Ocean velocity observations are presently not routinely assimilated in most ocean forecasting systems. This is largely due to the availability of NRT surface velocity observations which provided sufficient spatial coverage. This study looks at the impact of assimilating a novel surface velocity product into a multi model ensemble. Ocean surface velocities are derived from AVHRR SSTs using a maximum cross-correlation algorithm. The authors describe the methods used to process the data and outline the multi-model Ensemble Kalman filter used to assimilate the observations. I believe that this study is of interest to the community as it expands on the relatively limited field of ocean velocity assimilation. However, I feel the paper would benefit from significant revisions. In many places the explanations are unclear and in some places, there are contradictions in the text. In addition, many of the results need a clearer benchmark to allow meaningful conclusions to be drawn.

Main general comments:

1. The paper needs to have some context for the HF radar and renalaysis comparisons (in sections 4.2 and 4.3) by providing the statistics for the background multi model ensemble. In some places in the text these comparisons are alluded to (lines 499, 646, 818) but I can't find any assessment of the background ensemble outside of section 4.1

Response: The requested context has now been incorporated by updating Table 3 to include the background ensemble statistics with respect to the reanalysis dataset. This provides a direct reference for evaluating the improvement achieved by the analysis over the background and complements the comparisons presented in Sections 4.2 and 4.3.

In the original manuscript, the background ensemble was evaluated only during the assimilation process (Section 4.1), while the HF radar and reanalysis datasets were used as independent references for validating the analysis results. Following the reviewer's suggestion, we extended the evaluation by including the background versus reanalysis statistics.

A corresponding comparison with HF radar has not been added because the available HF radar observations have limited spatial coverage. These data are suitable for localized validation of the analysis but do not provide sufficient coverage for a representative and statistically robust assessment of the analysis over the study period. We therefore limited the additional background evaluation to the reanalysis dataset, which provides consistent spatial and temporal coverage across the entire domain.

2. The paper requires a more thorough and clear description of the MEnKF method. The authors mention an iterative strategy, but this isn't clearly defined in the text. Cycling aspects of the MEnKF are mentioned at times (e.g Figure 2) but it's unclear if and how this is applied in this study.

Response: The methodology section has been revised to provide a clearer description of the MEnKF framework and its implementation in this study. In particular, the description accompanying Figure 2 has been expanded to better explain the assimilation procedure. To improve clarity, Section 3.3.1 has been reorganized and expanded. The term *"iterative strategy"* has been replaced with *"sequential assimilation"* to more accurately represent the adopted framework.

The revised text now explicitly clarifies that assimilation was performed in repeated 3-hour forecast-analysis cycles using an 8-hour observation window, within which MCC-derived observations were assimilated sequentially according to their individual acquisition times rather than being aggregated into fixed-time averages. After all observations associated with a given analysis cycle were assimilated, the posterior ensemble state was used as the initial condition for the subsequent forecast cycle. This forecast-analysis cycling is part of the assimilation framework, while forecast evaluation itself is beyond the scope of the present study.

3. The paper would benefit from providing an example MCC EO image

Response: A representative example of the MCC-derived EO surface current field has been added to the revised manuscript (new Figure 3 in Section 3.1). This figure provides a visual example of the MCC-derived observations used in the data assimilation framework and improves the description of the EO retrieval process.

4. The observation errors for the HF radar should be provided

Response: The manuscript has been revised to provide the observation error information associated with the HF radar dataset and to clarify the quality control procedure applied prior to validation. The HF radar provides standard error estimates for both the zonal and meridional velocity components, which were used in the first stage of quality control. Observations with a reported standard error greater than 0.01 m/s for either velocity component were excluded to remove low-confidence retrievals. Second quality control step was then applied to remove spurious velocity fields by excluding observations with current speeds exceeding 2 m/s. To illustrate the spatial distribution and magnitude of the HF radar observation along with outliers, a new figure (Figure 9) has been added showing the study period-averaged velocity fields as both vector (quiver), magnitude (bar) plots, and mean error of the magnitudes (line). The description of the observation errors and the two-stage quality control procedure have been incorporated into Section 3.4.

5. More discussion/details on the handling of the different temporal windows for the observations is required. From my understanding, the temporal windows used to calculate the velocities can very from 1 hours to 24 hours. Are all these observations assumed to represent the 3 hourly mean currents? Are the handled in a different way during the data assimilation?

Response: The manuscript has been revised to clarify the treatment of the varying temporal windows associated with the MCC-derived observations. Each MCC retrieval represents the temporally averaged surface displacement estimated from an individual AVHRR image pair, with image separations ranging from approximately 1 to 24 hours depending on image availability, cloud contamination, preprocessing constraints, and spatial overlap requirements (see Section 3.1.1).

The revised text now explicitly distinguishes the three temporal scales used in the assimilation framework. The 3-hour interval defines the frequency at which the MEnKF analysis is updated. For each update, an 8-hour assimilation window is used to collect all available MCC-derived surface current products whose acquisition times fall within that window, thereby increasing observational availability under the irregular sampling characteristics of satellite imagery. The 1-24 hour temporal separation refers only to the image pairs used to derive each individual MCC current field and does not imply that all observations falling within a given assimilation window were processed using the same assimilation procedure, irrespective of their temporal separation.

The manuscript also clarifies that differences arising from the varying temporal separations were not modeled explicitly within the assimilation framework.

6. In several places there authors claim that current magnitudes are underestimated in the MEnKF (lines 535, 536) but in other places they state that the current magnitudes are overestimated in MEnKF (lines 561,685-686, 743,749,767). Please make sure the results are consistently and correctly represented

Response: The manuscript has been carefully reviewed, and all statements describing bias have been revised to ensure consistent interpretation throughout. In the revised manuscript, the bias definition is applied consistently throughout. Based on the updated analysis, the positive mean bias (Reanalysis − MEnKF), indicates a modest underestimation of current speed by the MEnKF relative to the NWS reanalysis. The same interpretation is maintained consistently in the validation, sensitivity analyses, discussion, and conclusions.

In addition, the text now explicitly notes that variations in bias among the observation-density and spatial-coverage classes reflect changes in assimilation performance under different observational conditions, while the overall tendency remains showing the underestimation relative to the reanalysis. This revision removes the previous ambiguity and ensures that all reported results are internally consistent.

7. In general, the conclusions drawn from the results seem too strong. Please ensure that the results are not overstated.

Response: The manuscript has been carefully reviewed to ensure that the conclusions remain fully aligned with the results presented. Statements in the Results, Discussion, and Conclusions sections have been revised to avoid overstating the findings and to more clearly distinguish the observed outcomes from broader implications. In particular, language implying general applicability or definitive performance improvements has been moderated, and the interpretation of the results is now framed within the scope of the study domain, observational datasets, and experimental configuration. The discussion has also been revised to better acknowledge the limitations associated with the selected study period, the use of archived forecast products, and the availability of EO-derived surface current observations. Overall, the revised manuscript presents a more balanced assessment of the proposed MEnKF framework while keeping the conclusions consistent with the evidence presented.

8. The discussion section would benefit from rationalisation. It is very long and not very coherent, in places it feels like statements are being repeated.

Response: The Discussion section has been carefully reviewed and revised to improve its overall organization and readability. Repetitive statements have been removed or consolidated, and several paragraphs have been reorganized to improve the logical flow between the interpretation of the assimilation results, the validation analyses, and the limitations of the proposed framework. The revised discussion places greater emphasis on interpreting the key findings while avoiding repetition of results already presented in the Results section, resulting in a more concise and coherent discussion.

Specific comments:

Introduction

9. Line 60 ". They describe the instantaneous motion of surface features" – given that they are the displacement between 2 sequential features, I don't think they can be considered instantaneous.

Response: We agree that the original wording was inaccurate. The MCC-derived currents do not represent instantaneous measurements, but temporally averaged surface displacements estimated from sequential AVHRR image pairs. The manuscript has been revised to describe the MCC-derived surface current observations accordingly, clarifying that they represent time-averaged displacement velocities derived over image-pair intervals ranging from approximately 1 to 24 hours

10. Line 99. These studies (Mirouze et al. 2024 and Waters et al. 2024) were assimilating synthetic data from a proposed satellite mission, real observation data of this type is not presently available. Please make sure this is clear in the text.

Response: The manuscript has been revised to clearly distinguish between synthetic and real satellite-derived surface current observations. It now explicitly states that the studies by Mirouze et al. (2024) and Waters et al. (2024) assimilated synthetic observations designed to emulate the proposed SKIM mission, rather than real observational data, which are not yet available. The revised text also clearly differentiates these synthetic datasets from the real EO-derived surface current observations used in the present study.

11. Line 114 “Despite these DA advances, only a limited fraction of available EO-derived ocean surface current observations” – This sentence is a bit misleading as in reality there are very limited NRT ocean surface currents available and the generally don’t provide the necessary spatial coverage for data assimilation. Mold this in context of SLA, based EO observations assimilation. NRT should be omitted.

Response: The manuscript has been revised to better place this statement in the context of established data assimilation systems based on sea level anomaly (SLA), sea surface temperature (SST), and in situ profile observations. The revised text now clarifies that, compared with these mature observing systems, EO-derived surface current observations remain limited in spatial and temporal coverage for assimilation applications. The unnecessary reference to near real-time (NRT) availability has also been removed.

12. Line 123 “Although EO satellite systems now provide near-global, high-temporal observations” – do you mean surface current velocities? I don’t think this is true.

The statement has been revised to avoid overstating the current capability of EO-derived surface current observations. The revised wording no longer implies the availability of near-global, continuously observed surface current velocities and instead describes the present capability as retrievals from multiple EO platforms with limited spatial and temporal coverage either due to clouds or trade-off with coarse spatial resolution.

Materials and methods

13. Line 159 – 161. “The selected period captures active mesoscale variability and evolving flow structures, offering a sufficiently complex dynamical regime for rigorous testing of ensemble behavior, covariance propagation, and incremental forecast skill.” – Do the authors have any evidence for this?

Response: The reviewer is correct that the original statement was not sufficiently supported. The manuscript has been revised to justify the selection of the study period citing oceanographic studies describing the winter circulation over the Galician shelf. The revised text cites the seasonal intensification of the Iberian Poleward Current (IPC), the occurrence of mesoscale eddies, and the influence of intermittent downwelling-favorable winds, which together characterize January as a period of enhanced dynamical variability (Peliz et al., 2005; Pingree and Le Cann, 1992; Torres and Barton, 2007; Teles-Machado et al., 2016). These references provide the physical basis for selecting the study period for evaluating the proposed assimilation framework.

14. Line 161 “The objective at this stage is not climatological generalization” –Can the authors please clarify what they mean by this.

Response: The phrase *"not climatological generalization"* was ambiguous and has been revised for clarity. The manuscript now explicitly states that the objective was not to assess performance across multiple seasons or circulation regimes, but to evaluate the assimilation framework during a representative period of dynamically active winter conditions over the Galician shelf.

15. Line 170 – Was does “harmonized” mean in this context? It’s not defined or explained until later in the text.

Response: The term *"harmonized"* was not sufficiently defined at its first occurrence. The manuscript has been revised to clarify that the preprocessing included interpolation of all model forecasts onto a common spatial grid, temporal alignment with the assimilation cycle, and standardization of the units and spatial domain. This clarification is now introduced before the detailed preprocessing methodology presented later in the manuscript.

16. Line 171 – What is the “associated climatology”? It hasn’t been defined yet.

Response: We agree that the phrase *"associated climatology"* was not defined yet at highlighted position in manuscript. To mitigate this ambiguity, we have revised the text of Sec. 2.2, and the climatology definition is now clearly defined in Sec. 3.2. and its usage in Sec. 3.3.2.2.

Edits: “*Forecasts spanning January-December 2020 and January 2021 were acquired for each forecasting system. The 2020 archive was used to construct model-specific climatological fields, whereas January 2021 forecast were used to generate the ensemble members employed in the assimilation experiments.*”

17. Table1. – More information is needed about the forecasts used in this study. What forecast lead times are use? How often are the forecasts initialised? My impression from the table is that both a global and regional version of NEMO are used. These should be separated in the table for clarity.

Response: Table 1 and the accompanying dataset description have been revised to provide additional information on the forecast systems used in this study. The revised table now includes the forecast lead times, forecast update frequency, atmospheric forcing, and data assimilation characteristics of each forecasting system. The two NEMO-based products, the CMEMS Global Ocean Physics Analysis and Forecast system (001_024) and the Mediterranean Sea Physical Analysis and Forecast system (006_013), are now presented separately to clearly distinguish their spatial domains and operational configurations. The manuscript also clarifies that the NWS reanalysis dataset was used solely for performance evaluation and was not included as an ensemble member.

Table 1. Summary of spatial resolution, latency, and temporal resolution of EO products and circulation models used in this study.

Dataset	Domain	Category	Role in This Study	Native DA Sys.	Spatial Res. (km)	Latency (h)	Temp. Res. (h)	Lead Time (day)
AVHRR	Regional	Satellite Data	Surface Current Retrieval	–	1.1	2–6	1–6	–
HFR	Coastal	In-Situ	Reference for Evaluation	–	5.5	–	1	–
HYCOM	Global	Forecast	Background Ensemble	NCODA (multivariate 3Dvar)	8.8	6	3	4
GRTOFS	Global	Forecast	Background Ensemble	RTOFS-DA (multivariate 3DVar)	8.8	12	3	8
MedSea Phys (006_013)	Regional	Forecast	Background Ensemble	OceanVar (3DVar)	4.2	6	1	4
CMEMS Glob Phys (001_024)	Global	Forecast	Background Ensemble	SAM2 (Kalman Filter)	9.25	12	3	9
NWS Reanalysis (004_009)	Regional	Reanalysis	Reference for Evaluation	NEMOVar (3DVar)	7	–	1	–

18. Line 197 – "The suitability of reanalysis products for evaluating EO-derived surface currents is well established." - can the reanalysis really be used to assess the ageostrophic currents? Since the reanalysis doesn't assimilate velocity observations directly, it may largely be useful for assessing geostrophic currents.

Response: We thank the reviewer for raising this point. We agree that the original statement was too broad and have revised the manuscript to better describe the role of the NWS reanalysis. Unlike products based solely on sea surface height assimilation, the CMEMS NWS reanalysis assimilates multiple complementary observations, including sea surface height, sea surface temperature, hydrographic profiles, HF radar measurements, and surface drifter observations. Consequently, the CMEMS NWS reanalysis represents both geostrophic and ageostrophic components of the surface circulation to the extent permitted by the assimilated observations and the underlying model dynamics. Nevertheless, it remains a model-constrained product and cannot be regarded as a complete representation of all high-frequency surface variability. The manuscript now explicitly clarifies that the NWS reanalysis is used as a spatially comprehensive, dynamically consistent reference for evaluating the large-scale and mesoscale structure of the MEnKF analyses, while acknowledging its limitations in representing the complete surface current field.

Edits: "*Ocean reanalysis products provide a physically consistent, dynamically constrained reference for evaluating the large-scale and mesoscale structure of surface circulation. The CMEMS NWS reanalysis assimilates multiple complementary observations, including sea surface height, sea surface temperature, hydrographic profiles, HF radar measurements, and surface drifter observations, enabling the representation of both geostrophic and ageostrophic components of the surface circulation. Although the resulting fields remain constrained by the underlying numerical model and may not fully resolve all high-frequency variability, they provide a spatially comprehensive reference for assessing the regional consistency of the MEnKF analyses.*"

19. Line 200- 204 – "More recently, Wang et al. (2023) reported high spatial correlations and relatively low RMSE between satellite-derived surface currents and CMEMS reanalysis fields, noting that residual discrepancies were often comparable to intrinsic model uncertainty rather than attributable solely to observational error."- In general, this paragraph is arguing that a reanalysis is a useful product for assessing the currents, however, this sentence appears to be counter to that argument. It is fine to include this, but it needs to be clear that this is a counter-argument in the phrasing of this text.

Response: We thank the reviewer for this important observation. We agree that the original statement overstated the suitability of reanalysis for evaluating total surface currents. The CMEMS NWS reanalysis assimilates sea surface height, sea surface temperature, and hydrographic profiles, but does not directly assimilate surface current velocity observations. Its skill is therefore primarily in representing the ageostrophic and geostrophic component of surface circulation, it remains a model-constrained product. Consequently, residual differences between EO-derived currents and the reanalysis may arise from uncertainties in both datasets rather than from observational error alone. The revised text now explicitly acknowledges this limitation while retaining the reanalysis as a valuable, spatially comprehensive reference for assessing the regional consistency of the MEnKF analyses.

Revised text: *" However, Wang et al. (2023) also reported that residual discrepancies between satellite-derived surface currents and NWS reanalysis were often comparable to the intrinsic uncertainty of the reanalysis itself, rather than being attributable solely to errors in the EO-derived observations. This provides an important caveat when interpreting reanalysis-based evaluations. Although the NWS reanalysis assimilates multiple complementary observations and represents both geostrophic and ageostrophic components of the surface circulation, it remains a dynamically constrained model product. Consequently, agreement with the reanalysis should be interpreted as evidence of consistency with a physically based reference rather than as a comparison against an error-free representation of the true ocean state."*

20. Section 2.2.4 - I don't think that at any point the authors explicitly list the number of ensemble members in their multi-model ensemble. Please make sure this is provided.

Response: We thank the reviewer for identifying this omission. The manuscript has been revised to explicitly state that the MEnKF framework employs a four-member heterogeneous multi-model ensemble. This information is now clearly provided in the summary of the ensemble configuration in Table 1 and is also explicitly described in Section 3.2.4.

21. Line 208 – In the text you refer to the "the CMEMS global analysis" but this isn't included in Table 1. Please make sure the text and table are consistent. It is also perhaps a little confusing that both the renanalysis and NEMO forecast model are CMEMS products using NEMO. I believe the are produced by different centres (UK Met Office and Mercator). A clearer naming convention would be beneficial.

Response: The manuscript has been revised to ensure consistency between the text and Table 1. The CMEMS Global Ocean Physics Analysis and Forecast system (001_024) is now explicitly listed in Table 1 as a background ensemble member.

To avoid ambiguity, the naming convention has also been standardized throughout the manuscript to clearly distinguish the three CMEMS products used in this study: (i) the CMEMS Global Ocean Physics Analysis and Forecast system (001_024), (ii) the Mediterranean Sea Physical Analysis and Forecast system (006_013), both used as background ensemble members, and (iii) the North-West European Shelf Reanalysis (004_009), which was used exclusively for independent validation. The revised text also clarifies that these are operationally distinct products with different domains and configurations.

22. Line 208- 210 – It is unclear to me whether the ensemble is made up for forecast outputs, or analysis outputs from the these models. Please make this clear. If forecasts are used the forecast lead times should be provided.

Response: The manuscript has been revised to explicitly clarify that the multi-model ensemble is constructed from the **forecast outputs** of the contributing operational systems rather than their internal analysis states. For each assimilation cycle, the forecast field corresponding to the analysis time was extracted from each forecasting system and used as an ensemble member within the MEnKF framework. The dataset description and Table 1 have also been updated to include the forecast update frequency and forecast lead times for each system.

23. Line 211 "All selected models provide surface current fields at spatial and temporal resolutions compatible with the underlying DA framework." What is that temporal and spatial resolution of the DA framework?

Response: The manuscript has been revised to explicitly define the spatial and temporal configuration of the data assimilation framework. Assimilation was performed on a common 0.11° spatial grid, consistent with the resolution of the MCC-derived surface current observations, with analyses conducted at 3-hour intervals. The revised text also clarifies that the forecast products from the contributing models were temporally aligned with the analysis cycle and interpolated onto the common grid prior to ensemble construction.

Methodology

24. Line 217 – "Harmonization" is used again but it hasn't yet been defined what this is.

Response: The manuscript has been revised to define the term *"harmonization"* at its first occurrence in the **Multi-model Ensemble Construction** section. The revised text explains that harmonization comprises the preprocessing applied prior to ensemble generation, including temporal alignment to the assimilation cycle, interpolation onto a common spatial grid, masking to a common study domain, and reconciliation of land-sea boundaries. With this definition introduced earlier in the manuscript, the subsequent use of the term is now unambiguous.

25. Line 217 – It would be useful for the reader if the authors included a sentence here to clarify that these components will be described in more details later in this section.

Response: During revision, the manuscript was reorganized to more clearly separate the dataset descriptions from the methodological procedures. The forecast products are now described in Section 2.2.4, while the ensemble construction, preprocessing, harmonization, and data assimilation methodology are presented in the dedicated subsections of Section

3. As these topics are now introduced in a more logical sequence, an additional forward reference is no longer necessary.

26. Figure 2 - I was very unclear on how the cycling aspects on the MEnKF framework worked and whether any cycling of the MEnKF was used data assimilation framework was applied in this study. Please include a more detailed description of this Figure.

Response: The description of Figure 2 has been expanded to more clearly explain the forecast-analysis cycling procedure adopted in this study. The revised text now describes the construction of the background ensemble at each 3-hour analysis cycle, the sequential assimilation of MCC-derived observations within the observation window, and the use of the updated ensemble state to initialize the subsequent cycle. The figure caption has also been revised to clarify the role of the archived forecast ensemble and the observation assimilation sequence. These revisions make the implementation of the MEnKF cycling framework explicit.

27. Line 228-231 "To mitigate data gaps and irregular acquisition timing inherent to NRT operations, a data-enrichment strategy was employed whereby all valid AVHRR image pairs acquired within a 24-hour window were systematically constructed. This enrichment strategy substantially increased the availability of short time-lag image pairs ($\Delta T < 3$ hour), critical for NRT ocean currents estimation." I found this difficult to understand, please rewrite. I found the explanation provided on lines 149-250 easier to understand.

Response: The description of the enrichment strategy has been rewritten to improve clarity. The revised text now explicitly explains that each AVHRR image was paired with all subsequent cloud-free images acquired within the following 24-hour period across the available NOAA and MetOp platforms. This multi-platform pairing generated multiple valid image combinations from the available observations, increasing the number of MCC-derived surface current retrievals and improving the availability of short time-lag image pairs. The revised description more clearly explains both the implementation and the rationale of the enrichment strategy.

28. Line 270 –Are missing vectors the same as missing pixels mentioned above? In figure 17, are the number of pixels therefore the number of "non fill-value" pixels? Please make this clear in the description of Figure 17.

Response: The manuscript has been revised to clarify that *missing vectors* refer to grid cells where no valid MCC-derived surface current retrieval is available because of cloud contamination, insufficient image overlap, low cross-correlation values, land masking, or quality-control filtering. These correspond to the locations previously referred to as *missing pixels*. The description of Figure 17 has also been revised to explicitly state that the reported pixel counts represent only valid MCC-derived current vectors after quality-control filtering and do not include placeholder (fill-value) grid cells introduced solely to maintain matrix dimensions during the assimilation process.

29. Line 271 - I have some concerns about the choice to fill missing data with very low velocity information. The surface currents in this area could be very strong but the data may be missing because of cloud cover. Assimilating these calm-wind condition values will systematically underestimate the currents. While this is addressed in this discussion, something should be said about this here.

Response: The manuscript has been revised to clarify that missing observations resulting from cloud cover or other retrieval limitations are **not** interpreted as weak surface currents. The placeholder value (0.001 m/s) is introduced solely to preserve matrix dimensional consistency during construction of the observation vector and is not treated as physical observation during assimilation. Missing observations remain identified by the original no-data mask and are assigned very large observation uncertainties, resulting in a negligible Kalman gain and effectively excluding them from the analysis update. Consequently, observational gaps are not filled by the placeholder values and therefore do not systematically bias the analysis toward weak currents. Instead, regions without valid MCC observations evolve according to the ensemble forecast and its forecast-error covariance. This clarification has been added to the methodology section to clearly distinguish the numerical treatment of missing values from their physical interpretation.

30. Line 290 – Need to state that this is Surface current velocity components.

Response: The manuscript has been revised to explicitly state that the variables refer to the zonal and meridional surface current velocity components, removing the ambiguity identified by the reviewer.

31. Line 291 –“ extracted at matching lead times” – what is this lead time?

Response: The text has been revised to remove the ambiguous reference to *"matching lead times."* It now explicitly states that the surface current velocity components (*uo*, *vo*) were temporally aligned by extracting the forecast fields valid at each 3-hour assimilation time.

32. Line 292 – “forecasts were resampled onto the highest-resolution regional grid using bilinear interpolation” – It would be useful to state the resolution used explicitly here.

Response: The text has been revised to explicitly state the spatial resolution used during preprocessing. It now reads: *"Forecast fields were interpolated onto a common 0.11° spatial grid, corresponding to the native resolution of the MCC-derived surface current observations, using bilinear interpolation."*

33. Line 309 – “and the resulting posterior ensemble was then advanced to initialize the subsequent forecast cycle. This approach ensures temporal consistency between asynchronous observations and forecast models.” - Please provide more explanation of this cycling aspect in the text.

Response: The manuscript has been revised to provide a clearer description of the forecast-analysis cycling procedure. The revised text now explicitly explains the 3-hour assimilation cycle, the 8-hour observation window, the sequential assimilation of MCC-derived observations according to their acquisition times, and the use of the updated analysis as the background state for the subsequent cycle. The description and caption of Figure 2 have also been expanded to more clearly illustrate the cycling procedure.

34. Line 318 – “the global analysis field” – how is this defined?

Response: The term *"global analysis field"* was ambiguous and has been revised. It was not intended to refer to a global ocean model or an external analysis product, but rather to the complete analysis state over the study domain obtained after applying the localized MEnKF updates. The revised text now explicitly states that the localized analysis increments are mapped back onto the full study-domain analysis field, which is then used in the subsequent assimilation cycle.

35. 318-320 – “The resulting analysis state was then used as the initial state for the next assimilation cycle if additional EO products remained available within the current time window, or otherwise adopted directly as the initial state for the forward forecast.” - were forward forecasts run in this study? Please make this clearer.

Response: No forward model integrations were performed in this study. The manuscript has been revised to clarify that the MEnKF framework was evaluated in a retrospective configuration using archived forecast products to construct the background ensemble at each assimilation time. The updated analysis state was conceptually treated as the initial condition for the subsequent assimilation cycle; however, no new forecast integrations were generated between cycles. The corresponding text has been revised to remove any implication that forward forecasting was performed.

36. Line 364 – Is anything done to account for the different deltaT used in the AVHRR pairing? Please clarify the temporal interpolation aspect.

Response: The manuscript has been revised to clarify how the varying image-pair time separations (ΔT) are handled within the assimilation framework. The MCC-derived surface currents were not temporally interpolated or normalized to a common averaging interval prior to assimilation. Instead, each retrieval was assimilated at its corresponding acquisition time and interpreted as the mean surface displacement over its native image-pair interval (approximately 1–24 hours). Observations were assimilated sequentially within the 8-hour assimilation window according to their acquisition times, without being converted to equivalent 3-hour mean currents. The influence of differing ΔT values was not modeled explicitly; instead, unreliable retrievals associated with unfavorable image pairs were mitigated through the MCC quality-control procedures, including cross-correlation thresholding, low-pass filtering, and other screening steps described in Section 3.1.1. The manuscript now explicitly acknowledges this as a practical limitation of the present framework. Additionally, time-lag based influence on ocean currents retrieval is discussed in detail in our companion paper (Baig at al., 2026) [under review].

37. Line 387 -388 - "Following the analysis update, Xa and Pa serve as the prior for the subsequent assimilation cycle, with the final analysis state mapped back to the native model grid to initialize the next forecast." – From my understanding this study is using historical forecasts from an archive, how can the analysis state be used to initialised the next forecast in this framework?

Response: The reviewer is correct. No forward model integrations were performed in this study, and the original wording reflected the generic forecast-analysis cycle of the EnKF rather than the retrospective implementation adopted here. The manuscript has been revised to remove references to forecast initialization and to clarify that archived forecast fields were used to construct the background ensemble at each assimilation time. The updated ensemble state is now described only within the context of the assimilation procedure, eliminating the implication that new forecasts were generated between assimilation cycles.

38. Line 423 – Please specify what the withheld observations are and how they are sampled.

Response: The term *"withheld observations"* was misleading and has been removed from the manuscript. No independent subset of observations was reserved for validation. Instead, the RMSE was computed between the current analysis field and the next available MCC-derived surface current observation within the sequential assimilation procedure, prior to assimilating that observation. The revised text now explicitly describes this procedure to avoid confusion with an independent validation experiment.

Results

39. Line 443 "By effectively integrating dynamically relevant information from EO data into the MEnKF, the system significantly enhances the representation of mesoscale circulation features." – which result demonstrates this?

Response: We thank the reviewer for this observation. We agree that the original statement was stronger than warranted by the analyses presented in the manuscript. Although the vector-field comparisons, statistical metrics, and complex correlation analysis indicate that the MEnKF analyses reproduce the dominant circulation patterns and exhibit improved agreement with the reference datasets, mesoscale circulation features were not evaluated explicitly. The statement has therefore been revised to avoid making an unsupported claim and now refers only to the improvements that are directly demonstrated by the quantitative validation results presented in Sections 4.2 and 4.3. The broader implications for representing mesoscale circulation are instead discussed qualitatively in the Discussion section.

Edits: *" By assimilating EO-derived surface current observations into the MEnKF, the analysis consistently reduces RMSE and MAE relative to the background ensemble and improves agreement with the independent HFR observations and the NWS reanalysis. These results indicate improved representation of the resolved surface circulation, while the implications for mesoscale circulation features are discussed in Section 5."*

40. Figure 6 – I assume these are the RMSE of the ensemble means? Please clearly define this. Same goes for line 450.

Response: The reviewer is correct. Figure 6 (now Figure 9) presents the RMSE of the background and analysis ensemble mean fields, but this was not stated explicitly in either the figure caption or the main text. Both have been revised to clarify that the RMSE is computed for the ensemble mean fields related to the corresponding MCC-derived surface current observations.

Revised Figure 6 (now figure 9) caption: Temporal evolution of RMSE between the ensemble mean and the MCC-derived surface current observations for the uo and vo velocity components during January 2021. Dashed lines represent the RMSE of the background ensemble mean relative to the observations, while solid lines denote the RMSE of the analysis ensemble mean after assimilation. Green-shaded regions highlight the reduction in RMSE achieved from the background to the final analysis state at each assimilation cycle. Gray bars indicate the number of assimilated EO-derived MCC observations available at each cycle

Edits: "*The temporal evolution of assimilation performance was assessed using RMSE values computed at each assimilation cycle. For each cycle, the background and analysis ensemble means were evaluated against the next available MCC-derived surface current observations before their assimilation. This provides a quantitative measure of how the DA framework modifies forecast skill over time while accounting for the irregular temporal availability of EO-derived observations.*"

41. In section 4.1 does the sampling of the reference data also vary temporally? This may have an impact on the results.

Response: In Section 4.1, the reference dataset used to compute the background and analysis RMSE consists of the MCC-derived surface current observations available within each assimilation cycle, rather than a fixed dataset. Consequently, the number and spatial distribution of reference observations vary throughout the study period owing to satellite sampling and cloud cover. The RMSE at each assimilation cycle is therefore calculated only from the observations available for that cycle. To make this explicit, the manuscript has been revised to clarify the RMSE calculation, and the corresponding figure now includes the number of available MCC observations (grey bars) for each assimilation cycle. This allows the temporal variability in sampling to be considered when interpreting the RMSE time series.

42. In figure 7a, the RMSE is low at 03:00 despite a limited number of EO images. I don't think this is discussed. Is there an explanation for this?

Response: The reviewer is correct that this feature was not discussed in the original manuscript. Section 4.1 has been expanded to address the relatively low RMSE observed at 03:00 UTC despite the smaller number of available EO observations. Section 4.1 has now been expanded to address this behavior. The revised discussion emphasizes that assimilation performance depends not only on the number of observations but also on their quality and spatial distribution. During the nighttime period (03:00 UTC), the absence of daytime solar heating produces more coherent SST patterns, which are generally more favorable for MCC feature tracking and lead to more reliable surface current retrievals despite the reduced number of image pairs. The revised manuscript also notes that the limited spatial coverage at 03:00 UTC may sample the flow field unevenly, contributing to the slightly higher errors in the vo component relative to the uo component, while the overall RMSE remains comparatively low.

43. Line 492 – "RMSE values range from 0.15–0.17 m/s and MAE from 0.10 to 0.13 m/s" These values don't seem to be entirely consistent with results in the tables provided with the Figures 8-10.

Response: The apparent inconsistency arose because the values reported at this point refer to the current speed, whereas the accompanying tables present the RMSE and MAE for the *uo* and *vo* velocity components separately, as well as the current speed magnitudes. This distinction was not explicitly stated in the original manuscript. The text has been revised to clarify that the reported RMSE and MAE values correspond to the current speed, removing the ambiguity.

44. Line 493 – Where the paper mentions the small positive bias, does this mean in the current speed?

Response: Yes. The reported positive bias refers to the current speed, not the individual velocity components. The manuscript has been revised to explicitly state this and avoid ambiguity.

45. Line 499-501: “Nevertheless, these results indicate a moderate level of consistency between the MEnKF analysis and HFR observations, demonstrating the effectiveness of the assimilation in constraining surface circulation when direct velocity measurements are available.” - without providing the baseline of the background statistics, the results don’t clearly demonstrate that the EO assimilation is constraining the surface circulation. This comment also applies to lines 646-647 and line 818.

Response: We agree that the original manuscript did not provide the corresponding background performance statistics, making it difficult to assess the impact of assimilating the EO-derived surface currents. To address this, we have added background performance metrics to Table 3 (provided below) and updated Section 4.3 to present both the background ensemble means and the assimilated analysis relative to the NWS reanalysis. The comparison shows that assimilation consistently improves the agreement with the reference dataset, reducing the current-speed RMSE from 0.29 to 0.20 m/s and the MAE from 0.21 to 0.15 m/s, with similar improvements for both the zonal and meridional velocity components. These results provide the quantitative baseline requested by the reviewer and demonstrate that assimilation improves the analyzed surface circulation relative to the background ensemble. In addition, the corresponding statements throughout the manuscript (Lines 499–501, 646–647, and 818) have been revised to avoid overstating the results and now state that the assimilation improves agreement with the reference datasets while reproducing the dominant circulation patterns, rather than implying a complete constraint of the surface circulation.

Table 3. Mean performance metrics of the MEnKF analysis relative to the NWS reanalysis surface currents product, quantified by RMSE, MAE, and bias, averaged across all assimilation cycles. The Scatter Index (SI) is reported against current magnitude.

Metric	Background Ensemble Mean			Assimilated Analysis Mean		
	uo	*vo*	Speed	*uo*	*vo*	Speed
RMSE (m/s)	0.33	0.32	0.29	0.22	0.23	0.20
MAE (m/s)	0.24	0.25	0.21	0.18	0.18	0.15
Bias (m/s)	-0.02	-0.06	0.11	0.048	-0.03	0.08
SI	-	-	0.57	-	-	0.66

46. Line 532 – “observation – analysis” – since the results are comparing the reanalysis and MEnKF this needs clarification

Response: The reviewer is correct that the terminology was ambiguous. In this context, "observation" referred to the CMEMS reanalysis used as the reference dataset rather than the assimilated EO observations. To avoid confusion with standard data assimilation terminology, the text has been revised to explicitly state that the differences are computed between the CMEMS reanalysis and the MEnKF analysis.

47. Line 547 – state here what the RMSE, bias and SI are calculated relative to.

Response: We have revised the text at this location to explicitly state how the bias is defined and to provide a consistent interpretation throughout the manuscript. The revised manuscript now states that RMSE, bias, and SI was calculated as the difference between the MEnKF estimate and the reference dataset (MEnKF − NWS Reanalysis).

48. Line 583 - The persistent negative bias suggests a consistent tendency in the assimilation system of the observations themselves." – this sentence isn't clear

Response: The original sentence has been removed and replaced with a clearer interpretation based on the revised analysis. After re-evaluation the script configuration and re-driving the results, the revised manuscript now states that the observed speed bias is small relative to the RMSE and MAE and indicates a slight overall underestimation of current speed relative to the NWS reanalysis because of its positive sign. To avoid overinterpretation, the Results section is limited to describing the observed bias characteristics, while the potential contributing factors; including differences between the EO-derived surface currents and the reanalysis reference, analysis vs reanalysis differences, and the influence of dynamical smoothing in the reanalysis; are discussed separately in the Discussion section. This revision removes the ambiguity in the original wording and provides a clearer distinction between the reported results and their interpretation.

49. Section 4.3.1 – In figure 17 and 18 how much are the results impacted by the number of samples in each bin? The results may not very robust if there are only a small number of samples in some of the bin. This should be taken into account when the authors draw their conclusions from the data.

We thank the reviewer for highlighting the importance of sample size when interpreting the binned sensitivity analyses. To address this concern, Figures 17 and 18 have been substantially revised. Figure 17 now reports the number of assimilation cycles contributing to each bin together with the corresponding ±1 standard error envelopes, allowing the statistical robustness of each estimate to be assessed directly. Figure 18 has also been redesigned to include a sample-count-weighted representation of the joint RMSE distribution, distinguishing systematic performance patterns from isolated low-error bins supported by only a few samples. The revised figures are included with the response.

The accompanying discussion has also been updated accordingly. Rather than emphasizing isolated minima, the revised interpretation now focuses on the observational regimes that are consistently supported by sufficient numbers of assimilation cycles. The analysis shows that the most robust assimilation performance occurs for intermediate observational densities (approximately 9–20 assimilated EO products) combined with moderate-to-high spatial coverage (approximately 3000–3500 valid observation pixels), where both the error metrics and the sample populations are favorable. Conversely, several bins exhibiting very low RMSE are now explicitly identified as being supported by only one or two assimilation cycles and are therefore not interpreted as statistically robust. Similarly, the apparent increase in SI at the highest observational densities is no longer interpreted as a meaningful trend because it is based on only two samples. These revisions ensure that the conclusions are supported by the statistical confidence associated with each observational regime rather than by isolated low-error cases.

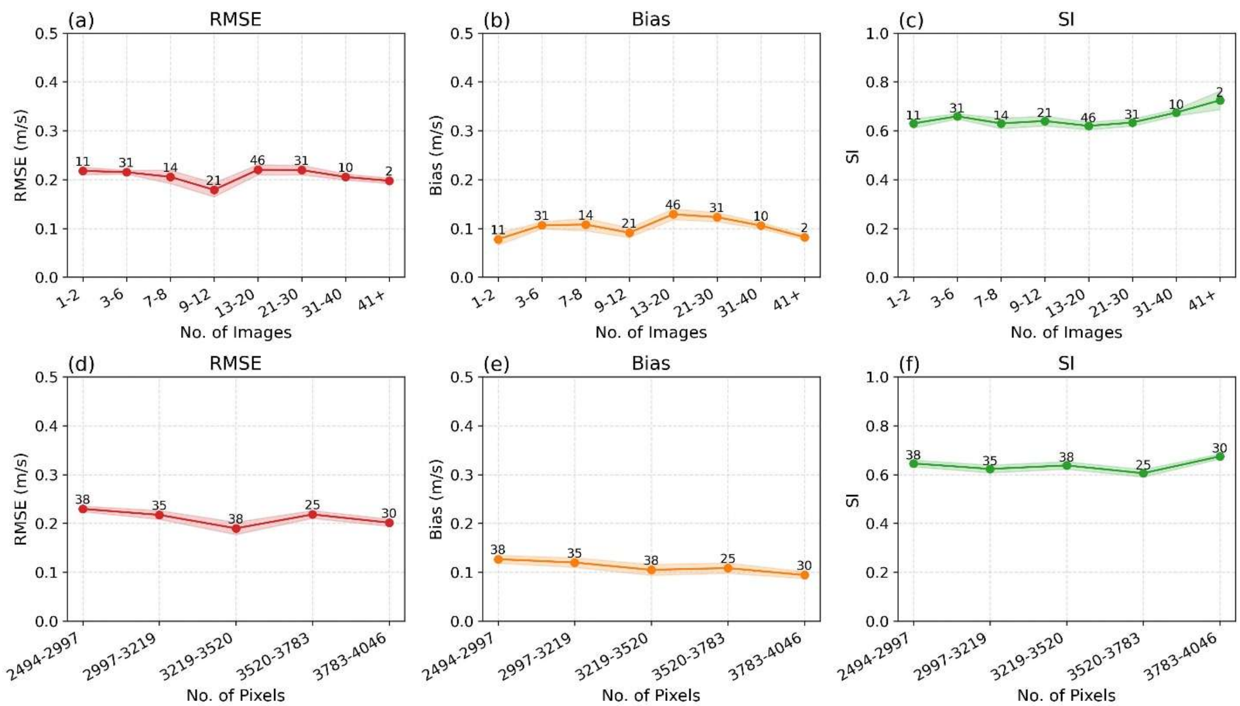

Figure 17: Sensitivity of MEnKF assimilation performance to EO-derived observation density and spatial coverage. Panels (a-c) show RMSE, bias, and Scatter Index (SI) for current speed as functions of observational density, expressed as the number of assimilated EO-derived current products per assimilation cycle. Panels (d-f) show the same metrics as functions of spatial coverage, expressed as the number of valid MCC-derived current vectors after quality-control filtering. Markers denote bin-averaged values, annotated with the corresponding sample counts (*n*), while shaded envelopes represent the ±1 standard error.

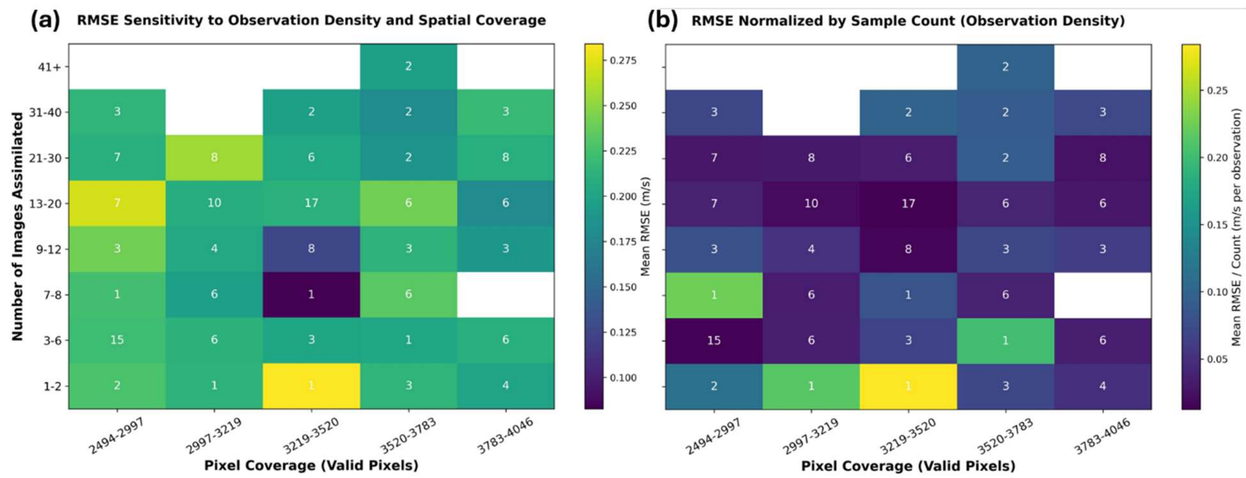

Figure 18: Joint sensitivity of MEnKF assimilation performance to observational density and spatial coverage. (a) Mean RMSE aggregated by binned number of assimilated images (y-axis) and valid observation coverage (x-axis), with color indicating average RMSE within each bin. (b) The values in (a) normalized by the number of assimilation cycles contributing to each bin, shown to illustrated how strongly each bin's RMSE estimate is influenced by its underlying sample size. White cells indicate bins with no available samples.

Discussion

50. Line 652 - (Wang et al., 2023; King et al., 2021; Vandenbulcke et al., 2017). – please check these references are being used correctly.

Response: Following the reviewer's comment, the supporting references have been reviewed and revised to ensure they directly support the statement. The previous citations have been replaced with studies specifically demonstrating improvements in upper-ocean circulation through the assimilation of satellite SST and coastal HF radar observations. The revised references are more closely aligned with the discussion presented in this section.

51. Line 666 – “The mean RMSE of ~0.18 m/s against HFR measurements, ~60% less than the internal misfit, confirms that the MEnKF achieves meaningful correction of the background ocean state” – What does the internal misfit mean? Where was this 60% shown?

Response: The term "internal misfit" referred to the mean analysis vs MCC-derived observations misfit (Mean RMSE ~0.50 m/s presented as black line) reported in Section 4.1. We agree that this terminology was insufficiently defined and that the basis of the reported 60% difference was not clearly presented. The comparison involved two different reference datasets with different error characteristics and sampling properties, making the percentage reduction potentially misleading. To avoid overinterpretation, this comparison has been removed, and the discussion now focuses on the HFR and NWS reanalysis as reference for validating the MEnKF performance.

52. Line 686 – “This is physically consistent with the known tendency of MCC-derived observations to capture energetic surface features, particularly in regions of strong thermal gradients” – this seems to contradict previous statements in the paper e.g line 512-513.

Response: The reviewer is correct that the original wording could be interpreted as inconsistent with the earlier discussion. Our intention was to distinguish between the ability of MCC-derived observations to identify dynamically active regions associated with strong thermal gradients and their known limitations in accurately estimating the magnitude of energetic surface currents. The text has been revised to make this distinction explicit and to avoid implying that MCC-derived observations fully resolve the intensity of mesoscale circulation features.

Edit: This behavior is consistent with the ability of MCC-derived observations to identify dynamically active regions associated with strong thermal gradients, while their velocity estimates may still underestimate the magnitude of energetic surface currents (Bowen et al., 2002).

53. Line 730 – “The temporal stability throughout January 2021 (RMSE ~0.6 m/s), maintaining stable performance across the variable winter time characteristics of the Galician shelf, including mesoscale eddy activity and intermittent poleward flow events” – This sentence isn’t clear.

Response: We agree that the original sentence was unclear. The text has been revised to explicitly state that the RMSE remained relatively consistent throughout the study period and despite the variable winter circulation conditions of the Galician shelf, including mesoscale eddy activity and intermittent poleward flow events, the performance of the framework remained stable.

54. Line 767 – would it be better to not assimilate the observations with low signal to noise (this would include the calm-weather values used to fill missing data)?

Response: The reviewer raises an important point. Low-quality MCC retrievals are already excluded prior to assimilation through the applied quality-control procedures, including the MCC cross-correlation threshold ($\rho \geq 0.6$) and subsequent spatial filtering, which remove unreliable and spatially inconsistent vectors before assimilation.

The near-zero values assigned to cloud-masked regions are not observations, but numerical placeholders introduced solely to maintain matrix dimensional consistency during construction of the observation vector. As described in the revised methodology, these entries were assigned very large observation uncertainties, effectively preventing them from influencing the analysis update. The manuscript has been revised to clarify this distinction and the role of the placeholder values within the assimilation framework.

Conclusion

55. The authors mention the calm-weather fill values in the conclusion, but I'm not clear why it is necessary to fill these values. Data assimilation algorithms can deal with irregularly spaced observations already.

Response: We thank the reviewer for this important observation. We agree that ensemble-based data assimilation methods can naturally accommodate spatially irregular observations and do not inherently require gap filling. The value previously referred to as a *calm-weather value* has therefore been clarified throughout the manuscript as a numerical placeholder used solely to preserve a consistent observation-vector structure during matrix construction and MEnKF computations, thereby avoiding numerical issues associated with variable observation dimensions. It is not treated as physical observation during assimilation. The placeholder entries remain identified by the original no-data mask and are assigned very large observation uncertainties, resulting in negligible Kalman gain and effectively excluding them from the analysis update.

The motivation for future gap reconstruction is therefore not to satisfy a requirement of the EnKF, but to increase the amount of physically meaningful observational information available for assimilation. As illustrated in the newly added Figure 4, MCC-derived surface current fields frequently contain extensive gaps caused by cloud cover, weak thermal contrast, and quality-control filtering, which reduce the observational constraint over dynamically important regions. Physically consistent gap-reconstruction methods could recover additional information from these incomplete fields, improve the spatial continuity of the assimilated observations, and provide a stronger observational constraint while remaining fully compatible with the EnKF framework. Such methods would also establish a foundation for future multi-sensor data fusion, enabling complementary EO observations to be integrated into a more spatially complete surface-current product prior to assimilation. Accordingly, the Conclusions section has been revised to identify advanced statistical and machine-learning-based gap reconstruction as an important direction for future work.

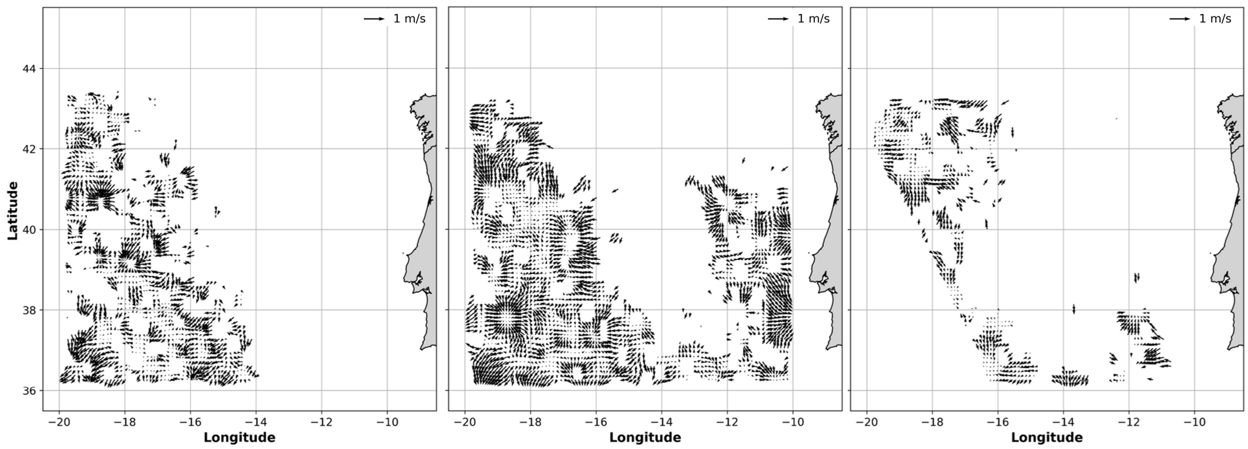

Figure 4. Representative examples of EO-derived surface current fields estimated using the MCC algorithm from AVHRR image pairs with short, moderate, and long temporal separations ranging from approximately 1 to 24 h over 01, 02, 03 January 2021 (left to right). Grey-shaded regions along the right-hand side of each panel indicate the Spanish coastline.

Reference

56. Vandenbulcke, L., Barth, A., and Beckers, J.-M.: Assimilation of High-Frequency Radar Surface Currents to Improve the Representation of 1055 Inertial Oscillations in a Coastal Ocean Model, Ocean Science, 13, 1–20, https://doi.org/10.5194/os-13-1-2017, 2017. – The DOI goes to a different paper.

Response: We thank the reviewer for identifying this error. The DOI has been corrected to reference the appropriate publication.

57. Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, pp. 343–367,https://doi.org/10.1007/s10236-003-0036-9, 2003 – This is provided in - references twice.

Response: The duplicate reference resulted from the same publication being included under two different bibliography entries. This has been corrected, and a single consistent reference is now used throughout the manuscript.