

Dear Editor,

Dear referees,

We thank you for your minor and technical comments which helped us to improve the quality of the paper.

Please find below our answers to your remarks (in blue). The proposed changes in the text are marked in red.

### Reply to Reviewer 1

Fig XX The schematic of the adopted methodology.”

To clarify, I suggest using a generic ‘circulation model’ instead of repeating MERCATOR and HYCOM in each panel

Our answer:

Following your suggestion, the schematic of the adopted methodology was modified as below:

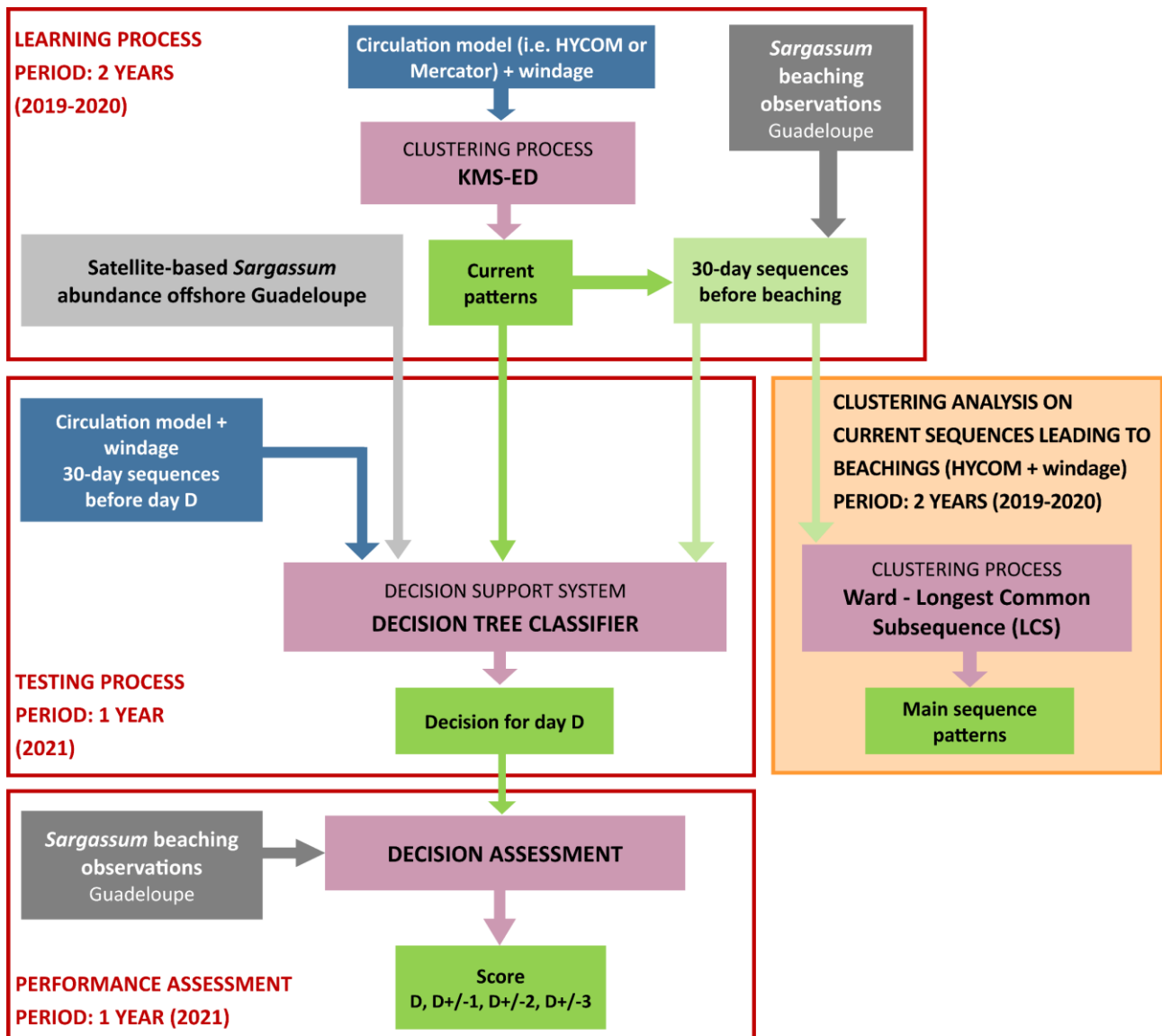


Figure 2: A schematic showing the overall methodology.

“Figure XX: The schematic of the Expert Distance process.”

Beware of typos :

distribution.

KL ‘divergence’ rather than convergence I guess.

Our answer:

Following your remark, this schematic was modified: “convergence” was replaced by “divergence”.

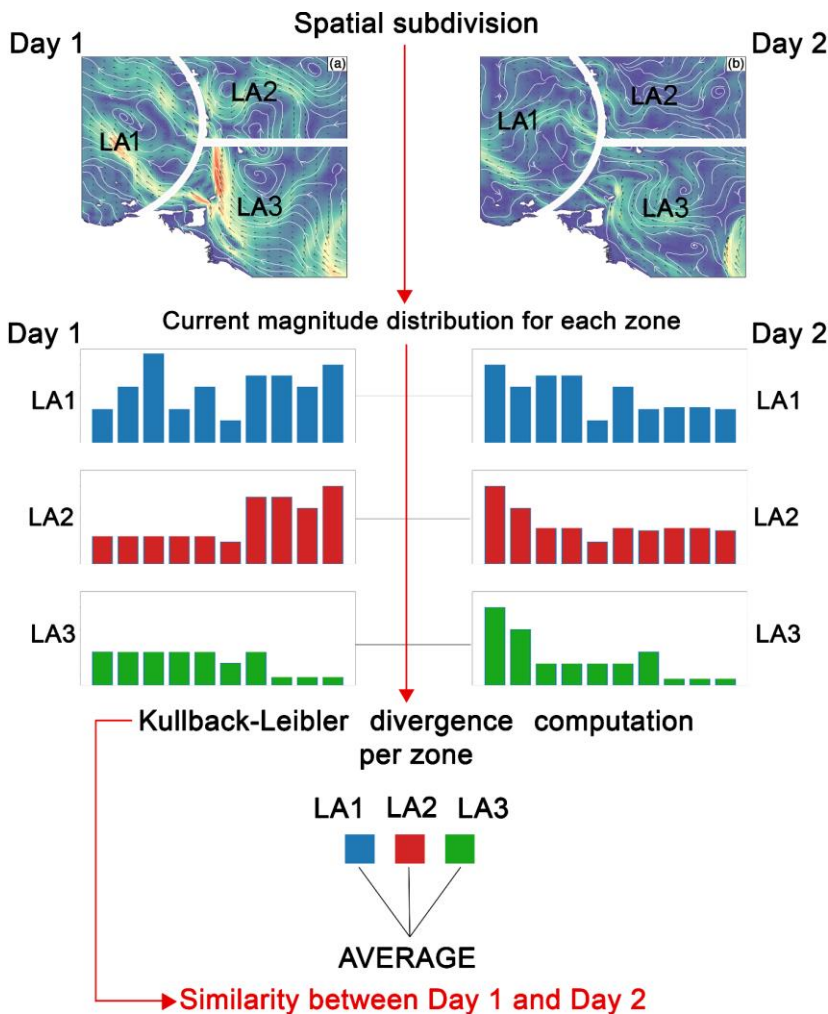


Figure 3: The schematic of the Expert Distance process.

Fig 11

As the monthly distribution of cluster MC1 and MC2 differ from HC1 and HC2, you should mention that the cluster number differ

Our answer:

To clarify this point, the following sentence was added L284:

*“The cluster numbering does not take into account these match percentages (e.g. MC1 and MC2 main patterns respectively differ from HC1 and HC2 patterns).”*

The caption of Fig 11 was also modified as below:

*“Figure 11: Representative elements of the clusters from HYCOM current data combined with ERA-5 windage (KMS-ED method with  $k = 4$ ): HC1 (day 29-04-2019) (a), HC2 (day 06-01-2020) (b), HC3 (day 04-05-2020) (c), HC4 (day 11-11-2019) (d). The HYCOM clusters numbering differs from the Mercator clusters numbering.”*

## Reply to Reviewer 2

” Compared to other conventional forecast...” to “Compared to other conventional forecasts...” ”

“predictive modelling methods requiring low computational costs are characterized by their flexibility” to “predictive modelling methods require low computational costs **and?** are characterized by their flexibility” add “and” or only the ones with low computational cost?

Our answer :

Following your remarks,

L78: ” *Compared to other conventional forecast...* ” was replaced to “*Compared to other conventional forecasts...* ”

”*predictive modelling methods requiring low computational costs are characterized*” was replaced by “*predictive modelling methods require low computational costs and are characterized*”

Section 2.1 suggested modification (in bold the specific changes made):

### 2.1 HYCOM surface current dataset

Daily (**12 UTC, i.e. Coordinated Universal Time**) surface current components from the 41-layer Hybrid Coordinate Ocean Model (HYCOM) **at 1/12-degree, global analysis** (HYCOM GLBy0.08 version), were examined. The HYCOM surface forcing including 10-m wind velocities are extracted from Climate Forecast System Version 2 (CFSv2). The Navy Coupled Ocean Data Assimilation (NCODA) system is used to assimilate available observational data: satellite altimeter sea surface height, satellite and in situ sea surface temperature, temperature vertical profiles and salinity vertical profiles (Cummings, 2005; Cummings and Smedstad, 2013; Helber et al., 2013). **The bathymetry** used is the GEBCO8 (Becker et al., 2009) with 30 arc second of resolution. The HYCOM GLBy0.08 grid resolution is 0.08 degree in longitude and 0.04 degree in latitude. To perform the present study, the native HYCOM fields were **first** interpolated on the Mercator uniform lon/lat 0.08-degree grid with a bilinear method. Putman et al. (2018) and Johns et al. (2020) used a previous version of HYCOM model including uniform lon/lat 0.08° scale grid to successfully simulate *Sargassum* trajectories.

Section 2.2 suggested modification (in bold the specific changes made, and underlined phrases to be revised):

### 2.2 Mercator surface current dataset

The daily (**12 UTC**) surface current components from the 50-layer PSY4V3R1 Mercator 1/12-degree 3D analysis system (Lellouche et al., 2018; Gasparin et al., 2019) were also analyzed. The atmospheric surface forcing **is** extracted from the 3-hourly ECMWF (European Centre for Medium-Range Weather Forecasts) IFS (Integrated Forecast System). Assimilated observational data types are quite similar to HYCOM model. Unlike the HYCOM GLBy0.08 native grid including higher resolution in latitude (i.e. 0.04 degree), the Mercator native grid is uniform in longitude and latitude with 0.08-degree **horizontal grid resolution**. This would suggest that HYCOM may better reproduce small scale patterns than Mercator. Moreover, as described by Lellouche et al. (2018), the Mercator bathymetry includes GEBCO8 data in regions shallower than 200 m and the coarse 1 arc-minute ETOPO1 data (Amante and Eakins, 2009) in regions deeper than 300 m. The complex bathymetry of the Lesser Antilles Arc studied here could be less realistic **in the Mercator than in the HYCOM fields**.

Our answer: following your remarks, these paragraphs were modified as below (specific changes in red)

### 2.1 HYCOM surface current dataset

Daily (**12 UTC, i.e. Coordinated Universal Time**) surface current components from the 41-layer Hybrid Coordinate Ocean Model (HYCOM) **at 1/12-degree, global analysis** (HYCOM GLBy0.08 version, **available at: <https://www.hycom.org/data/glby0pt08/expt-93pt0>, last access: 17 January 2022**), were examined. The HYCOM surface forcing including 10-m wind velocities are extracted from Climate Forecast System Version 2 (CFSv2). The Navy Coupled Ocean Data Assimilation (NCODA) system is used to assimilate available observational data: satellite altimeter sea surface height, satellite and in-situ sea surface temperature, temperature vertical profiles and salinity vertical profiles (Cummings, 2005; Cummings and Smedstad, 2013;

Helber et al., 2013). The **bathymetry** used is the GEBCO8 (Becker et al., 2009) with 30 arc second of resolution. The HYCOM GLBy0.08 grid resolution is 0.08 degree in longitude and 0.04 degree in latitude. To perform the present study, the native HYCOM fields were **first** interpolated on the Mercator uniform lon/lat 0.08-degree grid with a bilinear method. Putman et al. (2018) and Johns et al. (2020) used a previous version of HYCOM model including uniform lon/lat 0.08° scale grid to successfully simulate *Sargassum* trajectories.

## 2.2 Mercator surface current dataset

The daily (12 UTC) surface current components from the 50-layer PSY4V3R1 Mercator 1/12-degree 3D analysis system (Lellouche et al., 2018; Gasparin et al., 2019) were also analyzed (**available at: [https://resources.marine.copernicus.eu/product-detail/GLOBAL\\_ANALYSIS\\_FORECAST\\_PHY\\_001\\_024/DATA-ACCESS](https://resources.marine.copernicus.eu/product-detail/GLOBAL_ANALYSIS_FORECAST_PHY_001_024/DATA-ACCESS)**, last access: 17 January 2022).

The atmospheric surface forcing **is** extracted from the 3-hourly ECMWF (European Centre for Medium-Range Weather Forecasts) IFS (Integrated Forecast System). **This version of Mercator model includes assimilation of observational data quite similarly to HYCOM NCODA system (i.e. satellite altimeter sea surface height, satellite and in situ sea surface temperature, temperature vertical profiles and salinity vertical profiles).** Unlike the HYCOM GLBy0.08 native grid including higher resolution in latitude (i.e. 0.04 degree), the Mercator native grid is uniform in longitude and latitude with **0.08-degree horizontal grid resolution**. This would suggest that HYCOM may better reproduce small scale patterns than Mercator. Moreover, as described by Lellouche et al. (2018), the Mercator bathymetry includes GEBCO8 data in regions shallower than 200 m and the coarse 1 arc-minute ETOPO1 data (Amante and Eakins, 2009) in regions deeper than 300 m. The complex bathymetry of the Lesser Antilles Arc studied here could be less realistic **in the Mercator than in the HYCOM fields**.

“Assimilated observational data types are quite similar to HYCOM model.” This phrase is not very clear, with the assimilated observational data types are you referring to the Mercator dataset? If so, please specify whether the HYCOM and Mercator datasets include data assimilation.

Our answer:

To clarify this point, the sentence “*Assimilated observational data types are quite similar to HYCOM model.*” was replaced by “*This version of Mercator model includes assimilation of observational data quite similarly to HYCOM NCODA system (i.e. satellite altimeter sea surface height, satellite and in situ sea surface temperature, temperature vertical profiles and salinity vertical profiles).*”

The above phrase and section 2.3 suggested to be re-written for clarity in the following way:

“Surface wind influences the transport of floating seaweed rafts and a drag or windage coefficient must be added to the surface currents. The value of  $C_w = 0.01$  was used by Putman et al. (2018), Johns et al. (2020) and Berline et al. (2020). The use of other

windage values should be investigated in a further study. The daily 12 UTC surface wind data (at 1000 hPa) from the 31-km scale ERA-5 model were integrated with Mercator and HYCOM currents following this formula:

$$u_s(x, t) = u_m(x, t) + C_w u_w(x, t) \quad (1)$$

where  $u_s$  represents the oceanic surface currents with windage,  $u_m$  the oceanic surface currents velocity,  $C_w$  the windage and  $u_w$  the surface winds velocity. This approach is consistent with Putman et al. (2018) and Johns et al. (2020) studies.”

To:

“Surface wind influences the transport of floating seaweed rafts and a drag or windage coefficient must be added to the surface currents. Daily (12 UTC) from the **31-km horizontal resolution** ERA-5 model was used. The wind data was integrated with Mercator and HYCOM **ocean currents data** following this formula:

$$u_s(x, t) = u_m(x, t) + C_w u_w(x, t) \quad (1)$$

where  $u_s$  represents the oceanic surface currents with windage,  $u_m$  the oceanic surface currents velocity,  $C_w$  the windage and  $u_w$  the surface winds velocity. This approach is consistent with Putman et al. (2018) and Johns et al. (2020) studies. The value of  $C_w = 0.01$  was used, **following** Putman et al. (2018), Johns et al. (2020) and Berline et al. (2020). The use of other windage values should be investigated in a further study.”

### 2.3 ERA-5 dataset: surface winds

Surface wind influences the transport of floating seaweed rafts and a drag or windage coefficient must be added to the surface currents. *Daily 12 UTC fields from the hourly 31-km horizontal resolution ERA-5 reanalysis dataset (Hersbach et al. 2020) was used (ERA-5 reanalysis, available at: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels>, last access: 17 January 2022).* The wind data was integrated with Mercator and HYCOM ocean currents data following this formula:

$$u_s(x, t) = u_m(x, t) + C_w u_w(x, t) \quad (1)$$

where  $u_s$  represents the oceanic surface currents with windage,  $u_m$  the oceanic surface currents velocity,  $C_w$  the windage and  $u_w$  the surface winds velocity. This approach is consistent with Putman et al. (2018) and Johns et al. (2020) studies. The value of  $C_w = 0.01$  was used, following Putman et al. (2018), Johns et al. (2020) and Berline et al. (2020). The use of other windage values should be investigated in a further study.

Information on the ERA5 data still missing: is it the reanalysis dataset? Any references of the dataset? Is the temporal resolution daily or hourly and just the 12 UTC fields used? Lastly, please add here or in the Data availability section where this data was downloaded or obtained from.

Our answer:

The sentence L135 was modified as below:

*“Daily 12 UTC fields from the hourly 31-km horizontal resolution ERA-5 reanalysis dataset (Hersbach et al. 2020) was used (ERA-5 reanalysis, available at: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels>, last access: 17 January 2022).. “*

These references were added in the references section:

*ERA-5 reanalysis, <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels>, last access: 17 January 2022.*

*Hersbach, H., Bell, B., Berrisford, P., et al. : The ERA5 global reanalysis. Q. J. R. Meteorol. Soc., 146, 1999-2049. <https://doi.org/10.1002/qj.3803>, 2020.*

Also, this information is missing for the HYCOM and Mercator datasets (sections 2.1 and 2.2).

Our answer:

This information was added for the HYCOM dataset L111:

*“...global analysis (HYCOM GLBy0.08 version, available at: <https://www.hycom.org/data/glby0pt08/expt-93pt0>, last access: 17 January 2022),”*

This information was added for the Mercator dataset L122:

*“...were also analyzed (PSY4V3R1 Mercator 1/12-degree 3D analysis, available at: [https://resources.marine.copernicus.eu/product-detail/GLOBAL\\_ANALYSIS\\_FORECAST\\_PHY\\_001\\_024/DATA-ACCESS](https://resources.marine.copernicus.eu/product-detail/GLOBAL_ANALYSIS_FORECAST_PHY_001_024/DATA-ACCESS), last access: 17 January 2022).*

This reference was added in the references section:

*PSY4V3R1 Mercator 1/12-degree 3D analysis, [https://resources.marine.copernicus.eu/product-detail/GLOBAL\\_ANALYSIS\\_FORECAST\\_PHY\\_001\\_024/DATA-ACCESS](https://resources.marine.copernicus.eu/product-detail/GLOBAL_ANALYSIS_FORECAST_PHY_001_024/DATA-ACCESS), last access: 17 January 2022.*

**L128: “Ward's method for HAC” Please explain and add reference.**

The sentence “Besides the measures and the classes of distance between objects such as the Euclidean distance for K-means and the Ward's method for HAC, a new metric was also added (Biabiany et al. 2020)”

was modified like below (L154)

“Besides the measures and the classes of distance between objects such as the Euclidean distance for K-means and the Ward method which allows to identify homogeneous subsets of data (Ward, 1963), a new metric was also added.”

Please clarify phrasing, maybe split sentence in 2 or 3?

Our answer:

Following your suggestion, the sentence was splitted in 2, L159:

*“The Ward method allows to identify homogeneous subsets of data (Ward, 1963). Besides the measures and the classes of distance between objects such as the Euclidean distance for K-means and the Ward method, a new metric was also added.”*

The proposed tree in Fig. 5 was tested on the full year of 2021 except 31 December 2021 including missing data, i.e. in total 364 tests.

Phrase not clear, do you mean the 31st of December was not included because of missing data?

Our answer:

This sentence was replaced L223 by *“The proposed tree in Fig. 5 was tested on the full year of 2021 except 31 December 2021 which was not included because of missing data, giving a total of 364 tests”*

“the lack of study dealing” to “the lack of studies dealing”

Our answer:

Following your remark, *“the lack of study dealing”* was replaced by *“the lack of studies dealing”* (L228)

“90% of them remain” to “90% of the velocity values remain”

Our answer:

Following your remark, *“90% of them remain ”* was replaced by *“90% of the velocity values remain ”* (L230)

Thanks for the clarification, maybe to clarify further modify “A normal kernel was used to obtain these distributions.” to “A **Gaussian** kernel was **applied** to obtain these distributions.”

Our answer:

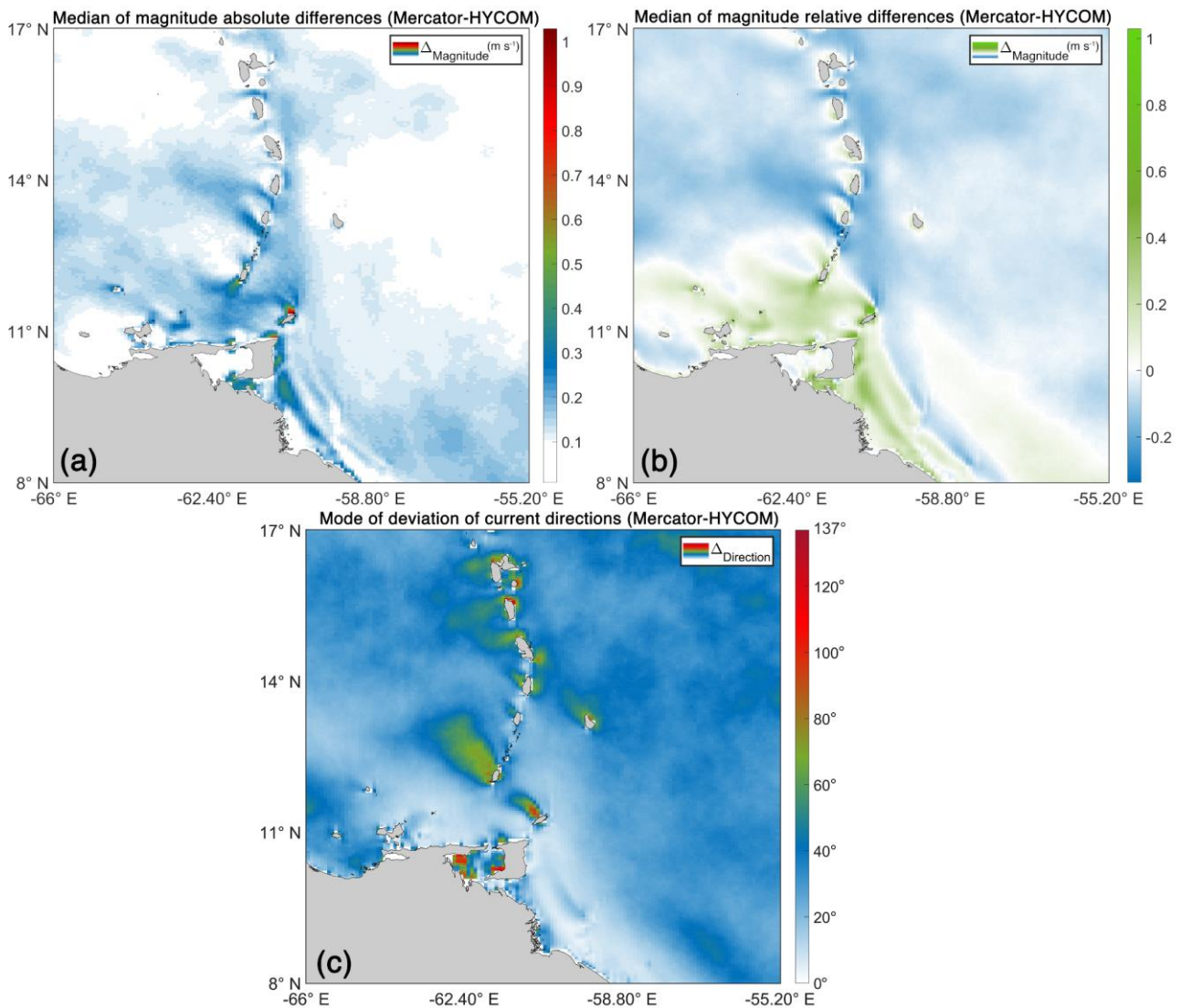
Following your remark, *“A normal kernel was used to obtain these distributions.”* was replaced by *“A Gaussian kernel was applied to obtain these distributions. ”* (L234)

Figure 8: Comparison between Mercator and HYCOM surface currents from 2019 to 2020 on the same 0.08° grid: (a) median of magnitude absolute differences (Mercator-HYCOM) in m s-1 and (b) median of magnitude relative differences (Mercator-HYCOM) in m s-1 and (c) mode of current direction differences (Mercator-HYCOM) in degree.

For subplot c), colormap a bit confusing as the white regions suggest not difference in direction, but those regions are actually different. Maybe the same colormap as subplot a) would be more appropriate?

Our answer:

Following your remark, the colormap of the subplot c (Fig 8) was modified as below.



**Figure 8: Comparison between Mercator and HYCOM surface currents from 2019 to 2020 on the same 0.08° grid: (a) median of magnitude absolute differences (Mercator-HYCOM) in m s<sup>-1</sup> and (b) median of magnitude relative differences (Mercator-HYCOM) in m s<sup>-1</sup> and (c) mode of current direction differences (Mercator-HYCOM) in degree.**

**L272-L273: “The monthly evolution of observed stranding days on the Guadeloupe coasts, the monthly evolution of Sargassum abundance over the Central Atlantic region**

**(SaWS, <https://optics.marine.usf.edu/projects/SaWS.html>)” I imagine it should be: “Guadeloupe coasts and the monthly evolution...”, to make clear you talking about two datasets. The observed stranding dataset is mentioned in the dataset section (section 2.4), but not the Sargassum abundance over the Central Atlantic region.**

Line 272, the sentences:

“The monthly evolution of observed stranding days on the Guadeloupe coasts, the monthly evolution of Sargassum abundance over the Central Atlantic region (SaWS, <https://optics.marine.usf.edu/projects/SaWS.html>) were also analyzed on the focused period 2019-2020 (Figs. 11 and 12). During these two years, the amount of Sargassum over the Central Atlantic region increased significantly from February to July, then decreased from July to November.”

were replaced L304 by:

“The monthly evolution of observed stranding days on the Guadeloupe coasts, the monthly evolution of Sargassum abundance in the area 30-100 km offshore Guadeloupe were also analyzed on the focused period 2019-2020 (Figs. 11 and 12). During these two years, the amount of Sargassum which may enhance the

beaching risk in Guadeloupe increased significantly from February to May, then decreased from May to November.”

I think my remark here was not clarified. Please re-check.

Our answer:

The observed stranding dataset is mentioned in the dataset section (section 2.4), and the Sargassum abundance in the dataset section (section 2.5).

L150

### **“2.5 Satellite-based offshore abundance of Sargassum**

*Sargassum satellite observations were included in the present decision support system. To quantify the abundance of Sargassum in an area of 100 km radius offshore Guadeloupe, the 7-day Floating Algae (FA) density fields derived from the Alternative Floating Algae Index (Wang and Hu, 2016) were analyzed. As described by Trinanes et al. (2021), the 7-day Floating Algae (FA) density fields are accumulated on 7 days and have a 0.1° resolution. Due to optical complexity in nearshore waters, the FA density fields are masked with missing values within 30 km from shoreline (Trinanes et al. 2021). The cumulative FA density values were added up in the area 30-100 km offshore Guadeloupe (Fig. 1) then averaged over the two years 2019 and 2020 for each day.”*

### **3. Technical corrections**

Some *Sargassum* in the Conclusions section forgot to put in italics.

Our answer: Following your suggestion, these “*Sargassum*” were italicized in the conclusion section.

L47-48: “current, and numerical models, both the roles of both subsurface nutrient supply and surface current transport were estimated.” to “current, and numerical models; the roles of both subsurface nutrient supply and surface current transport were estimated.”

Our answer:

Following your remark, “*current, and numerical models, both the roles of both subsurface nutrient supply and surface current transport were estimated.*” was replaced by “*current, and numerical models; the roles of both subsurface nutrient supply and surface current transport were estimated.*” (L47)

“During this period of 730 days, only 110 days of observed beaching were recorded” to “During this period of 730 days, only 110 days of *Sargassum* beaching were recorded”

Our answer:

Following your remark, “*During this period of 730 days, only 110 days of observed beaching were recorded*” was replaced by “*During this period of 730 days, only 110 days of Sargassum beaching were recorded*” (L145)

### **L151: Define all variables of equation 2!**

This sentence was added below the equation (2) :

where  $k$  is the number of clusters,  $C_j$  the set of days from the cluster  $j$ ,  $i$  a day from  $C_j$  and  $s(i)$  the silhouette index (Rousseeuw, 1987) value of day.

“...value of day.” to “..value of day  $i$ ”.

Our answer:

Following your remark, “...value of day.” was replaced by “..value of day  $i$ ”. (L180)



For the maximum velocity of 2.57 m/s for both models, is it exactly the same too? If so, why does in figure 7 the x-axis go up to 2.57 m/s for Mercator, but only 2.49 m/s for HYCOM?

Our answer:

The sentence “For both models HYCOM and Mercator, the maximum surface velocity is 2.57 m s<sup>-1</sup> and 90% of them remain below 0.65 m s<sup>-1</sup>” was replaced by:

*“The maximum surface velocity reaches 2.49 m s<sup>-1</sup> and 2.57 m s<sup>-1</sup>, respectively for HYCOM and Mercator. For both models 90% of the velocity values remain below 0.65 m s<sup>-1</sup> (the respective 90th centile values are respectively 0.6515 m s<sup>-1</sup> and 0.6458 m s<sup>-1</sup> for HYCOM and Mercator).” (L229)*

Table 6: Decision tree performance scores.

Please add details to the table caption. True/ negative positive of... and recall % referring to... Accuracy of... and ratio between... .

Our answer:

The caption of Table 6 was changed to:

*“Table 6: Decision tree performance scores: “True positive/negative” respectively refer to the number of observed beaching/non-beaching days predicted by the decision system; “Recall” refers to the ratio in percentage between these respective numbers of days and the total number of tests (i.e. 364 days); “Accuracy” corresponds to the number of days with a true prediction and its ratio in percentage over the total number of tested days.”*

Moreover, in section 3.5 you first introduce table 6, but further details are still missing. I assume that with true positive you mean beaching and true negative non-beaching days? Please specify this. Moreover, the values in table 6 for true positive is the number of days beaching is observed and the percentage, the percentage of days with beaching observed? The corresponding for non-beaching and true negative? Also, how is the accuracy calculated? And the ratio?

Our answer:

Following your remark, these sentences were added in section 3.5 (L339):

*“True positive/negative” respectively refer to the number of observed beaching/non-beaching days, predicted by the decision system. “Recall” refers to the ratio in percentage between these respective numbers of days and the total number of tests (i.e. 364 days). “Accuracy” corresponds to the number of days with a true prediction and its ratio in percentage was computed over the total number of tested days.”*

#### 4. Figures and tables

Figure 2: “The schematic of the overall methodology.” to “A schematic showing the overall methodology.”

Our answer:

Following your remark, the Figure 2 caption was modified: “The schematic of the overall methodology.” was replaced by “A schematic showing the overall methodology.”

Figure 4: “The schematic of the clustering process on the current sequences leading to beachings.” to “The schematic of the clustering process **used** on the **ocean** current sequences leading to beachings.”

Our answer:

Following your remark, the Figure 4 caption was modified: “The schematic of the clustering process on the current sequences leading to beachings.” was replaced by “The schematic of the clustering process **used** on the **ocean** current sequences leading to beachings.”