data

MDPI

Data Descriptor

A Combined HF Radar and Drifter Dataset for Analysis of Highly Variable Surface Currents

Bartolomeo Doronzo [1,2,*], Michele Bendoni [3], Stefano Taddei [2], Angelo Boccacci [3] and Carlo Brandini [1,4]

[1] Istituto di Scienze Marine ISMAR-CNR, 50019 Firenze, Italy; carlo.brandini@cnr.it
[2] Consorzio LaMMA, 57126 Livorno, Italy; taddei@lamma.toscana.it
[3] Istituto di Scienze Marine ISMAR-CNR, 19032 Lerici, Italy; michele.bendoni@cnr.it (M.B.); angelo.boccacci@cnr.it (A.B.)
[4] Consorzio LaMMA, 50019 Firenze, Italy
* Correspondence: bartolomeo.doronzo@cnr.it or doronzo@lamma.toscana.it

Abstract

This data descriptor presents the HF radar and drifter datasets, along with the methods used to process and apply them in a previously published study on the validation of surface current measurements in a region characterized by highly variable coastal dynamics. The data were collected in the framework of a large-scale Lagrangian experiment, which included extensive drifter deployment and the generation of virtual trajectories based on HF radar-derived flow fields. Both Eulerian and Lagrangian approaches were used to assess radar performance through correlation and RMSE metrics, with additional refinement achieved via Kriging interpolation. The validation results, published in Remote Sensing, demonstrated good agreement between HF radar and drifter observations, particularly when quality control parameters were optimized. The datasets and associated methodologies described here support ongoing efforts to enhance HF radar tuning strategies and improve surface current monitoring in complex marine environments.

Dataset: https://doi.org/10.5281/zenodo.15593063; https://doi.org/10.17882/105936

Dataset License: CC-BY 4.0

Keywords: drifter measurements; HF radar; surface currents; radar validation; Tuscany Archipelago

check for updates

Academic Editors: Milan S. Dimitrijević, Vladimir Sreckovic and Zoran Mijic

Received: 5 June 2025
Revised: 8 July 2025
Accepted: 10 July 2025
Published: 12 July 2025

Citation: Doronzo, B.; Bendoni, M.; Taddei, S.; Boccacci, A.; Brandini, C. A Combined HF Radar and Drifter Dataset for Analysis of Highly Variable Surface Currents. *Data* **2025**, *10*, 115. https://doi.org/10.3390/data10070115

1. Summary

The dataset described in this manuscript was collected during a large-scale Lagrangian oceanographic campaign conducted in October and November 2020 in the Tuscany Archipelago [1], a region characterized by highly variable surface currents. It includes measurements from 26 CODE-type drifters deployed in a square array and tracked via GPS with a 5 min sampling interval, as well as surface current data from a SeaSonde network consisting of three coastal stations.

This dataset contributes to the validation of remote sensing technologies for surface current monitoring in coastal marine environments, supporting cross-border and regional observational efforts within European frameworks such as CMEMS and the EuroGOOS HF Radar Task Team. By providing reliable HF radar and drifter data from this dynamic coastal region, it supports improvements in operational oceanographic services and coastal management.

Radar data were subjected to a series of quality control (QC) procedures designed to maximize the accuracy and consistency of surface current vectors. Two key post-processing filters, VART_QC (temporal variance filter) and MDFL_QC (median filter), as defined by CMEMS guidelines [2], identify and flag vectors exhibiting excessive temporal variability or significant deviation from the local median. Alternative thresholds to the default value of 1 m/s were applied to assess the sensitivity of the dataset to progressively stricter quality control criteria. Radar datasets corresponding to these thresholds are provided in the Supplementary Materials.

2. Data Description

The HF radar dataset includes hourly surface current measurements derived from a CODAR SeaSonde Direction Finding system operating at 13.5 MHz. The network consists of three stations, Livorno (LIVO), San Vincenzo (SVIN), and Isola del Tino (TINO), covering the Ligurian and Northern Tyrrhenian Seas. Radial velocities are collected on a polar grid with approximately 1.5 km radial resolution and 5° angular spacing, then combined into total velocity fields using a least-squares method. Measurements represent surface currents in the upper ~0.3–2.5 m of the water column and are aggregated over 1 h intervals centered on the cardinal hour. The data are georeferenced using EPSG:4326 (WGS84), with a geographic extent roughly between 41.5 and 44.5° N as well as 8.5 and 12.5° E.

Data are stored in NetCDF-4 classic model format and conform to CF-1.6 and Copernicus In Situ-TAC conventions [3]. Quality control procedures include several flagging tests: VART_QC (temporal variance filter), MDFL_QC (median filter), CSPD_QC (velocity threshold), OWTR_QC (over-water check), AVRB_QC (average bearing), and RDCT_QC (radial count) [4,5]. Both filtered and unfiltered versions of the radar data are available. Filtered datasets correspond to different configurations of the VART_QC and MDFL_QC filters, with thresholds ranging from 0.182 to 1 m/s. These variants allow users to assess the sensitivity of spatial coverage and data quality to different quality control settings.

The Lagrangian component of the dataset is based on 26 CODE-type surface drifters deployed on 8 October 2020 in a regular-grid configuration within the Tuscany Archipelago, as part of a larger multi-instrument experiment described in [6]. Drifters were tracked via GPS with a nominal sampling interval of 5 min. The drifter dataset spans approximately one month, although effective durations vary due to recovery operations, grounding events, or transmission interruptions. Each record includes time, latitude, longitude, and computed velocity components (u, v) obtained from the displacement between consecutive positions. Due to occasional gaps in GPS transmission, primarily caused by a firmware-induced standby behavior under low-motion conditions, a filtering and cleaning process was applied to remove intervals exceeding expected transmission time or associated with pre- and post-deployment phases. The final cleaned trajectories are provided in NetCDF format with metadata including deployment time, drifter ID, and geospatial bounds.

To enhance spatial and temporal continuity, Kriging interpolation [7,8] was applied to drifter velocities, generating regular-grid fields suitable for comparison with radar data. These interpolated products support applications such as trajectory modeling and coastal circulation studies.

Synthetic Lagrangian trajectories were also generated using the GNOME (General NOAA Operational Modeling Environment) model, driven by radar-derived surface velocity fields. These simulations complement the drifter data and are provided in NetCDF format with metadata including initialization time, reference drifter ID, and radar coverage boundary.

All data components, including raw and filtered radar fields, cleaned drifter tracks, Kriging-interpolated velocities, and synthetic trajectories, are available as NetCDF files with standardized metadata. Filtered radar data, interpolated drifter fields, and synthetic trajectories are included as Supplementary Materials to this article, while the complete radar dataset (including raw fields) is openly accessible via Zenodo and the European HFR THREDDS server. A summary of the dataset structure, instruments, and derived products is provided in Table 1.

Table 1. Summary of dataset contents, including HF radar system specifications, drifter deployment details, and processed products such as Kriged fields and synthetic trajectories.

Component	Description
HF radar sites	3 SeaSonde stations at TINO, LIVO, SVIN (13.5 MHz)
Radar data format	NetCDF with CMEMS-compliant metadata
Radar coverage area	Northern Tyrrhenian and Ligurian Sea (~80–100 km per station)
Drifter type	26 CODE-type drifters with 5 min GPS sampling
Deployment date	8 October 2020
Drifter data duration	Up to 1 month (variable per drifter)
Trajectory sampling interval	~5 min (variable due to firmware-induced standby mode)
Interpolated drifter velocities	Interpolated using Ordinary Kriging with an exponential variogram (range = 100, sill = 1, nugget = 0.1)
Synthetic GNOME trajectories	705 sub-tracks of 12 h each, based on radar-derived currents, provided as NetCDF
Real sub-trajectories (12 h)	Used for validation only; not included in Supplementary Materials
Auxiliary formats	All trajectory and velocity data available in NetCDF

Note: Filtered radar data, Kriging-interpolated drifter velocities, and synthetic trajectories are available as Supplementary Materials accompanying this article. The full radar dataset, including both filtered and unfiltered versions, is openly accessible via Zenodo and the European HFR THREDDS server (see Dataset section for links).

3. Methods

The methodological framework integrates Eulerian and Lagrangian approaches to assess the consistency between HF radar-derived currents and drifter observations. For Eulerian validation, surface current vectors derived from GPS-tracked, CODE-type drifters were projected onto the radar beam directions and compared with co-located radial velocity components. Statistical metrics such as Pearson correlation and root mean square error (RMSE) were computed across multiple stations and time intervals to quantify agreement and identify spatial or temporal patterns of deviation.

Lagrangian validation involved simulating virtual drifter trajectories using the GNOME (General NOAA Operational Modeling Environment) model [9], driven by hourly HF radar-derived velocity fields. The simulations were initialized at the start of each 12-h segment of the real drifter tracks and restricted to the radar coverage area, defined by the dynamically extracted footprint polygon. Radar data gaps shorter than 3 h were linearly interpolated, while longer gaps led to re-initialization from the next available drifter position. This approach yielded 705 matched pairs of observed and modeled sub-trajectories, enabling consistent statistical comparison across diverse current regimes.

To enhance the spatial coherence of radar velocities and improve comparison with in situ data, Ordinary Kriging was applied to interpolate radar-derived current fields at the exact times and locations of the drifter observations. Using an exponential variogram model with tuned parameters (sill = 1, range = 100, nugget = 0.1), interpolated current components (u, v) were estimated within a local buffer around each drifter position. This geostatistical method reduced sensitivity to local gaps or anomalies and was particularly effective in mitigating the effect of partial spatial coverage during specific time intervals.

Overall, the combination of direct velocity comparisons, trajectory-based validation, and spatial interpolation provides a robust, multi-dimensional assessment of the radar system's performance and supports its application in coastal circulation studies and operational modeling.

Further methodological details, including sub-track segmentation, trajectory matching, and validation procedures, are described in [10]. All Kriging-interpolated velocity fields, radar data, and synthetic trajectories are provided as Supplementary Materials accompanying this article.

4. User Notes

Users should note that radar data are sensitive to coastal geometry and signal interference. The dataset includes filtered and unfiltered versions of the radar outputs. Drifter trajectories contain occasional gaps due to communication issues. Recommended tools for analysis include GNOME, MATLAB (R2023a), and Python 3.12 with NetCDF libraries.

To assess the sensitivity of the dataset to different quality control configurations, three versions of the radar-derived total velocity fields were produced by varying the thresholds of the VART_QC (temporal variance filter) and MDFL_QC (median filter) parameters within the range from 0.182 to 1.0. These thresholds control the aggressiveness of the filtering process applied to the radial vectors before combining them into total velocity fields. Lower thresholds remove a greater number of vectors, potentially improving consistency with in situ drifter measurements while reducing the spatial coverage of the radar fields.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/data10070115/s1. The supplementary materials include the file **kriging_velocity_fields.nc**, containing Kriging-interpolated surface current velocity fields (u, v) in NetCDF4 format (approximately 20 MB); the folder **VirtualDrifterDataset/**, which contains synthetic drifter trajectories generated using GNOME in NetCDF4 format (approximately 40 MB); and the file **readme_supplementary_data.txt**, which provides documentation of the supplementary dataset in plain text format (approximately 8 KB).

Author Contributions: Conceptualization, B.D., C.B., S.T., and M.B.; methodology, B.D., C.B., S.T., and M.B.; software, B.D.; validation, B.D., S.T., M.B., and A.B.; formal analysis, B.D., C.B., S.T., and M.B.; investigation, B.D., C.B., S.T., and M.B.; resources, B.D., C.B., S.T., M.B., and A.B.; data curation B.D., S.T., and M.B.; writing—original draft preparation, B.D.; writing—review and editing, B.D., C.B., S.T., and M.B.; visualization, B.D., M.B., and S.T.; supervision, C.B.; project administration, C.B.; funding acquisition, C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The drifter dataset is publicly available and fully documented online: https://doi.org/10.17882/105936. The filtered HF radar surface current dataset, including quality-control thresholds, is openly accessible on Zenodo: https://doi.org/10.5281/zenodo.15593063. An extended set of raw HF radar surface current fields is available via the European HFR data-sharing THREDDS server, maintained by the EuroGOOS HFR Task Team: https://thredds.hfrnode.eu:8443/thredds/NRTcurrent/HFR-LaMMA/HFR-LaMMA_catalog.html (accessed on 9 July 2025).

Acknowledgments: Collection of data supported by the LaMMA Consortium and the IMPACT (IMpact of Ports on marine protected areas: Cooperative Cross-Border Actions, included in the program 'Interreg Italia-Francia Marittimo 2014–2020' (EU funded, PC Interreg VA IFM 2014–2020). Deployment logistics were carried out with the support of NATO-CMRE (SP, Italy), CNR-ISMAR

(Italy), SIO/UCSD (USA), and the Italian Coast Guard (Capitaneria di Porto of Livorno). In particular, we acknowledge Pierre Poulain (OGS), Luca Centurioni (Scripps), Marina Ampolo Rella (CMRE), and Maristella Berta (CNR-ISMAR), for the support given, in terms of experience, contribution, to the planning and involvement in the oceanographic campaign 'Drifter Demonstration and Research (DDR20)-Sea Trial Plan'.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

CMEMS	Copernicus Marine Environment Monitoring Service
CODE	Coastal Ocean Dynamics Experiment
CF	Climate and Forecast (metadata convention)
EPSG	European Petroleum Survey Group (standardized coordinate reference codes)
GNOME	General NOAA Operational Modeling Environment
GPS	Global Positioning System
HF	High Frequency
HFR	High-Frequency Radar
ID	Identifier
NetCDF	Network Common Data Form
NOAA	National Oceanic and Atmospheric Administration
QC	Quality Control
RMSE	Root Mean Square Error
SVIN	San Vincenzo (HF radar site)
LIVO	Livorno (HF radar site)
TINO	Isola del Tino (HF radar site)
u, v	Zonal (east–west) and meridional (north–south) velocity components

References

1. Doronzo, B.; Poulain, P.-M.; Taddei, S.; Boccacci, A.; Brandini, C. CODE drifter observations from the DDR20 experiment in the NW Mediterranean Sea. *SEANOE* **2025**. [CrossRef]
2. Copernicus Marine Environmental Monitoring Service (CMEMS). *QA Best Practices and Protocols on QC for Radial and Total HF Radar Data*; Copernicus Marine Environmental Monitoring Service: Ramonville-Saint-Agne, France, 2017.
3. Copernicus Marine In Situ TAC Data Management Team. Copernicus Marine In Situ TAC NetCDF Format Manual. Technical Report. 2023. Available online: https://archimer.ifremer.fr/doc/00488/59938/ (accessed on 9 July 2025).
4. Mantovani, C.; Corgnati, L.; Horstmann, J.; Rubio, A.; Reyes, E.; Quentin, C.; Cosoli, S.; Asensio, J.L.; Mader, J.; Griffa, A. Best practices on high frequency radar deployment and operation for ocean current measurement. *Front. Mar. Sci.* **2020**, *7*, 210. [CrossRef]
5. Cosoli, S.; Bolzon, G. A real-time and offline quality control methodology for SeaSonde high-frequency radar currents. *J. Atmos. Ocean. Technol.* **2012**, *29*, 1313–1328. [CrossRef]
6. Poulain, P.-M.; Centurioni, L.; Brandini, C.; Taddei, S.; Berta, M.; Menna, M. Relative dispersion and kinematic properties of the coastal submesoscale circulation in the southeastern Ligurian Sea. 2023. Available online: https://os.copernicus.org/articles/19/1617/2023/ (accessed on 9 July 2025).
7. Cressie, N.A.C. *Statistics for Spatial Data, Wiley Series in Probability and Statistics*; John Wiley & Sons: New York, NY, USA, 1993; pp. 1–900. [CrossRef]
8. Allard, D.; Chilès, J.-P.; Delfiner, P. Geostatistics: Modeling spatial uncertainty. *Math. Geosci.* **2013**, *45*, 377–380. [CrossRef]

9. Beegle-Krause, C.J. General NOAA Oil Modeling Environment (GNOME): A new spill trajectory model. In Proceedings of the International Oil Spill Conference Proceedings 2001, Tampa, FL, USA, 26–29 March 2001; pp. 865–871. [CrossRef]
10. Doronzo, B.; Bendoni, M.; Taddei, S.; Boccacci, A.; Brandini, C. Validating HF radar current accuracy via Lagrangian measurements and radar-to-radar comparisons in highly variable surface currents. *Remote Sens.* **2025**, *17*, 1243. [CrossRef]