https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Open Access Earth System Science Data Discussions

Daily Sargassum detection from OLCI images over the tropical Atlantic Ocean since 2018 through a multi-index approach

Marianne Debue[1], Thibault Guinaldo[2], Anastasia Tarasenko[3], Stéphane Saux-Picart[1]

[1]Météo-France, CNRS, Univ. Toulouse, CNRM, Centre d'Études en Météorologie Satellitaire, Lannion, France
[2]Météo-France, CNRS, Univ. Toulouse, CNRM, Toulouse, France
[3]CNES, Toulouse, France

Correspondence to: Marianne Debue (marianne.debue@meteo.fr)

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Abstract. This datapaper presents the methodology and the associated dataset of *Sargassum* detection from Sentinel-3 OLCI images developed at Météo-France. The dataset includes daily detections over the whole tropical Atlantic Ocean, the Caribbean Sea and the Gulf of Mexico, from May 2018 up to date. It is updated daily with the latest detections. *Sargassum* is detected through a physically-based approach relying on two indices (the AFAI and the MCI) and using adaptive thresholds. A comparison of this dataset with two existing ones reveals a reduction in masked areas and improved detection of *Sargassum*, particularly in areas of sunglint and at the edges of clouds. The entire dataset is freely available on the Odatis platform and can be downloaded in NetCDF format (DOI: 10.12770/1eb82d09-77ed-4f63-9f03-2c3516a9713d; Météo-France, 2026). Two variables are available: the status of detection and the deviation of the AFAI for pixels including *Sargassum*.

1 Introduction

Since 2011, recurrent massive strandings of *Sargassum* have affected the coasts of the Caribbean and West Africa (Wang et al., 2019), with major consequences for human health (Resiere et al., 2026), coastal ecosystems (Rodríguez-Martínez et al., 2025), and local economies (Rodríguez-Martínez et al., 2025). Mitigating these impacts requires rapid collection of stranded rafts before decomposition, which, in turn, depends on the ability to anticipate the timing and location of these strandings.

Forecasting systems are therefore essential and typically rely on numerical models simulating *Sargassum* drift driven by winds and ocean currents (Podlejski et al., 2023), sometimes also including *Sargassum* growth and mortality (Jouanno et al., 2021). These models, sensitive to initial conditions, require accurate and timely information on the location of *Sargassum* aggregations as input. Satellite remote sensing provides such information, offering near real-time, large-scale observations over far-off ocean regions (Debue et al., 2025b; Hu et al., 2025). Beyond forecasting, these observations are also valuable for operational purposes (e.g. guiding field sampling (Ody et al., 2019) or avoiding navigation hazards such as entanglement in the propellers) and research applications (e.g. estimation of *Sargassum* drift (Minghelli et al., 2021), evaluation of physical and biogeochemical models (Podlejski et al., 2023), understanding of the origin and variability of the phenomenon (Jouanno et al., 2025)).

Several monitoring and forecasting systems have been developed to provide information on *Sargassum* location and associated stranding risks (Putman et al., 2025). These systems differ in terms of data sources (e.g. satellite detection or in situ observations, models including only drift or also growth/decay), spatial coverage (e.g. regional to basin scale), update frequency, intended use (e.g. visualization, downloadable data) and accessibility (open or paid access). Among them, the *Sargassum* Watch System (Hu et al., 2016) (SaWS), developed by the University of South Florida, provides daily basin-scale satellite detection, transport modeling outputs and monthly reports based on peer-reviewed methods (Wang and Hu, 2016). However, this platform does not provide direct access to downloadable *Sargassum* aggregations location data.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Conversely, the SAMtool platform (CLS, 2026), developed by Collecte Localisation Satellites (CLS), allows data download but with a two-day delay. Furthermore, the methodology is not fully documented in the peer-reviewed literature, limiting reproducibility and comparison.

In 2019, the French government mandated Météo-France to implement an operational system for the detection and forecasting of *Sargassum* strandings affecting the French Caribbean islands, establishing a unique public institutional service at the European scale. Within this framework, Météo-France produces weekly or bi-weekly bulletins to inform authorities and the public about *Sargassum* strandings risks, including short-term forecasts (up to four days) and two-week trend. These bulletins rely on satellite-derived detections from multiple sensors, including Sentinel-3 Ocean and Land Color Instrument (S3-OLCI), combined with drift simulations using the MOTHY model (Debue et al., 2025a).

The dataset presented here consists of daily maps of *Sargassum* aggregation locations derived from S3-OLCI observations, covering the tropical Atlantic basin (latitude: [0° N, 30° N], longitude: [10° W, 100° W]) since 2018, at 0.0032° resolution, and provided in a standardized, downloadable format for use by both the scientific community and local authorities in charge of mitigating the consequences of strandings.

Given the wide range of scientific and operational applications of such observations, and the current lack of freely accessible, near real-time, and consistently processed and documented datasets of *Sargassum* aggregation locations, this dataset fills a critical gap. The Météo-France dataset is made publicly available here and enables applications including model initialization and validation, analysis of spatial and temporal variability, and the development of automated detection or forecasting approaches. This dataset will be appended in the future by other sensors used in operation.

2 Methods

2.1 Overview

The processing workflow is summarized in Fig. 1. It builds upon the approach of Wang and Hu (2016) and includes several adaptations, notably the combined use of two spectral indices and the implementation of adaptive thresholds to account for spatial variability in observation conditions. All the processing chain was implemented in Python. Performance-critical components, including the background estimation (see Sect. 2.5), were optimized using the numba compiler (Lam et al., 2015). All the figures in the Sect. 2 are based on the S3B-OLCI image from March 1st, 2026.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

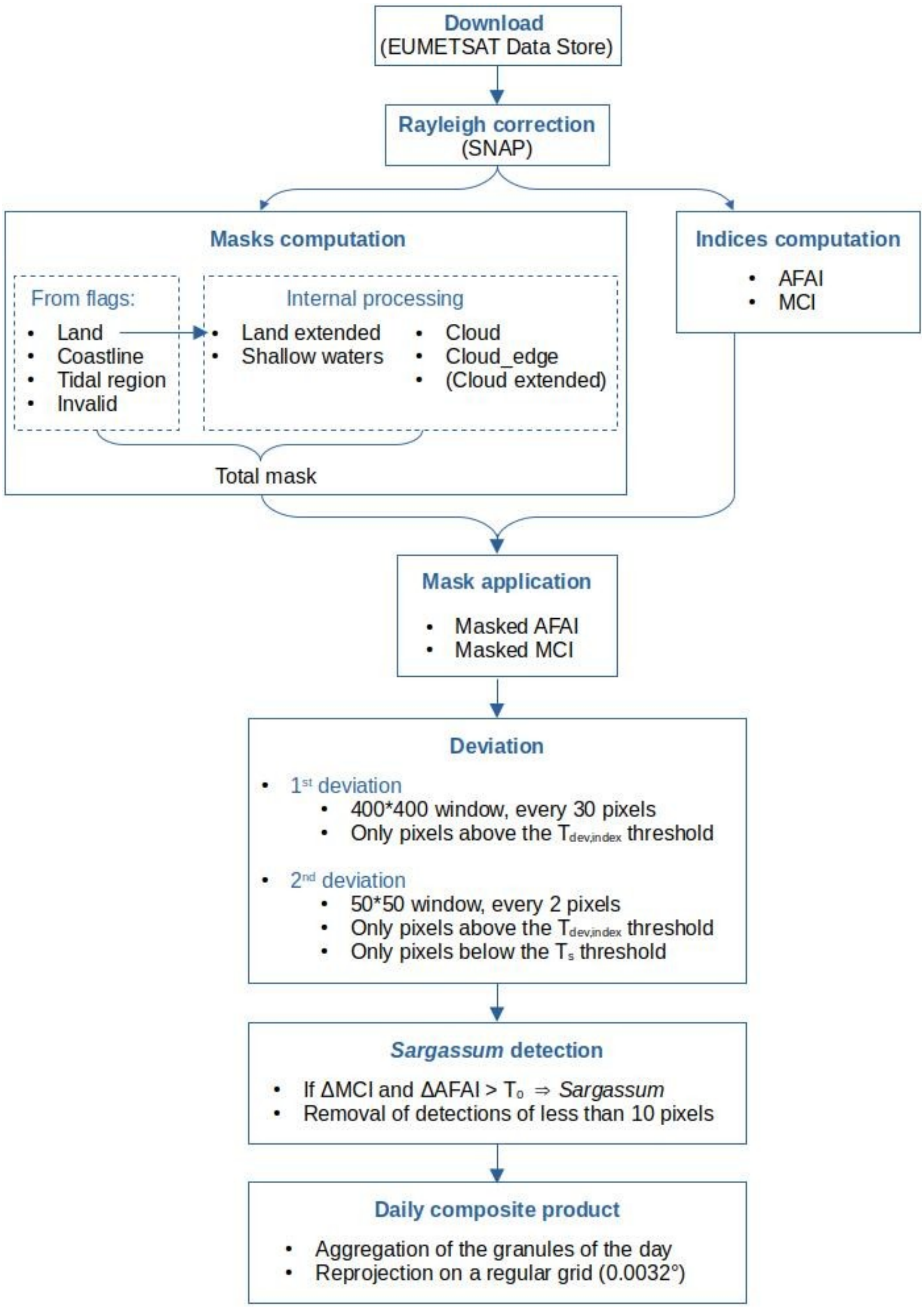

Figure 1: Work flow to produce daily composite product of *Sargassum* detections from S3-OLCI images. See text for more details.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

2.2 Data source and pre-processing

This dataset relies on S3-OLCI images. The S3 satellites are part of the Copernicus program and were launched in February 2016 (S3A) and April 2018 (S3B). They follow a polar orbit, enabling a revisit period of 1 to 2 days. On board, the OLCI multispectral sensor acquires images in 21 bands ranging from 400 to 1020 nm with a spatial resolution of 300 m at nadir.
S3A- and S3B-OLCI Level-1B images, providing Top Of Atmosphere (TOA) radiance, were obtained from the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) data store (Eumetsat, 2026) (collection ID EO:EUM:DAT:0409 and EO:EUM:DAT:0885). All granules acquired in the area of interest since the 8th of May 2018 were processed (approximately 30 * 365 granules per complete year).
Atmospheric correction was performed using the Sentinel Application Platform (SNAP) software (ESA, 2026) to derive bottom-of-Rayleigh reflectance, reducing the contribution of molecular scattering while preserving the signal from floating material. Even if the French Caribbean islands experience recurrent events of Saharan dust transport, no aerosols correction was performed as the correction often relies on the "black pixel" assumption which is not satisfied in the presence of *Sargassum*.

2.3 Masks computation

The detection of *Sargassum* relies on its high reflectance in the near-infrared (NIR) compared to that of water. However, land and clouds also exhibit this type of signal, which can lead to false detections (Podlejski et al., 2022). It is therefore necessary to filter them out. Several masks were thus computed for each granule.
Four masks ("land", "coastline", "tidal_region", and "invalid") were derived from S3-OLCI quality flags.
The land mask was further modified (Fig. 2) to reduce contamination from adjacent pixels and shallow bathymetry (Gower et al., 2013; Hu et al., 2023). A first dilation of 4 pixels (~1200 m) was applied to ensure that all land and water bodies less than 5 meters deep are properly masked. Shallow coastal waters between the two major islands of Guadeloupe were not covered by the 4-pixel dilation and were specifically masked by applying a 20-pixel dilation. The final land mask is the combination of these two masks.
Clouds were detected using reflectance at 865 nm (Fig. 3). To account for longitudinal gradients in illumination, an adaptive threshold was applied, defined as a linear function of longitude, ranging from 0.05 at the western boundary of the granule to 0.2 at the eastern boundary of the granule.
To reduce false detections near cloud edges, a cloud-edge mask was computed. This mask includes pixels located within 10 pixels of the cloud mask and exhibiting a reflectance difference between the 753.75 nm band and the 865 nm band greater than a threshold of 0.004. This threshold was determined empirically by adjusting its value to identify the setting that effectively masks cloud edges, thereby minimizing false detections, while keeping the masked areas to a minimum.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Open Access Earth System Science Data Discussions

A total mask was defined as the combination of all the masks described above and applied to subsequent processing steps. An extended cloud mask was also created by dilating the cloud mask by 10 pixels. This mask is not included in the total mask because detection is performed within it, but is used to implement a different threshold for cloud-edge detection (see Sect. 2.6).

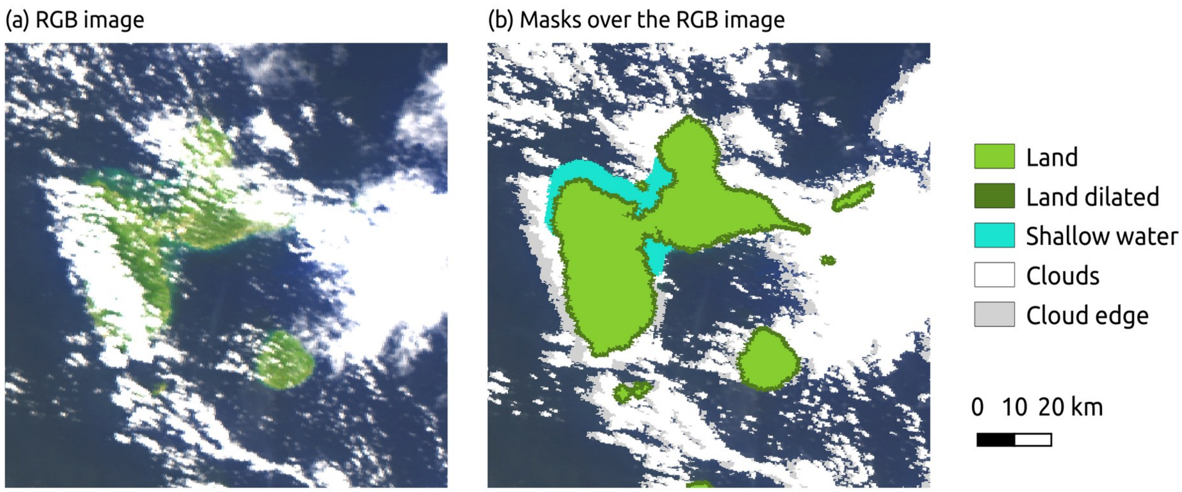

Figure 2: Land and cloud masks around Guadeloupe from S3-OLCI acquisition on 01/03/2026. (a) RGB (Red, Green, Blue) image; (b) Masks over the RGB image.

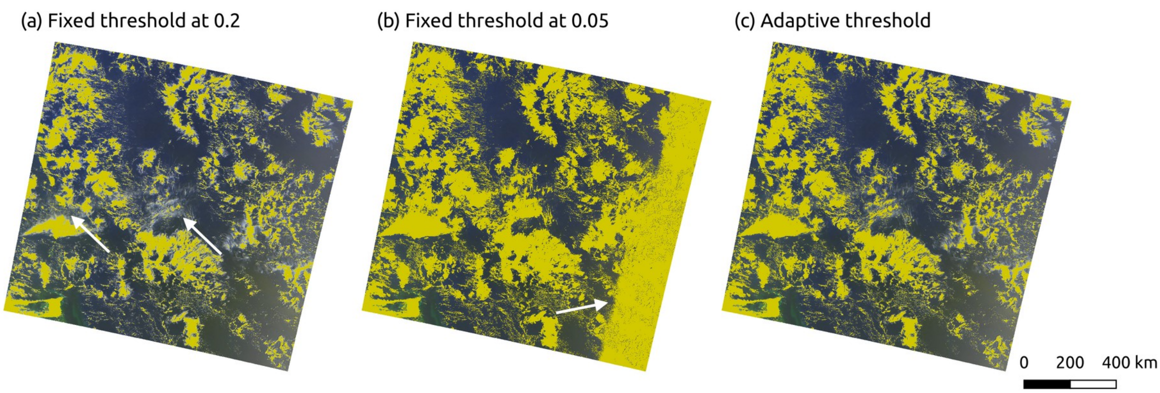

Figure 3: Cloud mask with different thresholds: (a) Fixed threshold at 0.2; some clouds are not masked in the West; (b) Fixed threshold at 0.05; the eastern part of the image is completely masked; (c) Adaptive threshold. Background: RGB from 01/03/2026. White arrows indicate masking defects.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Earth System Science Data Discussions

2.4 Spectral index computation

Floating *Sargassum* modifies the spectral reflectance of ocean surface pixels, in particular with an increase in reflectance in the red and the NIR. The detection of this algae is based on the computation of spectral indices specifically designed to highlight floating vegetation by measuring the reflectance jump around the red-edge (Hu, 2009).

Two spectral indices were computed: the Maximum Chlorophyll Index (MCI) (Gower et al., 2006) and the Alternative Floating Algae Index (AFAI) (Wang and Hu, 2016). Both follow a common formulation (Eq. (1)):

$$Index = R_{\lambda 2} - R_{\lambda' 2} \qquad (1)$$

$$\text{with } R_{\lambda' 2} = R_{\lambda 1} + \left(R_{\lambda 3} - R_{\lambda 1}\right) * \frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_1}$$

with $R_{\lambda 1}$, $R_{\lambda 2}$, and $R_{\lambda 3}$ the reflectance values at wavelength λ_1, λ_2 and λ_3, and [λ_1, λ_2, λ_3] = [681.25, 708.25, 753.75] nm for MCI, [λ_1, λ_2, λ_3] = [665, 753.75, 865] nm for AFAI.

The two indices exhibit complementary sensitivities to environmental conditions. AFAI tends to produce high values in coastal and turbid waters, while MCI is more sensitive to cloud contamination. As a result, relying on a single index may lead to environment-specific false detections. Combining both indices improves the robustness of the detection across a wide range of conditions (Fig. 4).

All masked pixels derived from the total mask (see Sect. 2.3) were excluded before the index computation.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

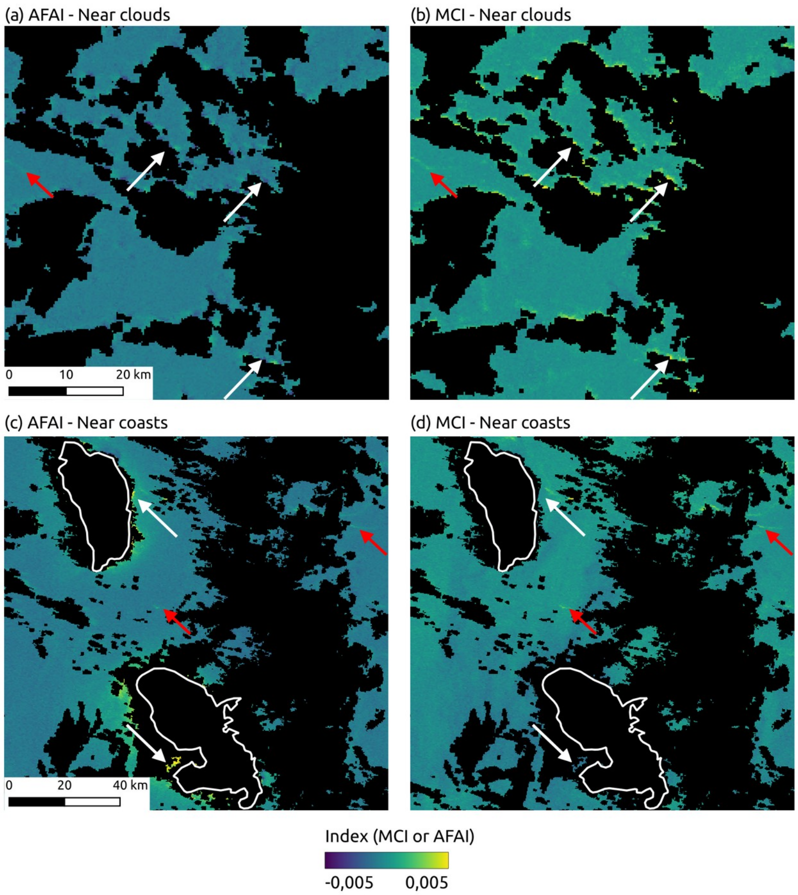

Figure 4: Sensitivity of indices to observation conditions. (a) AFAI near cloud edges; (b) MCI near cloud edges; (c) AFAI near coasts; (d) MCI near costs. The white arrows point to areas where there are differences in index sensitivity. The red arrows point to Sargassum rafts that have high values for both indices. Masked areas in black, shoreline in white.

2.5 *Sargassum* signal increase

Sargassum aggregations typically occupy only a fraction of a pixel and are often submerged. Their spectral signature is thus attenuated by the surrounding water (Qi and Hu, 2021). To increase the signal and account for varying environmental conditions, the detection relies on the identification of local anomalies relative to surrounding water.

In practice, a deviation (ΔIndex) is computed for each pixel and is defined as the difference between the index value of a pixel and the estimated background index value representative of surrounding water (Eq. (2)):

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

$$\Delta Index = Index_i - Index_{bg,i} \qquad (2)$$

where $Index_i$ is the MCI or AFAI value of the pixel i and is defined as the median index value of neighboring pixels of the pixel i assumed to be free of *Sargassum*.

To account for variability at multiple spatial scales, a two-step approach was implemented. First, a large window (400 × 400 pixels) was used to estimate large-scale background variability (e.g. due to sun glint or large-scale gradients). Second, a smaller window (50 × 50 pixels) was used to capture local variability (e.g. wave-induced patterns). To reduce computational cost, subsampling was applied, using one pixel out of 30 for the large window and one out of 2 for the small window.

To avoid artifacts in low-signal areas or highly turbid regions, characterized by low index values, only pixels with index values above a threshold $T_{dev,index}$ were considered in the background estimation. Due to discontinuities between OLCI cameras, in particular between the third and the fourth (Lamquin et al., 2020), corresponding to the 3305th column of the data (Fig. 5), this threshold was defined as a function of pixel column number (N_{col}). Constant thresholds were applied before column 3306, while linearly varying thresholds were used beyond this point to correct for striping effects:

If $N_{col} < 3306$, $T_{dev,MCI} = -0.001$ and $T_{dev,AFAI} = -0.0025$

If $N_{col} >= 3306$, $T_{dev,MCI}$ (resp. $T_{dev,AFAI}$) follows a linear relation with the column number, from -0.001 (resp. -0.0025) at column 3306 to 0.0005 (resp. -0.0015) at column 4865.

For the large-scale deviation, all pixels fulfilling this condition were included in the computation. For the small-scale deviation, in addition to this condition, only pixels whose MCI or AFAI were below the empirical threshold $T_s = 8.10^{-4}$ were included, other pixels being considered as including Sargassum and thus removed from the background computation.

These thresholds were defined through iterative evaluation across twenty images encompassing a wide range of environmental conditions to optimize the compromise between detection sensitivity and false positives.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

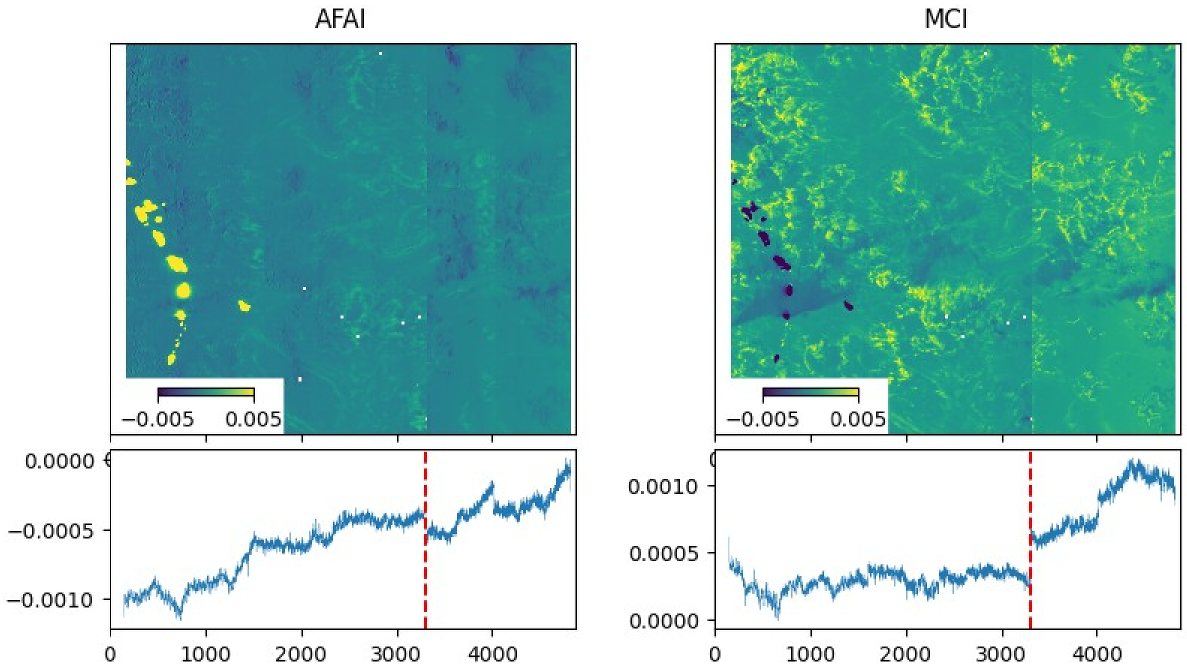

Figure 5: Stripe effect on the AFAI (left) and MCI (right). The upper images correspond to the index without masks to better see the stripes. The lower graphs are the median values of each column of the index after applying the total mask to highlight the increasing value without any land or cloud influence, in particular beyond the column 3305 (red vertical dotted line).

2.6 *Sargassum* detection

Following the deviations computations, each pixel is assigned two values: ΔMCI and ΔAFAI. A pixel was classified as containing *Sargassum* only if both values exceeded a detection threshold T_0 (Fig. 6). This threshold was set at 3.10^{-4} to reduce false detections associated with index-specific artifacts. The value was determined empirically to strike a balance between signal and noise, and corresponds to the 88th percentile of the ΔAFAI distribution. For pixels located within the extended cloud mask, a higher threshold ($T_0 = 6.10^{-4}$) was applied to limit false detection near clouds. A spatial filtering was applied to remove isolated detections smaller than 10 connected pixels.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Earth System Science Data
Open Access Discussions

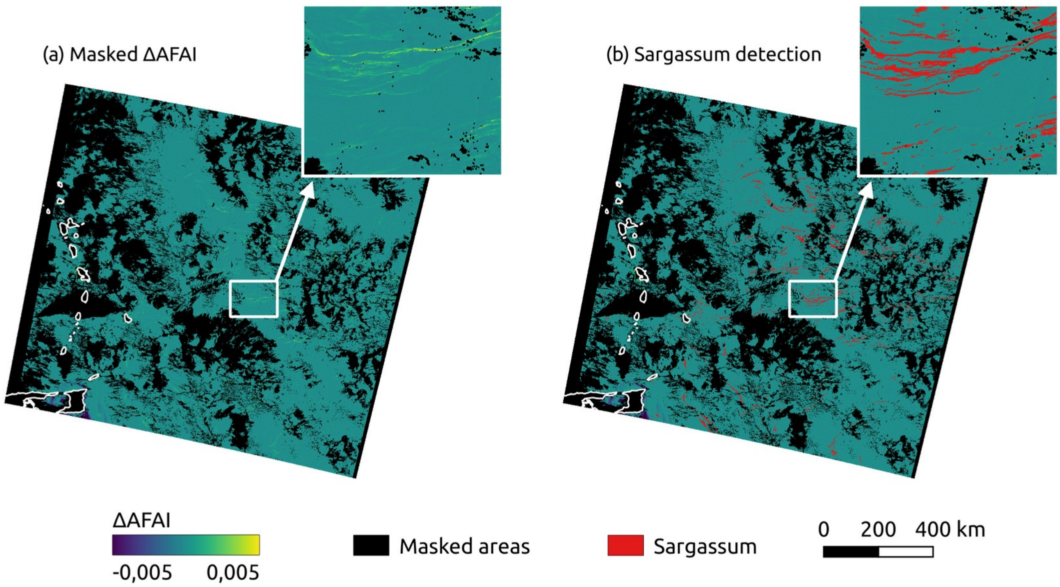

Figure 6: *Sargassum* detection from 01/03/2026. (a) Masked ΔAFAI; (b) *Sargassum* detection based on the comparison of ΔAFAI and ΔMCI to a threshold. Background: ΔAFAI; *Sargassum* detections in red, masked areas in black, shoreline in white.

2.7 Daily composite products

Daily composite maps were generated by aggregating detections from all available granules over the study area, covering the tropical Atlantic Ocean, Caribbean Sea and Gulf of Mexico. Data were reprojected onto a regular latitude-longitude grid with a spatial resolution of 0.0032°. In cases of overlapping observation (e.g. overlap between two sensor images or reprojection effects), pixel values were averaged.

The final product includes both binary detection maps (i.e. *Sargassum* detection mask) and associated ΔAFAI values. This choice is motivated by the possibility of retrieving biomass estimates such as the fractional coverage from this spectral index (Wang et al., 2018).

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Open Access Earth System Science Data Discussions

3 Results

3.1 Dataset overview

The dataset contains the detection of *Sargassum* from S3-OLCI images, over the tropical Atlantic Ocean, the Caribbean Sea and the Gulf of Mexico. Its characteristics are summarized in Table 1.

Table 1: Characteristics of the dataset

Satellite	Sentinel-3A and 3B
Sensor	OLCI
Temporal resolution	Daily
Spatial resolution	0.0032°
Area of interest	Latitude: [0° N, 30° N] Longitude: [10° W, 100° W] Pacific water excluded
Time coverage	08/05/2018 - Today
Coordinate Reference System	WGS84 (EPSG 4326)
Convention	Climate and Forecast (CF) (version 1.8) Attribute Convention for Data Discovery (ACDD) (version 1.3)
Format	Network Common Data Form (NetCDF)

Ultimately, the dataset will cover the period from May 8, 2018 to the present day, with a daily update. The detections of day D are available for download on day D+1 around 10 a.m. UTC. At the time of publication of this article, the archive for previous years is currently being compiled and only detections from 2022 to the present day are available. Once processing for each additional year is complete, the detections for that year will be added online. The complete archive should be reprocessed by the end of 2026.

One year's worth of data for the entire study area is approximately 5 GB and consists of 365 (or 366) records (one per day), each averaging 11 MB.

The dataset is freely accessible on the Odatis platform: https://www.odatis-ocean.fr/en/data-and-services/data-access/direct-access-to-the-data-catalogue#/metadata/1eb82d09-77ed-4f63-9f03-2c3516a9713d. It is indexed with the following DOI: 10.12770/1eb82d09-77ed-4f63-9f03-2c3516a9713d. No account creation is required to download the dataset. The download

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

is possible via a ftp or a https server. The products are in NetCDF format. The metadata of the files comply with the Climate and Forecast (CF) convention (version 1.8) and with the Attribute Convention for Data Discovery (ACDD) (version 1.3). The name of a file complies with the following naming convention:

YYYYMMDD120000-MF-L3S-Sargassum-AFAI-OLCI-fv01.0.nc

with :

- YYYYMMDD: Year, Month and Day
- 20000: Hour, Minute, Second; fixed to 120000 as the product is a composite containing all the granules for a day
- MF: Météo-France
- L3S: Indicator of Level-3 composite product
- Sargassum-AFAI-OLCI: Object studied, Index and Sensor
- fv01.0: version number
- nc: netCDF format

One product is composed of two variables:

- Status_of_detections: this variable can take four possible values:
    - 0: Detection performed, *Sargassum* detected
    - 1: No satellite data
    - 2: No detection performed (masked area)
    - 3: Detection performed, no *Sargassum* detected
- AFAI_deviation_sargassum: Deviation of the AFAI (ΔAFAI) only for pixels including *Sargassum* (see Sect. 2.5 for the definition of the deviation)

An example of a file from the dataset is displayed in Fig. 7.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Open Access Earth System Science Data Discussions

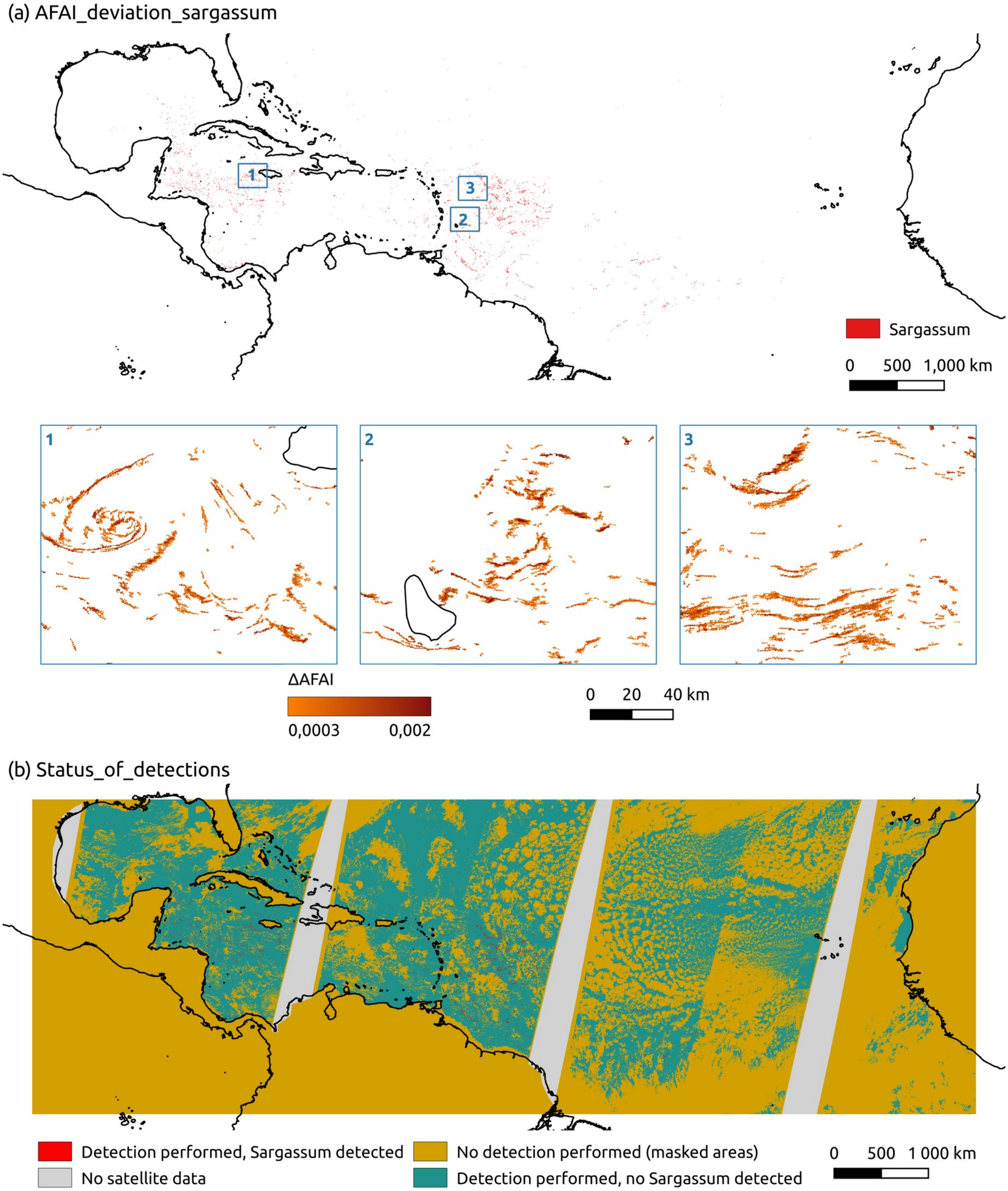

Figure 7: Example of the product from 01/03/2026. (a) Variable "AFAI_deviation_sargassum" corresponding to ΔAFAI for *Sargassum* pixels over the tropical Atlantic Ocean, the Caribbean Sea and the Gulf of Mexico, with three zoomed-in areas; (b) Variable "Status_of_detections" for the same date and studied area.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

3.2 Visual assessment

The performance of the algorithm was assessed through systematic visual inspection across the entire study area for more than fifty dates, including both individual granules (approximately thirty per date) and daily composite products.

In the open ocean, without particular observing conditions (e.g. clouds), *Sargassum* aggregations are consistently detected across a range of ΔAFAI values, suggesting sensitivity to aggregations of varying sizes (Fig. 8c).

The main contribution of the proposed algorithm here lies in its ability to detect *Sargassum* in areas typically considered challenging and often excluded by conventional masking and detection approaches. Three types of areas were considered here: sunglint-affected areas, coastal zones, and cloud-edge areas (Fig. 8).

As visible in Fig. 8d, detection can be performed in the eastern part of the image. The use of adaptive thresholds enables the detection of valid observations in the sunglint areas while maintaining a low level of visually identifiable false detections compared to the rest of the granule.

In coastal areas, detections remain consistent and spatially coherent (Fig. 8a). While some local false detections, due to adjacency effects, may occur, they remain limited compared to the overall gain in observable areas. This improvement is primarily allowed by the complementary advantages of the two spectral indices in coastal areas (see Sect. 2.6).

At cloud edges, some isolated false detections persist (Fig. 8b). However, these are spatially limited and represent a compromise for the increased spatial coverage achieved through reduced masking. As the main goal of these detections are for operational purpose and drift modeling, the compromise should favor the increase in *Sargassum* aggregations detections. This compromise is made possible by the combination of adaptive thresholds and the dual index approach.

No systematic dependency of detection performance on acquisition geometry or geographic location (i.e. between granules) was observed.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Open Access Earth System Science Data Discussions

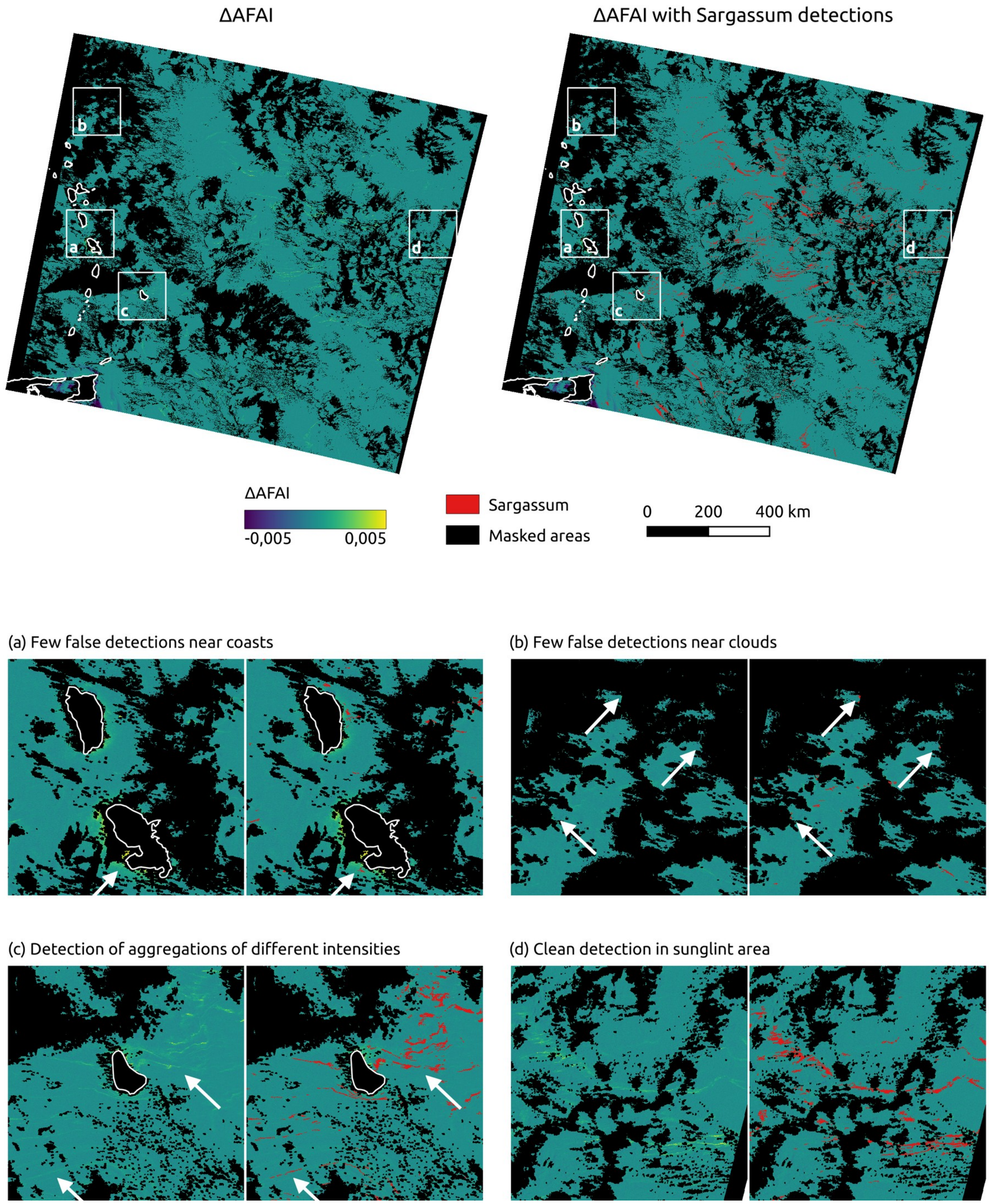

Figure 8: Visual assessment of the accuracy of the detections. Detections from 01/03/2026. Background: ΔAFAI; *Sargassum* detections in red; masked areas in black; shoreline in white.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

3.3 Comparison with existing *Sargassum* detection products

In the absence of extensive in situ validation, the performance and accuracy was further assessed through direct comparison with two existing *Sargassum* detection products: the one produced by the University of South Florida, which can be considered as a reference on the subject, and whose masked AFAI .png images are available on the SaWS platform (Hu et al., 2016), and the SAMTool product developed by CLS (CLS, 2026), downloadable in NetCDF format. These products are widely used and provide a relevant benchmark for comparison.

As visible on Fig. 9, the dataset presented in this study systematically exhibits a larger observable area. In particular, sunglint-affected regions in the eastern part of granules remain largely unmasked, and masking in coastal areas and cloud edges is reduced. These differences are particularly noticeable under high-reflectance conditions, where conventional approaches may mask large portions of the scene (Fig. 9b).

Despite the increased spatial coverage, no substantial increase in visually identifiable false detections is observed, suggesting that the proposed approach improves the balance between detection sensitivity and spatial completeness.

These results are confirmed by statistical analysis performed with the SAMTool dataset as the detections from the SaWS platform are not in open access (Fig. 10). For 100 dates selected at random in 2025, for the entire study area, the detections and masked area from the dataset presented here were reprojected onto the same grid as the SAMTool dataset. The number of pixels detected as *Sargassum* and the number of masked pixels were compared between the two datasets. Of the 100 dates, 17 show a higher number of detections by SAMTool. Upon observation, all these dates have evident false detections around Florida, at the edges of clouds, or off the African coast, which in the latter case can be explained by Saharan dust (Fig. 11). For the remaining 83 dates, the number of detected *Sargassum* is higher in the dataset developed here, up to more than 6 times for the most extreme case. In 50% of cases, the number of *Sargassum* pixels detected in this dataset is at least three times higher than that in the SAMTool dataset. Regarding masked pixels, in 100% of cases, their number is higher in the SAMTool dataset. The highest values correspond to complete masking of certain granules, likely due to a processing error.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

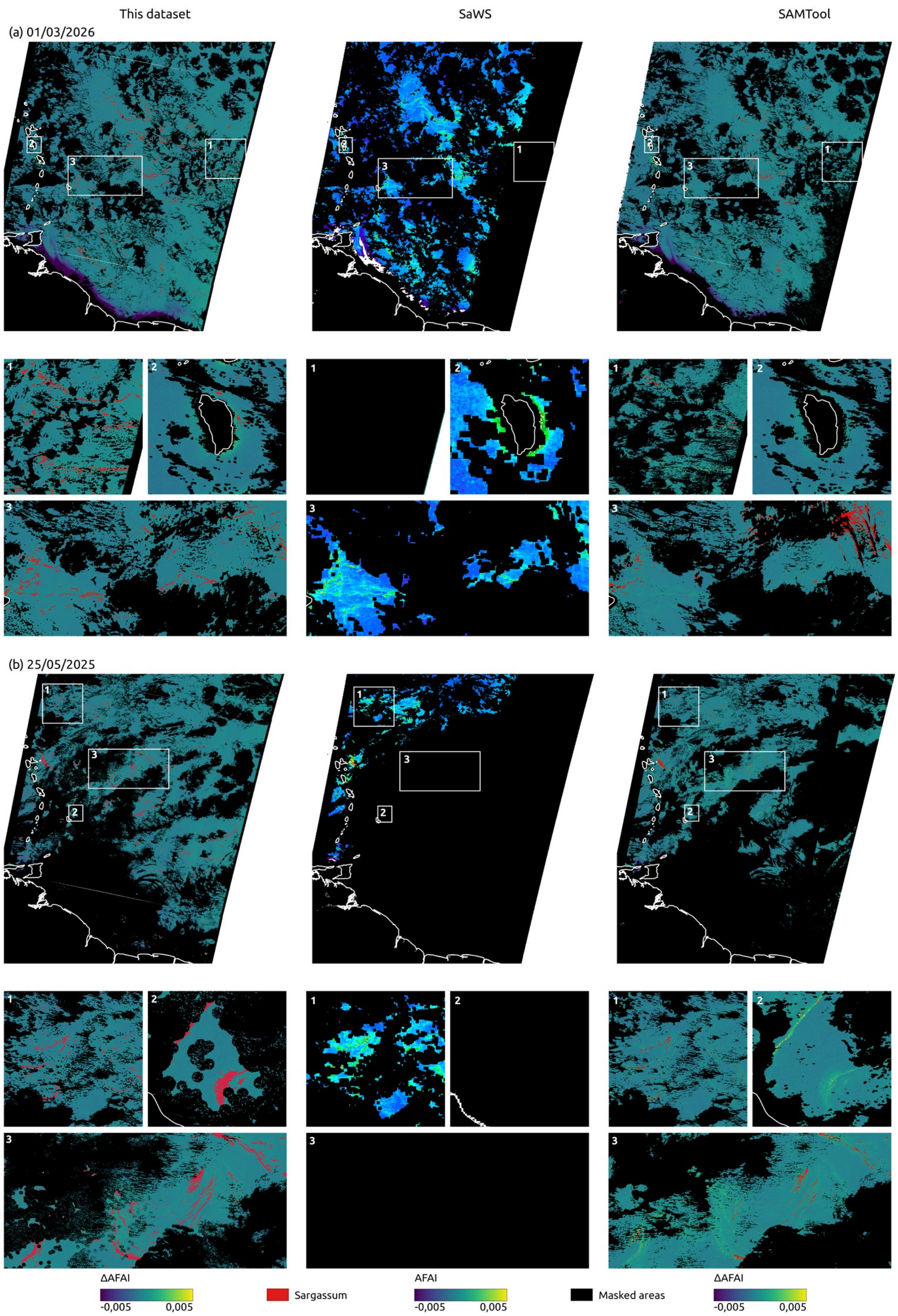

Figure 9: Comparison of the dataset to other Sargassum detection products. (a) Detection on 01/03/2026; (b) Detection on 25/05/2025. Left: This dataset; Middle: SaWS; Right: SAMTool. Detections are in red, masked areas in black, shoreline in white. AFAI in background. The SaWS product displays more masked areas. The SAMtool product also has more masked areas; some false detections are visible near clouds (a.3): some Sargassum are missed (b.2, b.3).

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

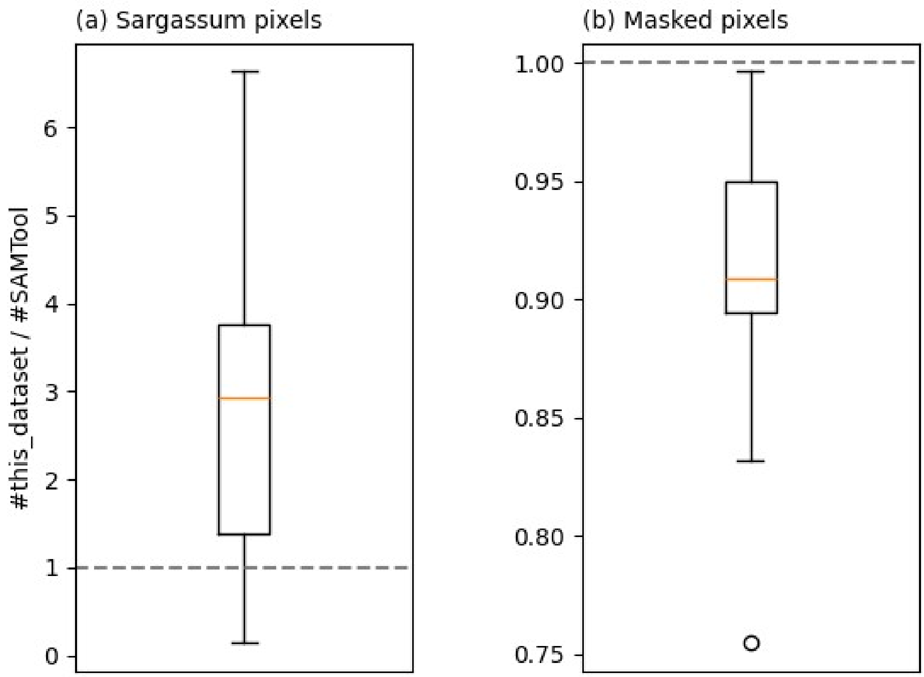

Figure 10: Statistical comparison of the dataset developed here and the SAMTool dataset. (a) Distribution of the ratio number of *Sargassum* detected pixels in this dataset divided by the number of *Sargassum* detected pixels in the SAMTool dataset; (b) Same with the number of masked pixels. A value higher than 1 indicates a higher number of detected/masked pixels in this dataset. Statistics are based on 100 randomly chosen dates in 2025, over the whole study area. The box extends from the first quartile to the third quartile of the data, with a line at the median. The whiskers extend from the box to the farthest data point lying within 1.5 times the inter-quartile range from the box. Beyond that, data are represented by circles.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

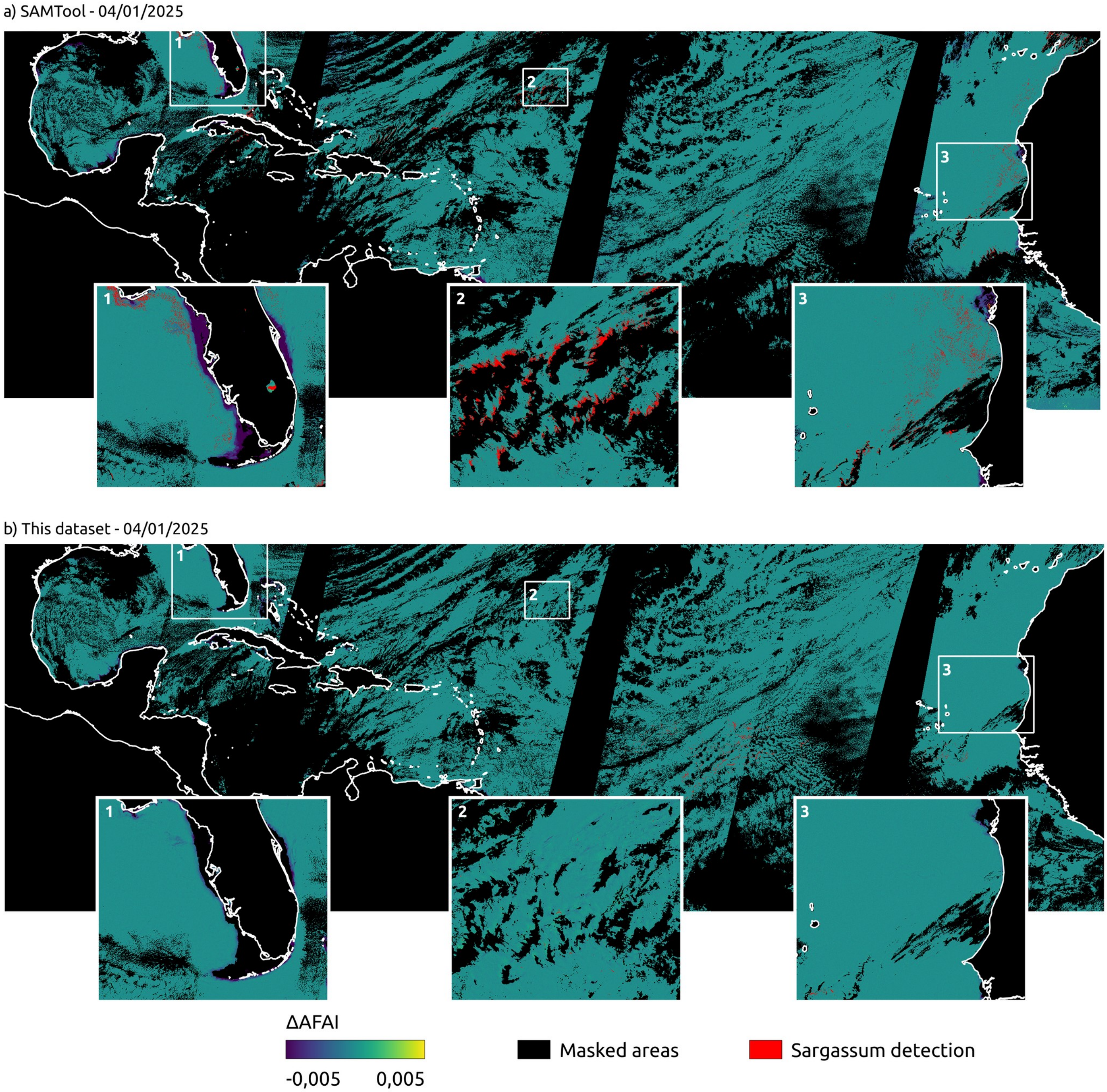

Figure 11: Example of a case where the number of detections by SAMTool is higher than that in this dataset. (a) SAMTool detections on 04/01/2025; (b) Detections from this dataset on 04/01/2025; Background: ΔAFAI. Detections are in red, masked areas in black, shoreline in white.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

4. Discussion

Satellite-based detection of *Sargassum* is subject to several well-known limitations common to floating algae detection algorithms (Debue et al., 2025b). These limitations arise from both sensor characteristics (e.g. spatial (Ody et al., 2019), spectral (Hu et al., 2015) and temporal resolutions (Ody et al., 2019; Hu, 2009)) and environmental or observing conditions (e.g. clouds (Ody et al., 2019; Minghelli et al., 2021; Wang and Hu, 2016), sunglint (Wang and Hu, 2016), size of the aggregations (Gower and King, 2011)).

Due to the spatial resolution of S3-OLCI at 300 m, aggregations below a theoretical threshold estimated at 250 m^2 in clear conditions cannot be detected (Qi and Hu, 2021). In addition, the multispectral sensor S3-OLCI offers a limited number of spectral bands limiting the discrimination of *Sargassum* from other floating materials, including other algae with similar spectral response. However, thanks to the S3 satellites' paired configuration, S3-OLCI data offer good temporal resolution, enabling near-daily monitoring of the tropical Atlantic.

To overcome the limitations inherent in a single sensor, composite products combining observations from multiple sensors are created to leverage the strengths of each sensor (Dierssen et al., 2023; Sun et al., 2024). Such products notably increase spatial coverage and thus improve the accuracy of the observations. However, the quality of a composite product depends on the quality of the observations from each sensor, so it remains essential to continue improving the detection algorithms for each sensor individually. Furthermore, composite products, which involve spatially and temporally resampling the data as well as intra-sensor calibration, are not always compatible with certain studies (such as those examining the relationship between *Sargassum* and temperature), which continue to require single-sensor products.

As mentioned earlier, *Sargassum* detection has significant limitations under certain observation conditions. Cloud cover is a major limitation, as detection in the visible and the NIR is not possible beneath clouds. This limitation can only be partially mitigated through temporal composite or the use of geostationary satellites (Minghelli et al., 2021; Schaeffer et al., 2023). Areas of sunglint or edges of clouds are challenging areas, where increased reflectance can lead to false detections. Near the coast, the signal can also be contaminated due to the proximity of land or shallow waters, which can also generate false detections (Gower and King, 2011). These areas are therefore generally masked, preventing detection in areas that may be quite large and likely to contain *Sargassum* (Hu et al., 2023). The detection algorithm developed here limits masked areas using two methods: adaptive thresholds and the combination of two spectral indices. Masking clouds using a threshold based on longitude accounts for the reflectance gradient in the images between east and west, thereby avoiding masking the eastern portion of the images affected by sunglint, thus increasing the coverage of detections. Defining a threshold based on the pixel column number makes it possible to account for the discontinuities between the OLCI cameras and thus limit false detections, which are caused by excessively low background values, themselves due to pixels with weak signals in areas of higher reflectance. The combination of two spectral indices, meanwhile, enables detections in nearshore areas thanks to the

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

difference in sensitivity of the two indices to observation conditions. Finally, combining the two methods, using a stricter threshold and calculating both indices at the edges of clouds, makes detections possible in these high-reflectance areas.

The empirical definition of these thresholds could be seen as a weakness of the method. However, they were defined through trial and error using a substantial number of images, distributed both spatially (across the entire tropical Atlantic) and temporally (over several years and several months), in order to account for, and thus be robust to, the diversity of observational conditions that may be encountered. These thresholds were then validated through a meticulous visual inspection of the detections generated daily across the entire Atlantic, confirming their robustness within the specific context of this method.

Finally, a common limitation in *Sargassum* remote sensing concerns the final step of validating detections. The most robust validation would be a comparison with ground-truth data, which is scarce due to the remote locations and the vast areas that would need to be covered to acquire such data (Wang and Hu, 2018; Wang and Hu, 2021). Several alternatives are therefore used, including visual inspection of the detections (Qi et al., 2020; Wang and Hu, 2016) or comparison with detections made using a sensor with higher spatial or spectral resolution (Descloitres et al., 2021; Wang and Hu, 2021). However, these comparisons require open access to other datasets and also have limitations, as the other datasets do not represent an absolute truth but also contain detection errors; furthermore, when different sensors are used, there may be a time lag that makes comparisons difficult due to the dynamics of the *Sargassum* (Wang and Hu, 2018; Wang and Hu, 2021).

Despite these various limitations, we believe that this dataset, through its open access and the reduction in masked areas improving detection, could be beneficial to the community working on *Sargassum*. Multiple uses can be envisaged, such as studying the ecological niche of *Sargassum*, training Artificial Intelligence (AI) detection models using the physical approach on which this dataset is based, or feeding drift or growth models for operational or research purposes. Thanks to its accessibility, it would allow all stakeholders who need detection data for their operational or research objectives to independently access such data. Given its spatial and temporal coverage, it would also help deepen our understanding of the phenomenon and its evolution.

5. Conclusion

This dataset contains *Sargassum* detections derived from S3-OLCI imagery across the entire tropical Atlantic since 2018. It was generated using a physical approach based on two spectral indices: the AFAI and the MCI. This dual-index approach, combined with the use of adaptive thresholds based on observation conditions, helps minimize the area of masked regions, thereby allowing detection algorithms to be applied to a larger portion of the study area. This increases the number of

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Sargassum detections, particularly near coasts, at the edges of clouds and in sunglint areas, without increasing the number of false detections.

Since locating *Sargassum* is a preliminary yet essential step, both for all research on this topic, ranging from the origins of the phenomenon to its intra- and inter-annual variability and future evolution, and for operational systems for predicting strandings, making this dataset freely available aims to enable the broadest possible community to continue deepening its understanding of the phenomenon.

Code and data availability

The dataset is freely accessible on the Odatis platform: https://www.odatis-ocean.fr/en/data-and-services/data-access/direct-access-to-the-data-catalogue#/metadata/1eb82d09-77ed-4f63-9f03-2c3516a9713d (Météo-France, 2026). It is indexed with the following DOI: 10.12770/1eb82d09-77ed-4f63-9f03-2c3516a9713d. The code is available upon request.

Author contributions

MD: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Supervision, Project Administration; TG: Conceptualization, Investigation, Data Curation, Writing – Review & Editing, Supervision, Project Administration; AT: Conceptualization, Software, Writing – Review & Editing; SSP: Conceptualization, Writing – Review & Editing, Supervision, Project Administration.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Disclaimer

Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Acknowledgements

The authors thank Frédéric Bacheviller, Olivier Peron and Gilles Montesquieu from Météo-France for their work to operationalize the processing algorithms The authors also thank Jean-François Piollé and Dominique Briand from Institut français de recherche pour l'exploitation de la mer (Ifremer) for their help in sharing the dataset on the Odatis platform. Finally, the authors would like to thank Eumetsat for providing access to the Sentinel-3 data and the Odatis platform for hosting the dataset.

References

CLS: SAMtool: Sargassum Detection, https://datastore.groupcls.com/products/samtool-sargassum-detection/, 2026.

Debue, M., Guinaldo, T., Ayache, D., Bacheviller, F., Barbier, S., Berthomier, L., Blot, E., Conseil, S., Daniel, P., Daniel, W., Durand, E., Gibier, F., Guevel, S., Guibert, F., Hernandez, J.-B., Jaron, O., Manhaval, A., Mauss, A., Montesquieu, G., Nicolas, C., Palany, P. Peron, O., Pfaffenzeller, I., Pradel, B., Saux-Picart, S., Tarasenko, A., Thomé, B. and Tournier, T.: Prévision des échouements de sargasses aux Antilles françaises, La Météorologie, 131, 53-60, https://doi.org/10.37053/lameteorologie-2025-0085, 2025a.

Debue, M., Guinaldo, T., Jouanno, J., Chami, M., Barbier, S., Berline, L., Chevalier, C., Daniel, P., Daniel, W., Descloitres, J., Gros-Desormeaux, J.-R., Lett, C. and Minghelli, A.: Understanding the Sargassum phenomenon in the Tropical Atlantic Ocean: From satellite monitoring to stranding forecast, Mar. Pollut. Bull., 216, p. 117923, https://doi.org/10.1016/j.marpolbul.2025.117923, 2025b.

Descloitres, J., Minghelli, A., Steinmetz, F., Chevalier, C., Chami, M. and Berline, L.: Revisited estimation of moderate resolution Sargassum fractional coverage using decametric satellite data (S2-MSI), Remote Sens., 13 (24), 5106, https://doi.org/10.3390/rs13245106, 2021.

Dierssen, H.M., Gierach, M., Guild, L.S., Mannino, A., Salisbury, J., Schollaert Uz, S., Scott, J., Townsend, P.A., Turpie, K., Tzortziou, M., Urquhart, E., Vandermeulen, R. and Werdell, P.J.: Synergies Between NASA's Hyperspectral Aquatic Missions PACE, GLIMR, and SBG: Opportunities for New Science and Applications, J. Geophys. Res. Biogeosciences, 128.10, e2023JG007574, https://doi.org/10.1029/2023JG007574, 2023.

ESA: SNAP, https://earth.esa.int/eogateway/tools/snap, 2026.

Eumetsat: Data Store. Eumetsat User Portal, https://data.eumetsat.int/search, 2026.

Gower, J.F.R. and King, S.A.: Distribution of floating Sargassum in the Gulf of Mexico and the Atlantic Ocean mapped using MERIS, Int. J. Remote Sens., 32 (7), 1917–1929, https://doi.org/10.1080/01431161003639660, 2011.

Gower, J.F.R., Hu, C., Borstad, G. and King, S.A.: Ocean color satellites show extensive lines of floating Sargassum in the Gulf of Mexico, IEEE Trans. Geosci. Remote Sens., 44 (12), 3619–3625, https://doi.org/10.1109/TGRS.2006.882258, 2006.

Gower, J.F.R., Young, E. and King, S.A.: Satellite images suggest a new Sargassum source region in 2011, Remote Sens. Lett., 4.8, pp. 764–773, https://doi.org/10.1080/2150704X.2013.796433, 2013.

Hu, C.: A novel ocean color index to detect floating algae in the global oceans, Remote Sens. Environ., 113.10, 2118–2129, https://doi.org/10.1016/j.rse.2009.05.012, 2009.

Hu, C., Feng, L., Hardy, R.F. and Hochberg, E.J.: Spectral and spatial requirements of remote measurements of pelagic Sargassum macroalgae, Remote Sens. Environ., 167, 229–246, https://doi.org/10.1016/j.rse.2015.05.022, 2015.

Hu, C., Murch, B., Barnes, B.B., Wang, M., Maréchal, J.-P., Franks, J., Johnson, D., Lapointe, B., Goodwin, D., Schell, J. and Siuda, A.: Sargassum watch warns of incoming seaweed, Eos, 97 (22), p. 10–15, 2016.

Hu, C., Zhang, S., Barnes, B.B., Xie, Y., Wang, M., Cannizzaro, J.P. and English, D.C.: Mapping and quantifying pelagic Sargassum in the Atlantic Ocean using multi-band medium-resolution satellite data and deep learning, Remote Sens. Environ., 289, p. 113515, https://doi.org/10.1016/j.rse.2023.113515, 2023.

Hu, C., Brian, B. B., Qi, L., Gower, J. F. R., Jiao, J. and Xie, Y.: Monitoring pelagic Sargassum in the Atlantic Ocean from space: Principles and practices, Harmful Algae, 144, p. 102840, https://doi.org/10.1016/j.hal.2025.102840, 2025.

Jouanno, J., Benshila, R., Berline, L., Soulié, A., Radenac, M.-H., Morvan, G., Diaz, F., Sheinbaum, J., Chevalier, C., Thibaut, T., Changeux, T., Menard, F., Berthet, S., Aumont, O., Ethé, C., Nabat, P. and Mallet, M.: A NEMO-based model of Sargassum distribution in the tropical Atlantic: description of the model and sensitivity analysis (NEMO-Sarg1.0), Geosci. Model Dev., 14.6, pp. 4069–4086, https://doi.org/10.5194/gmd-14-4069-2021, 2021.

Jouanno, J., Berthet, S., Muller-Karger, F., Aumont, O. and Sheinbaum, J.: An extreme North Atlantic Oscillation event drove the pelagic Sargassum tipping point, Commun. Earth Environ., 6.1, p. 95, https://doi.org/10.1038/s43247-025-02074-x, 2025.

Lam, S.K., Pitrou, A. and Seibert, S.: Numba: a LLVM-based Python JIT compiler, Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, SC15: The International Conference for High Performance Computing, Networking, Storage and Analysis, Austin Texas: ACM, pp. 1–6, https://doi.org/10.1145/2833157.2833162, 2015.

Lamquin, N., Clerc, S., Bourg, L. and Donlon, C.: OLCI A/B Tandem Phase Analysis, Part 1: Level 1 Homogenisation and Harmonisation, Remote Sens., 12.11, p. 1804, https://doi.org/10.3390/rs12111804, 2020.

Météo-France: Daily composite of Sargassum detection derived from OLCI observations, Ver. 1, Obtained from Ifremer/CERSAT, Plouzane, France, https://doi.org/10.12770/1eb82d09-77ed-4f63-9f03-2c3516a9713d, 2026.

Minghelli, A., Chevalier, C., Descloitres, J., Berline, L., Blanc, P. and Chami, M.: Synergy between Low Earth Orbit (LEO)—MODIS and Geostationary Earth Orbit (GEO)—GOES Sensors for Sargassum Monitoring in the Atlantic Ocean, Remote Sens., 13.8, p. 1444, https://doi.org/10.3390/rs13081444, 2021.

Ody, A., Thibaut, T., Berline, L., Changeux, T., André, J.-M., Chevalier, C., Blanfuné, A., Blanchot, J., Ruitton, S., Stiger-Pouvreau, V., Connan, S., Grelet, J., Aurelle, D., Guéné, M., Bataille, H., Bachelier, C., Guillemain, D., Schmidt, N.,

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Fauvelle, V., Guasco, S. and Ménard, F.: From In Situ to satellite observations of pelagic Sargassum distribution and aggregation in the Tropical North Atlantic Ocean, PLoS ONE, 14.9, p. e0222584, https://doi.org/10.1371/journal.pone.0222584, 2019.

Podlejski, W., Descloitres, J., Chevalier, C., Minghelli, A., Lett, C. and Berline, L.: Filtering out false Sargassum detections using context features, Front. Mar. Sci., 9, 960939, https://doi.org/10.3389/fmars.2022.960939, 2022.

Podlejski, W., Berline, L., Nerini, D., Doglioli, A. and Lett, C.: A new Sargassum drift model derived from features tracking in MODIS images, Mar. Pollut. Bull., 188, p. 114629, https://doi.org/10.1016/j.marpolbul.2023.114629, 2023.

Putman, N. F., Beyea, R. T., Ackerman, E. G., Trinanes, J., Le Hénaff, M., Hu, C. and Lumpkin, R.: Systems to monitor and forecast pelagic Sargassum inundation of coastal areas across the North Atlantic: present tools and future needs, Harmful Algae, 149, p. 102933, https://doi.org/10.1016/j.hal.2025.102933, 2025.

Qi, L. and Hu, C.: To what extent can Ulva and Sargassum be detected and separated in satellite imagery?, Harmful Algae, 103, 102001, https://doi.org/10.1016/j.hal.2021.102001, 2021.

Qi, L., Hu, C., Mikelsons, K., Wang, M., Lance, V., Sun, S., Barnes, B.B., Zhao, J. and Van der Zande, D.: In search of floating algae and other organisms in global oceans and lakes. Remote Sens. Environ., 239, 111659. https://doi.org/10.1016/j.rse.2020.111659, 2020.

Resiere, D., Florentin, J. and Névière, R.: Evaluating the ten-year health impact of hydrogen sulfide (H2S) and ammonia (NH3) exposure from sargassum seaweed invasions in the Caribbean: Public health implications, Harmful Algae, 152, p. 103027. https://doi.org/10.1016/j.hal.2025.103027, 2026.

Rodríguez-Martínez, R.E., Torres-Conde, E.G., Rosellón-Druker, J., Cabanillas-Terán, N. and Jáuregui-Haza, U.: The Great Atlantic Sargassum Belt: Impacts on the Central and Western Caribbean – A review, Harmful Algae, 144, 102838, https://doi.org/10.1016/j.hal.2025.102838, 2025.

Schaeffer, B.A., Whitman, P., Vandermeulen, R., Hu, C., Mannino, A., Salisbury, J., Efremova, B., Conmy, R., Coffer, M., Salls, W., Ferriby, H. and Reynolds, N.: Assessing potential of the Geostationary Littoral Imaging and Monitoring Radiometer (GLIMR) for water quality monitoring across the coastal United States, Mar. Pollut. Bull., 196, p. 115558, https://doi.org/10.1016/j.marpolbul.2023.115558, 2023.

Sun, Y., Wang, M., Liu, M., Li, Z.B., Chen, Z. and Huang, B.: Continuous Sargassum monitoring across the Caribbean Sea and Central Atlantic using multi-sensor satellite observations, Remote Sens. Environ. 309, 114223, https://doi.org/10.1016/j.rse.2024.114223, 2024.

Wang, M. and Hu, C.: 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, 2016.

Wang, M. and Hu, C.: On the continuity of quantifying floating algae of the Central West Atlantic between MODIS and VIIRS. Int. J. Remote Sens., 39.12, 3852–3869. https://doi.org/10.1080/01431161.2018.1447161, 2018.

https://doi.org/10.5194/essd-2026-589
Preprint. Discussion started: 24 August 2026

Wang, M. and Hu, C.: 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, 2021.

Wang, M., Hu, C., Barnes, B.B., Mitchum, G., Lapointe, B. and Montoya, J.P.: The great Atlantic Sargassum belt, Science, 365.6448, pp. 83–87, https://doi.org/10.1126/science.aaw7912, 2019.

Wang, M., Hu, C., Cannizzaro, J., English, D., Han, X., Naar, D., Lapointe, B., Brewton, R. and Hernandez, F.: Remote Sensing of Sargassum Biomass, Nutrients, and Pigments, Geophys. Res. Lett., 45.22, https://doi.org/10.1029/2018GL078858, 2018.