

A global gridded dataset of significant wave height via fusion of multi-mission altimetry and numerical hindcast

Haoyu Jiang ^{1,2*}, Hao Su ¹

¹College of Life Sciences and Oceanography, Shenzhen University, Shenzhen 518000, China

²Laboratory for Regional Oceanography and Numerical Modeling, Qingdao Marine Science and Technology Center, Qingdao 266000, China

Correspondence to: Haoyu Jiang (Haoyujiang@szu.edu.cn)

Abstract. Satellite altimeters provide long-term, high-accuracy observations of significant wave height (SWH) over the global ocean. However, their sparse spatial and temporal sampling introduces undersampling errors in wave climate analyses. Direct gridding of multi-mission altimeter data through spatiotemporal interpolation still results in low-accuracy global SWH fields due to this limitation. To overcome this challenge, we use SWH outputs from a WAVEWATCH III hindcast as a background field and apply an offline fusion with along-track, jointly calibrated altimeter observations from the Climate Change Initiative Sea State dataset. As a retrospective reconstruction, the offline fusion allows observations acquired both before and after a target time to contribute to the estimate, with their influence explicitly constrained by temporal separation. Validation against buoy measurements and withheld satellite data demonstrates that the fused gridded product achieves high accuracy. To address different application needs, we provide two versions of the fused dataset: 1) A “two-sat” version that incorporates data from only two satellites at any given time, designed for wave climate studies. This configuration is designed to reduce sampling-related temporal inhomogeneity by maintaining a relatively stable number of observations and broadly consistent orbital sampling characteristics throughout the record. 2) A “multi-sat” version that integrates data from as many altimeter missions as possible, intended to support applications such as the training of artificial intelligence-based wave models, where higher spatial and temporal accuracy is prioritized. The dataset is freely available at <https://doi.org/10.57760/sciencedb.29314> (Su and Jiang, 2025).

1 Introduction

Wind-generated surface gravity waves (hereafter referred to as waves) are a ubiquitous oceanic phenomenon and play a crucial role in air–sea interactions. A statistical understanding of wave behaviour, commonly referred to as the wave climate, and its variability under changing climate conditions is vital from both scientific and societal perspectives (e.g., ocean engineering, coastal protection, marine renewable energy, and climate assessment). As a result, wave climate has been extensively studied over the past two decades.

To investigate historical wave climate and its variability, long-term wave datasets are essential. Currently, three primary sources of wave data are commonly used in wave climate studies: in situ observations, satellite remote sensing, and Numerical Wave Models (NWMs), each with its own strengths and limitations. Among them, in situ wave buoys with long-term records are very limited in number and are predominantly located along the coastlines of developed countries. This sparse and uneven spatial distribution makes them unsuitable for large-scale wave climate analyses.

Satellite remote sensing, particularly satellite altimetry, provides high-accuracy measurements of global Significant Wave Height (SWH) and has accumulated over 30 years of observations (e.g., Zieger et al., 2009; Ribal and Young, 2019; Dodet et al., 2020). Systematic inter-calibrations of multiple past and present altimeter missions have been conducted by Ribal and Young (2019) and Dodet et al. (2020), resulting in consistent multi-mission SWH datasets spanning more than three decades.

These datasets have enabled comprehensive investigations of long-term global wave climate variability with respect to SWH (e.g., Young and Ribal, 2020; Timmermans et al., 2020).

However, satellite altimeters and other spaceborne wave-observing instruments, such as synthetic aperture radars operated in wave mode and wave spectrometers (e.g., Surface Waves Investigation and Monitoring instrument onboard China-France Oceanography Satellite), sample the ocean only along ground tracks or at spatially discrete footprints, resulting in sparse spatiotemporal coverage. Altimeters typically have an across-track spacing of 300–500 km and revisit periods of 10 days or more. This sparse sampling often fails to capture rapidly evolving sea states or extreme events, introducing significant undersampling errors in wave climate analyses, particularly in the estimation of high percentiles (e.g., 90th and 99th) and their trends. Such errors may lead to misleading conclusions (Jiang, 2020).

To support robust SWH-based wave-climate studies, a long-term, accurate, and seamless (gridded) wave dataset is essential. NWMs have been widely used to reconstruct historical global wave climate by forcing the models with historical wind fields (e.g., Sterl and Caires, 2005; Semedo et al., 2011, 2015; Fan et al., 2012; Stopa et al., 2014; Meucci et al., 2018). When forced with projected wind fields from climate models, NWMs can also be used to estimate future wave conditions (e.g., Fan et al., 2013, 2014; Semedo et al., 2013). However, NWMs are subject to both physical and numerical limitations. While they can reproduce the general patterns of wave height distributions, their ability to accurately simulate extreme wave events remains a subject of debate. In addition, the wind hindcasts used to drive NWMs often suffer from temporal inhomogeneities due to changes in the amount and quality of assimilated atmospheric observations over time, potentially introducing non-climatic artifacts into the wave hindcast data.

Another type of gridded wave dataset is the Level-4 product derived from multi-mission altimeter measurements through spatiotemporal interpolation, such as the *Global Ocean L4 SWH from NRT Satellite Measurements* (CMEMS, 2024, <https://doi.org/10.48670/moi-00180>). However, due to the severe undersampling inherent to satellite altimetry and the high spatiotemporal variability of ocean waves, such interpolation-based products still suffer from significant limitations. Specifically, they typically have coarse spatial and temporal resolutions (e.g., $2^\circ \times 2^\circ \times 1$ day). More importantly, the accuracy of these “independent” Level-4 products is often markedly lower than that of the original along-track observations, and in many cases, even lower than that of NWM outputs (see Section 3 for related results).

A more promising approach for generating a high-quality SWH dataset is to combine the seamless spatiotemporal coverage of NWM hindcasts with the higher accuracy of satellite altimeter observations, which generally have an RMSE of approximately 0.2 m when validated against offshore buoy measurements (Jiang, 2023). Such fusion or assimilation strategies aim to leverage the strengths of both data sources. One widely used example is the ERA5 wave reanalysis dataset (Hersbach et al., 2020), which is part of the ERA5 climate reanalysis covering the period from 1950 to the present. It provides hourly data for many wave parameters, including SWH, with a resolution $0.5^\circ \times 0.5^\circ \times 1$ h. In the ERA5 wave model, SWH observations from multiple altimeter missions, including ERS-1/2, ENVISAT, JASON-1/2, CRYOSAT-2, and SARAL, are assimilated using a spatial optimal interpolation method. This assimilation reduces errors in the NWM hindcast to some extent. Another example of wave reanalysis dataset is the WAVERYs, which also assimilates SWH observations from historical altimeter missions (Law-Chune et al. 2021). However, the “online” assimilation methodology in ERA5 and WAVERYs presents several limitations when used for historical SWH data reconstruction:

1) Lack of joint calibration: The altimeter SWH data used in ERA5 are not jointly calibrated across missions. Instead, observations from individual satellites are first regressed to match the model hindcast baseline before assimilation. While this helps mitigate inter-mission biases, it may also suppress real climate signals contained in the altimeter data due to biases in the model background.

2) Forward-only assimilation: The assimilation scheme in ERA5 considers only past and current SWH observations to constrain the future wave state. Due to the absence of a backward model (which is challenging for NWMs because the alleviation of the garden sprinkler effect is difficult to reverse), the assimilation scheme does not account for the influence of

future altimeter SWH observations on past wave states, thereby limiting the extent to which observational data can be fully utilized in reconstructing historical wave fields.

3) Uncertainty in spectral corrections: In “online” assimilation of NWMs, adjustments must be made to the shape of the directional wave spectra after the correction of SWH. Several approaches exist for correcting spectral shapes from SWH alone, but these methods can introduce additional uncertainty. In some cases, improper spectral corrections may even degrade the overall accuracy of the assimilated results.

To address the limitations outlined above, this study presents a dataset that fuses jointly calibrated satellite altimeter observations with NWM hindcast outputs. The proposed fusion framework addresses the above limitations in different ways. First, it uses the jointly calibrated CCI Sea State altimeter record as the observational reference, thereby avoiding calibration against the model background. Second, because the fusion is performed offline after the hindcast has been generated, both past and future altimeter observations can be used to constrain the SWH field at a given analysis time. Third, this retrospective constraint also reduces the dependence of the reconstruction on the forward spectral propagation in the wave model. Therefore, no adjustment of spectral shape is required, and the associated uncertainty is therefore avoided. The resulting dataset is therefore neither an observation-only gridded product nor an improved numerical simulation, but an offline fused SWH analysis constrained by both sources. This dataset effectively mitigates the undersampling issue inherent to altimetry while preserving the wave variability signals captured by the altimeters, making it well suited for SWH-based wave-climate studies. In addition, the rapid advancement of deep learning techniques has led to growing interest in data-driven wave modelling. These models typically require seamless and reliable datasets as “truth” for training. Therefore, beyond its value for wave climate analysis, the fused dataset introduced in this study also provides a valuable resource for a wide range of applications, including the development of deep learning-based wave models and SWH post-correction models.

This paper describes the production methodology, validation, and access information for the fused dataset. Section 2 introduces the jointly calibrated altimeter observations and NWM hindcasts used as inputs, along with the reference datasets employed for validation and comparison, including buoy measurements and a Level-4 gridded SWH product. Section 3 details the processing steps for generating the fused dataset. Section 4 presents its evaluation results, followed by a discussion in Section 5.

2 Data and Methods

2.1 Data

2.1.1 CCI-Sea State Dataset

The most important altimeter dataset used in this study is the Level-3 1-Hz significant wave height (SWH) product from the CCI (Climate Change Initiative) Sea State dataset (Dodet et al., 2020), with each record associated with its own time, latitude, and longitude. Here, “1-Hz” denotes along-track observations provided at approximately one measurement per second (corresponding to a spatial resolution of about 7 km). The CCI-Sea State project, developed by the European Space Agency, aims to provide a long-term and consistent satellite-based record of sea state. This dataset integrates observations from multiple radar altimeter missions, including ERS-1/2, TOPEX, JASON-1/2/3, ENVISAT, CRYOSAT-2, SARAL, and Sentinel-3A/3B/6. Rigorous quality control, cross-calibration, and denoising procedures are applied to enhance the accuracy and consistency of SWH measurements across different sensors. In particular, denoising is conducted using a non-parametric empirical mode decomposition (EMD) method to reduce random noise (Huang et al., 1998; Quilfen and Chapron, 2021), which effectively suppresses Gaussian noise while retaining small-scale variability in the data (Quilfen and Chapron, 2019, 2021).

Currently, the up-to-date publicly available version of the CCI-Sea State dataset is version 4 (v4). This dataset spans the period Aug 1991-Jan 2024. In previous versions, data of version 3 have undergone complete and consistent retracking of all included altimeters, ensuring greater uniformity across the dataset, whereas version 1 largely inherited the processing

approaches from the GlobWave project and did not apply this level of consistency. Compared to version 3, more satellites are included in v4, and SWH data from them have been more carefully re-edited and re-bias-corrected using altimeter tandem collocations, altimeter-buoy collocations, and altimeter cross-collocations in order to ensure better consistency between missions and instruments. This new altimeter dataset represents to date the longest continuous global wave climate data record. This dataset is now available at <https://data-cersat.ifremer.fr/data/ocean-waves/cci-seastate/v4/>, and a product user guide is available at <https://cciseastate.gitlab-pages.ifremer.fr/ccidoc/intro.html>.

2.1.2 WAVEWATCH III ST4 and ST6 Hindcasts

Assimilative wave reanalyses such as ERA5 are not suitable as background fields for the present study, because subsequent fusion with altimeter observations would lead to repeated use of the same observational information. Therefore, the SWH data from a WAVEWATCH III (WW3) hindcast based on the source term package 6 (ST6), without wave data assimilation, is adopted as the background field. The WW3-ST6 configuration is based on the physical parameterizations of Liu et al. (2021), forced by ERA5 10-m surface winds (with a resolution of $0.25^\circ \times 0.25^\circ \times 1\text{h}$). The model outputs are provided at $0.25^\circ \times 0.25^\circ$ spatial resolution and 3 h temporal resolution. Without the assimilation of wave observations, the hindcast SWHs show good agreement with measurements, with an overall RMSE of ~ 0.35 m (5–15% of SWH) when compared with altimeter observations in the open ocean. Further details and access to the dataset are provided in Liu et al. (2021).

To assess the sensitivity of the results to the choice of background field, an additional NWM hindcast based on WW3 source term package 4 (ST4) is also used in this study. This dataset, known as the Integrated Ocean Waves for Geophysical and Other Applications hindcast, is based on the parameterizations of Ardhuin et al. (2010) and is forced by global 10-m wind fields from the Climate Forecast System Reanalysis (CFSR). The WW3-ST4 hindcast has also been shown to perform well against both buoy and altimeter observations of SWH (Rascle and Ardhuin, 2013), and it is publicly available via the IFREMER FTP server (<ftp.ifremer.fr>).

It is noted that a more recent version of WW3-ST6, forced by ERA5 winds, has been released and provides improved performance (Alday et al., 2023). However, in the present study our objective is to assess the sensitivity of the fused product to plausible differences in the background field. For this purpose, a dataset with more distinct differences from WW3-ST6 was required. Therefore, we use the earlier CFSR-forced WW3-ST4 hindcast as a contrasting background field.

2.1.3 CMEMS Gridded SWH Dataset

The gridded multi-mission merged satellite SWH data product from CMEMS (Copernicus Marine Environment Monitoring Service) is also used here to show that it is difficult to generate a good-quality gridded SWH data product by using only altimeter data. CMEMS multi-mission gridded SWH merges multiple along-track altimetry SWH observations (Sentinel-6A, Jason-3, Sentinel-3A, Sentinel-3B, SARAL/AltiKa, Cryosat-2, CFOSAT, SWOT-nadir, HaiYang-2B and HaiYang-2C) and produces daily gridded data. Two SWH fields are produced, including a 2° “daily mean” field computed from all available level 3 along-track measurements from 00:00 UTC until 23:59 UTC and a 0.5° “instantaneous” field at 12:00 UTC using the same data source but accounting for their spatial and temporal proximity. However, after checking the data, it is found that the 2° “daily mean” field is not a “seamless” field with many grid points without data. Thus, only the “instantaneous” fields were used in this study. The data are available at <https://doi.org/10.48670/moi-00180> (CMEMS 2024).

2.1.4 Buoy Observations

The hourly SWH records derived from 20-min wave sampling periods obtained from the National Data Buoy Center (NDBC) are used as a reference in this study. We used observations from 50 NDBC buoys located more than 100 km offshore, covering the period 2010–2016, with buoy locations shown in Fig. 1. Although the NDBC data have undergone standard quality control, some spurious values remain; therefore, an additional screening was applied, retaining only records with $0.15\text{ m} <$

SWH < 15 m. It is noted that these buoy data are used only for independent validation and does not represent the temporal coverage of the fused dataset or a climatological reference period.

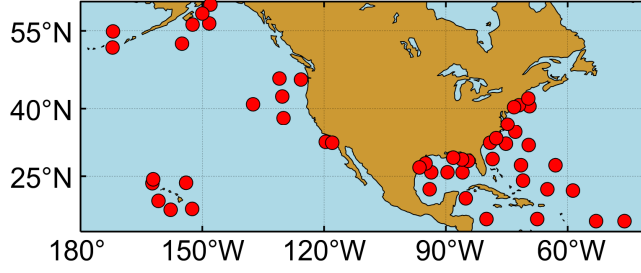


Fig.1. Locations of the 50 NDBC buoys used in this study.

2.2 Fusion Methods

Prior to the space-time weighted fusion, a simple global one-dimensional look-up-table correction is applied to reduce the systematic SWH-dependent bias of the WW3-ST6 background relative to the calibrated CCI Sea State altimeter observations. This correction is analogous to a nonlinear calibration. Collocated WW3-ST6 and CCI Sea State SWH pairs are grouped according to the WW3-ST6 SWH value, and the mean altimeter-minus-WW3 difference is calculated within each 0.1-m SWH bin. The resulting bin-mean correction is then applied to the WW3-ST6 SWH field and to the model-equivalent SWH values collocated with each altimeter record. For SWH values above 12 m, where collocated samples are relatively sparse, a simple linear-regression correction is used instead of fitting individual bins, in order to avoid overfitting in the high-wave range. All subsequent data-fusion procedures are therefore based on the look-up-table-corrected hindcast background rather than on the original hindcast.

The data-fusion method employed in this dataset is a space-time weighted fusion approach inspired by optimal interpolation and objective analysis, originally developed for meteorological and oceanographic data analysis (e.g., Bretherton et al., 1976). In this framework, sparse observations are used to correct a background field through distance- or covariance-dependent weights. The present study modifies this framework for SWH reconstruction to give strong local influence to calibrated altimeter SWH observations within the fused product with the corresponding formula expressed as follows:

$$A(i, j, t) = M(i, j, t) + \sum_{k=1}^N w(i, j, t)_k \cdot (O_k - M_k) \quad (1)$$

$$w(i, j, t)_k = \exp\left[-d_k^2(i, j, t) / 2R(i, j, t)^2\right] / \sum_{k=1}^N \exp\left[-d_k^2(i, j, t) / 2R(i, j, t)^2\right] \quad (2)$$

$$R(i, j, t) = \min[d_k(i, j, t)] \quad (k = 1, 2, 3, \dots, N) \quad (3)$$

$$d_k(i, j, t) = C + \sqrt{(S_k / S_l)^2 + (T_k / T_l)^2 + Q \left[\frac{(M_k - M(i, j, t))}{M_k} \right]^2} \quad (4)$$

$$d_k(i, j, t) = +\infty \quad \left(\begin{array}{l} (S_k < S_{thr} \ \& \ T_k < T_{thr} \ \& \ r_{i,j,k} > r_{thr}) \\ (S_k > S_{thr} \ or \ T_k > T_{thr} \ or \ r_{i,j,k} < r_{thr}) \end{array} \right)$$

where i, j , and t represent longitude, latitude, and time, respectively; M (model) and A (analysis) represent the NWM hindcast data before and after data fusion, respectively; $k \in \{1, 2, \dots, N\}$ represents the index of altimeter observations; O_k and M_k represent the values of altimeter-observed and corresponding collocated modelled SWH at the k -th altimeter record; w_k represents the weight factor at spatio-temporal location (i, j, t) for the k -th altimeter record; d_k represents the distance from the model location (i, j, t) to the k -th altimeter record in a super-space spanned by space, time, and SWH, which is defined in Eq.(4); R represents the distance to the nearest altimeter record; C is a constant to mitigate the impact of cases with very small distances; S and T represent the spatial distance and the time difference between (i, j, t) and the k -th altimeter record, respectively, and S_l and T_l represent coefficients that make S and T dimensionless; Q is also a scaling coefficient that ensures the quantities of SWH, time, and space are comparable; $r_{i,j,k}$ is the correlation coefficient (CC) between the NWM-hindcast SWHs time series at grid location

(i, j) and the time series at the location of the k -th altimeter record within the corresponding month; S_{thr} , T_{thr} , and r_{thr} are threshold values for S_k , T_k , and $r_{i,j,k}$, respectively: An observation is excluded from consideration if the space-time separation or the correlation between the model point and the observation point exceeds these thresholds.

195 This method implicitly gives strong confidence to calibrated altimeter SWH observations relative to the non-assimilative WW3 background. The uncertainty of the resulting product is therefore assessed empirically through independent buoy validation, leave-one-satellite-out altimeter validation, parameter-sensitivity tests, and background-sensitivity experiments, rather than through a formal analytical error-covariance model.

200 To reduce potential space-time representativeness errors, all altimeter observations located within 75 km of coastlines were excluded, as SWH can vary substantially over short distances in nearshore regions due to bottom-wave interactions and coastal sheltering.

Here, C , S_I , T_I , Q , S_{thr} , T_{thr} , and r_{thr} are set to 8, 50 km, 90 min, 8000, 2000 km, 24 h, and 0.7, respectively, based on a straightforward empirical tuning procedure which was conducted using the 2011 CCI-Sea State dataset. Specifically, Jason-2 data were withheld, and the remaining three altimeter missions (CryoSat-2, ENVISAT, and Jason-1) were fused with the WW3-ST6 hindcast data. The fused results were then validated against the independent Jason-2 observations, and the combination of tuning parameters (C , S_I , T_I , Q , S_{thr} , T_{thr} , r_{thr}) that yielded relatively small overall validation errors was selected. The error metrics considered in this evaluation include bias, RMSE, and correlation coefficient (CC):

$$Bias = \frac{1}{n} \sum_{i=1}^n (y_i - x_i) \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2} \quad (6)$$

$$CC = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (7)$$

where x and y denote the SWH from the dataset to be evaluated and the reference data (in this case, fused dataset without Jason-2 and Jason-2 from CCI-Sea State), respectively; n is the sample size, and the bars over them denote their mean values. It is worth noting that the fused results are not sensitive to these tuning parameters near their optimal values. Moderate changes in these tuning parameters only weakly affect the fused results. This is because most target grid points are simultaneously constrained by spatial separation, temporal separation, modelled-SWH similarity, background-field correlation, threshold-based observation selection, and the base-distance term. The fusion therefore mainly represents an interpolation of altimeter-minus-background increments among dynamically related observations, rather than an extrapolation controlled by a single parameter.

With respect to the construction of the weights, while conventional objective analysis and optimal interpolation methods often define weights primarily according to spatial separation, our approach introduces several additional terms to better reflect the characteristics of wave fields:

220 1) *Introduction of the temporal distance.* The distance definition incorporates the temporal separation between the model grid point to be corrected and the observation point. This accounts for the fact that the sea state at a given location is influenced not only by nearby observations in space but also by those in time. Consequently, both past and future observations are allowed to contribute to the correction at a given time, unlike most operational wave assimilation schemes, which generally exclude future data when updating past states.

225 2) *Introduction of SWH differences.* The difference in modelled SWH between the model grid point to be corrected and that with observation is also included. This modification addresses the anisotropic nature of storm-generated wave fields, particularly swell. Due to model arrival-time errors (Jiang et al., 2016), modelled wave fields may be spatially displaced relative to altimeter measurements. By assuming that “similar modelled SWHs within the same meteorological event imply similar corrections”, this SWH-based distance partly accounts for anisotropy, giving relatively higher weight to observations within the same storm wave system compared to those outside. The inclusion of the modelled-SWH difference term reduces

the global RMSE only slightly, by approximately 0.005 m, and this small reduction should not be over-interpreted as a major improvement in the domain-mean error metric. The term is used mainly because it improves the physical consistency of local correction increments under model-observation displacement conditions. When a modelled storm or swell system has an arrival-time or position error, geographically or temporally nearby observations may correspond to dynamically different sea states. The SWH-based distance helps reduce such inappropriate influences by assigning larger weights to observations with similar modelled SWH and smaller weights to observations that are likely outside the wave system.

3) *Introduction of a base distance C*. A constant base distance is included to mitigate cases where very small distances (approaching zero) would otherwise lead to unrealistically large weights. Without this adjustment, the fused results exhibit prominent along-track stripes corresponding to the altimeter ground tracks, which are physically implausible. During the tuning of C , visual inspection was first used to ensure the absence of such striping, followed by optimization to minimize RMSE.

It is noted, however, that no dedicated sea-ice-specific correction, dynamic ice-edge filtering scheme, or coastal-wave correction is applied in the present fusion procedure. Sea-ice-affected regions are treated according to the availability and quality control of the input WW3 and CCI Sea State data, together with the general observation-screening criteria described above. Additional uncertainties are therefore expected near sea-ice margins, where the WW3 background may be less accurate because of ice-wave interactions and possible ice-mask uncertainties, and where valid altimeter observations are generally sparser and may provide weaker local constraints. Similarly, although altimeter observations within 75 km of coastlines are excluded to reduce nearshore representativeness errors, fused SWH values near coastal regions should also be interpreted with caution because no dedicated coastal-wave correction is applied. Therefore, the product is primarily intended for open-ocean wave applications, and values near coastlines and sea-ice margins should be used with caution.

2.3 Two Types of Fused Data

During the data fusion process, the number of altimeter satellites used plays a critical role. In principle, fusing more satellites increases the spatiotemporal coverage of the observations, which should, in turn, improve the accuracy of the fused dataset. However, as pointed out by Jiang (2020), the growing number of altimeters over time also increases the likelihood of sampling extreme events. This may introduce undersampling biases in long-term climate analyses. Although the fusion process can mitigate this issue to some extent, it may still remain a potential source of error in climate-related applications.

To address this concern, we adopt a strategy similar to that used in the Maps of Sea Level Anomalies product of AVISO, generating two complementary datasets: a “*multi-sat*” product, which integrates observations from all available altimeter satellites in CCI-Sea State, and a “*two-sat*” product, which consistently fuses data from only two altimeters. The “*multi-sat*” product prioritizes accuracy by maximizing observational coverage, whereas the “*two-sat*” product is designed to reduce sampling-related temporal inhomogeneity by maintaining a relatively stable number of observations and broadly consistent orbital sampling characteristics. Here, sampling consistency refers to reducing temporal changes in observation density and orbital sampling characteristics and is distinct from the cross-mission calibration consistency already provided by the CCI Sea State dataset. From an application perspective, the “*multi-sat*” product is better suited for tasks that demand higher accuracy, such as model evaluation and the training of deep learning algorithms. In contrast, the “*two-sat*” product is more appropriate for long-term climate analyses, where temporal consistency is of greater importance.

For the “*two-sat*” product, the selection of satellites is guided by sampling characteristics: one satellite with an orbital inclination of approximately 66° is paired with another in a near-polar orbit to maintain broadly consistent orbital sampling characteristics over time (Table 1). In cases where a given satellite experiences temporary data gaps, observations from another satellite with a similar orbital configuration are preferentially used as substitutes. For example, in the CCI-Sea State v3 dataset, Jason-2 data were missing during several days in June 2017 and March 2019. In such instances, Jason-3 data were used as replacements for Jason-2. Although the CCI Sea State v4 record starts in August 1991, the fused product starts in October 1992 because, before this time, only one altimeter mission is available in the CCI record and the prescribed two-satellite

configuration cannot be constructed. For consistency, the multi-sat product is released over the same temporal period as the two-sat product.

Table 1. Altimeters used in different periods for the “two-sat” product.

Time	1992/10- 1995/5	1995/6 2002/1	2002/2- 2008/7	2008/8- 2010/6	2010/7- 2016/2	2016/3- 2023/12
Satellite1	TOPEX	TOPEX	JASON-1	JASON-2	JASON-2	JASON-3
Satellite2	ERS-1	ERS-2	ENVISAT	ENVISAT	CRYOSAT-2	CRYOSAT-2

3. Quality assessment of the Fused Dataset

3.1 Comparison against In-situ Data

The CMEMS gridded dataset, the WW3-ST6 dataset, and the fused dataset (“two-sat” version) were all bilinearly interpolated to each buoy location, and subsequently compared against the corresponding SWH of buoy observations. Figures 2a–c show scatter plots comparing buoy observations with the three datasets for instantaneous values at 12:00 UTC.

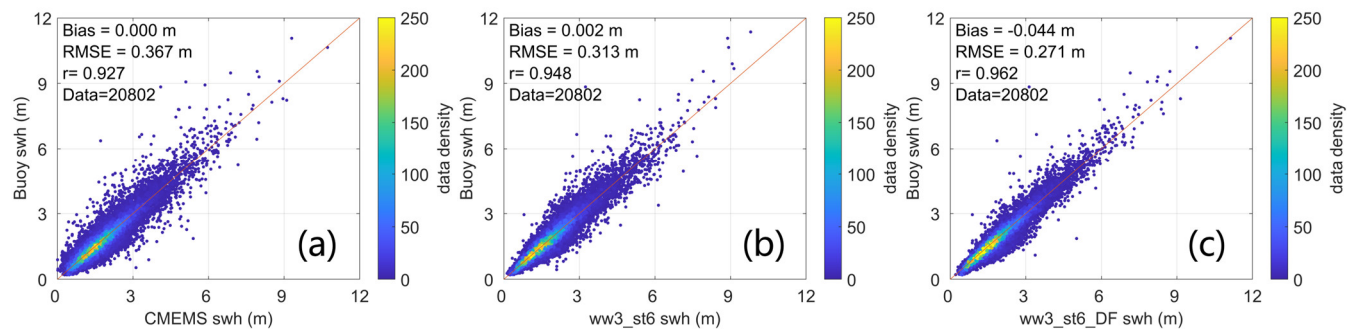


Fig. 2. Scatter plots comparing SWHs from buoy observations with those from (a) the CMEMS (Copernicus Marine Environment Monitoring Service) gridded dataset, (b) the WW3-ST6 dataset, and (c) the fused dataset (“two-sat” version) for instantaneous values at 12:00 UTC during the period 2010–2016. All gridded datasets were bilinearly interpolated to the buoy locations

The comparison of instantaneous values at 12:00 UTC between SWHs from the CMEMS gridded dataset and NDBC buoys (Fig. 2a) shows a bias of nearly zero, an RMSE of 0.37 m, and a CC of 0.93. Although these error metrics may not appear poor at first glance, they are substantially worse than the error metrics obtained from comparisons between along-track altimeter SWHs and buoys over the same period (e.g., Jiang, 2023). Moreover, they are significantly inferior to those from the NWM hindcast (WW3-ST6), which does not assimilate any observations (Fig. 2b, bias = 0.002 m, RMSE = 0.31 m, CC = 0.95). This finding supports our point raised in the Introduction Section that the accuracy of “independent” Level-4 gridded altimeter SWH products is often much lower than that of raw NWM outputs. However, after fusing WW3-ST6 with along-track altimeter observations, the RMSE further decreased to 0.27 m and the CC increased to 0.96, demonstrating the effectiveness of the data fusion approach.

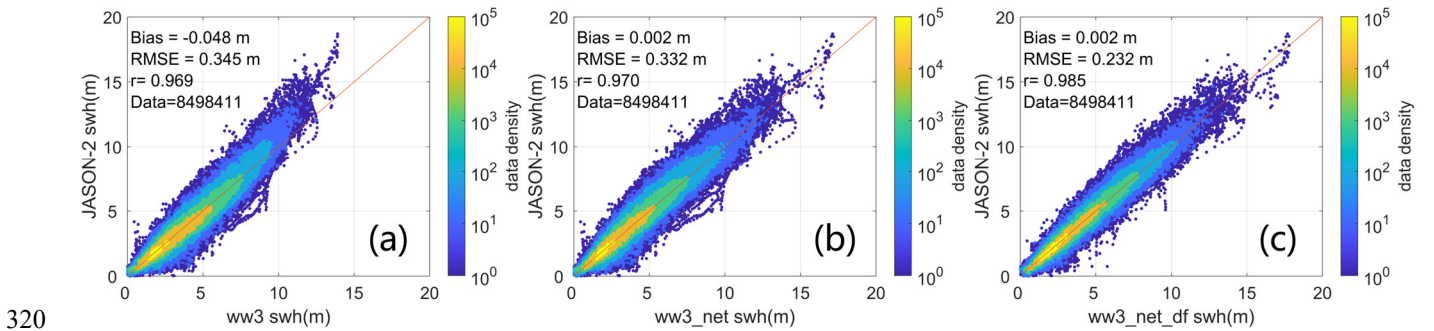
3.2 Comparison against Altimeter Measurements

We further evaluated the global performance of the fused dataset using altimeter observations from 2011. This year was selected because four altimeter missions were available simultaneously, enabling multiple leave-one-satellite-out and two-satellite-combination experiments. In each experiment, one or more satellite missions were excluded from the fusion and then used as withheld references for evaluation. In this validation, one satellite was excluded from the fusion, while the remaining altimeter missions were fused with WW3-ST6. The excluded satellite was then used as an independent reference for evaluation. For example, in Fig. 3, Jason-2 was withheld and subsequently compared against the fused and unfused datasets. It should be noted that, when Jason-2 is used as the withheld reference, Jason-2 observations are excluded from the space-time fusion but were included in the global one-dimensional look-up-table correction. The comparison is therefore independent with respect

305 to the fusion step, while the possible influence of the look-up-table step is expected to be minor because this global correction has a very small effect on RMSE and CC, as shown in Fig. 3b.

Figure 3a shows the comparison between the original WW3-ST6 hindcast and Jason-2. The results indicate that even without assimilation, WW3-ST6 already achieves a relatively high accuracy, with a bias <0.1 m, an RMSE of 0.34 m, and a CC of 0.97. Figure 3b presents the outcome after applying a one-dimensional look-up table correction to WW3-ST6 SWHs using CCI-Sea State data prior to fusion. Although this calibration reduces the bias to a smaller value, its effect on RMSE and CC is negligible, with only a 0.01 m decrease in RMSE and a 0.001 increase in CC. This suggests that the pre-fusion correction was not the main contributor to accuracy improvement. In contrast, the fused dataset (Fig. 3c) shows substantial enhancement relative to Jason-2, with the bias remaining negligible, the RMSE reduced by nearly one third to 0.23 m, and the CC increased by more than 0.015 to 0.985, highlighting the high accuracy of the fused product.

315 Tables 2–4 summarize the evaluation results of bias, RMSE, and CC, respectively. In each experiment, two of the four available satellites were fused with WW3-ST6, while the remaining satellites were used as independent references. It is noted that these experiments are intended to evaluate the stability and achievable accuracy of the fusion method under different observational configurations, rather than to provide a complete year-by-year assessment of the full record, thus, only one year of data is used.



320 **Fig. 3.** Evaluation of the fused dataset using Jason-2 as an independent reference in 2011. (a) Comparison between the original WW3-ST6 hindcast and Jason-2. (b) Results after applying a one-dimensional look-up table correction to WW3-ST6 SWHs using CCI-Sea State data prior to fusion. (c) Comparison between the fused dataset (WW3-ST6 + remaining altimeters) and Jason-2. The fusion substantially reduces the RMSE and improves the correlation compared to both the original and bias-corrected hindcast.

As shown in the tables, when any pair consisting of one satellite with a 66° inclination and one with a near-polar inclination was selected for fusion, all error metrics improved markedly compared with the original WW3-ST6 results. Specifically, the bias, which originally ranged from 0.03 to 0.06 m, was reduced to within 0.01 m. Such a small bias can be regarded as small relative to the mean SWH, but can be meaningful for long-term climate analyses. Meanwhile, the RMSE decreased from 0.36–0.39 m to 0.23–0.25 m, representing a reduction of more than 30%, and the CC increased from about 0.95 to 0.98, also reflecting a substantial improvement. These results demonstrate the robustness and accuracy of the fusion method under different satellite-combination configurations. In particular, the method consistently reduces bias and RMSE and increases CC relative to the original WW3-ST6 background, regardless of which pair of satellites is used for fusion and which mission is used as the withheld reference. Since the final fused dataset is generated using the same fusion framework, the same calibrated CCI Sea State observations, and the same WW3-ST6 background, these results provide strong evidence supporting the robustness and accuracy of the resulting dataset.

Table 2. Bias (in m) between WW3-ST6 and independent altimeter observations, as well as bias between the fused dataset and independent altimeter observations, in 2011. For each experiment, two satellites were selected for fusion with WW3-ST6, while the remaining satellites served as independent references.

Reference Mission	WW3-ST6	CRYOSAT-2 & JASON-1	CRYOSAT-2 & JASON-2	CRYOSAT-2 & ENVISAT	JASON-1 & JASON-2	JASON-1 & ENVISAT	JASON-2 & ENVISAT
CRYOSAT-2	-0.035	—	—	—	0.006	0.006	-0.000
JASON-1	-0.068	—	-0.012	-0.014	—	—	-0.011
JASON-2	-0.059	0.003	—	-0.006	—	0.002	—
ENVISAT	-0.034	0.007	0.000	—	0.008	—	—

Table 3. The same as Table 2, but for RMSE (in m).

Reference Mission	WW3-ST6	CRYOSAT-2 & JASON-1	CRYOSAT-2 & JASON-2	CRYOSAT-2 & ENVISAT	JASON-1 & JASON-2	JASON-1 & ENVISAT	JASON-2 & ENVISAT
CRYOSAT-2	0.368	—	—	—	0.234	0.238	0.240
JASON-1	0.388	—	0.250	0.245	—	—	0.244
JASON-2	0.385	0.250	—	0.244	—	0.243	—
ENVISAT	0.373	0.243	0.242	—	0.232	—	—

Table 4. The same as Table 2, but for CC.

Reference Mission	WW3-ST6	CRYOSAT-2 & JASON-1	CRYOSAT-2 & JASON-2	CRYOSAT-2 & ENVISAT	JASON-1 & JASON-2	JASON-1 & ENVISAT	JASON-2 & ENVISAT
CRYOSAT-2	0.953	—	—	—	0.981	0.980	0.980
JASON-1	0.953	—	0.980	0.981	—	—	0.981
JASON-2	0.952	0.979	—	0.980	—	0.980	—
ENVISAT	0.954	0.981	0.981	—	0.982	—	—

Comparison of the results in Tables 2–4 with those in Fig. 3 indicates that, when Jason-2 is used as the independent reference, the improvement in accuracy from fusing three satellites instead of two is marginal, with the RMSE reduced by only about 0.01 m and the CC increased by just 0.001. Similar results are obtained when other satellites are used as the reference, leading to the consistent conclusion that the additional reduction in errors is insignificant. This can be attributed to the fact that, after fusing two satellites, the overall accuracy of the fused dataset is already very close to the intrinsic accuracy of radar altimeter SWH measurements, which corresponds to a global RMSE of about 0.2 m according to previous studies (e.g., Ribal & Young, 2019; Jiang, 2020; Jiang, 2023). This further demonstrates the robustness of the fusion method.

Although these results suggest that adding more satellites beyond two does not substantially improve accuracy, we still generated a “multi-sat” version of the dataset. This is because, despite the diminishing marginal benefits in terms of accuracy, additional satellites can still provide slight improvements. Moreover, the inclusion of more satellites may help to homogenize the spatial distribution of errors—reducing potential regional biases, especially in gaps between the ground tracks of the fused satellites. Importantly, as long as the altimeter observations themselves are more accurate than the background field (a condition met by all current missions), increasing the number of fused satellites will not degrade the overall accuracy.

Figure 4 shows the spatial distribution of errors corresponding to the data presented in Fig.3. The upper and lower panels respectively display the comparisons between the WW3-ST6 and Jason-2 SWH data before and after data fusion. The three columns represent the spatial distributions of bias, RMSE, and CC. Before fusion, the data exhibit opposite bias patterns between the high latitudes of the Northern and Southern Hemispheres, with the maximum bias reaching nearly 0.5 m. Such large biases also lead to RMSE values exceeding 0.5 m in high-latitude regions. Areas with low CC values are mainly concentrated in low-latitude regions, where the mean SWH is smaller and its annual variability is limited. In high-latitude regions, although the absolute errors are larger, the mean SWH and its variability are also higher, allowing the CC to remain above 0.9. After fusion, except for a few areas near sea ice, the bias is largely reduced to within ± 0.1 m, the RMSE decreases to below 0.5 m, and the CC shows a clear improvement across both high- and low-latitude regions.

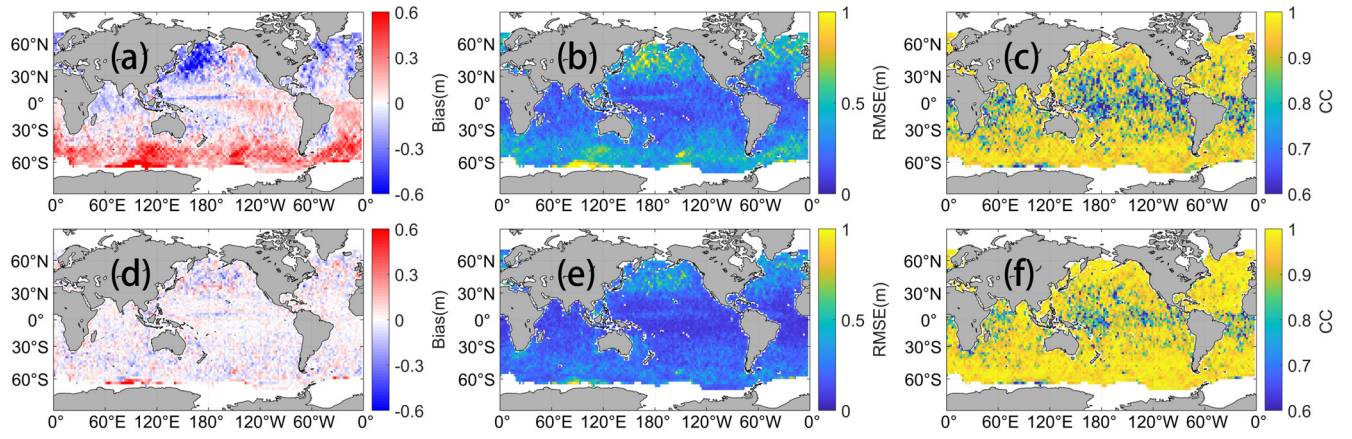


Fig. 4. Spatial distributions of SWH error metrics between WW3-ST6 and Jason-2 (a-c) before and (d-f) after fusing SWH data from CCI-Sea State in 2011: (a & d) bias, (b & e) RMSE, and (c & f) CC.

3.3 Diagnosis of mission-transition discontinuities

The purpose of the “two-sat” version is to provide a dataset with improved temporal homogeneity for wave climate studies. However, maintaining a constant number of fused altimeters does not by itself guarantee the homogeneity of the time series. The replacement of satellite missions may still introduce artificial discontinuities through differences in orbit, sampling geometry, instrument characteristics, or residual inter-mission calibration uncertainties.

Therefore, we analyzed the possible discontinuities associated with satellite mission transitions in the “two-sat” product. Fig. 5 shows the time series of monthly mean SWH from the WW3-ST6 hindcast background, the “two-sat” fused dataset, and the fused-minus-background difference during 1992–2023 for the global ocean and different latitude bands. The fused SWH generally follows the temporal variability of the WW3-ST6 background, including the pronounced seasonal cycle in the mid- and high-latitude bands, while applying a relatively stable positive or negative correction depending on region. The fused-minus-background difference also exhibits clear regional and seasonal characteristics, reflecting systematic differences between the altimeter-constrained analysis and the numerical background.

Importantly, these differences evolve smoothly across the major satellite-transition times marked by the vertical dashed lines. Although the magnitude and seasonality of fused-minus-background difference vary among latitude bands, particularly in the mid- and high-latitude oceans where wave variability is stronger, no abrupt shift appears to be synchronized with the replacement of satellite missions. This indicates that the two-satellite configuration does not introduce visually detectable step-like discontinuities in the monthly mean SWH record. Therefore, no obvious discontinuity synchronized with the mission transitions is observed in either the fused SWH or the difference time series, supporting the temporal homogeneity of the two-satellite product across mission transitions.

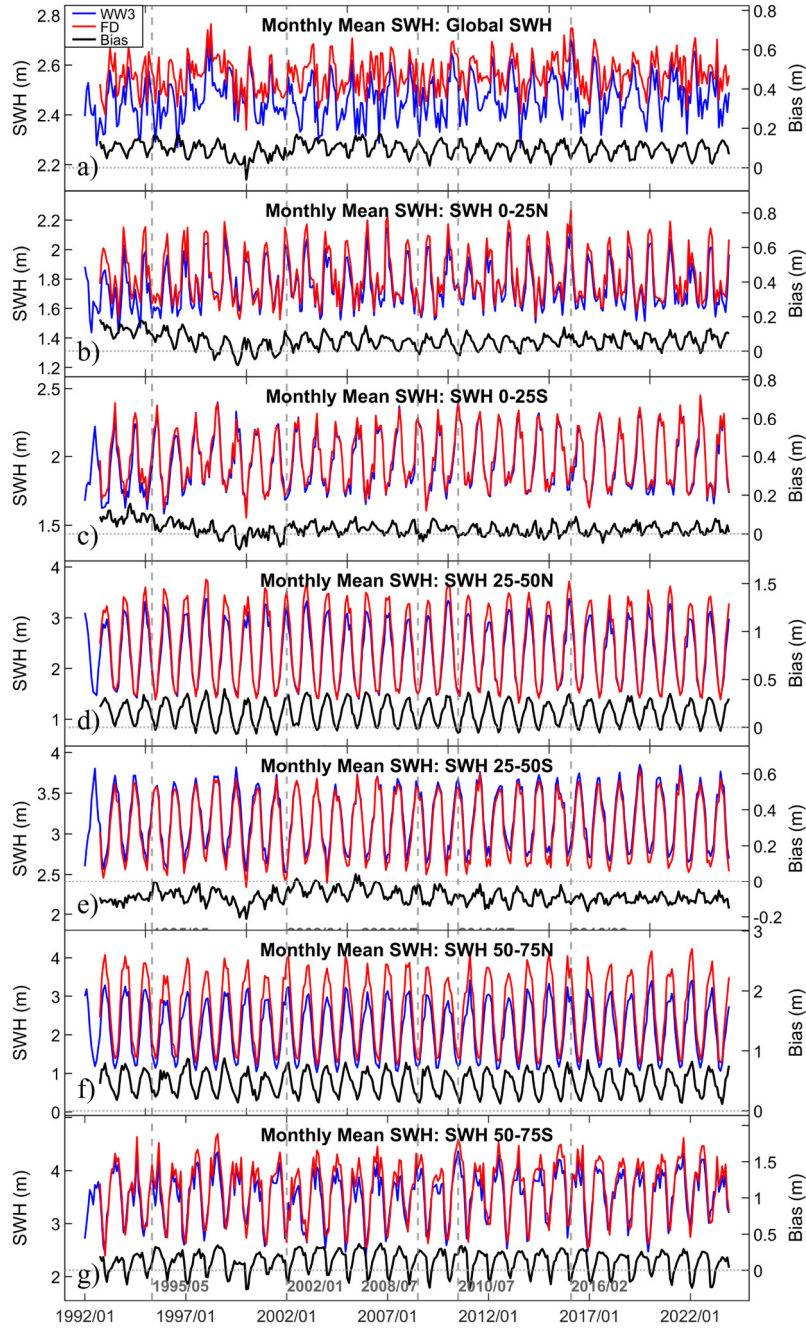


Fig. 5. Time series of monthly mean SWH from the WW3-ST6 hindcast background (blue), the “two-sat” fused dataset (red), and the fused-minus-background difference (black) during 1992–2023: (a) the global ocean, (b) 0–25°N, (c) 0–25°S, (d) 25–50°N, (e) 25–50°S, (f) 50–75°N, and (g) 50–75°S. Vertical dashed lines indicate the major satellite-transition times in the two-satellite configuration, as shown in Table 1.

4. Discussion on the Importance of Background Fields

Section 3 has demonstrated the high accuracy of the fused dataset, which results from the complementary strengths of two components: (1) the relatively sparse but highly reliable altimeter observations, and (2) the less accurate yet spatiotemporally “seamless” numerical wave hindcasts serving as the background field. Both components are indispensable. As shown previously, when no background field is used (which is equivalent to assuming a zero-valued background), the sparse altimeter observations alone cannot reconstruct a high-quality SWH field. This highlights the crucial role of the background field in ensuring the reliability and robustness of the fusion framework, but also suggests that the accuracy of the fusion product depends on the quality of the background field.

Because the present method does not prescribe a formal analytical error-covariance model, the uncertainty and robustness of the fused product are assessed empirically through independent validation and sensitivity experiments. To assess the impact of background-field selection, we conducted fusion experiments for January 2011 using two different hindcasts: WW3-ST6 forced by ERA5 10-m winds and WW3-ST4 forced by CFSR 10-m winds. Figure 6 presents the results before and after data fusion. The first row (Figs. 6a–c) shows the direct comparison between the original WW3-ST6 and WW3-ST4 hindcasts in terms of bias, RMSE, and CC. Despite employing different parameterizations and wind forcing, both hindcasts were tuned against realistic SWH conditions and thus exhibit reasonable consistency. The bias generally lies within ± 0.3 m, RMSE remains below 0.6 m in most regions (though reaching ~ 0.8 m in parts of the westerlies and near-polar regions), and CC exceeds 0.9 in most regions with significant wave variability. Lower CC values (< 0.8) appear around tropical swell pools (Chen et al., 2002; Jiang and Yang, 2022), mainly due to differences in source wind forcing and parameterizations that cause mismatches in swell arrival times (Jiang et al., 2016), combined with the inherently small SWH variability in these regions.

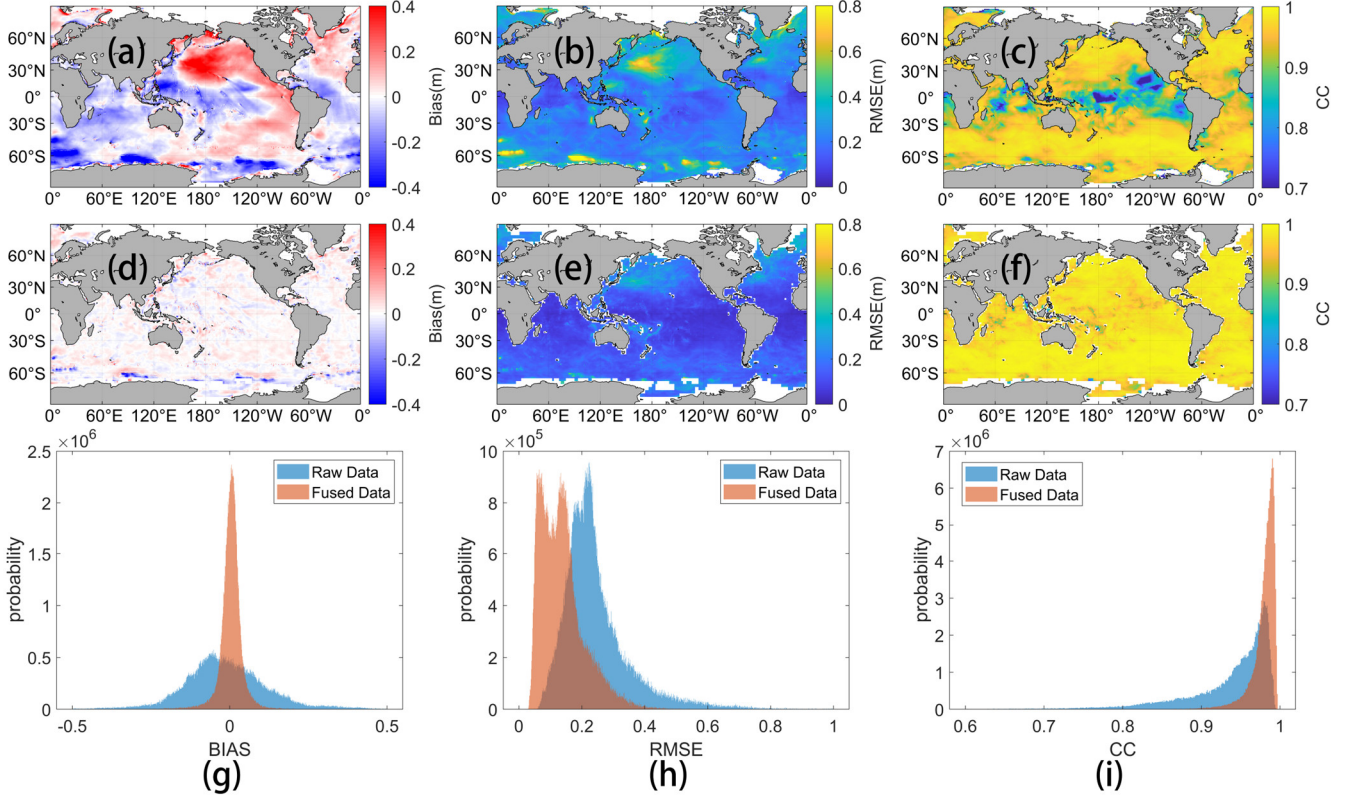


Fig. 6. Comparison between WW3-ST6 (forced by ERA5 winds) and WW3-ST4 (forced by CFSR winds) before and after data fusion is applied to both hindcasts using the same altimeter observations in January 2011. (a–c) Maps of bias, RMSE, and CC between the two hindcasts before data fusion. (d–f) Same as (a–c), but after fusion with altimeter observations. (g–i) Probability density functions of grid-point bias, RMSE, and CC, respectively, illustrating the overall improvement in consistency between the two hindcasts after fusion.

After fusion, the discrepancies between WW3-ST6 and WW3-ST4 are greatly reduced (Figs. 6d–f). With the exception of a few near-polar areas, the bias decreases to within ± 0.05 m almost everywhere. RMSE also shows a widespread reduction, with maximum values in the Northern Hemisphere westerlies not exceeding 0.4 m. CC is also markedly improved, exceeding 0.9 even within the “swell pools”, substantially enhancing the consistency between the two datasets. The probability density functions of the three error metrics across all global grid points (Figs. 6g–i) further illustrate the improved agreement after fusion.

It is noted that the errors induced by different background fields are much smaller than the errors of monthly mean values introduced by undersampling when no background field is used. Therefore, although the fused product is, as expected, not independent of the background field, the influence of the specific choice of background is not critical, provided that the background field offers a physically reasonable estimate of the large-scale SWH distribution. This also highlights the robustness of the proposed fusion approach across different model background fields. Moreover, the robustness with respect to the background further suggests that the altimeter observations are effectively assimilated into the fused dataset, making it well suited for applications such as the reconstruction and analysis of historical wave climate.

5. Applications of the Fused Dataset

5.1 Wave Climate Studies

One of the primary applications of the fused dataset is in SWH-based wave-climate studies. The fused product substantially reduces the undersampling errors that have long limited the use of altimeter data in SWH-based wave-climate studies. Jiang (2020) showed that, when only two altimeter missions are available, undersampling can introduce an RMSE of approximately 0.17 m in the monthly mean SWH on a coarse 2° grid (assuming the altimeter SWH is error-free). While such errors may not strongly impact climatological means, they can still introduce additional uncertainty in trend estimates. More importantly, even on such a coarse 2° grid, the undersampling effect can still cause systematic underestimations exceeding 0.1 m in the monthly 90th percentile, with RMSEs larger than 0.4 m, and even greater biases at higher percentiles: the monthly 99th percentile may suffer from systematic underestimations over 0.7 m and RMSEs over 0.8 m. These biases decrease as the number of altimeter missions increases, but if not explicitly corrected (which may be a difficult task), they can lead to artificial trend errors far larger than the true climate signal. This suggests that previous studies on extreme SWH climate based solely on along-track altimeter data (e.g., Izaguirre 2011; Patra et al. 2020; Takbash et al. 2019; Young and Ribal 2020) may be affected by such limitations, although the random error of altimeters may partially compensate for undersampling errors in extreme value analysis, which represents a topic worthy of further investigation in future work.

The fused dataset effectively mitigates these undersampling issues. In particular, it provides spatially complete SWH fields on a regular $0.5^\circ \times 0.5^\circ$ grid, thereby reducing the need to account for the strong grid-size dependence that arises when sparse along-track observations are directly gridded. Compared with the gridding of along-track data, the fused dataset also provides a more convenient and reliable basis for SWH-based wave-climate studies.

Figures 7a–d show the global mean SWH distribution for January 2011 derived from the fused dataset and the original CCI-Sea State L4 data, together with their absolute and relative differences (with the fused dataset re-averaged to 1° resolution for the difference calculation). In Fig. 7a, the distribution exhibits typical features of global wave fields in boreal winter, with high SWH in the Northern Hemisphere westerlies driven by extratropical storms. In comparison, the waves generated by the Southern Hemisphere westerlies are relatively smaller. During this month, the lowest mean SWH occurs in the region near the Arabian Sea in the Indian Ocean. The monsoon is inactive in this region and swells generated in the Southern Ocean cannot propagate into this area due to the blocking effects of the African continent and Madagascar. The original CCI-Sea State L4 data exhibit qualitatively consistent spatial patterns with the fused product (Fig. 7b). However, because of undersampling, the altimeter ground tracks remain clearly visible in the monthly mean fields of the CCI-Sea State L4 product. Figs. 7c and 7d show the absolute and relative differences between Figs. 7a and 7b, respectively. In regions with large mean SWH, undersampling associated with satellite orbits can lead to biases approaching ± 1 m. In regions with smaller mean SWH, although the absolute sampling error is typically within ± 0.3 m, this seemingly small absolute error may still correspond to a relative error of up to 20%. Such undersampling errors can exert a significant influence on certain types of SWH-based wave-climate studies.

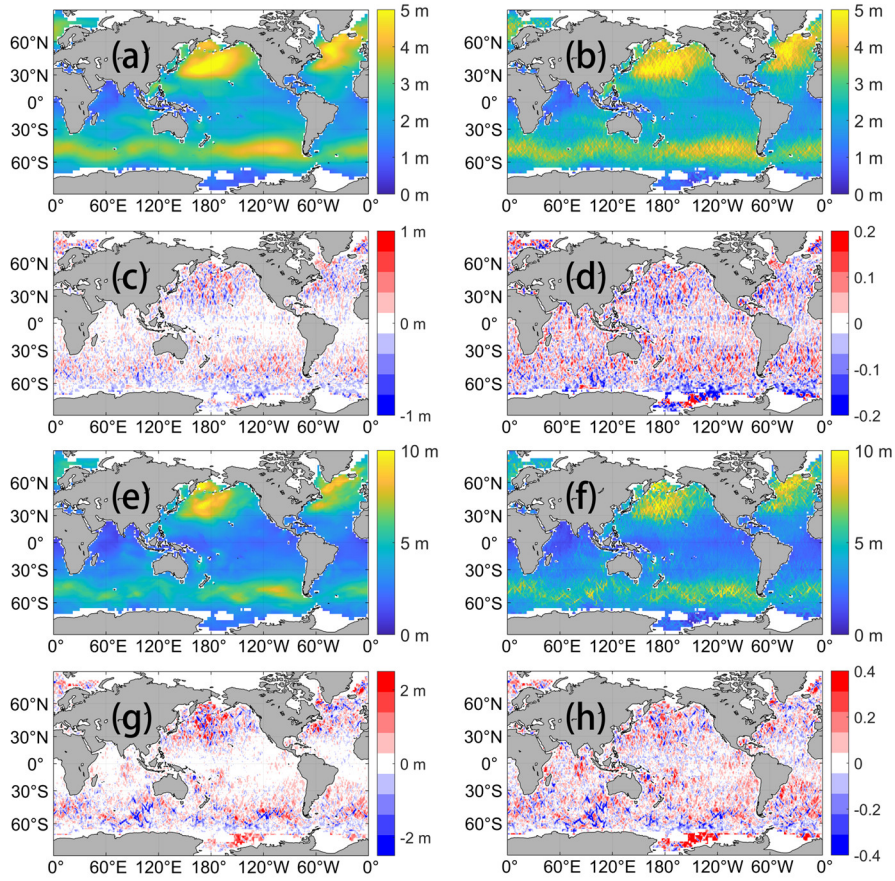


Fig. 7. Global distributions of the monthly mean SWH in January 2011 derived from (a) the fused dataset and (b) the original CCI-Sea State L4 dataset, as well as their (c) absolute difference and (d) relative difference. Panels (e–f) show the corresponding global distributions of the 95th-percentile SWH during the same period derived from (e) the fused dataset and (f) the original CCI-Sea State dataset, as well as their (g) absolute difference and (h) relative difference.

While the differences between the monthly mean SWH distributions in Figs. 7a and 7b are not particularly pronounced, Figs. 7e–h respectively present the global distribution of the 95th-percentile SWH in January 2011 derived from the fused dataset, the original CCI-Sea State data at 1° resolution, and their difference (with the fused dataset also re-averaged to 1° for comparison). The differences between Figs. 7e and 7f are even more pronounced. When a storm is successfully captured by the altimeter, the 95th-percentile SWH derived from Fig. 7f is significantly higher than that in Fig. 7e. Conversely, if the altimeter fails to sample the waves generated by a storm (particularly the high SWH values during the storm), the 95th-percentile SWH in Fig. 7f becomes notably lower than that in Fig. 7e. This issue makes the orbital patterns and data discontinuities in Fig. 7f more apparent. Such undersampling can easily lead to errors exceeding 2 m (Fig. 7g) in the monthly 95th-percentile SWH and relative errors of over 40% (Fig. 7h).

It is worth noting that in January 2011, four altimeter satellites were operating simultaneously, which considerably reduced undersampling errors compared with periods when only two satellites were in orbit. Nevertheless, even under these favourable conditions, such large errors still occur in the estimation of extreme wave heights, underscoring the necessity of data fusion.

A closer inspection further reveals that underestimation occurs more frequently than overestimation in the estimation of extreme SWH in Fig. 7f. This is because extreme wave events are generally less likely to be captured by the altimeter sampling, which has been pointed out by previous studies (Jiang, 2020). Jiang (2020) has further demonstrated that as the number of satellites increases, the likelihood of such overestimation gradually diminishes, which may, in turn, affect the analysis of long-term climate trends.

495 Previous studies have also pointed out that the use of a 1° grid tends to underestimate altimeter-based SWH percentiles, and that coarser grids of 2° or 3° may be more appropriate (Young and Ribal, 2019; Takbash et al., 2019; Jiang, 2020). However, coarser grids inevitably lead to the loss of spatial detail. In addition, as the computation area increases, the physical meaning of percentile estimates derived from altimeter observations becomes less clear: They will shift from the temporal percentiles of spatially averaged SWH within a given grid cell to the spatio-temporal percentiles of the broader region corresponding to the grid size. In contrast, percentiles computed from the fused dataset not only retain higher spatial resolution but also have a well-defined physical interpretation. For instance, when the 95th or 99th percentile is calculated at the native 0.5° resolution of the fused dataset, the resulting value corresponds directly to the extreme percentile of the time series of spatially averaged SWH within each 0.5° grid cell.

5.2 Studies of AI Wave Models

505 In recent years, with the rapid development of artificial intelligence (AI), data-driven modeling approaches have re-emerged as a promising alternative to traditional numerical simulations. Compared with numerical models, AI-based methods offer lower computational cost and higher efficiency, and they have already achieved remarkable success in weather forecasting. Large-scale AI models such as *Pangu*, *GraphCast*, *GenCast*, and *Aurora* have, in some cases, delivered forecasting skill surpassing that of numerical weather prediction (NWP) systems at a fraction of the computational expense. Similarly, we believe that AI-based wave models represent a promising future research direction.

A key requirement for training high-performing AI models is the availability of large, accurate, and well-structured datasets. Although ERA5 SWH data already provide a reasonable source of training samples, as noted in Section 1, ERA5 does not incorporate future observations to constrain past states, leaving room for improvement. In one of our recent studies, we used a five-year subset of the fused dataset generated using the present framework to train and fine-tune an AI wave model for global SWH (Wang and Jiang 2024). Hindcast experiments demonstrated that the AI model fine-tuned with the fused dataset not only outperformed the counterpart trained directly on ERA5 but also significantly exceeded the hindcast skill of WW3-ST6. This indicates that the fused dataset enables AI wave models to achieve, in terms of SWH hindcast accuracy, performance that rivals or even surpasses state-of-the-art NWMs (Fig. 8). Figure 8 is reproduced from an independent-test-year results of Wang and Jiang (2024), in which 2020 was withheld from model training and used exclusively for evaluation. It is noted that Wang and Jiang (2024) focused on AI model development and used the fused SWH fields only as training data. It did not provide the full long-term dataset, the two product configurations, the systematic validation, or the public data documentation presented here.

525 More details of this AI wave model and how this fused dataset is used can be found in Wang and Jiang (2024). Meanwhile, it is noted that we additionally provide the R value in Eq. 3, which represents the spatiotemporal distance to the nearest altimeter observation, to further adapt the fused dataset for AI training applications. Intuitively, a smaller R value indicates stronger influence from altimeter measurements, which allows it to be incorporated into the training loss function as a weight, so that data points closer to direct observations can be assigned higher importance. This R value and corresponding weighting strategy were also employed in Wang and Jiang (2024).

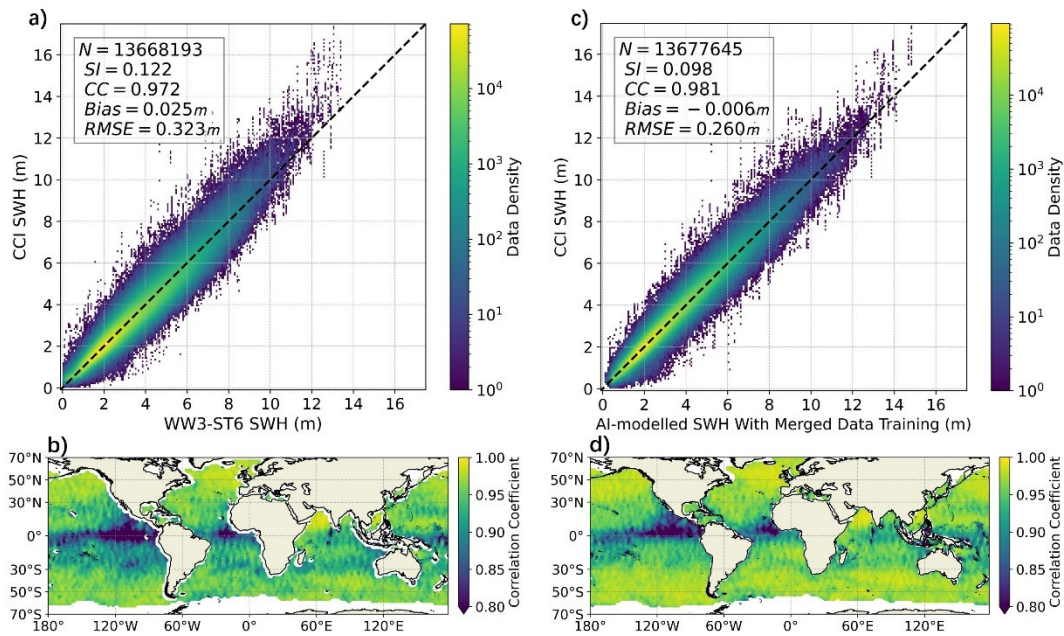


Fig. 8. Comparison between SWHs from numerical/AI wave models and CCI-sea state dataset for global ocean in 2020. (a) Scatter plot of SWHs from the NWM versus CCI-sea state. (b) Spatial distributions of CC between SWHs from NWM and CCI-sea state. (c, d) Same as (a, b), respectively, but for the AI wave model trained with ERA5 and subsequently fine-tuned using the 5-year fused dataset. Note that the year 2020 is excluded from the model training.

6. Data Description and Availability

The fused dataset is freely available at <https://doi.org/10.57760/sciencedb.29314> (Su and Jiang, 2025). Two versions are provided: (1) a *two-sat* version, which incorporates data from only two altimeters at any given time and is primarily intended for SWH-based wave-climate studies; and (2) a *multi-sat* version, which integrates observations from as many concurrent altimeter missions as possible and is designed to support applications such as the training of artificial intelligence-based wave models. The data are stored in standard NetCDF grid format, with SWH and the distance parameter R expressed as functions of latitude, longitude, and time. The main characteristics of the released fused SWH products are summarized in Table 5. Meanwhile, each NetCDF file corresponds to one calendar month, with global attributes documenting metadata such as the source WW3-ST6 hindcast file, the names of the fused satellites, the start and end time, the identifiers of the fused satellites, and the percentage contribution of each satellite dataset to the fusion.

Table 5. Main characteristics of the released fused SWH products.

Information	Description
Product version	Multi-sat / Two-sat (with the same format)
Time period	October 1992 – December 2023 *
Data coverage	Global
Spatial resolution	$0.5^\circ \times 0.5^\circ$
Temporal resolution	3 hours
Missing-value convention	NaN for land, masked coastal regions, and sea-ice-covered grid cells.
Single file volume	~ 480 MB (for one month per version)
Total dataset volume	~ 360 GB + 360 GB
Version of WW3 hindcast	v7.00, ST6 source-term package, ERA5 10-m wind forcing, $0.25^\circ \times 0.25^\circ$, 3 h.
Altimeter data fused	CCI Sea State v4, Level-3 1-Hz along-track SWH, Daily Multi-Sensor product.
Available variables	Fused SWH value (m), Distance-related parameter indicating the proximity to the nearest altimeter constraint (R value), Latitude, Longitude, Time

* The CCI Sea State v4 record starts in August 1991, but the released fused products start in October 1992 because the two-sat configuration requires two simultaneously available altimeter missions.

7. Summary

Satellite altimeters can provide long-term, high-accuracy observations of SWH across the global ocean. However, their inherently sparse spatial and temporal coverage introduces undersampling errors, which can limit the reliability of wave climate analyses, particularly for extreme wave climate. To address this limitation, this study developed a global gridded SWH dataset by applying an offline fusion framework to combine WAVEWATCH III hindcast fields with along-track, jointly calibrated altimeter observations from the CCI Sea State dataset. Unlike online wave assimilation systems designed primarily for dynamically consistent reanalysis, the offline fusion framework used here aims to maximize the utilization of historical observations for retrospective SWH reconstruction. To cater to different applications, two versions of the fused dataset are provided: (1) a *two-sat* version, which incorporates data from only two satellites at any given time and is optimized for SWH-based wave-climate studies, ensuring broadly consistent orbital sampling characteristics throughout the record by reducing sampling-related temporal inhomogeneity; and (2) a *multi-sat* version, which assimilates observations from as many altimeter missions as possible and is intended for applications such as training artificial intelligence-based wave models, where high spatial and temporal fidelity is critical.

Validation against buoy measurements and withheld satellite data shows that the resulting fused gridded product achieves high accuracy, demonstrating that combining the spatial and temporal continuity of numerical hindcasts with the accuracy of altimeter observations provides a more reliable gridded SWH field than either source alone. Experiments using different background hindcasts show that the fused results are not strongly dependent on the specific numerical background field, provided that the background offers a physically reasonable estimate of the large-scale SWH distribution. Although the fused product inevitably retains some dependence on the background model, the discrepancies between products generated from different hindcasts are effectively reduced after fusion.

Although this high-accuracy gridded dataset offers many practical advantages, it also has some limitations compared with along-track altimeter observations. These limitations arise from the design of the fusion framework. Because the fused fields are generated on the spatial grid and temporal resolution of the numerical hindcast and involve space-time interpolation, small-scale signals present in the original along-track observations are smoothed. Therefore, the product is not intended for studies that require the full fine-scale variability of altimeter measurements, such as analyses of regional-scale SWH gradients or other small-scale wave-field features. In addition, because the WW3-ST6 background hindcast used here does not explicitly account for wave–current interactions, the fused product should not be used as a primary dataset for diagnosing such interactions, even though some wave-current-interaction signals are present in the original altimeter data. Moreover, this dataset is intended for analyses based only on total SWH, thus, other wave parameters from WW3 are not jointly adjusted within the present fusion framework, caution is warranted in applications that require strict internal spectral consistency among SWH and other wave parameters. For such applications, the consistency assumptions and associated uncertainties should be evaluated according to the objective.

Despite these limitations, this fused dataset already offers substantial utility across a wide range of applications. Future improvements can be made by incorporating additional satellite missions and extending the temporal coverage of the dataset. In particular, including more recent altimeter and wave-scatterometer observations, such as those from Haiyang-2 and CFOSAT, may further improve the multi-sat product and support more robust analyses of long-term wave-climate variability and trends.

Author contributions

Conceptualization: HJ

Methodology: HS

Investigation: HS, HJ

590 Visualization: HS
Supervision: HJ
Writing—original draft: HJ, HS
Writing—review & editing: HJ

595 **Competing interests**

All authors declare that they have no competing interests.

Acknowledgments

This work was jointly supported by the National Natural Science Foundation of China (42376172), the Guangdong Basic and
600 Applied Basic Research Foundation (2026B1515020068), the Shenzhen Science and Technology Program (JCYJ20250604182014019).

Code and data availability

Input datasets used for fusion: The CCI sea state product is available at <https://data-cersat.ifremer.fr/data/ocean-waves/cci-seastate/v4/>. WAVEWATCH III ST4 and ST6 hindcasts are available from Raschle and Ardhuin (2013) and Liu et al. (2021),
605 respectively. Reference datasets used for validation and comparison: The gridded multi-mission merged satellite SWH data product from CMEMS is available from <https://doi.org/10.48670/moi-00180> (CMEMS 2024). The NDBC buoy data are available from <https://www.ndbc.noaa.gov/>. The AI wave model is available from Wang and Jiang (2024). The final fused products released: The fused dataset is available at <https://doi.org/10.57760/sciencedb.29314> (Su and Jiang 2025).

610

Reference

- Alday, M., Accensi, M., Ardhuin, F., and Dodet, G.: A global wave parameter database for geophysical applications. Part 3: Improved forcing and spectral resolution, *Ocean Modell.*, 166, 101848, doi:10.1016/j.ocemod.2021.101848, 2021.
- Ardhuin, F., Rogers, E., Babanin, A. V., Filipot, J.-F., Magne, R., Roland, A., van der Westhuysen, A., Queffelecoulou, P., Lefevre,
615 J.-M., Aouf, L., and Collard, F.: Semiempirical dissipation source functions for wind-wave models: part I, definition, calibration and validation, *J. Phys. Oceanogr.*, 40, 1917–1941, doi:10.1175/2010JPO4324.1, 2010.
- Bretherton, F. P., Davis, R. E., and Fandry, C. B.: A technique for objective analysis and design of oceanographic experiments applied to MODE-73, *Deep-Sea Research*, 23, 559–582, 1976.
- CMEMS (Copernicus Marine Service Information), Global Ocean L 4 Significant Wave Height From Nrt Satellite
620 Measurements, Marine Data Store [dataset], doi: 10.48670/moi-00180, 2024.
- Chen, G., Chapron, B., Ezraty, R., and Vandemark, D.: A global view of swell and wind sea climate in the ocean by satellite altimeter and scatterometer, *J. Atmos. Oceanic Technol.*, 19, 1849–1859, doi:10.1175/1520-0426(2002)019<1849:AGVOSA>2.0.CO;2, 2002.
- Dodet, G., Piolle, J.-F., Quilfen, Y., Abdalla, S., Accensi, M., Ardhuin, F., Ash, E., Bidlot, J.-R., Gommenginger, C., Marechal,
625 G., Passaro, M., Quartly, G., Stopa, J., Timmermans, B., Young, I., Cipollini, P., and Donlon, C.: The Sea State CCI dataset v1: towards a sea state climate data record based on satellite observations, *Earth Syst. Sci. Data*, 12, 1929–1951, doi:10.5194/essd-12-1929-2020, 2020.
- Fan, Y., Lin, S., Held, I. M., Yu, Z., and Tolman, H. L.: Global ocean surface wave simulation using a coupled atmosphere-wave model, *J. Clim.*, 25, 6233–6252, doi:10.1175/JCLI-D-11-00621.1, 2012.

- 630 Fan, Y., Held, I. M., Lin, S., and Wang, X.: Ocean warming effect on surface gravity wave climate change for the end of the 21st century, *J. Climate*, 26, 6046-6066, doi:10.1175/JCLI-D-12-00410.1, 2013.
- Fan, Y., Lin, S., Griffies, S. M., and Hemer, M. A.: Simulated Global Swell and Wind Sea Climate and Their responses to Anthropogenic Climate Change at the End of the 21st Century, *J. Clim.*, 27, 3516-3536, doi:10.1175/JCLI-D-13-00198.1, 2014.
- 635 Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. Roy. Soc. A-Math. Phys.*, 454, 903–995, doi:10.1098/rspa.1998.0193, 1998.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De
640 Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, *Q. J. Roy. Meteor. Soc.*, 146, 1999–2049, doi:10.1002/qj.3803, 2020.
- Izaguirre, C., Méndez, F. J., Menéndez, M., and Losada, I. J.: Global extreme wave height variability based on satellite data,
645 *Geophys. Res. Lett.*, 38, L10607, doi:10.1029/2011GL047302, 2011.
- Jiang, H. and Yang, Z.: A revisit of global wind-sea and swell climate and variability using multiplatform altimeters, *Remote Sens. Environ.*, 271, 112922, doi:10.1016/j.rse.2022.112922, 2022.
- Jiang, H.: Evaluation of altimeter undersampling in estimating global wind and wave climate using virtual observation, *Remote Sens. Environ.*, 245, 111840, doi:10.1016/j.rse.2020.111840, 2020.
- 650 Jiang, H.: Random, Environmental, and Representativeness Errors in Ocean Remote Sensing Versus In Situ Data: An Example of Wave Heights From Altimeters, *IEEE Trans. Geosci. Remote Sens.*, 61, 1–13, doi:10.1109/TGRS.2023.3285348, 2023.
- Jiang, H., Babanin, A. V., and Chen, G.: Event-based validation of swell arrival time, *J. Phys. Oceanogr.*, 46, 3563–3569, doi:10.1175/JPO-D-16-0208.1, 2016.
- Law-Chune, S., Aouf, L., Dalphiné, A. et al. WAVERYS: a CMEMS global wave reanalysis during the altimetry period,
655 *Ocean Dynamics*, 71, 357–378, doi:10.1007/s10236-020-01433-w, 2021.
- Liu, Q., Babanin, A. V., Rogers, W. E., Zieger, S., Young, I. R., Bidlot, J.-R., and Durrant, T. H.: Global wave hindcasts using the observation-based source terms: Description and validation, *J. Adv. Model. Earth Sys.*, 13, e2021MS002493, doi:10.1029/2021MS002493, 2021.
- Meucci, A., Young, I. R., and Breivik, Ø.: Wind and Wave Extremes from Atmosphere and Wave Model Ensembles, *J. Climate*,
660 31, 8819–8842, doi:10.1175/JCLI-D-18-0217.1, 2018.
- Patra, A., Min, S.-K., and Seong, M.-G.: Climate variability impacts on global extreme wave heights: Seasonal assessment using satellite data and ERA5 reanalysis, *J. Geophys. Res.-Oceans*, 125, e2020JC016754, doi:10.1029/2020JC016754, 2020.
- Quilfen, Y. and Chapron, B.: On denoising satellite altimeter measurements for high-resolution geophysical signal analysis,
665 *Adv. Space Res.*, 68, 875–891, doi:10.1016/j.asr.2020.01.005, 2021.
- Quilfen, Y. and Chapron, B.: Ocean surface wave-current signatures from satellite altimeter measurements, *Geophys. Res. Lett.*, 46, 253–261, doi:10.1029/2018GL081029, 2019.
- Rascle, N. and Ardhuin, F.: A global wave parameter database for geophysical applications. Part 2: Model validation with improved source term parameterization, *Ocean Modell.*, 70, 174-188, doi:10.1016/j.ocemod.2012.12.001, 2013.
- 670 Ribal, A. and Young, I. R.: 33 years of globally calibrated wave height and wind speed data based on altimeter observations, *Sci. Data*, 6, 77, doi:10.1038/s41597-019-0083-9, 2019.

- Semedo, A., Sušelj, K., Rutgersson, A., and Sterl, A.: A global view on the wind sea and swell climate and variability from ERA-40, *J. Clim.*, 24, 1461-1479, doi:10.1175/2010JCLI3718.1, 2011.
- 675 Semedo, A., Vettor, R., Breivik, Ø., Sterl, A., Reistad, M., Soares, C. G., and Lima, D.: The wind sea and swell waves climate in the Nordic seas, *Ocean Dynam.*, 65, 223-240, doi:10.1007/s10236-014-0788-4 , 2015.
- Semedo, A., Weisse, R., Behrens, A., Sterl, A., Bengtsson, L., and Günther, H.: Projection of global wave climate change towards the end of the Twenty-First Century, *J. Climate*, 26, 8269-8288, doi:10.1175/JCLI-D-12-00658.1, 2013.
- Sterl, A. and Caires, S.: Climatology, variability and extrema of ocean waves: The web-based KNMI/ERA-40 wave atlas, *Int. J. Climatol.*, 25, 963–977, doi:10.1002/joc.1175, 2005.
- 680 Stopa, J. E. and Cheung, K. F.: Periodicity and patterns of ocean wind and wave climate, *J. Geophys. Res.-Oceans*, 119, 5563–5584, doi:10.1002/2013JC009729, 2014.
- Su, H. and Jiang, H.: Significant wave height fusing multi-mission altimetry and WW3 hindcast, V3. Science Data Bank, , doi:10.57760/sciencedb.29314, 2025.
- Takbash, A., Young, I. R., and Breivik, Ø.: Global Wind Speed and Wave Height Extremes Derived from Long-Duration
685 Satellite Records, *J. Climate*, 32, 109–126, doi:10.1175/JCLI-D-18-0520.1, 2019.
- Timmermans, B. W., Gommenginger, C. P., Dodet, G., and Bidlot, J.-R.: Global wave height trends and variability from new multi-mission satellite altimeter products, reanalyses and wave buoys, *Geophys. Res. Lett.*, 47, e2019GL086880, doi:10.1029/2019GL086880, 2020.
- Wang, X. and Jiang, H.: Physics-guided deep learning for skillful wind-wave modeling, *Sci. Adv.*, 10, eadr3559, doi:10.1126/sciadv.adr3559, 2024.
- 690 Young, I. R. and Ribal, A.: Multiplatform evaluation of global trends in wind speed and wave height, *Science*, 364, eaav9527, doi:10.1126/science.aav9527, 2020.
- Zieger, S., Vinoth, J., and Young, I. R.: Joint calibration of multiplatform altimeter measurements of wind speed and wave height over the past 20 years, *J. Atmos. Ocean. Technol.*, 26, 2549–2564, doi:10.1175/2009JTECHA1303.1, 2009.