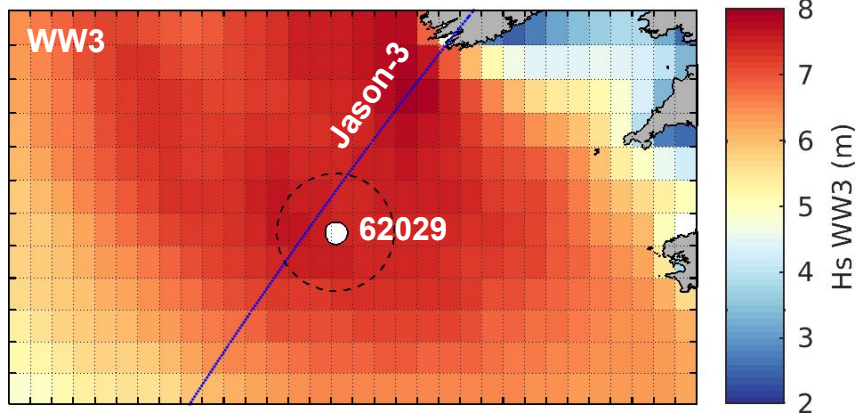
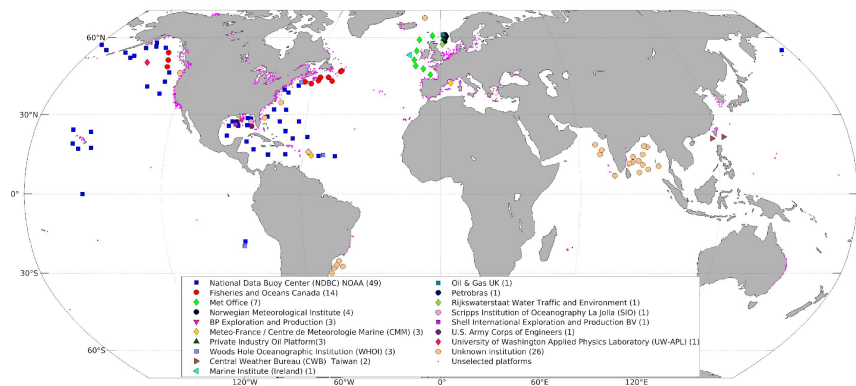


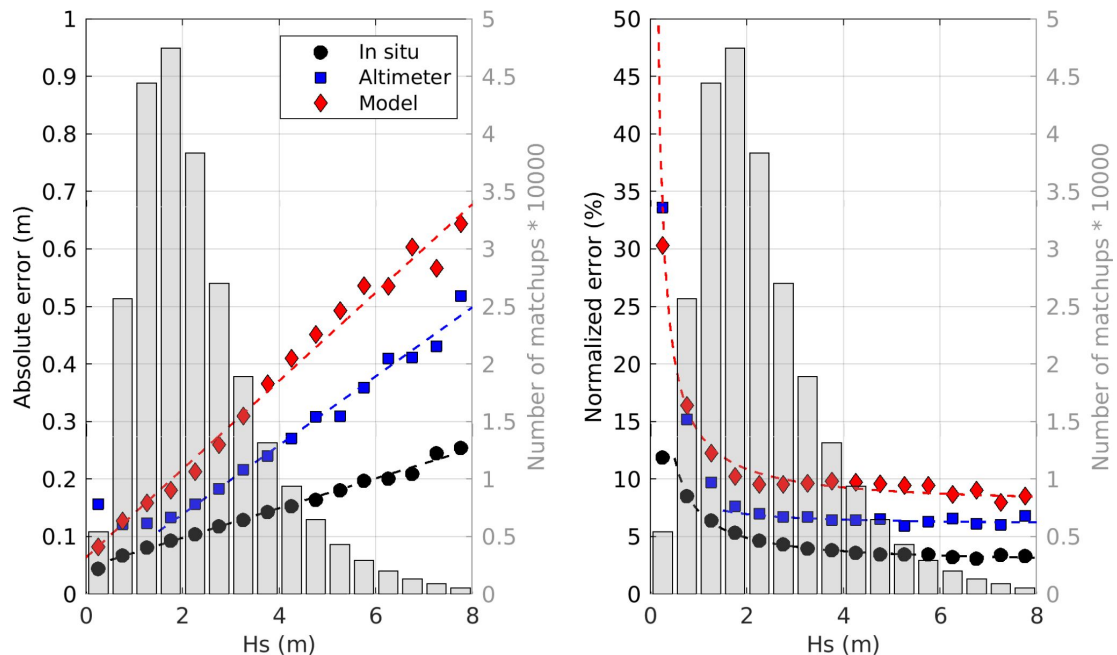
Error estimation of buoy, altimeter, and model significant wave height from triple collocation technique

G. Dodet¹, J.R. Bidlot², M. Accensi¹, M. Alday¹, S. Abdalla², J.F. Piolle¹, F. Ardhuin¹



- **Buoys :** CMEMS In Situ TAC (122 platforms)
- **Altimeter :** ESA Sea State CCI (10 missions)
- **Model :** WW3 Ifremer-ERA5 T475 (global hindcast)

1992-2018



Context

- Ocean wave information is critical for climate studies, safety at sea, marine engineering, coastal risk management...
- Characterizing the error of wave measurement systems is needed to optimize the combined use of multiple source data (e.g. data assimilation of satellite data in NWP model)
- Conventional (dual comparison) error assessment techniques assume that a reference system (e.g. buoys) is free of errors, which is not the case
- Triple collocation (TC) technique (Stoffelen, 1998) can be used to estimate error characteristics in 3 collocated datasets, e.g. in situ, satellite and model
- Several authors have used TC to investigate errors in wind speed (Abdalla and De Chiara, 2017), wind stress (Portabella and Stoffelen, 2008), wave height (Janssen et al. 2007), SST (Xu and Ignatov, 2016), soil moisture (Gruber et al., 2016)...
- Major difficulty with TC is to get a sufficiently large collection of independent collocated data

Method & Data

Triple collocation technique

In simple words

In triple collocation, three (ideally) independent data sets are brought together, so three scatter plots can be made. The plot with the least scatter obviously denotes the two systems that agree most, while the worst scatter plot, indicates that the excluded and third measurement system is the best performing (Vogelzang and Stoffelen, 2003).

True value : T
dataset 1 (ref) : $X_1 = T + e_1$
dataset 2 : $X_2 = a_2 + b_2 T + e_2$
dataset 3 : $X_3 = a_3 + b_3 T + e_3$

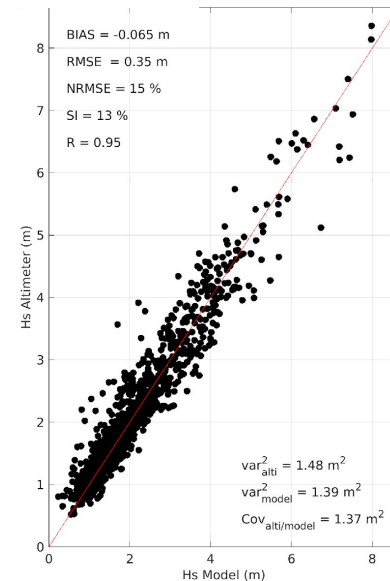
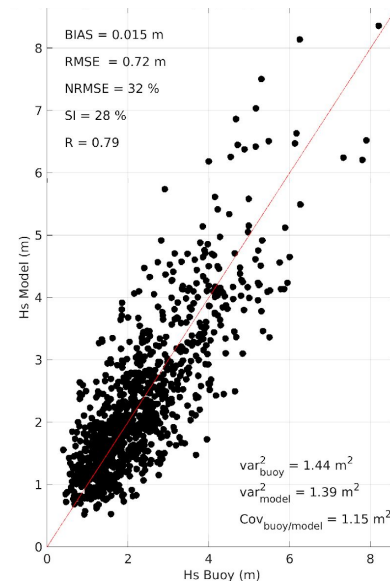
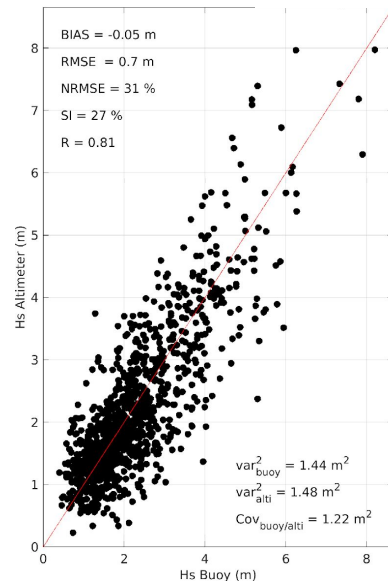
Following Stoffelen (1998), calibration coefficients and random errors e_i in each measurement system are estimated statistically from the mean ($\langle \rangle$), variance (C_{ii}) and co-variance (C_{ij}) of / between each dataset

$$a_2 = C_{23} / C_{13} \quad b_2 = \langle X_2 \rangle - a_2 \langle X_1 \rangle$$
$$a_3 = C_{32} / C_{12} \quad b_3 = \langle X_3 \rangle - a_3 \langle X_1 \rangle$$

$$\sigma_1^2 = C_{11} - (C_{12} C_{13}) / C_{23}$$
$$\sigma_2^2 = C_{22} - (C_{12} C_{23}) / C_{13}$$
$$\sigma_3^2 = C_{33} - (C_{23} C_{13}) / C_{12}$$

Assumptions

- Reference system is unbiased
- Linear calibration is sufficient over the whole range of values
- Errors e_i are uncorrelated



Dataset #1: In situ platforms (1)

CMEMS In Situ TAC dataset

Wave data and metadata from 1180 platforms are currently available from the CMEMS In Situ TAC dataset. Auxiliary data and additional quality flags have been derived for each in situ platform.

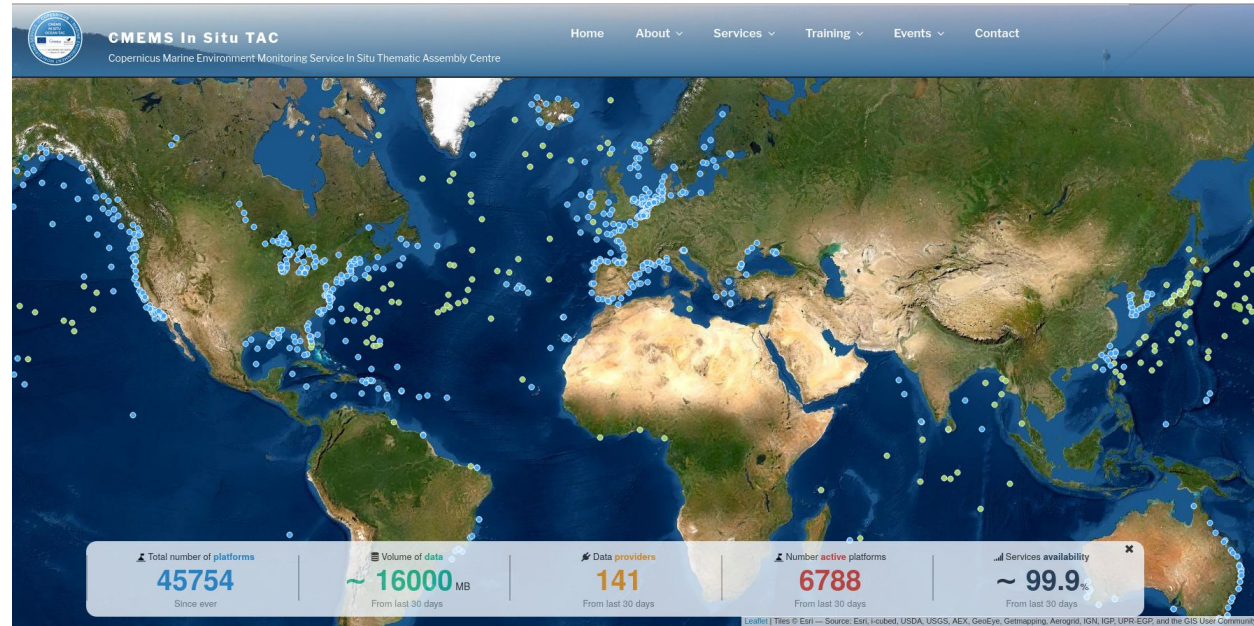
Auxiliary data

- distance to the coast (OSM polygons)
- depth at platform location (GEBCO 2019)
- time-varying Hs resolution
- platform position statistics
- quality flags (see below)

Quality control

The following flags are applied :

- quality flags provided by CMEMS
- & data provider
- flags on Hs range (0-30m)
- flags on stuck values (>36hrs)
- flags on low resolution acquisition ($\geq 0.5\text{m}$)



<http://www.marineinsitu.eu/dashboard/>

Dataset #1: In situ platforms (2)

CMEMS In Situ TAC dataset

Wave data and metadata from 1180 platforms are currently available from the CMEMS In Situ TAC dataset. Auxiliary data and additional quality flags have been derived for each in situ platform.

Auxiliary data

- distance to the coast (OSM polygons)
- depth at platform location (GEBCO 2019)
- time-varying H_s resolution
- platform position statistics
- quality flags (see below)

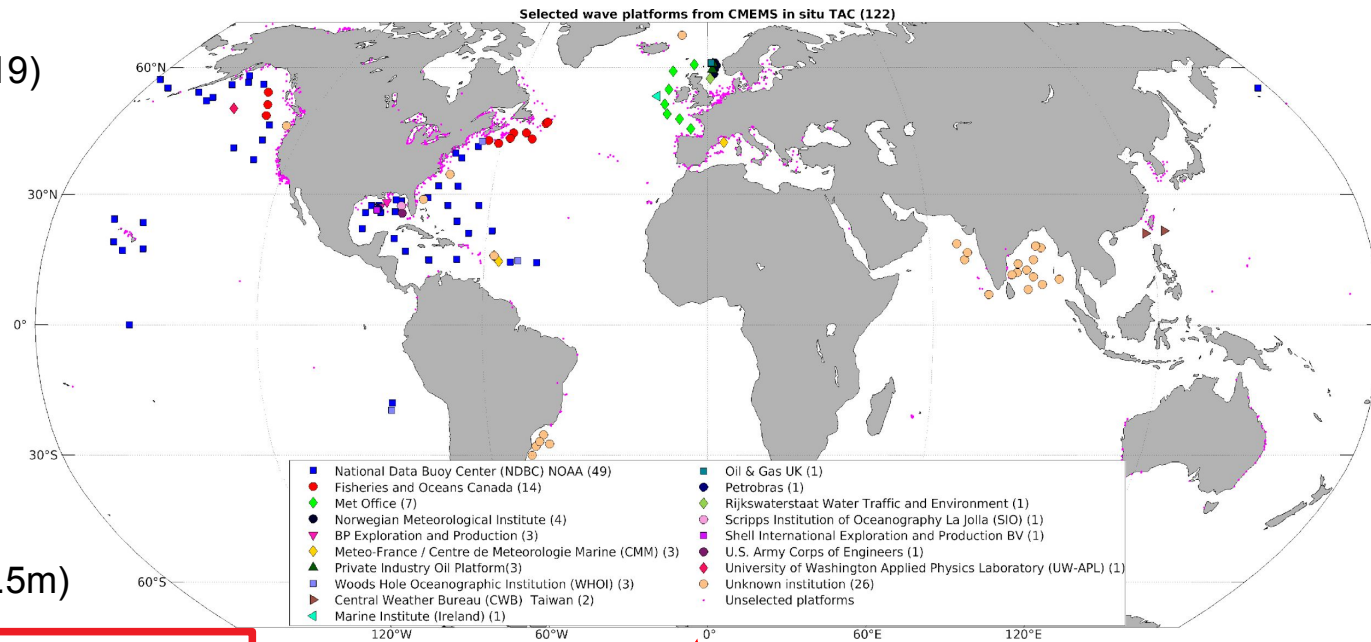
Quality control

The following flags are applied :

- quality flags provided by CMEMS
- & data provider
- flags on H_s range (0-30m)
- flags on stuck values (>36hrs)
- flags on low resolution acquisition ($\geq 0.5m$)

Platform selection

- distance to the coast > 100km
- depth > 50m
- record length covers at least 1 year within 1992-2018



Dataset #2: Satellite Altimetry

ESA Climate Change Initiative (CCI)

The objective of the CCI programme is to realise the full potential of the long-term global Earth Observation archives as a significant contribution to the ECV databases required by the UNFCCC

<https://climate.esa.int/>

Sea State CCI dataset V1 (Dodet et al., 2020)

Mission	Instrument	Covered period	Repeat period (days)	Altitude (km)	Inclination (°)	Source product
ERS-1	RA	1991–2000	35	785	98.52	OPR [ESA/F-PAF]
TOPEX	NRA	1992–2006	10	1336	66	MGDR [CNES]
ERS-2	RA	1995–2011	35	785	98.52	OPR [ESA/F-PAF]
GFO	GFO-RA	1998–2008	17	800	108	GDR/POE [NOAA]
JASON-1	Poseidon-2	2001–2013	10	1336	66	GDR vE [AVISO]
ENVISAT	RA-2	2002–2012	35	799	98.55	GDR v2.1 [ESA/F-PAC]
JASON-2	Poseidon-3	2008–2019	10	1336	66	GDR vD [AVISO]
CRYOSAT-2	SIRAL	2010–Ongoing	369	717	92	IGDR [NOAA]
SARAL	AltiKa	2013–Ongoing	35	785	98.55	GDR [AVISO]
JASON-3	Poseidon-3B	2016–Ongoing	10	1336	66	GDR vD [AVISO]

3 products:

- L2P: multi-mission along-track significant wave height
- L3: daily merged multi-mission along-track significant wave height
- L4 : merged multi-mission monthly gridded significant wave height



available on <https://catalogue.ceda.ac.uk/>

climate change initiative

esa

Essential Climate Variables

 antarctic ice sheet cci	 sea ice cci
 ice sheets greenland cci	 sea level cci
 land cover cci	 sea level budget closure cci
 land surface temperature cci	 sea state cci
 ocean colour cci	 snow cci
 ozone cci	 soil moisture cci
 permafrost cci	 sst cci
 salinity cci	 water vapour cci

Dataset #3: Wave model

WW3 T475 parameterization (Alday et al. 2021)

Forcings

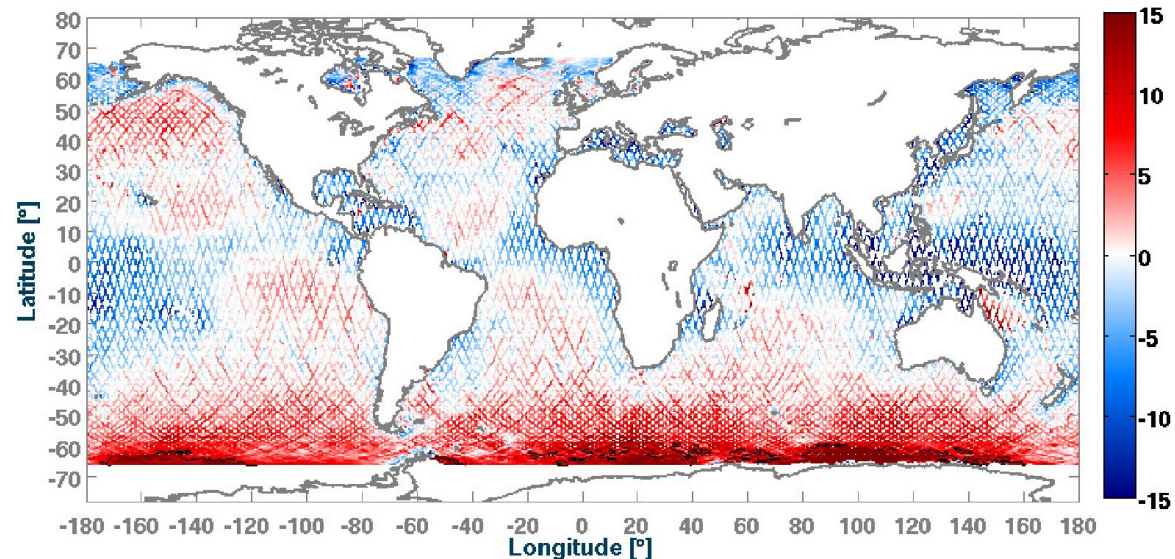
Wind : ERA-5
Currents: CMEMS-Globcurrent
Ice: Ifremer SSMI-derived daily product
(Girard-Ardhuin and Ezraty, 2012)
Iceberg: Ifremer-Altiberg icebergs distribution
database (Tournadre et al., 2015).

Resolution

Spatial: from 0.5° in global grid to 0.05° for
nested grids (multi-grid strategy)
Spectral: 24 directions and 36 frequencies from
 0.037 to 0.71 Hz

Time / Spatial Coverage

1993-2020
 78°S to 83°N



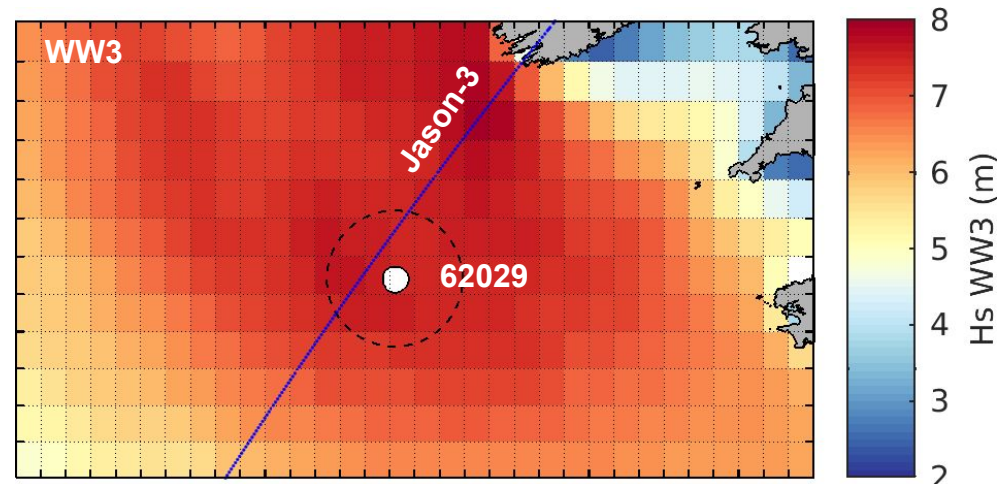
Normalized mean bias (NMB) for 1-year averaged between WW3 T475 and altimeter data, for year 2011 (Alday et al, 2021).

See presentation from Alday et al on Tuesday, 27 April 2021, 13:47 Session EGU21-15014

Triple collocation dataset

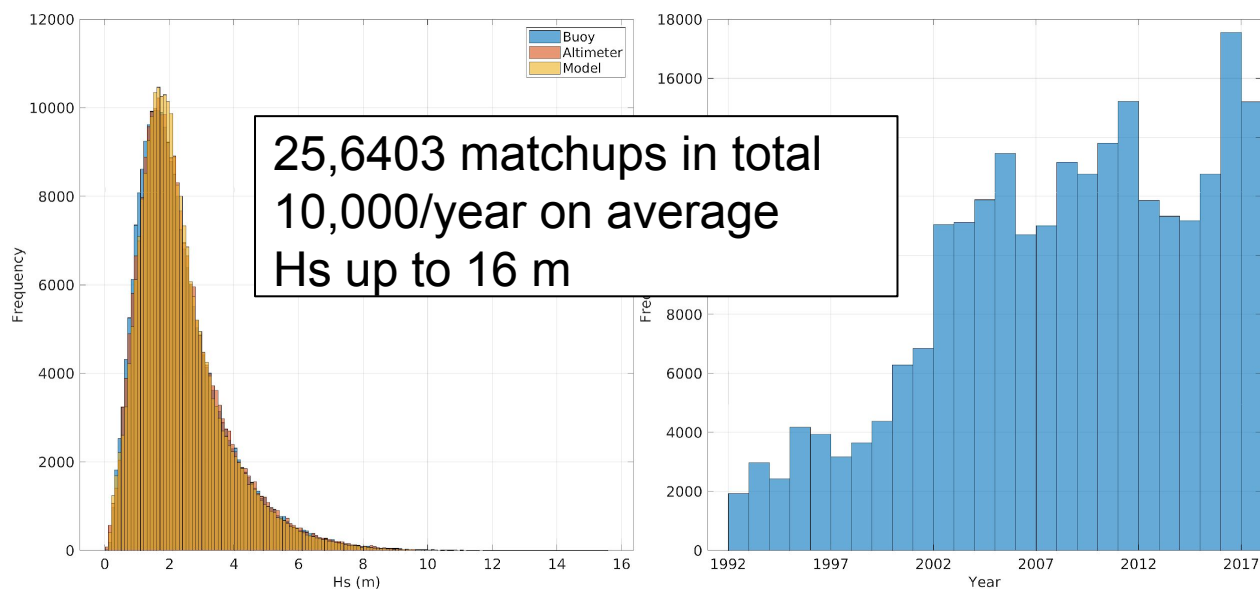
Altimeter – in situ platform matchups

Matchups between altimeter and in situ platform records are computed within 100-km distance and 1-hr time window. Along-track altimeter Hs are smoothed with a 50-km running average and in situ observations are smoothed with a 2-hr running average, in order to filter the high-frequency variability that cannot be resolved by a model with 0.5° grid spacing. 3-hourly model outputs are then (linearly) interpolated at the buoy location and at the time of the satellite-buoy intersection.



Matchup database

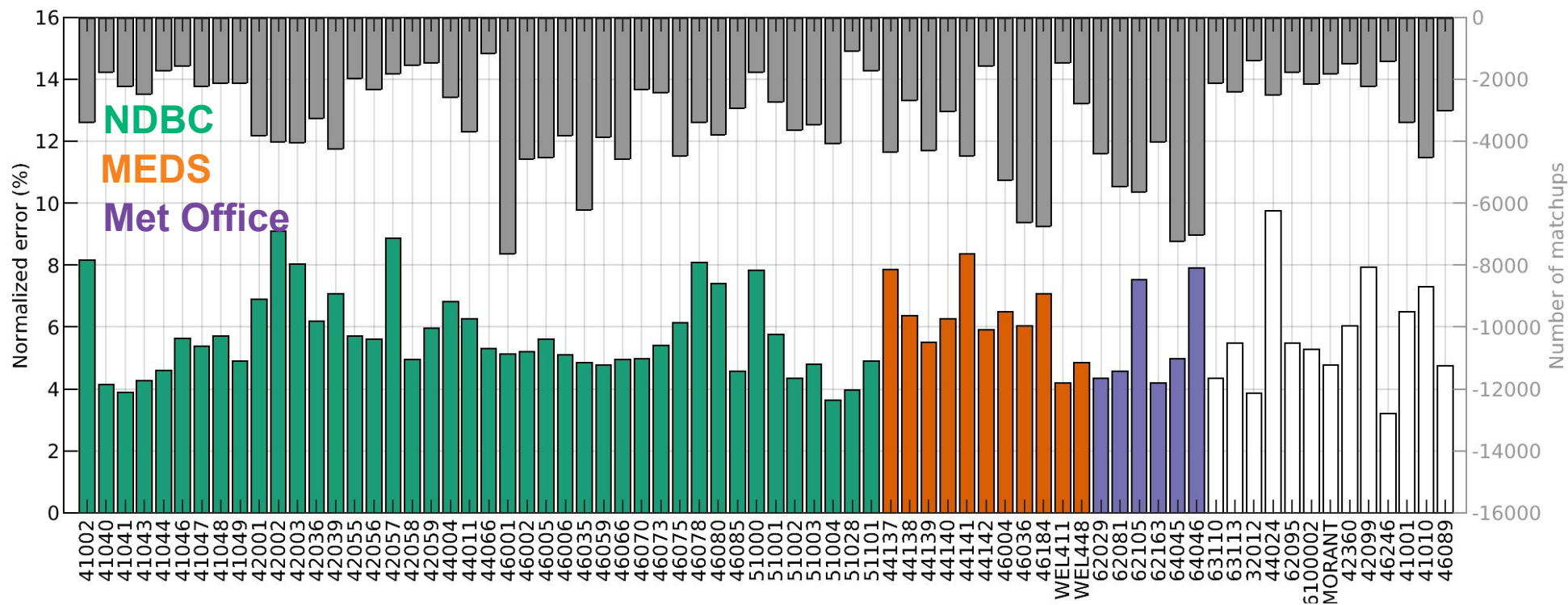
Mission	N years	N matchups
ERS-1	5	6758
TOPEX	13	26222
ERS-2	16	31804
GFO	9	21686
ENVISAT	10	29591
JASON-1	11	37333
JASON-2	10	40415
CRYOSAT-2	8	25736
SARAL	6	24281
JASON-3	3	12577



Results

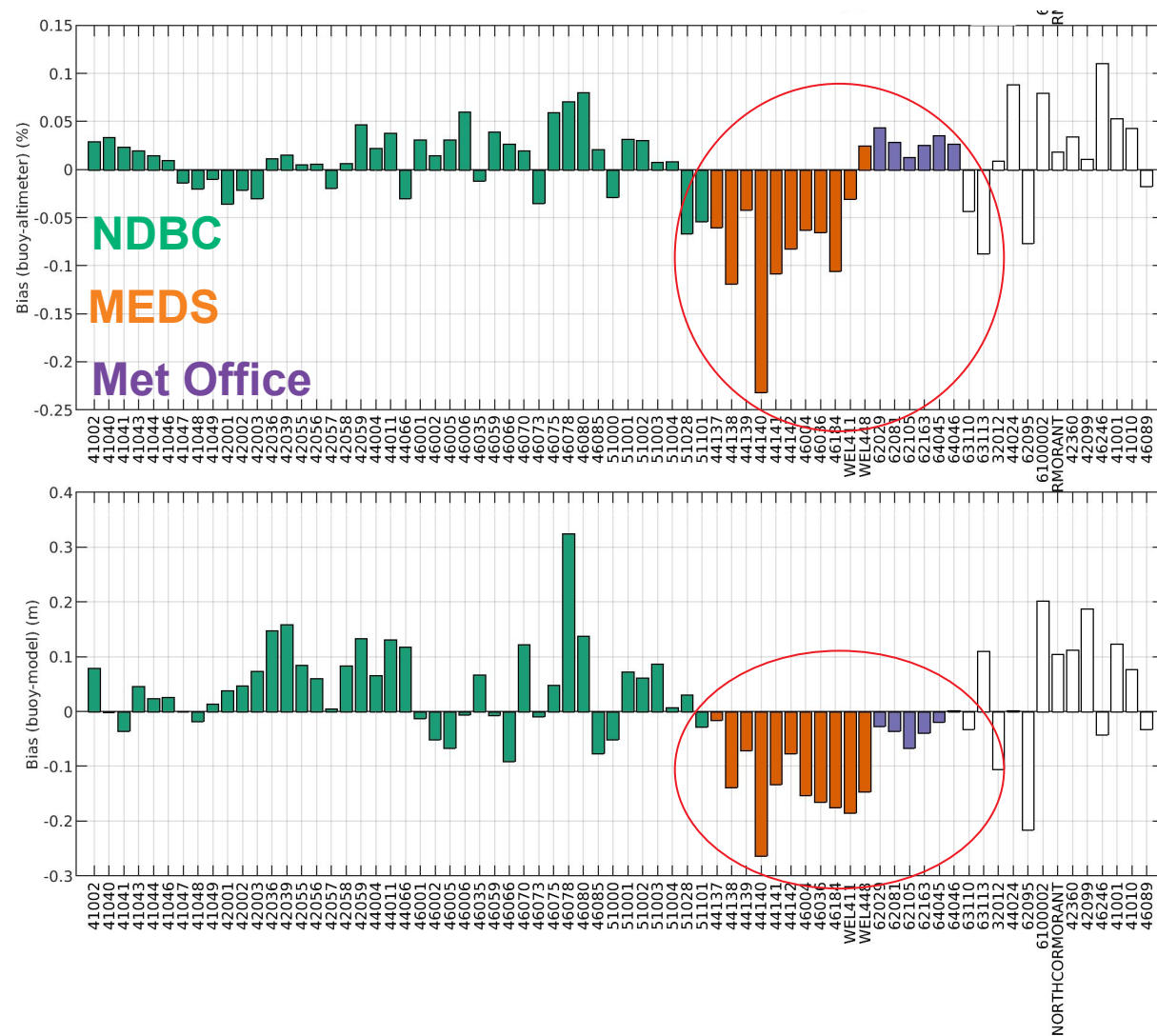
In situ platform errors

- Normalized error are computed for 72 platforms (platform with matchups < 1000 or with Hs resolution $\geq 0.5\text{m}$ are rejected)
- Normalized errors show inter-platform variability with values ranging between 3-10%.
- Lowest normalized error (3.4%) is obtained for platform 46246 (OWS Papa) in North Pacific (DWR MKIII 0.9m diameter buoy owned and maintained by University of Washington Applied Physics Laboratory)
- No specific trends are identified among the three main data providers (NDBC, MEDS and Met Office)

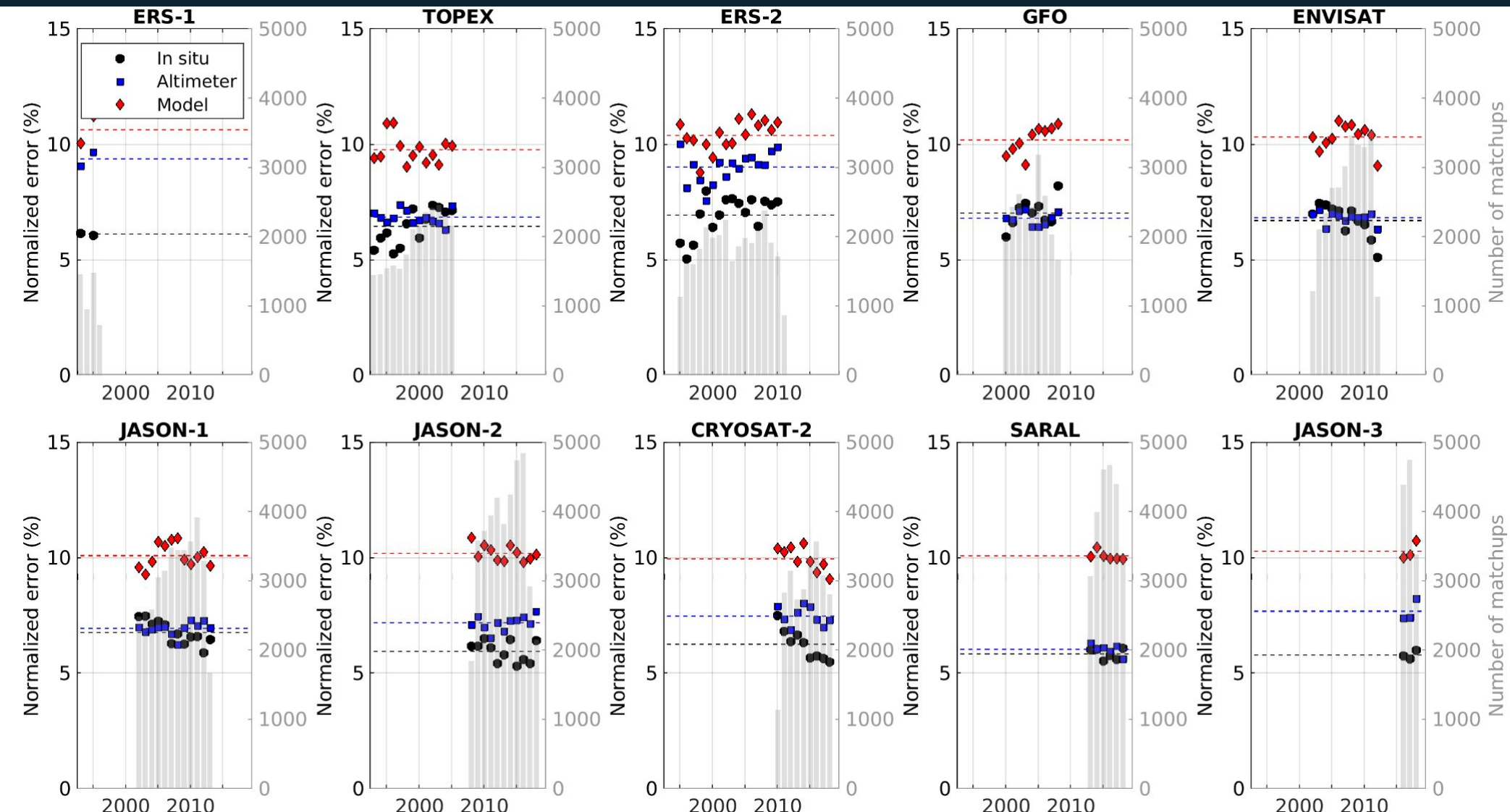


In situ platform errors

- If we look at the bias between in situ platforms and altimeter and between in situ platforms and model we see a consistent negative bias of the MEDS network. This corroborates the finding of Durrant et al (2009) who found that MEDS buoys were underestimating H_s relative to NDBC buoys by about 10%
- Met Office in situ platforms show a consistent positive bias relative to altimeter data and a consistent negative bias relative to model data, which may result from a positive bias of the model relative to the altimeter in European seas

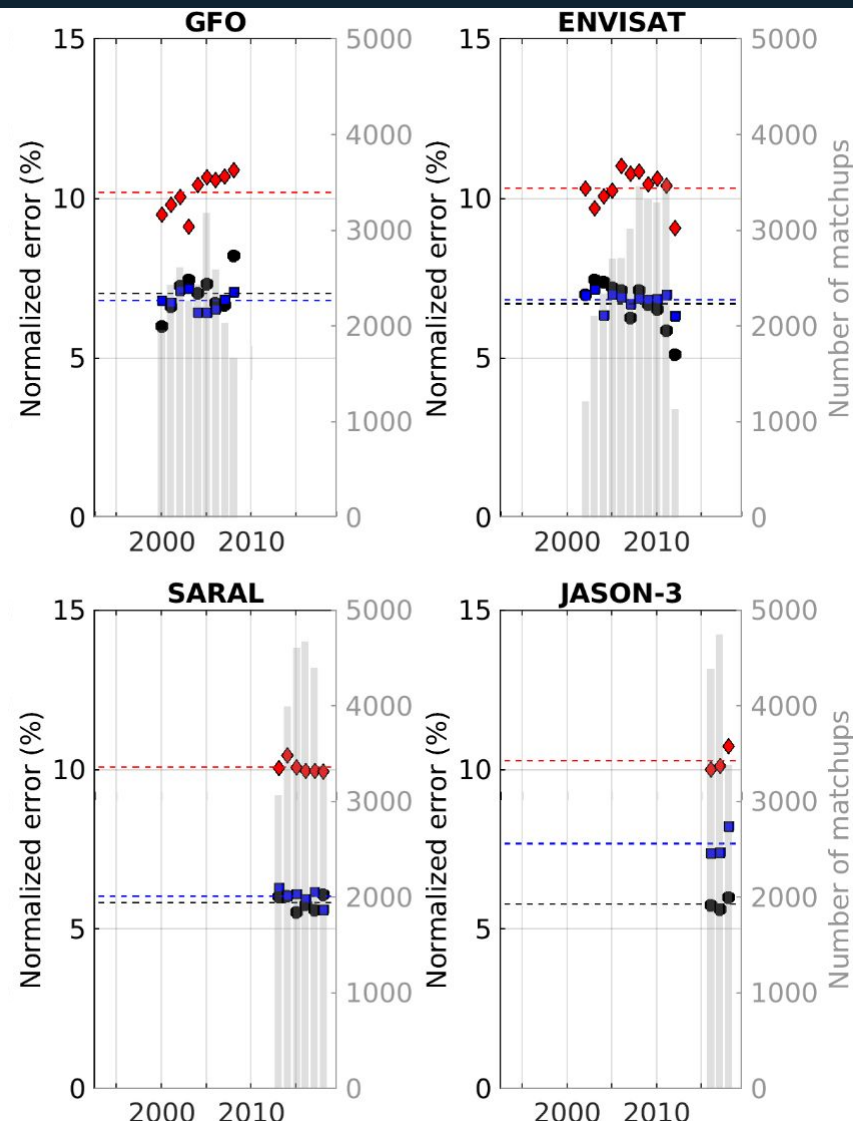


Altimeter errors



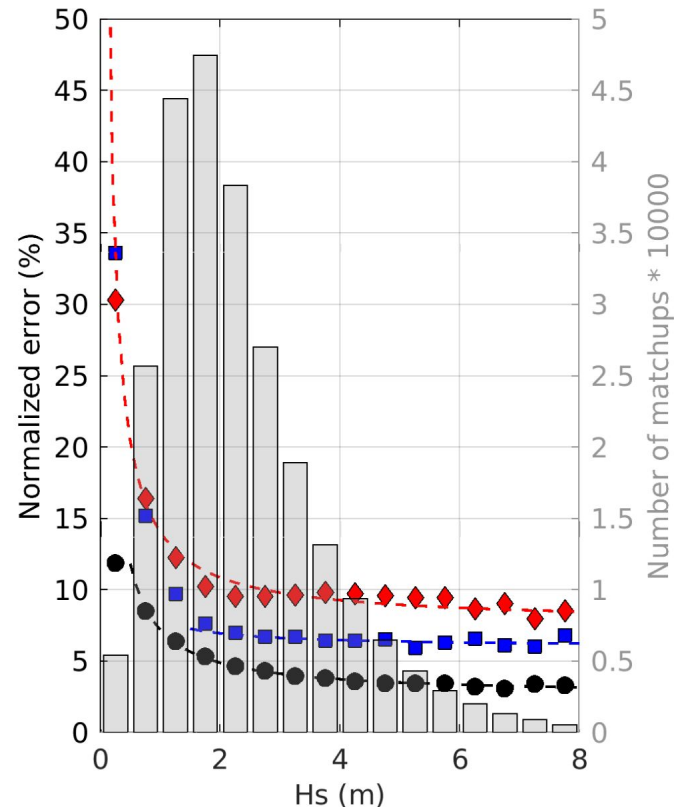
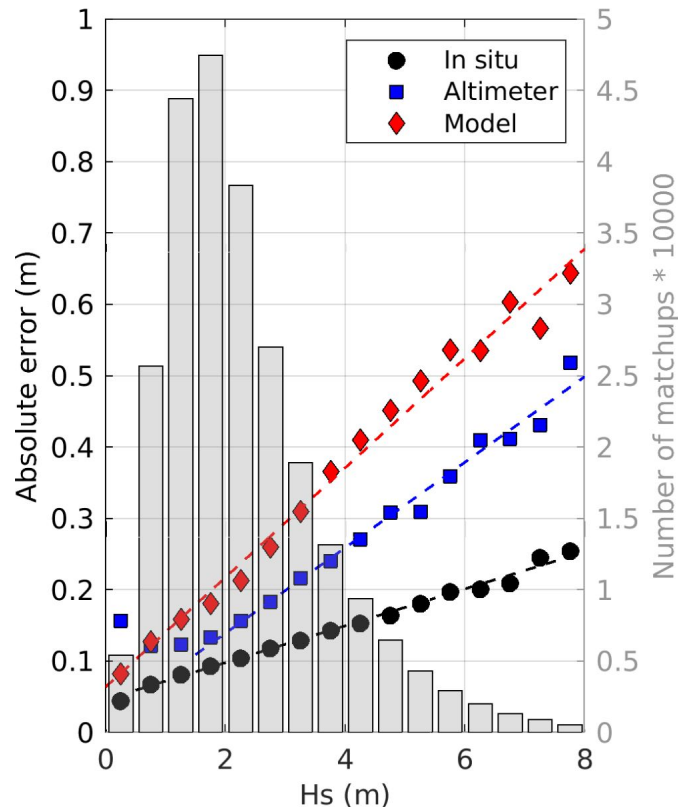
Altimeter errors

- Time-series of number of matchups and normalized errors at yearly timescale are computed for each mission (years with N matchups < 1000 are rejected)
- Trends in number of matchups can be attributed to the increasing number of in situ measurements
- Inter-mission variability in number of matchups can be attributed to difference in satellite orbit inclination
- Significant inter-annual variability in the normalized error of altimeter records (blue squares) can be attributed to several factors: varying available situ platforms from year to year, issues with the radar altimeter system affecting the quality of the measurements, inter-annual variability of the wave climate
- Largest mean normalized errors (blue dashed line) are obtained for ERS-1 and ERS-2 (9.29% and 8.90%, respectively), while lowest mean normalized error is obtained for SARAL (5.75%)
- The second lowest mean normalized error (6.69%) and standard deviation (0.26%) is obtained for the ENVISAT mission
- Differences in the mean normalized error between each mission can be attributed to the different radar technologies and differences in the satellite orbits and observed wave climate



Hs-dependance of measurement errors

- Errors in the 3 measurement system are estimated as a function of H_s by computing triple collocation results for bins of 0.5-m width, from 0 to 8 m. To obtain enough data in each bin, we assume that all altimeter and all in situ platforms share similar error characteristics, respectively, so that we can merge them.
- A quasi-linear relationship between the absolute error and H_s is obtained
- At low H_s (<1m), altimeter errors are the highest (0.15 m, 34%) and in situ errors are the lowest (0.03 m, 12%).
- For H_s comprised between 0.5-8 m, measurements errors are systematically lower for in situ measurements, followed by altimeters and model data.
- The slope of the error function is twice as high for altimeter (0.06) and model (0.07) data as for in situ measurements (0.03)



Conclusions

Conclusions

- A triple collocation database between altimeter (ESA CCI Sea State dataset), in situ platforms (CMEMS INSTAC) and wave model (WW3 T475) has been produced to investigate error characteristics in wave products
- In situ measurements, as a whole, show the lowest mean normalized errors (6.6%), followed by combined altimeter records (7.2%) while errors of the model (9.6%) are the highest among the three datasets
- Results show some inter-platform and inter-mission error variabilities, which can be attributed to the different measurement devices and radar technologies, as well as the observed wave climate characteristics
- Absolute errors of the three measurement system all show a linear relationship with sea state conditions for $H_s > 1.5\text{m}$, with lowest slope for in situ platform and steepest slope for model
- At low sea state ($H_s < 1\text{ m}$), altimeter data show the largest errors, followed by model and in situ platform measurements

References

- Abdalla, S., Chiara, G.D., 2017. Estimating Random Errors of Scatterometer, Altimeter, and Model Wind Speed Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10, 2406–2414. <https://doi.org/10.1109/JSTARS.2017.2659220>
- Alday, M., Ardhuin, F., Accensi, M., Dodet, G., 2021. A global wave parameter database for geophysical applications. Part 3: improved forcing and spectral resolution. *Earth and Space Science Open Archive*. <https://doi.org/10.1002/essoar.10505476.3>
- Dodet, G., et al. 2020. The Sea State CCI dataset v1: towards a sea state climate data record based on satellite observations. *Earth System Science Data* 12, 1929–1951. <https://doi.org/10.5194/essd-12-1929-2020>
- Durrant, T.H., Greenslade, D.J.M., Simmonds, I., 2009. Validation of Jason-1 and Envisat Remotely Sensed Wave Heights. *J. Atmos. Oceanic Technol.* 26, 123–134. <https://doi.org/10.1175/2008JTECHO598.1>
- Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., Wagner, W., 2016. Recent advances in (soil moisture) triple collocation analysis. *International Journal of Applied Earth Observation and Geoinformation, Advances in the Validation and Application of Remotely Sensed Soil Moisture - Part 1* 45, 200–211. <https://doi.org/10.1016/j.jag.2015.09.002>
- Janssen, P.A.E.M., Abdalla, S., Hersbach, H., Bidlot, J.-R., 2007. Error Estimation of Buoy, Satellite, and Model Wave Height Data. *J. Atmos. Oceanic Technol.* 24, 1665–1677. <https://doi.org/10.1175/JTECH2069.1>
- Portabella, M., Stoffelen, A., 2009. On Scatterometer Ocean Stress. *Journal of Atmospheric and Oceanic Technology* 26, 368–382. <https://doi.org/10.1175/2008JTECHO578.1>
- Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research: Oceans* 103, 7755–7766.
- Vogelzang, J., Stoffelen, A., 2012. Triple collocation (NWPSAF Technical report No. NWPSAF-KN-TR-021), EUMETSAT. <https://doi.org/10.13140/RG.2.2.30926.66888>
- Xu, F., Ignatov, A., 2016. Error characterization in iQuam SSTs using triple collocations with satellite measurements. *Geophysical Research Letters* 43, 10,826–10,834. <https://doi.org/10.1002/2016GL070287>

Thank you !