in La Parguera

The persistence of occurrence for each class (mode) was calculated for each pixel to determine areas where sargassum classes have been present the most. Given that areas where sargassum accumulates the most are often sources of Sbt, we defined a Sargassum Accumulation Hotspot (SAH) as an area where fresh sargassum and Sbt persisted throughout the time series. SAHs were manually delineated in GEE and spatial statistics were used to assess their temporal and spatial dynamics. We included in the analysis only images that had equal to or <10 % of cloud cover for a given SAH.

3. Results

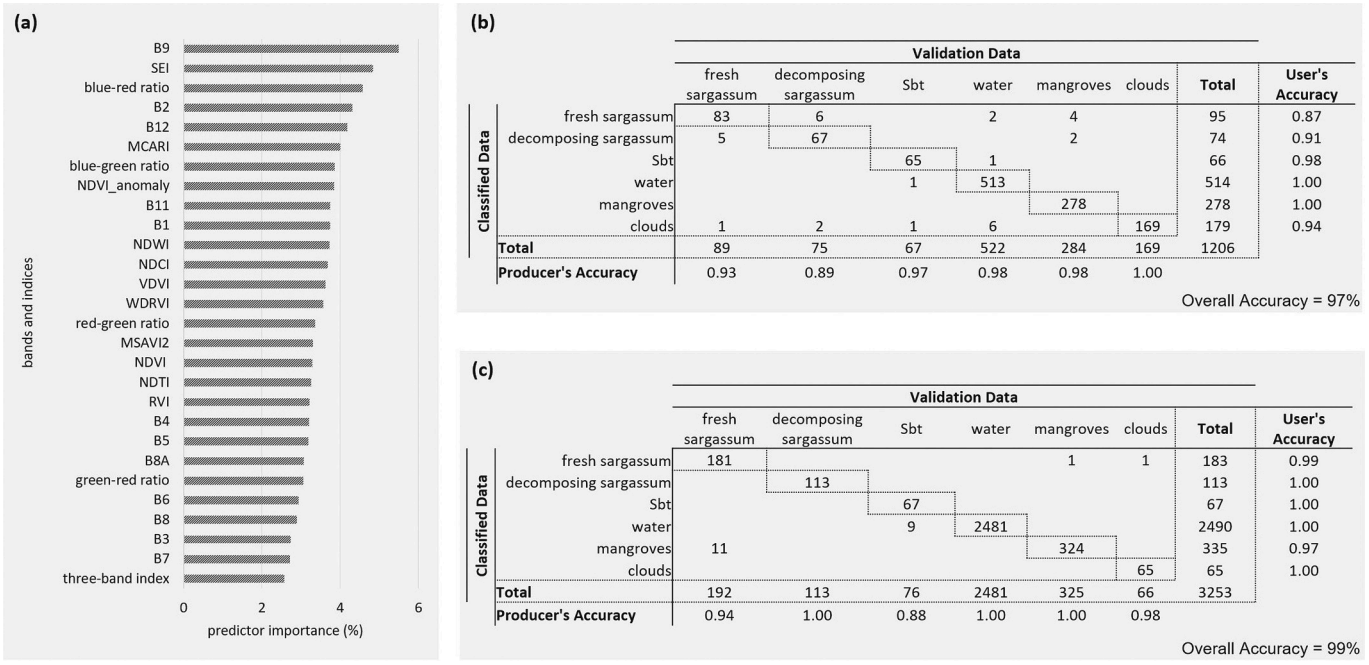
3.1. Classification model and accuracy assessment

The multi-index approach developed in this study incorporated a combination of 12 of the 13 MSI spectral bands, and the 16 vegetation and seaweed, water, and water quality indices described in Table 3. Other band, band ratios, and indices were originally included as inputs of the classification model, however they were eliminated due to low relative variable importance, these were: B10, difference vegetation index (Xue and Su, 2017), NIR-red band ratio (Gholizadeh et al., 2016), floating algae index (FAI) (Hu, 2009), alternative floating algae index (AFAI) (Wang and Hu, 2016), and normalized difference red-red edge, and red-red edge band ratio (Avdan et al., 2019). FAI is defined as the difference between reflectance at the vegetation red-edge and a linear baseline between the red band and short-wave infrared band (Hu, 2009), while the AFAI uses the same FAI design but with different spectral bands examining the red-edge reflectance of floating vegetation (Wang and Hu, 2016).

Table 3
Names, description and calculi of the predictors used in the RF classification model (Avdan et al., 2019; Gholizadeh et al., 2016; Siddiqui et al., 2019; Xue and Su, 2017).

Predictors		Description/equations adjusted for Sentinel-2 MSI bands
Spectral bands	B1, B2, B3, B4, B5, B6, B7, B8, B8A, B9, B11, B12	TOA reflectance of the Level-1C MSI product
Vegetation and seaweed indices	Normalized difference vegetation index (NDVI)	$\frac{B8-B4}{B8+B4}$
	Seaweed enhancing index (SEI)	$\frac{B8-B11}{B8+B11}$
	Ratio vegetation index (RVI)	$\frac{B4}{B8}$
	Modified secondary soil-adjusted vegetation index (MSAVI2)	$0.5*\left[(2B8+1)-\sqrt{(2B8+1)^2-8(B8-B4)}\right]$
	Visible-band difference vegetation index (VDVI)	$\frac{(2B3-B4-B2)}{2B3+B4+B2}$
	Wide dynamic range vegetation index (WDRVI)	$\frac{0.2B8-B4}{0.2B8+B4}$
	Modified chlorophyll absorption ratio index (MCARI)	$\frac{1.5*[2.5(B8-B4)-1.3(B8-B3)]}{\sqrt{(2B8+1)^2-(6B8-5B4)-0.5}}$
	NDVI anomaly	The NDVI anomaly consist of the difference between NDVI pixels values in each image within the image collection and the average of NDVI pixel values of the sargassum free baseline. In the absence of MSI imagery from before 2011, the NDVI baseline of sargassum free images was created from four images that visually had no or negligible sargassum in any of its forms.
Water index	Normalized difference water index (NDWI)	$\frac{B3-B8}{B3+B8}$
Water quality indices	Blue-green ratio	$\frac{B2}{B3}$
	Green-red ratio	$\frac{B3}{B4}$
	Blue-red ratio	$\frac{B2}{B4}$
	Normalized difference chlorophyll index (NDCI)	$\frac{B3-B5}{B3+B5}$
	Normalized difference turbidity index (NDTI)	$\frac{B4-B3}{B4+B3}$
	Red-green ratio	$\frac{B4}{B3}$
	Three-band index	$\frac{B4*B3}{B2}$

The overall variable importance score (Fig. 3), shown in the form of percent of the total importance, indicated that B9 was the most important variable for predicting the classes used in this study, followed by SEI, blue-red ratio, and B2 (Table 3). When looking individually at the



(b)

Classified Data \ Validation Data	fresh sargassum	decomposing sargassum	Sbt	water	mangroves	clouds	Total	User's Accuracy
fresh sargassum	83	6		2	4		95	0.87
decomposing sargassum	5	67			2		74	0.91
Sbt			65	1			66	0.98
water			1	513			514	1.00
mangroves					278		278	1.00
clouds	1	2	1	6		169	179	0.94
Total	89	75	67	522	284	169	1206	
Producer's Accuracy	0.93	0.89	0.97	0.98	0.98	1.00		

Overall Accuracy = 97%

(c)

Classified Data \ Validation Data	fresh sargassum	decomposing sargassum	Sbt	water	mangroves	clouds	Total	User's Accuracy
fresh sargassum	181				1	1	183	0.99
decomposing sargassum		113					113	1.00
Sbt			67				67	1.00
water			9	2481			2490	1.00
mangroves	11				324		335	0.97
clouds						65	65	1.00
Total	192	113	76	2481	325	66	3253	
Producer's Accuracy	0.94	1.00	0.88	1.00	1.00	0.98		

Overall Accuracy = 99%

Fig. 3. Analysis summarizing the performance of the RF classification model. (a) Importance of the predictors used in the model. Indices acronyms are described in Table 3. Confusion matrix for the (b) field data and (c) interpreted validation datasets. The diagonal values in the central part of the tables refers to the observations correctly classified by the model.

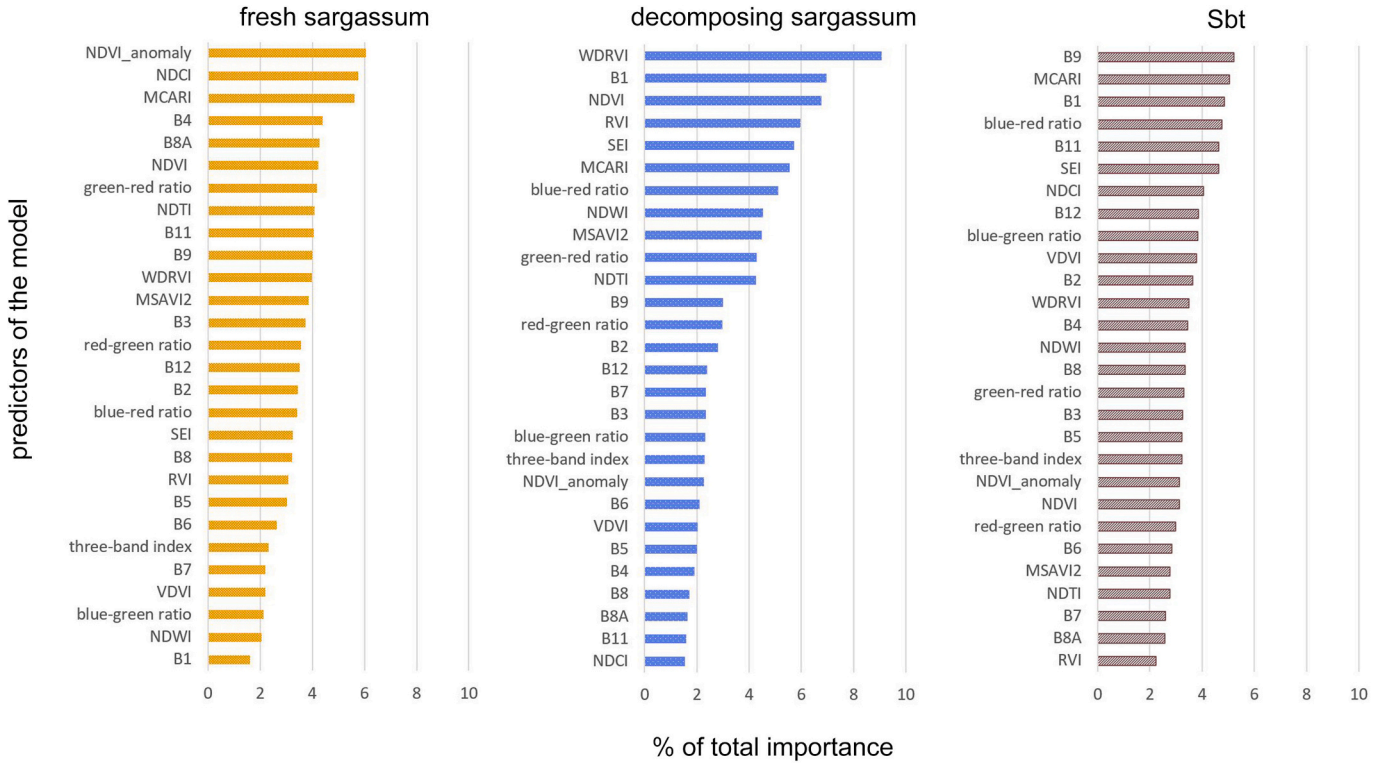


Fig. 4. Importance of the predictors used in the RF model for fresh sargassum, decomposing sargassum, and Sbt. Indices acronyms are described in Table 3.

variable importance for the fresh sargassum class (Fig. 4), NDVI anomaly, NDCI, and MCARI were the three most relevant inputs. The most important predictor for decomposing sargassum was predominantly WDRVI, although B1 and other vegetation and seaweed indices played an important role. For the Sbt class, the percent of total importance of the predictors was distributed more evenly than in fresh and decomposing sargassum, meaning that no input was predominantly more important than the others. A combination of spectral bands, and vegetation, seaweed, and water quality indices (e.g., B9, MCARI, B1, blue-red ratio, B11, and SEI) composed the top predictors of this class.

The confusion matrixes for the field and interpreted validation datasets show high agreement between the pixels classified by the model

and the validation data. The overall mapping accuracy for the field and interpreted datasets was 97 % and 99 %, respectively. Focusing on the sargassum classes and from the standpoint of the creator of the model, the producer's accuracies indicate that fresh sargassum, decomposing sargassum and Sbt had high probabilities (≥88 %) of being classified as such in the model. On the other hand, user's accuracy is a measure of the reliability of the map (Congalton, 1991) and also had high probabilities (≥87 %) for the three sargassum classes, meaning that the model can accurately predict the cover type on the ground.

3.2. Spatial-temporal dynamics of sargassum classes

Fresh sargassum was found either in the form of long mats in open water or accumulated at some of the cays and fringing mangrove forest. It tended to accumulate within the sections of the shoreline that face the east and southeast. Decomposing sargassum was primarily detected within enclosed small coves along the shoreline also oriented towards the east and southeast. Similarly, Sbt was associated with areas where decomposing sargassum was found, as well as waters surrounding some of the cays where sargassum accumulates (Fig. 2).

Hotspots of sargassum accumulation in La Parguera were identified by looking at the persistence of occurrence of fresh sargassum and Sbt for the entire time series (Fig. 5). There were three areas along the shoreline where the presence of these two classes coincided and therefore were identified as SAHs, these were Isla Cueva, Isla Guayacán, and La Pitahaya (León-Pérez et al., in press). The presence of decomposing sargassum during summer months in both field observations and classified imagery confirmed this determination.

Temporal dynamics of the three sargassum classes within each SAH are shown in Fig. 6. Fresh and decomposing sargassum followed a seasonal pattern, covering a greater area from May to August, although this pattern becomes less evident as the SAH is located further west (from top to bottom in Fig. 6). The portion of the classified area with fresh and decomposing sargassum also becomes less from east to west. Decomposing

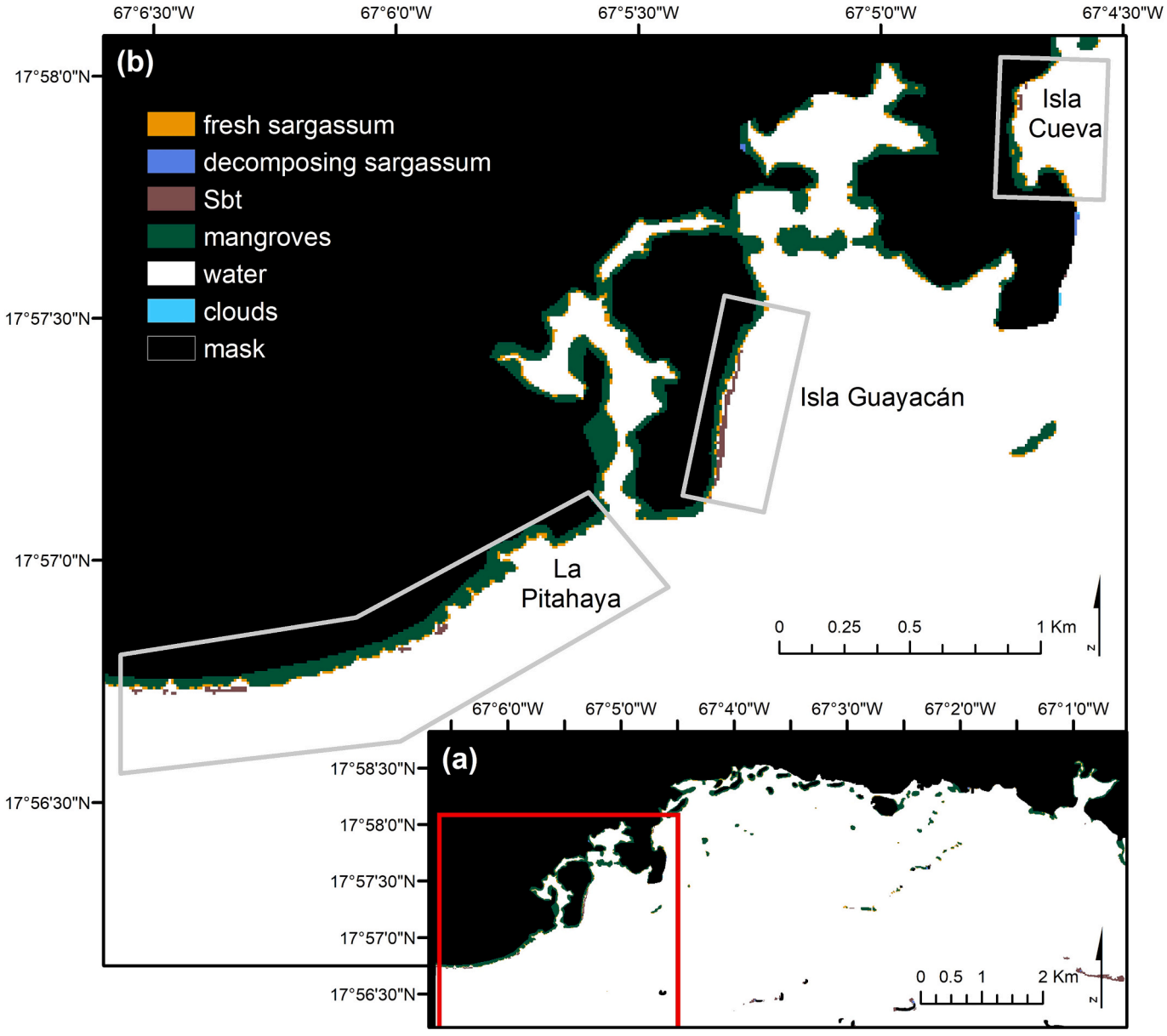


Fig. 5. (a) Persistence of occurrence (mode) of each class in La Parguera, Puerto Rico during the study period. (b) Zoomed area of (a) (red rectangle) showing the three sargassum accumulation hotspots identified (grey rectangles): Isla Cueva, Isla Guayacán, and La Pitahaya. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

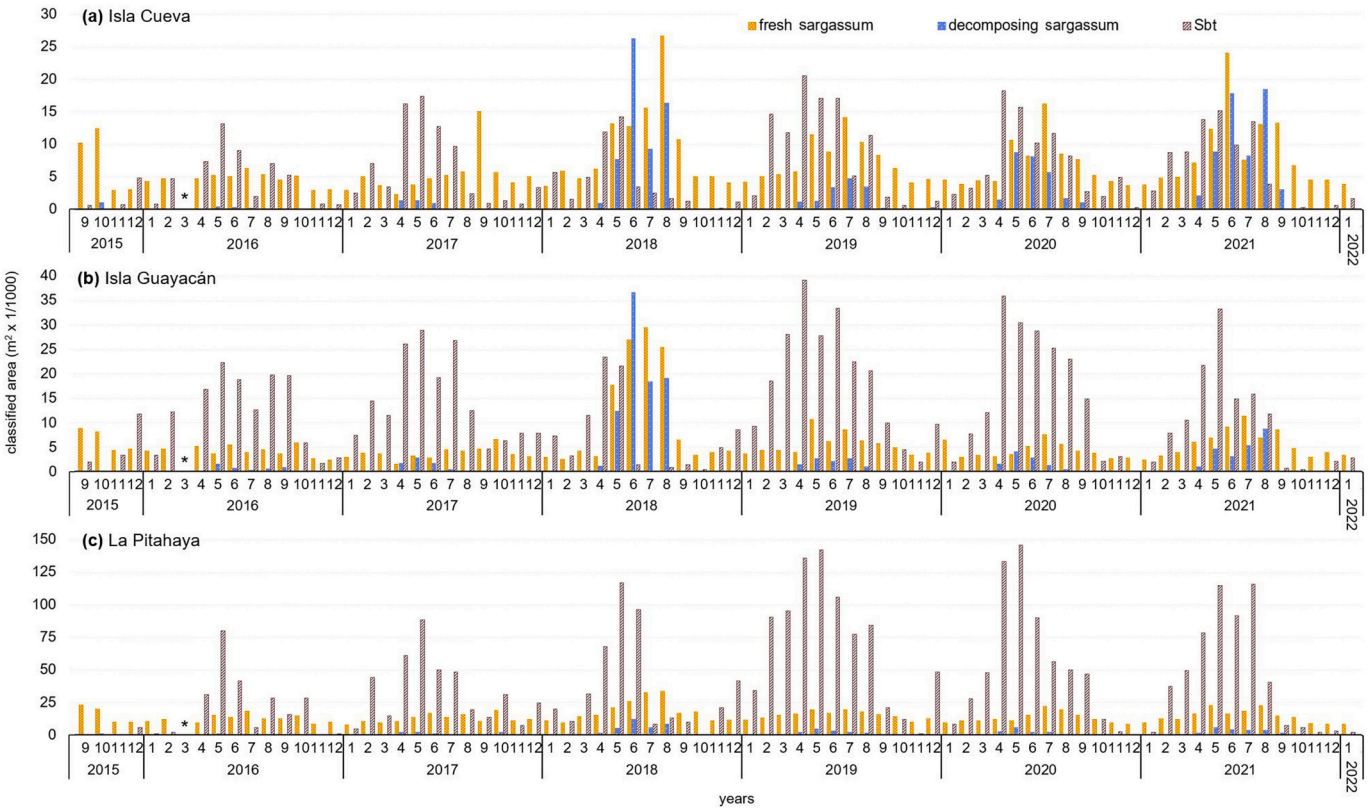


Fig. 6. Time series of the monthly average area covered by the three sargassum classes in (a) Isla Cueva, (b) Isla Guayacan and (c) La Pitahaya SAHs. Note: Graphs have different y-axis maximums. *Month with no data.

sargassum is mainly present from April to September. Sbt also follows a seasonal pattern, however the curve starts to rise earlier in the years, mostly in February and March, and lasting, on many occasions, to August and September. In contrast to fresh and decomposing sargassum, Sbt represent a bigger percent of the classified area in the two westward SAHs.

The maximum area covered by fresh and decomposing sargassum during the time series occurred in Summer 2018 for all SAHs (León-Pérez et al., in press), except for Isla Cueva SAH that also had a high value during Summer 2021. In general, after the 2018 sargassum influx event, the maximum area covered by Sbt in all SAHs increased and remained higher than years before.

4. Discussion

4.1. Performance of the classification model

RF classification models have been widely used for solving a variety of classification scenarios (Sannigrahi et al., 2022; Zhang and Yang, 2020) including sargassum detection (Cuevas et al., 2018). Since the decision of assigning a class to a certain pixel is purely made by the RF model based on training data, it reduces the biases associated with the use of threshold-based methods such as the NDVI, SEI, FAI and AFAI (Hu, 2009; Siddiqui et al., 2019; Wang and Hu, 2016). Being interested in the performance of the model, an all-at-once approach was used, meaning that we included the maximum number of predictors that the GEE processing capacity allowed and that contributed to model performance. While Cuevas et al. (2018) found that seven predictors (e.g., four vegetation and one seaweed indices, and two MSI spectral reflectance bands) provided a robust way to detect fresh sargassum in Landsat 8 images off the coast of Mexico, our results show that twenty-eight predictors enabled us to account for the differences in spectral signatures of the different sargassum classes with very high classification accuracies (Fig. 3).

One of the most relevant predictors of the model, mainly for detecting fresh sargassum, was the NDVI anomaly. The concept behind incorporating this predictor was to capture sargassum as an anomaly between images with sargassum and images without sargassum. Although initially we intended to calculate this anomaly from FAI or AFAI values, both predictors ranked low in the variable importance analysis and removing them improved the overall accuracy. FAI probably ranked low because it incorporates B10, a MSI spectral band that seems to bring a lot of noise into the model. On the other hand, even though AFAI does not incorporate B10, we were surprised that it did not provide as much importance as expected to the model. The combination of the other predictors contained within the model probably provided the needed information for a good performance. Consequently, the sargassum anomaly was computed from NDVI values, which is the most important predictor for the fresh sargassum class. This result opens the possibility for an alternate approach to traditional threshold-base methods for detecting sargassum in coastal and open waters.

Meanwhile, it is important to consider potential sources of error and limitations of the method described here. The use of drone photographs and videos helped overcome limitations in the amount of field data collected for training and validation of the model, in particular for areas not easily accessible by boat or walking. More field data coinciding with Sentinel-2 passes would provide more data for verification of the accuracy of the model.

Regarding spatial resolution, the high-resolution Sentinel-2 MSI data used in this study provided the resolution necessary to detect sargassum accumulations on the shoreline in contrast with coarse resolution sensors such as the Medium Resolution Imaging Spectrometer (MERIS), the Moderate Resolution Imaging Spectroradiometer (MODIS), the Visible Infrared Imaging Radiometer Suite (VIIRS), among others, used to monitor large-scale sargassum distribution across the ocean (Hu et al., 2016; Wang et al., 2019; Wang and Hu, 2016). Nevertheless, pixels located in the land-water edge where sargassum tends to accumulate

could have a combination of classes (e.g., pixel contamination) potentially confusing the model. Mangrove and seaweed have a similar spectral signature (Siddiqui et al., 2019), but the likelihood of misclassification was reduced by incorporating various vegetation and seaweed indices in the RF model. Errors in classification could also occur for fresh sargassum floating in open water. Given that sargassum moves in the form of thin mats, often narrower than Sentinel-2 MSI spatial resolution, the detection rate will be lower (Shin et al., 2021), and sargassum coverage will be underestimated (Wang and Hu, 2021). Higher spatial resolution sensors such as Planet Scope/Dove and WorldView-II sensors may be better at detecting smaller sargassum (Wang and Hu, 2021) and shoreline features, but they provide limited long-term data availability. For example, stages of sargassum decomposition above and within the water line at the beach should be better captured with a higher spatial resolution sensor. Still, we believe the use of Sentinel-2 MSI imagery provides the compromise between adequate spatial resolution and imagery availability for assessing long-term temporal dynamics of sargassum accumulations on the shoreline.

It is interesting to note that our method detected fresh sargassum during dry season months, when accumulations of fresh sargassum on the shoreline is not expected in MSI detectable quantities. We went to the field to corroborate the observations and realized that sunken sargassum from at least the previous year was misclassified as live floating fresh sargassum. This mainly occurred within sargassum accumulation hotspots when the water was clear (Figs. 6 and 7b), leading to overestimations of fresh sargassum during dry season months. Ideally, we would have trained the model to include an additional class of submerged sargassum, but data limitations prevented this from happening.

Another source of error is due to the unintended incorporation of training data from a non-harmonized Sentinel-2 image of March 23, 2022 into the RF model. After January 25, 2022, Sentinel-2 MSI scenes had their digital number range shifted by 1000 and although GEE provided a new corrected harmonized image collection, it was not used for image processing. Fortunately, the March 23, 2022 image was only used to train three classes (e.g., Sbt, water and mangroves) and the three of them had high user and producer accuracies (≥97 %). We speculate that this might be because only one of the 17 images used for training had this error and because that image was used to train only three classes. Future classification efforts should use the harmonized Sentinel-2 MSI dataset.

4.2. Spatial-temporal dynamics of SAHs in La Parguera

The method presented here has the power to assess the temporal and spatial dynamics of the different stages of the sargassum decaying process in a coastal area. Regarding living, floating sargassum, our results are in accordance with previous studies that show a seasonal pattern with maximum coverage during summer months (Rodríguez-Martínez et al., 2019; Wang et al., 2019; Wang and Hu, 2016). Our results also found evidence of the massive sargassum influx event of Summer 2018 (Wang et al., 2019), when SAHs in La Parguera experienced the maximum coverage of fresh sargassum.

The sargassum decaying process starts when organic material from sargassum is trapped against the shoreline or restricted with diminished water circulation (van Tussenbroek pers. com.). Our results demonstrate that this process starts to occur quickly (in less than a month), since decomposing sargassum follows the same seasonal pattern as fresh

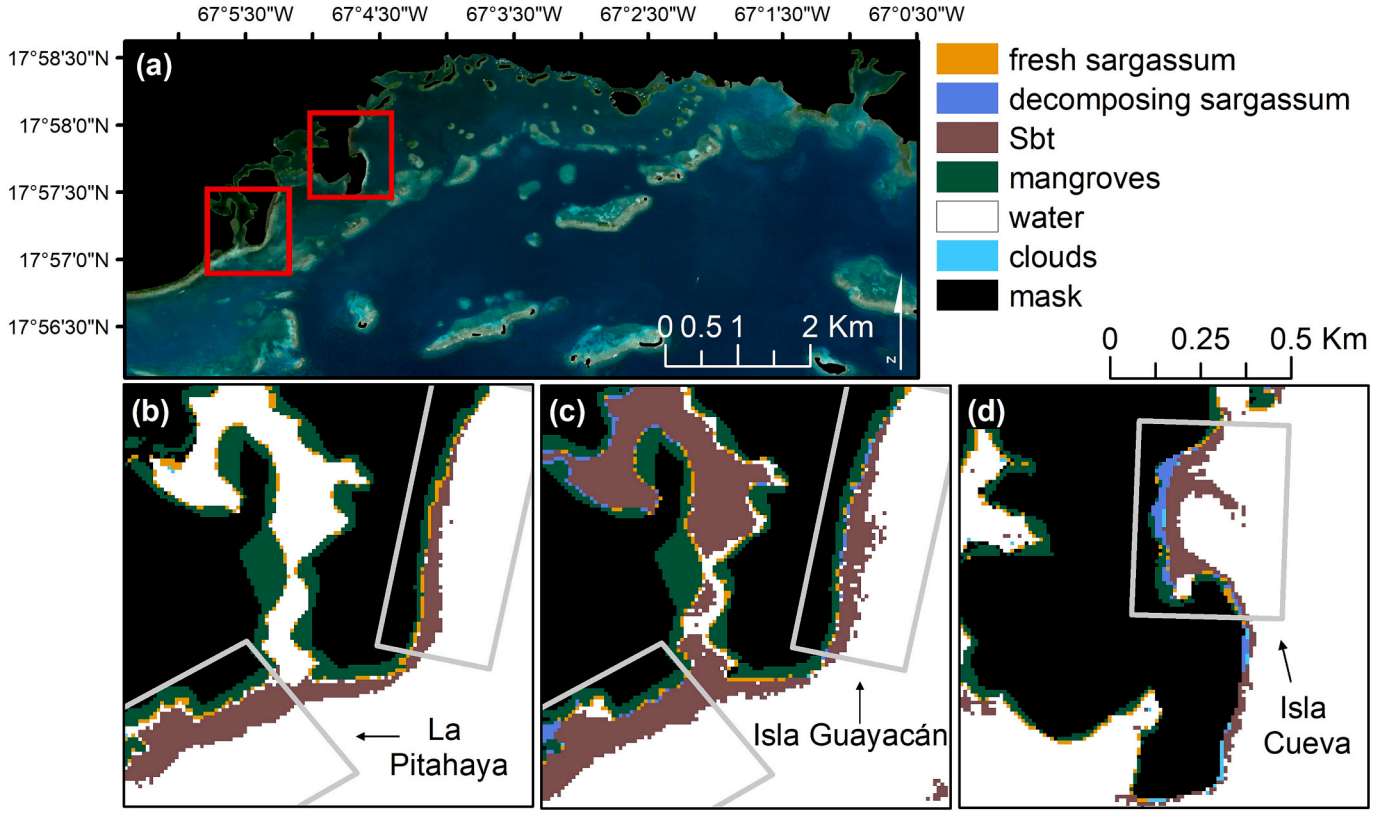


Fig. 7. Examples of classification products in La Parguera's SAHs. (a) La Parguera and the spatial extents of b and c (left red rectangle), and d (right rectangle). (b) February 22, 2019: Sunken sargassum classified by the RF model as fresh sargassum at Isla Guayacán and La Pitahaya SAHs, and contribution of Sbt from Isla Guayacán into La Pitahaya SAH. (c) June 1, 2021: Contribution of Sbt from Isla Guayacán and inner lagoons into La Pitahaya SAH. (d) June 1, 2021: Sbt moving away from the shore at Isla Cueva SAH and towards the Southwest. Red star: Location of a former mangrove canal that led to Monsio José Lagoon, which is currently blocked due to recurrent sargassum accumulations and lack of maintenance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

sargassum and is only present mostly from April to September when the amount of fresh sargassum is greatest. It is important to consider that the coverage of the decomposing sargassum class represents an underestimate of sargassum decomposition. This class was trained with pixels from sargassum decaying areas that appeared whitish, meaning that decomposition that looks spectrally different (e.g., below water decomposition and anaerobic decomposition occurring above and within the water line in cays) was not detected in this class. Sargassum decomposition within mangrove roots was also undetected by the method used because the mangrove canopy obstructs the satellite nadir observation.

The area covered by Sbt, on the other hand, starts to increase months after the sargassum influx event usually at the beginning of next year, but follows a seasonal pattern as well. Therefore, in contrast to decomposing sargassum, Sbt takes more time to develop and persists for longer periods, usually >7 months per year (Fig. 6). It was also observed that, in general, the area covered by Sbt increased after the 2018 event and remained higher throughout the rest of the time series (Fig. 6 and León-Pérez et al., in press). This finding is consistent with van Tussenbroek et al. (2017) observations, where a year after a sargassum influx event a decrease in water transparency persisted suggesting that the imported organic material and nutrients were not removed from the system. During our last field visit in March 2022, we observed impressive accumulations of shredded sargassum on the seabed within SAHs. This was observed about 5 months after the 2021 sargassum influx event, the second most massive influx event in our time series. Therefore, it is likely that the imported organic material from at least the previous year, 2021, had not been flushed from the system when in April 2022 the next sargassum influx began. As van Tussenbroek et al. (2017) suggested, detrimental consequences to coastal stability and other ecosystem services are expected given that the frequency of sargassum influxes is higher than the time of recovery of the system.

The proportion of Sbt compared to the other sargassum classes varied spatially between SAHs. We believe two factors contributed to these variations, the geomorphology of each SAHs, and the prevailing current (Hernández-Guerra, 2000) and wind direction (WindFinder.com GmbH and Co. KG, 2022). The following information may help explain why Sbt represents a bigger portion of the classified area for SAHs located westward. Isla Cueva SAH is a confined area that faces East (Fig. 5), which allows accumulation of fresh sargassum and formation of decomposing sargassum. Under the dominant wind direction (south-southeast), sargassum within this SAH is unlikely to be flushed. On the contrary, Isla Guayacán and La Pitahaya SAHs are less confined and more open areas, where fresh sargassum may be able to flush with the influence of northeast winds that occur less often (WindFinder.com GmbH and Co. KG, 2022). La Pitahaya is the SAH oriented more southward of the three SAHs, and therefore the one that experiences the less coverage of fresh sargassum throughout the study period. Although the geomorphological characteristics of Isla Cueva SAH prevent the natural removal of fresh and decomposing sargassum, a current that appears to come from a former canal within this SAH (Fig. 7d), flushes Sbt away from the shore and eventually Southwest with the dominant current direction. In contrast, Sbt observed in La Pitahaya SAH may be a consequence of the cumulative influence of Sbt generated in situ and Sbt coming from Isla Guayacán SAH (Fig. 7b and c). La Pitahaya SAH also receives the Sbt that flushes from the channel between Isla Guayacán and La Pitahaya, probably coming from smaller sargassum accumulation areas within the inner lagoons (Fig. 7c).

4.3. Advantages and future applications of the classification model

The method developed provides significant benefits for the detection of sargassum in nearshore environments. Thanks to the use of GEE, the complete archive of Sentinel-2 imagery was accessed and processed in the cloud without the need to download it and process it locally as in traditional approaches (Cuevas et al., 2018; Shin et al., 2021; Wang and Hu, 2021). GEE provided the processing capacity to incorporate twenty-eight predictors into a RF model of 150 trees. By adjusting few model parameters, this model could be easily adapted to other locations and even, to some extent, other satellite data (e.g., Landsat, Planet Dove). Equally important are the facts that both GEE and Sentinel-2 datasets are freely accessible, and that the Sentinel-2 dataset is continuously updated within GEE. The combination of all these facts opens the possibility of providing decision makers with a timely and practical nearshore sargassum monitoring tool.

Although Sentinel-2 MSI does not provide the ideal spatial resolution for sargassum detection in nearshore environments (Wang and Hu, 2021), the current study shows that useful information can be derived from using this satellite constellation. Different from other very high-resolution sensors such as Planet Dove, Sentinel-2 MSI also provides a long-term dataset for the assessment of spatiotemporal dynamics of SAHs. This information can be beneficial for: (1) determining the location of SAHs along a coastline and the contribution of Sbt from each, (2) prioritizing mitigation interventions to minimize impacts, and (3) assessing the placement of boom barriers and plan the removal of sargassum confined within them to prevent further Sbt generation.

5. Conclusion

In this article we addressed the need for a comprehensive detection and assessment of the spatial-temporal dynamics of sargassum accumulations and its decaying stages on and near the shoreline. We developed a method able to accurately detect fresh sargassum, decomposing sargassum, and sargassum brown tide from high resolution satellite imagery. Combining Sentinel-2 MSI reflectance bands with vegetation, seaweed, water, and water quality indices in a Random Forest model proves to be a powerful approach for discriminating between water and sargassum classes, providing significant benefits compared to traditional approaches using single indices and threshold-based methods. The results also open possibilities for using anomaly calculations, such as NDVI anomaly, for the detection of sargassum.

The practicality of this method was proven by detecting and assessing the spatial-temporal dynamics of three sargassum accumulation hotspots in La Parguera, Puerto Rico. We found that since 2015, sargassum has been present along the coast of La Parguera every month either as fresh and decomposing, or as Sbt. The geomorphological characteristics of the shoreline along with prevailing currents and wind direction are factors that drive where along the shoreline sargassum accumulation hotspots form and where Sbt reaches.

The identification of sargassum in its different facets provided by this work is a major step towards understanding the impacts of these events on natural and social systems and for prioritizing mitigation strategies to overcome large sargassum influxes. We will build on this effort by incorporating additional training data from coastal areas with diverse characteristics, exploring integrating other sources of satellite data, and providing a user-friendly interface for its use.

CRediT authorship contribution statement

Mariana C. León-Pérez: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Anthony S. Reisinger:** Methodology, Software, Writing – review & editing. **James C. Gibeaut:** Conceptualization, Methodology, Resources, Writing – review & editing, Project administration, Funding acquisition.

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 shared through the Gulf of Mexico Research Initiative Data and Information Cooperative (GRIIDC)

Acknowledgments

This publication was made possible by the NOAA, Office of Education Educational Partnership Program award (NA16SEC4810009). Its contents are solely the responsibility of the award recipient and do not necessarily represent the official views of the U.S. Department of Commerce, NOAA.

We are grateful to Dr. Philippe Tissot and Dr. Brigitta van Tussenbroek for their suggestions in the development of this research. We thank Jenniffer Pérez-Pérez, a NOAA Center for Earth System Sciences and Remote Sensing Technologies scholar, for sharing her field data and for her assistance in additional field data collection. We also thank Jan P. Zegarra, Efrain Figueroa, and collaborators from the "Developing Decision-Making Tools for Sargassum Management in Coastal Areas" project funded by Sea Grant Puerto Rico, Dr. Roy Armstrong and Dr. William Hernández, for their support in field data collection. Finally, the authors would like to thank two anonymous reviewers for their contributions.

References

Arellano-Verdejo, J., Lazcano-Hernandez, H.E., Cabanillas-Terán, N., 2019. ERISNet: deep neural network for sargassum detection along the coastline of the mexican caribbean. PeerJ 2019, 1–19. https://doi.org/10.7717/peerj.6842.

Avdan, Z.Y., Kaplan, G., Goncu, S., 2019. Monitoring the water quality of small water bodies using high-resolution remote sensing data. Int. J. Geo-Information 8.

Azanza Ricardo, J., Pérez Martín, R., 2016. Artículo Original Impacto de la Acumulación de Sargazo del Verano del 2015 sobre las Tortugas Marinas de Playa La Barca, Península de Guanahacabibes. Rev. Investig. Mar. 36, 52–60.

Cabanillas-Terán, N., Hernández-Arana, H.A., Ruiz-Zárate, M.-Á., Vega-Zepeda, A., Sanchez-Gonzalez, A., 2019. Sargassum blooms in the Caribbean alter the trophic structure of the sea urchin Diadema antillarum. PeerJ 7, e7589. https://doi.org/10.7717/peerj.7589.

Chen, D.M., Stow, D., 2002. The effect of training strategies on supervised classification at different spatial resolutions. Photogramm. Eng. Remote. Sens. 68, 1155–1161.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B.

CRFM, 2016. Model Protocol for the Management of Extreme Accumulations of Sargassum on.

Cuevas, E., Uribe-Martínez, A., Liceaga-Correa, M. de los Á., 2018. A satellite remote-sensing multi-index approach to discriminate pelagic Sargassum in the waters of the Yucatan Peninsula, Mexico. Int. J. Remote Sens. 39, 3608–3627. https://doi.org/10.1080/01431161.2018.1447162.

Dierssen, H.M., Chlus, A., Russell, B., 2015. Hyperspectral discrimination of floating mats of seagrass wrack and the macroalgae Sargassum in coastal waters of Greater Florida Bay using airborne remote sensing. Remote Sens. Environ. 167, 247–258. https://doi.org/10.1016/j.rse.2015.01.027.

Donges, N., 2022. Random Forest Classifier: A Complete Guide to How It Works in Machine Learning (WWW Document). Builtin. https://builtin.com/data-science/random-forest-algorithm. (Accessed 20 December 2019).

European Space Agency (ESA), 2019. SENTINEL-2. URL. https://sentinel.esa.int/web/sentinel/missions/sentinel-2. (Accessed 15 June 2019) (WWW Document).

Gavio, B., Santos-Martinez, A., 2018. Floating Sargassum in Serranilla Bank, Caribbean Colombia, may jeopardize the race to the ocean of baby sea turtles. Acta Biológica Colomb. 23, 311–313. https://doi.org/10.15446/abc.v23n3.68113.

Gholizadeh, M.H., Melesse, A.M., Reddi, L., 2016. A comprehensive review on water quality parameters estimation using remote sensing techniques. Sensors 16. https://doi.org/10.3390/s16081298.

Glynn, P.W., 1973. Ecology of a Caribbean coral reef. The Porites reef-flat biotope: Part I. Meteorology and hydrography. Mar. Biol. 20, 297–318.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Gower, J., Young, E., King, S., 2013. Satellite images suggest a new Sargassum source region in 2011. Remote Sens. Lett. 4, 764–773. https://doi.org/10.1080/2150704X.2013.796433.

Hernández, W.J., Morell, J.M., Armstrong, R.A., 2022. Using high-resolution satellite imagery to assess the impact of Sargassum inundation on coastal areas. Remote Sens. Lett. 13, 24–34. https://doi.org/10.1080/2150704X.2021.1981558.

Hernández-Guerra, A., 2000. Water masses and circulation in the surface layers of the Caribbean at 66W. Geophys. Res. Lett. 27, 3497–3500.

Hu, C., 2009. A novel ocean color index to detect floating algae in the global oceans. Remote Sens. Environ. 113, 2118–2129. https://doi.org/10.1016/j.rse.2009.05.012.

Hu, C., Murch, B., Barnes, B., Wang, M., Maréchal, J.-P., Franks, J., Johnson, D., Lapointe, B., Goodwin, D., Schell, J., Siuda, A., 2016. Sargassum watch warns of incoming seaweed. Eos (Washington DC) 97, 1–14. https://doi.org/10.1029/2016eo058355.

Kjerfve, B., 1981. Tides of the Caribbean Sea. J. Geophys. Res. 86, 4243–4247.

Miller, G.L., Lugo, A.E., 2009. Guide to the Ecological Systems of Puerto Rico. U.S. Department of Agriculture, Forest Service, International Institute of Tropical Forestry, San Juan, P.R.

NOAA, 2019. Tides and Currents Playa de Ponce (TEC4733)/San Juan, La Puntilla (9755371). URL. https://tidesandcurrents.noaa.gov/noaatidepredictions.html?id=TEC4733&units=standard&bdate=20190608&edate=20190609&timezone=LST&clock=12hour&datum=MLLW&interval=hilo&action=dailychart. (Accessed 20 June 2019) (WWW Document).

Oxenford, H.A., Cox, S.-A., van Tussenbroek, B.I., Desrochers, A., 2021. Challenges of turning the Sargassum crisis into gold: current constraints and implications for the Caribbean. Phycology 1, 27–48. https://doi.org/10.3390/phycology1010003.

Pittman, S.J., Hile, S.D., Jeffrey, C.F.G., Clark, R., Woody, K., Herlach, B.D., Caldow, C., Monaco, M., Appeldoorn, R.S., 2010. Coral reef ecosystems of Reserva Natural de La Parguera (Puerto Rico): spatial and temporal patterns in fish and benthic communities (2001–2007). In: NOAA Technical Memorandum NOS NCCOS 107.

Puerto Rico Climate Change Council (PRCCC) Working Group 1, 2013. Geophysical and chemical scientific knowledge. In: Jacobs, K.R., Terando, A., Diaz, E. (Eds.), Puerto Rico's State of the Climate 2010–2013: Assessing Puerto Rico's Social-Ecological Vulnerabilities in a Changing Climate. Puerto Rico Coastal Zone Management Program, Department of Natural and Environmental Resources, NOAA Office of Ocean and Coastal Resource Management, San Juan, P.R., pp. 21–84

Resiere, D., Mehdaoui, H., Névière, R., Mégarbane, B., 2019. Sargassum invasion in the Caribbean: the role of medical and scientific cooperation. Rev. Panam. Salud Pública 43, 1. https://doi.org/10.26633/rpsp.2019.52.

Resiere, D., Mehdaoui, H., Florentin, J., Gueye, P., Lebrun, T., Blateau, A., Viguier, J., Valentino, R., Brouste, Y., Kallel, H., Megarbane, B., Cabie, A., Banydeen, R., Neviere, R., 2020. Sargassum seaweed health menace in the Caribbean: clinical characteristics of a population exposed to hydrogen sulfide during the 2018 massive stranding. Clin. Toxicol. 1–9 https://doi.org/10.1080/15563650.2020.1789162.

Rodríguez-Martínez, R.E., Medina-Valmaseda, A.E., Blanchon, P., Monroy-Velázquez, L. V., Almazán-Becerril, A., Delgado-Pech, B., Vásquez-Yeomans, L., Francisco, V., García-Rivas, M.C., 2019. Faunal mortality associated with massive beaching and decomposition of pelagic Sargassum. Mar. Pollut. Bull. 146, 201–205. https://doi.org/10.1016/j.marpolbul.2019.06.015.

Sannigrahi, S., Basu, B., Sarkar, A., Pilla, F., 2022. Development of automated marine floating plastic detection system using Sentinel-2 imagery and machine learning models. Mar. Pollut. Bull. 178 https://doi.org/10.1016/j.marpolbul.2022.113527.

Schowengerdt, R.A., 1983. Techniques for Image Processing in Remote Sensing. Academic Press Inc, New York, p. 249p.

Shin, J., Lee, J.S., Jang, L.H., Lim, J., Khim, B.K., Jo, Y.H., 2021. Sargassum detection using machine learning models: a case study with the first 6 months of goci-ii imagery. Remote Sens. 13, 1–20. https://doi.org/10.3390/rs13234844.

Siddiqui, M.D., Zaidi, A.Z., Abdullah, M., 2019. Performance evaluation of newly proposed seaweed enhancing index (SEI). Remote Sens. 11, 1434.

Su, L., Huang, Y., 2019. Seagrass resource assessment using WorldView-2 imagery in the Redfish Bay, Texas. J. Mar. Sci. Eng. 7, 98. https://doi.org/10.3390/jmse7040098.

van Tussenbroek, B.I., Hernández Arana, H.A., Rodríguez-Martínez, R.E., Espinoza-Avalos, J., Canizales-Flores, H.M., González-Godoy, C.E., Barba-Santos, M.G., Vega-Zepeda, A., Collado-Vides, L., 2017. Severe impacts of brown tides caused by Sargassum spp. on near-shore Caribbean seagrass communities. Mar. Pollut. Bull. 122, 272–281. https://doi.org/10.1016/j.marpolbul.2017.06.057.

Valdés-Pizzini, M., Schärer-Umpierre, M., 2014. People, Habitats, Species, and Governance: An Assessment of the Social-Ecological System of La Parguera, Puerto Rico.

Wang, M., Hu, C., 2016. Mapping and quantifying Sargassum distribution and coverage in the Central West Atlantic using MODIS observations. Remote Sens. Environ. 183, 350–367. https://doi.org/10.1016/j.rse.2016.04.019.

Wang, M., Hu, C., 2021. Satellite remote sensing of pelagic Sargassum macroalgae: the power of high resolution and deep learning. Remote Sens. Environ. 264, 112631 https://doi.org/10.1016/j.rse.2021.112631.

Wang, M., Hu, C., Barnes, B.B., Mitchum, G., Lapointe, B., Montoya, J.P., 2019. The great Atlantic Sargassum belt. Science 80-.). 365, 83–87.

WindFinder.com GmbH & Co. KG, 2022. Wind & weather statistics La Parguera/Magueyes. URL. https://www.windfinder.com/windstatistics/la_parguera_magueyes_islands. (Accessed 20 July 2022) (WWW Document).

Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: a review of developments and applications. J. Sensors 2017, 17.

Zhang, F., Yang, X., 2020. Improving land cover classification in an urbanized coastal area by random forests: the role of variable selection. Remote Sens. Environ. 251 https://doi.org/10.1016/j.rse.2020.112105.

Zhang, S., Hu, C., Barnes, B.B., Harrison, T.N., 2022. Monitoring Sargassum inundation on beaches and nearshore waters using PlanetScope/Dove observations. IEEE Geosci. Remote Sens. Lett. 19 https://doi.org/10.1109/LGRS.2022.3148684.

Data references

León-Pérez, M.C., Reisinger, A.S., Gibeaut, J., 2022. Pelagic Sargassum Accumulations Along the Shoreline and Nearshore Waters in La Parguera, Puerto Rico. Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Harte Research Institute, Texas A&M University–Corpus Christi. https://doi.org/10.7266/gw5dnwpd.