https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

Quantitative Evaluation of Mesoscale Eddies in the North Atlantic Using Satellite Altimetry and Ocean Reanalyses

Paolo Mauriello[1,2], Gregory C. Smith[3], Andrea Storto[1], Chunxue Yang[1]

[1]Institute of Marine Science, National Research Council of Italy, Rome, Italy
[2] Department of Physics "E.R. Caianiello", University of Salerno, Fisciano, Italy
[3] Meteorological Research Division, Environment and Climate Change Canada (ECCC), Dorval, Canada

Correspondence to: Paolo Mauriello (pmauriello@unisa.it – paolo.mauriello@artov.ismar.cnr.it)

Abstract. Ocean mesoscale eddies (10–250 km) play a key role in transporting heat, momentum, and nutrients. Accurately evaluating their representation in ocean reanalysis becomes important given that ocean reanalysis is widely used in ocean and climate research due to their temporal and spatial consistent coverage. This study aims to assess two global ocean reanalysis products—GLORYS12V1 (1/12° resolution) and GLORYS2V4 (1/4° resolution) available from the Copernicus Marine Service—against satellite altimetry observations. We have used both AVISO SSALTO/DUACS (1/4°) and the higher-resolution AVISO SWOT MIOST Science product (1/8°), as reference datasets that allow us to understand better the advantage of high-resolution altimetry data and have a fair evaluation for eddy-resolving ocean reanalysis (1/12°). The evaluation approach is based on a feature-based eddy-verification method to compare eddy properties for example amplitude, radius, centroid, shape using a cost-function metric, Probability of Detection (POD) and False Alarm Ratio (FAR) are then used to quantify the reanalysis skill. GLORYS12V1 demonstrates better agreement with observations than GLORYS2V4, especially for eddies with amplitudes >10 cm. For both DUACS and SWOT MIOST Science, the POD increases by more than 30% when moving from GLORYS2V4 to GLORYS12V1, while the FAR also rises by about 20%, mainly due to the detection of more small and weak eddies. The hit-cost metrics also improve with SWOT, the total hit cost decreases by more than 9%, and the amplitude error is reduced by about 6%. Small increases in radius and distance between centroids are still observed, but these are smaller in the SWOT comparison, showing better compatibility with the high-resolution reanalysis. Overall, POD values remain above 50% and reach around 60% with SWOT, while FAR stays below 30%. These results underline the added value of high model resolution for the representation mesoscale eddies in ocean reanalyses and highlight the value of wide-swath altimetry for model verification.

1 Introduction

Ocean mesoscale features with spatial scales from tens to hundreds of kilometers and time scales from weeks to months were first observed in the early 1930s by Columbus Iselin and colleagues during Montauk-to-Bermuda Cruises, as in Cullen (2005). Since 1992, the launch of satellite altimetry missions has enabled global, high-resolution monitoring of sea level and ocean currents, significantly improving our understanding of ocean features, particularly mesoscale dynamics, as in La Traon and

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

Morrow (2001). Over the past decades, numerous studies have explored the dynamics and impacts of eddies, given their significant roles in transporting water masses, heat, salt and nutrients (Rhines and Young, 1982; Richardson, 1983; Stephens & Marshall, 1999; Chenillat et al., 2016; Yang et al., 2022). Recognizing their important role, several studies in recent years have investigated the predictability of eddies (Xu et al., 2018; Thoppil et al., 2021).

However, satellite altimetry only measures the surface signature of the eddy features. In contrast, ocean reanalyses produced by combining numerical models with observations through data assimilation schemes provides the evolution of four-dimensional ocean state estimates, enabling us to study the ocean meso-scale eddy structures not only at the surface but the subsurface. Therefore, it is important to evaluate how well ocean reanalyses represent mesoscale eddies to better understand the uncertainty inherent in these products. In recent years, the spatial resolution of ocean models has increased allowing a greater representation of smaller ocean features. The first aim of this study is to evaluate the ability of ocean reanalyses with different spatial resolutions (1/4° and 1/12°) to represent mesoscale eddies in the North Atlantic.

Several studies have been performed to investigate the ocean meso-scale eddy properties (population, duration, lifetime) in numerical models and reanalyses, such as de Vos et al. (2018), Mason et al. (2019), Xie et al. (2020), Bonaduce et al. (2021), Smith and Fortin (2022). Studies have also focused on understanding the impact of data assimilation (Cipollone et al., 2017; de Vos et al., 2018). The detection methods used in these studies are usually based on the Okubo-Weiss parameter, the wavelet analysis of the relative vorticity field and geometry of the streamlines (Souza et al., 2011). Another very popular technique is carried out by Chelton et al. (2011) with satellite altimetry. The py-eddy-tracker package developed by Mason et al. (2014) is widely used in the community to explore the eddy properties at regional and global scales as in de Vos et al. (2018) and López-Álzate et al. (2022). This detection and tracking method allow us to evaluate the statistical properties of ocean meso-scale eddies such as population, lifetime and tracks. However, a feature detection method is needed to evaluate the accuracy of the eddy feature detected at specific locations. To fill the gap, Smith and Fortin (2022) developed a verification method to evaluate the capability of ocean reanalyses in representing the eddy features at exact locations and time detected in satellite altimetry data. In this study we will adopt this method to evaluate the representation of eddies in two ocean reanalyses, thereby opening the door to a greater exploitation of ocean reanalyses.

A reference dataset is needed to validate and evaluate the performance of ocean numerical models in representing meso-scale eddies. The finer spatial resolution of ocean reanalyses creates a problem when using the 1/4° AVISO DUACS data set as a reference, because this dataset cannot capture many small-scale features. Amores et al. (2018) show that signals smaller than about 50 km can be aliased into longer wavelengths in satellite altimetry. This means that some of the small-scale variability found in high-resolution reanalyses may not be shown correctly in the AVISO product. For this reason, we did not use the ¼° AVISO DUACS product in this study. Instead, we used the AVISO DUACS dataset at 1/8°, which offers a finer spatial resolution and allows a better comparison with high-resolution ocean reanalyses. The AVISO DUACS 1/8° satellite altimetry dataset, available from 1993 to the present at the global scale, has been widely used as a benchmark to validate and assess numerical models and ocean reanalyses in many studies.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

The availability of observations from the SWOT (Surface Water and Ocean Topography) wide-swath satellite altimetry mission provides another possible reference dataset for evaluating higher-resolution ocean reanalyses. SWOT was launched in December 2022 and offers wide-swath measurements of sea-surface height, which allows observations of ocean features at much finer spatial scales than traditional altimetry. However, there are some important limitations. First, because the mission is very recent, the time coverage of SWOT data is limited. Second, the primary purpose of SWOT is data assimilation for models, not necessarily validation, this can affect its suitability as an independent verifying dataset. Third, SWOT-derived gridded products are not fully independent of existing satellite altimetry, as they rely on a combination of SWOT wide-swath observations and multi-mission nadir altimetry used to constrain large-scale signals. Since similar nadir altimetry observations are assimilated in ocean reanalysis products, SWOT products cannot be considered a fully independent reference dataset. In this study, satellite datasets are therefore used primarily as complementary evaluation references rather than strictly independent verification datasets. The AVISO DUACS 1/8° product is not independent from the reanalyses, as the latter assimilate along-track altimeter data from the same satellite missions used in DUACS. Similarly, the SWOT 1/8° gridded product is partially dependent on conventional altimetry through its multi-mission background and mapping methodology, even though it introduces additional fine-scale information from the SWOT wide-swath measurements that is not directly assimilated in the reanalysis products considered here. For these reasons, we use both the AVISO DUACS 1/8° dataset and the SWOT 1/8° dataset to evaluate ocean reanalysis products at different spatial resolutions. This dual approach allows us to assess the representation of mesoscale eddies in reanalyses while accounting for the limitations related to data dependence. In particular, the comparison with SWOT provides insight into how well reanalyses capture fine-scale variability revealed by wide-swath altimetry, while the DUACS-based evaluation offers continuity with the pre-SWOT period. The combination helps us test the performance of reanalyses in representing mesoscale eddies while highlighting the need for high-resolution altimetry data for validating high-resolution ocean reanalyses or numerical models in general. This comparison also provides an estimate of uncertainty in the verification metrics when using AVISO DUACS 1/8° in the pre-SWOT period. An evaluation over the full period of the reanalyses (1993 – present) will be the focus of a future study.

In the following sections, we will introduce in detail the data and methods we use in this study and present the results of the evaluation (Section 2) and finally provide assessment results in terms of ocean reanalyses in representing ocean meso-scale eddies (Section 3).

2 Data and Methods

In this study, we use two Global Ocean Reanalysis datasets, GLORYS12V1 and GLORYS2V4, and two standard satellite-altimetry gridded-observational products, AVISO DUACS L4 product and SWOT MIOST L4 product. Detailed descriptions are provided below.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

2.1 Satellite Altimetry Products

The AVISO DUACS global 1/8° daily gridded product provides sea level anomaly (SLA), absolute dynamic topography (ADT), and geostrophic velocities using measurements from various satellite altimetry missions. These data are processed by the CNES/CLS DUACS system, which merges observations from several altimeters: Sentinel-6A, Sentinel-3A/B, Jason-3, CryoSat-2, SARAL/AltiKa, TOPEX/Poseidon, Jason-1/2, and Envisat. The product is distributed by CMEMS with global coverage. It is widely used in ocean forecasting and variability studies (Taburet et al., 2019; Ballarotta et al., 2019). Although the nominal grid is 1/8°(≈12.5 km), the effective resolution is lower because the traditional optimal interpolation (OI) used by DUACS smooths small-scale structures. These choices are necessary to merge observations from different satellite missions and to reduce along-track noise, but they also remove part of the small-scale variability. The effective resolving capability is typically between 50 and 100 km, depending on latitude and observing constellation. This means that the product cannot fully represent many mesoscale eddies smaller than ~50 km and cannot capture sub-mesoscale signals, as documented in the CMEMS Sea Level Product User Manual (Issue 1.0, 2025). Previous studies show that typical mesoscale eddies at mid-latitudes have a radius of about 80 km, while altimetry products can effectively resolve eddies with radii of about ~50 Km, as described in Chelton et al. (2011b) and Ballarotta et al. (2019). This suggests that part of the small-scale mesoscale variability is filtered out or underestimated. The formal mapping error fields included in the product highlight this limitation and depend on satellite sampling and the spatial scales resolved by the mapping procedure.

To address the spatial-resolution limitations of standard DUACS mapped products, we also consider the experimental global 1/8° SWOT MIOST gridded sea-level dataset distributed by AVISO/DUACS (CLS and CNES, 2025). This product combines observations from the contemporary nadir altimeter constellation (including Jason-3, Sentinel-3A/B, Sentinel-6A, SARAL/AltiKa, CryoSat-2 and Haiyang-2) together with SWOT Level-3 measurements, incorporating both the nadir altimeter and the Ka-band Radar Interferometer (KaRIn). The dataset is generated using the MIOST (Multi-Scale Inversion of Ocean Surface Topography) framework, a multiscale inversion approach designed to improve the reconstruction of sea surface height fields by representing multiple dynamical modes of ocean variability, in contrast to classical single-scale optimal interpolation methods used in DUACS that are known to smooth mesoscale variability and limit effective resolution (Ubelmann et al., 2015; Rogé et al., 2017; CLS and CNES, 2025). By adapting spatial and temporal correlation scales to local ocean dynamics and assimilating high-resolution SWOT observations, MIOST aims to enhance the effective resolution of mapped products and reduce interpolation errors, particularly in energetic regions. SWOT (Surface Water and Ocean Topography), launched on 16 December 2022 as a joint mission between NASA and CNES with contributions from CSA and UKSA, represents a major advance in satellite altimetry by providing two-dimensional, wide-swath measurements of sea surface height (NASA SWOT Science Team, 2024). The KaRIn instrument enables observations of ocean variability at much shorter spatial scales than conventional nadir altimeters, with mission requirements targeting wavelengths of approximately 15 km and larger, and an effective resolving capability often discussed in the range of ~15–30 km depending on the diagnostic metric (Fu and Ubelmann, 2014; SWOT Science Team, 2024). However, it is important to emphasize that the SWOT MIOST product remains a Level-4

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

mapped field provided on a daily 1/8° grid. Consequently, it should not be interpreted as a direct representation of the native KaRIn sampling, but rather as an enhanced reconstruction that still involves significant spatial smoothing and interpolation (CLS and CNES, 2025). In addition, several limitations must be considered. First, KaRIn measurements are affected by significant noise at small spatial scales, requiring strong filtering prior to mapping, which inevitably removes part of the fine-scale signal. Second, the reconstruction skill depends on the mapping methodology and on the temporal and spatial sampling of the observing system, implying that some of the smallest detected features may be either smoothed or contaminated by noise. Third, the dataset remains relatively recent and its temporal coverage is still limited compared to long-term DUACS products. Finally, recent updates of the SWOT MIOST product (e.g., v2.0.1) include improved KaRIn noise treatment and additional dynamical components such as internal-tide modes, highlighting that the representation of small-scale variability is still an active area of development and subject to methodological uncertainties (CLS and CNES, 2025).

2.2 Global Ocean Reanlysis Configurations

To evaluate the impact of the spatial resolution on reanalysis skills, we used two Global Ocean Reanalysis products, distributed by CMEMS: GLORYS2V4 and GLORYS12V1. These products are part of the CMEMS multi-year reanalysis portfolio and are designed to deliver consistent, observation-constrained reconstructions of the ocean state over the satellite altimetry era (1993–present).

GLORYS12V1 is an eddy-resolving reanalysis, developed by Mercator Ocean International. It provides a global ocean reanalysis at 1/12° horizontal resolution (~8 km), with 50 vertical levels. The system is based on the NEMO ocean general circulation model (v3.1) coupled with the LIM2 sea-ice model and uses Arakawa C-grid discretization, as explained in Madec, G., & the NEMO System Team (2024). Data assimilation follows the SAM2 scheme, a SEEK-based reduced-order Kalman filter, combined with First Guess at Appropriate Time (FGAT), Incremental Analysis Update (IAU), and a 3DVAR bias correction applied to temperature and salinity (Lellouche et al., 2019). The reanalysis assimilates along-track delayed-time sea-level anomaly (SLA) from all altimetric satellites, AVHRR-only Reynolds SST (1/4°), in-situ temperature and salinity profiles from the CORA v4.1 database, and sea-ice concentration from CERSAT (Szekely et al., 2019). The atmospheric forcing is provided by ERA5 from 1 January 2019 (ERA-Interim up to the end of 2018), with 3-hourly and daily fluxes including corrections for precipitation and radiative components. For this study, the period analyzed spans 1st August 2023 to 1st May 2025. For both comparisons, the study requires the use of the interim extension for GLORYS12V1 (cmems_mod_glo_phy_myint_0.083deg_P1D-m): daily mean fields from the interim archive, covering the period from 1st August 2023 to 1st May 2025.

GLORYS2V4 is based on the NEMO v3.1 model framework coupled with the LIM2 thermodynamic–dynamic sea ice model, configured on the ORCA025 grid with a nominal horizontal resolution of 1/4° and 75 vertical levels. The vertical grid is refined in the upper ocean, with layer thickness starting at 1 m to resolve near-surface stratification and progressively increasing to about 200 m in the deepest layers (down to 5900 m). The system is forced at the surface using bulk formulae with atmospheric fields from ERA-Interim. For the present study, the temporal coverage selected for GLORYS2V4 spans from 1st August 2023

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

to 1st May 2025, matching the period used for the satellite datasets and allowing direct comparison with GLORYS12V1. All details on the main parameters of the two reanalysis products are included in Table 1.

Parameters	GLORYS12V1	GLORYS2V4
Model version + grid version	LIM2 EVP NEMO 3.1, driven at the surface by ECMWF ERA Interim reanalysis until end of 2018 and ERA5 reanalysis thereafter + ORCA Grid (1/12°)	LIM2 EVP NEMO 3.1 + ORCA1/4°
Domain	Global	Global
Horizontal resolution	1/12°: 4320 X 2041 grid points	1/4°: 1440 X 1021 grid points
Vertical sampling	50 z-levels	75 z-levels
Numerical technique	Primitive equations with finite differences: on Arakawa C native grid	Same as GLORYS12V1
Time integration	Leapfrog explicit, barotropic/baroclinic splitting; baroclinic step: 300 s	Leapfrog explicit, barotropic/baroclinic splitting; baroclinic step: 450 s
Prognostic variables	3D currents, potential temperature, salinity, turbulent kinetic energy (TKE), sea surface height (SSH)	Same variables as GLORYS12V1
Geophysical variables	Bathymetry from ETOPO2 plus some local smoothing to accommodate large tidal currents	Same as GLORYS12V1
Horizontal diffusion	horizontal Laplacian on trace variables, bi-Laplacian on momentum variables	Same as GLORYS12V1
Surface scheme	ERA-Interim until 2018, ERA5 from 2019, bulk CORE formula with corrections for precipitation and radiative fluxes	No surface nudging precipitation, flux correction Climatological runoff + ice shelf and iceberg melting
Turbolent mixing	TKE scheme with mixing length set to 10 m	Same as GLORYS12V1
Assimilated Satellite variables	Reynolds 0.25° AVHRR-only SST, Delayed Time SLA from all altimetric satellites, in situ T/S profiles from Copernicus Marine CORAv4.1 database, CERSAT Sea Ice Concentration	SST, SLA, T/S profiles, SIC

Table 1. Parameters used in GLORYS12V1 and GLORYS2V4. Both the Reanalysis dataset used the same assimilated scheme: SAM2, a large-scale bias correction 7-day assimilation window Merge MDT (observations + model) Reynolds SST, CORA.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

2.3 Verification Period and Domain

As mentioned in 2.2, the analysis was conducted over the period from 1st August 2023 to 1st May 2025. This period corresponds to the availability of daily outputs from both reanalysis systems, and it ensures the overlap with the DUACS 1/8° observational datasets and provides a sufficient window for an eddy statistics evaluation. An evaluation of eddy statistics over the full period of the reanalysis products is left to a future study. Here our focus is to demonstrate the robustness of the approach and assess the uncertainty associated with the DUACS 1/8° product. Before the application of the verification method, all products are previously interpolated to a 1/8° regular grid (DUACS 1/8° and SWOT MIOST Science grids), and high-pass filtered using a Bessel filter with a 500 km cutoff, as described in Mason et al. (2014), to isolate the mesoscale signals. The same parameters are also used for the eddy detection algorithm. The verification domain is defined as the region bounded by 29.87° N – 59.87° N in latitude and 80.13° W – 30.13° W for longitude, to include the Gulf Stream influence region. This region is known for its intense mesoscale activity and high eddy kinetic energy, making it particularly suitable for validating eddy-resolving capabilities of ocean analysis and reanalysis systems.

2.4 Eddy verification methodology

The detection and tracking of mesoscale eddies in this study follows the standard procedure implemented in PyEddyTracker described in Mason et al. (2014), Pegliasco et al. (2020) and adopted in Smith and Fortin (2022) to develop a robust verification methodology. It is based on identifying closed contours in daily fields of Sea Level Anomaly (SLA) or Sea Surface Height (SSH), and using the associated geostrophic velocities to classify the rotation polarity. We have adopted the approach described by Pegliasco et al. (2020), where the detection is performed on SSH (or Absolute Dynamic Topography) fields rather than SLA to reduce the risk of misidentifying Gulf Stream meanders as eddies, an issue commonly encountered when using SLA in strongly sheared regions. A closed contour is “accepted” as an eddy if it satisfies the following criteria:

- It must contain a single maximum (for Anticyclonic) or minimum (for Cyclonic),
- The number of enclosed grid cells must be between 5 and 2000, corresponding to an effective radius of at least 25 km.
- The shape must be roughly circular, with a contour area not deviating more than 55% from a perfect circle.
- Pixel count 5 pixels = Imin ≤ I ≤ Imax = 2000 pixels: have between 5 and 2000 pixels.
- The amplitude, computed as the difference between the contour extremum and the outermost valid contour, must be at least twice the contour interval (here, 0.2 cm contour interval implies a minimum of 0.4 cm amplitude).

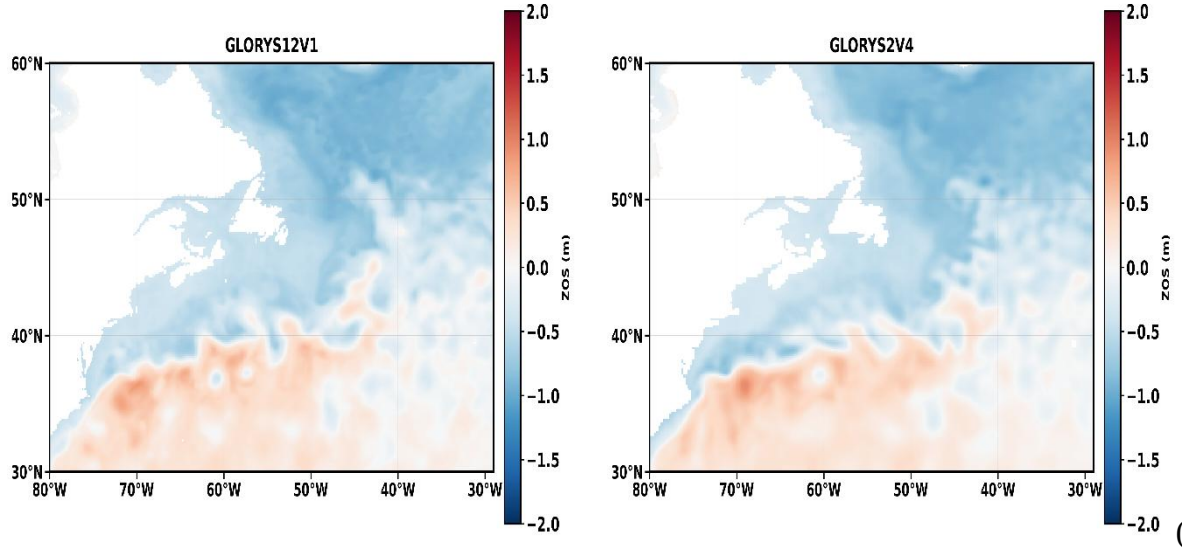
Following the detection, eddies are tracked in time using cost function C, which links eddies on consecutive days based on similarity in amplitude, radius, and position. For a particular eddy, all eddies within a radius of 125 km are identified. The cost function, adopted for each pair of eddies, is based on the difference in amplitude A, in effective radius R and distance D, and it is defined as:

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

$$C = \sqrt{\frac{(A_1 - A_2)^2}{A_1^2} + \frac{(R_1 - R_2)^2}{R_1^2} + \frac{(D)^2}{(125\ km)^2}} \quad (1)$$

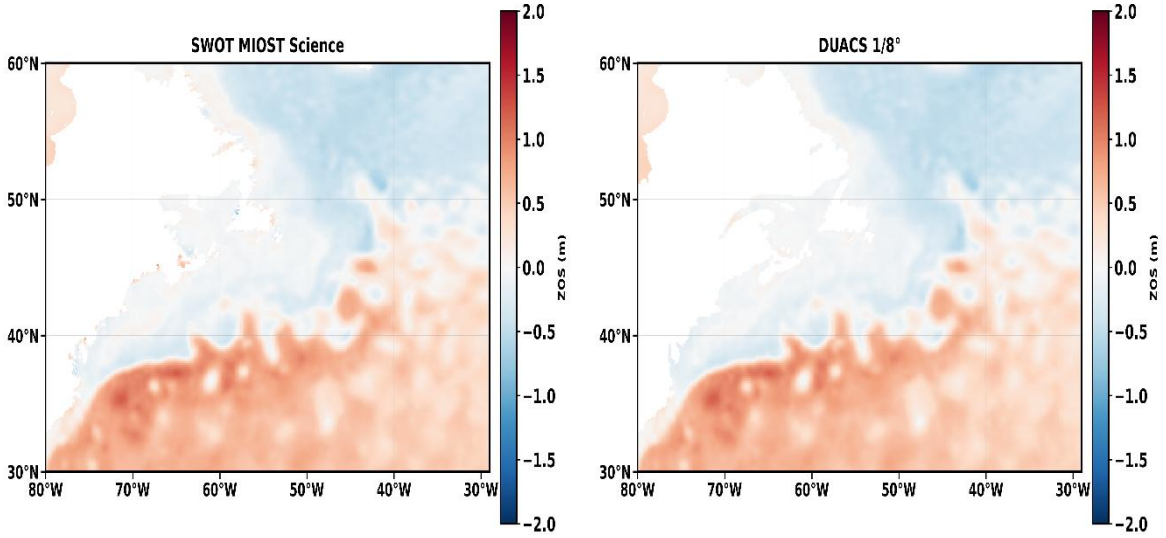
Where the subscripts 1 and 2 are referred to, the two eddies being considered.

A minimum lifetime threshold prevents the identification of spurious eddies. Some studies have employed a minimum eddy lifetime of 28 days as described in Mason et al (2014). Here we have adopted a short lifetime threshold: 4 days, assuming that there are no observational gaps for each eddy between two consecutive days. A lifetime of 7 days is also used for certain evaluations to demonstrate the sensitivity to this choice. The main impact of the use of a large minimum eddy lifetime threshold, like in Mason et al. (2014), is the reduction of the number of ‘‘small” eddies (amplitude less than 10 cm and radius less than 50 km) and spurious features. Eddy properties such as effective radius, amplitude, trajectory, and lifetime are computed for each track. The effective radius is defined as the radius of a circle with the same area enclosed by the eddy contour. The amplitude is determined as the difference between the eddy maxima/minima SSH and the SSH of the last contour enclosing the eddy.



(a) (b)

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

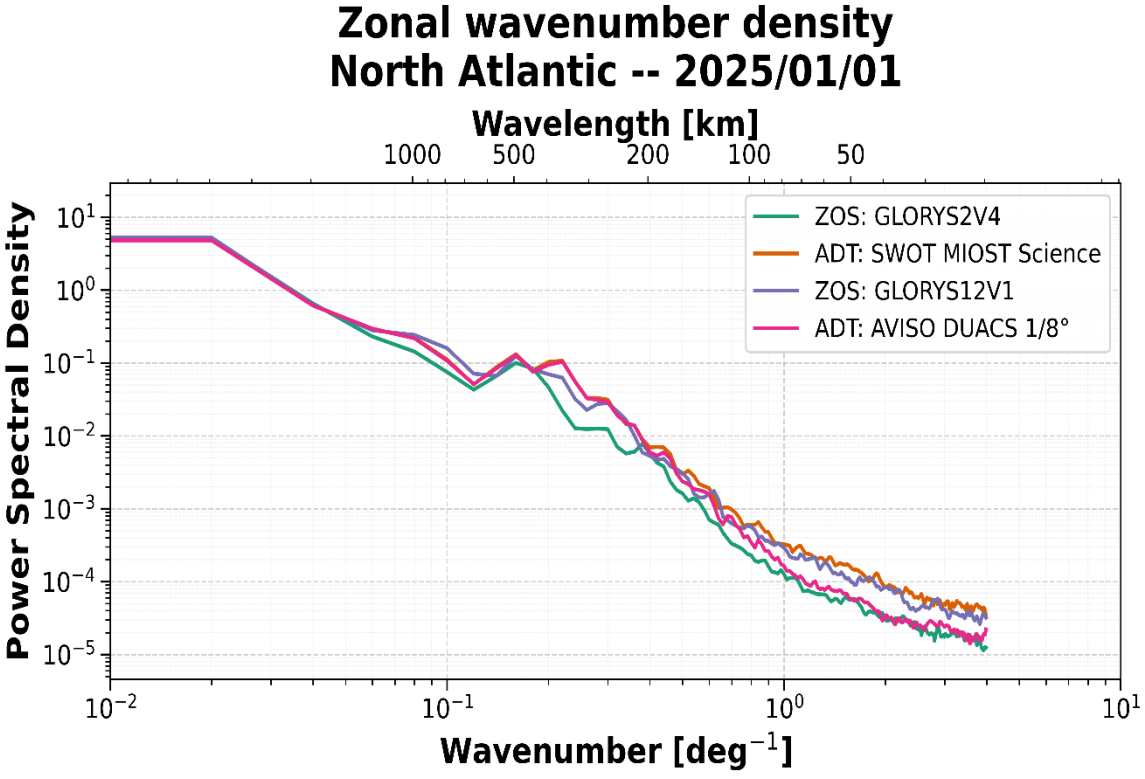


(c) (d)

Fig.1. Daily ADT (or SSH) fields Bessel filtered (500 km) on 1 January 2025 for the four datasets considered in this study: (a) GLORYS12V1 reanalysis, (b) GLORYS2V4, (c) SWOT MIOST Science and (d) AVISO DUACS 1/8°. All fields have been re-mapped onto the same regular 1/8° grid to allow direct comparison. The figure shows the different levels of mesoscale variability captured by each product over the selected North Atlantic region.

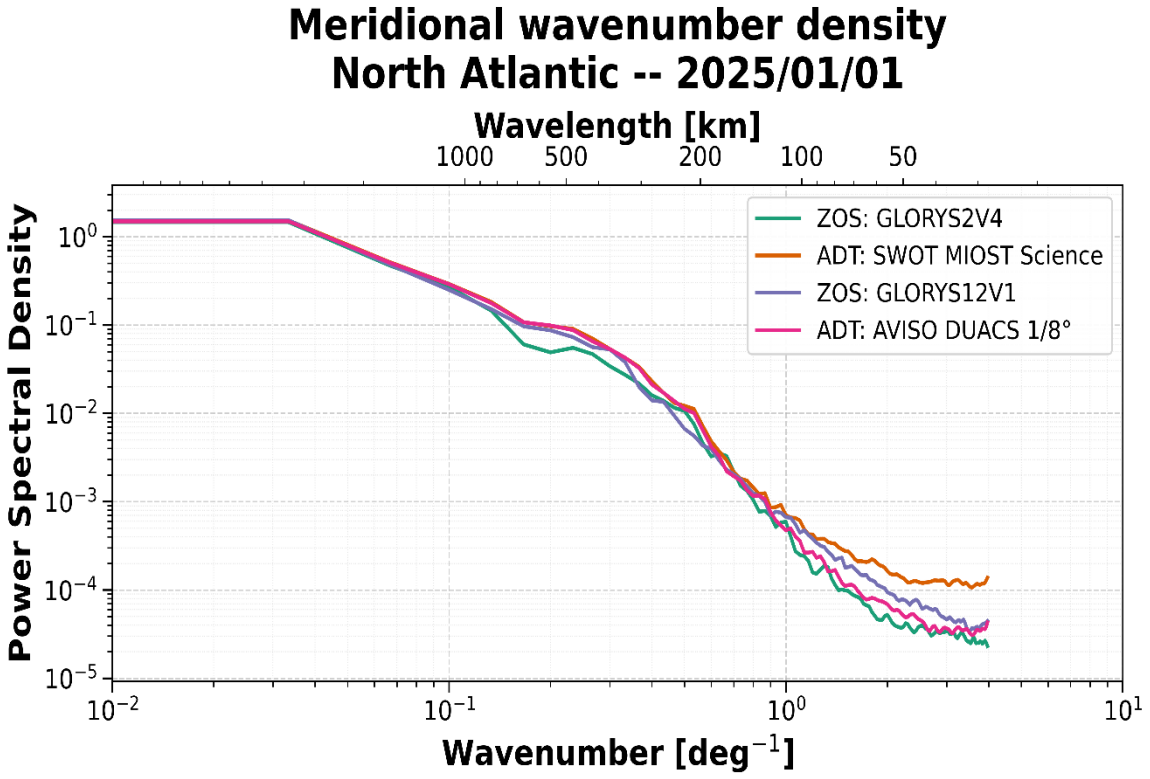
This consistent methodology for all datasets ensures that the detected eddies are comparable and that any differences in eddy representation between reanalysis or between them and observations are due to the intrinsic quality of the reanalysis systems and not to processing artifacts (Ciani et al., 2019; Droghei et al., 2018). First, an analysis of the power spectral analysis for the four products considered in this study, using 1st January 2025 as a representative test case. All datasets were remapped onto the same regular 1/8° grid, which is the native grid of both SWOT MIOST Science and AVISO DUACS 1/8°. In Fig.1 is shown the ADT or SSH fields from AVISO DUACS 1/8° (a), the two reanalysis datasets (b), (c) and SWOT MIOST (d). In analogy to the approach adopted in Droghei et al. (2018) and Ciani et al. (2019), we selected this region for the analysis (e.g. North-West Atlantic) and computed the one-dimensional wavenumber spectra through fast Fourier transform (FFT) to quantify the scales effectively resolved by each dataset.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026



(a)

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026



(b)

Fig.2. Power spectral density (PSD) of ADT (ZOS) for the four datasets on 1 January 2025, computed over the selected North Atlantic region. All products were re-mapped onto the same 1/8° regular grid. Panel (a) shows the PSD along the zonal (ky) direction, and panel (b) shows the PSD along the meridional (kx) direction. The comparison includes SWOT MIOST Science (orange), AVISO DUACS 1/8° (pink), GLORYS2V4 (green), and GLORYS12V1 (blue). At large scales, the four spectra are similar, indicating consistent representation of mesoscale patterns. At smaller scales (≈50–100 km), SWOT MIOST Science shows the highest spectral energy, greater than both reanalyses and AVISO, especially in the meridional direction, suggesting a better ability to preserve fine-scale variability. No evidence of aliasing is visible in the MIOST spectra, which decreases smoothly at high wavenumbers.

Fig. 2 reports the comparative spectral analysis of the four SSH fields, on longitudinal and latitudinal directions, respectively. At large spatial scales (> 200 km), all spectra exhibit comparable energy levels, confirming a consistent representation of the dominant mesoscale structures. Progressively approaching smaller scales, approximately below 100 km (wavenumber $\gtrsim 1$ per degree), the SWOT MIOST spectrum separates from the AVISO DUACS one and follows more closely the spectral slope of the higher-resolution GLORYS12V1 reanalysis. At scales below ≈50 km (wavenumbers > 2–3 deg^{-1}), all products show a rapid loss of spectral energy, with GLORYS2V4 decreasing the fastest, while SWOT MIOST maintains the highest energy level, also greater than GLORYS12V1. The difference is even clearer in the meridional direction (ky). Here, the SWOT MIOST spectrum is higher than all other products across most wavenumbers. This suggests that SWOT MIOST is better at representing small-scale structures, especially in the north–south latitudinal direction. Secondly, as an example for the same day, in Fig. 3, we can see the identification of anticyclonic and cyclonic eddies for 1 January 2025 using the eddy identification

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

and tracking technique described before, for the four different products. Larger eddies are concentrated along the path of the Gulf Stream.

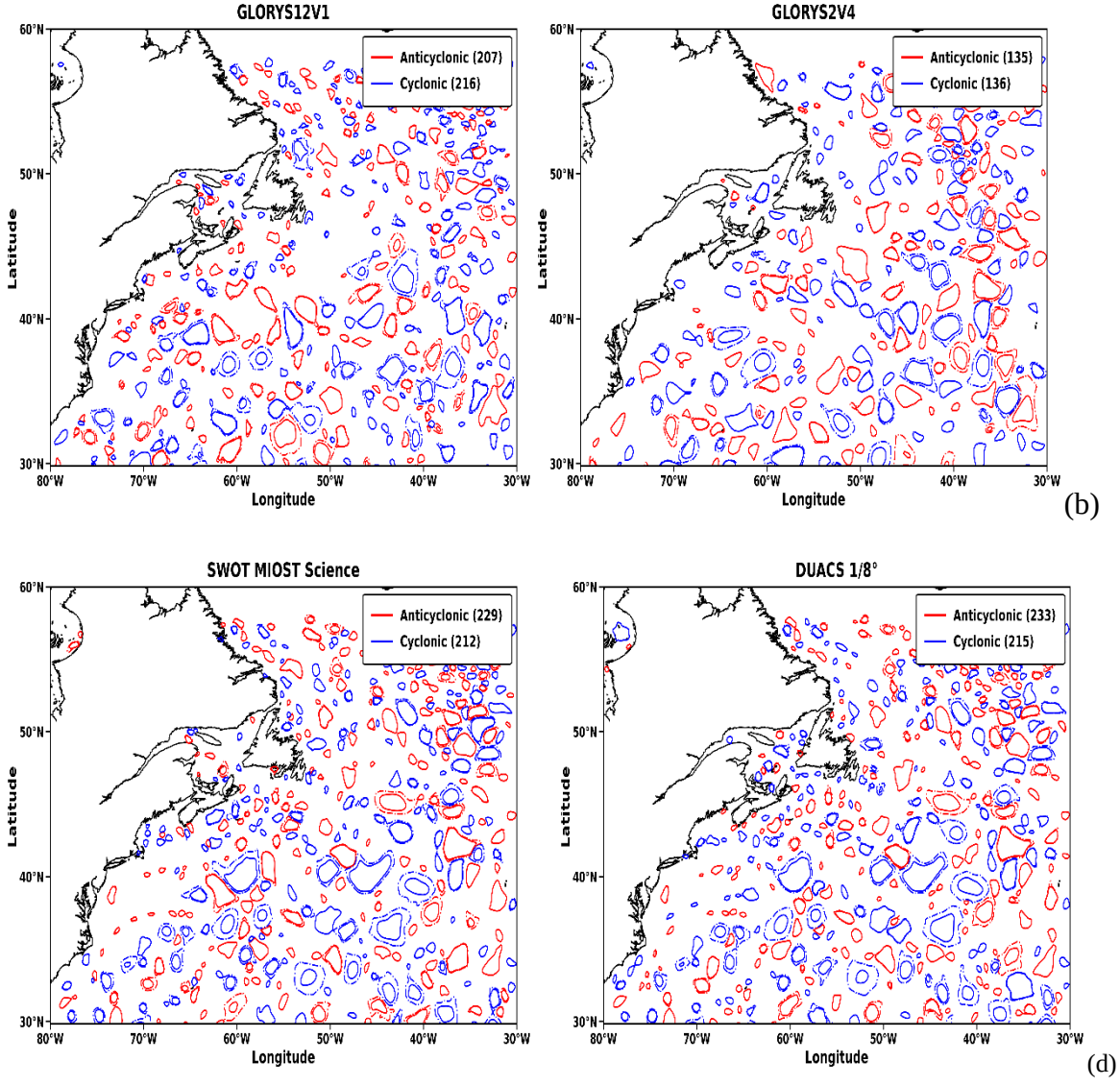


Fig.3. Eddy identification in North-Atlantic area, after a Bessel filtering of SSH field (cm) with a cutoff of 500 km, valid for 1st January 2025, for: (a) GLORYS12V1 (re-gridded onto 1/8° regular grid), (b) GLORYS2V4 (re-gridded onto 1/8° regular grid), (c) SWOT MIOST Science, (d) AVISO DUACS 1/8°.

In Fig. 4, we can see the dependence between radius and amplitude, in which we note that eddies display similar properties in terms of effective radius and amplitude. Many small eddies are detected with a radius below 50 km and amplitude less than 5 cm. In Fig. 5, we can see that most eddies have amplitudes lower than 10 cm, and a lifetime lower than 21 days. The minimum

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

radius and amplitude are respectively: 25 km and 0.4 cm. It can be noted, generally, that larger amplitude eddies are also found to have larger radius (Fig. 4). The eddy identification and tracking procedure are applied for each product and produce plots about physical features and timeseries of anticyclonic and cyclonic eddies.

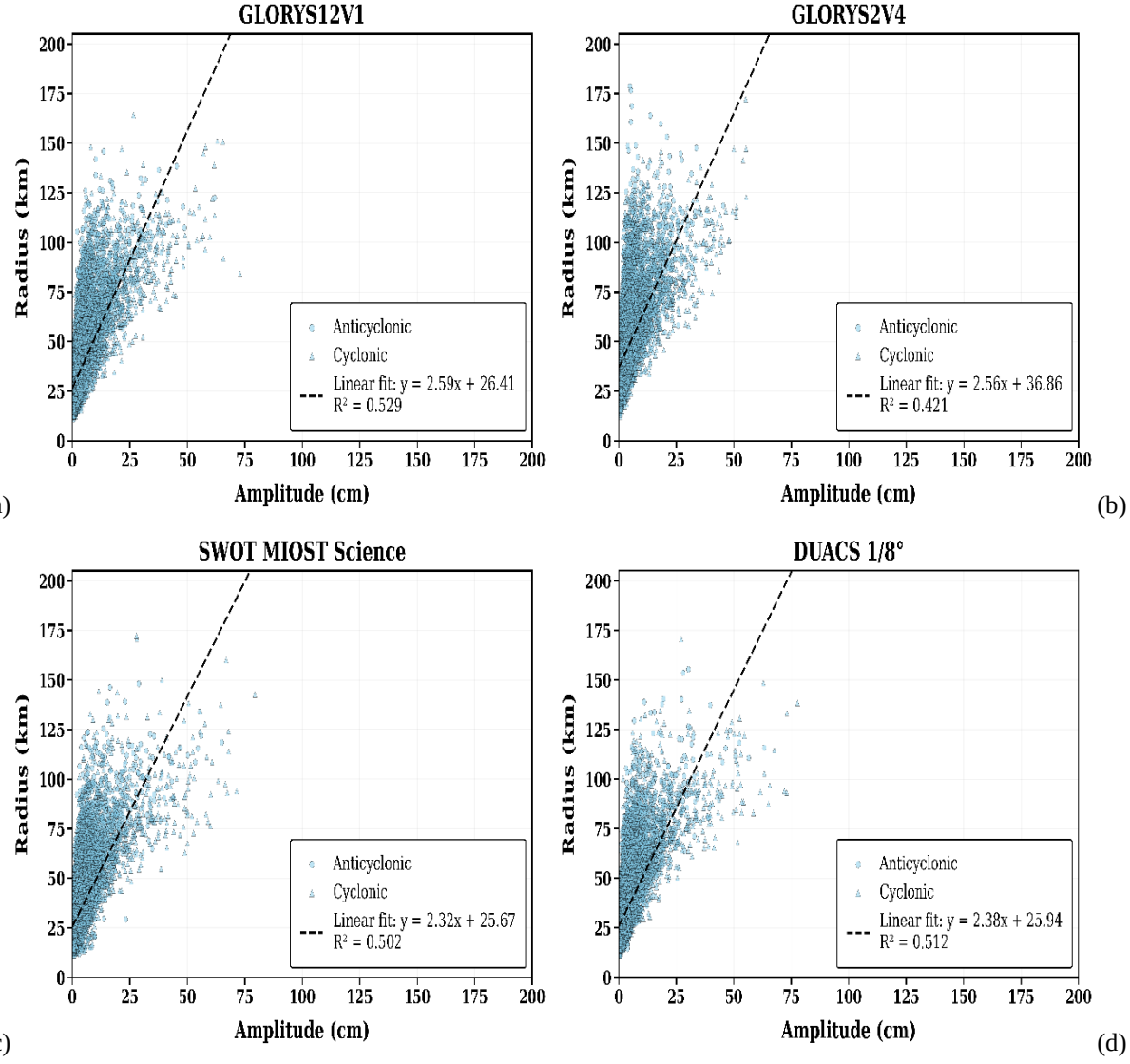


Fig. 4. Scatter plots of detected eddies, with a minimum lifetime threshold of 4 days, between Radius and Amplitude, from 2023-08-01 to 2025-05-01, for (a) GLORYS12V1 (remapped onto AVISO grid), (b) GLORYS2V4 (remapped onto AVISO grid), (c) SWOT MIOST Science, and (d) AVISO DUACS 1/8°. The eddy amplitude and the eddy radius are calculated at the trace level, by averaging the amplitude and radius over the entire lifetime of the eddy.

The same cost function (1) used for the identification and tracking procedure is also used to match the eddies between the reanalysis and the observed eddy (satellite). This function allows us to see how the eddies can be matched even if they have different values for radius, amplitude, mean centroid position, and shape. Although one or more of these parameters can vary,

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

for an eddy to be considered matched to the corresponding satellite dataset, the cost function must be minimized, provided that the distance between the centers of mass does not exceed 125 km (is set to 125 km, i.e., that's slightly greater than the average diameter of eddies in the Northwest Atlantic region; see Ballarota et al., 2019).

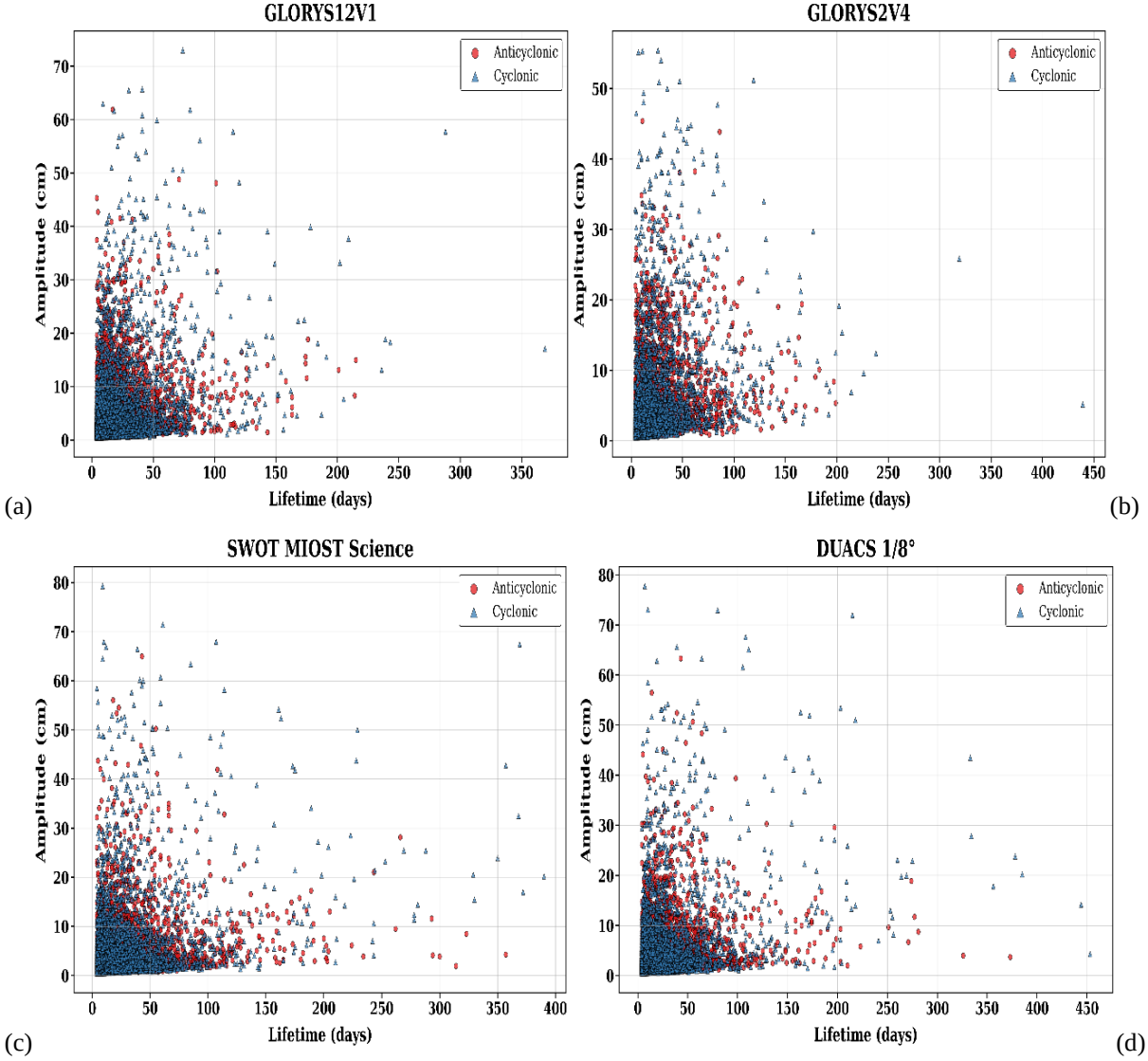


Fig.5. Scatter plots of identified eddies, with a lifetime threshold of 4 days, between Amplitude vs Lifetime of eddies for (a) GLORYS12V1, (b) GLORYS2V4, (c) SWOT MIOST Science, (d) AVISO DUACS 1/8°. Each eddy is counted only once. Most eddies have an amplitude of less than 10 cm. However, their mean lifetime is lower than 50 days. The amplitude is calculated at the trace level, by averaging the amplitude over the entire lifetime of the eddy.

Therefore, following the application of the matching algorithm, an eddy identified in both the reanalysis (i.e. GLORYS2V4 and GLORYS12V1) and verifying datasets (AVISO DUACS and SWOT MIOST products) is referred to as ''hit''. An eddy in the verifying dataset with no corresponding eddy in the reanalysis is referred to as a ''miss", whereas an eddy present in a

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

reanalysis but not in the verifying dataset, is deemed as a ‘’false alarm’’. In Fig. 6(a) and in Fig. 6(b), we see an example of eddy matching procedure for both comparisons. The statistics for hits, misses and false alarms can be tabulated over the study area and over the full evaluation period to provide an indication of the skill of two reanalysis products in representing the
280 “observed eddies”. It is interesting to consider what fraction of the observed eddies are captured (not excluded) by reanalysis products, using the Probability of Detection (POD), defined as:

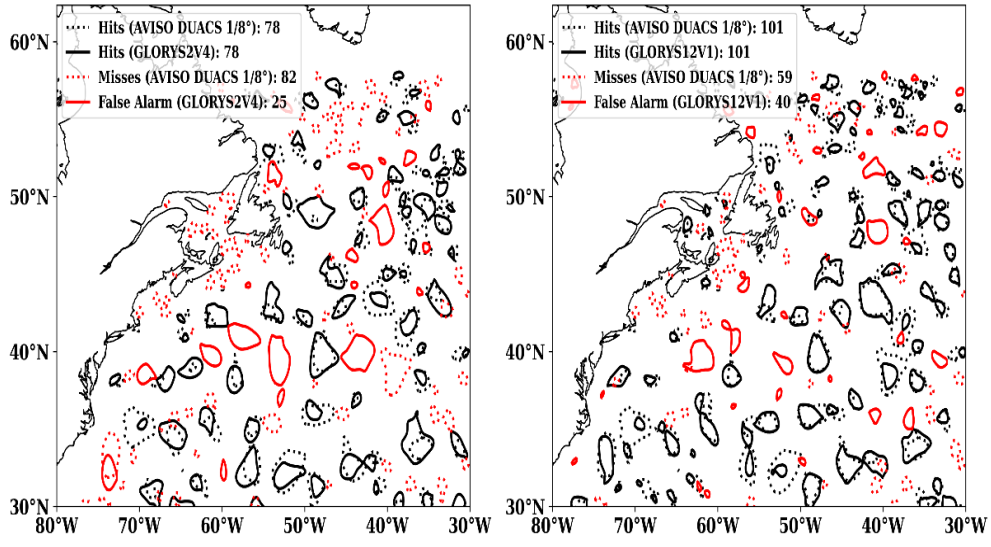
$$POD = \frac{hits}{(hits + misses)} \quad (2)$$

The POD is the number of matched eddies over the total number of eddies present in the satellite product.

Similarly, one can consider what fraction of the predicted eddies in the model analyses did not occur using the False Alarm
285 Ratio (FAR) defined as:

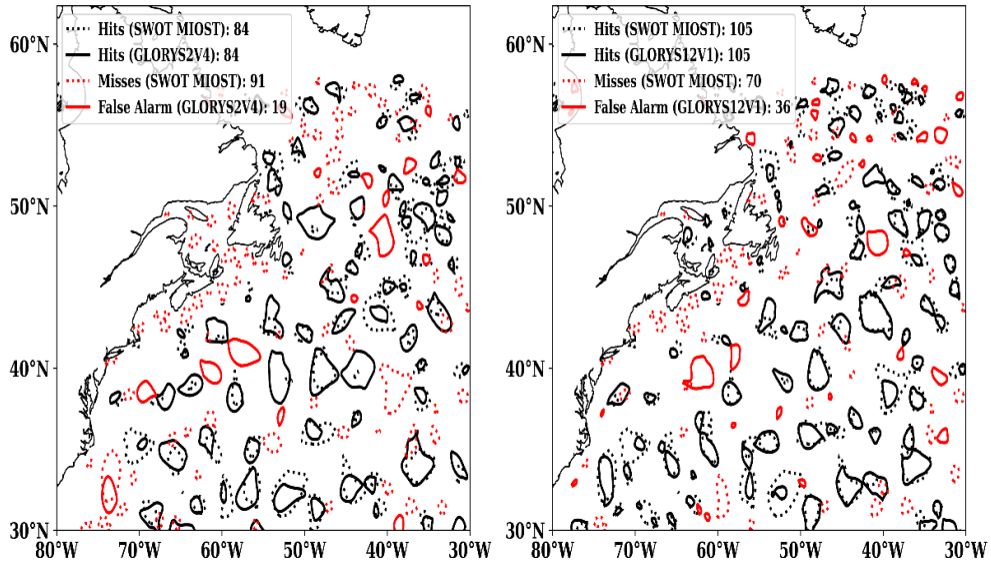
$$FAR = \frac{false\ alarms}{(false\ alarms + hits)} \quad (3)$$

Both POD and FAR have values between [0,1] with perfect values of 1 and 0, respectively.



(a)

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026



(b)

Fig.6. The first one (a) is an example showing Cyclonic eddy matching statistics for 1st August 2023 for GLORYS2V4 (left), and GLORYS12V1 (right), compared with AVISO SSALTO/DUACS 1/8°. Eddies for which a corresponding match is found (hits) between AVISO and GLORYS2V4/GLORYS12V1 are shown as dotted and solid black contours. Eddies identified in AVISO SSALTO/DUACS (satellite data: source for the comparison) with no corresponding eddy in GLORYS2V4/GLORYS12V1 are deemed misses (red dotted contours). Eddies identified in GLORYS2V4/GLORYS12V1 but without a corresponding eddy to AVISO SSALTO/DUACS are deemed false alarms. The second one (b) is an example showing Cyclonic eddy matching statistics for 1st August 2023 for GLORYS2V4 (left), and GLORYS12V1 (right), compared with SWOT MIOST Science.

3.Results

In this section, we present the results from the application of the eddy identification, tracking and matching algorithm described above. First, the results of the eddy identification in the four sets of analyses are presented in Section 3.1. We then examine the matching statistics (hits, false alarms, misses, POD and FAR) in terms of eddy radius and amplitude in Section 3.2. We also evaluate errors in the properties of matched eddies due to the relative contributions to the matching cost function (hits cost, radius, amplitude, and distance between eddy centroids) in Section 3.3.

3.1 Results of eddy identification

The properties of identified eddies for the four datasets are quite similar (see Fig. 4)) over the entire period. In Fig. 4, we see similar trends in the correlation between radius and amplitude. Many small eddies, with a radius lower than 60 km (not far from the threshold of detection) and an amplitude lower than 10 cm, have been identified. In Fig. 5 we see that most have amplitudes below 10 cm, but at the same time there is a positive correlation between amplitude and lifetime, whereby as amplitude increases, so does the average survival time of the eddies. The number of identified eddies is sensitive to the choice of the minimum eddy lifetime threshold, as noted in Section 2.4. The total number of detected eddies is higher for both satellite

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

products as compared to GLORYS2V4 and GLORYS12V1, for both anticyclonic and cyclonic eddies (see Table 4 and Table 5). Consistent with its higher spatial resolution, GLORYS12V1 has a higher number of eddies than GLORYS2V4.

3.2 Eddy verification results

Following the tabulation of hits, misses, and false alarms over the respective evaluation period (1 year and 9 months for both comparisons), the probability of detection (POD) and the false alarms ratio (Eqs. (2) and (3)), whose results are shown in Fig. 7(a) and Fig. 7(b) and Table 4 for comparison with AVISO DUACS 1/8°, and in Fig. 7(c) and Fig. 7(d) and Table 5 for SWOT MIOST comparison. The statistics are aggregated, meaning that they are evaluated by adding up all the counts of hits, misses, and false alarms for the entire verification period, for both cyclonic and anticyclonic eddies, and then calculating the POD and FAR with these accumulated counts.

	Anticyclonic	% w.r.t. AVISO	Cyclonic	% w.r.t. AVISO
DUACS 1/8°	131738 (4 days) 122555 (7 days)	- -	132210 (4 days) 123149 (7 days)	- -
GLORYS2V4	72815 (4 days) 67557 (7 days)	55.27% 55.12%	78737 (4 days) 73596 (7 days)	59.55% 59.76%
GLORYS12V1	104965 (4 days) 94012 (7 days)	79.68% 76.71%	115175 (4 days) 104269 (7 days)	87.11% 84.67%

Table 2: total number of anticyclonic and cyclonic eddies identified in the region of study over the evaluation period for AVISO DUACS 1/8°, GLORYS2V4, and GLORYS12V1. The percentage of eddies identified in GLORYS2V4 and GLORYS12V1 with respect to AVISO is also provided. The numerical value in the first row of each column refers to the 4-day minimum lifetime threshold, while the numerical value in the second row refers to the 7-day minimum lifetime threshold.

	Anticyclonic	% w.r.t. SWOT	Cyclonic	% w.r.t. SWOT
SWOT MIOST	143296 (4 days) 131872 (7 days)	- -	146349 (4 days) 135459 (7 days)	- -
GLORYS2V4	80325 (4 days) 74435 (7 days)	56.05% 56.44%	86695 (4 days) 80899 (7 days)	59.23% 59.72%
GLORYS12V1	115114 (4 days) 103071 (7 days)	80.33% 78.16%	126213 (4 days) 114239 (7 days)	86.24% 84.33%

Table 3: total number of anticyclonic and cyclonic eddies identified in the region of study over the evaluation period for SWOT MIOST, GLORYS2V4, GLORYS12V1. The percentage of eddies identified in GLORYS2V4 and GLORYS12V1 with respect to SWOT MIOST is also provided. The numerical value in the first row of each column refers to the 4-day minimum lifetime threshold, while the numerical value in the second row refers to the 7-day minimum lifetime threshold.

Statistics	GLORYS12V1	GLORYS2V4	Diff.	Diff.%
POD (4 days) (7 days)	0.586 0.617	0.447 0.463	0.139 0.154	31.09% 33.26%
FAR (4 days) (7 days)	0.329 0.234	0.231 0.195	0.098 0.039	42.42% 20.00%

Table 4: aggregated eddy verification statistics for the Probability of Detection (POD) and False Alarm Ratio (FAR) evaluated over the period, from 1st August 2023 to 1st May 2025, for the comparison with AVISO DUACS 1/8°. The numerical value in the first row

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

of each column refers to the 4-day minimum lifetime threshold, while the numerical value in the second row refers to the 7-day minimum lifetime threshold.

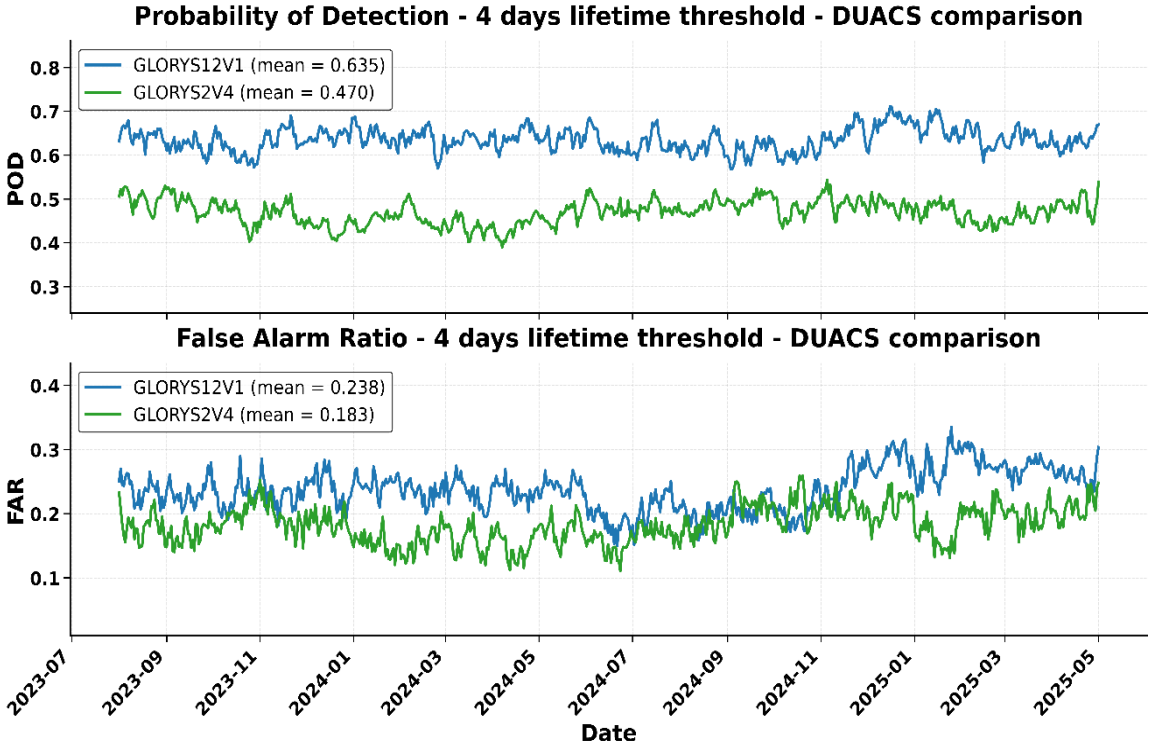
Statistics	GLORYS12V1	GLORYS2V4	Diff.	Diff.%
POD (4 days) (7 days)	0.627 0.618	0.460 0.464	0.167 0.154	36.28% 33.18%
FAR (4 days) (7 days)	0.245 0.235	0.197 0.195	0.048 0.04	24.36% 20.51%

Table 5: aggregated eddy verification statistics for the Probability of Detection (POD) and False Alarm Ratio (FAR) evaluated over the period 1st August 2023 to 1st May 2025, for the comparison with SWOT MIOST. The numerical value in the first row of each column refers to the 4-day minimum lifetime threshold, while the numerical value in the second row refers to the 7-day minimum lifetime threshold.

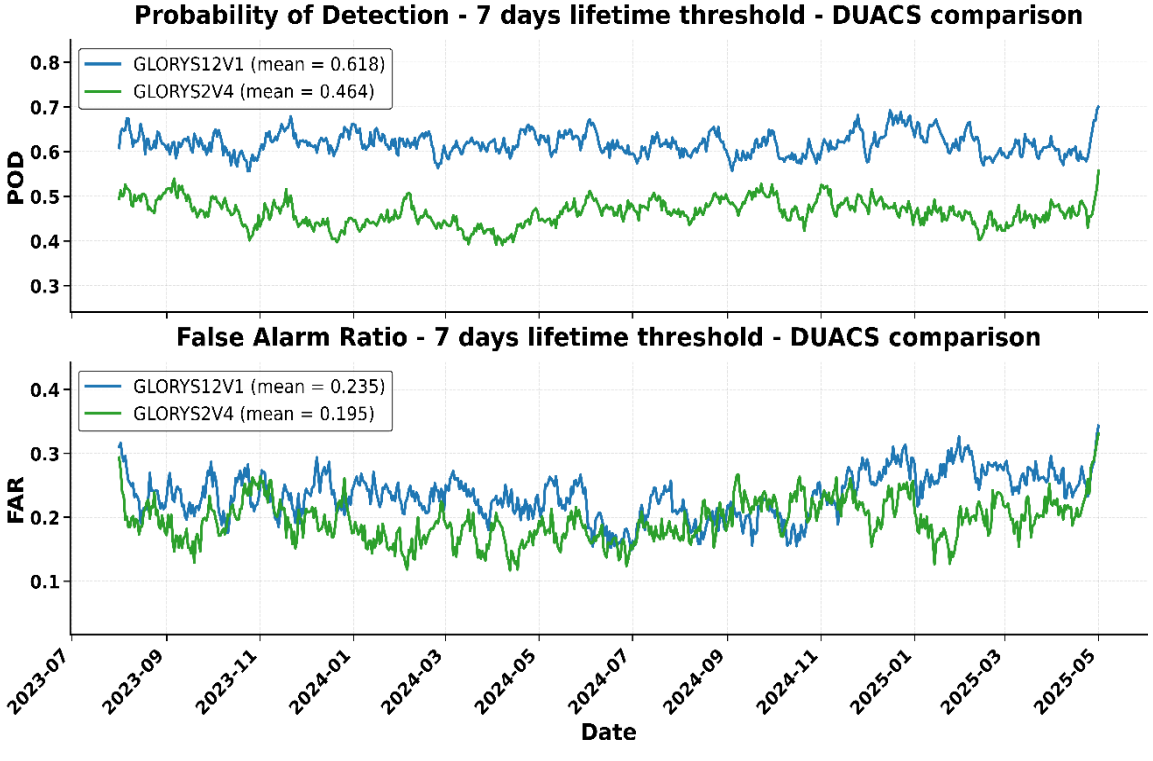
The daily POD and FAR fluctuate but remain stationary over time, with no apparent seasonal cycle. An exception is observed for GLORYS2V4, which exhibits a notable FAR peak in 2024 (between 2024/07 to 2024/09), potentially linked to changes in its data assimilation configuration or observation sampling strategy, as shown in Fig. 7(a) and in Fig. 7(b). As noted in 2.4, the calculation is also applied to the minimum threshold of 7 days without any observational gaps for each eddy within this time interval. In this case, we demonstrate not only how the sensitivity of the metrics changes as the minimum lifetime varies, but also how the verification metrics used, tend to plateau and reduce the differences between the two reanalysis products considered in the study.

On average values, GLORYS12V1 achieves a POD exceeding 50% (higher than 60% for SWOT MIOST comparison) and a FAR lower than 30% for 7-day minimum lifetime thresholds for eddies, except for the comparison with AVISO DUACS 1/8°, with a lifetime threshold of 4 days, in which the FAR is around 32.9%. GLORYS12V1 shows an improvement over GLORYS2V4, that's greater than 30% for POD, but a high value for FAR increasing, that's 42.42% for comparison with lifetime threshold of 4 days, while for 7-days lifetime threshold decreases until the 20% (Table 4). The same verification approach was applied to the SWOT MIOST that exhibits a generally higher mean eddy lifetime, particularly anticyclonic eddies, leading to a different balance of hits, misses, and false alarms (Table 5). GLORYS12V1 tends to align more closely with SWOT MIOST observations than GLORYS2V4, with values higher than 60%, for both lifetime thresholds, with a greater relative gain in POD (+36.3%) and a lower FAR increasing (24.36%%) than the AVISO DUACS comparison, for the minimum lifetime of 4 days, and a relative gain in POD (+33.2%) and FAR (20.5%%) for a minimum lifetime of 7 days (Table 3 and Fig. 6). This suggests that the high-resolution SWOT observations still emphasize mesoscale structures that are more faithfully reproduced by the eddy-resolving GLORYS12V1 system.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

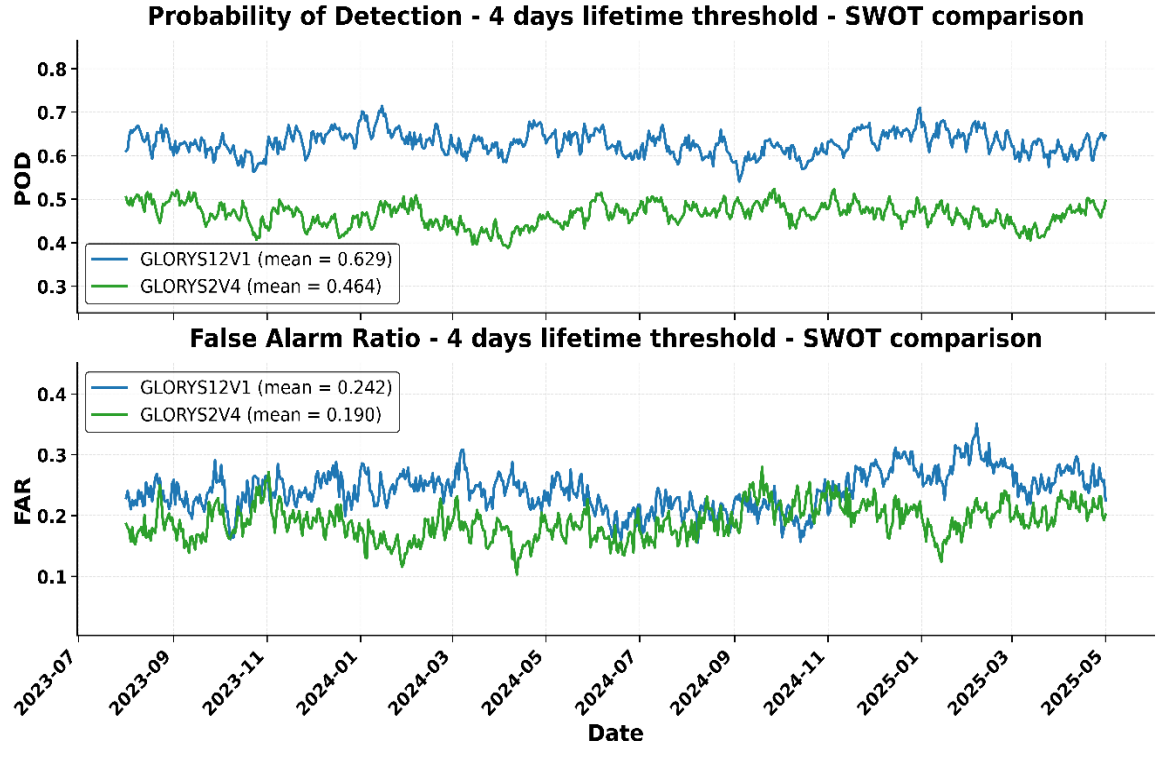


(a)

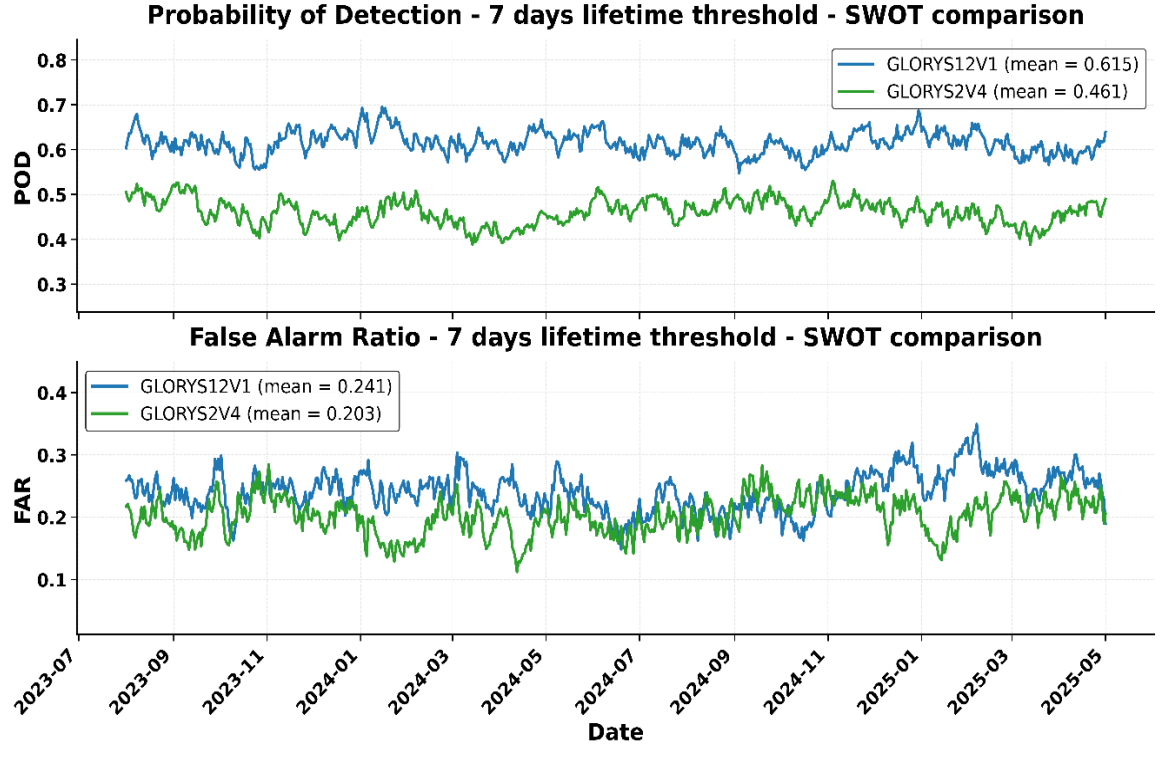


(b)

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026



(c)



(d)

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

Fig.7. Timeseries of matching statistics for GLORYS2V4 and GLORYS12V1, compared with AVISO SSALTO/DUACS, over the period from 1st August 2023 to 1st May 2025, for two different minimum lifetime thresholds: 4 days (a), 7 days (b). In the bottom: timeseries of matching statistics for the same re-analyses compared with SWOT MIOST Science, over the period from 1st August 2023 to 1 May 2025, for the different minimum lifetime thresholds: 4 days (c) and 7 days (d).

This indicates that, specifically with regard to GLORYS2V4, the POD values are identical in both the comparison with AVISO and the comparison with SWOT (this is even more evident when considering eddies with a minimum lifetime of 7 days). This confirms that GLORYS2V4 detection capability is already sufficient for the AVISO validation and that the additional eddies detected in SWOT do not include any eddies that are also detected in GLORYS2V4. Above all, this is evident for a minimum lifetime of 7 days, in which similar POD values are observed even for GLORYS12V1. This suggests that the remaining differences may be linked to the presence of spurious features (small eddies, meanders), and that, therefore, it would be necessary to introduce an effective methodology capable of eliminating these ones. However, the overall POD values, that range between 58-62%, for both comparisons with GLORYS12V1, show that a significant proportion of observed eddies (around 40%) remains misplaced. This highlights the presence of limitations in their high-resolution applications, such as scale lower than mesoscale or trajectory prediction of eddies, where a more complete eddy representation is essential. When using larger values for the minimum eddy lifetime threshold, as noted before, the improvements in POD are no longer present, while reductions in FAR become greater (this means that the small eddies are the most affected for short minimum lifecycle). GLORYS12V1 FAR always stays higher than the GLORYS2V4 FAR. This is likely due to the higher ability of GLORYS12V1 to identify more eddies, including some that later become false alarms because they do not find a match in the satellite datasets. Even so, when moving from AVISO DUACS to SWOT MIOST, FAR values become lower, and POD becomes slightly higher. This means that the comparison with SWOT MIOST gives a better agreement overall, with more matched eddies and fewer false alarms.

To further assess the robustness of these findings, a block bootstrap resampling (95% confidence intervals) was performed using 4-day and 7-day blocks, consistent, respectively, with the imposed minimum lifetime thresholds. Results for both SWOT MIOST Science and AVISO SSALTO/DUACS 1/8° (Tables 6 and 7) confirm that GLORYS12V1 systematically outperforms GLORYS2V4 in terms of POD (≈0.62–0.63 vs ≈0.46–0.47), with a statistically significant difference (~0.15–0.16) that is insensitive to block length. This gain is consistently associated with a moderate increase in FAR (~0.035–0.047), confirming a robust trade-off between detection sensitivity and reliability. The agreement between the two observational references strengthens confidence in the statistical significance of the results, although it does not fully remove uncertainties related to the definition and detectability of small-scale or short-lived eddy structures.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

Metric	GLORYS12V1	GLORYS2V4	Difference (12V1 − 2V4)
POD (Mean)	0.628 0.615	0.468 0.466	0.160 0.149
POD (95% CI)	[0.621, 0.635] [0.606, 0.623]	[0.459, 0.476] [0.454, 0.478]	[0.150, 0.171] [0.136, 0.162]
FAR (Mean)	0.236 0.240	0.190 0.206	0.046 0.035
FAR (95% CI)	[0.233, 0.242] [0.233, 0.248]	[0.184, 0.197] [0.196, 0.215]	[0.040, 0.052] [0.026, 0.044]

Table 6. Bootstrap resampling for SWOT MIOST Science comparison, for a minimum lifetime threshold of 4 days a block length of 4 days (first row), and a minimum lifetime threshold of 7 days and a block length of 7 days (second row).

Metric	GLORYS12V1	GLORYS2V4	Difference (12V1 − 2V4)
POD (Mean)	0.632 0.619	0.472 0.469	0.161 0.149
POD (95% CI)	[0.625, 0.639] [0.611, 0.626]	[0.462, 0.481] [0.456, 0.482]	[0.155, 0.167] [0.138, 0.161]
FAR (Mean)	0.233 0.236	0.186 0.201	0.047 0.035
FAR (95% CI)	[0.230, 0.237] [0.228, 0.246]	[0.181, 0.192] [0.191, 0.212]	[0.040, 0.054] [0.024, 0.046]

Table 7. Bootstrap resampling for DUACS 1/8° comparison, for a minimum lifetime threshold of 4 days days a block length of 4 days (first row), and a minimum lifetime threshold of 7 days and a block length of 7 days (second row).

A final evaluation is also performed, taking into account the trade-off between sensitivity and reliability, using sample lengths of 20 days (equal to the average lifetime of all eddies detected in the various datasets). The results, which have not been included, show numerical values quite similar to those of the previous two block length choices (Table 6 and Table 7). These values already tended to deviate more significantly from the average values calculated and included in Table 3 and Table 4. Therefore, more for the sake of completeness than for any other reason, the calculation has been included only for durations of 4 and 7 days.

In general, it can be observed that, as the minimum lifetime threshold increases, the FAR and POD values of each reanalysis systems, from both satellite comparisons, also tend to converge toward similar levels, suggesting that long-lived eddies are more robust and easier to match across all datasets. This is shown in Fig.8 and Fig.10: dependence of the eddy matching statistics on eddy properties, as effective eddy radius and amplitude. The highest rates of misses and false alarms are associated with small eddies, specifically for those with radii below 50 km and amplitudes below 5 cm. For this population, the POD falls

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

well below 50%, and the FAR often exceeds 30%, reflecting the challenge both reanalysis face in capturing the weakest mesoscale eddies, as shown in Fig.9 and Fig.11.

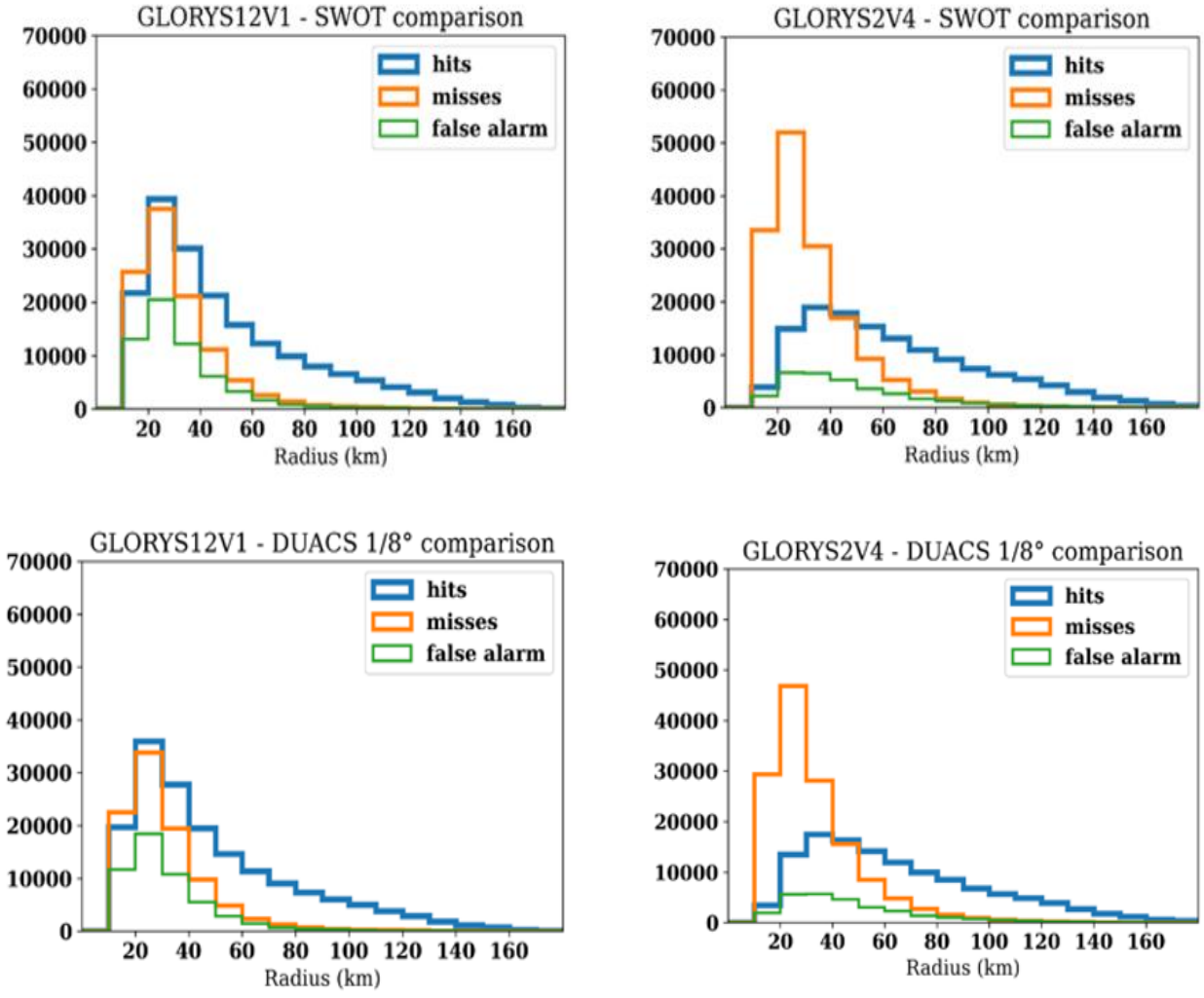


Fig. 8. On the top: eddy matching histograms for the total number of hits (blue), misses (orange), false alarms (green) for GLORYS12V1 (right) and GLORYS2V4 (left), compiled as a function of radius (km) over the evaluation period, for DUACS 1/8° comparison, with a lifetime threshold of 4 days. In the bottom: eddy matching histograms compiled as a function of radius (km) over the evaluation period, for comparison with SWOT MIOST Science. The upper row shows the same histograms of the total number of hits (blue), misses (orange), false alarms (green) for GLORYS2V4 (right) and GLORYS12V1 (left).

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

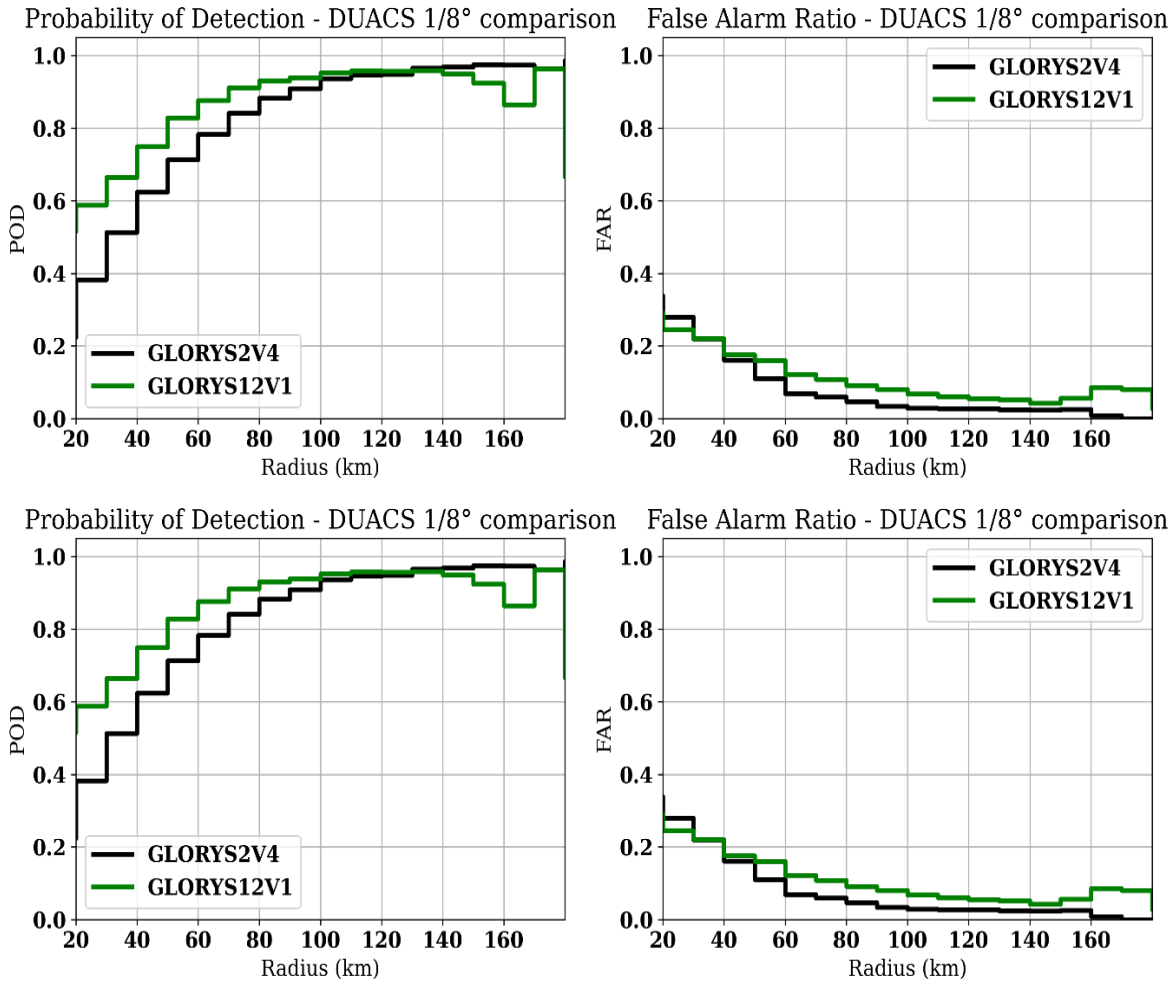


Fig.9. On the top: histograms of the probability of detection (POD; left) and the false alarm ratio (FAR; right), over the evaluation period, evaluated as a function of radius (km) for comparison with DUACS 1/8°, with a lifetime threshold of 4 days. In the bottom: histograms of the POD and FAR as a function of radius (km), for the comparison with SWOT MIOST Science, with a lifetime threshold of 4 days.

For the AVISO DUACS 1/8° comparison, in Fig. 8 and in Fig. 10, false alarms dominate the small radius and small amplitude bins, with FAR exceeding 30%, underscoring the challenge of verifying the smallest features in this gridded product. The FAR remains greater than 20% until 50-55 km for the radius. The comparison with the SWOT MIOST dataset reinforces these findings: both GLORYS12V1 and GLORYS2V4 continue to show they accurately represent eddies with radii below a lower threshold: 30 km. For these smaller eddies, the POD remains below 50% and the FAR exceeds 30%, as showed in Fig. 9. Indeed, the detection skill improves significantly for eddies larger than 30 km, where POD surpasses 50% and FAR drops below 30%, indicating a greater meso-scale matching capability for both reanalyses. For this comparison, GLORYS12V1 shows better consistency with the SWOT product, compared to its performance compared with AVISO. Indeed, differences in

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

POD and FAR remain most pronounced for small-radius and small-amplitude eddies. The largest proportion of misses is consistently associated with small eddies, specifically those with radii below 50 Km and amplitudes under 4 cm. In terms of false alarms, the differences between the two comparison datasets are more pronounced. Conversely, in the SWOT MIOST comparison, false alarms remain lower than both misses and hits across all radii and amplitudes, even at the smallest scales. This suggests SWOT captures the smaller eddies more robustly than the conventional AVISO product, potentially identifying features that AVISO-based comparison under-represents. However, a distinct behavior emerges when considering GLORYS12V1 against the SWOT MIOST product. In this case, the imbalance between misses and hits is significant only for very small eddies (until 3 cm in eddies amplitude), while for eddies exceeding 4–5 cm amplitude, the POD rises above 60%. As noted in Section 2.1, Ballarota et al. (2019), estimated the eddy feature radius in the AVISO gridded product at mid-latitudes to be about 50 Km.

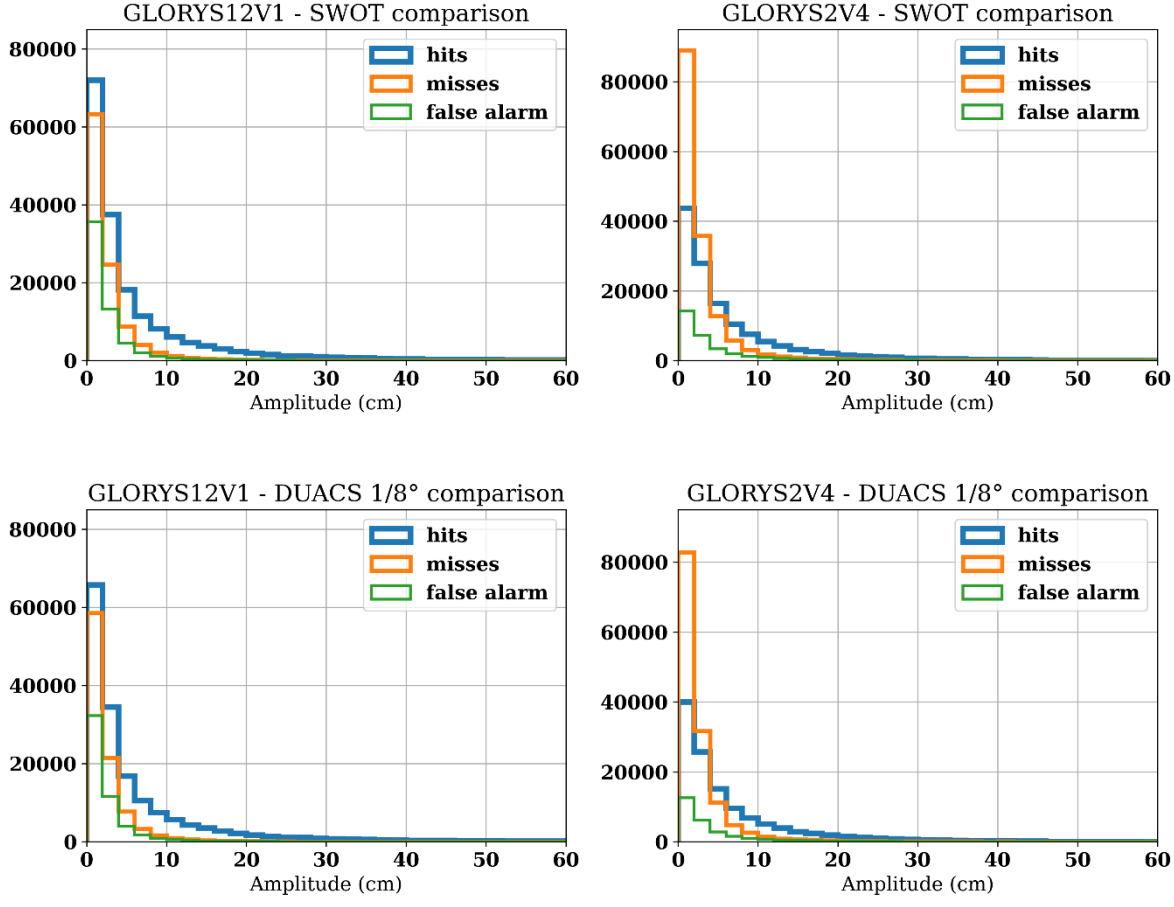


Fig.10. On the top: eddy matching histograms for the total number of hits (blue), misses (orange), false alarms (green) for GLORYS2V4 (right) and GLORYS12V1 (left), compiled as a function of amplitude (cm) over the evaluation period, for DUACS 1/8° comparison. In the bottom: eddy matching histograms compiled as a function of amplitude (cm) over the evaluation period, for comparison with SWOT MIOST Science. The upper row shows the same histograms of the total number of hits (blue), misses (orange), false alarms (green) for GLORYS2V4 (right) and GLORYS12V1 (left).

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

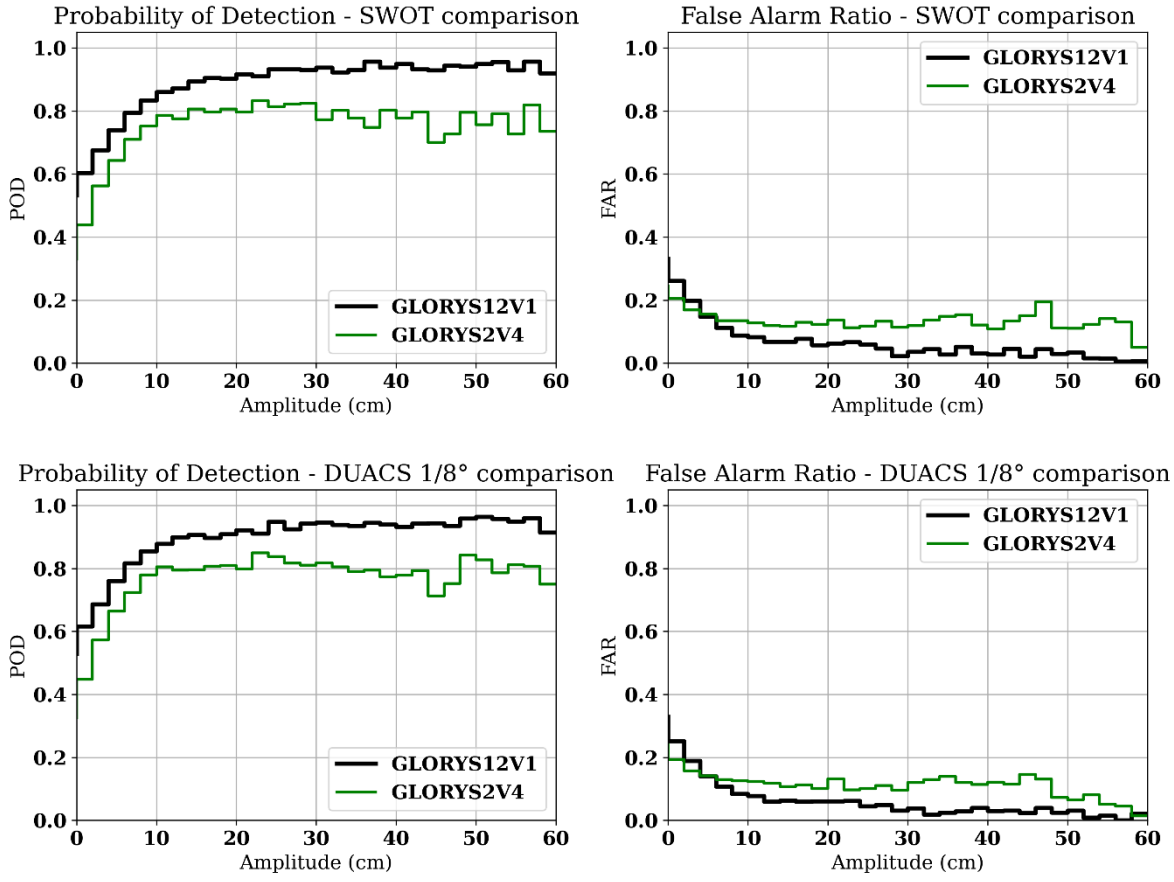


Fig. 11. On the top: histograms of the probability of detection (POD; left) and the false alarm ratio (FAR; right), over the evaluation period, evaluated as a function of amplitude (cm) for comparison with DUACS 1/8°. In the bottom: histograms of the POD and FAR as a function of amplitude (cm), over the evaluation period, for the comparison with SWOT MIOST Science.

Therefore, it can be concluded that many of these small eddies are likely to be spurious features (mesoscale features) that do not provide useful information regarding the quality of the analyses. It should be noted that the number of small eddies is quite sensitive to the minimum lifetime threshold. The presence of such features could stem from factors such as the incomplete along-track nadir altimeter coverage in the AVISO gridded analyses, spurious features linked to the covariance structures in the SAM2 data assimilation scheme (Lellouche et al., 2018) used in GLORYS12V1 and GLORYS2V4, enhanced variability in the Gulf Stream, where standard deviations in SSH can exceed 0.2 m in GLORYS2V4, the global SSH rise observed in GLORYS2V4 since 1993, tracking a realistic trend of ~3 mm/year but accompanied by a growing standard deviation of 0.03 m, increasing with this trend, in the post-2020 (Copernicus Marine Service, Bourdallé-Badie R. & Drévillon M., 2024).

For eddies larger than these limits, both comparisons show strong agreement. For radii > 100 km, the POD reaches ~90% and FAR falls to around 10–15%. In the SWOT-based comparison, the performance is even better: FAR drops below 10%, while POD rises above 80–95% for amplitudes greater than 10 cm. The differences between the two reanalyses become clearer for

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

intermediate scales. GLORYS12V1 shows a 15–20% higher POD than GLORYS2V4 for amplitudes above 20 cm, while FAR increases only slightly (5–10%). These improvements are more visible in SWOT one, likely because it captures energetic eddies more effectively. For AVISO-based comparison, improvements appear mainly for radii above 80 km, while with SWOT they also appear at smaller ranges (40–80 km), where GLORYS12V1 identifies more eddies than GLORYS2V4.

An additional perspective on the differences between GLORYS12V1 and GLORYS2V4 can be obtained by examining the integrated matching statistics as a function of eddy radius and amplitude (Fig. 12 and Fig. 13). These curves show how hits, misses, and false alarms are distributed, and they help identify the size range where most errors occur. This cumulative analysis allows us to answer questions such as: Below what radius do 90% of misses? What percentage of hits occur above 10 cm amplitude? Below radius do 90% of hits occur? The first question provides insight into the radius threshold for filtering out spurious eddy detections, while, the second question, quantifies the system skill in detecting energetic eddies, beyond the "noise" of weaker eddies. The third question allows us to answer, up to what surface extent of eddies do most of the shortcomings of reanalysis systems occur?

For the AVISO comparison, from Fig. 11, we can see that:

- 90% of false alarms are concentrated below 52.55 Km radius for GLORYS12V1 and 83.70 Km for GLORYS2V4, yielding a 31.15 Km gain for the higher-resolution GLORYS12V1 in reaching the same performance level. Regarding Amplitude, 71.6% of GLORYS12V1 hits and 68.8% of GLORYS2V4 hits occur below 10 cm.

So, following the equation (4):

$$N_{hits\%} = (GLORYS12V1_{hits\%} - GLORYS2V4_{hits\%})/GLORYS12V1_{hits\%} \quad (4)$$

We can estimate that the 28.4% of GLORYS12V1 hits are above 10 cm versus 31.2% for GLORYS2V4, meaning GLORYS12V1 detects 9.85% more energetic eddies (amplitude > 10 cm).

For the SWOT comparison, following the same Eq. (4), and using the results of Fig.12, we can estimate:

90% of false alarms fall below 53.10 km radius for GLORYS12V1 and 83.25 km for GLORYS2V4, representing a 30.15 Km gain for GLORYS12V1 at the same threshold. In terms of amplitude, 79.05% of GLORYS2V4 hits and 81.31% of GLORYS12V1 hits lie below 10 cm.

Thus, from Eq. (4), the 18,69% of GLORYS12V1 hits are above 10 cm versus 20,95% for GLORYS2V4, showing a -12.09% for GLORYS12V1.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

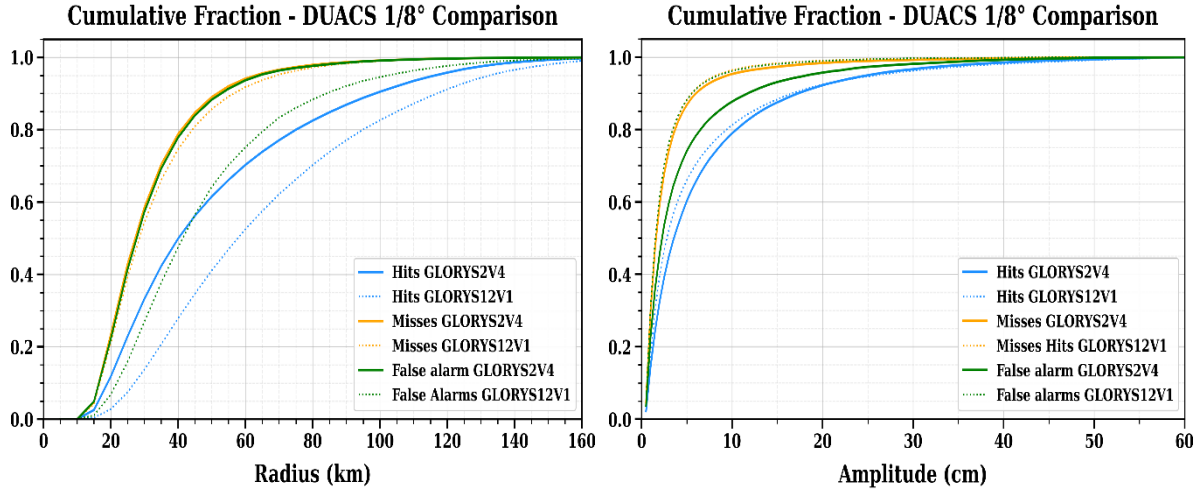


Fig. 12. Cumulative distribution of hits (blue), misses (orange) and false alarms (green) for GLORYS2V4 (solid) and for GLORYS12V1 (dotted) as a function of radius (left) and amplitude (right), for the comparison with AVISO DUACS 1/8°, with a lifetime threshold of 4 days.

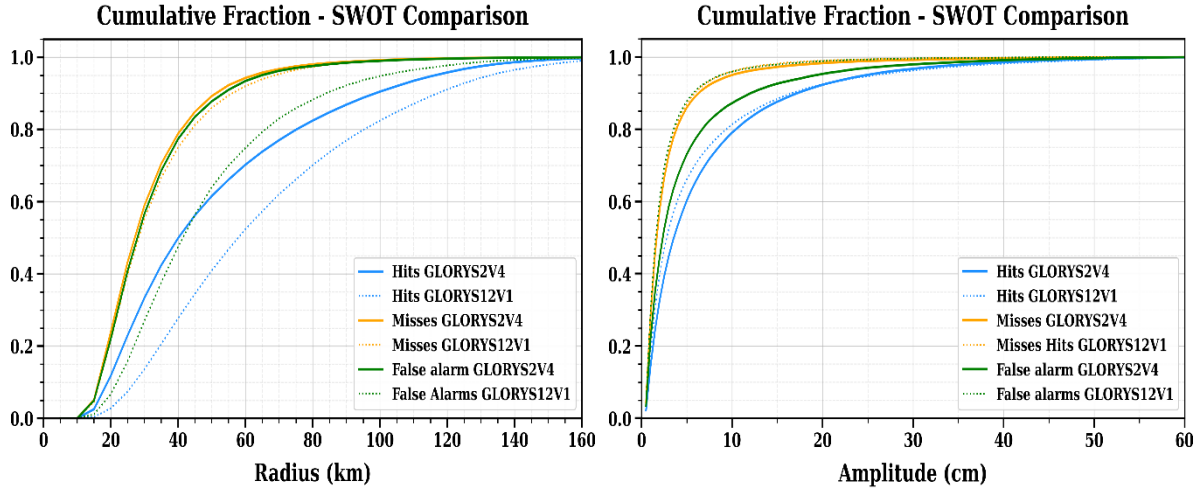


Fig.13. Cumulative distribution of hits (blue), misses (orange) and false alarms (green) for GLORYS2V4 (solid) and for GLORYS12V1 (dotted) as a function of radius (left) and amplitude (right), for the comparison with SWOT MIOST Science, with a lifetime threshold of 4 days.

A clear pattern emerges from comparing AVISO with SWOT: the SWOT evaluation is more consistent, with fewer false alarms at all scales and a higher fraction of matched eddies. This shows that SWOT captures small-scale structures better than AVISO, making it a more reliable reference. However, it also confirms a key weakness of GLORYS12V1: although it detects more eddies overall, it also produces more unmatched eddies (small eddies), which increases the false alarms rate. These cumulative plots show that both reanalysis systems still struggle to match very small eddies (radius < 50 km and amplitude < 4–5 cm), and that many of these features are likely not true eddies but artefacts in either the satellite products or the model.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

This suggests that future verification studies should apply filters to remove or down-weight small eddies, so that evaluation focuses on dynamically meaningful ocean structures.

3.3 Evaluation of errors in eddy properties

Determining whether an observed eddy has a corresponding match in a reanalysis dataset is important, but it does not fully capture how accurately that eddy is represented. As illustrated in Fig. 5, and explained in Section 2.4, some matched eddies exhibit significant differences in terms of radius, amplitude, shape, position, compared to the satellite data, despite correctly identifying the same features. These discrepancies are expected to vary between the two reanalysis systems given their differing spatial resolutions, GLORYS12V1 (1/12°), while GLORYS2V4 (1/4°), and thus their capacity to resolve mesoscale variability. So, we try to answer the question about the possibility to detect differences in the properties of matched eddies between the two reanalysis systems used in this study. To quantify these differences, we analyze the errors in radius, amplitude, and centre displacements for each "matched" eddy (hit). For each of these properties, we compute binned, normalized histograms that illustrate the distribution of radius error (difference between reanalysis and satellite effective radius), amplitude error (difference in SSH extremum relative to the outer contour), separation distance between eddies centroids.

Unlike the tracking cost function (Eq. 1), these quantities are presented without the normalization factors (i.e., not scaled by A_1, R_1, or the 125 km distance), to retain their physical units (km or cm). While this approach no longer allows a direct assessment of their relative contribution to the overall cost function, it provides a clearer, physically interpretable picture of where and how the reanalyses diverge from satellite data (facilitates a physical interpretation of the histograms). The resulting normalized histograms (normalized by the total number of hits) are shown in Fig. 14(a) and Fig. 14(b). The distributions of the cost function and associated errors for matched eddies reveal only subtle differences between GLORYS2V4 and GLORYS12V1 (see Fig. 14(a) and Fig.14(b)), especially in terms of radius, cost function, and amplitude. While no improvements are noticeable at all for the distance of the centroids for both comparisons (both with AVISO and SWOT MIOST) with a slight deterioration in the case of the comparison with AVISO SSALTO/DUACS. Indeed, from Table 8, the mean cost function values, for DUACS 1/8° comparison, are 1.74 for GLORYS2V4 and 1.60 for GLORYS12V1, indicating a reduction in the cost function for each matched eddy for the eddy-resolving (GLORYS12V1) system. Likewise, the mean errors in effective radius (25.83 km vs. 26.63 km) and centroid distance (56.53 km vs. 57.47 km) are nearly identical, with no statistically significant differences between the two reanalysis systems. While there is a little difference in terms of error of the amplitude.

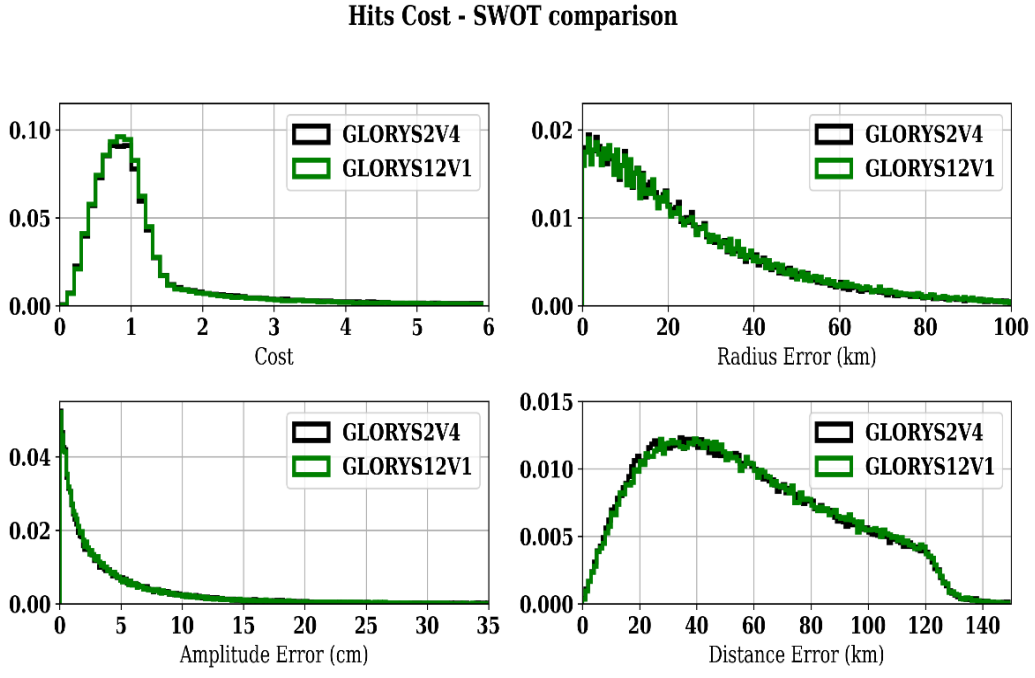
https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

Mean error	Cost	Radius (Km)	Amplitude (cm)	Distance (Km)
GLORYS2V4	1.74	25.83	5.34	56.53
GLORYS12V1	1.60	26.63	5.05	57.47
% diff.	-8.05%	+3.09%	-5.43%	+1.66%

Table 8. Error contributions to the cost function for all hits evaluated over the period 1st August 2023 to 1st May 2025, for DUACS comparison. Mean values for the cost function as well as contributions from radius (km), amplitude (cm), and distance errors (km). The percentage difference between GLORYS2V4 and GLORYS12V1 is provided in the bottom row.

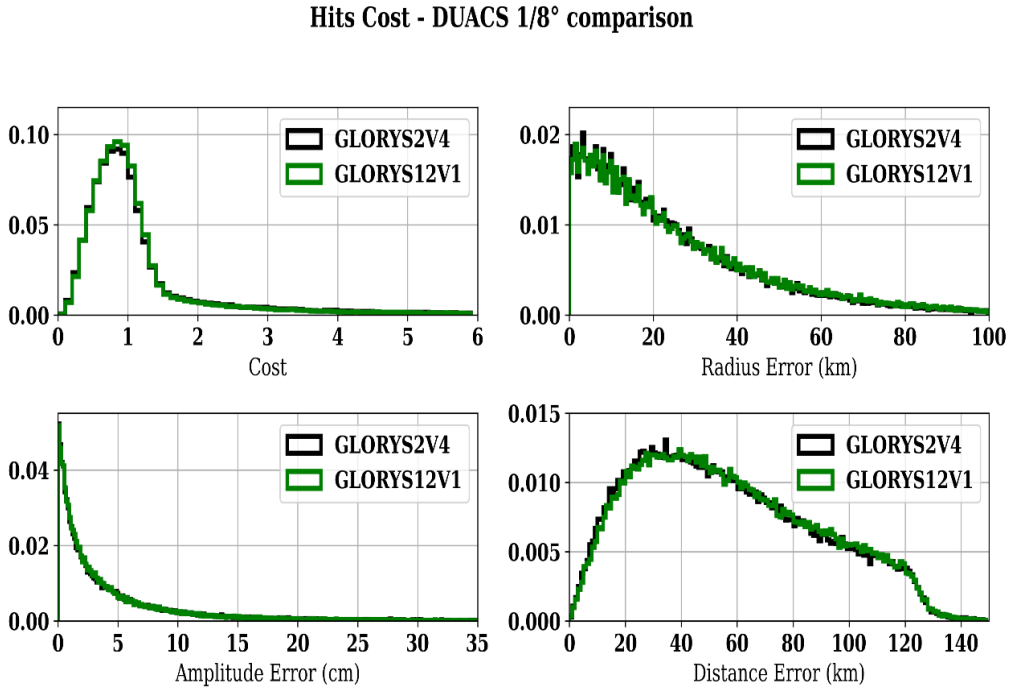
Mean error	Cost	Radius (Km)	Amplitude (cm)	Distance (Km)
GLORYS2V4	1.77	26.09	5.38	56.95
GLORYS12V1	1.60	26.51	5.04	57.65
% diff.	-9.6%	+1.61%	-6.32%	+1,22%

Table 9. Error contributions to the cost function for all hits evaluated over the period 1st August 2023 to 1st May 2025. Mean values for the cost function as well as contributions from radius (km), amplitude (cm), and distance errors (km). The percentage difference between GLORYS2V4 and GLORYS12V1 is provided in the bottom row.



(a)

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026



(b)

Fig. 14. Normalized histograms of values for matched eddies between SWOT MIOST Science and GLORYS2V4 (black line) or GLORYS12V1 (green line), in (a). While in (b), the comparison between DUACS 1/8° and GLORYS2V4 (black line) or GLORYS12V1 (green line). The distribution of values for the total cost is shown in the top panel on the left. Also shown are the normalized distributions of differences in terms of radius in the top on the right, amplitude in the bottom on the right, and distance between eddy positions, in the bottom on the right.

For SWOT comparison (see Table 9), the differences between the two reanalyses become more pronounced than in the AVISO comparison, for cost function and for amplitude. The mean cost function is 1.77 for GLORYS2V4 and 1.60 for GLORYS12V1, corresponding to an 9.6% difference, crossing from GLORYS2V4 to GLORYS12V1. The amplitude error is 5.38 cm versus 5.04 cm (6.32% in difference, crossing from GLORYS2V4 to GLORYS12V1). The effective radius error and the centroid distance error remain almost identical (a difference of –1.61% and -1.22%, respectively). These results indicate that, in terms of cost function and amplitude, GLORYS12V1 gains more advantage when evaluated against SWOT MIOST Science than when compared with the AVISO DUACS product (Table 7). The improvement is more evident in the overall cost function and to a lesser degree in amplitude errors, while distance errors and effective radius showed a lower decrease (especially in terms of distance between centroids). As a result, the advantage of GLORYS12V1 becomes more apparent primarily when using SWOT product, which preserves more mesoscale and sub-mesoscale variability despite the re-gridding in the AVISO DUACS map for both reanalysis systems.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

4. Summary and Discussions

We have applied a feature-based verification methodology for evaluating ocean re-analysis products, following the verification approach developed and adopted by Smith et Fortin (2022), and we have used the eddy detection and tracking procedures described in Mason et al. (2014) and Pegliasco et al. (2020). This methodology is based on prior work in eddy tracking and leverages open-source tools (pyEddyTracker) to provide a robust, reproducible framework for assessing the skill of reanalysis datasets in representing observed mesoscale features. As a first demonstration, we evaluated eddy properties in two global reanalysis systems of different spatial resolutions, GLORYS2V4 and GLORYS12V1, from 1st August 2023 to 1st May 2025 for both comparisons.

The analysis demonstrates statistically significant improvements in the representation of eddies by the higher-resolution GLORYS12V1 system, particularly when compared against SWOT MIOST, whose enhanced observational capability provides a more reliable benchmark for validating mesoscale variability. The verification results show clear differences between the two reanalysis and the satellite products used as references. When compared with AVISO DUACS 1/8°, GLORYS12V1 has a higher Probability of Detection (POD), around 60%, while GLORYS2V4 stays near 45%. This means an improvement of about 31–33% when moving from GLORYS2V4 to GLORYS12V1. However, the False Alarm Ratio (FAR) also increases, with values rising by about 42% for a minimum lifetime of 4 days and by 20% for 7 days. The comparison with SWOT MIOST shows better overall performance. GLORYS12V1 keeps POD values always above 60% for both 4-day and 7-day lifetimes, and the improvement over GLORYS2V4 is about 36% and 33%, respectively. The FAR is around 23–24% for GLORYS12V1, slightly higher than for GLORYS2V4, which stays just below 20%. Small eddies play an important role in these results. With AVISO DUACS, many small eddies are missed, especially those with radii smaller than 50 km. With SWOT, this problem is smaller: POD stays close to or above 50% for most radius classes and drops below 50% only for amplitudes of 4 cm or less. In both satellite datasets, the highest FAR values appear for small or weak eddies (radii below 50 km for AVISO and 40 km for SWOT; amplitudes up to 5–6 cm for AVISO and 4 cm for SWOT).

The results suggest that many small eddies are either noisy or not physically robust, which makes them difficult for the reanalysis systems to match correctly. As a consequence, the number of false alarms increases and the overall skill decreases, even for high-resolution systems such as GLORYS12V1. This highlights the need for a simple and objective approach to limit the influence of very small and weak eddies in both satellite products and reanalyses, in order to obtain more reliable verification metrics. It should be noted, however, that this limitation mainly applies to the AVISO DUACS 1/4° product used in earlier studies. As described in Amores et al. (2018), the minimum resolvable eddy radius in the North Atlantic for the 1/4° AVISO analysis is around 50 km, implying that eddies below this scale are poorly resolved and may be aliased into larger features, leading to an overestimation of eddy counts near the mesoscale threshold. This result cannot be directly extrapolated to higher-resolution products. The AVISO DUACS 1/8° datasets, and in particular the SWOT MIOST Science product, benefit from a finer spatial grid, an increased number of satellite observations, and more advanced interpolation techniques. The Copernicus Marine Service Synthesis Quality Overview Document for Sea Level Products CMEMS-SL-SQO-008-033-068

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

describes how the 1/8° DUACS products are designed to improve the representation of mesoscale variability relative to the 1/4° analysis, while also noting that their effective resolution remains constrained by altimeter sampling and mapping uncertainties. The use of the SWOT MIOST Science dataset as an additional reference is therefore intended to explore whether the inclusion of wide-swath SWOT observations and the Merged Interpolation of Observations with Spatial and Temporal constraints (MIOST) methodology can improve the detection and characterization of smaller eddies. Nevertheless, even with these enhanced products, caution is required when interpreting results close to the effective resolution limit, as very small and weak eddies may still be affected by noise and mapping artefacts.

Despite these issues, the results show the added value of the eddy-resolving GLORYS12V1 reanalysis compared with the eddy-permitting GLORYS2V4 (despite of remapping onto 1/8° regular grid: AVISO DUACS 1/8° and SWOT MIOST Science grid resolution). For both satellite datasets, GLORYS12V1 gives higher POD values, with improvements that are statistically significant at the 95% level, although it also produces a small increase in FAR. With DUACS 1/8°, GLORYS12V1 reaches POD values that are 10–15% higher and FAR values about 10% higher for eddies with amplitudes above 10 cm. With SWOT MIOST Science, the improvement is even larger, with 15–20% higher POD and only a 5% higher FAR for the same amplitude range. False alarms occur mainly for small eddies. With AVISO DUACS, 90% of all false alarms appear for radii smaller than 54.55 km in GLORYS12V1 and 85.70 km in GLORYS2V4. With SWOT, these values are 53.10 km and 83.25 km. These reductions show that GLORYS12V1 better limits false alarms for small radii. However, GLORYS12V1 does not detect a larger fraction of the most energetic eddies. For amplitudes above 10 cm, in GLORYS12V1 only 28.4% of its hits exceed this threshold with AVISO comparison, compared with 30.9% for GLORYS2V4. With SWOT, the values are 18.7% and 20.95%, showing small decreases of –3.3% and –12.05%. Overall, these results confirm that the AVISO SWOT MIOST Science product is a better reference for evaluating both reanalysis products. Its higher resolution preserves more small-scale variability, which helps to better separate the performance of GLORYS12V1 and GLORYS2V4, especially for intermediate and energetic eddies. At the same time, SWOT also highlights the difficulties in verifying very small features, where noise and unresolved variability remain a challenge. Both reanalyses agree better with SWOT than with AVISO DUACS, thanks to SWOT's ability to resolve smaller scales. However, this advantage is partly reduced by the presence of many small eddies, some of which may not be real eddies. These small features increase false alarms and lower POD, and they can influence the verification even after all datasets are remapped onto the 1/8° grid. To address this, future work should develop strategies to filter or downweigh small eddies, ensuring metrics reflect dynamically relevant scales. Moreover, extending the analysis to include the "coincidence status" framework, as explained in Ma et al. (2020), classifying eddies as matched, split, merged, missed, or artifacts, could offer a more nuanced assessment, improving the robustness and diagnostic power of this verification approach.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

Data availability

The dataset used in this study are available from the following sources. DUACS 1/8° global gridded altimetry product is provided by the Copernicus Marine Service and can be accessed at: Global Ocean Gridded L 4 Sea Surface Heights And Derived Variables Reprocessed 1993 Ongoing | Copernicus Marine Service. The SWOT MIOST Science, an experimental high-resolution altimetry gridded product, is available in: TDS Catalog. The Global Ocean Reanalysis products are also provided by Copernicus Marine Service, and they are available respectively in: Global Ocean Physics Reanalysis | Copernicus Marine Service (GLORYS12V1), Global Ocean Ensemble Physics Reanalysis | Copernicus Marine Service (GLORYS2V4).

Code availability

The eddy detection methodology is based on the publicly available Python package py-eddy-tracker (Mason et al., 2014; Faghmous et al., 2015; Pegliasco et al.,2020), available at: https://py-eddy-tracker.readthedocs.io. The analysis scripts and the methodology used in this study were provided by the second author (Smith and Fortin, 2022) and further developed by the first author. These scripts are inspired by and partially derived from the methods implemented in py-eddy-tracker but include additional modifications specific to this study. The custom code developed for the analysis is available from the corresponding author upon request.

Competing interest

At least one of the (co-)authors serves as editor for the special issue to which this paper belongs.

Author Contribution

PM developed the analysis code and implemented the methodology with contributions from GCS, CY and AS. The methodology is built upon initial scripts provided by GCS.

PM performed the data analysis and led the writing of the manuscript. Significant contributions to the interpretation of results and manuscript preparation were provided by AS, GCS and CY. All authors reviewed and approved the final version of the manuscript.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

References

[1] Abernathey, R., & Haller, G. (2018). Transport by Lagrangian vortices in the eastern Pacific. *Journal of Physical Oceanography*, 48, 667–685.

[2] Amores, A., Jordà, G., Arsouze, T., & Le Sommer, J. (2018). Up to what extent can we characterize ocean eddies using present-day gridded altimetric products? *Journal of Geophysical Research: Oceans*, 123, 7220–7236.

[3] **Ballarotta, M., et al. (2019). DUACS DT2018: Improved altimeter sea level products for monitoring mesoscale ocean variability. *Ocean Science*, 15, 1091–1109. https://doi.org/10.5194/os-15-1091-2019.**

[4] Ciani, D.; Santoleri, R.; Liberti, G.L.; Prigent, C.; Donlon, C.; Buongiorno Nardelli, B. Copernicus Imaging Microwave Radiometer (CIMR) Benefits for the Copernicus Level 4 Sea-Surface Salinity Processing Chain. Remote Sens. 2019, 11, 1818. https://doi.org/10.3390/rs11151818.

[5] **Ciani, D., Fanelli, C., and Buongiorno Nardelli B. "Estimating ocean currents from the joint reconstruction of absolute dynamic topography and sea surface temperature through deep learning algorithms." *Ocean Science* 21.1 (2025): 199-216**.

[6] **Chelton, D. B., Gaube, P., Schlax, M. G., Early, J. J., & Samelson, R. M. (2011a). The influence of nonlinear mesoscale eddies on near-surface oceanic chlorophyll. *Science*, 334(6054), 328–332**.

[7] **Chelton, D. B., Schlax, M. G., & Samelson, R. M. (2011b). Global observations of nonlinear mesoscale eddies. *Progress in Oceanography*, 91(2), 167–216**.

[8] Chenillat, F., Franks, P. J. S., & Combes, V. (2016). Biogeochemical properties of eddies in the California Current System. Geophysical Research Letters, 43, 5812–5820. https://doi.org/10.1002/2016GL069009.

[9] Cullen, V. L. (2005). Columbus Iselin and oceanographic research in the western North Atlantic. Oceanography, 18(1), 30–33.

[10] **CLS, CNES. (2025). DUACS Experimental Products Handbook: SSALTO/DUACS Level-4 Sea Surface Height and Derived Variables Gridded Products – Including SWOT MIOST Science. Version 1.2, April 2025. Available at: https://www.aviso.altimetry.fr.**

[11] Conti, D., Orfila, A., Mason, E., Sayol, J. M., Simarro, G., & Balle, S. (2016). An eddy tracking algorithm based on dynamical systems theory. *Ocean Dynamics*, 66, 1415–1427.

[11] Copernicus Marine Service (2024). CMEMS-GLO-PUM-001-031: Global Ocean Multi-Model Reanalysis Product GLORYS2V4 (GREP).

[12] Copernicus Marine Service (2024). CMEMS-GLO-QUID-001-030: Quality Information Document for Global Ocean Reanalysis Product GLORYS12V1.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

[13] Copernicus Marine Service (2024). CMEMS-GLO-QUID-001-031: Global Ocean Multi-Model Reanalysis Product GLORYS2V4 (GREP).

[14] **Copernicus Marine Service (CMEMS), 2024. *GLOBAL_MULTIYEAR_PHY_ENS_001_031: Global Ocean Reanalysis Ensemble Product (GREP), Quality Information Document (QUID), Issue 1.3.* Mercator Ocean International. Bourdallé-Badie R., et al.**

[15] Copernicus Marine Service, 2025. Product User Manual for Sea Level L4 NRT and MY Products. CMEMS-SL-PUM-008-046-047-060-068, Issue 1.0.

[16] Copernicus Marine Service, 2025. Synthesis Quality Overview Document for Sea Level Products. CMEMS-SL-SQO-008-033-068, Issue 14.0.

[17] Droghei, R., Buongiorno Nardelli, B., and Santoleri, R.: A new global sea surface salinity and density dataset from multivariate observations (1993–2016), Front. Mar. Sci., 5, 84, https://doi.org/10.3389/fmars.2018.00084, 2018. a

[18] EO Portal, 2024. *SWOT (Surface Water and Ocean Topography) Satellite Mission*. Available at: https://www.eoportal.org/satellite-missions/swot (accessed 15 April 2026).

[19] Fouchet, E. et al. (2025). Comparison of a global high-resolution ocean data assimilation system with SWOT observations.

[20] Fu, L.-L., and Ubelmann, C., 2014. *On the transition from profile altimeter to swath altimeter for observing global ocean surface topography*. Journal of Atmospheric and Oceanic Technology

[21] La Traon, P. Y., & Morrow, R. (2001). Ocean currents and eddies. In Satellite altimetry and earth sciences (pp. 171–210). Academic Press. https://doi.org/10.1016/B978-012431551-4/50007-0

[22] Lellouche, J. M., et al. (2013). Evaluation of global monitoring and forecasting systems at Mercator Océan. *Ocean Science*, 9, 57–81.

[23] **Lellouche, J. M., et al. (2018). Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1⁄12° high-resolution system. *Ocean Science*, 14, 1093–1126.**

[24] Leroux, S., Brankart, J.-M., Albert, A., Brodeau, L., Molines, J.-M., Jamet, Q., Le Sommer, J., Penduff, T., & Brasseur, P. (2022). Ensemble quantification of short-term predictability of the ocean dynamics at a kilometric-scale resolution: A Western Mediterranean test case. *Ocean Science*, 18, 1619–1644.

[25] **Ma, C., Guo, X., Zhang, H., Di, J., & Chen, G. (2020). An investigation of the influences of SWOT sampling and errors on ocean eddy observation. *Remote Sensing*, 12(2682).**

[26] Madec, G., & Imbard, M. (1996). A global ocean mesh to overcome the North Pole singularity. *Climate Dynamics*, 12, 381–388.

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

[27] Madec, G., & the NEMO System Team. (2024). NEMO Ocean Engine Reference Manual. Zenodo. https://doi.org/10.5281/zenodo.1464816.

[28] Martin, M. et al. (2024). Data assimilation schemes for ocean forecasting: state of the art.

[29] **Mason, E., Pascual, A., & McWilliams, J. C. (2014). A new sea surface height–based code for oceanic mesoscale eddy tracking.** ***Journal of Atmospheric and Oceanic Technology*****, 31(5), 1181–1188**.

[30] **Mason, E., et al. (2017). Subregional characterization of mesoscale eddies across the Brazil-Malvinas confluence.** ***Journal of Geophysical Research: Oceans*****, 122, 3329–3357**.

[31] **Pegliasco, C., Chaigneau, A., Morrow, R., & Dumas, F. (2020). Detection and tracking of mesoscale eddies in the Mediterranean Sea.** ***Advances in Space Research*****, 66(11), 2480–2500**.

[32] **Pegliasco, C. et al. (2022). META3.1exp: A new global mesoscale eddy trajectory atlas derived from altimetry**.

[33] Rhines, P. B., & Young, W. R. (1982). Homogenization of potential vorticity in planetary gyres. Journal of Fluid Mechanics, 122, 347–367. https://doi.org/10.1017/S0022112082002250.

[34] Richardson, P. L. (1983). Eddy kinetic energy in the North Atlantic from surface drifters. Journal of Geophysical Research: Oceans, 88(C7), 4355–4367. https://doi.org/10.1029/JC088iC07p04355.

[35] Rogé, M., Ubelmann, C., and Le Traon, P.-Y., 2017. Using a dynamical advection constraint to improve the resolution of surface geostrophic currents in altimetry. Journal of Atmospheric and Oceanic Technology, 34(11), 2545–2559.

[36] Sangrà, P. et al. (2009). The Canary Eddy Corridor: A major pathway for long-lived eddies in the subtropical North Atlantic.

[37] **Smith, G. C. & Fortin, A.-S. (2022). Verification of eddy properties in operational oceanographic analysis systems.** ***Ocean Modelling*****, 172, 101982**.

[38] SSALTO/DUACS (2021). Experimental Product Handbook. Issue: 1.8, Date: 09/10/2021, Nomenclature: SALP-MU-P-EA-23172-CLS.

[39] **SSALTO/DUACS (2024). Experimental Product Handbook: Gridded Sea Level Heights and geostrophic velocities computed with SWOT Level-3 products (using both KaRIn and nadir instruments).**

[40] Stephens, C., & Marshall, J. (1999). Dynamical influence of air–sea buoyancy fluxes on the thermohaline circulation. Journal of Physical Oceanography, 29(12), 3042–3056. https://doi.org/10.1175/1520-0485(1999)029<3042:DIOASB>2.0.CO;2.

[41] **SWOT Science Team, 2024.** ***SWOT Science Data Products User Handbook*****. NASA/JPL.**

https://doi.org/10.5194/egusphere-2026-2257
Preprint. Discussion started: 23 April 2026

[42] Taburet, G., et al. (2019). DUACS DT2018 sea level anomalies and geostrophic currents. EUMETSAT Ocean and Sea Ice SAF, CMEMS Documentation.

[43] **Thoppil, P. et al. (2021). Ensemble forecasting greatly expands the prediction horizon for ocean mesoscale variability**.

[44] UNESCO (2024). Ocean Decade. Retrieved from: https://www.unesco.org/en/decades/ocean-decade.

[45] Ubelmann, C., Cornuelle, B., and Fu, L.-L., 2015. Dynamic mapping of along-track ocean altimetry: performance from observing system simulation experiments. Journal of Atmospheric and Oceanic Technology, 32(9), 1717–1739.

[46] Walsh, K., 1997. Objective detection of tropical cyclones in high-resolution analyses. Mon. Weather Rev. 125, 1767–1779.

[47] Wang, G., Zhao, B., Qiao, F., Zhao, C., 2018. Rapid intensification of super typhoon Haiyan: the important role of a warm-core ocean eddy. Ocean Dyn. 68, 1649–1661.

[48] Wu, C., Lee, C., Lin, I., 2007. The effect of the ocean Eddy on tropical cyclone intensity. J. Atmos. Sci. 64, 3562–3578. http://dx.doi.org/10.1175/JAS4051.1.

[49] Xu, X., McClean, J. L., & Moore, A. M. (2018). Assessing the predictability of mesoscale eddies in the Southern California Current System using an ensemble of regional ocean model simulations. Ocean Modelling, 122, 35–50. https://doi.org/10.1016/j.ocemod.2017.12.005.

[50] Yang, C., Liu, Y., Testut, C. E., Benkiran, M., Chikhar, K., Smith, G. C., & Fortin, A.-S. (2022). Verifying ocean mesoscale eddy properties using satellite altimetry and global ocean reanalyses.

[51] Yang, C. et al. (2024). Gathering users and developers to shape together the next-generation ocean reanalyses: Ocean reanalyses workshop of the European Copernicus Marine Service.

[52] Zadra, A., McTaggart-Cowan, R., Vaillancourt, P.A., Roch, M., Bélair, S., Leduc, A.M., 2014. Evaluation of tropical cyclones in the Canadian global modeling system: Sensitivity to moist process parameterization. Mon. Weather Rev. 142, 1197–1220.

[53] Zhang, Z., Wang, W., Qiu, B., 2014. Oceanic mass transport by mesoscale eddies. Science 345 (6194), 322–324.