World Meteorological Organization &

Intergovernmental Oceanographic Commission (of UNESCO)

JOINT WMO/IOC TECHNICAL COMMISSION FOR OCEANOGRAPHY AND MARINE METEOROLOGY

Data Buoy Cooperation Panel Thirty Seventh Session 1-4 November 2022

DBCP-38/Doc. 9.8

Submitted by:
T. Carval

22.10.2022

DOI 10.13155/91109
Version 1.0

AGENDA ITEM 9: **Recommendations by the Task Teams**

AGENDA ITEM 9.5.2: **Recommendations by Task Team Buoy Data Management Centers**

SUMMARY

This document provides a report by the Task Team on Buoy Data Management Centers including actions/decisions required. Report will focus on challenges, opportunities and risks, and derived recommendations.

(Draft text for inclusion in the final report)

A. INTRODUCTION/SUMMARY[1]:

(maximum length half (1/2) a page)

In the Marine Climate Data System (MCDS) scheme, two Drifting Data Buoys Global Data Assembly Centre (DDB GDAC), have been established; lead by "Fisheries and Oceans Canada" (DFO) and Coriolis (French organization including Météo-France and Ifremer).

Both GDACs acquire data circulated on the Global Telecommunication System (GTS) of WMO. Additionally, Coriolis acquires data from Copernicus Marine service. The GDACs aim to consolidate near-real-time and delayed-mode data, to deliver a comprehensive best version archive of data and metadata.

Both GDACs routinely compare GTS bulletin headings and data volume received and have ways to make data available to requesters (DFO: through an offline form request system, France: through the Copernicus Marine Environment Monitoring Service: http://www.marineinsitu.eu/access-data/.

In addition, a public FTP server distributing data and metadata as one file per drifting buoy is operated in France: ftp://ftp.ifremer.fr/ifremer/dbcp-drifter/ (NetCDF format), and Canada distributes data grouped in monthly files on their ftp server: ftp://ftp.meds-sdmm.dfo-mpo.gc.ca/pub/dribu_bufr in CSV format.

The two DDB GDAC are working together to establish DDB GDAC organization document and user manual, to formulate exchanges and best practices between the two GDACs.

[1] Half a page or less of Summary

B. ACTIONS/DECISIONS[2] REQUIRED:

(a) Adopt draft Action(Decision)[2] 9.5.2/1 —*Action(Decision) title;*

- Drifter DACs are to forward at regular intervals (either in near-real-time or in delayed mode) the Iridium SBD drifter messages to the Coriolis DAC.
- Participants to try out the new services provided by the GDACs and provide feedback as needed.

C. RECOMMENDATIONS[3] :

(a) Adopt draft Recommendation[3] 0.0.0/1 — *Recommendation title;*

- What, To who, Timeline
- Rational

[2] An Action/Decision is an item directly related to DBCP and on which DBCP can action or decide directly. Details on rational for the action/decision should be included in the Background under Draft Actions/Decisions.

[3] A Recommendation involves proposed action(s)on another body outside of DBCP (e.g. SOT, JCOMM, WMO, IOC, CBS etc.). Details on rational for the Recommendation should be included in the Background under Draft Recommendation.

C. BACKGROUND INFORMATION (not to be included in the session report):

References (if any):

1. *[Link to the full report on the website]*

2. .....

Draft Actions/Decisions

- Participants to review the two GDAC documents (Drifter data management and Proposal for drifting buoy metadata in the WIGOS Metadata Standard) and send comments to the authors by 1 November 2022.

Draft Recommendations

[Comment: Details on main points and arguments leading to formulation of draft actions/decision presented in this document]

French Drifting Buoys Global Data Assembly Center report

The migration to a GDAC for Drifting Data Buoys, through the MCDS implementation phase II, involves Météo-France, Ifremer (Coriolis Data Center) and Fisheries and Oceans Canada.

A daily collation and archiving of buoy reports from the global ocean is performed by Météo-France. Collaboration within the Coriolis project (www.coriolis.eu.org), with OceanOPS and also CLS-Argos are main aspects of this FR DDB GDAC, beside regular exchanges with other data centres, measurement teams and agencies, and with users.

The global data buoy callation activity is strongly supported by EU Copernicus service.

Data products

Copernicus Marine service (https://marine.copernicus.eu) aggregates drifting buoys observations from DB-GDAC and WMO-GTS.

Copernicus In Situ service (https://insitu.copernicus.eu) supports delayed mode activity on drifting buoy data (data retrieval "archeology" and scientific assessment).

Drifting buoys global dataset

- INSITU_GLO_NRT_OBSERVATIONS_013_030 in situ data product
  Real-time drifting buoys: the last 30 days of 1.490 active drifting buoys
  One NetCDF file per day and per platform
  From ERDDAP server:
  https://nrt.cmems-du.eu/erddap/tabledap/copernicus_GLO_insitu_nrt_DB.html
  From ftp server:
  ftp://***@nrt.cmems-du.eu/
  Core/INSITU_GLO_NRT_OBSERVATIONS_013_030/nrt/latest/*/*_TS_DB*.nc

- History drifting buoys: all observations from 14.015 drifting buoys
  One NetCDF file per drifting buoy
  ftp://***@nrt.cmems-du.eu/
  Core/INSITU_GLO_NRT_OBSERVATIONS_013_030/nrt/history/DB/*.nc

Drifting buoys surface currents product

The surface currents product derived from drifting buoys data are produced in real-time and delayed mode by EU Copernicus Marine service.

Surface current product, real time

Since the 1st of January 2002, Coriolis (Météo-France and Ifremer) produces a **weekly surface current data** calculated from SVP drifter tracks.

This product is distributed by Copernicus Marine service as:

- INSITU_GLO_UV_NRT_OBSERVATIONS_013_048 product
  Active buoys (the last 30 days)
  ftp://nrt.cmems-du.eu/Core/INSITU_GLO_UV_NRT_OBSERVATIONS_013_048/drifter/latest
  YYYYMMDD/GL_TS_DC_*_YYYYMMDD.nc

- All buoys ("history")
  ftp://nrt.cmems-

du.eu/Core/INSITU_GLO_UV_NRT_OBSERVATIONS_013_048/drifter/history
GL_TS_DC_*.nc

The Copernicus NetCDF CF format is documented in:

- *Copernicus Marine in situ NetCDF format reference manual* *https://doi.org/10.13155/59938*

Surface current products, delayed mode

For the Global Ocean delayed mode in-situ observations of surface currents, use the following links to access to:

- *Rio Marie-Hélène, Etienne Hélène (2019). Copernicus Global Ocean delayed mode in-situ observations of ocean surface currents http://doi.org/10.17882/41334*
- Product user manual: http://doi.org/10.13155/41257
- Quality Information Document: http://doi.org/10.13155/41256

This product is updated once a year.

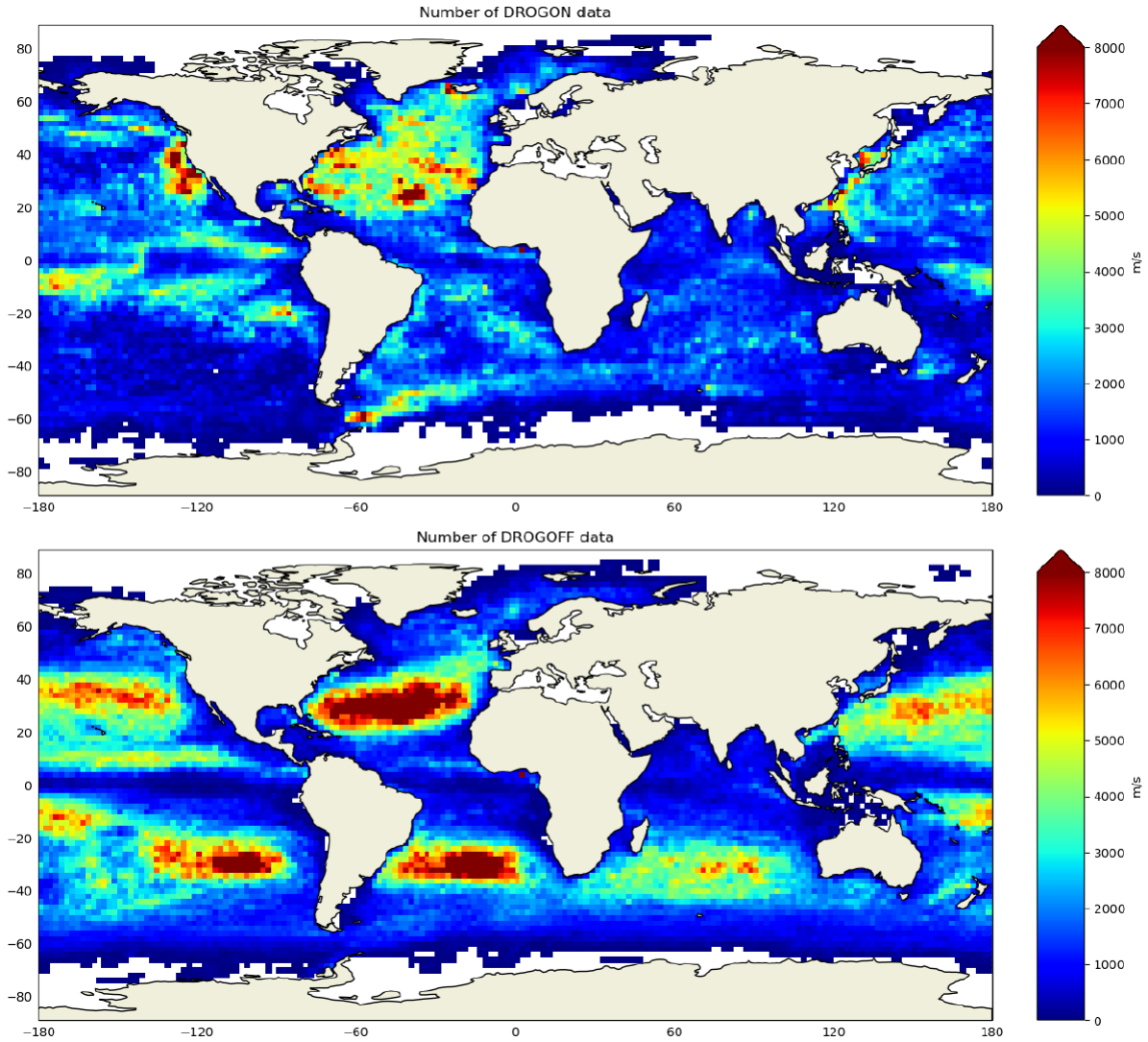


*Figure 1: 2*2° bin mean number of "drogued" (upper) and undrogued (lower) drifter measurements in 1990 - 2021 period*

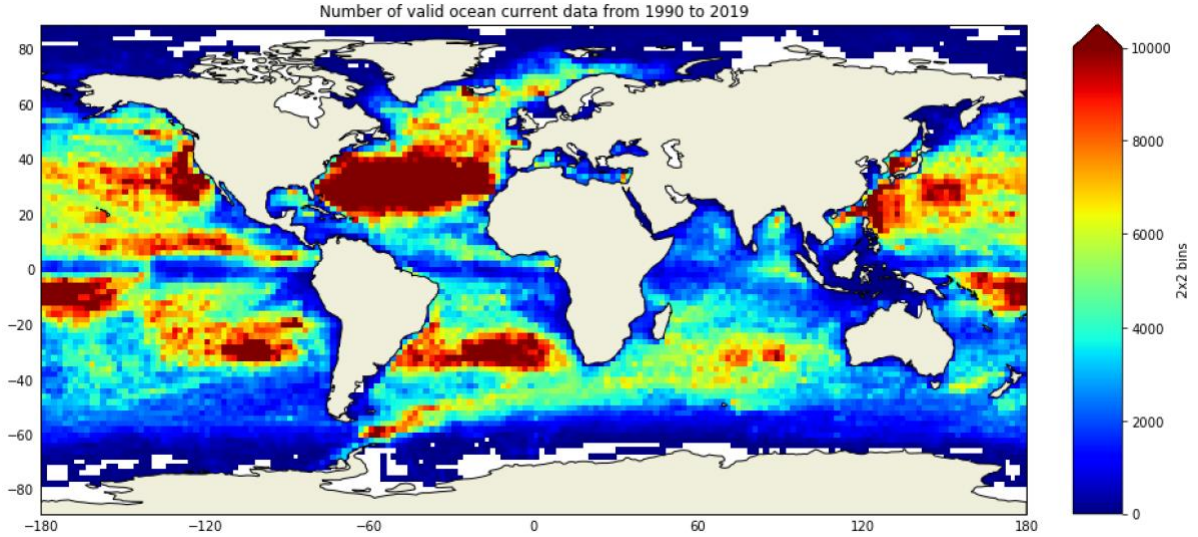


*Figure 2: 2*2° bin mean number of number of velocity measurements from 1990 to August 2021*

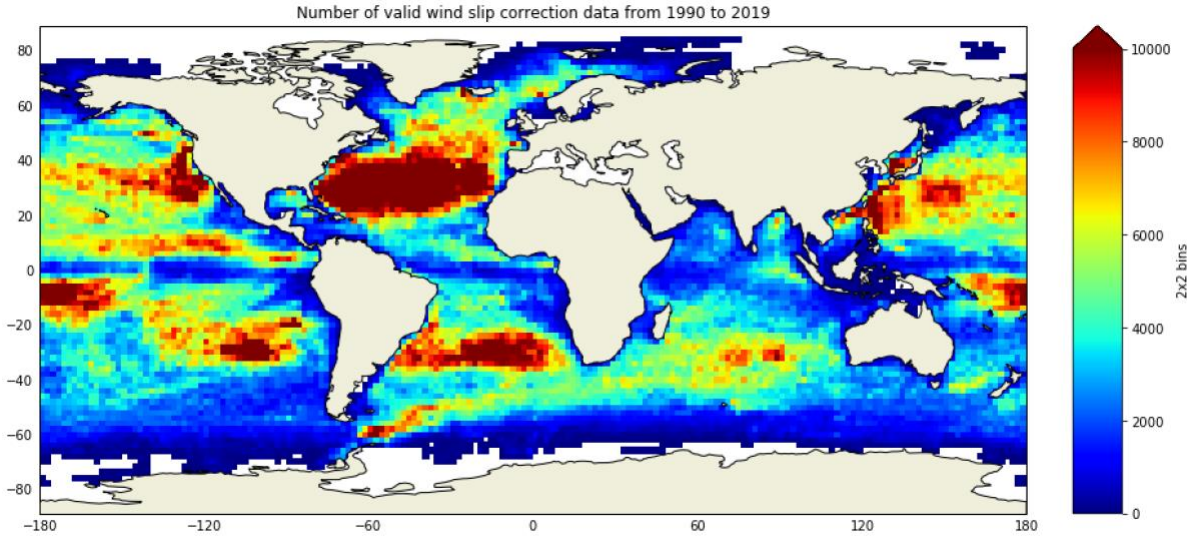


*Figure 3: 2*2° bin mean number of number of wind slip correction data from 1990 to August 2021*

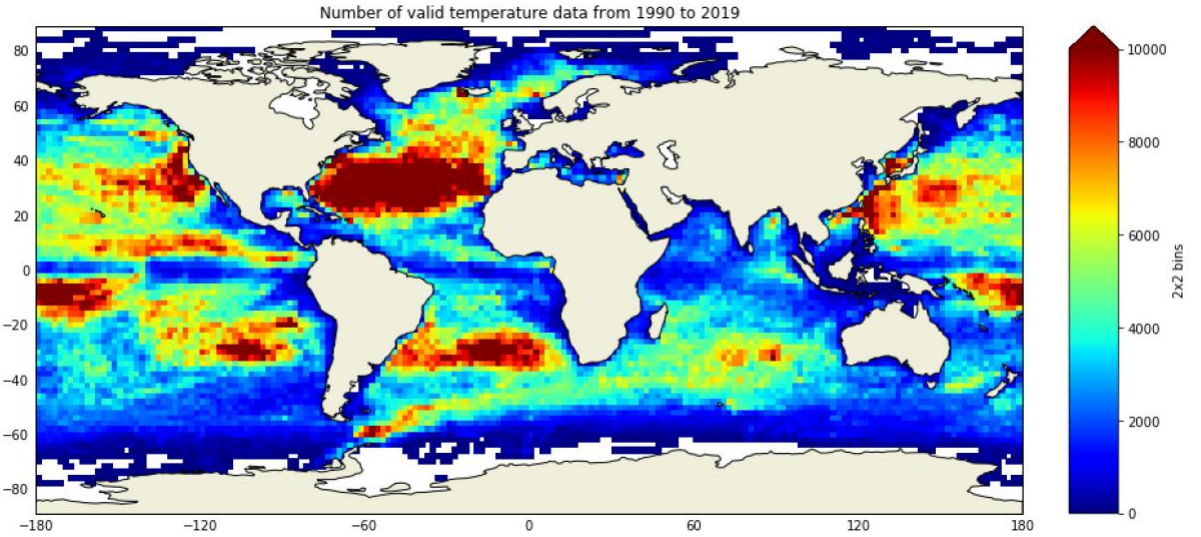


*Figure 4: 2*2° bin mean number of number of temperature data from 1990 to August 2021*

Fixed and drifting buoys Global Ocean Wave Observations Reanalysis

Copernicus Marine In Situ updates twice a year the global reanalysis of wave data observed from fixed or drifting buoys.

- Global Ocean Wave Observations Reanalysis https://doi.org/10.17882/70345
- Quality information document for reprocessed In Situ waves https://doi.org/10.13155/58696

This product integrates observations aggregated and validated from the Regional EuroGOOS consortium (Arctic-ROOS, BOOS, NOOS, IBI-ROOS, MONGOOS) and Black Sea GOOS as well as from National Data Centers (NODCs), JCOMM global systems (OceanSITES, DBCP) and the Global telecommunication system (GTS) used by the Metoffices. The INSTAC relies on observing systems maintained by institutes that are not part of the INSTAC and CMEMS service is not contributing to the maintenance and setting up of the observing systems it uses.

Data are distributed on full level (no interpolation). They are available in a dedicated directory to waves (INSITU_GLO_WAV_REP_OBSERVATIONS_013_045) of CMEMS Dissemination System in one or two files per platform.

The wave REP product is a global product and provides two kind of files: one with integrated parameters computed from the wave spectrum (e.g. significant wave height, peak period, mean direction) or zero crossing parameters (e.g. maximum wave height, mean height) and another one with the spectral information when available (scalar spectrum and directional functions like mean direction and angular spreading depending on the frequency). The files with integrated parameters contain also other physical and meteorological variables measured by the same platform. The complete list of variables distributed by the In Situ TAC can be found in the Copernicus Marine in situ TAC physical parameters list (https://doi.org/10.13155/53381).

This product is updated twice a year after a validation process carried out at regional level and described in detail in the Quality Information Document CMEMS-INS-QUID-013-045.

Since December 2020, files with wave spectra are included in the wave REP product.

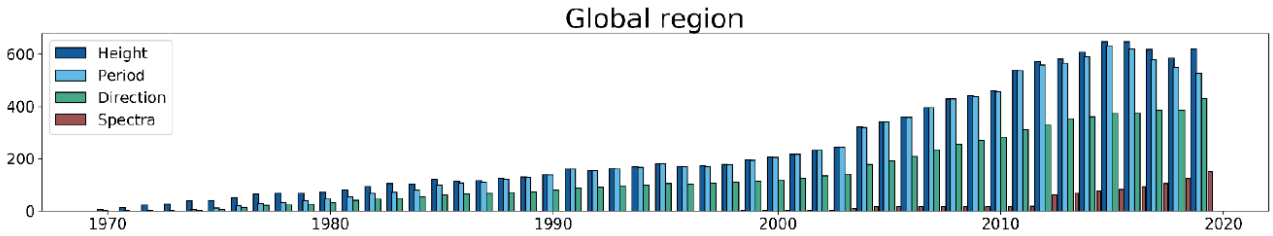


Figure 5: evolution of number of wave platforms in Copernicus product INSITU_GLO_WAV_REP_OBSERVATIONS_013_045

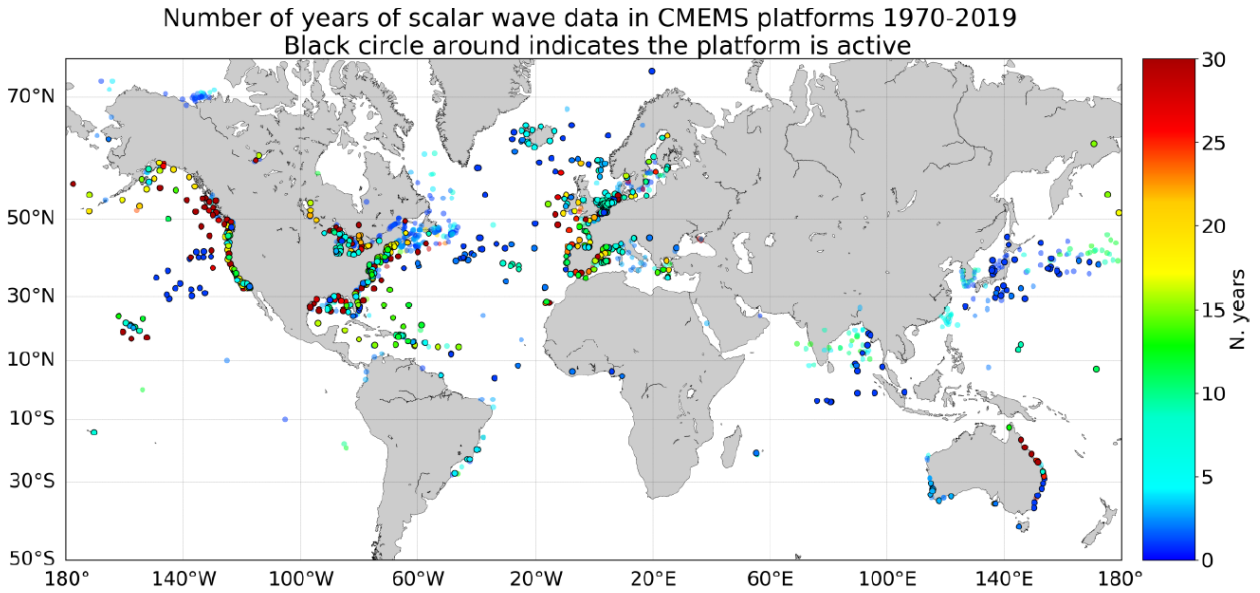


Figure 6: number of years of scalar wave data for years 1970 - 2020

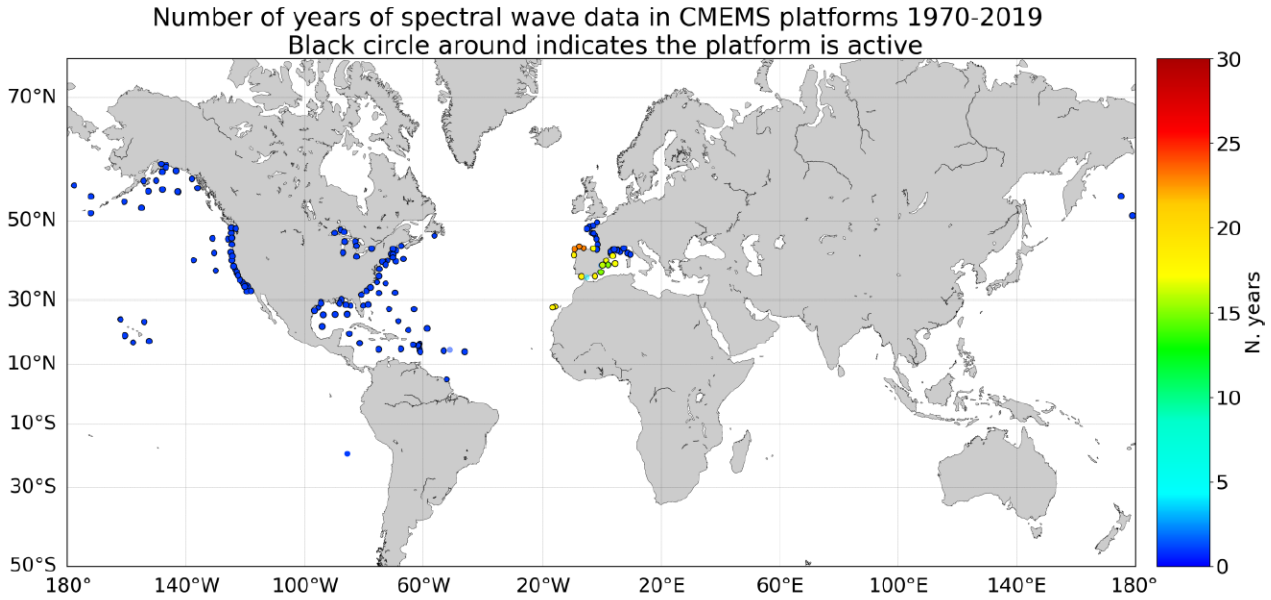


Figure 7: number of years of spectral wave data for years 1970 - 2020, black circle around indicates active platforms

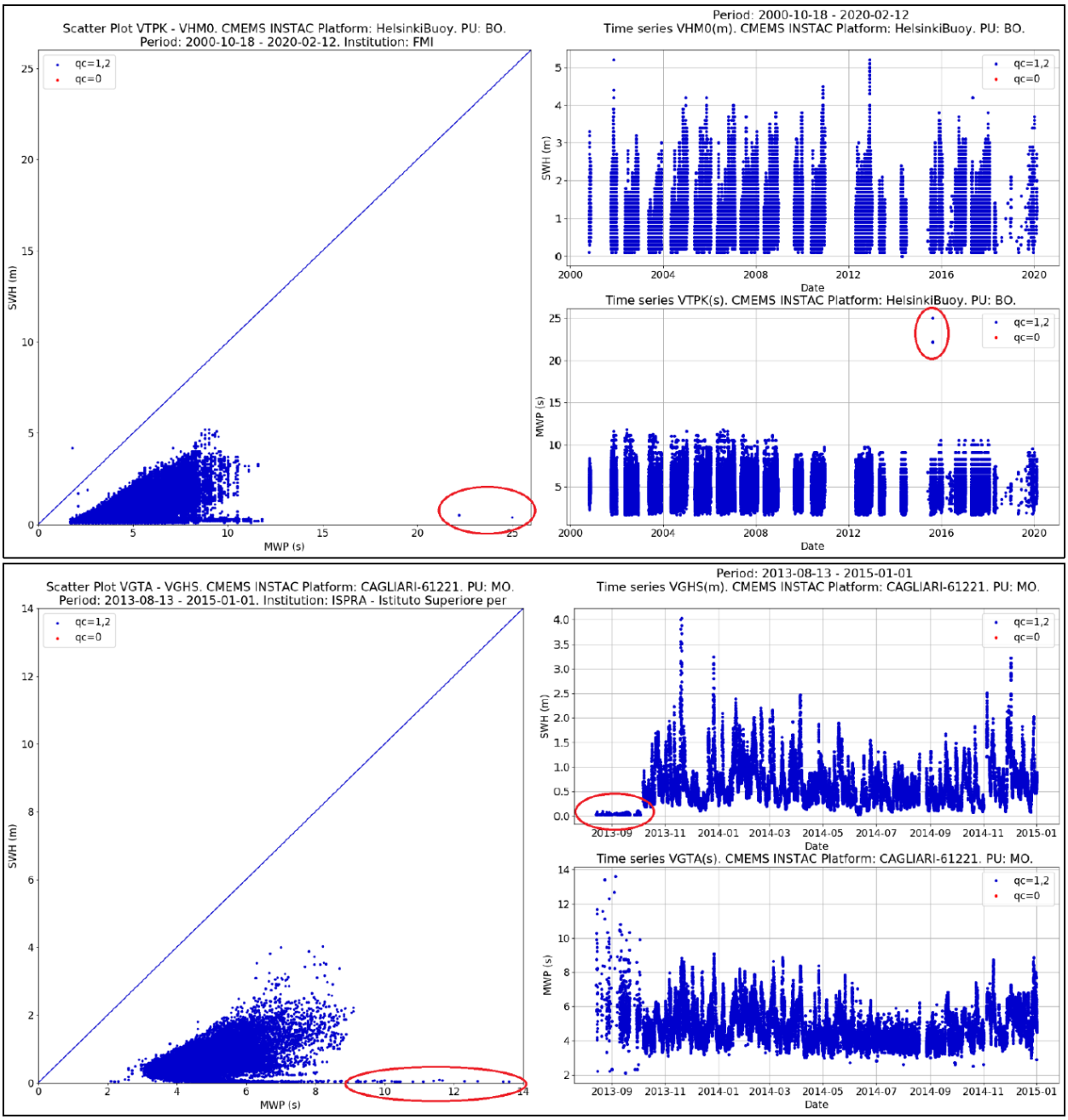


Figure 8: example of visual inspection graph with detection of spikes in Helsinki buoy (Baltic sea, upper figure) and wrong data in an anomalous period due to a malfunction in Cagliari buoys (Mediteranean sea)

Météo-France QCTools

Météo-France operates quality control (QC) procedures on drifting buoys data. Buoy data QC tools developed by Météo-France are available on the Internet (http://esurfmar.meteo.fr/qctools/) to help buoy operators to check their own buoys: monthly statistics carried out by four meteorological centres for individual buoys; plots of data and differences with model outputs; blacklists of buoys reporting dubious air pressure values or being perhaps ashore can be seen.

DAC – Data Assembly Centre activity

Coriolis/Météo France operates a Data Assembly Centre for drifting buoys.
The DAC function is to:

- Receive original drifting buoys data and preserve them
- Aggregate drifting buoys data and metadata into standardized NetCDF-CF files
- Decode and apply quality control in real-time
- Distribute quality controlled data on GDAC and GTS
- Manage updates (calibration, reprocessing) in delayed mode

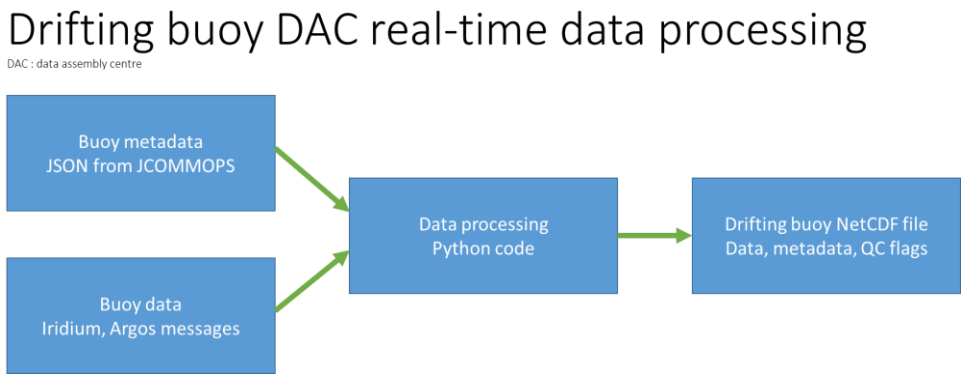
In September 2021, 851 drifting buoys with Iridium-SBD communication are managed by Coriolis-DAC, metadata are provided by JCOMMOPS.

Fisheries and Oceans Canada sends its drifters iridium data to Coriolis DAC.

Coriolis DAC may act as a DAC for SVP Iridium buoy data providers. This is realistic if the complexity and the number of type of buoys is not too big. Coriolis may also act as the DAC for orphan buoys whose data circulate on GTS only.

The drifting buoy data processing chain is freely available from:

- **Drifting buoys DAC data processing chain version 1.0**
  http://doi.org/10.17882/51148



GDAC – Global Data Assembly Centre activity

Coriolis operates a Global Data Assembly Centre for drifting buoys.
The GDAC function is to:

- Aggregate real-time and delayed-mode NetCDF-CF files provided by DACs
- Check the NetCDF-CF compliance of the DACs files
- Preserve NetCDF-CF drifting buoys files
- Distribute files on multiple channels
  FTP, HTTP, ERDDAP

The Drifting buoys GDAC MCDS is documented with:

- **Drifting buoys GDAC organization**

- **Drifting buoys GDAC NetCDF data and metadata format version 1.0**
  http://doi.org/10.13155/52037

- **Drifting buoys DAC data quality control manual version 1.0**
  http://doi.org/10.13155/52040

- **Drifter metadata in the WIGOS standard**
  https://doi.org/10.5281/zenodo.1406121

Drifting Buoys GDAC ftp server

The Drifting Buoys GDAC activity started in August 2018.

ftp://ftp.ifremer.fr/ifremer/dbcp-drifter

- gdac/active real-time drifting buoys
- gdac/history delayed-mode drifting buoys
- gdac-index.csv index of all GDAC drifting buoys

The gdac/active and gdac/history directories are populated with one file per buoy. Each file contains the buoy data and metadata (from OceanOPS).

The gdac/active is populated with iridium data received from Coriolis DAC (851 active Iridium-SBD drifting buoys in September 2021). The objective is to extend it to all history and active buoys managed by Meteo-France.

Historical data of drifting buoys with no access to original telemetry data may be recovered from AOML or MEDS historical GTS database.

The C-RAID project described in the following chapter is working on retrieving and reprocessing original drifting buoys data.

Drifting buoys Coriolis DAC ERDDAP access

The GDAC NetCDF files are also distributed on ERDDAP server

- DBCP Coriolis DAC ERDDAP server

Drifting buoys GDAC DOI access

A snapshot of the whole GDAC content is performed regularly (quarterly). The snapshot is preserved and published with a unique DOI: Digital Object Identifier.

- DBCP drifting buoys data and metadata from Global Data Assembly Centre (DBCP GDAC) https://doi.org/10.17882/57247
  - The DOI should systematically used for DBCP data citation: crucial for efficient bibliographic surveys.
  - The DOI and its associated fragment identifies a specific snapshot: crucial for reproducibility of result cited in scientific publications.

Statistics on drifting buoy data collated from GTS, Ifremer, SHOM

With Copernicus Marine funding, Coriolis (Météo-France, Ifremer) collates, quality controls, archives and distributes marine in situ data circulating on GTS or deployed by French and European institutions. Most of these data come from observation networks: drifting buoys, Argo floats, moorings, vessels, gliders or sea-mammals.

Observation networks

In 2022, Coriolis data centre managed 58.934 distinct platforms including 25.478 drifting buoys.

To measure the impact of the observations available from these platforms on models using in situ data, we calculate a platform-day histogram.
Definition of platform-day: a platform that reported at least one observation on a day: +1.

In 2022, drifting buoys represent 50% of the platform-day observations.

The diagram below illustrates the importance of drifting buoys along animal, Argo float, mooring, glider and vessel observation networks.

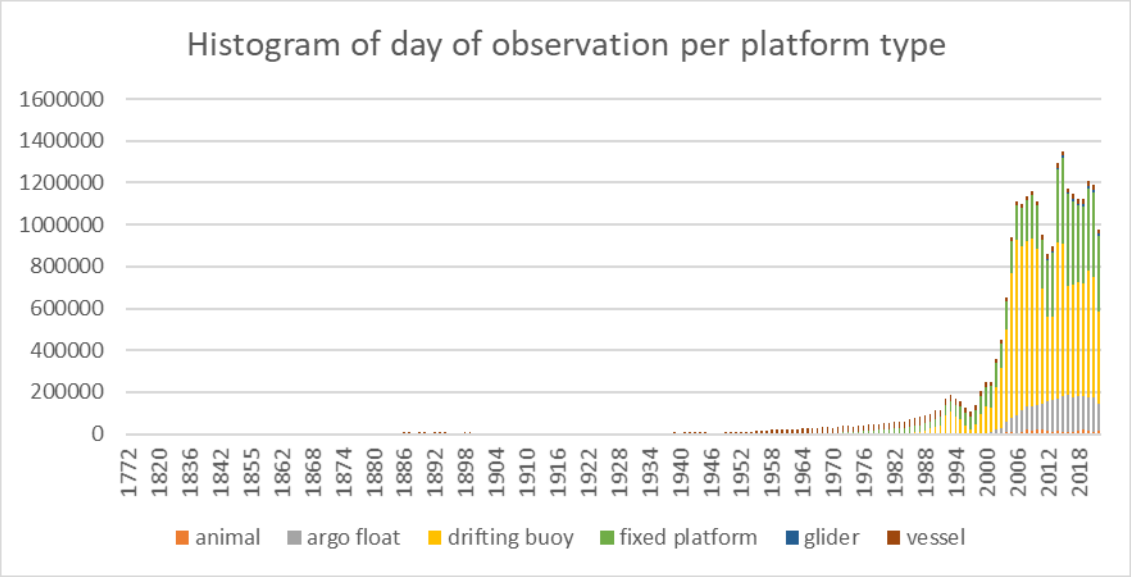


Figure 9: observations in platform-day, from year 1772 to 2022

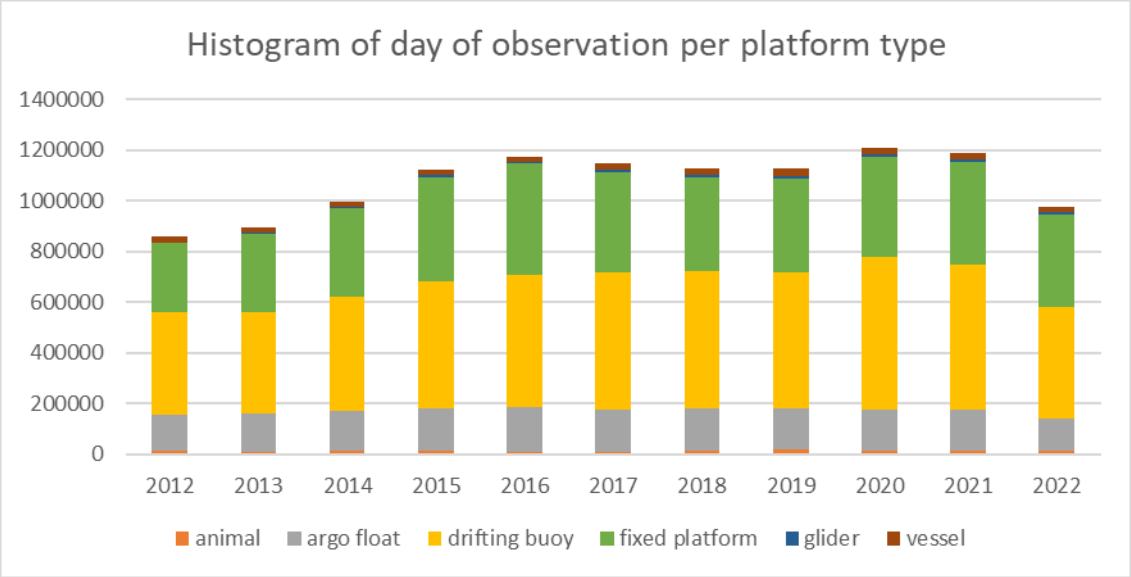


Figure 10: the last 10 years of observations in platform-day (year 2022 is not over)

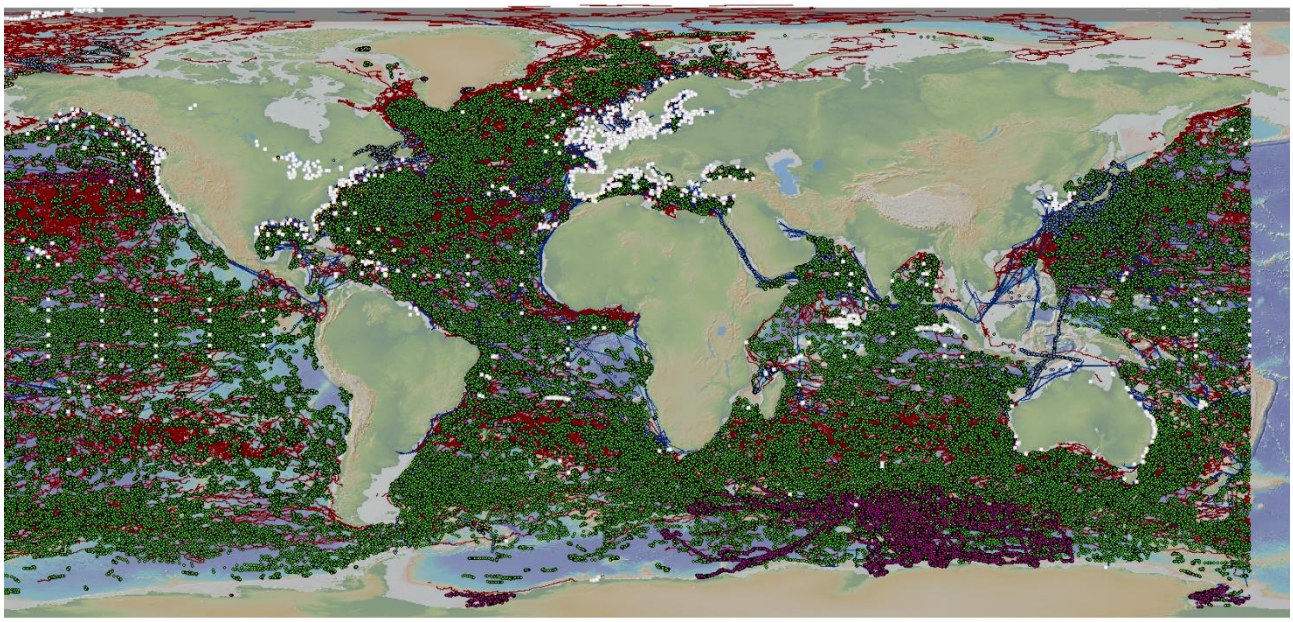


Figure 11: observation networks geographic coverage, 2021 (one year). The drifting buoys trajectories (red lines) are somewhat hidden by Argo profiles (green dots)

Drifting buoys observation network

In 2022, Coriolis Data Center collated data from **25 478 drifting buoys** (+79% from 2021), **2 700** were active during the last 12 months. The 79% increase in number of drifting buoys is due to the C-RAID project, a reprocessing and data rescue of 20 000 drifting buoys (in 2022, 10 000 drifting buoys were retrieved with C-RAID project).

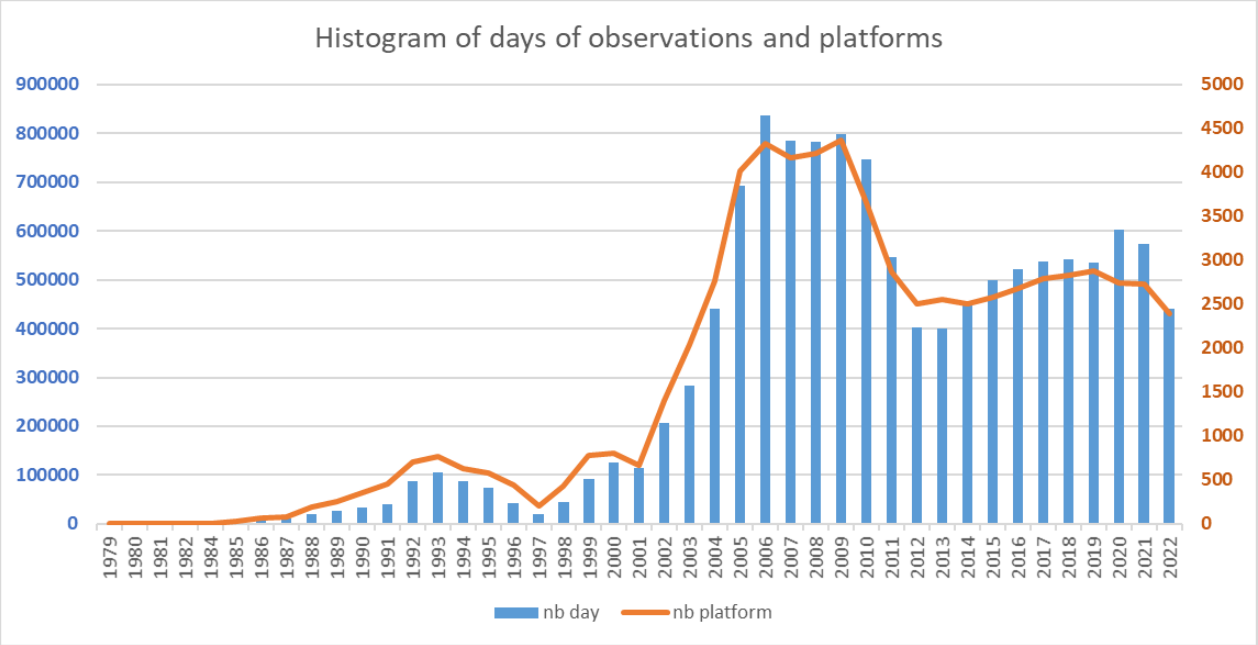


Figure 12: histogram of drifting buoys observations since 1979

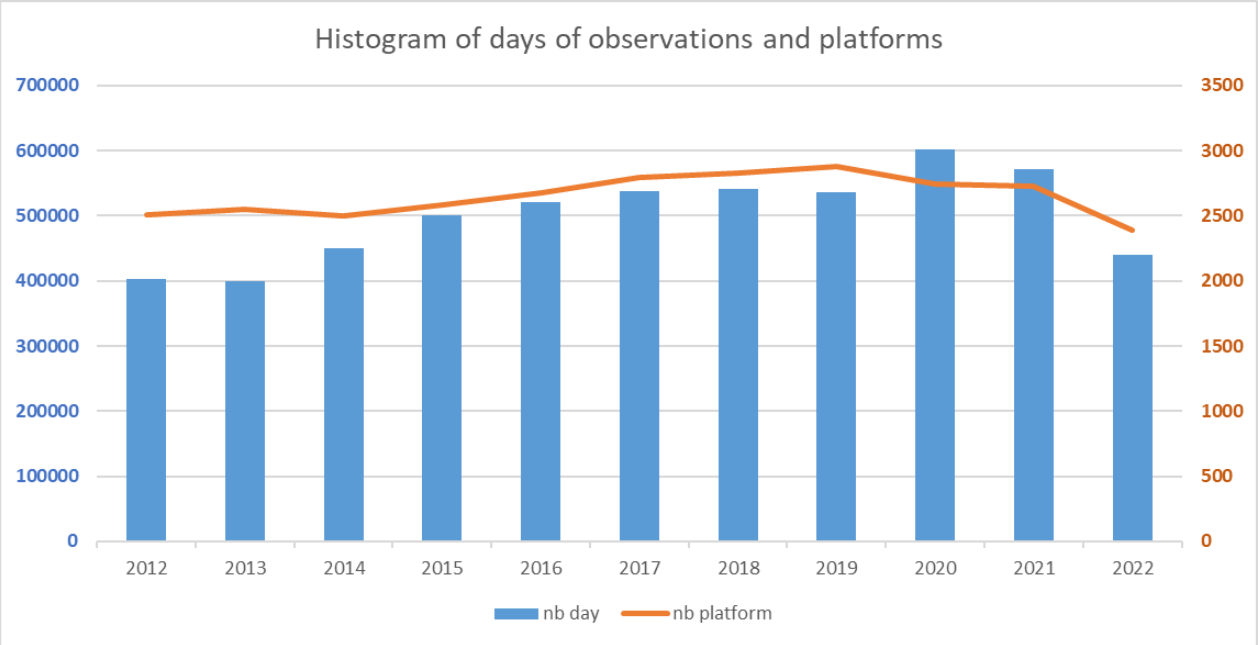


Figure 13: drifting buoys observations yearly distribution (year 2022 is not over) blue: platform-day, orange: number of active platforms (One day of observation from one buoy = +1)

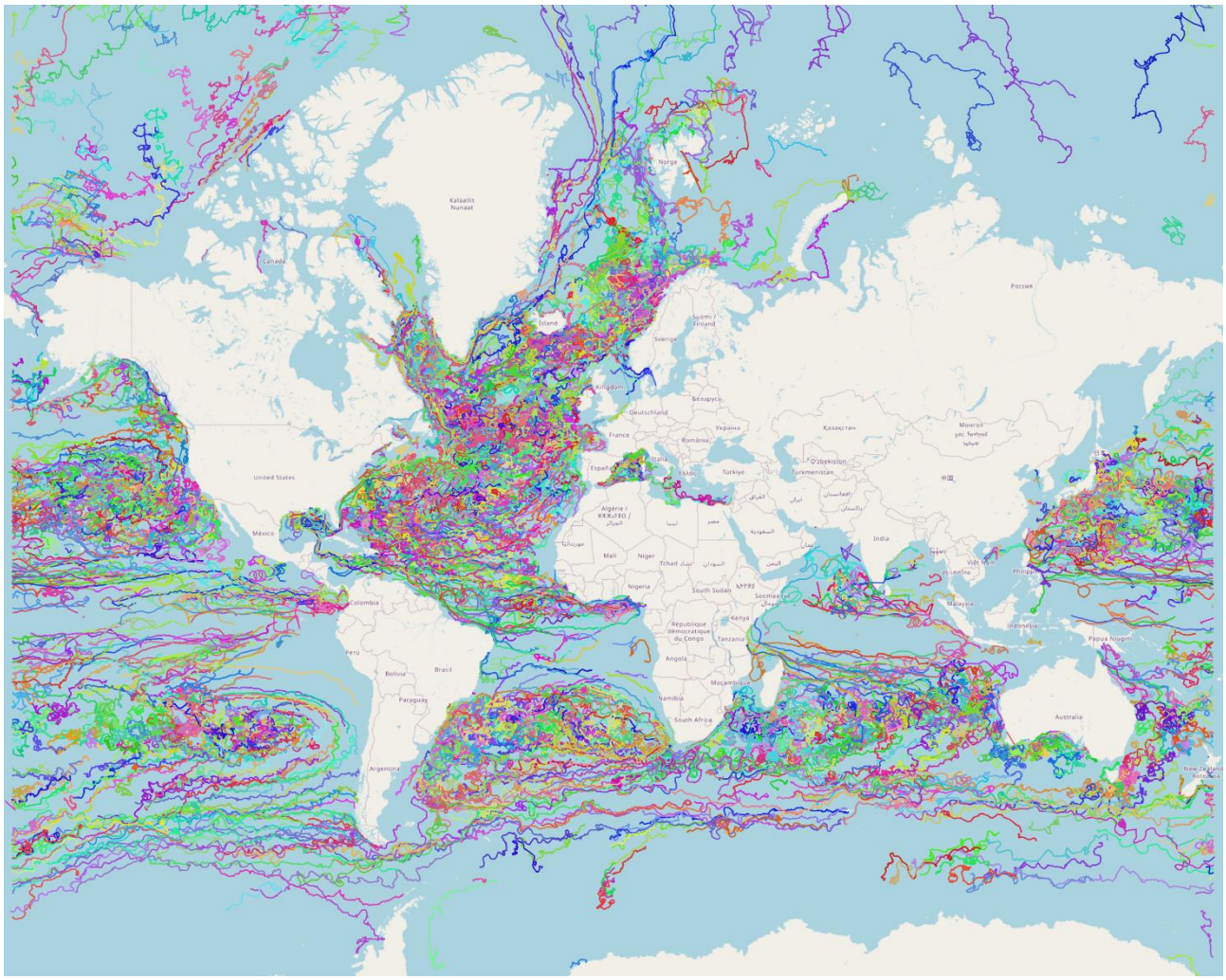


Figure 14: drifting buoys trajectories distributed by Coriolis - Copernicus in 2022 (one color one buoy)

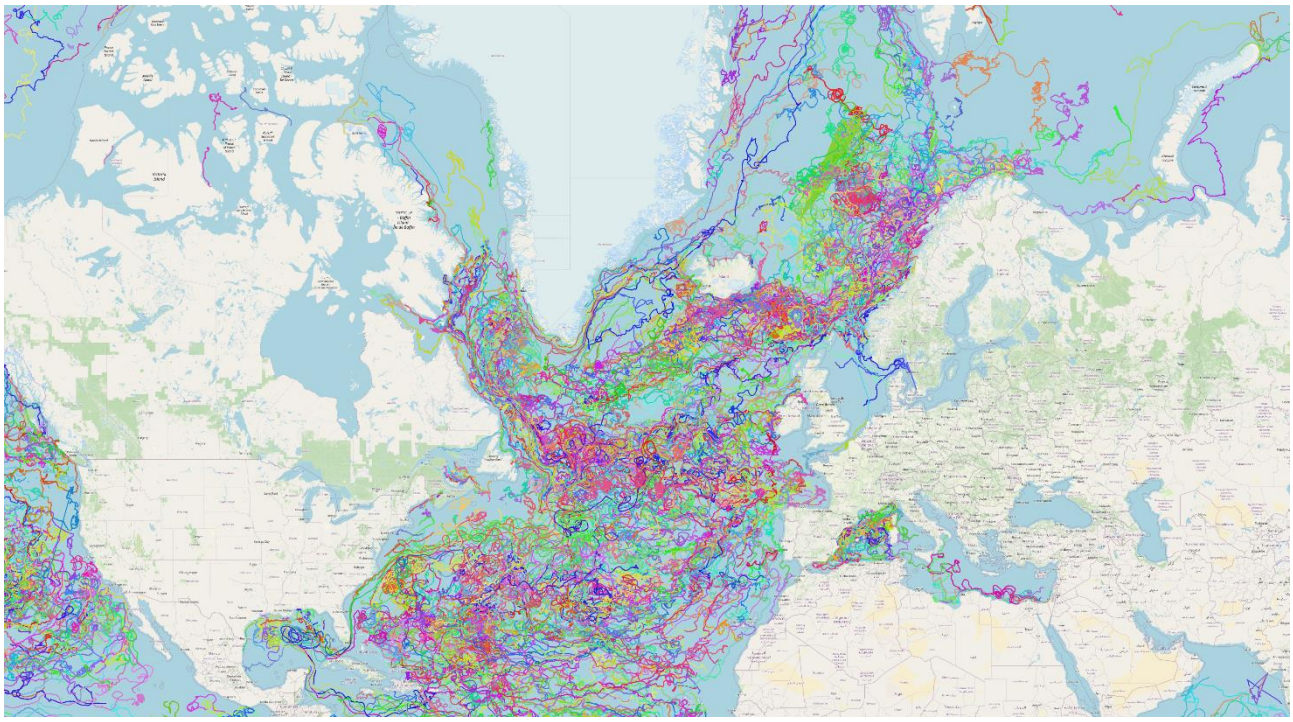


Figure 15: drifting buoys trajectories, zoom on North Atlantic, 2022 activity

DAC GDAC Objectives 2022/2023

- Consolidate the European DAC activity
- Initiate the activity of new DACs
- Organize the data management of orphan drifting buoys (no DAC), in particular from C-RAID project

C-RAID project contribution to Drifting Buoys GDAC

A data rescue project: C-RAID - Copernicus Reprocessing and Access Improvement for Drifter data https://doi.org/10.17882/77184

C-RAID project is a global reprocessing of drifting buoys data and metadata: 25 000 drifting buoys, deployed between 1979 and 2018. The data of 12 000 drifting buoys deployed between 1997 and 2010 have been delayed mode processed (including comparison with ERA5 reanalysis). The project is continuing in 2021-2022 as "C-RAID phase 2" to reprocess drifting buoys data deployed between 1979 - 1996 and 2011 - 2018.

C-RAID deliverables

1. An improved drifting buoys data archive
2. FAIR interfaces to drifting buoys data : Web data discovery for human users, API data discovery/subsetting/download services

By "Improved drifting buoy data record" we mean:

- Missing datasets and parts of datasets recovered (data rescue)
- Homogeneous & rich metadata and data format (NetCDF-CF)
- Homogeneous expert QC on marine and atmospheric data
- Matchup ERA5 data (temperature, atm pressure, wind)

C-RAID phase1 data, metadata and documentation are now published on

- https://doi.org/10.17882/77184

FAIR data commitment

Findable	DOI published on DataCite & Google Schema.org, Link with bibliography and authors (ORCID)
Accessible	One click download, anonymous access
Interoperable	CF and SeaDataNet standards, QC documented
Re-usable	CC-BY license

C-RAID stakeholders

- **Ocean-Atmosphere community**
  C-RAID higher resolution data would complement ICOADS buoys data and improve Copernicus ERA reanalysis.
- **Ocean community**
  Temperature for satellite calibration/validation, model and reanalyses validation or forcing
  - **CMEMS**: Global "OSTIA" L4REP and L4NRT use drifter SST to force the analyses in complement of satellites with a maximum nb of sensors, to solve mesoscale structures.
  - The drifter SST data are assimilated by CMEMS MFCs to constrain the surface fields. They are also used to validate surface fields.
    C3S: Global ESA SST CCI and C3S L4REP use in situ drifter SST to validate fields built from satellite only data
    "climat" type product = stable in time

C-RAID data rescue activity

As part of C-RAID project, CLS reprocessed the Argo locations of 3000 buoys trajectories with a Kalman filter. On a majority of drifters, it provides a huge increase in the quality of trajectories, in particular on poor satellite coverage positions.

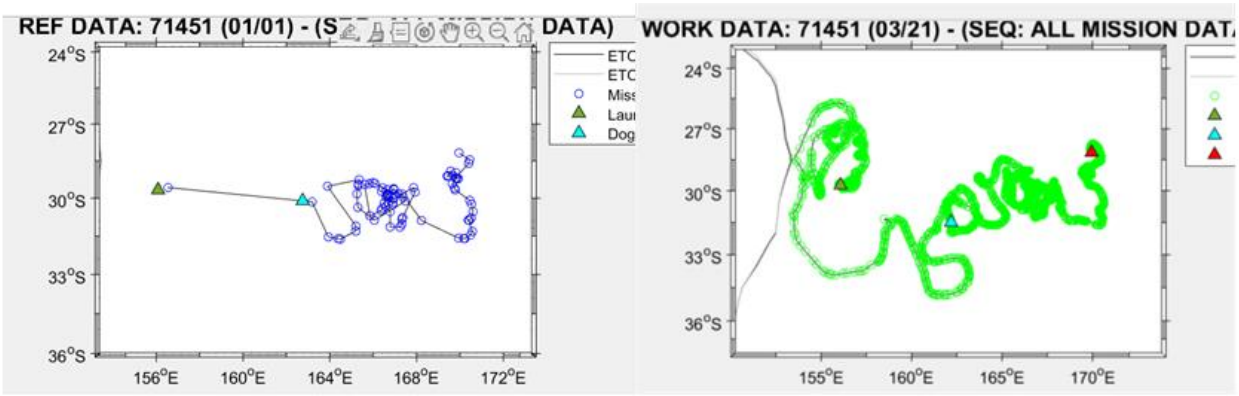


Figure 16: example of data rescue on buoy 71451: a huge amount of Argos locations were recovered with the Kalman filter (on the left, before Kalman, on the right after Kalman)

The reprocessing of the original Argos data recovered a significant buoys data observations.

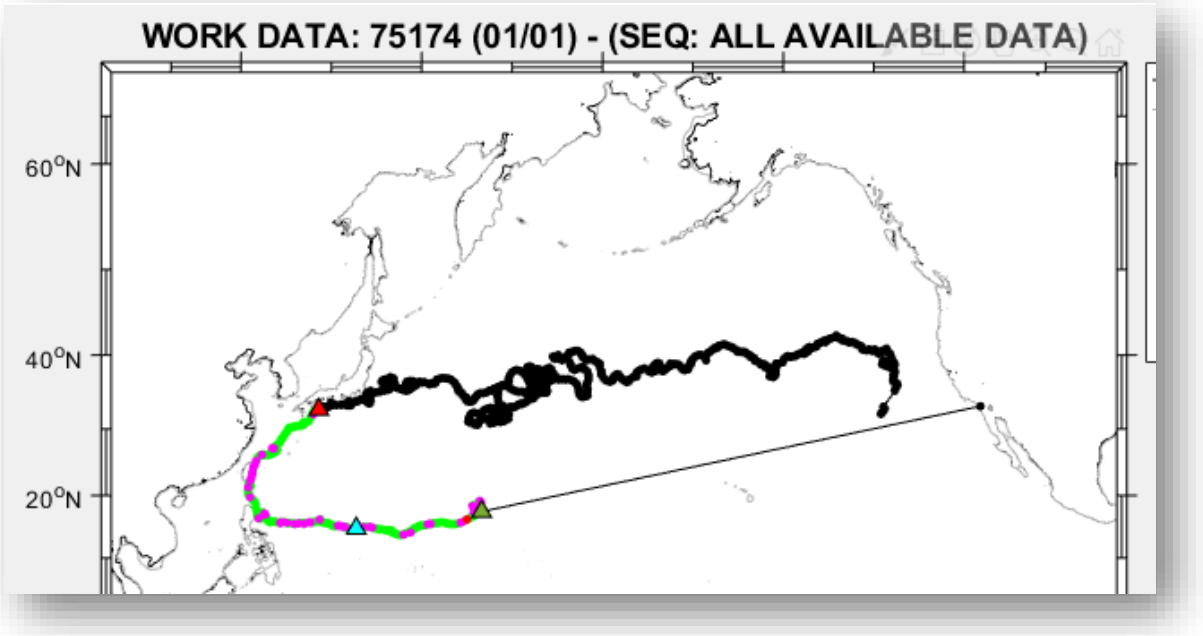


Figure 17: example of buoy 75174: the black and purple dots were missing in the buoys data repositories

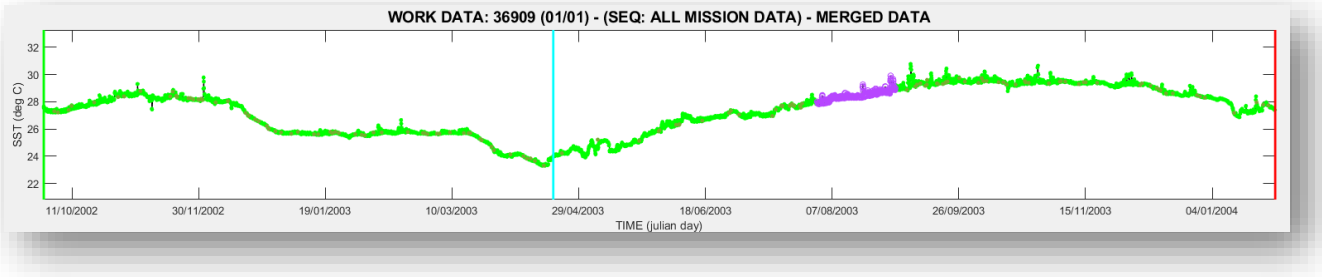


Figure 18: example of buoy 36909: the purple dots were missing in the buoys data repositories

C-RAID machine learning activity

The C-RAID dataset contains the data from 10 000 drifting buoys that have been scientifically inspected. It is fit for machine learning training.

An initial experiment was performed to detect the drifting buoys loss of floating anchor.

The question: is a buoy drifting with the nominal 15 meter deep current ?

A machine learning solution:

- train a model with 3500 buoys (with anchor loss detected by an expert) and ERA5 (wind, temperature,…)
- Run the model on 500 buoys and check the result with the expert decision

With XGBoost machine learning library, we reached a 98% of positive detections.

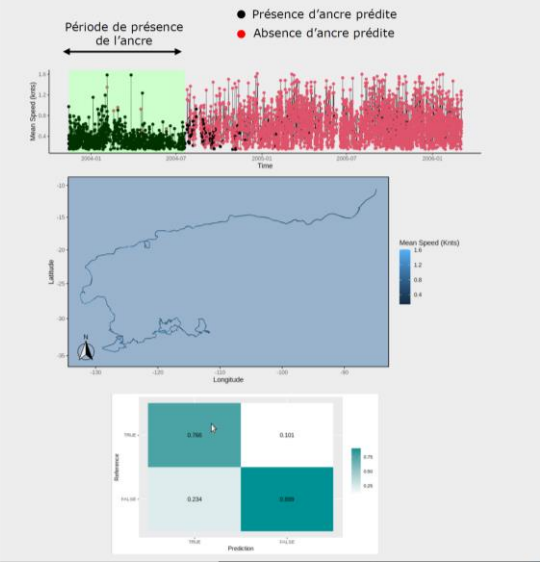


Figure 19: Machine learning technique to detect drifting buoy loss of floating anchor

References

Poli, P., D.P. Dee, R. Saunders, V.O. John, P. Rayer, J. Schulz, K. Holmlund, D. Coppens, D. Klaes, J.E. Johnson, A.E. Esfandiari, I.V. Gerasimov, E.B. Zamkoff, A.F. Al-Jazrawi, D. Santek, M. Albani, P. Brunel, K. Fennig, M. Schröder, S. Kobayashi, D. Oertel, W. Döhler, D. Spänkuch, and S. Bojinski, 2017: Recent Advances in Satellite Data Rescue. Bull. Amer. Meteor. Soc., 98, 1471–1484, https://doi.org/10.1175/BAMS-D-15-00194.1

WMO, 2016. The Global Observing System for Climate: Implementation Needs. GCOS- No. 200. https://library.wmo.int/doc_num.php?explnum_id=3417