https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

Near-Real-Time Assimilation of Satellite-Derived Ocean Surface Currents Using a Multi-Model Ensemble Kalman Filter

Shahbaz Baig[1, 2], Waqas A. Qazi[2], and Rafia Mumtaz[1]

[1]School of Engineering and Computer Science, National University of Sciences and Technology, H-12, Islamabad, 44000, ICT, Pakistan
[2]MetOcean, 26 Nikou Pattichi street , 3071, O. M. Offshore Monitoring Ltd., Limassol, 4102, Cyprus

Correspondence: Rafia Mumtaz (rafia.mumtaz@seecs.edu.pk)

Abstract. Accurate near-real-time (NRT) estimation of ocean surface currents remains challenging due to sparse in-situ observations and structural model uncertainties. Most operational systems primarily assimilate altimeter-derived geostrophic currents, which omit ageostrophic contributions from wind forcing, coastal processes, and transient mesoscale dynamics. Direct assimilation of satellite-derived ocean surface currents therefore provides a pathway to improve the dynamical consistency of NRT surface current estimates, particularly in regions of highly variable circulation where accurate knowledge of the evolving ocean state is critical for marine operations. We present an end-to-end framework for direct assimilation of high-resolution satellite-derived surface current fields into a Multi-model Ensemble Kalman Filter (MEnKF). Surface currents are retrieved using an adaptive, constrained Maximum Cross-Correlation (MCC) algorithm applied to sequential AVHRR thermal imagery. The Earth Observation (EO)-derived currents are then integrated into a heterogeneous ensemble of global and regional forecasts to explicitly account for structural model uncertainty. Evaluation against coastal HF-Radar observations and regional reanalysis confirms statistically significant improvements over background forecasts. Under optimal observational conditions, the lowest RMSE (0.18 m/s) occurs when 9–12 EO-derived surface current products contribute to each assimilation cycle, accompanied by improved directional consistency relative to reanalysis data. Sensitivity analysis reveals that performance is driven by observational density and spatial representativeness, with maximum skill achieved at intermediate densities of 8-12 images per assimilation cycle. This framework provides a scalable, physically consistent pathway for improving NRT predictions in data-sparse regions.

Keywrods: Ocean Surface Currents, Multi-model forecast, Data Assimilation, Ensemble Kalman Filter (EnKF), Satellite-Derived Currents, Earth Observation (EO), AVHRR, Sea Surface Temperature (SST), Maximum Cross Correlation (MCC), HF Radar

1 Introduction

Ocean surface currents play a fundamental role in the Earth's climate system, mediating air-sea fluxes of heat, momentum, and gases. They are also critical for a wide range of operational applications, including marine navigation, search and rescue operations, oil spill trajectory forecasting, and ecosystem management (Rusvan et al., 2024). Accurate and timely estimation of surface current fields is therefore essential for both scientific understanding and societal needs (Eriksson et al., 2018; Baig et al., 2024).

In-situ measurement systems, such as, drifters, moorings, and coastal High-Frequency Radar (HFR) networks, provide direct, high-resolution current measurements but are geographically constrained by sparse spatial coverage and logistical limitations (Georges et al., 1998). Numerical ocean forecast models offer continuous spatiotemporal predictions but their effective resolution is restricted by computational trade-offs, uncertainties in atmospheric forcing, and imperfect parameterizations. These limitations hinder their ability to capture mesoscale and sub-mesoscale dynamics that dominate surface transport processes (Griffies et al., 2019; Hewitt et al., 2020).

Satellite-based Earth Observation (EO) provides a critical bridge between sparse in-situ measurements and imperfect numerical forecasts, offering spatially extensive and temporally recurrent observations of the ocean physical state across a wide range of environmental conditions, though with an inherent trade-off between revisit frequency and spatial resolution. Among the established EO approaches, satellite altimetry has become the primary observational input into operational forecasting systems, providing Sea Surface Height (SSH) measurements from which geostrophic surface currents are routinely derived (Le Traon et al., 2019). While this strategy effectively represents large-scale, quasi-balanced flow, it is fundamentally limited in coastal and frontal regimes. The altimetry-derived currents capture only the geostrophic component of total surface velocity and do not resolve ageostrophic contributions from wind-driven Ekman transport, tides, or coastal boundary effects (Waters et al., 2024; Moore et al., 2019). In shelf seas and frontal zones, where ageostrophic motions constitute a dynamically significant fraction of the velocity field, altimeter-constrained analyses systematically underestimate total surface transport, leaving a substantial component of the circulation observationally unconstrained (King et al., 2021; Matthews and Emery, 2009). In contrast, Sea Surface Temperature (SST) observations offer a complementary and more observationally complete perspective on surface circulation. SST-derived currents estimates are kinematic in nature, capturing total displacement of surface thermal features regardless of dynamical balance driving that motion, and therefore implicitly include both geostrophic and ageostrophic contributions that altimetry cannot resolve (Matthews and Emery, 2009; Gonzalez Haro et al., 2020). The Advanced Very High-Resolution Radiometer (AVHRR) is particularly well suited for this purpose, delivering near-global SST coverage at $\sim$1 km spatial resolution with revisit intervals of 1 to 6 hours when combining data from all available AVHRR sensors (NOAA, 2024a). Furthermore, in comparison with ocean color or SAR (Synthetic Aperture Radar) imagery, SST-based current retrieval benefits from persistent thermal gradients and broader spatial coverage, making it especially attractive for NRT marine applications in coastal and shelf environments (Röhrs et al., 2021; Choi et al., 2021).

A range of feature-tracking techniques have been developed to estimate surface currents through sequential satellite imagery, including Maximum Cross-Correlation (MCC), wavelet-based approaches, gradient thresholding, inversion equations, single tracer methods, and data-driven techniques such as Neural Networks (NN) (Emery et al., 1986; Kelly and Strub, 1992; Isern-Fontanet et al., 2017; Sun et al., 2017). Among these, MCC remains one of the most widely adopted feature-tracking method due to its conceptual simplicity, computational efficiency, and demonstrated robustness in resolving mesoscale surface dynamics, including fronts, eddies, and shelf-edge dynamics (Heuze et al., 2017; Carvajal et al., 2016; Qazi et al., 2014; Yang et al., 2015). Although these techniques provide spatially extensive and temporally frequent estimates of surface flow, EO-derived surface currents represent velocity fields inferred from tracer displacement between two sequential images. They describe the instantaneous motion of surface features but do not explicitly enforce the governing momentum balance or other physically inconsistent constraints embedded in numerical ocean circulation models nor do they provide prognostic capability for forward prediction (Piracha et al., 2023). Achieving a reliable representation of ocean surface currents remains a persistent challenge, particularly when observations are spatially fragmented, temporally sparse, or confined to the surface layer with limited capacity to resolve the full three-dimensional structure of the flow (Evensen, 2006; Moore et al., 2019).

Data Assimilation (DA) provides a framework to combine these observations with circulation models, allowing EO-derived velocity estimates to inform flow analysis while preserving physical consistency. By combining observations with model forecasts, DA reduces state uncertainty while ensuring that resulting estimates remaining compatible with the governing equations of the model and its resolved spatial scales (Evensen, 2003a; Alley et al., 2019; Kalnay, 2002). This approach underpins modern operational ocean forecasting systems across a wide range of applications, including circulation prediction, coastal hazard assessment, and extreme-event forecasting (Robinson and Pierre, 2000; Sun et al., 2018; Wang and Pan, 2021). Methodologically, DA approaches are commonly classified into variational schemes (e.g., 3D-Var, 4D-Var) and sequential methods based on the Kalman Filter and its extensions (Evensen, 2006; Moore et al., 2019). Among sequential methods, the Ensemble Kalman Filter (EnKF) is particularly advantageous for oceanographic applications. It estimates flow-dependent error covariances through Monte Carlo sampling, avoiding the development of computationally intensive adjoint models required by variational schemes, and flexibly accommodates asynchronous, spatially sparse, and intermittently available observations (Evensen, 2009; Bocquet and Farchi, 2023). To further address uncertainties associated with individual model formulation, Multi-Model Ensemble Kalman Filters (MEnKF) combine multi-source models within a unified ensemble, thereby providing a principled treatment of structural model uncertainty through adaptive ensemble weighting (Xue and Zhang, 2014; Bach and Ghil, 2023). By leveraging complementary strengths across heterogeneous models with varying fidelity and resolution, MEnKF systems have demonstrated consistent improvements over both individual models and unweighted multi-model combinations across deterministic and probabilistic forecasting contexts (Hagedorn et al., 2005; Bach and Ghil, 2023).

Historically, ocean DA efforts have focused predominantly on assimilating scalar variables such as SST and sea level anomaly (SLA). Direct assimilation of EO-derived surface current vectors has remained comparatively limited and fragmented (Park and Xu, 2017; Sakov et al., 2012; Miyazawa et al., 2013; Li and Toumi, 2017; Friedemann, 2022). Existing studies have largely been confined to regional domains, single-model ensemble configurations, and indirect use of satellite information. When satellite-derived currents are considered, they typically rely on coarse-resolution products, simulated SKIM (Sea surface KInematics Multiscale monitoring) like synthetic observations, or geostationary satellite estimates, rather than high-resolution currents retrieved directly from polar-orbiting EO sensors (Ohishi et al., 2022; Baduru et al., 2019; Mirouze et al., 2024). Recent studies have demonstrated the potential of assimilating high-resolution current observations within single-model frameworks. HFR surface velocities have been assimilated into regional coastal models such as WMOP (Western Mediterranean OPerational system) in Ibiza Channel, reducing drifter-model trajectory separation by 53% after 48 hours and clearly improving mesoscale circulation forecasts (Hernandez-Lasheras et al., 2021). Ensemble-based regional systems have also assimilated in-situ and satellite SST to improve upper-ocean current estimates, for example in the northern Indian Ocean, demonstrating that

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

assimilation of combined SST and salinity profiles can substantially correct large-scale current biases (Baduru et al., 2019), More recently, observing system simulation experiments within global eddy-permitting models have assessed the impact of assimilating satellite total surface current velocity observations, indicating significant reductions in surface velocity RMSE particularly in tropical regions (Mirouze et al., 2024; Waters et al., 2024). While these studies represent important advances, they remain limited to single-model based ensemble frameworks and mostly regional or single-configuration applications, offering only integration of partially available surface circulation observations and limited generalizability across regions.

Multi-model DA studies, by contrast, have primarily focused on assimilating in-situ observations such as surface drifters, with satellite data incorporated only indirectly through each model's internal DA approach. For instance, during the Grand Lagrangian Deployment (GLAD) experiment, Coelho et al. (2015) demonstrated improved short-range forecasts in the Gulf of Mexico by combining several operational models with drifter observations. Bach and Ghil (2023) formalized the MEnKF as a minimum-variance linear unbiased estimator and showed systematic performance gains in idealized and ocean-inspired experiments. In both cases, however, the multi-model filter itself never ingested EO derived surface current products.

Our initial work (Baig et al., 2024), presented a prototype for an NRT DA framework capable of directly assimilating HFR captured velocity vectors into multi-model forecasts, highlighting practical challenges such as data gaps, varying input dimensions, error handling, and maintaining dynamical consistency. Mirouze et al. (2024) assessed the impact of assimilating simulated EO data alongside classical observations, showing significant RMSE reduction relative to control runs assimilating only conventional datasets (altimeter, SST, and hydrographic profiles). Collectively, these studies have confirmed the feasibility and benefits of incorporating EO-derived currents to improve NRT forecasting, while underscoring ongoing challenges such as sparse velocity observations, error specification, patchy coverage, and maintaining dynamical consistency between velocity, SSH, and hydrographic fields.

Despite these DA advances, only a limited fraction of available EO-derived ocean surface current observations have been assimilated into operational prediction systems, which continue to rely predominantly on coarser-resolution products and a narrow set of state variables (Mirouze et al., 2024; Lahoz and Schneider, 2014; Chabot et al., 2015). This under-utilization constrains the effective exploitation of surface velocity information and leaves key challenges insufficiently addressed, including the treatment of heterogeneous model forcings, inconsistencies associated with Stokes drift representation across models, and the estimation of optimal ensemble weights under NRT constraints (Bach and Ghil, 2023). To date, no study has implemented a framework for the direct NRT assimilation of high-resolution two-dimensional (2D) EO-derived total surface currents, such as those obtained from AVHRR-based MCC estimates, into a heterogeneous ensemble composed of both global and regional ocean forecast models. Although EO satellite systems now provide near-global, high-temporal observations and ensemble DA methodologies have matured substantially, but their integration within a unified operational framework remains unaddressed, particularly in coastal regions lacking HFR coverage and highly active environments where coarse-resolution models inadequately resolve mesoscale and sub-mesoscale circulation.

This study addresses these gaps by developing, to our knowledge, the first operationally oriented NRT data assimilation framework that directly integrates high-resolution, 2D EO-derived ocean surface currents into a heterogeneous ensemble of global and regional ocean forecast systems. Specifically, it:

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

1. Exploits multi-platform AVHRR-SST imagery to derive surface currents using an adaptive fully automated MCC approach to provide NRT ocean surface currents (Baig et al., 2025, Under Review);

2. Formulates an MEnKF specifically tailored for EO-driven surface current assimilation into ensemble forecast models;

3. Demonstrates the operational feasibility of integrating EO-driven surface currents into an automated, end-to-end NRT assimilation workflow, in which the system ingests newly available satellite observations every 3 hours to deliver frequent corrections of modeled surface currents under variable atmospheric and data-availability conditions.

The resulting architecture provides a scalable and physically consistent pathway for improving short-term ocean surface current forecasts across regional and global domains, enabling seamless integration with multi-model ensemble forecasting systems.

2 Materials and Methods

This section presents the methodological framework developed for the assimilation of EO-derived ocean surface currents into a multi-source forecasting system using MEnKF approach. The methodology is structured to introduce the study area, describe data acquisition and pre-processing, and detail the systematic DA and analysis procedures.

2.1 Study Area

The study focuses on the Galician Coast of Spain, a region characterized by complex mesoscale circulation and exceptionally high biological productivity driven by recurrent coastal upwelling (Navarro et al., 2014; Bao et al., 1997; Otero et al., 2008). This area has also served as the testbed for our earlier work on MCC-based surface current estimation from SST observations Baig et al. (2025, Under Review), thereby ensuring methodological continuity and providing a well-characterized physical context for the present analysis.

The Galician Coast offers an ideal setting for evaluating the assimilation of EO-derived surface currents due to the presence of the OPERA (OPErational RAdars for research in marine sciences) HFR network, delivering continuous, NRT surface current observations for validation (Alonso-martirena et al., 2011; Martinez Casas et al., 2024). The inclusion of dynamically active coastal features such as the Ría de Arousa and the nearby port of Vigo further underscores the ecological and socioeconomic importance of accurately resolving surface circulation in this region (Villasante et al., 2021).

As illustrated in Fig. 1, the study domain spans the rugged continental shelf and adjacent open Atlantic waters, capturing a wide range of dynamical regimes. It also shows the spatial coverage of the selected datasets during January 2021, which forms the temporal focus of the analysis.

To assess the methodological feasibility of the proposed MEnKF framework under operationally relevant conditions, experiments were conducted over a continuous one-month period (January 2021). This duration was selected to evaluate high-frequency assimilation cycling of NRT surface current products while maintaining computational tractability across multiple ensemble configurations and sensitivity experiments. The selected period captures active mesoscale variability and evolving

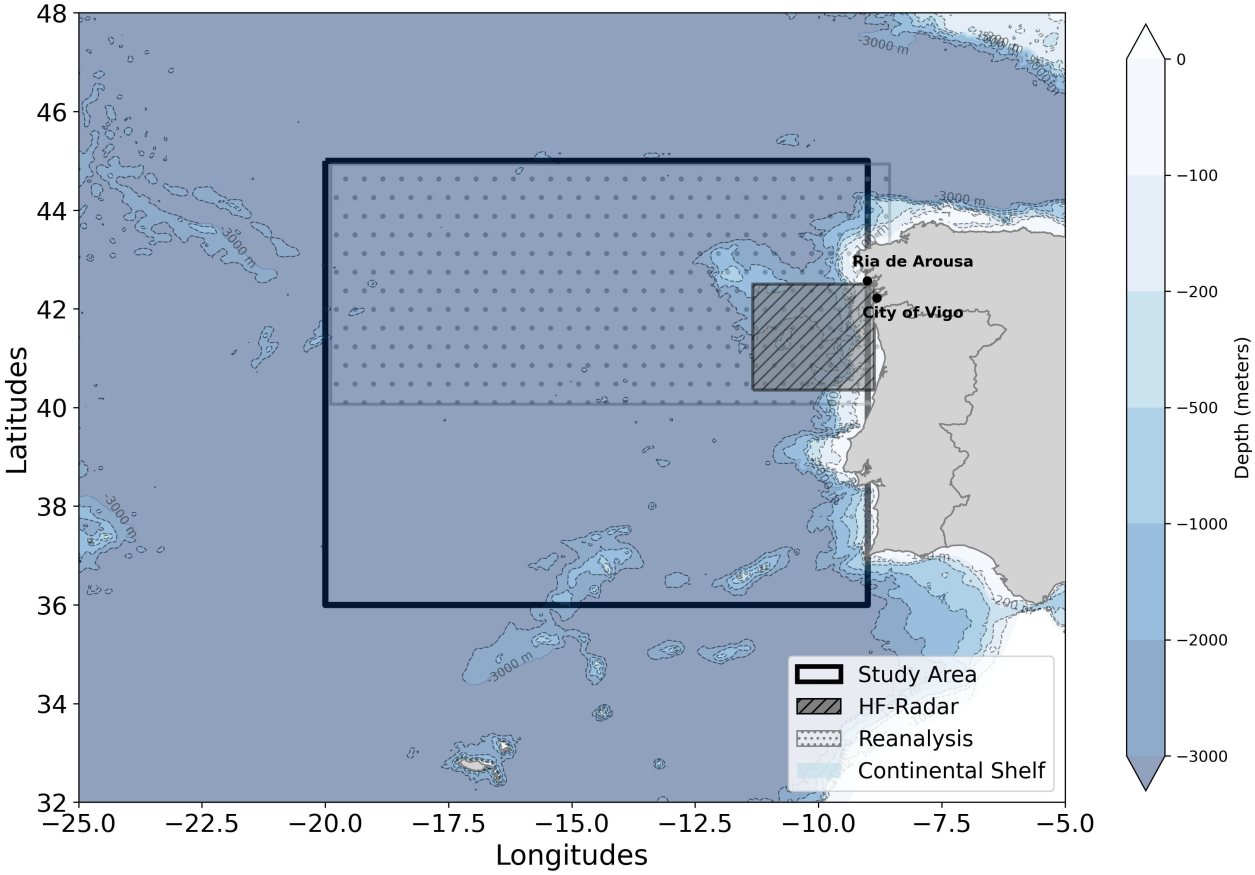


Figure 1. Study area along the Galician coast, Spain. The black box denotes the model domain, the diagonally shaded gray region indicates HFR coverage, and the light gray dotted region represents the spatial extent of the reanalysis product. The background shading contours illustrates bathymetric features derived from GEBCO Compilation Group (2025) highlighting the continental shelf configuration.

flow structures, offering a sufficiently complex dynamical regime for rigorous testing of ensemble behavior, covariance propagation, and incremental forecast skill. The objective at this stage is not climatological generalization, but rather a controlled assessment of system stability, observational impact, and short-range forecast improvement.

2.2 Observational and Model Datasets

The proposed framework integrates four distinct data streams: satellite-based AVHRR-SST observations, coastal HFR measurements, ocean reanalysis products, and multi-source ocean forecast models. A summary of all selected datasets and their key characteristics is provided in Table (1).

The analysis focuses on January 2021, selected from a curated archive of long-range multi-model forecasts and climatological fields generated within the European Commission-funded Horizon-2020 project "EcoSail" (EcoSail Consortium, 2018).

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

These datasets were systematically acquired, preprocessed, and quality-controlled providing a dynamically consistent and temporally continuous ensemble foundation for the present experiments. The availability of harmonized multi-model forecasts and associated climatology enables robust characterization of background error structures and facilitates controlled assessment of assimilation performance. Moreover, the archived time-series structure further supports future seasonal-scale evaluation and longer-term variability analyses, building directly upon the configuration established here.

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

Dataset	Category	Spatial Res. (km)	Delay (h)	Temporal Res. (h)
AVHRR	Satellite Data	1.1	2–6	1–6
HFR	Coastal Radar Observations	5.5	–	1
HYCOM	Forecast Model	8.8	6	3
GRTOFS	Forecast Model	8.8	12	3
NEMO	Forecast Model	4.4–28	6–12	1–3
NWS Reanalysis	Reanalysis	7	–	1

2.2.1 Satellite Data (AVHRR)

Full Resolution Area Coverage (FRAC) Level 1B products from NOAA and EUMETSAT (MetOp) satellites equipped with AVHRR sensors were acquired via the NOAA Comprehensive Large Array-data Stewardship System (CLASS) archive (NOAA, 2024b). The multi-platform AVHRR product are available at spatial resolution of approximately 1.1 km and high temporal sampling at typically less than 1 hour when combining all available platforms. This high temporal resolution feature make it well suited to provide for NRT estimation of mesoscale surface current features.

2.2.2 High-Frequency Radar (HFR) Surface Currents Observations

Hourly HFR current observations at approximately 5.5 km spatial resolution, with coverage extending up to 200 km offshore, were acquired for validation of the proposed multi-model DA framework. HFR data temporally aligned with the satellite observations were obtained via the Intecmar THREDDS data server (Intecmar, 2023). The MCC-derived current observations have already been validated against HFR observations and documented in our earlier work (Baig et al., 2025, Under Review).

2.2.3 CMEMS Reanalysis Products

A regional reanalysis product from the Copernicus Marine Environment Monitoring Service (CMEMS) for the North-West European Shelf (NWS) was acquired, offering 7 km spatial resolution (CMEMS, 2024). This reanalysis product is based on the Nucleus for European Modeling of the Ocean (NEMO) model coupled with the NEMOVAR 3DVar assimilation system, which

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

integrates satellite and in-situ observations, including SST, vertical temperature and salinity profiles. The product delivers hourly, daily, and monthly averaged fields of ocean temperature, salinity, currents, sea level, mixed-layer depth, and bottom temperature across 24 vertical levels. CMEMS updates this product biannually to ensure quality and consistency, making it suitable for regional ocean circulation studies and validation purposes (CMEMS, 2024).

In this study, the CMEMS reanalysis product is used as a physically consistent benchmark for evaluating the performance of the MEnKF system. By integrating multiple observation types within a dynamically constrained framework, modern ocean reanalysis are widely used to assess the skill, spatial structure, and error characteristics of DA and prediction systems, particularly in regions or periods with sparse in-situ observations. The suitability of reanalysis products for evaluating EO-derived surface currents is well established. Heuze et al. (2017) employed reanalysis surface currents to optimize MCC-based surface current retrievals, demonstrating their value as a dynamically consistent reference for both retrieval assessment and assimilation impact. Penny et al. (2019) highlighted the effectiveness of ocean reanalysis as evaluation tools for ocean prediction systems due to their coherent assimilation of SLA, SST, and hydrographic observations. 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. Similar conclusions were drawn by Rahman et al. (2024), under observationally sparse conditions for evaluating ocean prediction systems.

2.2.4 Multi-source Ocean Current Forecast Models

To explicitly represent structural model uncertainty within a multi-model framework, a heterogeneous ensemble was generated from multi-source operational regional and global ocean forecast models. Based on dataset availability during the study period, the ensemble includes the CMEMS global analysis (Le Galloudec et al., 2022), the Iberian-Biscay-Ireland (IBI) regional model (Lecci et al., 2023), the HYbrid Coordinate Ocean Model (HYCOM) (Wallcraft et al., 2003), and the Global Real-Time Ocean Forecast System (GRTOFS) (NWC, 2023). A summary of these models and their key characteristics is provided in Table 1. All selected models provide surface current fields at spatial and temporal resolutions compatible with the underlying DA framework. The diversity in dynamical cores, forcing strategies, and spatial resolutions across the selected models enables a robust representation of structural uncertainty and allows systematic exploitation of complementary model strengths within the multi-model ensemble DA framework.

3 Methodology

The major components of the MEnKF framework include: (i) EO-based surface current retrieval, (ii) construction and harmonization of a multi-model ensemble, (iii) the localized 2D MEnKF implementation, and (iv) multi-tiered validation. The workflow, illustrated in Fig. 2, presents the flow of information through different components of the framework.

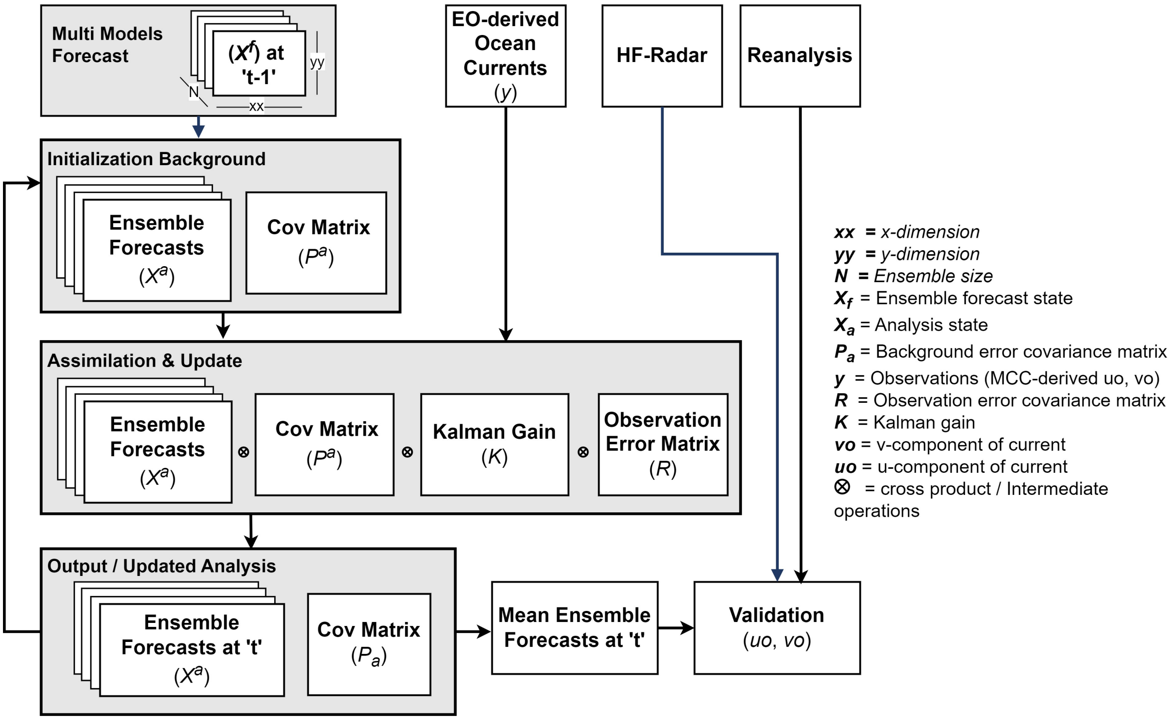


Figure 2. Schematic overview of the MEnKF-based data assimilation framework for EO-derived ocean surface currents.

3.1 EO-Derived Surface Current Retrieval

Surface current vectors are retrieved from AVHRR-SST imagery using an automated, adaptive MCC algorithm. A concise description of the methodological approach used is given below, while comprehensive algorithmic treatment and extended validation are presented in our recent work (Baig et al., 2025, Under Review). Surface current estimates were derived from thermal infrared (TIR) imagery acquired by the AVHRR sensors aboard multiple NOAA and EUMETSAT (MetOp) platforms. The processing chain for MCC algorithm was explicitly designed to assist NRT operations and consists of the following steps:

1. *AVHRR preprocessing:* Level-1 AVHRR data were radiometrically calibrated, reprojected, and resampled to a common spatial resolution (i.e, 1.1 km). Cloud-contaminated pixels were removed using available quality flags, and image normalization was applied to ensure radiometric consistency across sensors and acquisition times.

2. *Multi-platform data enrichment:* 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.

3. *Maximum Cross-Correlation (MCC) feature tracking:* Surface current vectors were estimated using an MCC-based feature-tracking algorithm applied to each sequential image pair. The implementation uses an adaptive, constraint-based framework that automatically adjusts key operational parameters, particularly the search-window, thereby eliminating manual tuning and improving robustness across varying flow regimes and observation geometries. A fixed template-window of 10×10 pixels was employed, while the search radius was adaptively defined as a function of time-lag (ΔT) and an estimated regional maximum surface current velocity of approximately 1 m/s. For short time-lags ($\Delta T < 3$ hours), a minimum search radius of 10 pixels was enforced to ensure stable feature matching.
4. *Quality control and post-processing:* Retrieved velocity vectors were subjected to a two-stage quality control procedure. Vectors with a maximum cross-correlation coefficient below $\rho_{\max} = 0.6$ were rejected. A 3×3 low-pass spatial filter was subsequently applied to suppress isolated outliers and enhance spatial coherence.
5. *Output and validation:* The final product consists of 2D surface current fields representing the mean flow over the time interval ΔT. Each retrieval includes zonal (uo) and meridional (vo) components, the associated maximum correlation coefficient. and spatial and temporal metadata. Independent validation against HFR observations indicated typical RMSE of approximately 0.11 m/s, with complex correlation magnitudes mostly ranging around 0.75 for short time-lag image pairs.

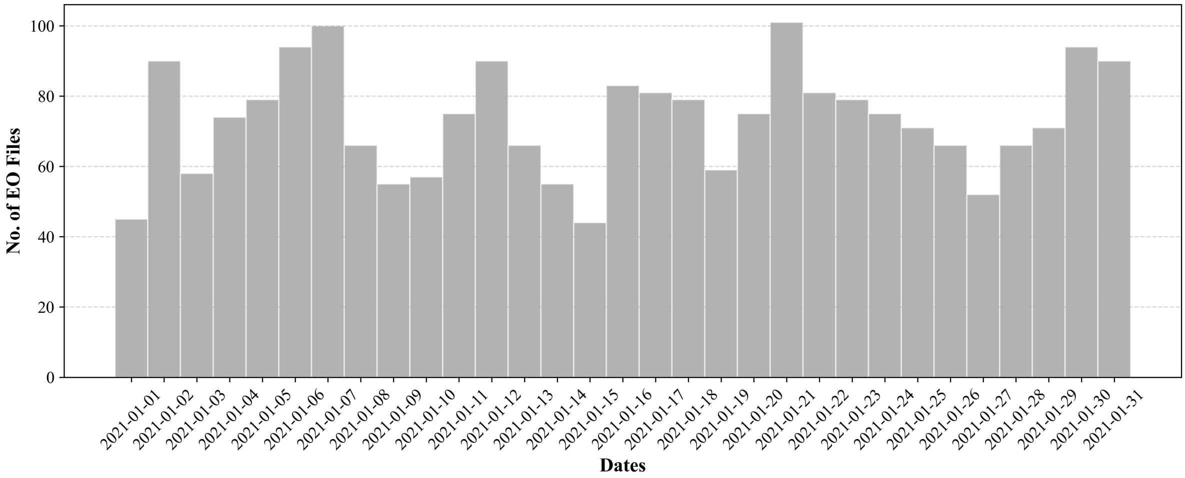


Figure 3. Daily count of MCC-derived ocean surface products retrieved from AVHRR sequential thermal imagery over the Galician coast during January 2021, illustrating the temporal variability in EO data availability for assimilation across the study period.

3.1.1 MCC-derived ocean surface current products preparation

Figure 3 presents the daily number of MCC-derived ocean surface current products available for assimilation throughout the study period. Current retrievals were generated by pairing each AVHRR image with all subsequent images acquired within a 24-

hour lag window, maximizing the number of valid velocity estimates under variable cloud-cover and overpass conditions (Baig et al., 2025, Under Review). Consequently, the daily product count reflects the availability of cloud-free AVHRR overpasses within each 24-hour window. Prior to assimilation, these EO-derived products were subjected to a multi-level quality-control and preprocessing workflow summarized in Table (2). To avoid poorly constrained analysis updates and ensure consistency across assimilation cycles, products containing less than 10% valid ocean pixels were discarded. Additionally, current retrievals derived from image pairs with time-lags shorter than 1 hour were excluded, as empirical inspection revealed increased noise, clustered distribution, and reduced reliability in this regime (Baig et al., 2025, Under Review). Figure 4 illustrates the spatial distribution of MCC-derived surface current frequency, highlighting regions with higher observational density for selected time span.

Missing data are ubiquitous in satellite remote sensing and geophysical time series data, as cloud contamination, quality-control screening, and sensor swath geometry generate spatially clustered gaps rather than random omissions (Kim et al., 2022; Weiss et al., 2022; Nguyen et al., 2020). Conventional spatial or temporal interpolation is frequently employed to bridge these gaps. However, such approaches tend to smoothen sharp gradients, attenuate extremes, introduce artificial spatial coherence, and sometimes fail to fill large gaps, thereby biasing statistical properties and degrading downstream applications (Beckers and Rixen, 2003; Alvera-Azcarate et al., 2005). Alternatively, substituting constant fill values, such as zero velocity, persistence from the last valid observation in the time series, or a climatological mean, avoids interpolation artifacts but introduces its own limitations such as an exact zero velocity is physically unrealistic for most oceanic regions and propagated through covariance estimation or linear algebra operations can induce rank deficiencies, ill-conditioned systems, and distorted spatial statistics (Evensen, 2003b; Cressie, 1993).

To mitigate these limitations and ensure dimensional consistency across DA cycles, missing MCC-derived surface current vectors were assigned a near-zero "calm-weather value of 0.001 m/s" for both velocity components. This value is physically justified: calm-wind conditions (Beaufort force 0–1), associated with surface currents typically <0.1 m/s, are observed in nature (Donelan, 1990). The adopted fill value therefore represents a physically plausible lower-bound state rather than a numerically artificial constant. At the same time, the non-zero magnitude prevents exact singularities in matrix computations while remaining negligible relative to typical MCC-derived current magnitudes (on the order of 10^{-3} of observed values), thereby preserving the covariance structure of the EO-derived fields without introducing spurious correlations (Baig et al., 2024).

Because AVHRR acquisitions occur along variable swath-widths, each EO-derived current image was further padded to the maximize the spatial extent and filled using the same calm-weather baseline to ensure dimensional consistency across assimilation cycles. A no-data mask generated prior to gap handling was subsequently applied during evaluation, restricting validation strictly to grid points supported by valid EO-derived current observations and thereby preserving the integrity of the MCC signal throughout the analysis.

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

Table 2. Summary of the filtering steps applied to EO-derived ocean surface currents.

Step	Description	Count	Percentage (%)
1	Available images before filtering	2,271	100.0
2	Images with $> 10\%$ of valid pixels per scene	1,796	79.0
3	Image pairs with $\Delta T \leq 1$ hour (from step 2)	89	4.0
4	Final EO-derived current products	1707	75.0

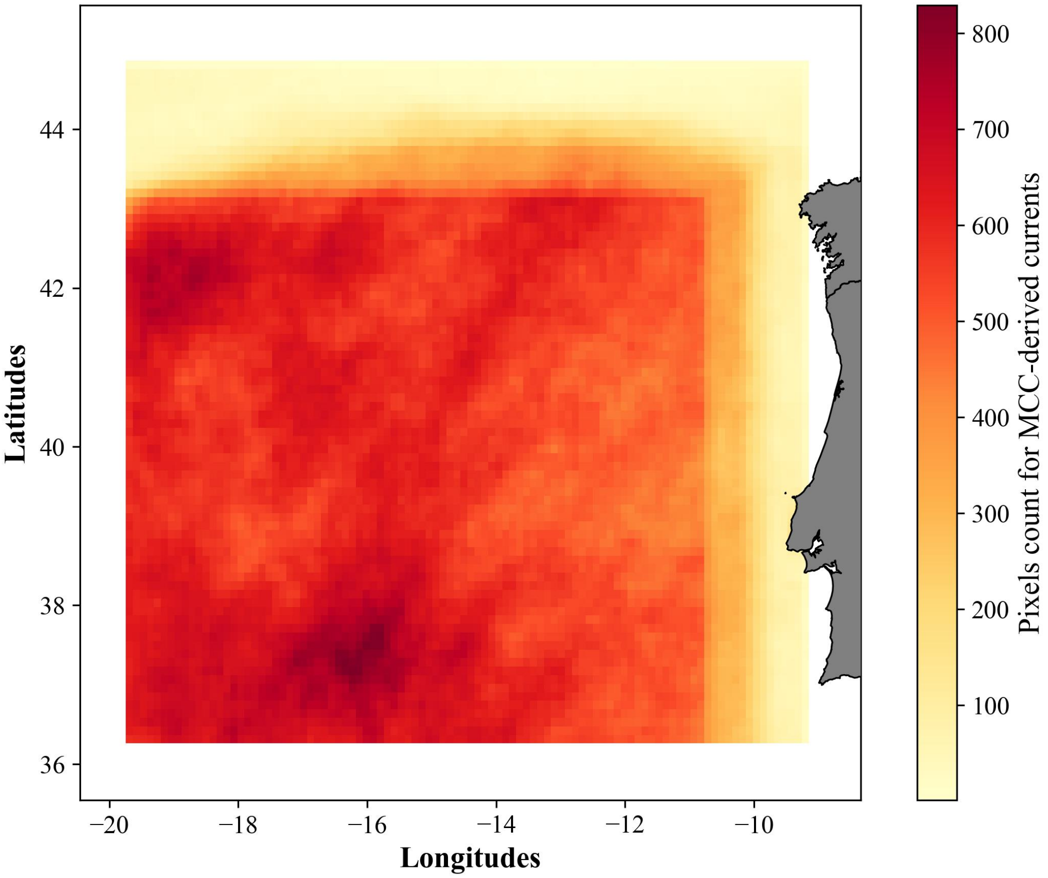


Figure 4. Spatial coverage density of MCC-derived surface current observations during the selected time span, after performing all preprocessing and filtering steps

3.2 Multi-model Ensemble Construction

The multi-model ensemble was constructed from archived ocean surface current forecasts spanning over January–December 2020–2021. Data from 2020 was used to develop the climatology while 2021 data were employed to generate the ensemble forecasts, forming the predictive ensemble backbone for the assimilation experiments. The ensemble comprises a heterogeneous combination of global (HYCOM, NEMO, GRTOFS) and regional (NEMO) ocean circulation models sourced from operational forecasting systems, as described in Sec. 2.2), thereby capturing structural uncertainty across model physics, spatial resolution, and boundary forcing.

To ensure consistency across ensemble members, all model outputs were processed through a harmonization workflow designed to reconcile differences in grid structure, spatial resolution, temporal sampling, and domain coverage. Current velocity components (uo, vo) were extracted at matching lead times and standardized to a common 3-hour temporal reference. Global forecasts were resampled onto the highest-resolution regional grid using bilinear interpolation, preserving flow continuity while avoiding the introduction of spurious gradients. All ensemble members were subsequently masked to the study domain with additional preprocessing applied to reconcile with land-sea masks and remove spurious velocities in shallow or poorly resolved regions. This harmonized construction ensures that inter-member variability predominantly reflects true structural model differences rather than numerical inconsistencies, thereby providing a statistically coherent representation of background error for the MEnKF framework.

3.3 Data Assimilation Framework

MCC-derived surface currents are assimilated for both velocity components (uo, vo) following the quality control procedures outlined in Sec. (3.1) and summarized in Table (2).

3.3.1 Assimilation window and iterative strategy

Assimilation occurs in 3-hourly cycles using an 8-hour observation window extending from 2 hours before to 6 hours after each selected analysis time, as illustrated in Fig. 5. For example, the observations available between 01:00 and 09:00 UTC were assimilated for the 03:00 UTC analysis. This assimilation window was chosen based on the assumption that surface ocean currents remain approximately coherent over a period of up to three hours, with an additional three-hour forward allowance to accommodate the nowcast time frame. Furthermore, this window ensures a sufficient number of observations for assimilation, enabling effective exploitation of the proposed method while facilitating assessment of the observational frequency at which EO products deliver optimal performance. All EO-derived ocean current observations associated with a given analysis timestamp were assimilated, 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.

To mitigate spurious long-range correlations arising from the limited ensemble size, a localized MEnKF formulation was employed. For each observation, an overlapping 10×10 grid-point subdomain was defined, restricting covariance updates to physically meaningful velocity field correlations while maintaining computational tractability in the matrix operations. This subdomain size was selected to reflect the characteristic correlation scales of mesoscale variability in the region, consistent with estimates reported in prior regional studies (Otero et al., 2008; Nogueira et al., 2007). Sensitivity tests confirmed that smaller domains under-represented coherent flow features, whereas larger domains increased both computational cost and spurious long-range correlations. Assimilation was performed independently for both uo and vo velocity components, and ensemble-mean analysis increments were subsequently inserted back into the global analysis field. 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.

This windowed, model-as-ensemble MEnKF formulation enables efficient ingestion of high-resolution EO-derived surface currents while preserving the ocean flow consistency of the contributing models. By combining spatial localization with a heterogeneous multi-model as prior, the framework effectively exploits complementary model skill across coastal and open-ocean regimes, yielding improved NRT surface current estimates over the full study domain.

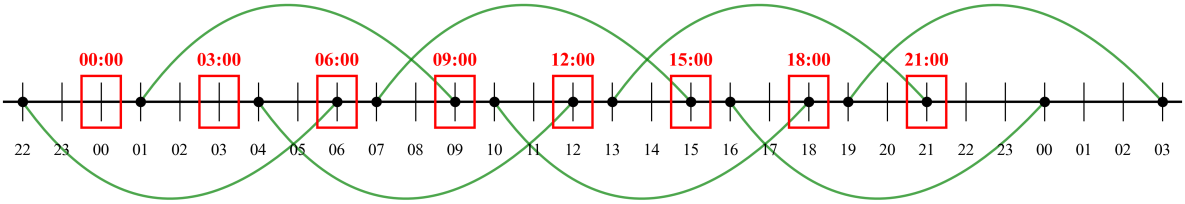


Figure 5. Schematic illustration of the assimilation time window associated with each 3-hourly assimilation cycle. Red boxes denote the analysis times at 3-hour intervals, while green trajectories indicate the corresponding temporal windows from which observations are selected for assimilation in each cycle.

3.3.2 Multi-model Ensemble Kalman Filter formulation

The Multi-model Ensemble Kalman Filter (MEnKF) frames forecasts from N distinct deterministic ocean circulation models as member of a unified prior ensemble, all defined within a common reference state space (see Sec. 3.2). The MEnKF follows the standard EnKF DA procedure comprising three sequential steps: background (prior), analysis update, and posterior, as formulated in Evensen (2003a) and applied in (Bach and Ghil, 2023; Baig et al., 2024).

3.3.2.1 State Vector and Prior

The prior ensemble state for member i at cycle k is defined as:

$$X^f_{k,i} = \mathrm{M}_i\left(X^a_{k-1,i}\right) + w_{k,i}, \qquad w_{k,i} \sim \mathcal{N}(0, Q_i), \qquad i = 1, \ldots, N \tag{1}$$

where M_i denotes the transition operator of the i-th ensemble member, $X^a_{k-1,i}$ is its analysis state from the preceding cycle, w_k, i is the process noise vector, and Q_i is the corresponding model error covariance, reflecting the structural uncertainty associated with each constituent model. Unlike classical sequential EnKF implementations, no explicit transition operator M is applied here, as ensemble members consist of independently produced ocean circulation model outputs from heterogeneous sources rather than perturbations of a single propagated model state, as detailed in Sec. (2.2.4). Consequently, w_k, i and Q_i

are not explicitly constructed but are implicitly represented by the structural spread among heterogeneous ensemble members arising from differences in model formulation, numerical discretization, and atmospheric forcing (Evensen, 2003a; Bach and Ghil, 2023). The ensemble spread therefore serves as the effective background error representation in place of an explicit Q_i specification.

3.3.2.2 Background Error Covariance

The background error covariance P^f is estimated as an ensemble error covariance, estimated from the spread of forecast anomalies across ensemble members. For each constitutional model, anomalies are computed relative to its own monthly climatology mean state $\bar{X}^c$, defined as the time-averaged output of that model for the corresponding calendar month derived from archived forecasts (see Sec. 3.2). This model-specific climatology referencing ensures that systematic biases and seasonal characteristics intrinsic to each model are removed prior to ensemble aggregation, so that P^f reflects physically meaningful inter-member spread rather than inter-model climatological offsets.

The ensemble perturbation matrix is:

$$\mathbf{X}'^{f}_{k,i} = X^{f}_{k,i} - \bar{X}^{c}_{i,m} \tag{2}$$

where $\bar{X}^{c}_{i,m}$ denotes the climatological mean of the i-th ensemble member computed for calendar month m, derived from its archived model outputs.

The background error covariance is then estimated as:

$$P^{f}_{k} = \frac{1}{N-1}\sum_{i=1}^{N} \mathbf{X}'^{f}_{k,i}\mathbf{X}'^{f}_{k,i}{}^{T} = \frac{1}{N-1}\sum_{i=1}^{N}\left(X^{f}_{k,i} - \bar{X}^{c}_{i,m}\right)\left(X^{f}_{k,i} - \bar{X}^{c}_{i,m}\right)^{T} \tag{3}$$

This formulation follows the Ensemble Optimal Interpolation (EnOI) approach of Oke et al. (2002), in which background error statistics are estimated from a stationary ensemble of anomalies relative to a climatological mean rather than from the instantaneous ensemble spread. In contrast to flow-dependent EnKF implementations where $\bar{X}^f$ varies at each cycle, the use of $\bar{X}^c$ provides a dynamically consistent and temporally stable baseline that stabilizes variance amplitudes and reduces susceptibility to sampling noise and variance underestimation arising from limited ensemble size (Bannister, 2008). While this covariance estimation strategy shares characteristics with EnOI (Oke et al., 2002), the present framework retains a flow-dependent character through the use of heterogeneous multi-model forecasts, whose inter-model spread introduces varying covariance at each cycle which is absent in single-model EnOI implementations.

3.3.2.3 Observation Operator and Error Covariance

The observation operator H maps the model state to observation space, incorporating spatial and temporal interpolation, projection to velocity components, and sub-setting to the observation grid. Observation errors are specified through the diagonal covariance matrix $\mathbf{R_j}$, derived from MCC retrieval quality.

The cross-correlation coefficient ρ_j obtained during MCC based feature matching, serves as a proxy for retrieval uncertainty, with higher ρ_j indicating lower observation error and lower ρ_j indicating increased uncertainty (Bowen et al., 2002):

$$\mathbf{R_j} \propto \frac{1}{\rho_j}, \quad j = 1, \ldots, p \tag{4}$$

where j indexes each individual observation, ranging from 1 to (p), the total number of observations available at assimilation cycle k.

3.3.2.4 Kalman Gain and Analysis Update

The innovation vector $\mathbf{d}_K$, representing the departure of the background from the prior state projected into observation space, is:

$$\mathbf{d_k} = y_k - HX_k^f \tag{5}$$

The innovation covariance S_k, which quantifies the total uncertainty in observation space and constitutes the matrix inverted in the Kalman gain computation, is:

$$S_k = HP_k^f H^T + R_k \tag{6}$$

The Kalman gain K_k optimally weights the relative contributions of the background state and the observations according to their respective error covariances:

$$K_k = P_k^f H^T S_k^{-1} = P_k^f H^T \left(HP_k^f H^T + R_k \right)^{-1} \tag{7}$$

The analysis state for each ensemble member i is obtained by correcting the prior forecast with the innovation weighted by the Kalman gain:

$$X_{k,i}^a = X_{k,i}^f + K_k \mathbf{d}_k = X_{k,i}^f + K_k \left(y_k - HX_{k,i}^f \right) \tag{8}$$

The analysis error covariance is updated using the simplified form, valid under the assumption of an optimal Kalman gain:

$$P_k^a = (I - K_k H) P_k^f \tag{9}$$

Following the analysis update, X^a and P^a 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.

3.3.3 2-Dimensional Computational Framework

A key methodological consideration in this implementation is the dimensionality at which the assimilation update is performed. Many previous EO-based data assimilation studies represent the model state as a one-dimensional (1D) vector prior to computing the Kalman gain and applying the analysis update (Oke et al., 2002; Kerry et al., 2016; Baduru et al., 2019). This formulation is not only computationally efficient but also consistent with the canonical vector-space representation of the ensemble Kalman filter, facilitating straightforward implementation, multivariate state handling, and integration with established data assimilation frameworks. These advantages have made 1D representations a standard choice in both operational and research applications.

Despite these strengths, representing spatially continuous geophysical fields in 1D can introduce limitations. Vectorization removes the explicit spatial organization of the model grid, requiring the covariance structure to be inferred in an abstract index space rather than directly in physical space. As a result, covariance estimates may conflate physically related and unrelated grid points, limiting their ability to capture the anisotropic and spatially structured error characteristics of mesoscale ocean dynamics (Hoteit et al., 2005; Nerger et al., 2012).

These challenges are further amplified by the relatively small ensemble sizes typically used in regional ocean data assimilation compared with the high dimensionality of the model state. Under such conditions, the sample covariance becomes rank-deficient and strongly affected by sampling noise, leading to spurious long-range correlations in the analysis update (Houtekamer and Mitchell, 2001; Bergemann and Reich, 2010). Covariance localization is therefore essential, but it is most physically meaningful when applied in the native spatial domain, where geographic distance directly reflects physical proximity and correlation structure. In a purely 1D representation, this spatial context must be reconstructed from the index space, making localization less transparent and potentially less robust. Moreover, although the updated state can be reshaped back onto the 2D grid, any distortion of spatial structure introduced during the assimilation step cannot be recovered at that stage.

To address these limitations, the present framework performs all assimilation operations directly within localized two-dimensional (2D) subdomains. For each observation, an overlapping 10×10 grid-point window is defined, within which the ensemble perturbations, forecast error covariance (P_k^f), innovation covariance (S_k), Kalman gain (K_k), and analysis update ($X_{k,i}^a$) are computed as native 2D matrix operations. This domain-localized strategy follows established approaches in the EnKF literature (Janjić et al., 2011; Bergemann and Reich, 2010).

Operating in 2D preserves the spatial covariance structure between neighboring grid points throughout the assimilation cycle, allowing the Kalman gain to reflect physically meaningful correlation length scales rather than artifacts of vectorization. The updated subdomains are subsequently reintegrated into the full model grid to produce a spatially coherent analysis field. This approach aligns with grid-aware and domain-localized EnKF formulations and is particularly well suited to EO-derived observations with inherently two-dimensional spatial structure, such as swath-based surface current estimates (Evensen, 2003a; Hoteit et al., 2010).

3.4 MEnKF System Validation

The performance of the MEnKF framework is initially evaluated by calculating the RMSE between the analysis fields and withheld observations. RMSE is calculated only at grid cells where valid observations are available at each assimilation iteration, thereby avoiding biases that could arise from interpolation or data gaps. Time series of RMSE for both uo and vo components are examined to assess temporal consistency and the effectiveness of the analysis improvements, particularly in relation to variations in observation density and spatial coverage.

Independent validation is subsequently carried out using surface current measurements for January 2021, from coastal HFR systems. Standard verification metrics widely adopted in operational oceanography, including RMSE, Mean Absolute Error (MAE), Scatter Index (SI), Bias, and complex correlation coefficients (Qazi et al., 2014; Röhrs et al., 2021; Heuze et al., 2017), are computed at 3-hourly analysis intervals throughout the selected study period. Prior to validation, HFR observations are subjected to rigorous quality control and data cleaning, applying established standard error thresholds (0.01 m/s) for both velocity components (Baig et al., 2025, Under Review).

During the experimental period, HFR coverage over the study region remained spatially limited, with significant data gaps and limited temporal overlap with the EO-derived surface current fields. To complement this and enable validation against a spatially more comprehensive dataset, the MEnKF analyses are further compared with regional CMEMS ocean reanalysis products. This product provides hourly surface current fields with approximately 50% spatial overlap of the study region. Reanalysis currents are interpolated onto the same reference grid as the MEnKF outputs, and the same suite of performance metrics was applied both temporally and spatially, offering a dynamically consistent and complementary benchmark for evaluation.

4 Results

Gridded ocean surface currents, derived from sequential AVHRR-SST imagery using the MCC technique, with time-lags ranging from 1 to 24 hours, were assimilated in the proposed MEnKF system. The assimilation was performed for the uo and vo velocity components, employing a time window strategy to maximize observational density for each assimilation update.

Overall, the results demonstrate that assimilating EO-derived surface currents yields systematic and consistent improvements in skill across integrating with heterogeneous multi-model ensemble. By effectively integrating dynamically relevant information from EO data into the MEnKF, the system significantly enhances the representation of mesoscale circulation features.

4.1 MEnKF iterative performance assessment

The temporal evolution of assimilation performance is evaluated using RMSE diagnostics computed at each cycle with reference observations. This analysis provides a quantitative assessment of how the DA framework performs in reference to background and final analysis along successive cycles and varying observational availability across spatial and temporal domains.

Figure 6 presents the temporal evolution of the RMSE for uo and vo surface current components for whole study span, shown in panels (a) and (b), respectively. The dashed gray line reflects the forecast model error (RMSE) at the first assimilation step (background), while solid black line represents the corrected RMSE at final step (analysis) for each individual DA cycle. Green shading denotes difference between background and analysis states (improvement). The gray-shaded bars on the secondary axis indicates the number of EO-derived current products assimilated during their respective assimilation cycle.

The background RMSE exhibits pronounced episodic spikes reaching up to $\sim$1.75 m/s for uo and $\sim$3.0 m/s for vo, while the analysis RMSE remains comparatively stable throughout the study period, consistently staying near $\sim$0.5 m/s for both components across the majority of assimilation cycles. A clear reduction between background and analysis RMSE is observed across the majority of cycles, most prominently where EO observation counts exceed two to three images per cycle. A particularly prominent concurrent spike in both background and analysis RMSE for uo and vo is evident around 2021-01-16, where the available EO observations reduce to ≤ 2 images and the difference between background and analysis RMSE is negligible, indicating an effectively absent assimilation update. Analogous episodes of negligible background-to-analysis improvement are apparent near 2021-01-10 and 2021-01-31, where EO observation counts are similarly low, though the corresponding background RMSE spikes at these dates remain notably smaller than those observed at 2021-01-16. Notably, the relationship between observation count and RMSE reduction is not strictly monotonic: several cycles with comparable EO observation counts yield markedly different magnitudes of background error reduction across the study period.

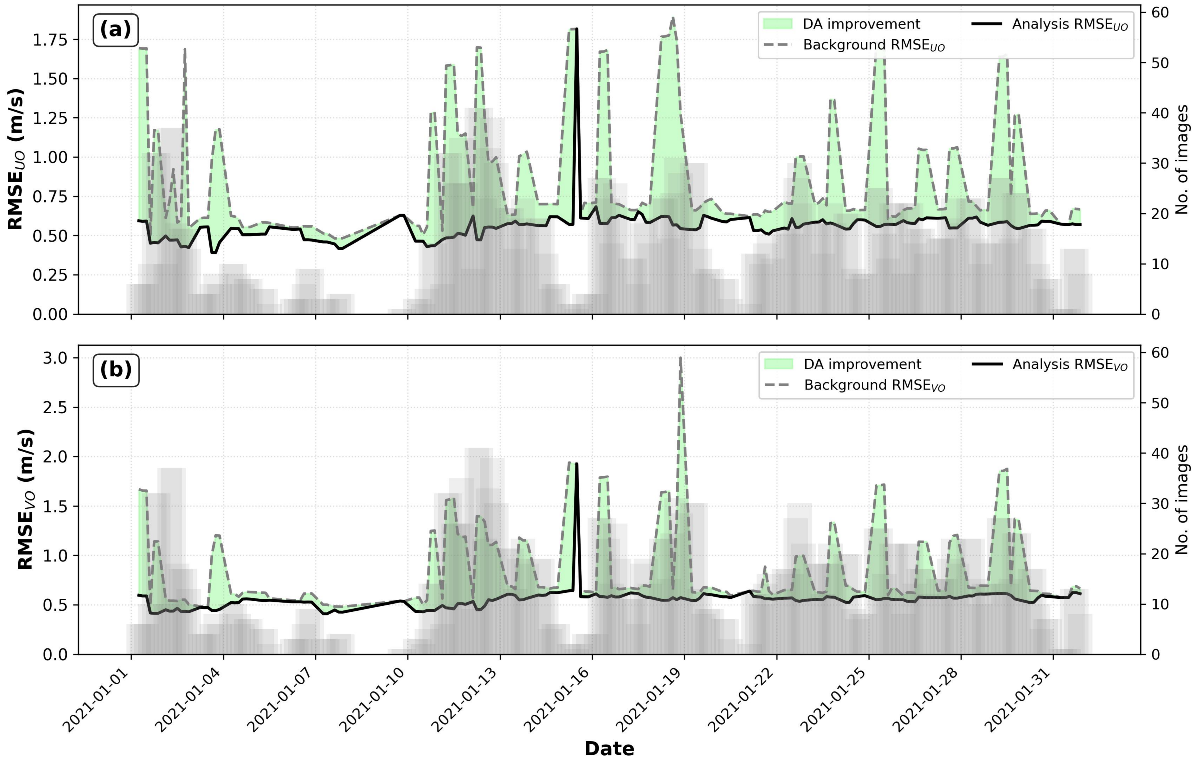


Figure 6. Background versus analysis RMSE for the uo and vo velocity components during January 2021. Dashed lines represent the background RMSE and solid lines are the analysis RMSE for each cycle. Green-shaded regions indicate RMSE performance from background to final analysis step. Gray bars are illustrating the count of assimilated EO images.

Figure 7 further examine the MEnKF performance at two temporal scales: panel 7-a shows RMSE averaged over 3 hourly timestamp for whole month, while panel 7-b presents daily aggregated RMSE values, both plotted against the number of assimilated EO products. The 3 hourly timestamp-level view provides a high-frequency diagnostic to see how MEnKF performed

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

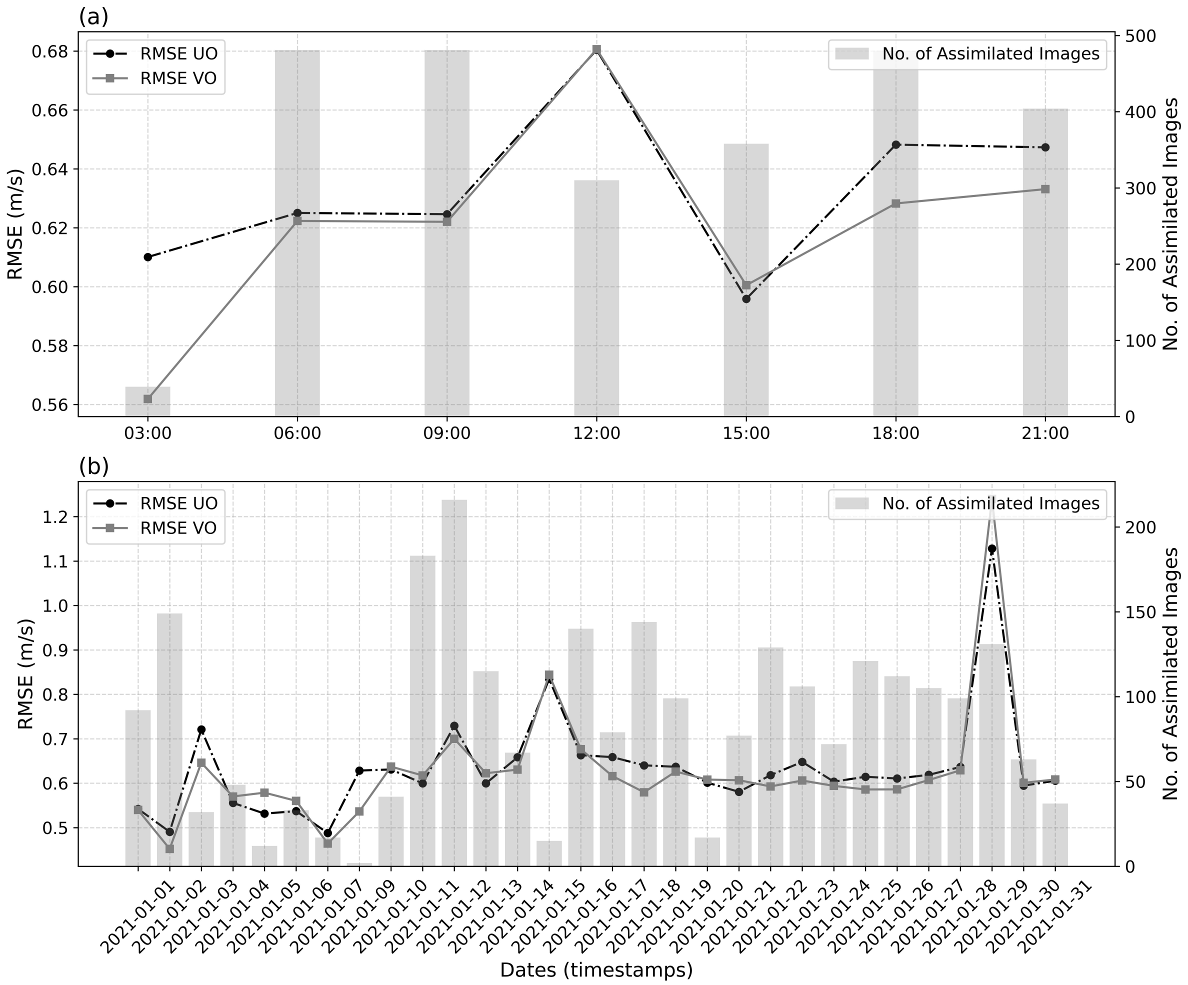


Figure 7. MEnKF performance relative to observational availability. (a) RMSE grouped by assimilation timestamp. (b) RMSE aggregated by date. Bars denote the total number of assimilated EO images within each grouping.

during different time of the day, whereas daily aggregation suppresses transient fluctuations and reveals broader performance trends across the study period.

In Fig. 7, both panels a & b show relatively consistent RMSE for both current components, with a few discrepancies appearing at certain points. Panel 7-a reveals overall diurnal pattern for MEnKF assimilation on selected 3 hour timestamp, in reference to total number of MCC-derived ocean products assimilated. RMSE values range from 0.56-0.68 m/s, with both components exhibiting similar trends: relatively stable performance during early morning hours (03:00 – 09:00), a pronounced degradation peaking around noon (12:00) with RMSE reaching 0.68 m/s, followed by improvement during afternoon and evening hours (15:00–21:00). This diurnal pattern correlates inversely with EO data availability, as indicated by the gray bars

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

showing the number of assimilated images. The midday degradation coincides with reduced image availability (approximately 300 images) compared to early morning and evening periods (450 – 500 images). Panel 7-b presents daily mean RMSE variations throughout the study span, showing generally consistent performance (RMSE ∼0.6 m/s) for most of the month. A significant anomaly occurs on January 27, where RMSE spikes to 1.15 m/s for both velocity components, despite maintaining moderate image assimilation levels (∼200 images). This suggests that data quality or oceanographic conditions, rather than data quantity alone, also influence assimilation performance. The overall monthly mean RMSE pattern demonstrates the MEnKF-MCC approach maintains robust performance across varying data availability, through periods with fewer assimilated images (e.g., January 5-7, 9-11) show slightly elevated RMSE values.

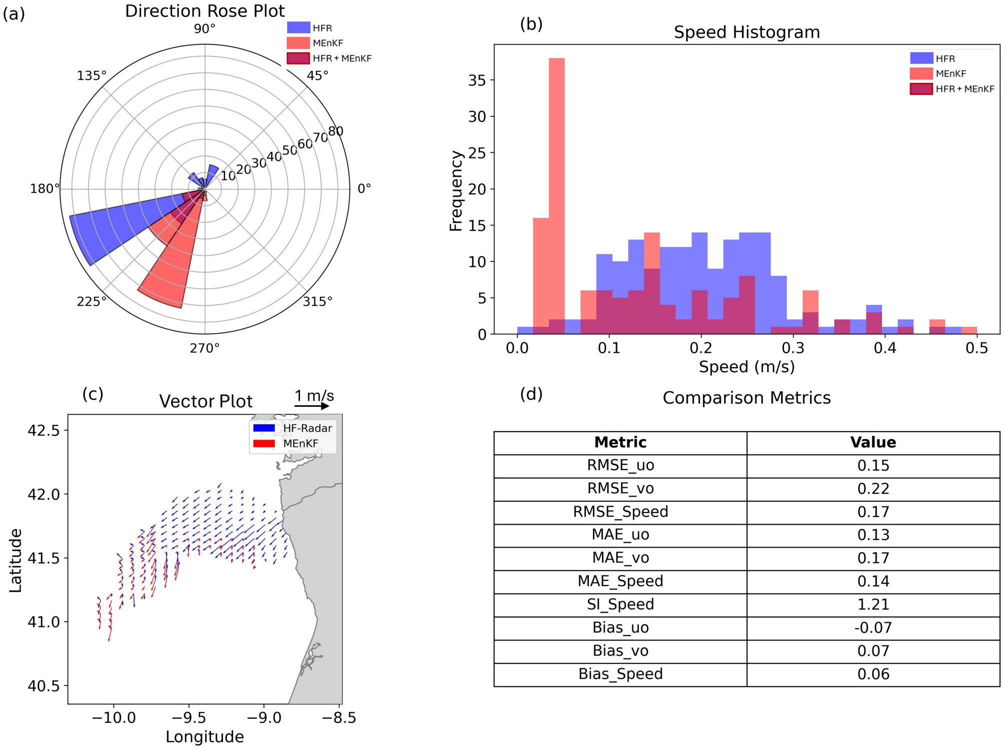
4.2 Independent Validation Using HFR Observations

Direct validation of the MEnKF analyses against coastal HFR surface current observations was performed using spatially and temporally collocated current fields. Representative examples comparing HFR and MEnKF fields are shown in Fig. 8 – 10. Directional consistency and magnitude agreement between MEnKF analyses and HFR observations are assessed through polar diagrams, bar plots, and vector plots (panels a-c). Panel (d) summarizes quantitative performance metrics, including RMSE, MAE, bias, and scatter index (SI). Across these timestamps, RMSE values range from 0.15–0.17 m/s and MAE from 0.10 to 0.13 m/s, indicating low magnitude errors and close agreement with the HFR measurements. A small positive bias (HFR - MEnKF) is consistently observed, while the SI reveals minor misalignments in current vectors, likely reflecting differences in measurement techniques between the datasets and the influence of coastal-scale dynamics (Qazi et al., 2014). The bar plots of speed distributions show that the MEnKF analysis successfully captures the full range of HFR-observed velocities. Pronounced peaks at lower speeds suggest that, during these periods, the MCC-derived currents are dominated by calm-weather conditions. This behavior is likely driven by limited EO coverage (see Fig. 4), which is also reflected in the vector plots where HFR data coverage is sparse. 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.

4.3 Extended Validation Using NWS Regional Reanalysis Products

The North West Shelf (NWS) regional ocean reanalysis product served as an independent reference to enable a systematic assessment of MEnKF performance across the selected study domain. Three representative examples from complete set (175) of analysis are presented through vector plots (Fig. 11, 13, and 15) and statistical comparisons (Fig. 12, 14, and 16) to evaluate the behavior of MEnKF performance characteristics in reference to reanalysis, providing broader spatial and temporal coverage as compared to HFR data.

From these examples, a broad directional agreement between MEnKF and reanalysis is evident across regions of appreciable current magnitude, with localized directional misalignment attributable to the spatial mismatch between the dynamically balanced reanalysis fields and cloud-induced gaps in the MCC-derived velocity retrievals. In most portions of the domain where vectors are prominent for comparison in both datasets, the MEnKF analysis captures the general directional pattern but sys-



Metric	Value
RMSE_uo	0.15
RMSE_vo	0.22
RMSE_Speed	0.17
MAE_uo	0.13
MAE_vo	0.17
MAE_Speed	0.14
SI_Speed	1.21
Bias_uo	-0.07
Bias_vo	0.07
Bias_Speed	0.06

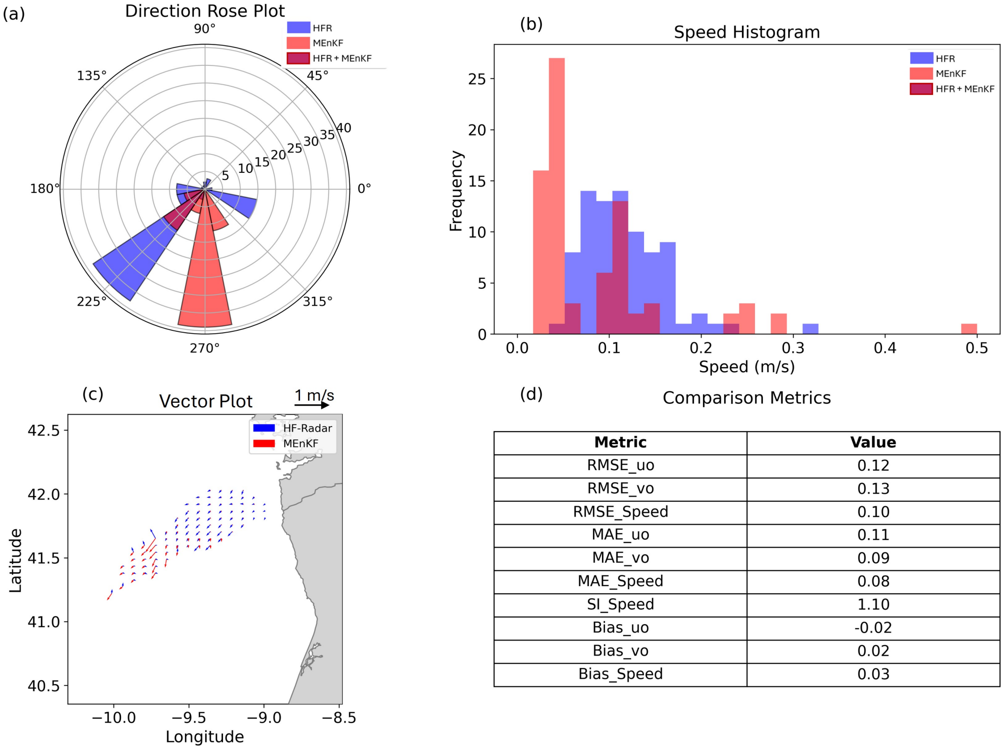
Figure 8. Representative example of MEnKF analysis compared against HFR observed surface currents at 2021-01-01 06:00 UTC, illustrating spatial patterns of current magnitude and vector orientation alongside the corresponding statistical performance metrics.

tematically underestimates current magnitudes, consistent with the inherent limitation of MCC-derived retrievals in resolving energetic mesoscale current structures (Bowen et al., 2002). A consistent pattern is observed across all comparison cases, where variations in the directional coherence and magnitude agreement systematically reflect differences in observational availability and spatial coverage of the assimilated EO products at each cycle.

The statistical comparisons support the spatial interpretations across all three cycles. Directional rose plots confirm cycle-level coherence between MEnKF and the reanalysis, while speed percentile curves show close agreement below approximately the 80th percentile ($\lesssim$0.25 m/s). Beyond this range, a systematic divergence emerges, with reanalysis capturing higher-speed features that are partially underestimated by the MEnKF analysis (Bowen et al., 2002).

Quantitative error metrics show moderate variability across these cycles, with RMSE_{Speed} ranging from 0.15 to 0.21 m/s, MAE_{Speed} from 0.13 to 0.15 m/s, and SI_{Speed} from 0.95 to 1.80. The elevated SI values reflect the high relative variability of speed residuals under the low mean current speeds characteristic of the January 2021 study period, rather than proportionally large absolute errors.

The aggregated metrics from the complete set of MEnKF and reanalysis comparison (175 pairs) are presented in Table 3 to summarize the MEnKF performance. This analysis yield RMSE values of 0.235 and 0.251 m/s for uo and vo, respectively,



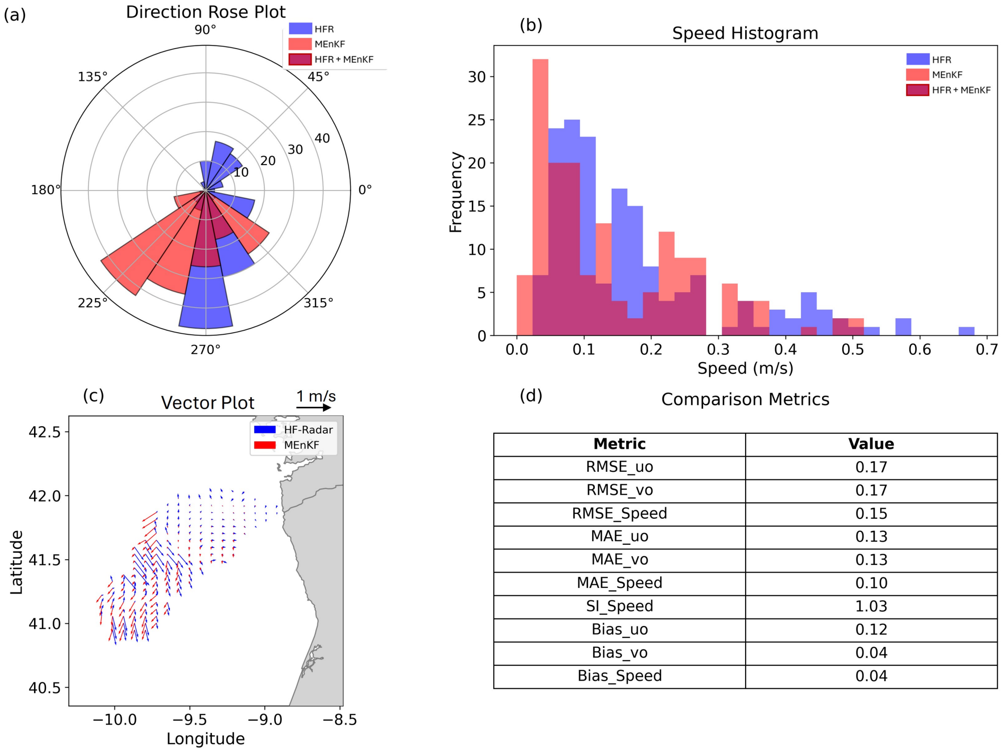
Metric	Value
RMSE_uo	0.12
RMSE_vo	0.13
RMSE_Speed	0.10
MAE_uo	0.11
MAE_vo	0.09
MAE_Speed	0.08
SI_Speed	1.10
Bias_uo	-0.02
Bias_vo	0.02
Bias_Speed	0.03

Figure 9. Representative example of MEnKF analysis compared against HFR observed surface currents at 2021-01-01 09:00 UTC, illustrating spatial patterns of current magnitude and vector orientation alongside the corresponding statistical performance metrics.

and 0.208 m/s for current speed RMSE, with an overall SI of 0.652. These aggregated errors are higher than those from the representative examples, reflecting the broader temporal coverage and wider range of dynamical regimes included in the full assessment. Across all three quantities, the mean MAE is consistently lower than the mean RMSE, indicating that intermittent larger errors contribute disproportionately to the RMSE, while the MAE more closely represents the typical error magnitude. The relatively small difference between RMSE and MAE suggests a moderately heavy-tailed error distribution without systematic dominance by extreme outliers.

Mean bias values (observation − analysis) are small for both velocity components. A slight negative zonal bias (+0.042 m/s) indicates a marginal underestimation of westward flow by the analysis, while the meridional bias (−0.077 m/s) suggests a modest overestimation, consistent with an effectively unbiased vo field in practical terms. The current speed bias of 0.083 m/s is comparatively larger and reflects a underestimation of total current magnitude as also revealed from statistical and vector plots presented above. This behavior indicates a MEnKF analyses slightly underestimates current speeds relative to the reference. The mean SI of 0.652 indicates that the MEnKF analysis reproduces approximately 65% of the variability resolved by the reanalysis product, suggesting that the dominant mesoscale circulation patterns and their temporal evolution are broadly repro-

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026



Metric	Value
RMSE_uo	0.17
RMSE_vo	0.17
RMSE_Speed	0.15
MAE_uo	0.13
MAE_vo	0.13
MAE_Speed	0.10
SI_Speed	1.03
Bias_uo	0.12
Bias_vo	0.04
Bias_Speed	0.04

Figure 10. Representative example of MEnKF analysis compared against HFR observed surface currents at 2021-01-01 12:00 UTC, illustrating spatial patterns of current magnitude and vector orientation alongside the corresponding statistical performance metrics.

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	uo	vo	Speed
RMSE (m/s)	0.235	0.251	0.208
MAE (m/s)	0.187	0.196	0.160
Bias (m/s)	0.042	-0.077	0.083
SI	-	-	0.652

duced across the study domain, while a residual fraction of the reanalysis variability remains unresolved by the assimilation system.

Building on these results, the following sections present a comprehensive, domain-wide skill assessment to examine how performance varies as a function of observational product count and spatial coverage. This extended analysis provides a more granular and systematic basis for evaluating the robustness and operational readiness of the proposed MEnKF framework.

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

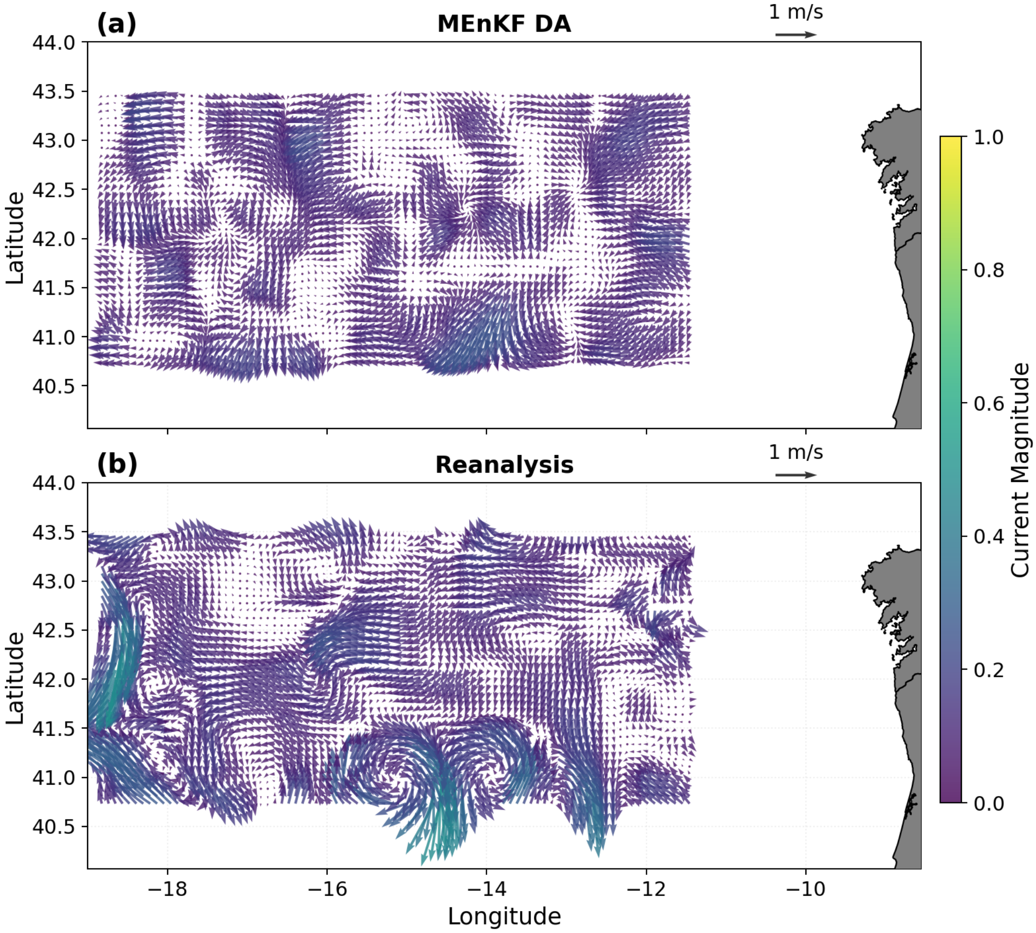
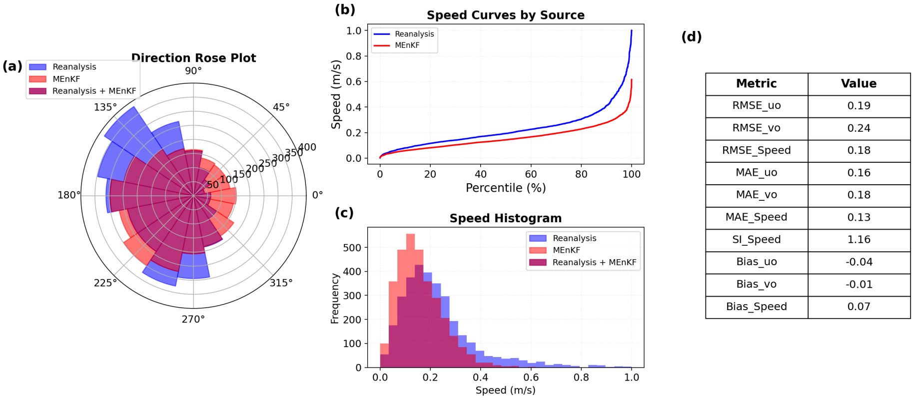


Figure 11. Surface current velocity fields for the assimilation cycle 04012021T21:00 UTC, after assimilating 7 EO-derived ocean products. (a) MEnKF analysis (b) reanalysis. Color indicates current magnitude.

4.3.1 MEnKF sensitivity to observation density and spatial coverage

545 To assess how the temporal and spatial representativeness of EO-derived surface current observations influences assimilation performance, MEnKF sensitivity was examined as a function of both the number of assimilated EO images and the associated valid pixel coverage. Figure 17 summarizes the RMSE, bias, and SI, aggregated within discrete bins of image count (panels a-c) and spatial coverage (panels d-f). Mean values are shown with the ±1 standard error envelopes to illustrate sampling robustness.

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026



Metric	Value
RMSE_uo	0.19
RMSE_vo	0.24
RMSE_Speed	0.18
MAE_uo	0.16
MAE_vo	0.18
MAE_Speed	0.13
SI_Speed	1.16
Bias_uo	-0.04
Bias_vo	-0.01
Bias_Speed	0.07

Figure 12. Statistical comparison of MEnKF analysis and reanalysis surface currents for assimilation cycle of 04012021T21:00 UTC, after assimilating 7 EO-derived ocean products. (a) Direction rose of current directions. (b) Percentile curves of cumulative speed distributions. (c) Speed frequency histograms. (d) Quantitative error metrics. Colors indicate reanalysis (blue), MEnKF (red), and their overlap (purple).

Sensitivity of image count (observational density):

RMSE exhibits a complex, non-monotonic response to increasing image count, revealing distinct optimal and sub-optimal observational regimes. Starting from low image counts (1–2 images), RMSE decreases progressively as observational density increases, reaching a pronounced minimum of ∼0.178 m/s at 9–12 images per assimilation cycle. This represents the optimal observational regime, where ensemble constraint is most effectively balanced. RMSE then increases sharply at 13–20 images (∼0.220 m/s), indicating performance degradation despite increased data availability. This counterintuitive behavior suggests that this specific density range introduces problematic characteristics, potentially including redundant or spatially correlated observations, dynamically inconsistent scenes, or ensemble filter saturation effects. Beyond this regime, RMSE gradually recovers at higher image counts (31–40 images: ∼0.205 m/s; 41+ images: ∼0.198 m/s), though performance does not return to optimal 9-16 images level. The uncertainty envelopes narrow progressively with increasing image count, confirming improved statistical stability despite the non-monotonic RMSE pattern.

Bias remains consistently negative across all bins, indicating systematic overestimation of current magnitude. This bias magnitude shows an oscillating behavior: starting at −0.08 m/s for very low counts (1–2 images), increasing to −0.108 m/s at 7-8 images, improving to −0.09 m/s at the optimal 9–12 images regime, then degrading sharply to −0.125 m/s at 13–20 images, coinciding with the RMSE maximum. This pattern reinforces that 13–20 image regime is particularly problematic, exhibiting both elevated random errors (RMSE) and amplified systematic errors (bias).

SI remains relatively stable (0.62 – 0.66) across most image count bins, showing only modest fluctuations until very high observational densities. The sharp increase to ∼0.725 at 41+ images indicates that extensive observational coverage increasingly samples diverse flow regimes, including weak-current (as highlighted by HFR vs model data comparison) where relative

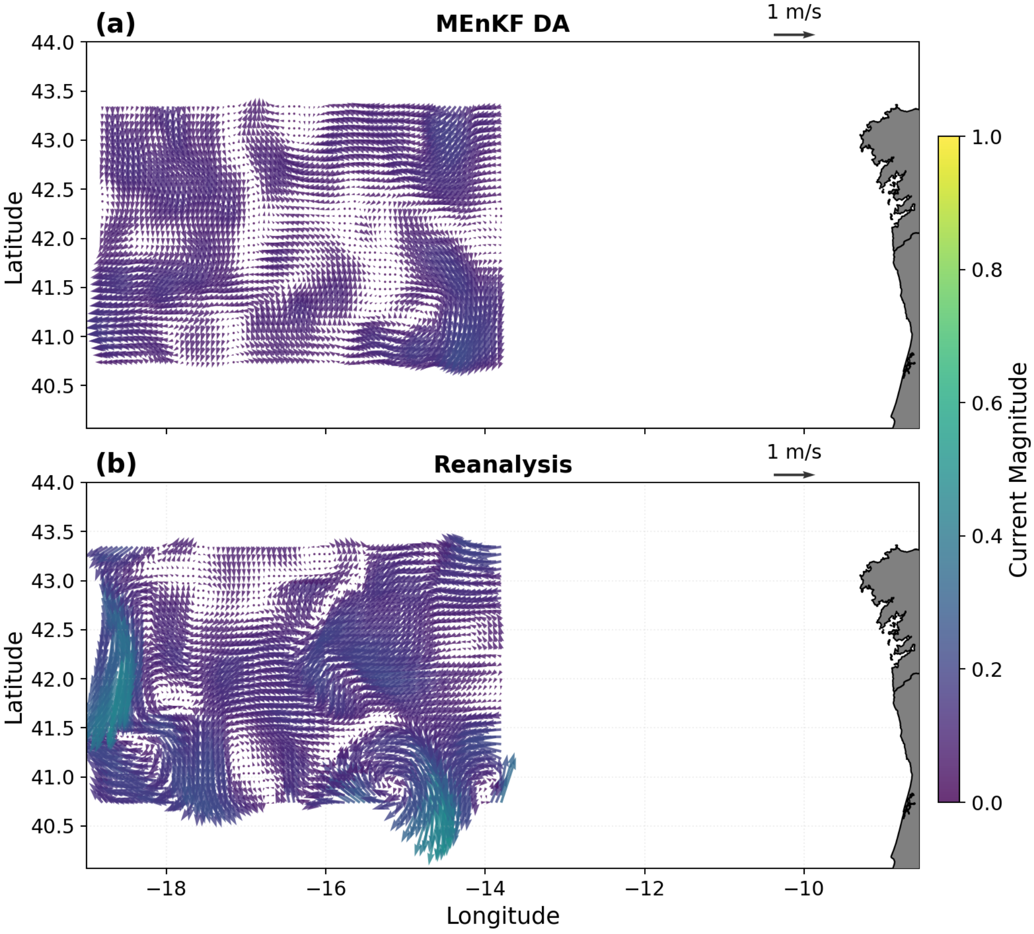
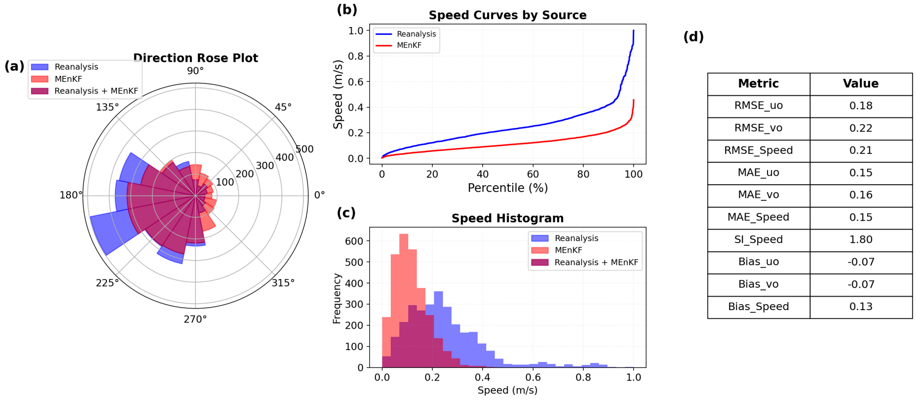


Figure 13. Surface current velocity fields for the assimilation cycle of 07012021T18:00 UTC, after assimilating 4 EO-derived ocean products. (a) MEnKF analysis and (b) reanalysis. Color indicates current magnitude.

absolute errors translate to large relative variance, inflating the skill score. The SI minimum occurs at 13-20 images (∼0.62), whereas the lowest RMSE is observed at 9–12 images; these metrics individually point towards different optimal observational density regimes.

Sensitivity to spatial coverage (pixel counts):

RMSE exhibits a non-monotonic but structured response to increasing pixel coverage. At the lowest coverage levels (<2900 pixels), RMSE is elevated (∼0.23 m/s), indicating limited spatial coverage constraint on the analysis. A pronounced reduction in RMSE occurs at intermediate pixel counts (3000–3100 pixels), where RMSE reaches a minimum (∼0.18 m/s). Beyond this range, RMSE increases slightly before stabilizing at 0.20–0.21 m/s to the highest coverage levels. This pattern suggests an



Metric	Value
RMSE_uo	0.18
RMSE_vo	0.22
RMSE_Speed	0.21
MAE_uo	0.15
MAE_vo	0.16
MAE_Speed	0.15
SI_Speed	1.80
Bias_uo	-0.07
Bias_vo	-0.07
Bias_Speed	0.13

Figure 14. Statistical comparison of MEnKF analysis and reanalysis surface currents for assimilation cycle of 07012021T18:00 UTC, after assimilating 4 EO-derived ocean products. (a) Direction rose of current directions. (b) Percentile curves of cumulative speed distributions. (c) Speed frequency histograms. (d) Quantitative error metrics. Colors indicate reanalysis (blue), MEnKF (red), and their overlap (purple).

optimal observational window at intermediate-to-high spatial coverage, where the analysis is well-constrained. The marginal error increases at the highest spatial coverages implies that beyond a certain threshold, additional observations may introduce redundancy or inconsistencies without yielding further improvements.

Bias remains consistently negative across all coverage bins, indicating a systematic overestimation of the assimilated current magnitude relative to the reanalysis. The bias magnitude is largest at low pixel counts and decreases as coverage increases, ranging from -0.13 to -0.09 m/s. The persistent negative bias suggests a consistent tendency in the assimilation system of the observations themselves.

SI values range between 0.59 and 0.67, with no clear monotonic trend with increasing pixel count. The minimum SI occurs at lower-to-intermediate coverage, while higher SI values are observed at both moderate and high coverage levels. Since SI is normalized by the mean reference magnitude, its variability is likely influenced by the diversity of the sampled flow regimes. For instance, the inclusion of weak-current or transitional regions in high-coverage scenes can inflate relative errors without signaling a degradation of absolute assimilation skill. Overall, the SI remains within a relatively narrow range, indicating stable relative performance across spatial coverage regimes.

Joint sensitivity to image count and spatial coverage:

Figure 18 illustrates the combined influence of image count and spatial coverage on MEnKF RMSE through a 2D sensitivity analysis. The heatmap further reveals complex, non-monotonic dependencies where optimal performance emerges from specific combinations of observational density and spatial coverage, rather than from simple tuning of either parameter in isolation. The lowest RMSE values ($\sim$0.075–0.10 m/s, dark purple) occur at 7-8 images with 3219–3520 valid pixels, representing the optimal balance between ensemble constraint and dynamical consistency. This regime provides sufficient observational in-

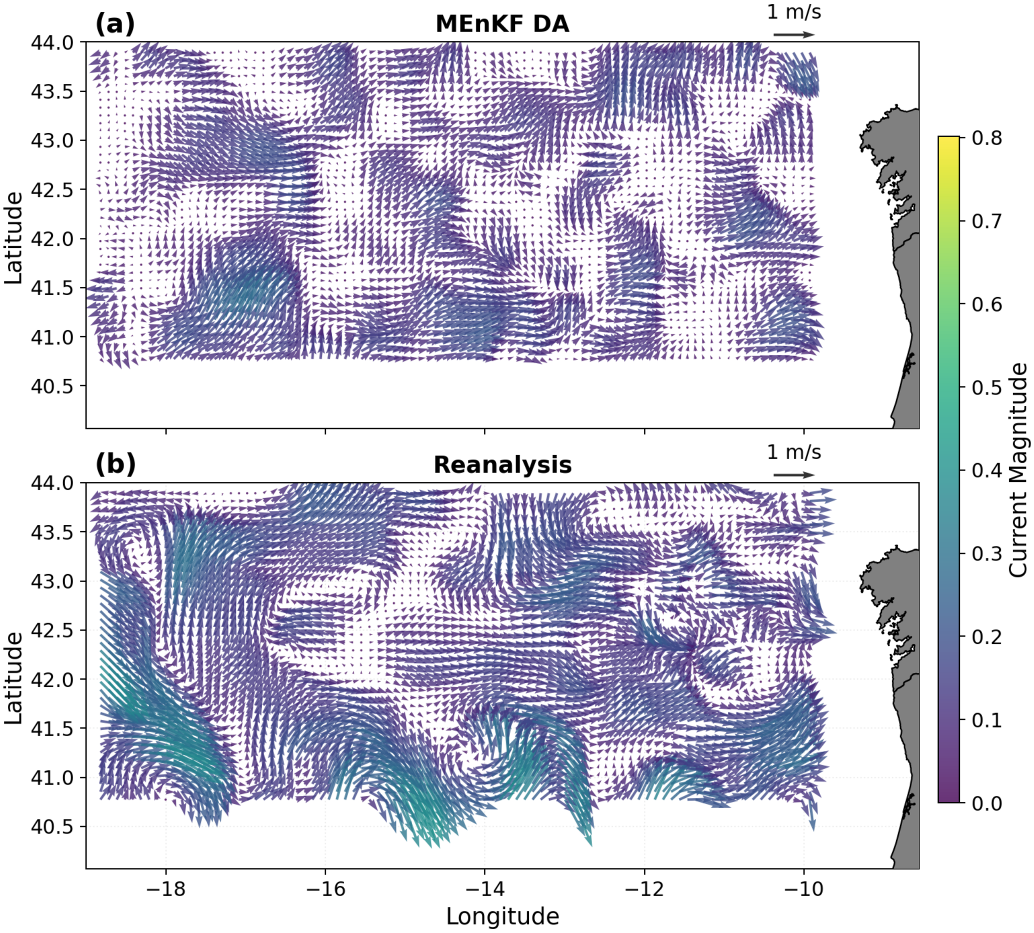
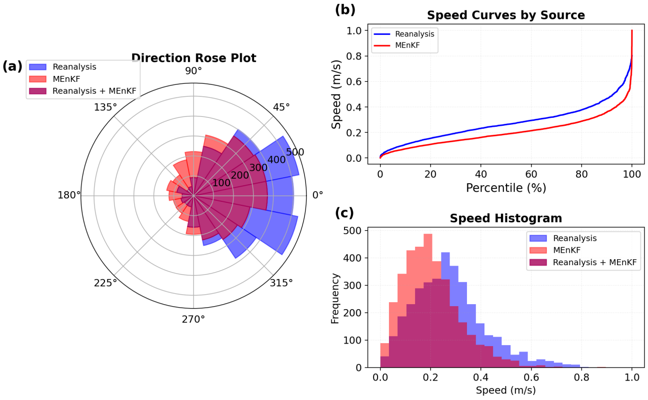


Figure 15. Surface current velocity fields for the assimilation cycle of 19012021T12:00 UTC, after assimilating 13 EO-derived ocean products. (a) MEnKF analysis (b) reanalysis. Color indicates current magnitude.

formation to reduce analysis uncertainty while maintaining coherent flow-field representation. Moderately good performance (RMSE ∼0.15–0.175 m/s, blue-green) extends to adjacent bins, including 9–12 images with similar pixel coverage, confirming a stable optimal zone rather than a single critical point.

At very low image counts (1-2), RMSE remains elevated (0.20 – 0.28 m/s, yellow-green) across nearly all pixel coverage levels, demonstrating that insufficient temporal sampling cannot be compensated by spatial extent alone. The particularly high RMSE (0.275 m/s, bright yellow) at 1–2 images with 3219–3520 pixels indicates that broad spatial coverage with minimal temporal constraint also produces poorly conditioned analyses, preventing the MEnKF from effectively adjusting model uncertainty due to insufficient observations. At very high image counts (41+ images), RMSE stabilizes at moderate levels (0.19–0.22 m/s), confirming asymptotic saturation where maximum data volume yields inferior performance compared to the optimal 7–12 image regime. For fixed image counts, RMSE generally improves with increasing pixel coverage up to ∼3500

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026



(d)

Metric	Value
RMSE_uo	0.26
RMSE_vo	0.26
RMSE_Speed	0.19
MAE_uo	0.20
MAE_vo	0.20
MAE_Speed	0.15
SI_Speed	0.95
Bias_uo	0.11
Bias_vo	-0.01
Bias_Speed	0.07

Figure 16. Statistical comparison of MEnKF analysis and reanalysis surface currents for assimilation cycle of 19 January 2021 (12:00 UTC), after assimilating 13 EO-derived ocean products. (a) Direction rose of current directions. (b) Percentile curves of cumulative speed distributions. (c) Speed frequency histograms. (d) Quantitative error metrics. Colors indicate reanalysis (blue), MEnKF (red), and their overlap (purple).

pixels before saturating or slightly degrading, suggesting that excessively broad spatial sampling may introduce dynamically heterogeneous regions that dilute overall skill.

To comprehensively characterize the agreement between MEnKF analyses and reanalysis currents, complex correlation statistics were computed for all 175 assimilation cycles with sufficient spatial overlap. The complex correlation magnitude $|\rho|$ measures the strength of the linear relationship between vector fields, accounting for both magnitude and direction, while the phase angle 'ϕ' represents the mean counterclockwise angular deviation of reanalysis currents relative to the MEnKF estimates (Qazi et al., 2014; Bowen et al., 2002).

Figure 19 shows the distribution of complex correlation magnitudes (Fig. 19-a) and phase differences (Fig. 19-b) across all overlapping vectors. The correlation magnitude exhibits a relatively narrow distribution with a pronounced peak between 0.41 and 0.43, yielding a mean value of 0.48. The distribution declines rapidly towards higher values (>0.50), with fewer than 5% of cases exceeding |ρ| = 0.52. This concentration indicates consistent moderate-strength agreement between the vector fields magnitudes across the study period.

The phase difference distribution (Fig. 19-b) characterizes the cycle-level directional relationship between MEnKF and reanalysis current vectors, where the phase angle ϕ represents the mean rotation of MEnKF estimates relative to the reanalysis ($\phi = \theta_{MEnKF} - \theta_{Reanalysis}$). The distribution is concentrated near 0°, indicating that the two fields are predominantly co-directional across the majority of assimilation cycles, with secondary concentrations between 30°–60° and around 300°–330° (equivalent to $-60°$ to $-30°$). The mean phase angle of ϕ = 5.97° indicates that MEnKF current vectors are on average rotated 5.97° counterclockwise relative to reanalysis, a negligible systematic offset with no operationally significant directional consequence. The standard deviation of 33.58° reflects moderate cycle-level directional scatter. The distribution shows minimal

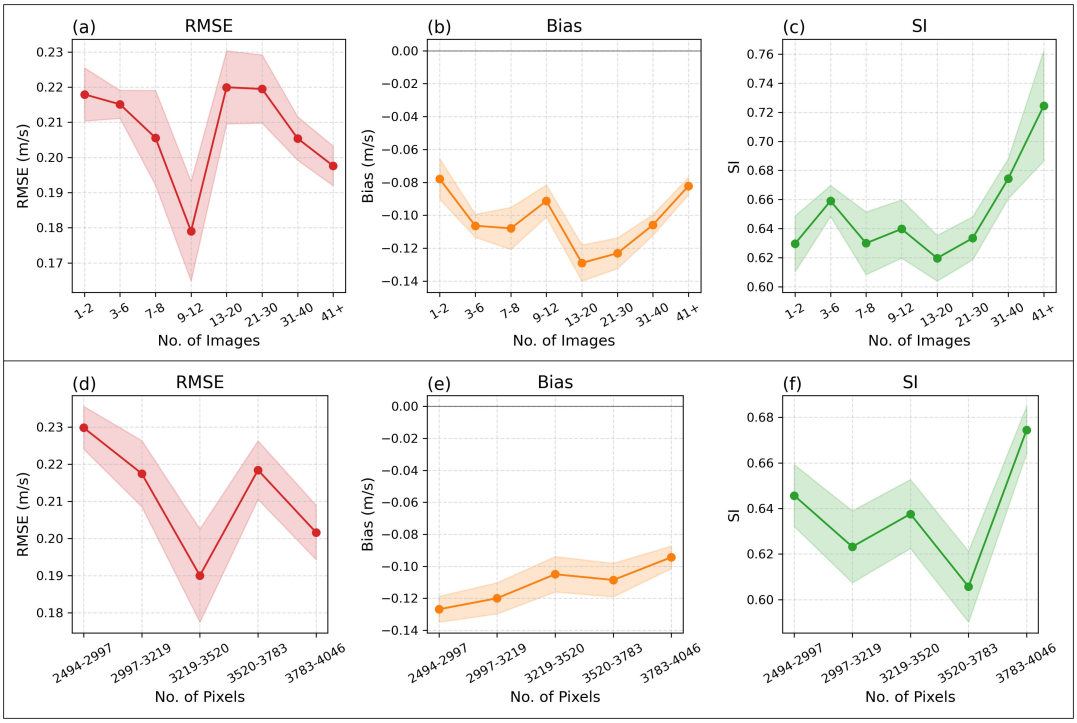


Figure 17. Sensitivity of MEnKF assimilation skill to observational coverage. Panels (a-c) show RMSE, Bias, and SI as functions of temporal data availability (number of assimilated images per cycle). Panels (d–f) present the same metrics versus spatial coverage, binned by pixel count. Markers present bin averaged values; shaded envelopes indicate ±1 standard error.

occurrences in the 90°–270° sector, confirming that near-perpendicular or opposing flow directions between the two fields are rare across the study period.

The principal findings can be summarized as follows:

- The lowest RMSE and MAE for both velocity components and current speed occur in the 8–12 assimilated images bin (RMSE_{speed} = 0.18 m/s; MAE_{speed} = 0.14 m/s), identifying this range as an optimal observational regime for NRT MEnKF operation.
- The close correspondence between RMSE and MAE across all bins indicates that performance improvements reflect systematic reduction in mean error rather than sensitivity to isolated extremes.
- Bias remains small relative to RMSE and MAE across all observational regimes, with a consistent negative bias in current speed in the best performing bins. This is likely attributable to differences in dynamical smoothing and observational constraints between EO-derived currents and the reanalysis reference.
- The SI for speed reaches its minimum value (0.62) in the 12–20 images bin, highlighting robust relative error characteristics even outside the regime minimum absolute error.

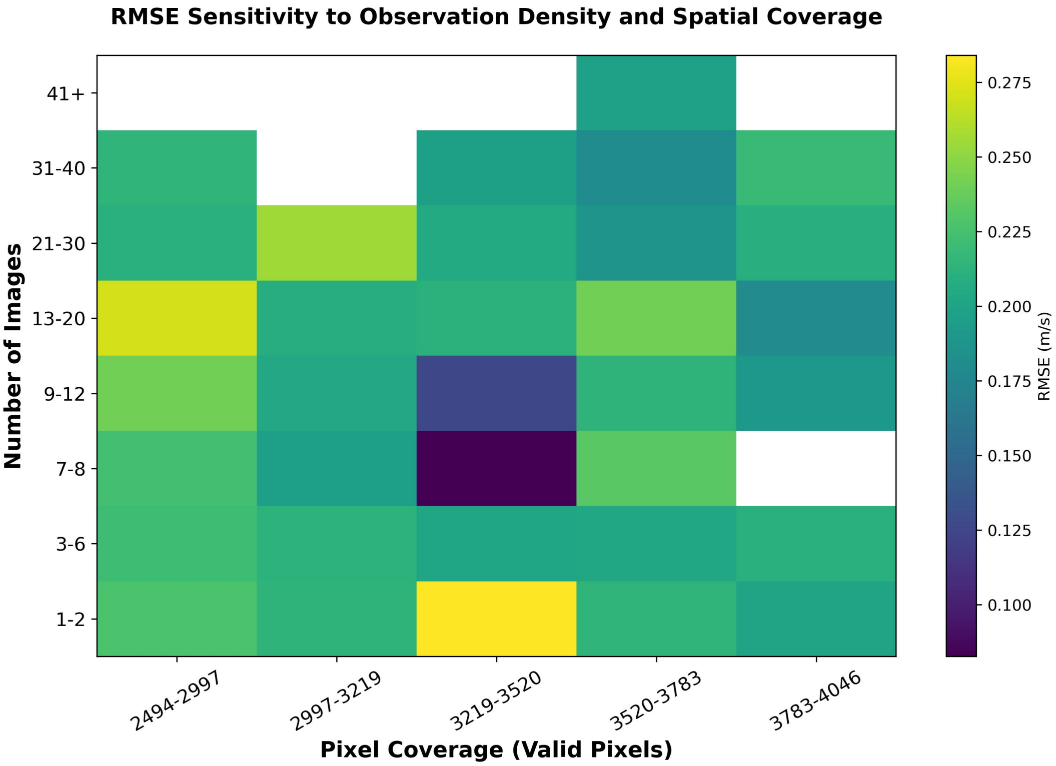


Figure 18. Joint sensitivity of RMSE to observation density and spatial coverage. The heatmap shows the mean RMSE aggregated by binned numbers of assimilated images (y-axis) and quartiles of pixel coverage (x-axis). Color intensity denotes the average RMSE, highlighting the combined influence of data availability and spatial coverage on assimilation performance. White color indicates data gaps.

- Increasing the number of assimilated EO products beyond an intermediate threshold yields diminishing returns, underscoring that spatial representativeness and information content are more important than raw data volume for effective surface current assimilation.
- Complex correlation analysis confirms coherent vector agreement between MEnKF analyses and reanalysis currents, with moderate correlation magnitudes and small mean phase offsets. This indicates that the assimilation preserves dominant flow directions and mesoscale structures, with residual discrepancies primarily reflecting structural differences between EO-derived surface signals and the dynamically smoothed reanalysis reference.

5 Discussion

This study evaluates the performance of NRT assimilation of MCC-derived surface currents within the MEnKF framework. Across the validation metrics and evaluation strategies considered, the assimilated outputs consistently outperform the corresponding reference data, indicating that satellite-derived velocity estimates provide dynamically meaningful constraints when

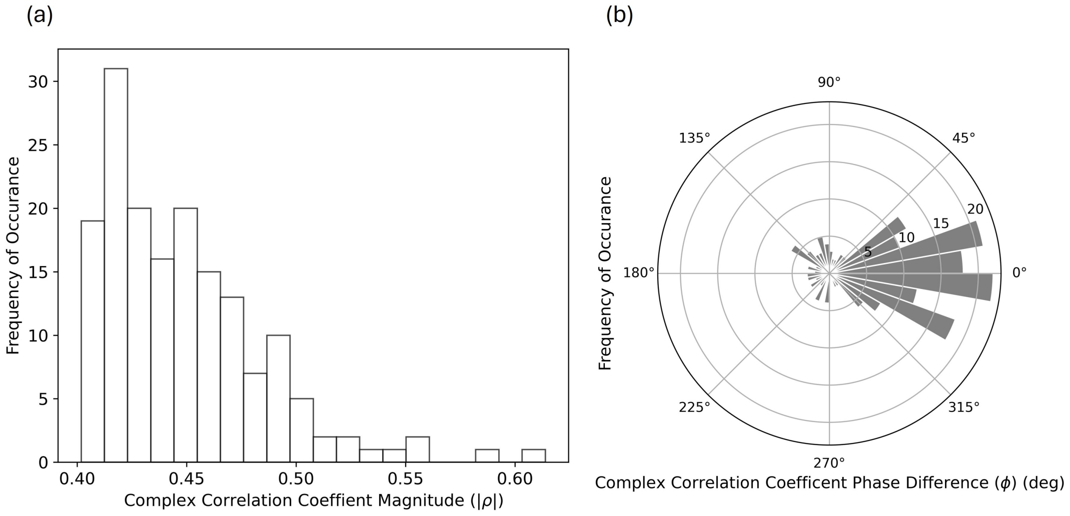


Figure 19. Histogram of complex correlation magnitudes and phase differences between MEnKF and Reanalysis currents computed for Jan. 2021 data. (a) Histogram of complex correlation magnitude; (b) phase difference, expressed as counterclockwise from East.

incorporated into a multi-model ensemble DA system. Performance, however remains sensitive to observational characteristics and prevailing environmental conditions.

These results align with previous work demonstrating that satellite SST observations can substantially improve upper-ocean circulation estimates when assimilated within statistically consistent frameworks (Wang et al., 2023; King et al., 2021; Vandenbulcke et al., 2017). The present analysis complements those studies by examining performance under operational NRT constraints, quantifying sensitivity to observation density, and documenting systematic patterns that emerge across iterative assimilation cycles.

Assimilation cycles with comparable numbers of EO products sometimes produce substantially different levels of background error reduction, indicating that assimilation efficiency depends not only on observational density, but also on observation quality and dynamical consistency with the ensemble background (Oke et al., 2008). Intermittent degradation episodes are plausibly linked to reduced SST feature-tracking reliability under unfavorable atmospheric conditions including persistent cloud cover and weakened thermal gradients, both of which increase uncertainty in MCC-derived velocities, or to transient dynamical regimes insufficiently represented within the ensemble spread (Evensen, 2003a; Talagrand, 1997). The mean analysis observation misfit (∼0.50 m/s) should therefore be understood as a compound product of retrieval uncertainty, ensemble representativeness limitations, insufficient number of EO-products, and the structural low-energy bias introduced by the calm-weather gap-fill strategy, the latter of which is examined in detail below.

Moving from internal diagnostics to independent skill assessment, the HFR-based validation provides a more informative assessment of genuine forecast improvement. 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 well beyond what the self-consistency diagnostic alone implies. This value lies within the characteristic velocity range of Galician shelf circulation (Otero et al., 2008). Comparison with independent validation benchmarks from analogous coastal and shelf DA systems, Oke et al. (2008) reported 0.20–0.35 m/s for a coastal EnKF system and Kerry et al. (2016) obtained 0.20 – 0.30 m/s for East Australian shelf reanalysis, places the present approach among the better-performing configurations in reference to the literature. Although the spatial converge of overlapping HFR and MEnKF data limits the geographic extent of this comparison to a sub-region of the domain, the subsequent extended comparison against NWS reanalysis product confirms that the skill generalizes across the full study region.

Across all 175 assimilation cycles, the average RMSE values of 0.235 and 0.251 m/s for uo and vo respectively, and a current speed RMSE of 0.208 m/s, demonstrate that the MEnKF reproduces the dominant mesoscale circulation patterns with consistent domain-wide skill. The overall SI of 0.652 indicates that the framework captures $\sim$65% of the variability resolved by the reanalysis, a competitive result for an EO-driven ensemble assimilation system operating under moderate data gaps (Wang et al., 2023; King et al., 2021). Moreover, the reanalysis product was not used at any stage to configure, calibrate, or constrain the assimilation framework, meaning that any alignment between the MEnKF analysis and the reanalysis reflects genuine dynamical skill rather than an artifact of shared climatological assumptions. Explicit tuning toward the reanalysis climatology could improve this metric superficially but would simultaneously compromise the independence of the validation, rendering the comparison uninformative as a true skill assessment. The bias structure provides additional diagnostic insight. The near-zero meridional bias (+0.029 m/s) confirms an effectively unbiased vo analysis, while the modest negative zonal bias (-0.050 m/s) and more pronounced speed bias (-0.103 m/s) indicate a systematic tendency of the MEnKF analysis to overestimate current magnitudes relative to the reanalysis product. This is physically consistent with the known tendency of MCC-derived observations to capture energetic surface features, particularly in regions of strong thermal gradients, where instantaneous feature displacement may exceed the smoothed velocities represented in reanalysis products that are subject to temporal and spatial averaging (Bowen et al., 2002; Heuze et al., 2017).

The stable, low-variance trajectory of the analysis RMSE throughout January 2021, despite pronounced episodic spikes in the background error, demonstrates that the MEnKF maintains robust error-reduction capacity across a range of observational densities and dynamical conditions. However, periods of marginal or absent improvement highlight the sensitivity of ensemble-based assimilation to observation sparsity, a structural vulnerability that is amplified by the current calm-weather fill strategy. Replacing the static fill with physically informed, data-driven gap-filling would reduce the proportion of uninformative observations entering the update step, potentially narrowing both the internal misfit and the residual against independent measurements, and is identified as a priority development for future implementations.

A central outcome of this analysis is the non-monotonic dependence of assimilation performance on observational density and spatial coverage. RMSE and MAE for the uo and vo current components, as well as current speed, reach minimum values at intermediate observational densities, with optimal performance occurring at approximately 7–12 EO-derived surface current observations per cycle and spatial coverage of 3000–3700 valid pixels. In this regime, the assimilation effectively constrains mesoscale surface circulation while avoiding degradation associated with either insufficient sampling or problematic high-density configurations. At lower observational densities (1-6 images < 2900 pixels), elevated errors (RMSE $\sim$0.20 – 0.23

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

m/s) reflect inadequate spatiotemporal constraint on the ensemble, consistent with established limitations in surface current assimilation where sparse observations cannot sufficiently reduce background uncertainty (Breivik and Sætra, 2001).

Conversely, RMSE values of ∼0.22–0.27 m/s observed at intermediate image counts (12–20 images) with low spatial coverage despite increased data availability, suggest that observational density alone is an insufficient predictor of assimilation skill. The joint sensitivity analysis (Fig. 18) indicate that this degradation likely reflects multiple interacting factors. Spatially clustered observations (e.g., Fig. 4) within a limited domain can introduce correlated observation errors that compromise the independent assumptions underlying ensemble update (Janjić et al., 2018). When image pairs span extended temporal windows, the assimilated velocity fields may sample dynamically inconsistent flow states arising from tidal evolution or diurnal heating cycles, introducing conflicting contrasts into the ensemble update (Bannister, 2008). Additionally, finite ensemble size limits the system's capacity to accommodate dense and spatial heterogeneous observational input without accumulating spurious correlations (Evensen, 2003a). A partial recovery in assimilation skill is observed at very high image counts (>30 images, RMSE ∼0.20 m/s), which is consistent with a statistical averaging effect across a larger number of observations partially compensating for individual observation noise. However, skill does not recover to the level achieved in the optimal low-density regime (9–12 images), suggesting that beyond a certain threshold, additional observations introduce redundancy and sample increasingly heterogeneous flow conditions that the ensemble cannot effectively reconcile. The SI values remain relatively stable between 0.62 and 0.67 across most observational density bins, confirming that the relative error structure of the analysis is consistently maintained even where absolute error reductions saturate, with the modest increase at the highest image counts reflecting broader dynamical sampling rather than a fundamental improvement in ensemble correction capacity.

Temporal performance analysis reveals predictable diurnal patterns strongly correlated with satellite availability and atmospheric conditions affecting SST retrieval quality. The midday degradation (RMSE ∼0.68 m/s at 12:00) coincides with reduced image availability (300 images), relative to morning and evening periods (450 – 500 images), reflecting both decreased observational input and reduced SST feature contrast arising from diurnal warm layer development and increased atmospheric water vapor loading during midday hours (Minnett, 2003; Donlon et al., 2002). On the Galician shelf specifically, solar-driven diurnal SST variability can locally suppress the thermal gradient signatures that MCC relies upon for feature tracking, particularly during low-wind winter conditions when the diurnal warm layer is not broken down by surface mixing (Donlon et al., 2002). Conversely, improved performance during morning (0:600 – 09:00) and evening (18:00 – 21:00) periods is consistent with higher image availability and more favorable thermal infrared retrieval conditions, including reduced solar contamination and a spatially coherent cool skin layer that enhance thermal gradient contrast (Emery et al., 1986; Bowen et al., 2002). The temporal stability throughout January 2021 (RMSE ∼0.6 m/s), maintaining stable performance across the variable wintertime characteristics of the Galician shelf, including mesoscale eddy activity and intermittent poleward flow events (Peliz et al., 2002). The isolated spike on 27 January represents a clear departure from this baseline and is consistent with a transient dynamical event such as mesoscale eddy passage or frontal intensification, both of which are documented wintertime features of this region (Peliz et al., 2002), temporarily exceeding the predictive capacity of the ensemble background (Evensen, 2003a). The comparable performance of uo and vo velocity components across all temporal scales confirms that the assimilation framework does not introduce systematic directional directional bias into the corrected velocity fields. The semidiurnal tidal dominance of the Galician shelf as described by (Nogueira et al., 2007), further provides a regional physical basis for the dynamically inconsistent velocity constraints identified in the joint sensitivity analysis, particularly for image pairs spanning lag intervals approaching or exceeding a full tidal cycle of ∼12.5 hours.

Systematic bias characteristics provided important insights into assimilation behavior and observation-model relationships. Bias remains consistently negative across all observational regimes (magnitude -0.08 to -0.14 m/s), indicating persistent overestimation of current magnitude by MEnKF analysis relative to the reanalysis reference. Critically, bias magnitude is substantially smaller than RMSE and MAE, confirming that assimilation errors are dominated by random rather than systematic components, a desirable characteristic for operational systems where unbiased estimates support statistical forecasting and ensemble prediction. The reduction in bias magnitude from sparse to optimal observational densities demonstrates that adequate sampling helps constrain systematic deviations, yet incomplete elimination even at optimal regimes indicates structural differences between the assimilated system and reference. The physical explanation of this behavior is: SST-derived geostrophic currents emphasize surface-intensified, energetic mesoscale signals that may amplify velocity magnitudes under thermal wind balance assumptions, directly capturing sharp thermal gradients and frontal features from high-resolution satellite observations. In contrast, reanalysis products represent dynamically smoothed, depth-averaged velocities constrained by subsurface observations and model physics that naturally dampen surface variability and moderate extremes. Importantly, the modest and relatively stable bias across observational regimes indicates that MEnKF updates preserve large-scale dynamical balance without introducing unrealistic current amplification. This confirms that ensemble-based covariance structures approximately represent physical relationships between observed SST patterns and underlying velocity fields while moderately favoring energetic surface signals over smoothed reference estimates.

The moderate complex correlation magnitude (mean $|\rho|$ = 0.48) represents a reasonable level of agreement given the structural differences between the compared products. Several factors contribute to this moderate, rather than high, correlation. First, the region exhibits strong spatial contrasts in circulation intensity: energetic boundary currents and mesoscale features dominate the shelf edge and slope, whereas the interior shelf is characterized by comparatively weak and variable residual flows (Peliz et al., 2002; Otero et al., 2008). In these quiescent interior regions, MCC-derived velocity signal is often comparable to retrieval uncertainty arising from thermal noise in AVHRR imagery, residual atmospheric correction errors, and sub-pixel feature displacement uncertainty is of comparable magnitude to the current signal itself (Bowen et al., 2002; Emery et al., 1986). This results in a reduced signal-to-noise ratio in the derived currents, which naturally limits the attainable vector correlation with independent datasets, because these weak-flow regions occupy a substantial fraction of the study domain, their contribution lowers the domain-averaged $|\rho|$ below the level that would be observed over energetic shelf-edge regions alone. Second, the systematic overestimation of current magnitudes by MEnKF analysis relative to reanalysis product introduces magnitude discrepancies that reduce vector correlation even where directional agreement is strong. Third, differences in spatial resolution between MEnKF analyses and reanalysis products can produce spatial phase shifts in mesoscale features such as eddies and fronts, that would further degrade point-wise vector correlations even when both products capture similar circulation structures at slightly displace locations (Kerry et al., 2016; Oke et al., 2008). The concentration of correlation magnitudes within a narrow range (0.40–0.50) rather than broad dispersion indicates that these limitations are relatively consistent across time and space, without catastrophic failures or exceptional successes dominating the distribution (Evensen, 2003a). This temporal stability confirms that the assimilation framework maintains reliable baseline performance across variable observational conditions and flow regimes encountered during the study period. The absence of very low correlations ($|\rho| < 0.4$) demonstrates that even under suboptimal observational configurations, the MEnKF preserves sufficient constraint to avoid divergence from physically realistic circulation bounds.

The mean complex correlation phase angle $\phi = 5.97°$ confirms a negligible systematic counterclockwise rotation of MEnKF current vectors relative to the reanalysis reference, consistent with no operationally significant directional bias in the assimilation output. The moderate phase scatter (std = 33.58°) reflects representativeness differences between the two systems (Janjić et al., 2018). MCC-derived currents are kinematic displacements sensitive to the instantaneous SST gradient field at the time of image acquisition, whereas reanalysis velocities represent smoothened and depth averaged fields constrained by a broader observing system. These two representations of the same physical quantity can diverge in direction over short timescales without either being erroneous. The semidiurnal tidal forcing that characterizes the Galician shelf drives energetic rotary current signatures that can be an additional source of phase scatter (Otero et al., 2008). MCC image pairs spanning tidal phases inconsistent with the tidal interpolation applied within the reanalysis system can produce cycle-level directional offsets that manifest as scatter rather than systematic bias, directly consistent with the near-zero mean phase observed here. The near-absence of occurrences in the 90°–270° sector confirms that near-perpendicular or opposing flow directions between the two fields are rare, and that the MEnKF analysis maintains broadly consistent flow orientation with reanalysis reference across the majority of assimilation cycles.

The calm-weather fill value (0.001 m/s), while numerically stable and physically plausible as a quiescent baseline, constitutes the principal source of structural uncertainty in the assimilated observation operator. With data gaps of 10–30% across the domain, a non-trivial portion of the assimilated observation field comprised near-zero placeholders rather than genuine EO retrieval. These gaps arise from multiple source, e.g, persistent cloud cover blocking thermal infrared retrieval, insufficient thermal contrast between adjacent water masses reducing MCC tracking reliability, and AVHRR swath geometry leaving portions of the domain unsampled within any given assimilation window (Bowen et al., 2002; Minnett, 2003). Beyond magnitude underestimation, the spatial interleaving of filled and valid grid points damps local velocity gradients, attenuating sharp dynamical structures that carry the strongest assimilation signal and subtly distorting the spatial covariance structure on which the ensemble update relies (Cressie, 1993; Beckers and Rixen, 2003). Although validation metrics were restricted to grid points supported by valid EO observations, the direct influence of the fill strategy on ensemble state evolution and error covariance propagation cannot be entirely decoupled from the reported performance. More sophisticated gap-handling approaches, including masked assimilation, EOF-based reconstruction, or stochastic regularization, could reduce this uncertainty but would introduce additional computational and methodological complexity. The calm-weather baseline therefore represents a pragmatic compromise between physical realism, numerical stability, and operational feasibility, while identifying a clear direction for methodological refinement in future work.

Overall, the obtained results establish the MEnKF framework as a practical, operationally viable approach for NRT surface current estimation. The integration of SST-derived velocity observations with multi-model ensemble representation suc-

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

cessfully balances physical consistency against computational efficiency, achieving accuracy levels suitable for maritime applications while maintaining the low latency required for operational implementation. This validated framework provides a foundation for future enhancements including multi-sensor integration and coupled ocean-atmosphere assimilation strategies.

6 Conclusion and Future Work

This study demonstrates the operational feasibility of assimilating EO-derived surface current observations within a MEnKF framework to produce physically consistent, NRT ocean surface current analyses across the Galician shelf. By combining heterogeneous model ensembles with MCC-derived surface current observations from multi-platform AVHRR observation within a 3-hour assimilation cycle, the framework addresses a critical gap in operational ocean forecasting: the systematic underutilisation of EO-derived surface current observations that capture both geostrophic and ageostrophic surface flow components.

Independent validation against HFR measurements and reanalysis values collectively confirm that the framework delivers quantifiable and physically meaningful improvements over the background state, with consistent skill across a broad range of observational densities and circulation conditions. The stable analysis RMSE trajectory throughout the study time span (January 2021), despite pronounced episodic spikes in the background error, confirms that the multi-model ensemble structure provides robust error-reduction capacity even under variable and constrained observational input. The sensitivity analysis establishes that assimilation skill is governed by the spatial distribution and quality of EO observations rather than their volume alone, a finding with direct implications for operational observation selection strategy design. Optimal performance is consistently associated with appropriate observational densities and adequate spatial coverage, while performance degradation at higher densities reflects the compounding effects of spatially clustered observations, extended temporal windows spanning multiple tidal cycles, and finite ensemble size. Systematic diurnal performance variations, with degraded skill during midday hours and improved performance during morning and evening periods, are physically consistent with known AVHRR retrieval characteristics and Galician shelf tidal structure, confirming that these performance patterns are interpretable and therefore manageable within an operational context.

Residual uncertainties associated with MCC retrieval characteristics, the data gaps fill strategy, and resolution differences between compared products are physically interpretable and define the current performance envelope of the proposed MEnKF framework. The sensitivity assessment approach developed here, characterizing assimilation performance as a function of observational density, spatial coverage, and temporal sampling, provides a transferable diagnostic methodology applicable to ensemble DA systems beyond the present regional configuration and observation type. The quantified performance boundaries and identified sensitivity regimes establish a principled basis for extending the framework to additional EO-derived velocity products, alternative ensemble configurations, and broader spatial and seasonal domains, supporting the wider integration of kinematic surface current observations into operational ocean forecasting pipelines for applications including maritime safety, coastal hazard response, and environmental monitoring.

The prototype MEnKF configuration reported in this study establishes the methodological foundation required for subsequent multi-seasonal and multi-year extensions. Future work will focus on extending the framework to broader spatial and tem-

https://doi.org/10.5194/egusphere-2026-1683
Preprint. Discussion started: 7 April 2026

poral scales, including global implementations and subsequent multi-seasonal evaluation to assess performance across contrasting regions. Methodological improvements will include systematic testing of varying EO observation densities and sampling strategies, refinement of background error covariance through hybrid or machine-learning-informed approaches, and further improvement in adaptive localization to better account for flow anisotropy and evolving spatial correlation scales. The assimilation of additional EO-derived velocity products from complementary sensors including Synthetic Aperture Radar (SAR) and geostationary ocean color imagery, will be explored as a mechanism to enhance observational coverage and improved robustness under data gaps. Building on these observational extensions, machine learning will be incorporated at multiple levels of the assimilation framework, including adaptive ensemble member weighting based on flow-dependent performance, intelligent observation selection informed by retrieval quality metrics, regime-dependent bias correction, and dynamic parameter optimization to enhance assimilation efficiency and accuracy. Furthermore, advanced statistical and machine learning based gap filling techniques will be explored as replacements for the calm-weather fill value, providing more physically realistic observation and background states to improve framework robustness across diverse meteorological and oceanographic conditions. Finally, coupling the assimilated surface currents with Lagrangian trajectory models search-and-rescue operations, oil spill transport forecasting, and marine debris tracking, will provide a direct pathway from the present methodological demonstration to quantified operational impact and decision support.

Data availability. All datasets used in this study are publicly available, except in-house archived forecast data used for deriving climatology, which has not submitted to public repository. This dataset can be provided upon request to the corresponding author. Full details of the datasets, including key characteristics and retrieval procedures, are described in Sec. 2.2.

Author contributions. Shahbaz Baig led the study, developed the methodology, performed the analysis, and drafted the manuscript. The second author, Waqas A. Qazi, provided conceptual and technical guidance, data interpretation, and critical revisions of the manuscript. The third author, Rafia Mumtaz, provided conceptual guidance and supervision, contributed to project design, data interpretation, and manuscript review. All authors discussed the results and approved the final version of the manuscript.

Competing interests. The authors report that there are no competing interests to declare.

Acknowledgements. This research received partial funding from the European Union's Horizon Europe programme under the call HORIZON-CL5-2022-D6-01 (Grant Agreement No. 101077026, SafeNav). This work was also supported by the European Union's Horizon 2020 programme under Grant Agreement No. 820593 (EcoSail). The views and opinions expressed are those of the authors and do not necessarily reflect those of the European Union or CINEA, which cannot be held responsible for them. .

References

Alley, R. B., Emanuel, K. A., and Zhang, F.: Advances in weather prediction, Science, 363, 342–344, https://doi.org/10.1126/science.aav7274, 2019.

Alonso-martirena, A., Sanchez, J., Fernandez, V., Whelan, C., and Pederson, L.: Marine Information Services associated with HF radar observing networks, working Paper, 2011.

Alvera-Azcarate, A., Barth, A., Rixen, M., and Beckers, J.-M.: Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: DINEOF algorithm, Ocean Modelling, 9, 325–346, https://doi.org/10.1016/j.ocemod.2004.08.001, 2005.

Bach, E. and Ghil, M.: A multi-model ensemble Kalman filter for data assimilation and forecasting, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022MS003123, 2023.

Baduru, B., Chattopadhyay, R., Ravichandran, M., et al.: Ensemble Based Regional Ocean Data Assimilation System for the Indian Ocean: Implementation and Evaluation, Ocean Modelling, 144, 101 508, https://doi.org/10.1016/j.ocemod.2019.101508, 2019.

Baig, S., Qazi, W. A., and Mumtaz, R.: Near-Real-Time Earth Observation Data Assimilation for Ocean Currents, in: International Geoscience and Remote Sensing Symposium, pp. 5980–5984, IEEE, Athens, ISBN 979-8-3503-6032-5, https://doi.org/10.1109/IGARSS53475.2024.10641867, 2024.

Baig, S., Qazi, W. A., and Mumtaz, R.: Evaluating the Maximum Cross-Correlation Method for Near-Real-Time Ocean Currents Estimation Using Multi-Platform AVHRR Thermal Imagery, European Journal of Remote Sensing, –, –, under Review, 2025.

Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, Quarterly Journal of the Royal Meteorological Society, 134, 1951–1970, https://doi.org/10.1002/qj.339, 2008.

Bao, R., Varela, M., and Prego, R.: Mesoscale distribution patterns of diatoms in surface sediments as tracers of coastal upwelling of the Galician Shelf (NW Iberian Peninsula), Marine Geology, 144, 117–130, https://doi.org/10.1016/S0025-3227(97)00080-7, 1997.

Beckers, J.-M. and Rixen, M.: EOF Calculations and Data Filling from Incomplete Oceanographic Datasets, Journal of Atmospheric and Oceanic Technology, 20, 1839–1856, https://doi.org/10.1175/1520-0426(2003)020<1839:ECADFF>2.0.CO;2, 2003.

Bergemann, K. and Reich, S.: A Localization Technique for Ensemble Kalman Filters, Quarterly Journal of the Royal Meteorological Society, 136, 701–707, https://doi.org/10.1002/qj.591, 2010.

Bocquet, M. and Farchi, A.: Introduction to the principles and methods of data assimilation in the geosciences, Cerea, https://cerea.enpc.fr/HomePages/bocquet/teaching/assim-mb-en.pdf, 2023.

Bowen, M. M., Emery, W. J., Wilkin, J. L., Tildesley, P. C., Barton, I. J., and Knewtson, R.: Extracting Multiyear Surface Currents from Sequential Thermal Imagery Using the Maximum Cross-Correlation Technique, Journal of Atmospheric and Oceanic Technology, 19, 1665–1676, 2002.

Breivik, O. and Sætra, O.: Real time assimilation of HF radar currents into a coastal ocean model, Journal of Marine Systems, 28, 161–182, https://doi.org/10.1016/S0924-7963(01)00002-1, 2001.

Carvajal, G. K., Wozniak, M., Heuze, C., Eriksson, L. E. B., Johan, K., and Rydberg, B.: Assessment of satellite and ground-based estimates of surface currents, in: International Geoscience and Remote Sensing Symposium, pp. 4675–4678, IEEE, ISBN 978-1-5090-3332-4, https://doi.org/10.1109/IGARSS.2016.7730220, 2016.

Chabot, V., Nodet, M., Papadakis, N., and Vidard, A.: Accounting for observation errors in image data assimilation, Tellus, Series A: Dynamic Meteorology and Oceanography, 67, https://doi.org/10.3402/tellusa.v67.23629, 2015.

Choi, J. M., Kim, W., Hong, T. T. M., and Park, Y.-G.: Derivation and Evaluation of Satellite-Based Surface Current, Frontiers in Marine Science, 8, https://doi.org/10.3389/fmars.2021.695780, 2021.

CMEMS: Atlantic- European North West Shelf- Ocean Physics Reanalysis | Copernicus Marine Service, https://doi.org/https://doi.org/10.48670/moi-00059, 2024.

Coelho, E. F., Hogan, P., Jacobs, G., Thoppil, P., Huntley, H., Haus, B., Lipphardt, B., Kirwan, A., Ryan, E., Olascoaga, J., Beron-Vera, F., Poje, A., Griffa, A., Özgökmen, T., Mariano, A., Novelli, G., Haza, A., Bogucki, D., Chen, S., Curcic, M., Iskandarani, M., Judt, F., Laxague, N., Reniers, A., Valle-Levinson, A., and Wei, M.: Ocean current estimation using a Multi-Model Ensemble Kalman Filter during the Grand Lagrangian Deployment experiment (GLAD), Ocean Modelling, 87, 86–106, https://doi.org/10.1016/j.ocemod.2014.11.001, 2015.

Cressie, N. A. C.: Statistics for Spatial Data, John Wiley & Sons, New York, revised edn., ISBN 978-0-471-00255-0, 1993.

Donelan, M. A.: Air–sea interaction, in: The Sea: Ocean Engineering Science, edited by LeMehaute, B. and Hanes, D. M., vol. 9, pp. 239–292, Wiley, New York, 1990.

Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., and Murray, M. J.: Toward Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research, Journal of Climate, 15, 353–369, https://doi.org/10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2, 2002.

EcoSail Consortium: Eco-friendly and customer-driven Sail Plan Optimisation Service, European Commission, Horizon 2020 Framework Programme (Grant Agreement No. 820593), https://doi.org/10.3030/820593, 2018.

Emery, W. J., Thomas, A. C., Collins, M. J., Crawford, W. R., and Mackas, D. L.: An objective method for computing advective surface velocities from sequential infrared satellite images, Journal of Geophysical Research: Oceans, 91, 12 865–12 878, https://doi.org/10.1029/jc091ic11p12865, 1986.

Eriksson, L. E. B., Ye, Y., Jonasson, L., Qazi, W., Mao, W., Wang, H., Möller, J., Lemmens, K., and Dokken, S.: EONav - Copernicus Data in Support of Maritime Route Optimization, in: International Geoscience and Remote Sensing Symposium, pp. 5606–5609, https://doi.org/10.1109/IGARSS.2018.8519477, 2018.

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, 2003a.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003b.

Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, Springer, Berlin, Heidelberg, 1st edn., ISBN 978-3-540-38300-6, https://doi.org/10.1007/978-3-540-38301-3, 2006.

Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, Springer, Berlin, Heidelberg, 2nd edn., ISBN 978-3-642-03710-8, https://doi.org/10.1007/978-3-642-03711-5, 2009.

Friedemann, S.: Ensemble-based data assimilation for large scale simulations, Theses, Université Grenoble Alpes [2020-....], https://theses.hal.science/tel-03852854, 2022.

GEBCO Compilation Group: GEBCO 2025 Grid, https://doi.org/10.5285/37c52e96-24ea-67ce-e063-7086abc05f29, 2025.

Georges, T. M., Harlan, J. A., Leben, R. R., and Lematta, R. A.: A Test of Ocean Surface-Current Mapping with Over-the-Horizon Radar, Transactions on Geoscience and Remote Sensing, 36, 101–110, https://doi.org/10.1109/36.655321, 1998.

Gonzalez Haro, C., Isern-Fontanet, J., Tandeo, P., and Garello, R.: Ocean Surface Currents Reconstruction: Spectral Characterization of the Transfer Function Between SST and SSH, Journal of Geophysical Research: Oceans, 125, e2019JC015 958, https://doi.org/10.1029/2019JC015958, 2020.

Griffies, S. M., Adcroft, A. J., Hallberg, R. W., Holland, M. M., et al.: Challenges and Prospects in Ocean Circulation Models, Frontiers in Marine Science, 6, 65, https://doi.org/10.3389/fmars.2019.00065, 2019.

Hagedorn, R., Doblas-Reyes, F. J., and Palmer, T. N.: The rationale behind the success of multi-model ensembles in seasonal forecasting, Tellus A: Dynamic Meteorology and Oceanography, 57, 219–233, https://doi.org/10.3402/tellusa.v57i3.14657, 2005.

Hernandez-Lasheras, J., Mourre, B., Orfila, A., Santana, A., Reyes, E., and Tintoré, J.: Evaluating high-frequency radar data assimilation impact in coastal ocean operational modelling, Ocean Science, 17, 1157–1175, https://doi.org/10.5194/os-17-1157-2021, 2021.

Heuze, C., Carvajal, G. K., and Eriksson, L. E.: Optimization of sea surface current retrieval using a maximum cross-correlation technique on modeled sea surface temperature, Journal of Atmospheric and Oceanic Technology, 34, 2245–2255, https://doi.org/10.1175/JTECH-D-17-0029.1, 2017.

Hewitt, H. T., Bell, M. J., Chassignet, E. P., Czaja, A., et al.: Resolving and Parameterising the Ocean Mesoscale in Earth System Models, Current Climate Change Reports, 6, 137–152, https://doi.org/10.1007/s40641-020-00160-0, 2020.

Hoteit, I., Pham, D. T., and Blum, J.: A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical Pacific, Journal of Marine Systems, 36, 101–127, https://doi.org/10.1016/S0924-7963(02)00129-X, 2005.

Hoteit, I., Cornuelle, B., Kim, S.-J., Forget, G., Kohl, A., and Stammer, D.: Data Assimilation in Oceanography: Current Status and New Directions, in: Ocean Modeling in an Eddying Regime, edited by Hecht, M. W. and Hasumi, H., vol. 177 of *Geophysical Monograph Series*, pp. 237–292, American Geophysical Union, Washington, DC, https://doi.org/10.1029/177GM16, 2010.

Houtekamer, P. L. and Mitchell, H. L.: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation, Monthly Weather Review, 129, 123–137, https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2, 2001.

Intecmar: Galician HF-Radar Observation Thredd Server, https://opendap.intecmar.gal/thredds/catalog/radarhf/Galicia/LS/catalog.html, accessed: 2023-07-01, 2023.

Isern-Fontanet, J., Ballabrera-Poy, J., Turiel, A., and Garcia-Ladona, E.: Remote sensing of ocean surface currents: A review of what is being observed and what is being assimilated, Nonlinear Processes in Geophysics, 24, 613–643, https://doi.org/10.5194/npg-24-613-2017, 2017.

Janjić, T., Nerger, L., Albertella, A., Schröter, J., and Skachko, S.: On Domain Localization in Ensemble-Based Kalman Filter Algorithms, Monthly Weather Review, 139, 2046–2060, https://doi.org/10.1175/2011MWR3552.1, 2011.

Janjić, T., McLaughlin, D., Dance, S. L., Losa, S. N., Nichols, N. K., Potthast, R., Waller, J. A., and Weston, P.: On the Representation Error in Data Assimilation, Quarterly Journal of the Royal Meteorological Society, 144, 1257–1278, https://doi.org/10.1002/qj.3130, 2018.

Kalnay, E.: Historical overview of numerical weather prediction, Cambridge University Press, ISBN 0521791790, https://doi.org/10.1017/CBO9780511802270.002, 2002.

Kelly, K. A. and Strub, P. T.: Comparison of velocity estimates from advanced very high resolution radiometer in the coastal transition zone, Journal of Geophysical Research: Oceans, 97, 9653–9668, https://doi.org/10.1029/92JC00734, 1992.

Kerry, C., Powell, B., Roughan, M., and Oke, P.: Development and evaluation of a high-resolution reanalysis of the East Australian Current region using the Regional Ocean Modelling System (ROMS) and Ensemble Optimal Interpolation (EnOI), Geoscientific Model Development, 9, 3779–3801, https://doi.org/10.5194/gmd-9-3779-2016, 2016.

Kim, H.-Y., Park, J. H., Kim, Y.-J., et al.: Spatial Gap-Filling of GK2A Daily Sea Surface Temperature (SST) around the Korean Peninsula Using Meteorological Data and Regression Residual Kriging (RRK), Remote Sensing, 14, 5265, https://doi.org/10.3390/rs14205265, 2022.

King, R. R., Ubelmann, C., Le Sommer, J., et al.: Assimilating Realistically Simulated Wide-Swath Altimeter Observations in an Ocean Forecast System, Ocean Science, 17, 1791–1810, https://doi.org/10.5194/os-17-1791-2021, 2021.

Lahoz, W. A. and Schneider, P.: Data assimilation: making sense of Earth Observation, Frontiers in Environmental Science, 2, 1–28, https://doi.org/10.3389/fenvs.2014.00016, 2014.

Le Galloudec, O., Law Chune, S., Nouel, L., Fernandez, E., Derval, C., Tressol, M., Dussurget, R., Biardeau, A., and Tonani, M.: For Global Ocean Physical Analysis and Forecasting Product, Product User Manual, pp. 1–41, https://doi.org/10.48670/moi-00016, 2022.

Le Traon, P.-Y., Reppucci, A., Alvarez Fanjul, E., Aouf, L., Behrens, A., Benkiran, M., Bertino, L., Brando, V. E., et al.: From Observation to Information and Users: The Copernicus Marine Service Perspective, Frontiers in Marine Science, 6, 234, https://doi.org/10.3389/fmars.2019.00234, 2019.

Lecci, R., Drudi, M., Grandi, A., Cretì, S., and Clementi, E.: For Mediterranean Sea Physical Analysis and Forecasting Product, CMEMS, pp. 1–22, 2023.

Li, Y. and Toumi, R.: A balanced Kalman filter ocean data assimilation system with application to the South Australian Sea, Ocean Modelling, 116, 159–172, https://doi.org/10.1016/j.ocemod.2017.06.007, 2017.

Martinez Casas, D., Viqueira, J. R., Villarroya, S., Cotos, J. M., and Flores, J.: Interoperable Access and Interactive Exploration of HF Radar Data, IEEE Access, 12, 79 445–79 457, https://doi.org/10.1109/ACCESS.2024.3408682, 2024.

Matthews, D. K. and Emery, W. J.: Velocity observations of the California Current derived from satellite imagery, Journal of Geophysical Research: Oceans, 114, https://doi.org/10.1029/2008JC005029, 2009.

Minnett, P. J.: Radiometric measurements of the sea-surface skin temperature: the competing roles of the diurnal thermocline and the cool skin, International Journal of Remote Sensing, 24, 5033–5047, https://doi.org/10.1080/0143116031000095880, 2003.

Mirouze, I., Remy, E., Lellouche, J.-M., Bricaud, C., Law-Chune, S., Levier, B., and Huck, T.: Assimilation of satellite-derived total surface current velocities in the Mercator Ocean operational system: an observing system simulation experiment, Geoscientific Model Development, 17, 1955–1981, https://doi.org/10.5194/gmd-17-1955-2024, 2024.

Miyazawa, Y., Murakami, H., Miyama, T., Varlamov, S. M., Guo, X., Waseda, T., Sil, S., and Agency, E.: Data Assimilation of the High-Resolution Sea Surface Temperature Obtained from the Aqua-Terra Satellites (MODIS-SST) Using an Ensemble Kalman Filter, Remote Sensing, 5, 3123–3139, https://doi.org/10.3390/rs5063123, 2013.

Moore, A. M., Martin, M. J., Akella, S., Arango, H. G., Balmaseda, M., Bertino, L., Ciavatta, S., Cornuelle, B., Cummings, J., Frolov, S., Lermusiaux, P., Oddo, P., Oke, P. R., Storto, A., Teruzzi, A., Vidard, A., and Weaver, A. T.: Synthesis of ocean observations using data assimilation for operational, real-time and reanalysis systems: A more complete picture of the state of the ocean, Frontiers in Marine Science, 6, 1–6, https://doi.org/10.3389/fmars.2019.00090, 2019.

Navarro, J. S., Jimenez, P. L., Gil, M. I. R., Serna, D., and Fanjul, E. A.: The OPERA project : An operational HF radar network in the Spanish coast, in: Encuentro Oceonografia Fisica Espanola, p. 12, 2014.

Nerger, L., Schröter, J., and Hiller, W.: On the influence of model nonlinearity and localization schemes on EnKF performance, Quarterly Journal of the Royal Meteorological Society, 138, 1977–1991, https://doi.org/10.1002/qj.1922, 2012.

Nguyen, H., Van Nguyen, H., Bui, Q.-T., et al.: A Machine Learning Approach for Remote Sensing Data Gap-Filling with Open-Source Implementation: An Example Regarding Land Surface Temperature, Surface Albedo and NDVI, Remote Sensing, 12, 3865, https://doi.org/10.3390/rs12233865, 2020.

NOAA: AVHRR Guide book, https://www.aev.class.noaa.gov/glossary/AVHRR.htm, 2024a.

NOAA: Comprehensive Large Array-data Stewardship System (CLASS), https://www.avl.class.noaa.gov, accessed: 2024-02-01, 2024b.

Nogueira, E., Ibanez, F., and Figueiras, F. G.: Spatial variability of the barotropic M2 constituent tidal current over the Rías Baixas Galician shelf (NW Spain), Estuarine, Coastal and Shelf Science, 74, 631–648, https://doi.org/10.1016/j.ecss.2007.05.007, 2007.

NWC, N.: Global Real-Time Ocean Forecast System, https://polar.ncep.noaa.gov/global/, 2023.

Ohishi, S., Miyazawa, Y., Varlamov, S. M., Guo, X., Miyama, T., et al.: An Ensemble Kalman Filter System with the Stony Brook Parallel Ocean Model v1.0, Geoscientific Model Development, 15, 8395–8416, https://doi.org/10.5194/gmd-15-8395-2022, 2022.

Oke, P. R., Allen, J. S., Miller, R. N., Egbert, G. D., and Kosro, P. M.: Assimilation of surface velocity data into a primitive equation coastal ocean model, Journal of Geophysical Research: Oceans, 107, 3122, https://doi.org/10.1029/2000JC000511, 2002.

Oke, P. R., Schiller, A., Griffin, D. A., and Brassington, G. B.: The Bluelink Ocean Data Assimilation System (BODAS), Ocean Modelling, 21, 46–70, https://doi.org/10.1016/j.ocemod.2007.11.002, 2008.

Otero, P., Ruiz-Villarreal, M., and Peliz, A.: Variability of river plumes off Northwest Iberia in response to wind events, Journal of Marine Systems, 72, 238–255, https://doi.org/10.1016/j.jmarsys.2007.05.016, 2008.

Park, S. K. and Xu, L.: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III), vol. III, Springer International Publishing, Cham, ISBN 978-3-319-43414-8, https://doi.org/10.1007/978-3-319-43415-5, 2017.

Peliz, A., Rosa, T. L., Santos, A. M. P., and Pissarra, J. L.: Fronts, jets, and counter-flows in the Western Iberian upwelling system, Journal of Marine Systems, 35, 61–77, https://doi.org/10.1016/S0924-7963(02)00076-3, 2002.

Penny, S. G. et al.: Observational needs for improving ocean and coupled data assimilation and prediction systems, Frontiers in Marine Science, 6, 391, https://doi.org/10.3389/fmars.2019.00391, 2019.

Piracha, A. et al.: Using Satellite Observations of Ocean Variables to Improve Estimates of Mixed-Layer Density Flux, Frontiers in Marine Science, 10, 1020 153, https://doi.org/10.3389/fmars.2023.1020153, 2023.

Qazi, W. A., Emery, W. J., and Fox-Kemper, B.: Computing ocean surface currents over the coastal california current system using 30-min-lag sequential SAR images, Transactions on Geoscience and Remote Sensing, 52, 7559–7580, https://doi.org/10.1109/TGRS.2014.2314117, 2014.

Rahman, R. et al.: Evaluation of sea surface temperature from ocean reanalyses over the Indian Ocean, Frontiers in Marine Science, 11, 1461 696, https://doi.org/10.3389/fmars.2024.1461696, 2024.

Robinson, A. R. and Pierre, F. J. L.: Overview of Data Assimilation, Tech. rep., Harvard University, Cambridge, Massachusetts, 2000.

Röhrs, J. et al.: Surface Currents in Operational Oceanography: Key Applications, Mechanisms, and Observation Techniques, Journal of Operational Oceanography, 14, s1–s26, https://doi.org/10.1080/1755876X.2021.1903221, 2021.

Rusvan, A., Maricar, F., Thaha, M., and Paotonan, C.: Evaluation of Tidal Energy Potential Using a Two-Way Tidal Energy Model, Civil Engineering Journal, 10, 3011–3033, https://doi.org/10.28991/CEJ-2024-010-09-016, 2024.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., and Korablev, A.: TOPAZ4: An Ocean–Sea Ice Data Assimilation System for the North Atlantic and Arctic, Ocean Science, 8, 633–656, https://doi.org/10.5194/os-8-633-2012, 2012.

Sun, M., Yin, X., Yang, Y., and Wu, K.: An effective method based on dynamic sampling for data assimilation in a global wave model, Ocean Dynamics, 67, 433–449, https://doi.org/10.1007/s10236-017-1030-y, 2017.

Sun, M., Yang, Y., Yin, X., and Du, J.: Data assimilation of ocean surface waves using Sentinel-1 SAR during typhoon Malakas, International Journal of Applied Earth Observation and Geoinformation, 70, 35–42, https://doi.org/10.1016/j.jag.2018.04.004, 2018.

Talagrand, O.: Assimilation of observations, an introduction, in: ECMWF Seminar on Data Assimilation, pp. 59–68, ECMWF, 1997.

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

Villasante, S., Tubio, A., Ainsworth, G., Pita, P., Antelo, M., and Da-Rocha, J. M.: Rapid Assessment of the COVID-19 Impacts on the Galician (NW Spain) Seafood Sector, Frontiers in Marine Science, 8, https://doi.org/10.3389/fmars.2021.737395, 2021.

Wallcraft, A., Carroll, S. N., Kelly, K. A., and Rushing, K. V.: Hybrid Coordinate Ocean Model (HYCOM) Version 2.1 User's Guide, Tech. Rep. ADA588120, Naval Research Laboratory, Stennis Space Center, MS, https://apps.dtic.mil/sti/pdfs/ADA588120.pdf, nRL Technical Report, 2003.

Wang, G. and Pan, Y.: Phase-resolved ocean wave forecast with ensemble-based data assimilation, Journal of Fluid Mechanics, 918, 1–27, https://doi.org/10.1017/jfm.2021.340, 2021.

Wang, W. et al.: Ocean surface currents estimated from satellite remote sensing data on a global isotropic hexagonal grid, International Journal of Digital Earth, 16, 1–24, https://doi.org/10.1080/17538947.2023.2192003, 2023.

Waters, J., Martin, M. J., Bell, M. J., King, R. R., Gaultier, L., Ubelmann, C., Donlon, C., and Van Gennip, S.: Assessing the potential impact of assimilating total surface current velocities in the Met Office's global ocean forecasting system, Frontiers in Marine Science, Volume 11 - 2024, https://doi.org/10.3389/fmars.2024.1383522, 2024.

Weiss, J., Miralles, D. G., Gentine, P., et al.: CLIMFILL v0.9: a Framework for Intelligently Gap Filling Earth Observations, Geoscientific Model Development, 15, 4569–4592, https://doi.org/10.5194/gmd-15-4569-2022, 2022.

Xue, L. and Zhang, D.: A multimodel data assimilation framework via the ensemble Kalman filter, Water Resources Research, 50, 4197–4219, https://doi.org/10.1002/2013WR014525, 2014.

Yang, H., Arnone, R., and Jolliff, J.: Estimating advective near-surface currents from ocean color satellite images, Remote Sensing of Environment, 158, 1–14, https://doi.org/10.1016/j.rse.2014.11.010, 2015.