Spaceborne SAR wave mode data as big data for global ocean wave observation

Bingqing Huang[a,b], Xiao-Ming Li[a]

[a] Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, No. 9 Dengzhuang South Road, Beijing, China

[b] University of Chinese Academy of Sciences, No. 19(A) Yuquan Road, Beijing, China

Abstract

The Wave Mode (WM) is a unique imaging mode of spaceborne synthetic aperture radar (SAR), which is dedicatedly designed for global ocean surface wave observations. Since it was firstly achieved in the ERS-1/SAR satellite starting from 1991, the WM data over global oceans have been continuously acquired by the ESA's SAR instruments onboard the satellites of ERS-2 (1997 – 2011), ENVISAT (2002 – 2012), Sentinel-1A(2014--)/1B(2016--) and the Chinese GaoFen-3 SAR (2016--), for nearly 30 years. Therefore, spaceborne SAR WM data have been a 'big data', which however, has not been widely exploited for global ocean wave measurements. In this study, we demonstrate its application on global ocean wave measurements based on the ten-year Advanced SAR (ASAR) WM data.

1 Introduction

Ocean wave is a key parameter of marine-meteorology interface, which plays an important role in interactions of ocean, atmosphere and cryosphere. Some evidences have shown that ocean waves might also be one of the scenarios of global ocean climate change[1]. The spaceborne active microwave sensors, such as the Radar Altimeter (RA) and the Synthetic Aperture Radar (SAR) are considered to be the most effective techniques for ocean wave observation in the global scale as they are independent of sunlight and weather conditions. By measuring the distance to the sea surface in nadir, RA is available to measure significant wave height and has been providing global ocean wave measurements for decades[2]. Different from RA, the SAR records the radar backscatter of sea surface in side-look and can provide two-dimensional ocean surface images.

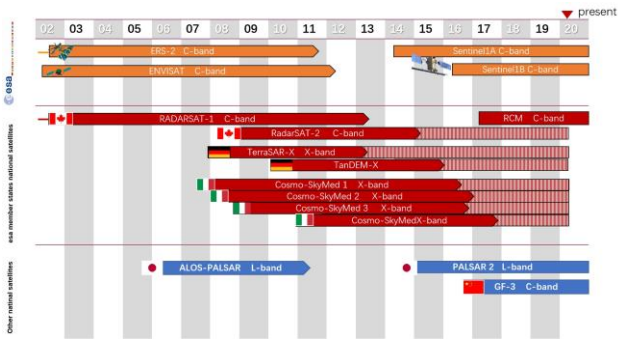


Figure 1 Spaceborne SAR missions that are widely used for ocean observation

Following the Seasat launched in 1978[3], there are dozens of spaceborne SAR satellites have been providing massive and valuable data over the global ocean, some of which are presented in **Figure 1.** In principle, the spaceborne SAR data have exhibited the so-called "5V" characteristics of big data, i.e. volume, velocity, variety, veracity and value. However, compared with optical remote sensing data (e.g., LandSat and MODIS data) that have been widely used as big data for earth observation, most of studies and applications on ocean based on spaceborne SAR data are still on a stage of "case studies". The reason is mainly due to the complicated imaging mechanisms of spaceborne SAR on various oceanic and atmospheric dynamics, which consequently leads to difficulties of image interpretations and quantitative retrievals. On the other hand, irregular acquisitions (i.e. incontinuous spatial and temporal coverages) of spaceborne SAR data also limit its applications on studying the dynamic ocean.

Among these SAR satellites listed in Figure 1, the ESA's missions of ERS-1/2 SAR, Envisat/ASAR and Sentinel-1A/B, as well as the Chinese mission Gaofen-3 can provide data by a unique imaging mode called 'Wave mode' (WM). Although these data are in small spatial coverages of approximately 5 km by 10 km, which are continuously acquired over the global ocean and are particularly powerful for global ocean surface wave observations.

The spaceborne SAR WM data have been available since 1991 and will be continuously acquired over at least for

another decade. Therefore, the spaceborne WM data have characteristics of EO (earth observation) big data, at least with respect to its volume. In principle, one can derive two-dimensional (2D) wave spectra from these WM data, and then widely used sea states parameters, e.g., significant wave height (SWH) and mean wave period (MWP) can be calculated by integrating the 2D spectra. However, to retrieve ocean wave information from SAR image is rather complicated, as the imaging of ocean waves by SAR is generally a nonlinear process[4-5]. The moving of the sea surface during the imaging time of SAR results in loss of high-frequency information of ocean waves, as well as image distortion. Thus, the prior information, such as the first guess spectrum [6] obtained from the numerical wave model (e.g., the WAM model[7]) must be introduced to retrieve the full directional ocean wave spectrum[8-9]. By degrading the nonlinear imaging process, a quasi-linear retrieval method that needs no prior information was proposed and was used for generating the ASAR WM Level2 product [10], which yields retrievals of swell spectrum, more accurately called retrievals of ocean wave components imaged by SAR. Alternatively, parametric methods can derive directly integral wave parameters, SWH and MWP, from SAR images without needing prior information during retrieval. Different from the quasi-linear retrieval method, this method derives the full sea state parameters by establishing relations between radar backscatter and other parameters of SAR images and ocean wave parameters [11-13].

Maybe because of the complicated SAR imaging mechanism of sea surface, the SAR WM data have not been exploited widely for global ocean wave measurements though the 'big data' has formed. In this paper, we present a novel study on investigating global ocean waves based on the ENVISAT/Advanced SAR (ASAR) wave mode (WM) data to demonstrate the possibility of using spaceborne SAR as big data for global ocean observations.

2 Datasets and Methods

2.1 Datasets

2.1.1 ASAR WM data

The ASAR WM data used in this paper were acquired between December 2002 and April 2012, which are all the available data during its lifetime from March 2002 to May 2012. The ASAR WM data were acquired every 100 km along the track with image coverage ranging from 6 km × 5 km to 10 km × 5 km. The ASAR WM data, provided in single complex format (i.e. real part R_e and imaginary part I_m), record both magnitude and phase information of the returned radar signals, therefore the SAR image intensity (I) could be calculated as $I = R_e^2 + I_m^2$. By applying the provided calibration factor, the normalized radar cross section σ_0 is obtained and then is used for sea state parameter retrieval. Except for the ASAR WM data, the GaoFen-3 WM data were also used to demonstrate the process of the nonlinear retrieval method, with the swath width of 5 km ×5 km, and the spatial resolution of 10m.

2.1.2 In situ buoy data

Data of 649 ECMWF buoys and 252 NDBC buoys were collected from the GlobWave data portal for validation and calibration of the ASAR derived SWH and MWP data. The time span of the buoy data consists with the ASAR WM data's, i.e. from December 2002 to April 2012. However, the EMCWF buoy data of the "mean wave period" is the averaged of all wave period records within a duration, generally 20 minutes, whereas the ASAR derived MWP is the up-crossing wave period (denote $ASAR\_T_{m02}$). Therefore, we used the NDBC buoy spectrum data to calculate the up-crossing wave period (denote $NDBC\_T_{m02}$).

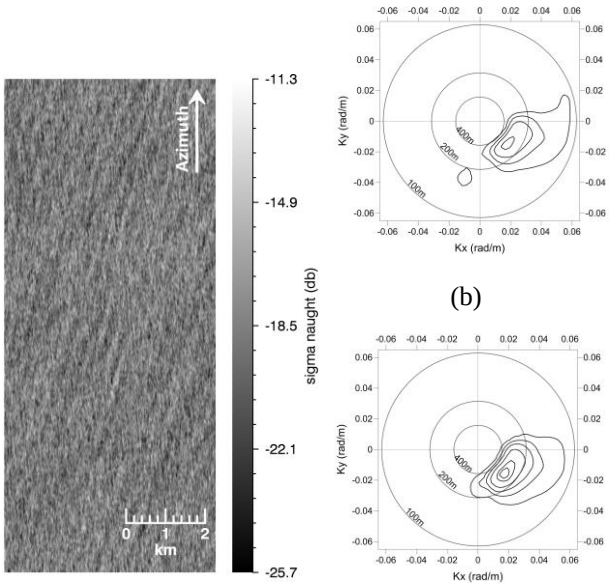
2.1.3 Radar Altimeter (RA) Data

The RA SWH data used in this study is the JASON-1 SWH data, which were also accessed from the GlobWave data portal. The JASON-1 SWH data are used for cross-validation with the ASAR derived SWH (denoted $ASAR\_H_s$ hereafter). The time span of collected JASON-1 data is from December 2002 to December 2011, which is slightly shorter than the period of ASAR WM data, as we found some JASON-1 SWH data are spurious in year 2012.

2.2 Methods

2.2.1 The Max-Planck Institute (MPI) approach of SAR-ocean wave spectrum retrieval

The MPI approach is a nonlinear retrieval of deriving two-dimensional ocean wave spectrum from spaceborne SAR data. By introducing the first guess spectrum, which is generally obtained by running a wave model such as the WAM model, the lost high-frequency information is compensated, and the 180° wave propagation ambiguity can be also solved.



(a) (c)

Figure 2 (a) The GF-3 SAR WM data (VV polarization) near the Hawaii, (b) the synchronous directional ocean wave spectrum ran by the WAM model at the nearest grid of the SAR image, (c) the retrieved ocean wave spectrum using the MPI approach.

The **Figure 2** shows a case of ocean wave spectrum retrieval from the GF-3 WM data using the MPI approach. The **Figure 2** (a) shows the GF-3 WM image acquired near the Hawaii, from which we can see the clear ocean swell pattern. The **Figure 2** (b) shows the synchronous directional ocean wave spectrum ran by the WAM model in the closest grid to the GF-3 SAR image, which was used as the first guess spectrum during the retrieval. The **Figure 2** (c) shows the retrieved ocean wave spectra based on the MPI approach. By comparing the first guess spectrum and the retrieved ocean wave spectrum, we can see that the retrieval did not make too much change in spectrum shapes. However, the retrieval did change the swell peak wave energy, as presented in the comparisons of one-dimensional wave spectra in **Figure 3**. In the figure, the red, blue and green lines represent the one-dimensional spectra of the SAR retrieval, the first guess WAM model and the buoy, respectively. The corresponding SWH is calculated by integrating the spectra. Both the one-dimensional spectrum and the integrated parameters indicate that the ASAR retrieved results are closer to the buoy measurements compared with the WAM model predicted.

Though the MPI method could derive the full ocean wave spectra from SAR images, the method relies on the prior information, which makes it not appropriate for long-term WM data processing. Thus, we decided to use the parametric model to reprocess the ASAR WM data, which could derive the full sea sate parameters without using the prior information.

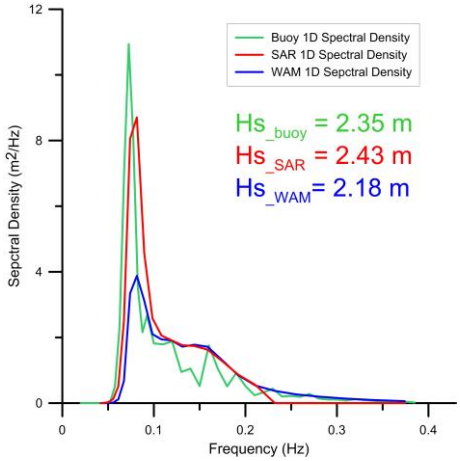


Figure 3 Comparison of the one-dimensional spectra

2.2.2 The CWAVE_ENV method of deriving integral ocean wave parameters from SAR WM data

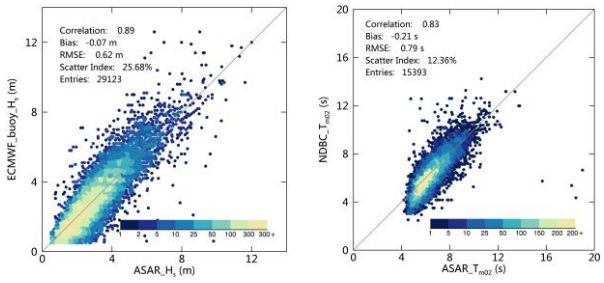
The CWAVE_ENV parametric model was used to derive integral wave parameters from the ASAR WM data. Detailed description of the algorithm development is not given here, while one can refer it to [12].

Before applying the parametric model to the ASAR WM data, one should notice that some phenomena that are not related with ocean surface waves, such as oil spill, atmospheric feature and bright targets presenting in ASAR imagettes can influence the retrieved results. To ensure 'purity' of the data, we used the 'homogeneity parameter' [14] for quality control of the ASAR WM data. ASAR imagettes with the homogeneity parameter greater than 1.05 are excluded from processing. Furthermore, some ASAR WM data were acquired experimentally with HH (Horizontal-Horizontal) polarization or with incident angle of 33° are also excluded from processing. After the pre-processing, there are approximately 6.48 million ASAR imagettes are available for producing global ocean wave parameters.

3 Validation and calibration

3.1 Validation and Calibration of ASAR retrieval by in situ buoy measurements

The collocation between the buoy data and the ASAR WM data following the criteria that the differences in spatial and temporal are less than 100 km and 0.5 h, respectively. The range of SWH is limited from 0.5 m to 30 m and the MWP's from 2 s to 20s for comparisons. Finally, there are 29,123 pairs of SWH data and 15,393 pairs of MWP data were collocated with in situ buoys. Comparisons of the ASAR derived SWH and MWP with the in-situ buoy measurements are shown in the diagrams presented in **Figure 4** (a) and (b), respectively, where the color presents the density of data pairs.



(a) (b)

Figure 4 (a) comparison between the ASAR-derived SWH and the ECMWF buoy data. (b) Comparison between the ASAR-derived MWP and the NDBC buoy data

From the **Figure 4** (a) we can see that the ASAR-derived SWH has a good agreement with the buoy measurements with the correlation coefficient of 0.89. The bias and RMSE are of -0.07 m and 0.62 m, respectively, which are

close to the results of previous validation based on two-month dataset [12]. The S.I. of 25.68% is rather higher, which maybe because that the ECMWF buoy measurements are provided by hourly while the ASAR data were acquired instantaneously. As shown in the **Figure 4** (b), the ASAR-derived MWP also has a good consistency with the buoy measurements with the correlation coefficient of 0.83. The bias and the RMSE of -0.21 s and 0.79 s, respectively, and the S.I. is only 12.36%.

Though the ASAR-derived results generally have good agreements with the buoy measurements, the ASAR-derived SWH is overestimated in the range where the SWH is lower than 2.5 m and turns to be underestimated with the increasing of SWH. The similar trend could also be found in **Figure 4** (b), where the ASAR-derived MWP is overestimated when the MWP is lower than 7 s. Comparisons of the ASAR derived SWH and MWP with buoy measurements suggest that the retrieval results have rooms of improvement. Therefore, we further used the buoy measurements to calibrate the ASAR retrieved sea state parameters.

The Reduce Major Axis (RMA) regression [15-16] method was used in the calibration process, as the buoy data are not totally free of errors [17]. The RMA regression method could take account into errors both from buoy measurements and ASAR retrievals. To avoid influences induced by outliers when building up the calibration, we applied the Tukey fences[18] and the Robust Regression method [19] to remove outliers in collected data pairs. The definition of Tukey fences is shown below.

$$\begin{aligned} IQR &= Q3 - Q1 \\ lower\ fence &= Q1 - 1.5IQR \\ Upper\ fence &= Q3 + 1.5IQR \end{aligned} \tag{3}$$

Where, the $Q1$, $Q2$ and $Q3$ represent the first, second and third quartiles, respectively. IQR is the interquartile range, which could be calculated by $Q1$ and $Q3$. The data smaller than lower fence or larger than higher fence are detected as outliers. The Robust Regression is insensitive to outliers, as it can assign data with different weight during iterations. The bigger the residual is, the smaller the weight is assigned in the next iteration of regression. In this paper, we used 0.15 as the threshold of weight and the collocation data pairs with weights lower than this threshold are marked as outliers. By applying the Tukey fences and the Robust regression method, 1,034 and 423 outliers were detected in SWH and MWP data pairs, respectively. 28,089 pairs of SWH data and 14,970 pairs of MWP data remained eventually for calibration. The **Figure 5** (a) and (b) show the scatter diagrams of ASAR-derived SWH and MWP against the buoy measurements after removing outlier, where the cross signs represent the detected outliers. By applying the RMA regression method, the calibrate formula is consequently derived as follows.

$$Calibrated\_ASAR\_H_s = 1.140 * ASAR\_H_s - 0.402 \tag{4}$$

$$Calibrated\_ASAR\_T_{m02} = 1.268 * ASAR\_T_{m02} - 1.887 \tag{5}$$

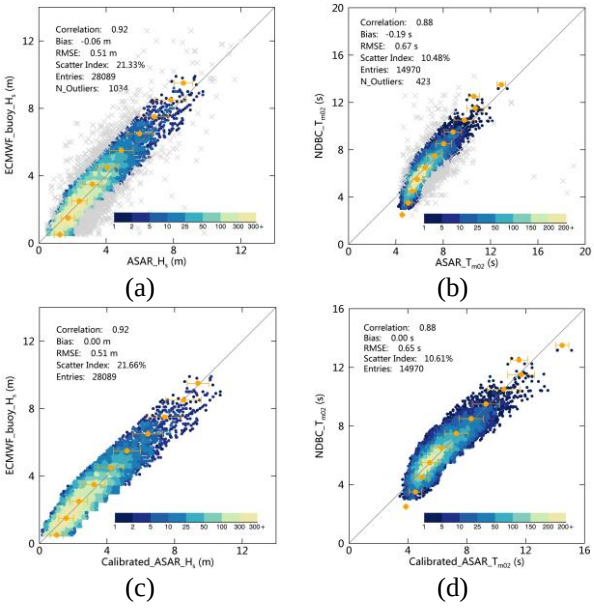


Figure 5 Comparisons of ASAR-derived SWH (a) and MWP (b) against the buoy measurements. The gray cross symbols in (a) and (b) are the detected outliers. Comparisons of the calibrated ASAR-derived SWH (c) and MWP (d) against the buoy measurements.

As shown in the **Figure 5** (a)-(d), the bias of SWH comparison is improved from -0.06 m to 0.00 m, and the bias of MWP is improved significantly from -0.19 s to 0.00 s, which suggests that the applied calibrations perform well for the ASAR-retrieved wave parameters. Though other three statistical parameters are not improved after calibration, the error bars overlaid on the scatter diagrams suggest that the overestimation and underestimation of the ASAR retrievals are significantly improved, as the mean values of each bins almost lie on the 1:1 diagonal line.

3.2 Validation

We further compared the calibrated ASAR-derived SWH with the calibrated Jason-1 RA data for cross-validation. By applying the same collocation criteria mentioned above and the Tukey fence to the ASAR and JASON-1 collocations, 22,862 pairs of SWH data are collected. The scatter diagram of the comparison is shown in **Figure 6**. It shows the calibrated ASAR SWH has a good agreement with the calibrated JASON-1 SWH, with the bias and RMSE of

0.18 m and 0.53 m, the correlation coefficient of 0.93 and the S.I. of 16.64%.

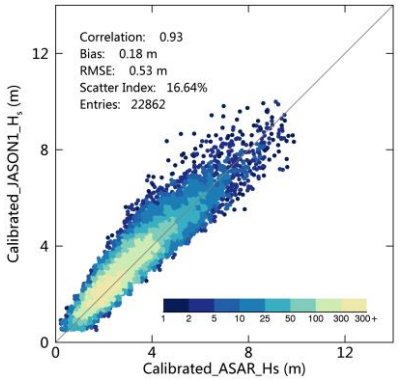


Figure 6 Comparison of the calibrated ASAR SWH with the calibrated RA SWH.

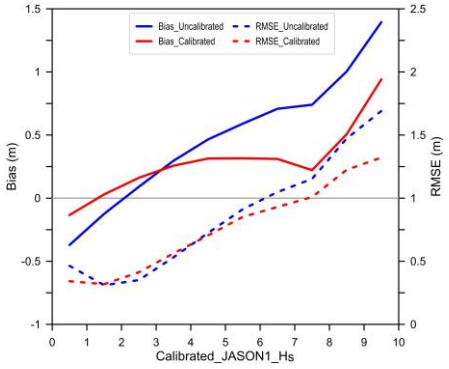


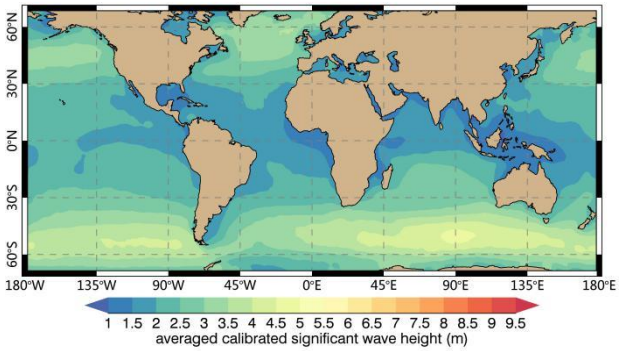
Figure 7 Variations of the ASAR derived SWH Bias and RMSE along with the calibrated RA SWH. The blue lines show the variations of bias and RMSE of uncalibrated ASAR-SWH and the red lines are the variations of calibrated results.

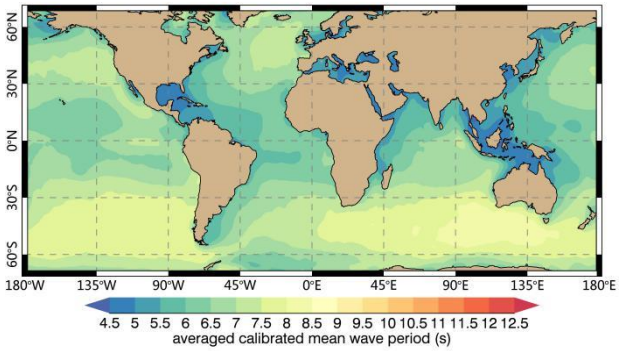
The **Figure 7** shows the changes of bias and RMSE of uncalibrated and calibrated ASAR-derived SWH against the calibrated JASON-1 SWH, where the solid and dash lines represent the bias and RMSE, and the red and blue lines refer to calibrated and uncalibrated ASAR results, respectively. The bias of the uncalibrated ASAR SWH increases along with sea states, which changes from negative values to positives ones when the SWH at approximately 2 m. The calibration process significantly reduces the bias to be less than 0.15 to 0.2 m from low to high sea state (at approximately 8 m) and more important is that the bias has less dependence on sea state. For very high sea state (SWH>9 m), the bias is of approximately 10% of the total SWH. RMSE of the calibrated ASAR SWH varies from 0.25 m to 1.20 m and it is particularly reduced for the sea higher than very rough sea state (above approximately 5 m).

4 Statistical analysis of the global ocean waves

Following the validation and calibration of the ASAR derived SWH and MWP, in this part, we used this 10-year global sea state dataset to conduct a preliminary analysis of global ocean wave characteristics.

As the sea ice in the polar regions may influence the retrieval of sea state parameters, the statistical analysis limits to latitudes lower than 70 degrees. The valid SWH is in the range between 0.5 m and 30 m, and the valid MWP is between 2 s and 20 s. By allocating the calibrated ASAR-derived SWH and MWP from each WM data to 2°x2° grids, the mean calibrated ASAR-derived sea state parameters are calculated, as shown in the maps of **Figure 8**. (a) and (b) show the mean characteristics of the calibrated ASAR-derived SWH and MWP, respectively. We further analyzed the extreme sea state by calculating the 99th percentile values of the calibrated SWH and MWP, as shown in **Figure 8** (c) and (d), respectively. No surprising that the extreme SWH (of approximately 10 m) occurs in the high latitudes' regions in the both Hemispheres, where the westerlies are dominant. Noting in the Northern Indian Ocean, the 99th percentile of SWH is also high, about 6 m. Comparing the patterns of extreme MWP to those of SWH, we found an interesting "spatial shifting" phenomenon. In the Northern Hemisphere, when the extreme SWH appears in the North Pacific and North Atlantic between 30°N and 60°N, the extreme patterns of MWP in the same region shift towards southeast beyond 30°N, indicating the extreme swell (with high MWP of approximately of 12 s) generated by North Pacific storms propagate southeast. The same feature is also observed in the Southern Hemisphere. The extreme SWH spans the 30°S and 60°S belt, while the extreme MWP clearly spans beyond 30°S further to north.





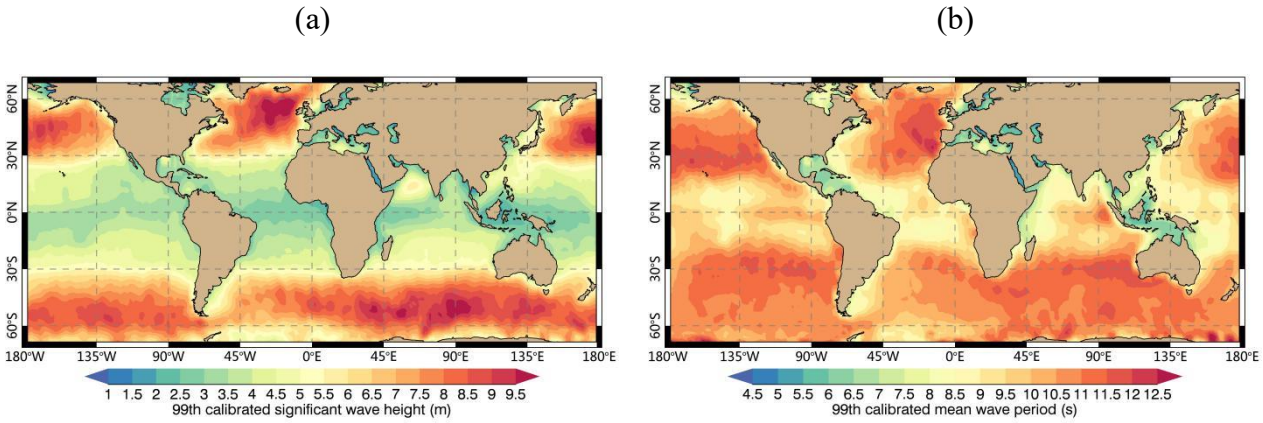


Figure 8 Global maps of (a) the mean calibrated SWH, (b) the mean calibrated MWP, (c) the 99th percentile of calibrated SWH over 10 years, and (d) the 99th percentile of calibrated MWP based on the ten-year ASAR WM data.

5 Conclusion

In this study, we presented the Cal/Val of the ASAR-derived integral ocean wave parameters in terms of SWH and MWP using the previously developed parametric model, based on the ten-year WM data. Preliminary analysis of the statistical characteristic of global ocean waves based on these data is also presented.

By comparing the ASAR-derived results with the in situ buoy data, good agreements are achieved with correlation coefficient of 0.89 and 0.83 for SWH and MWP. The RMSE and Bias of SWH comparisons are of -0.07 m and 0.62 m, respectively. The RMSE and Bias of MWP comparisons are of -0.21 s and 0.79 s, respectively.

By applying the Turkey fences and robust regression, outliers of the comparisons are removed. Then the RMA regression is applied to calibrate the ASAR retrievals using the buoy data. The Bias of ASAR-retrieved SWH and MWP are improved significantly, from -0.07 m to 0.00 m and from 0.19 s to 0.00 s, respectively. The most significance of the calibration process is to improve both the overestimation and underestimation trends in SWH and MWP. The cross-validation of calibrated ASAR-derived SWH also shows a good consistency with the JASON-1 RA measurements, with correlation coefficient of 0.93, the Bias and RMSE is 0.18 m and 0.53 m, respectively.

Based on the ten-year ASAR WM sea state parameters, we conducted the preliminary analysis of mean and extreme sea state of global ocean waves. Distinct features of mean and extreme sea state in the global oceans are found, which suggests the great potential of using this dataset for further analysis.

Acknowledgements It is acknowledged that the used ASAR wave mode Level 1B data to produce this dataset are kindly provided by the European Space Agency (ESA) under the framework of Dragon 4 Program. The buoy dataset and the Altimeter dataset are kindly provided by the GlobWave project. The study is partially supported by the National Key Research and Development Project (2018YFC1407100) of China.

6 Literature

[1] Young, I. R.: Zieger, S.: Babanin, A. V.: Global trends in wind speed and wave height. Science, 2011

[2] Ribal, A.: Young, I. R.: 33 years of globally calibrated wave height and wind speed data based on altimeter observations. Scientific Data, 2019

[3] Vesecky, J.F.: Stewart, R.H.: The observation of ocean surface phenomena using imagery from the SEASAT synthetic aperture radar: An assessment. J. Geophys., 1982

[4] Alpers, W.: Rufenach, C. L.: The effect of orbital motions of Synthetic Aperture Radar imagery of ocean waves., IEEE Trans. Antennas Propag., 1979

[5] Hasselmann, K.: Raney, R. K.: Plant, W. J.: Alpers, W.: Shuchman, R. A.: Lyzenga, D. R., Rufenach, C. L.: Tucker, M. J.: Theory of Synthetic Aperture Radar ocean imaging: A MARSEN view. J. Geophys. Res. 1985

[6] Huang, B.Q.: Zeng, K.: He, M-X.: On the effects of the region for obtaining first guess spectra on the directional ocean wave spectra retrieved by SAR with MPI algorithm. Periodical of Ocean University of China. 2017

[7] WAMDI GROUP.: The WAM model a third generation ocean wave prediction model. J. Phys. Oceanogry. 1984

[8] Hasselmann, K.: Hasselmann, S.: On the nonlinear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum. J. Geophys. Res.,1991

[9] Hasselmann, S.: Brüning, C.: Hasselmann, K.: Heimbach, P.: An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra. J. Geophys. Res., 1996

[10] Engen, G.: Johnson, H.: SAR-ocean wave inversion using image cross spectra. IEEE Trans. Geosci. Remote Sens., 2000

[11] Schulz-Stellenfleth, J.: König, Th.: Lehner, S.: An empirical approach for the retrieval of integral ocean wave parameters from synthetic aperture radar data. J. Geophys. Res., 2007

[12] Li, X. M.: Lehner, S.: Bruns.: Ocean wave integral parameter measurements using Envisat ASAR wave mode Data. IEEE Trans. Geosci. Remote Sens., 2011

[13] Stopa, J. E.: A. Mouche.: Significant wave heights from Sentinel-1 SAR: Validation and applications. J. Geophys. Res., 2017

[14] Schulz-Stellenfleth, J.: Lehner, S.: Measurement of 2-D sea surface elevation fields using complex Synthetic Aperture Radar data. IEEE Trans. Geosci.Rem. Sens. 2004

[15] Trauth, M. MATLAB Recipes for Earth Sciences. 4th Edn., Springer., 2015

[16] Zieger, S.: Vinoth, J.: Young, I. R.: Joint calibration of multiplatform altimeter measurements of wind speed and wave height over the past 20 years. J. Atmos. Ocean. Tech. 2009

[17] Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. J. Geophys. Res., 1998

[18] Tukey, J.: Exploratory Data Analysis., Addison-Wesley. 1977

[19] Rousseeuw, P.: Leroy, A.: Robust Regression and Outlier Detection 3rd edn, John Wiley and Sons, 1996