A satellite remote-sensing multi-index approach to discriminate pelagic Sargassum in the waters of the Yucatan Peninsula, Mexico

Article *in* International Journal of Remote Sensing · June 2018

DOI: 10.1080/01431161.2018.1447162

3 authors:

Eduardo Cuevas
CONACYT-Universidad Autónoma del Carmen
32 PUBLICATIONS 142 CITATIONS

Abigail Uribe-Martínez
Universidad Autónoma del Carmen
8 PUBLICATIONS 39 CITATIONS

Liceaga Maria
Center for Research and Advanced Studies of the National Polytechnic Institute
13 PUBLICATIONS 30 CITATIONS

Eduardo Cuevas, Abigail Uribe-Martínez & María de los Ángeles Liceaga-Correa (2018) A satellite remote-sensing multi-index approach to discriminate pelagic Sargassum in the waters of the Yucatan Peninsula, Mexico, International Journal of Remote Sensing, 39:11, 3608-3627, DOI: 10.1080/01431161.2018.1447162
To link to this article: https://doi.org/10.1080/01431161.2018.1447162

A satellite remote-sensing multi-index approach to discriminate pelagic *Sargassum* in the waters of the Yucatan Peninsula, Mexico

Eduardo Cuevas[‡], Abigail Uribe-Martínez, and María de los Ángeles Liceaga-Correa

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Departamento de Recursos del Mar, Laboratorio de Percepción Remota y SIG

Carretera Antigua a Progreso km 6, C. P. 97310, Tel. 52(999)9429463 Mérida, México. E-mail: maria.liceaga@cinvestav.mx

Corresponding Author: **María de los Ángeles Liceaga-Correa**. Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Departamento de Recursos del Mar, Laboratorio de Percepción Remota y SIG. Carretera Antigua a Progreso km 6, C. P. 97310, Mérida, México. E-mail: maria.liceaga@cinvestav.mx

Word count: 8,214

[‡] *New Affiliation*: CONACYT-Universidad Autónoma del Carmen, Centro de Investigación de Ciencias Ambientales. C. 56 No. 4, Esq. Avenida Concordia Col. Benito Juárez, C. P. 24180, Cd. Del Carmen, Campeche, Mexico. OrcID: 0000-0003-3814-7211

A satellite remote-sensing multi-index approach to discriminate pelagic *Sargassum* in the waters of the Yucatan Peninsula, Mexico

Abstract

Recently, the need for quantitative information on the spatiotemporal distribution of floating macroalgae, particularly *Sargassum* spp., has grown because of blooms of these species in the Gulf of Mexico and Caribbean Sea. Remote sensing is one of the most frequently used tools to assess pelagic *Sargassum* distribution. The purpose of this study was to implement a methodological approach to detect floating algae in an efficient and replicable manner at a moderate cost. We analyzed Landsat 8 imagery, from which we calculated four vegetation indices and one floating-algae index to implement a supervised classification, together with the bands 2 and 5, using the Random Forest algorithm. The analysis was performed monthly from 2014 to 2015 for the northeastern Yucatan Peninsula, Mexico, with a total of 91 analyzed images. The quantitative performance metrics of the classifier (overall, Kappa and Tau) were greater than 80%, whereas bands 2 and 5 as well as atmospherically resistant vegetation index made the greatest contributions to the classifications. During summer 2015, more than 4,000 ha of *Sargassum* coverage per image were observed, which was substantially greater than that over the rest of the period. This approach constitutes a transferable alternative for the systematic detection of *Sargassum*, which enables a quantitative semi-automated time series comparison.

Keywords: *Sargassum*, macroalgae, Landsat, Random Forest, remote sensing

Introduction

The pelagic *Sargassum* genus include two species (from now on abbreviated as *spp.*, as it is referred to abbreviate more than one unspecified species) of Phaeophyta macroalgae that are irregularly distributed in configurations whose lengths range from 50 cm to several kilometres (Butler et al. 1983). *Sargassum* is widely distributed along the Gulf of Mexico and the Caribbean Sea, where biological assemblages of high economic and ecological value to Mexico, the United States, and Cuba are found (Thiel and Gutow 2005).

The marine communities composed of *Sargassum* and its occupants are part of intricate trophic chains and host different commercial fish species during their early stages of development, e.g., *Coryphaena*, *Abudefduf*, and *Caranx* (South Atlantic Fishery Management Council 2002; Vandendriessche et al. 2007), as well as endangered species and other species of high ecological value, such as sea turtles and pelagic seabirds (Mansfield and Putman 2003; Moser and Lee 2012; Witherington, Hirama, and Hardy 2012).

Although *Sargassum* represents important ecological and economic interests, the anomalous and excessive growth of this macroalga has been recorded in the Gulf of Mexico, the Caribbean Sea, and the Western Central Atlantic (WCA) since 2006 (Gower et al. 2006; Gower and King 2008; Gower and King 2011; Gower, Young, and King 2013; Wang and Hu 2016). This excessive growth has generated massive off-shore *Sargassum* shoals that negatively affect coastal towns, especially the tourism and service sectors as well as certain coastal ecosystems and endangered species (Smetacek and Zingone 2013; Gavio, Rincón-Díaz, and Santos-Marín 2015; Maurer, De-Neef, and Stapleton 2015; Schell, Goodwin, and Siuda 2015; van-Tussenbroek et al. 2017).

The first studies of *Sargassum* communities were conducted from oceanographic cruises in the Atlantic Ocean (Stoner 1983; Huffard et al. 2014). As a technological alternative, remote sensing has been used as a major tool in studies of *Sargassum* worldwide, including its detection at the surface of the water and on the sea bed (less than 10 m in depth) (Hu, Hardy, and Hocberg 2015; McCarthy 2016; Wang and Hu 2016, Xing et al. 2017).

The initial studies that employed remote sensing techniques used low spatial resolution multispectral imaging sensors with high temporal resolution to assess algae blooms in the Yellow Sea (Hu 2009), the Gulf of Mexico (Gower and King 2011), the northern region of the Gulf of

Mexico (specifically, the area affected by the 2010 oil spill) (Hu, Hardy, and Hochberg 2015; Hu et al. 2016), and the WCA (Wang and Hu, 2016)

In this context, it is important to understand the spatial-temporal dynamics of floating *Sargassum* lines to design and implement strategies to address anomalous events as well as to advance our understanding of their origins, under which physical and biological conditions they occur, and the amount of *Sargassum* expected on the beach, among other aspects.

In the Gulf of Mexico and the Lesser Antilles, formal initiatives have been aimed at providing near real-time information on the distribution patterns of pelagic *Sargassum* using Moderate Resolution Imaging Spectroradiometer (MODIS) data with long term and detailed temporal coverage (Optical Oceanography Laboratory, University of South Florida). With regard to the north of the Gulf of Mexico in particular, one initiative uses medium resolution satellite images (Landsat) (Webster and Linton 2013) to provide semi-quantitative information products to assess and quantify spatial and temporal patterns.

In Mexico, specifically in Quintana Roo (Mexican Caribbean), different strategies have been implemented to address this issue. However, baseline data on the spatial-temporal patterns and dynamics of this macroalgae in adjacent waters are lacking.

The objective of this study was to expand existing methodological approaches through the implementation of a low cost and replicable approach with a relatively low computing demand to detect and monitor pelagic *Sargassum* in the waters of the Yucatan Peninsula, Mexico. This study will enable the analysis of spatial-temporal dynamics of pelagic *Sargassum* to inform natural resource managers as well as planners to act in case of contingencies.

In the next sections, we present a methodological approach for the detection and assessment of *Sargassum* in a quantitative, robust, and semi-automated manner along with an evaluation of its accuracy.

Materials and Methods

Study area

For this study, we selected the northeastern region of the Yucatan Peninsula, Mexico because it is the major entry point of water into the Gulf of Mexico. This region is known for the Cabo Catoche seasonal upwelling and is characterized by strong sea currents as well as associated biological and ecological processes (Merino 1997; Sheinbaum et al. 2002; Schmitz et al. 2005; Rousset and Beal 2010). In addition, the area is located to the north of the beaches in the state of Quintana Roo, Mexico (van Tussenbroek et al. 2017), where large *Sargassum* shoals were recorded in 2014 and 2015 (Figure 1). As Hu et al. (2016) recognized, knowledge about the abundance of *Sargassum* in the eastern region of the Gulf of Mexico is limited; therefore, studying the *Sargassum* of that region is strategic for generating information about this macroalga.

Image selection and preprocessing

Given the size of the *Sargassum* lines reported in this area, which MODIS or the Medium Resolution Imaging Spectrometer (MERIS) cannot detect, and considering the advances achieved by similar studies using Landsat (Hardy 2014; Hu et al. 2016), we used Landsat 8 (L8) Operational Land Imager (OLI) images with a 30 m spatial resolution. We selected four scenes for the images covering the area of interest (paths 018 and 019 as well as rows 044 and 045), took the first image acquired each month for the four scenes along the studied period, and built a set of 91 L8 images for the years 2014 and 2015 (five images were missed from the source archive at the time of this search in July 2016).

During this period of time, a *Sargassum* bloom was reported in the Caribbean (Maurer, De-NEef & Stapleton, 2015; Azanza-Ricardo and Pérez-Martín, 2016). Therefore, by selecting this time frame, we ensured that *Sargassum* would be present in at least some of the images, thereby allowing the implementation of this approach and the spatial-temporal assessment of the massive bloom of the macroalga along the coasts of the Mexican Caribbean.

Satellite images were downloaded from the United States Geological Survey (USGS, 2016), and all were subject to an atmospheric correction (see Hu, 2004 and GRASS Development Team, 2017).

A land mask was applied to each image based on the Global Administrative Area polygons (Hijmans et al. 2016). Pixels whose configuration in the quality band (QB) indicated the presence of clouds, cirrus, snow, and ice were masked so that only the pixels whose QB configuration indicated possible water and unaffected pixels remained active (USGS, 2014).

Vegetation indices

After a literature review (and given the basic recommendations by Hu, Hardy, and Hochberg, 2015), a set of five vegetation indices with the potential to contribute to *Sargassum* detection was defined: Normalized difference vegetation index (NDVI), Atmospherically resistant vegetation index (ARVI), Soil-adjusted vegetation index (SAVI), Enhanced vegetation index (EVI) and Floating algae index (FAI) (Equations (1)-(6) in Table 1).

To make these calculations using the L8 images, we used band 2 (0.452 – 0.512 μm) as Blue (B), band 4 (0.636 – 0.673 μm) as Red (R), band 5 (0.851 – 0.879 μm) as Near InfraRed (NIR), band 6 (1.566 – 1.651 μm) as Short Wave InfraRed (SWIR) (Table 1).

Supervised image classification

For each image, a classification model was generated based on the selected training sites using a Random Forest algorithm (Liaw and Wiener 2002). This classifier consists of a combination of classification trees, in which each one is generated from a set of random sites independently sampled from the entry set, and each tree casts a single vote for the most popular type of entry vector (Breiman 1999; Pal 2005). This machine learning algorithm is regarded as one of the most efficient algorithms in terms of prediction accuracy, speed, and efficiency for large databases. Although the individual contribution of each index is combined during the classification, in terms of detecting *Sargassum* aggregations, we assumed that the multiple criteria that this algorithm are given with the different indices is an advantage. Furthermore, the algorithm is also well recognized because it offers an intuitive approach to assess the importance of each independent variable used in the model (Crisci, Ghattas, and Perera 2012).

Different remote sensing studies have used Random Forest because it provides greater accuracy than other machine learning algorithms (Akar and Güngör 2012). In this regard, Random Forest has been used for acoustic data (Lucieer et al. 2012), to map marsh vegetation (van Beijma, Comber, and Lamb 2014), and even as a prediction model for harmful algal blooms (Kehoe et al. 2012), among other applications. This commercial algorithm is accessible to any person interested in detecting floating *Sargassum*; hence, the replicability of this method is ensured.

The classification inputs were the five vegetation indices obtained, i.e., NDVI, ARVI, SAVI, EVI and FAI, combined with bands 2 (blue, 0.452 - 0.512 µm) and 5 (SWIR, 0.851 - 0.879 µm). All of the data contributed by the indices individually and as a set as well as additional information derived from the contrast between the NIR and the blue bands were used to distinguish floating algae (Hu 2009, Xing et al. 2017). For the classification parametrization, we set the number of variables in the random subset at each node (*m*) as 3, and the number of trees in the forest (*k*) as 1,000, with replacement of samples at each step.

To provide the classifier with a wider spectral context of the analyzed image, we included training sites for the other objects present in the scene; hence, we expected the classifier to have more chances of correctly classify *Sargassum*. Hence, sets of training sites were visually defined (polygons in vector format) for the classes of *Sargassum*, clouds, blue sea (open sea), sea with sun glint, and cloud shadows (shadows) based on the different composites from the generated indices (three at a time, one for each channel) and a false color composite (bands 5,2,1) of the images.

We obtained between 38 and 123 training sites (polygons) per image (mean 69), where 33.0% of them were labeled *Sargassum*. The latter training sites varied from four to 100 pixels in size depending on the cover extension of the *Sargassum* in each image. The rest of the classes were trained with a maximum of 250 pixels each.

With these training sites, we expected to capture the spectral variation in the classes within each assessed individual image, considering the influence on the image that different data acquisition conditions had (solar light angle of incidence, sensor angle, and even the presence of mist) (Wang and Hu 2016). Bands 5 and 2 as well as the index values were extracted from these polygons for their corresponding images.

The original individual classifications with five categories were reclassified to contain only the *Sargassum* class, which was subsequently converted into polygons in vector format so that the rest of the classes were omitted. In addition, rasters were generated with the probability of each pixel of being *Sargassum*. We also obtained values of the general errors in the classification determined via cross validation as well as scores of the importance of the variables to the general classification and for the *Sargassum* class exclusively.

A graphical assessment of the importance values for the five indices and bands under consideration was conducted to determine the main decision tree criteria for the five classes, particularly for the *Sargassum* class.

Supervision and quantification of Sargassum

As part of the supervision process, the polygons identified as *Sargassum* were subjected to visual quality control based on the visual compositions used before (i.e., the indices and bands). During the supervision process, errors of commission were corrected by the operator; however, omission errors were not corrected, primarily because of the significant manual digitization effort that this procedure entails. The *Sargassum* polygons obtained after this visual supervision were considered the final detection in each image.

The *Sargassum* probability (i.e., a raster of probabilities derived from the classification) was analyzed, both in the original classification and in the final detection to determine the range of probability values over which greater intervention was required to correct errors of commission.

The quantification is reported as hectares of *Sargassum* detected for each month and year.

Sargassum classification accuracy

A quantitative validation of the final *Sargassum* polygons was performed. Based on the analysis of the probability of each polygon being *Sargassum*, the probability quartiles were defined as tiers in the distribution of the verification sites in a uniform, stratified scheme. A total of 120 verification points per image were defined (i.e., 30 points per probability quartile), and the verification points were randomly distributed among the pixels in each quartile.

The validation design included only the classified images with detected *Sargassum*. If no *Sargassum* was observed within a scene, then it was excluded from the validation. A person external to the supervision of the classification categorized the validation points (in vector format) using the composites of false color indices with the bands 5, 2, and 1 as a reference. The points were labeled *Sargassum* or *non Sargassum*, and both the probabilities of *Sargassum* class and the final classification were invisible to the person who performed the validation.

A confusion matrix was generated from the polygons obtained from the final classification (*Sargassum* or *non Sargassum*), and the class assigned to the verification (a = true positives, b = false positives, c = false negatives and d = true negatives) (Green et al. 2000; Anderson, Lew, and Peterson 2003).

In addition, the quantitative measures of the intrinsic performance of the classifier were calculated (overall, omission error, and commission index) (Anderson et al. 2003) as were the Kappa and Tau accuracy metrics (Green et al. 2000; see Equations (7)-(11) in Table 2).

Results

Maps of the monthly distribution and coverage of pelagic *Sargassum* were generated for the northeast Yucatan Peninsula using the above described semi-automated detection protocol and Landsat 8 OLI images for 2014 and 2015.

Image analysis

Of the 91 analyzed images, 54 (60.0%) showed evidence of *Sargassum*; thus, they were classified. Likewise, 31 (34.0%) did not show evidence of the presence of *Sargassum*, whereas the remaining six (6.0%) exhibited high cloud coverage (> 80.0%), making detection impossible.

A gap of five images exited for the studied area and period that were not available in the USGS archive.

Bands 2 and 5 along with the ARVI had the greatest importance, on average, in the decision making of the classification trees (Figure 2(*a*)). Importantly, this participation is relative to all classes considered in the classification of each image and that their variability range was wide with an interquartile range (50.0% of the data) that overlapped with the inputs from the rest of the indices. On the other hand, when we only considered the *Sargassum* class, NDVI and SAVI were the indices with the greatest importance (Figure 2(*b*)).

The average global classification accuracy value derived from the cross validation as part of the Random Forest classification process was 97.7% (± 2.2%).

The supervision of the direct result of the classification resulted in the elimination of polygons erroneously labeled *Sargassum* in 60.0% of the classified images. This circumstance occurred most frequently in cloud edges and cloud shadows.

Distribution of the* Sargassum *probability values

In general, for the distribution of the *Sargassum* assignment probability values in a complete scene, we noted that a great proportion of the set of pixels in the image corresponded to low or null probabilities. A large proportion of pixels also showed a 0.1 to 10.0% probability of being *Sargassum*, and only a few of these were finally classified as such. In contrast, we found pixels with a high probability of being *Sargassum* (> 90.0%) but that also showed greater probabilities of being other classes; hence, these pixels were not classified as *Sargassum*.

Most errors of commission detected in the supervision after classification were found in the group of pixels with a *Sargassum* assignment probability between 10.0% and 80.0%. When assessing the distribution of the probability values to be assigned as *Sargassum* in the polygons removed during supervision, we observed that the highest percentage of changes (after the elimination of the errors of commission) occurred in pixels with *Sargassum* probability values between 10.0% and 70.0%; thus, a greater confusion was accompanied by a lower assignment probability. However, most pixels with a high probability of being *Sargassum* (> 80.0%) were maintained as such after supervision; thus, few errors of commission occurred at those probability levels.

In terms of the distribution of the *Sargassum* probability values of the pixels that effectively remained as such (distributed in light gray in Figure 3(*a*)), 50.0% of the pixels presented classification probability values between 65.0% and 100.0%.

Figure 3(*b*) shows the effect of removing the polygons wrongly classified as *Sargassum*; the average of probabilities increased, and the interquartile range decreased. Furthermore, 50.0% of the data were distributed between 77.0% and 98.0% probabilities of being classified as *Sargassum*. This pattern was consistent for all of the classified scenes, and the highest incidence of refinement was found in pixels with medium probability values, which also increased the centrality of the probability of the pixels correctly classified as *Sargassum*.

Quantification of Sargassum

The temporal distribution pattern of *Sargassum* in the area differed for 2014 and 2015. In 2015, *Sargassum* coverage was four times larger than that in 2014, which is quantitative evidence of its massive growth in this region (Figures 4 and 5).

In 2014, *Sargassum* drifts were scattered throughout the study area with higher concentrations along the continental slope. In 2015, the aggregation of large amounts of *Sargassum* was evident, and it followed the pattern of surface flows, particularly along the edge of the continental shelf. On the continental shelf, the presence of *Sargassum* was minimal in 2014 but was significant in 2015.

In 2014, two periods of greater *Sargassum* coverage were recorded; the first occurred in August, and the second occurred in November and December with values ranging from 1,000 to 4,000 ha for each event, respectively. In 2015, the first evidence of floating *Sargassum* aggregations of approximately 1,000 ha in size were observed in March; later that year, however, the maximum macroalga accumulations were recorded as more than 6,000 ha of *Sargassum* detected during July and August (Figure 5). During September and October of that same year, the amount of *Sargassum* decreased considerably with a coverage of less than 1,000 ha.

During the periods of maximum *Sargassum* aggregation (November and December 2014) as well as in July and August 2015, the paths along which the maximum values were detected diverged. The maximum values were first found in path 18 and later in path 19, suggesting a response to the direction of the surface currents in the area.

Regarding the temporal differences in *Sargassum* coverage between months, sudden changes were noted from November 2014 to January 2015 and from July to September 2015, during which the differences in the area of *Sargassum* coverage between months was greater (Figure 5).

Of the *Sargassum* coverage values detected in 2014 and 2015, increases between 50.0% (August 2014) and up to 400.0% (June-July 2015) were recorded for a single scene, providing evidence about the anomalous nature of the event recorded in 2015.

Sargassum classification accuracy

Based on the confusion matrix for the 2,040 assessed validation points, different metrics were estimated to judge the efficiency of the proposed approach to detect floating algae with L8 images (Table 3).

The global performance value of the classification was 93.4%, which includes the errors of omission and commission. The rate of omission or false negatives was 19.0%. Thus, the percentage of detected *Sargassum* was underestimated by the classifier (Table 2). Finally, the rate of errors of commission, or false positives, was 2.7%.

The classification accuracy metrics showed values greater than 80.0%, which represents a satisfactory level of performance using the methodological approach presented here. With a Kappa value of 81.2%, it is assumed that the classification process avoided this percentage of errors relative to the errors that a completely random classification would have generated. In addition, the Tau value indicates that 97.5% more pixels were correctly classified than the result expected from a random classification. These accuracy metric values support the calculated *Sargassum* coverage estimates.

Discussion

Based on the available literature and to the best of our knowledge, this study is the first to assess *Sargassum* via remote sensing using an integrated multi index approach as well as the first to document the 2015 *Sargassum* bloom along the Yucatan Peninsula.

Quantification of Sargassum in waters of the Yucatan Peninsula

The major increases in *Sargassum* coverage over the study area were recorded during autumn-winter 2014 and summer 2015. For the greater part of 2014, no large *Sargassum* aggregations were detected, but from August to December of that year, a considerable increase in coverage was recorded. Wang and Hu (2016) reported the same pattern for the WCA, and an increase in *Sargassum* coverage was noted again in the first quarter of 2015. The months of the greatest

Sargassum presence in our study area were from June to September of 2015, which coincided with the same period reported for the WCA, with remnants recorded until November of that year.

The temporal coincidence of the greater *Sargassum* coverage values in our study area and the WCA suggests that these *Sargassum* lines in the Mexican Caribbean did not originate in southern regions of the Caribbean Sea because blooming apparently occurred at the same time. In contrast, the bloom might have originated closer to our study region, with the peculiarity that the coverage of *Sargassum* in the assessed paths showed a time lag. Specifically, the eastern path ranked first in terms of the increase in coverage, followed by the neighboring path to the west, which followed the direction of the surface currents and the prevailing winds (i.e., from east to west; Enríquez, Mariño-Tapia, and Herrera-Silveira 2010; Reyes-Mendoza et al. 2016).

Regarding of the use of Landsat imagery to detect *Sargassum* coverage, we related our results to those of Hu et al. (2016); although our study differs in the years analyzed, it was the only study that used the same input data until today. They detected *Sargassum* in the area affected by the Deep-Water Horizon oil spill of 2010 in the northern region of the Gulf of Mexico, and they reported a coverage of thousands of hectares for the first trimester of the year. This order of magnitude of *Sargassum* coverage is consistent with our results.

Accuracy of Sargassum detection

Importantly, a significant percentage of actual floating *Sargassum* corresponds to patches of minimal size, which are difficult to detect via remote sensing data such as Landsat (with a 30 m spatial resolution). For a pixel to be defined as *Sargassum* using indices such as NDVI and FAI, Hu, Hardy, and Hochberg (2015) reported that it should at least cover between 1.0% and 2.0% of the area. For that pixel to be differentiated from other floating objects and detected as *Sargassum*, it should cover between 20.0% and 30.0%.

The presence and location of *Sargassum* lines are in constant motion; therefore, collecting field data to calibrate multispectral images for their detection is highly complex and costly. However, the current approach supports training site selection using index composites as well as the probability of each pixel of being classified as *Sargassum*. We used these by-products as an input to obtain a quantitative criterion for the interpretation of the resulting classification and the accuracy assessment. This methodological approach is the first to present metrics (overall performance, omission error, commission index, Kappa, and Tau) of its accuracy to detect *Sargassum* using L8 images, thereby providing a baseline for comparing with other detection methods.

In general, the values of all metrics indicated satisfactory performance, and it is worth noting the rate of omission errors (19.0%), which was highly relevant to this approach because *Sargassum* omissions were not corrected. Therefore, this value was directly estimated from the results of the classifier. The rate of this type of error was higher than that of the errors of commission (about 2.0%), which were corrected.

Regarding the Kappa and Tau metrics, Green et al. (2000) suggested acceptable accuracy values between 60.0% and 80.0% for handling and reference purposes; these values were fully realized in this study. In addition, these authors also expressed the need for greater accuracy when performing quantitative assessments such as a change detection analysis, for which they suggest values of approximately 90.0%.

Some areas that were not classified as *Sargassum* using our methodological approach might have been omitted because the surface coverage of *Sargassum* within a pixel was minimal. Thus, the spectral response did not reach the *Sargassum* detectability values and might have been confused with another object in the image. In the supervision of the *Sargassum* classification directly derived from the classification process, the areas suspected of presence were not added, thereby underestimating the area covered.

In terms of the general errors of omission, when using Landsat images, Hu, Hardy, and Hochberg (2015) reported uncertainty values about 30.0% when estimating *Sargassum* coverage. This uncertainty value and a correction factor they applied were obtained for the northern Gulf of Mexico where the characteristics of the lines of *Sargassum* differed from those in the Caribbean. Therefore, it is not well established whether this factor can be accurately applied to our known underestimations; however, it represents a reference for studies using remote sensing to detect *Sargassum*.

The application of vegetation indices to detect floating *Sargassum* involved the inevitable risk of detecting other floating objects that have a radiometric response in the same range of the light spectrum (Hu, Hardy, and Hochberg 2015, Xing et al. 2017). In particular, spectral confusion of *Sargassum* with *Syringodium* and bacteria of the genus *Trichodesmium* has been reported. Other objects, which certain indices might detect in the same radiometric range as *Sargassum*, correspond to floating litter. However, Hu, Hardy, and Hochberg (2015) acknowledge that no large aggregations of litter have been reported in the Gulf of Mexico.

As mentioned above and recognized by Hu et al. (2016) and Wang and Hu (2016), it is difficult to assert that all of the objects detected in this study correspond to *Sargassum*. Considering the data in the available literature, however, this study shares the same theoretical basis of the red edge in the reflectance of vegetation in the NIR, which has proven useful for the detection of floating vegetation, including *Sargassum*.

Furthermore, considering that *Trichodesmium* bacteria strongly respond to the blue band (Hu, Hardy, and Hochberg 2015) and that the indices generated from the red edge play a more important role in classification than band 2 (blue), the contribution of this band is not so dominant to suggest the significant expression of *Trichodesmium*, even when its importance is relevant. Moreover, these bacteria are likely rare in this region; therefore, *Syringodium* is one of the objects with the greatest probability of being confused with *Sargassum*.

However, these same authors reported a high reflectance of *Syringodium* in bands 2 and 5 of the Landsat images. Considering that the response of pixels detected as *Sargassum* in the spectral window of the image derived from the ratio between bands 2 and 5 had lower average values than the red-edge indices (i.e., FAI, NDVI, SAVI, and ARVI), it is assumed that most of the pixels classified as *Sargassum* are in fact *Sargassum*.

In addition, the dates during 2015 in which the greatest amounts of floating objects assumed to be *Sargassum* were detected coincide with the arrival dates of these macroalgae along the coasts of the Mexican Caribbean as well as with the temporal patterns detected by Wang and Hu (2016) in the WCA. In this context, a high probability exists that most objects identified as *Sargassum* were correctly classified.

Multi-index methodological approach

Reported specialized image processing for *Sargassum* detection addresses several issues regarding satellite image conditions, although they have great potential for automatization. Wang and Hu (2016) implemented an effective procedure to detect *Sargassum* using MODIS images to systematize the detection process for thousands of images. Their approach was highly effective and has many applications and uses. However, it requires a level of technical specialization and computer equipment that could be important constraining factors for researchers with limited resources and decision makers who require access to simple and reliable evaluation strategies. These groups have the greatest need for attainable approaches that enable the assessment of the presence and distribution of lines of floating macroalgae in their study areas.

Given the need to know the spatial-temporal patterns and the origins of algal blooms, we proposed an approach to detect *Sargassum* in a low cost, efficient and technologically viable alternative that considers different indices that can detect floating vegetation. Because our approach includes the use of well known, standardized, and effective informatics tools (e.g., a machine learning family classifier; Akar and Güngör, 2012), it is more advantageous and convenient. As with every supervised classification algorithm, the contribution of diverse inputs constitutes one of the most robust aspects of the development of decision trees. Therefore, seven different sources were used for *Sargassum* detection in contrast to a single source.

Several studies have sought to define threshold values for certain indices (NDVI, FAI, NDVI, among others) to identify pelagic *Sargassum* (Hu 2009; Gower and King 2011; Gower, Young, and King 2013; Hu, Hardy, and Hochberg 2015; Hu et al. 2016, Xing et al. 2017). Nevertheless, several conditions such as fog, haze, and sun glint are present in the L8 images that do not allow the direct optimal separability of *Sargassum* from the remaining objects in the image; thus, this multi-index supervised classification was applied.

Most vegetation indices used in this approach have been tested for use to detect *Sargassum* and other floating macroalgae (Hu 2009; Wang and Hu 2016; Xing and Hu 2016) and are considered viable alternatives given the scarcity of hyperspectral products that decisively resolve the radiometric confusion of *Sargassum* with other objects. In general, the indices assessed in this study showed a narrower range of values in the *Sargassum* pixels than the bands alone, which suggests that the indices more directly captured the objects assumed to be *Sargassum* in the images.

The similarity of the index values between the classes was evident, as were the variations in the conditions within each of the assessed scenes. In this sense, the proposed approach makes use of the separability of all of the indices combined with the classification algorithm working in a seven-dimensional space to address each situation represented by an image in a particular way (Figure 6).

Regarding the separability of the *Sargassum* class, the most important indices were NDVI and SAVI, followed by band 2. Although the contribution of the latter to the classification of objects other than *Sargassum* suggests that this element is not diagnostic *per se*, it can be used to detect *Sargassum* in association with another band such as band 5.

The functionality of this methodological approach to detect pelagic *Sargassum* was corroborated along with the possibility of transference to and implementation in any other area of interest. Standard remote sensing as well as geographic information system applications and analyses were used, and they can be replicated using a wide range of spatial analysis software. This method further represents a consistent, accessible, and versatile alternative approach that can be adopted and implemented in other regions and by other groups.

The adopted methodological approach includes both automated stages and supervision conducted by operators. Moreover, relatively low cost tools are used to implement algorithms for image processing and classification. We sought a cost-effective balance in which the supervision added value to the final result. The cost is minimal given that a quantitative coverage can be obtained without efforts to digitize the *Sargassum* coverage, a task that can be arduous and is subject to human error. In addition, this supervision is aided by diagnostic tools; therefore, only a minimal level of experience is required to identify the objects in the images.

Users must be previously trained to correctly identify *Sargassum* in satellite images using color composites as a reference (Wang and Hu, 2016). Following our approach, index composites are incorporated into the identification task, enabling a clearer and more precise definition of the training sites and the supervision of results. White pixels indicate a high value for the three indices, thereby allowing a color spectrum to distinguish *Sargassum* from its context.

As Wang and Hu (2016) documented, the intervention of a *Sargassum* observation operator constitutes an alternative to improve detection accuracy. Although this method implies an additional cost, it might be significantly more cost-effective than manually digitizing *Sargassum* polygons (*e.g.,* Xing et al. 2017), thereby increasing the analytical capacity in terms of the number of images.

In this study, personnel inexperienced with satellite image analysis were successfully trained to define the training sites. With the help of false color composites, calculated indices, and probabilities obtained using the classification algorithm, the polygons erroneously assigned as *Sargassum* by the classifier were identified and eliminated. The cost involved in this operation was profitable in terms of correcting and increasing the level of accuracy of the images with a minimum time investment because false assignments were minimal and well identified. The performance of the people trained to classify the images was verifiable through the accuracy metrics of the images.

Conclusions

This approach and its associated results represent an alternative methodological frame for the systematic and transferable detection of *Sargassum*.

We found that combining vegetation indices in a classification process is a more robust way to detect *Sargassum* in L8 images than using a single index. In addition, although the levels of confusion between the object of interest, the *Sargassum*, and other objects such as clouds, cloud shadows, and sun glint constitute an important challenge, this approach proposes a semi-automated classification process that provides the user with a margin to adjust to the particular conditions of each image.

These results are the first such assessment of the southeastern part of the Gulf of Mexico and the Mexican Caribbean. This quantification represents a baseline reference and it provides a further understanding of *Sargassum* coverage and dynamics.

Given the possibility that similar massive blooms of this macroalga will occur in the future, this method represents an easily transferable approach that will support the monitoring and management of coastal areas affected by anomalous events.

The present work assessed *Sargassum* lines in Mexican territorial waters from the Gulf of Mexico to the Mexican Caribbean in 2014 and 2015; however, the approach used in this study is applicable to any region or moment of interest for which L8 images are available.

Acknowledgments

This work was supported by the Hydrocarbon Fund of the National Council for Science and Technology (SENER-CONACyT Hidrocarburos), project No. 201441, and the funding source did not participate in the design of this assessment. This study is part of project No. 201441 "Implementation of oceanographic observation networks (physical, geochemical, ecological) to generate scenarios in the face of possible contingencies related to hydrocarbon exploration and production in the deep waters of the Gulf of Mexico". Thanks to Pronatura Peninsula de Yucatan for its operative support through our collaboration on the project "Assessment and monitoring of ecosystems for immature sea turtles in Yum Balam" (SAM Fund/FMCN A1605007 MEX008-024). We are particularly grateful to Chuanmin Hu, Ph.D. of the College of Marine Science at the University of South Florida for his support and advice concerning the structuring and instrumentation of the satellite image atmospheric correction process. We thank Robert Hardy for sharing his experience on *Sargassum* detection. We especially thank Héctor Hérnandez Nuñez for preprocessing the images. We also thank David Espinosa Puch for his key contribution to the analysis of the images as part of an internship.

References

Akar, Ö. & Güngör, O. (2012). Classification of multispectral images using Random Forest algorithm. *Journal of Geodesy and Geoinformation*, 2, 105 – 112.

Anderson, R. P., Lew, D. & Peterson, A. T. (2003). Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, 162, 211 – 232.

Azanza-Ricardo, J. & Pérez-Martín, R. (2016). Impacto de la acumulación de sargazo del verano del 2015 sobre las tortugas marinas de Playa La Barca, península de Guanahacabibes. *Revista Investigaciones Marinas*, 1, 54 – 62.

Breiman, L. (1999). Random forests—random features. Technical Report 567, Statistics Department, University of California, Berkeley (ftp://ftp.stat.berkeley.edu/pub/users/breiman).

Butler, J. N., Morris, B. F., Cadwallader, J. & Stoner, A. W. (1983). Studies of *Sargassum* and the *Sargassum* Community. *Bermuda Biological Station, Special Publication No.22*.

Crisci, C., Ghattas, B. & Perera, G. (2012). A review of supervised machine learning algorithms and their application to ecological data. *Ecological Modelling*, 240, 113 – 122.

Enriquez, C., Mariño-Tapia, I. J. & Herrera-Silveira, J. A. (2010). Dispersion in the Yucatan coastal zone: Implications for red tide events. *Continental Shelf Research*, 30, 127 – 137.

Gavio, B., Rincón-Díaz, M. N. & Santos-Martínez, A. (2015). Massive quantities of pelagic *Sargassum* on the shores of San Andres Island, Southwestern Caribbean. *Acta Biológica Colmbiana*, 1, 239 – 241.

Gower, J., Hu, C., Borstad, G. & King, S. (2006). Ocean color satellites show extensive lines of floating Sargassum in the Gulf of Mexico. *IEEE Transactions on Geoscience and Remote Sensing*, 12, 3619 – 3625.

Gower, J. & King, S. (2008). Satellite images show the movement of floating *Sargassum* in the Gulf of Mexico and Atlantic Ocean. *Nature proceedings*, http://dx.doi.org/10101/npre.2008.1894.1.

Gower, J. F. R. & King, S. A. (2011). Distribution of floating *Sargassum* in the Gulf of Mexico and the Atlantic Ocean mapped using MERIS. *International Journal of Remote Sensing*, 7, 1917 – 1929.

Gower, J., Young, E. & King, S. (2013). Satellite images suggest a new *Sargassum* source region in 2011. Remote Sensing Letters, 8, 764 – 773.

GRASS Development Team, (2017). Geographic Resources Analysis Support System (GRASS) Software, Version 7.2. Open Source Geospatial Foundation. Electronic document:. http://grass.osgeo.org

Green, E. P., Mumby, P. J., Edwards, A. J. & Clark, C. D. (2000). Field survey: Building the link between image and reality. En: A. J. Edwards (Ed.) Remote sensing handbook for tropical coastal management. *Coastal Management Sourcebooks 3, UNESCO,* Paris. Pp. 57 – 65.

Guyot, G. & Gu, X. F. (1994). Effect of radiometric corrections on NDVI-Determined from SPOT-HRV and Landsat-TM data. *Remote Sensing of the Environment*, 49, 169 – 180.

Hardy, R. F. (2014). Assessments of Surface-pelagic drift communities and behavior of early juvenile sea turtles in the Northern Gulf of Mexico. Tesis de Maestría en Ciencias. College of Marine Science, University of South Florida. #1569947. St. Petersburg, FL. 131 p.

Hijmans, R.J., J. Kapoor, J. Wieczorek, N. Garcia, A. Maunahan, A. Rala, A. Mandel. (2016). GADM database of Global Administrative Areas. Downloaded de: http://www.gadm.org on 11/2016

Hu, C., Chen, Z., Clayton, T. D., Swarzenski, P., Brock, J. C. & Muller-Karger, F. E. (2004). Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: Initial results from Tampa Bay, FL. *Remote Sensing of Environment*, 93, 423 – 441.

Hu, C. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, 113, 2118 – 2129.

Hu, C., Feng, L., Hardy, R. F. & Hochberg, E. J. (2015). Spectral and spatial requirements of remote measurements of pelagic *Sargassum* macroalgae. Remote Sensing of Environment, http:/dx.doi.org/10.1016/j.rse.2015.05.022.

Hu, C., Hardy, R., Ruder, E., Geggel, A., Feng, L., Powers, S., Hernandez, F., Graettinger, G., Bodnar, J. & McDonald T. (2016). *Sargassum* coverage in the northeastern Gulf of Mexico during 2010 from Landsat and airborne observations: Implications for the Deepwater Horizon oil spill impact assessment. *Marine Pollution Bulletin*, 107, 15 – 21.

Huete, A. R. (1988). A soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25, 295 – 309.

Huffard, C. L., von Thun, S., Sherman, A. D., Sealey, K. & Smith Jr., K. L. (2014). Pelagic *Sargassum* community change over a 40-year period: temporal and spatial variability. *Marine Biology*, 161, 2735 – 2751.

Kaufman, Y. J. & Tanré, D. (1992). Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 2, 261 – 270.

Kehoe, M., O'Brien, K., Grinham, A., Rissik, D., Ahern, K. S. & Maxwell, P. (2012). Random forest algorithm yields accurate quantitative prediction models of benthic light at intertidal sites affected by toxic Lyngbya majuscule blooms. *Harmful Algae*, 19, 46 – 52.

Liaw, A. & Wiener, M. (2002). Classification and regression by randomForest. *R News* (ISSN 1609-3631) 2/3, 18 – 22.

Lucieer, V., Hill, N. A., Barrett, N. S. & Nichol, S. (2012). Do marine substrates 'look' and 'sound' the same? Supervised classification of multibeam acoustic data using autonomous underwater vehicle images. *Estuarine, Coastal and Shelf Science*, 2012, 1 – 13.

McCarthy, S. (2016). Automated *Sargassum* Detection for Landsat Imagery. 2016 Ocean Science Meeting AGU/ASLO/The Oceanography Society. New Orleans, U. S. A. [online] (https://agu.confex.com/agu/os16/preliminaryview.cgi/Paper89731.html) October, 2016.

Mansfield, K. L. & Putman, N. F. (2003). Oceanic habits and habitats: *Caretta caretta*. En: J. Wyneken, K. J. Lohmann, J. A. Musick (Eds). The Biology of Sea Turtles Volume III. CRC Press Taylor & Francis Group, Boca Raton, E. E. U. U. Pp: 189 – 210.

Maurer, A. S., De-NEef, E. & Stapleton, S. (2015). *Sargassum* accumulation may spell trouble for nesting sea turtles. Frontiers in Ecology, 2015, 394 – 396.

Merino, M. (1997). Upwelling on the Yucatan Shelf: hydrographic evidence. *Journal of Marine Systems*, 13, 101 – 121.

Moser, M. L. & Lee, D. S. (2012). Foraging over *Sargassum* by western north Atlantic seabirds. *The Wilson Journal of Ornithology*, 1, 66 – 72.

Optical Oceanography Laboratory, University of South Florida. Satellite-based Sargassum Watch System (SaWS). Consulted on February, 2016 website: http://optics.marine.usf.edu/projects/SaWS.html

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 1, 217 – 222.

Reyes-Mendoz, O., Mariño-Tapia, I., Herrera-Silveira, J., Ruíz-Martínez, G., Enriquez, C. & Largier, J. L. (2016). The effects of wind on upwelling off Cabo Catoche. *Journal of Coastal Research*, 3, 638 – 650.

Rousset, C. & Beal, L. M. (2010). Observations of the Florida and Yucatan Currents from a Caribbean Cruise ship. *Journal of Physical Oceanography*, 40, 1575 – 1581.

Schell, J. M., Goodwin, D. S. & Siuda, A. N. S. (2015). Recent *Sargassum* inundation events in the Caribbean. *Oceanography*, 3, 8 – 10 http://dx.doi.org/10.5670/oceanog.2015.70.

Sheinbaum, J., Candela, J., Badan, A. & Ochoa, J. (2002). Flow structure and transport in the Yucatan Channel. *Geophysical Research Letters*, 3, 10-1 – 10-4.

Schmitz, W.J., Biggs, D.C. Lugo-Fernandez, A., Oey, L., & Sturges, W. (2005). A Synopsis of the Circulation in the Gulf of Mexico and its Continental Margins. En: Circulation in the Gulf of Mexico: Observations and Models. Eds: Wilton Sturges, Alexis Lugo-Fernandez. *Geophysical Monograph Series 161.*

Smetacek, V. & Zingone, A. (2013). Green and Golden seaweed tides on the rise. *Nature*, 504, http://dx.doi.org/10.1038/nature12860.

South Atlantic Fishery Management Council. (2002). Fishery management plan for *Sargassum* habitat of the South Atlantic region. Second Revised Final. NMFS/NOAA/. Charleston, USA. 153p.

Stoner, A. W. (1983). Pelagic *Sargassum*: Evidence for a major decrease in biomass. *Deep-Sea Research*, 4A, 469 – 474.

Thiel, M. & Gutow, L. (2005). The ecology of rafting in the marine environment: I. The floating substrata. *Oceanography and Marine Biology: An Annual Review*, 42, 181 – 264.

USGS. (2014). Landsat 8 operational land imager (OLI) Quality Assessment (QA) band support. User Guide, Landsat land data operational product evaluation Toolbelt. Version 1.2, December 2014. Department of the Interior, U. S. Geological Survey. 17p.

USGS. (2016). Landsat 8 Operational Land Imager images courtesy of the U.S. Geological Survey. Downloaded on April, 2016 from https://earthexplorer.usgs.gov/

USGS. (2017). Product Guide, Landsat surface reflectance-derived spectral indices. Vesion 3.5, April 2017. Department of the Interior, U. S. Geological Survey. 31 p.

Van Beijma, S., Comber, A. & Lamb, A. (2014). Random forest classification of salt marsh vegetation hábitats using quad-polarimetric airborne SAR, elevation and optical RS data. *Remote Sensing of Environment*, 149, 118 – 129.

Vandendriessche, S., Messiaen, M., O'Flynn, S., Vincx, M. & Degraer, S. (2007). Hiding and feeding in floating seaweed: Floating seaweed clumps as possible refuges or feeding grounds for fishes. *Estuarine, Coastal and Shelf Science*, 71, 691 – 708.

van-Tussenbroek, B., 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. Marine Pollution Bulletin, 122, 272 – 281.

Wang, M. & Hu, C. (2016). Mapping and quantifying *Sargassum* distribution and coverage in the Central West Atlantic using MODIS observations. *Remote Sensing of Environment*, 183, 350 – 367.

Webster, R. K. & Linton, T. (2013). Development and implementation of *Sargassum* Early Advisory System (SEAS). *Shore & Beach*, 3, 1 – 6.

Witherington, B., Hirama, S. & Hardy, R. (2012). Young sea turtles of the pelagic *Sargassum*-sominated drift community: habitat use, population density, and threats. *Marine Ecology Progress Series*, 463, 1 – 22.

Xing, Q. & Hu, C. (2016). Mapping macroalgal blooms in the Yellow Sean and East China Sea using HJ-1 and Landsat data: Application of a virtual baseline reflectance height technique. *Remote Sensing of Environment*, 178, 113 – 126.

Xing, Q., Guo, R.,Wu, L., An, D., Cong, M.,Qin, S., Li, X. (2017). High-Resolution Satellite Observations of a New Hazard of Golden Tides Caused by Floating Sargassum in Winter in the Yellow Sea. *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1815-1819. doi: 10.1109/LGRS.2017.2737079

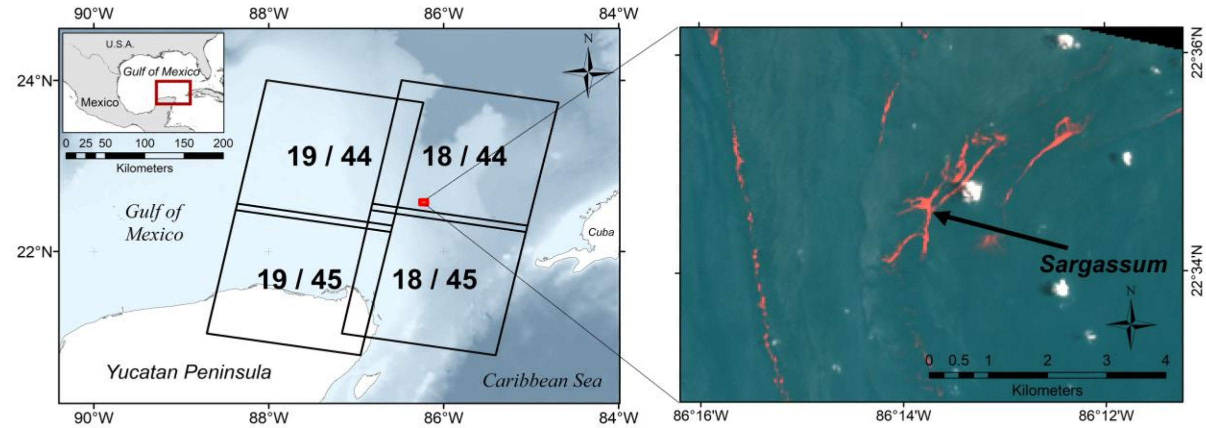

Figure 1. Study area located in the northeastern region of the Yucatan Peninsula, México (*a*). Black-lined polygons delimit the assessed scenes, and in a red square the small area represented in (*b*); (*b*) *Sargassum* driftlines are indicated by a black arrow, as seen in a color composition (bands 5, 2, 1) from a satellite image from Path 18 Row 45, July 2015.

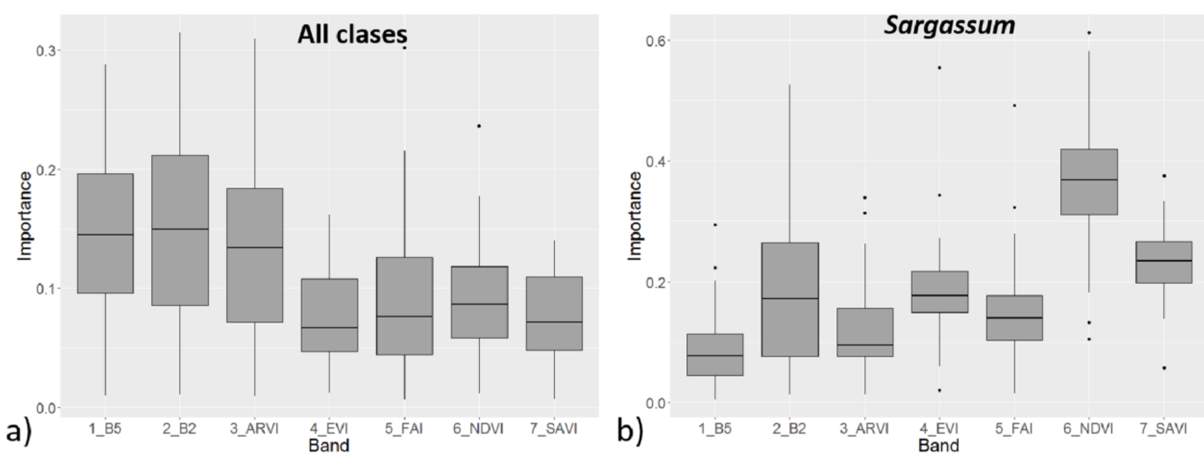

Figure 2. The importance of each of the indices and bands used for the supervised classification (Random Forest) of 54 Landsat images for all classes considered (*a*) and for the *Sargassum* class (*b*). The importance of a variable for a random forest model is estimated by looking at how much prediction error increases when the testing data for that variable is permuted while all others are left unchanged (Liaw & Wiener, 2002).

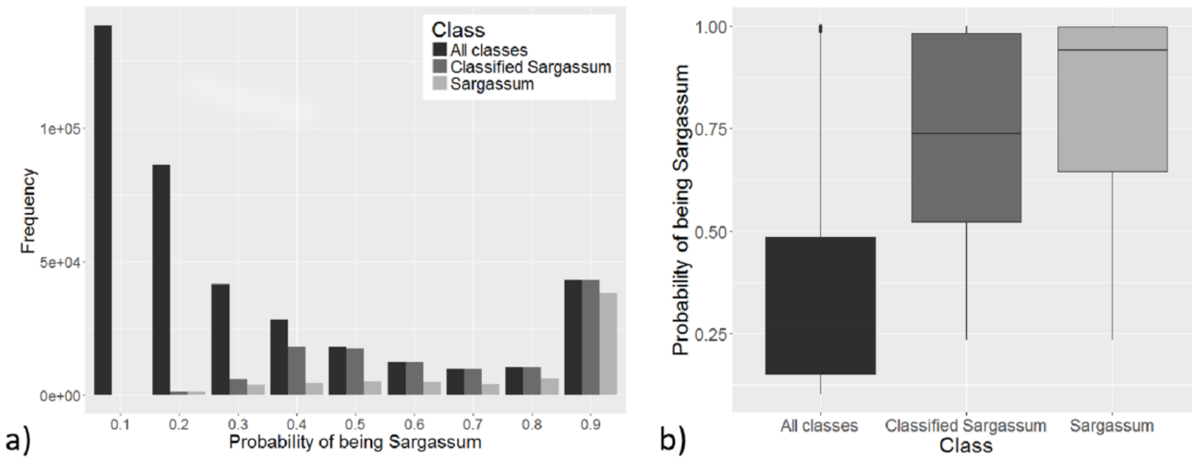

Figure 3. (*a*) Distribution of the frequencies of probability values of being *Sargassum* for the pixels in the whole scene (black) as well as the values for the pixels classified as *Sargassum* by the algorithm (dark gray) and those correctly assigned as *Sargassum* and refined by supervision (light gray). (*b*) The central tendency and dispersion values are presented for the three sets of pixels analyzed. Pixels with a probability lower than 0.1% were removed.

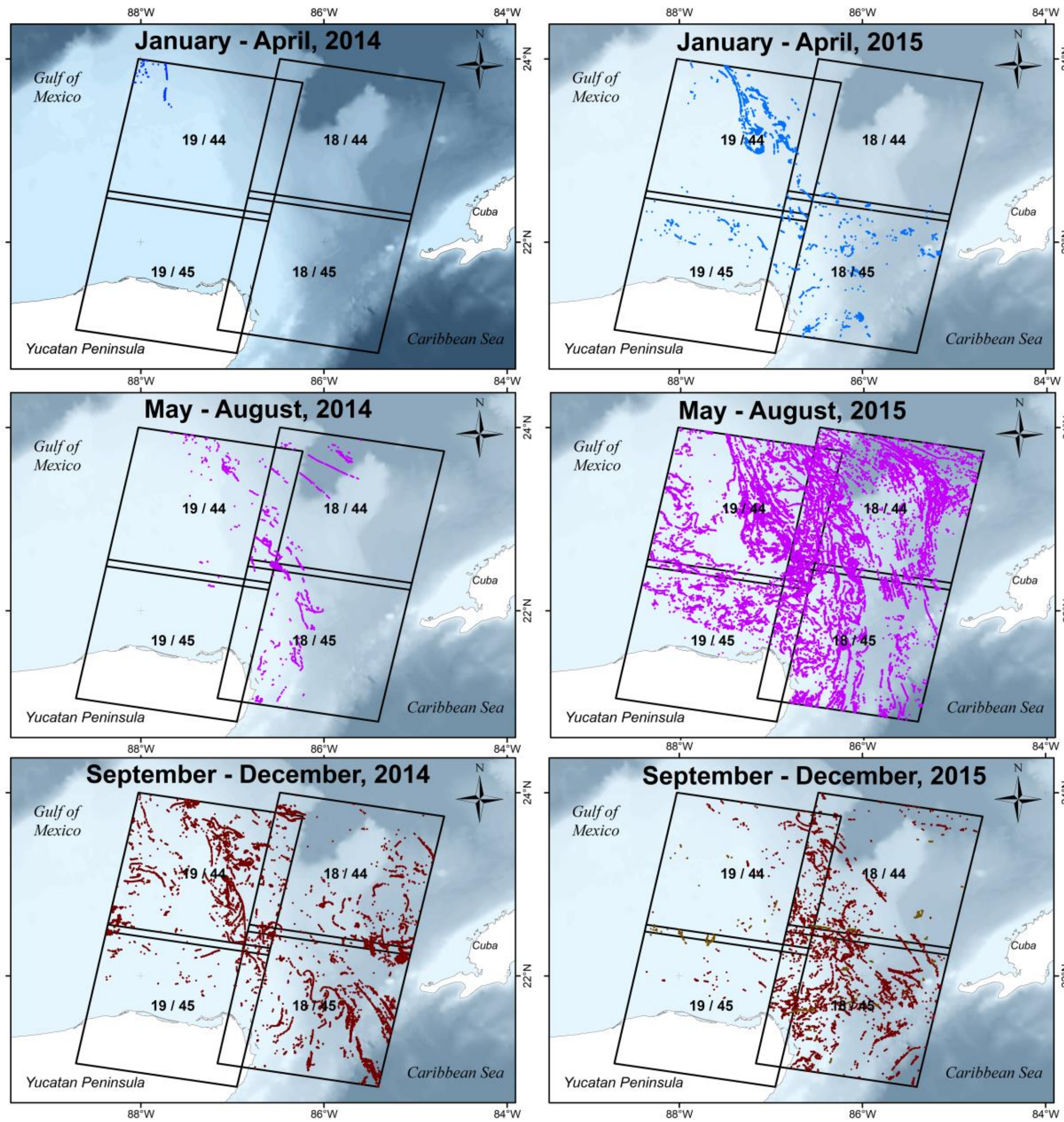

Figure 4. Spatial distribution of *Sargassum* drifts (*Sargassum* spp.) in the northeastern region of the Yucatan Peninsula, Mexico during 2014 and 2015. Each frame shows a four-month composite of all of the *Sargassum* detected during the correspondent period. Year 2014 is on the left side, and year 2015 is on the right side. (a) January to April, 2014; (b) May to August, 2014; (c) September to December, 2014; (d) January to April, 2015; (e) May to August, 2015; (f) September to December, 2015.For presentation purposes, the lines that delimit the polygons are 1-point thick.

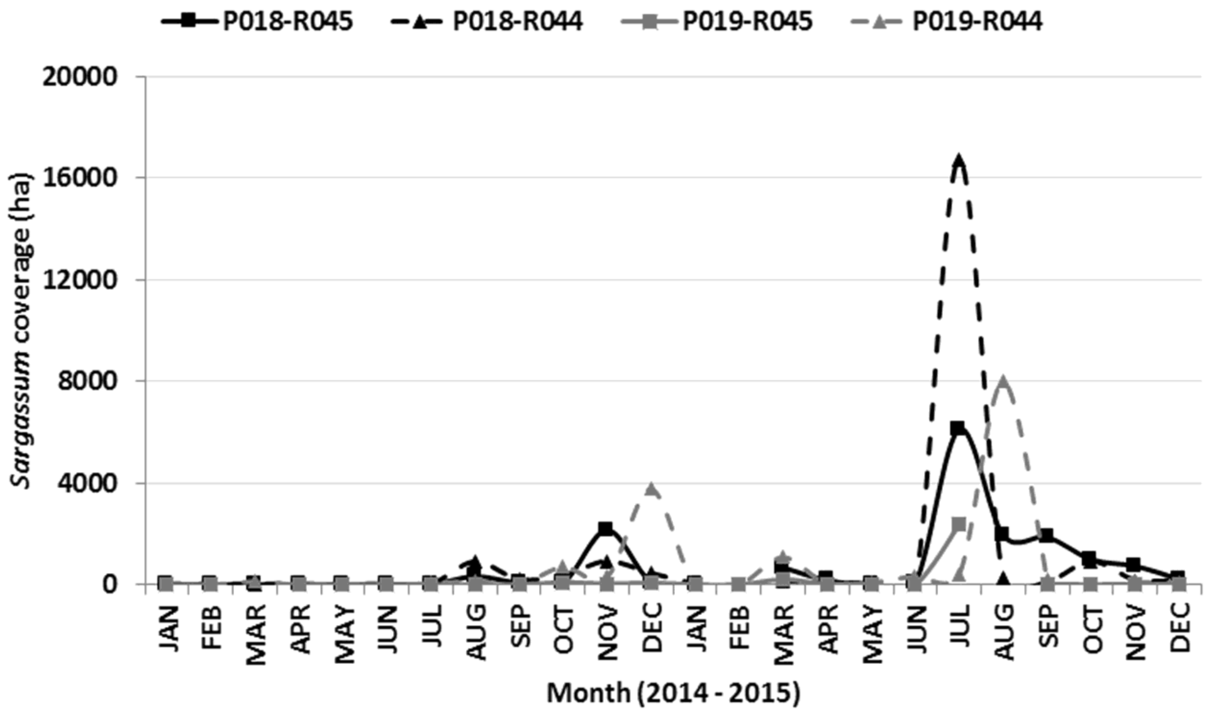

Figure 5. Temporal pattern of the cumulative coverage of the *Sargassum* drift lines in the study area in 2014 and 2015. [Figure in grayscale]

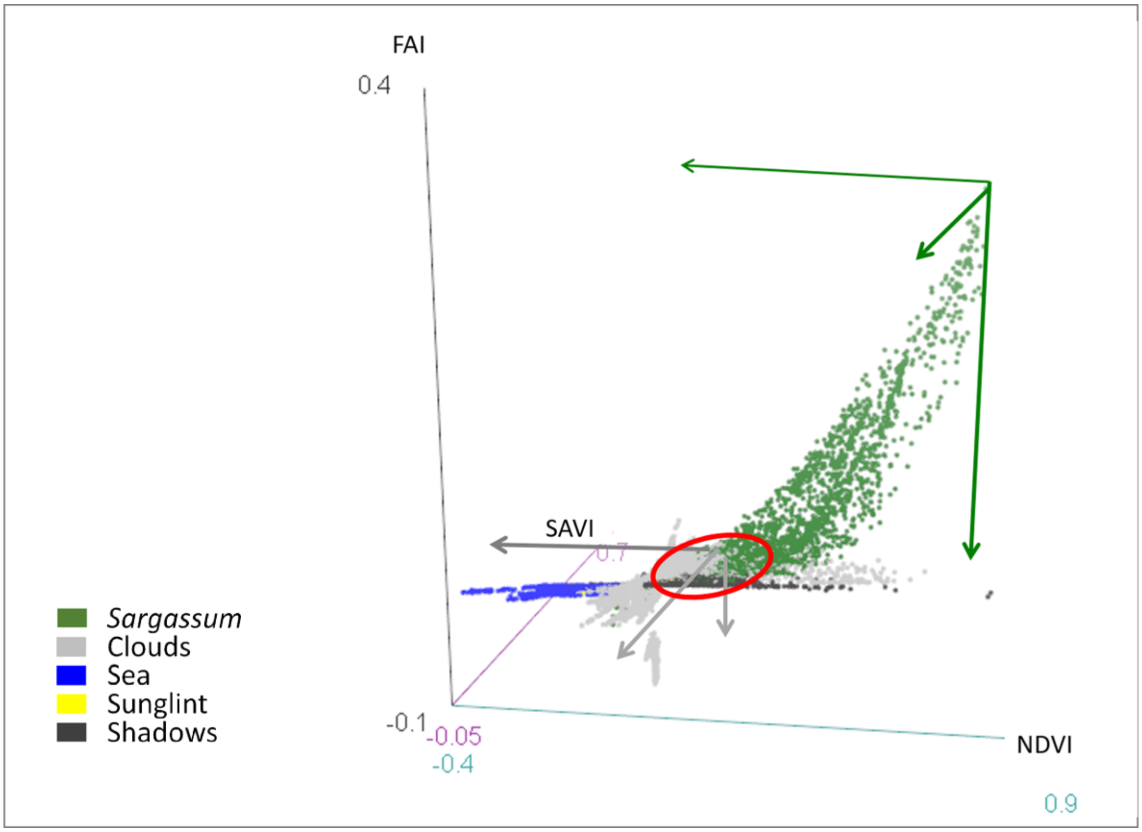

Figure 6. A simplified 3D representation of the informational elements and multi-index criteria provided for the classification algorithm (Random Forest) used in the proposed classification approach, which maximizes the separability between classes.

Table 1. Synthesis of the vegetation index calculi and the parameters used for the equations.

Vegetation index	Equation	Parameter	Author
NDVI	NDVI = $(\rho_{\mathrm{NIR}} - \rho_{\mathrm{R}}) / (\rho_{\mathrm{NIR}} + \rho_{\mathrm{R}})$ Equation (1)	ρ = atmospherically corrected reflectance band; NIR*:* Near infra-red band, R*:* Red band; B*:* Blue band;	Guyot and Gu 1994
ARVI	ARVI = $(\rho_{\mathrm{NIR}} - \rho_{\mathrm{R,B}}) / (\rho_{\mathrm{NIR}} + \rho_{\mathrm{R,B}})$ Equation (2)		Kaufman and Tanré 1992
SAVI	SAVI = $[(\rho_{\mathrm{NIR}} - \rho_{\mathrm{R}}) / (\rho_{\mathrm{NIR}} + \rho_{\mathrm{R}} + L)] (1 + L)$ Equation (3)	L = 0.5, is the context-dependent adjustment factor, which in this case is assumed to be the marine water in which *Sargassum* floats.	Huete 1988
EVI	EVI = G $(\rho_{\mathrm{NIR}} - \rho_{\mathrm{R}}) / (\rho_{\mathrm{NIR}} + C_{1,\mathrm{R}} - C_{2,\mathrm{B}} + L)$ Equation (4)	G is a gain factor (2.5); $C_{1,R}$ (6.0) and $C_{2,B}$ (7.5) are the aerosol resistance coefficients.	USGS, 2017
FAI	FAI = $\rho_{rc,NIR} - \rho'_{\mathrm{rc,NIR}}$ Equation (5) Where $\rho'_{\mathrm{rc,NIR}} = \rho_{\mathrm{rc,R}} + (\rho_{\mathrm{rc,SWIR}} - \rho_{\mathrm{rc,R}}) (\lambda_{\mathrm{NIR}} - \lambda_{\mathrm{R}}) / (\lambda_{\mathrm{SWIR}} - \lambda_{\mathrm{R}})$ Equation (6)	$\rho'_{\mathrm{rc,NIR}}$ is the base reflectance in the NIR band derived from a linear interpolation between the red (R) and the short-wave infrared (SWIR).	Hu 2009

Table 2. Quantitative metrics used in this study (adapted from Anderson et al. 2003) (Green et al. 2000) and their values calculated for all of the validated images with objects classified as *Sargassum* in the study area (N = 17).

Metric	Equation	Parameter

Overall performance (correct classification rate)	(a + d) / (a + b + c + d) Equation (7)	a = true positives;
		b = false positives;
		c = false negatives; and,
		d = true negatives.

Omission error (false negative rate)	c / (a + c) Equation (8)	
Commission index (false positive rate)	b / (b + d) Equation (9)	
Kappa	$K = \frac{N\sum_{i=1}^{r} X_{i,i} - \sum_{i=1}^{r}(X_{i,+}\, X_{+,i})}{N^2 - \sum_{i=1}^{r}(X_{i,+} \cdot X_{+,i})}$ Equation (10)	where *r* is the number of rows in the confusion matrix; $X_{i,i}$ is the number of observations in row *i* and column *i*; $X_{i,+}$ y $X_{+,i}$ are the marginal totals of row *i* and column *i*, respectively; and *N* is the total number of validation points.
Tau	$\tau = \frac{P_o - P_r}{1 - P_r}$, where $P_r = (\frac{1}{N^2}) \sum_{i=1}^{M} n_i\, X_i$ Equation (11)	where P_o is the global accuracy; *M* is the number of classes; *i* is the i^{th} class; *N* is the total number of validation points; n_i is the total of row i; and X_i is the diagonal value of class i.

Table 3. Confusion matrix developed from the points defined to quantitatively validate the *Sargassum* classifications.

Classification	Visual Validation		
Semi-Automated	*Sargassum*	Non *Sargassum*	
Sargassum	392	42	434
Non *Sargassum*	92	1,514	1,606
	484	1,556	2,040