



Article

Estimation of Significant Wave Heights from ASCAT Scatterometer Data via Deep Learning Network

He Wang ^{1,2,*} , Jingsong Yang ^{2,3} , Jianhua Zhu ¹, Lin Ren ^{2,3}, Yahao Liu ^{4,5,6}, Weiwei Li ¹ and Chuntao Chen ⁷

¹ National Ocean Technology Center, Ministry of Natural Resources, Tianjin 300112, China; besmile@263.net (J.Z.); 13821227862@163.com (W.L.)

² State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China; jsyang@sio.org.cn (J.Y.); renlin210@sio.org.cn (L.R.)

³ Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519082, China

⁴ CAS Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China; yhliu@qdio.ac.cn

⁵ Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao 266071, China

⁶ Laboratory for Ocean and Climate Dynamics, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China

⁷ School of Ocean, Yantai University, Yantai 264005, China; chenchuntao@ytu.edu.cn

* Correspondence: wanghe_sio@126.com; Tel.: +86-22-2753-6513

Abstract: Sea state estimation from wide-swath and frequent-revisit scatterometers, which are providing ocean winds in the routine, is an attractive challenge. In this study, state-of-the-art deep learning technology is successfully adopted to develop an algorithm for deriving significant wave height from Advanced Scatterometer (ASCAT) aboard MetOp-A. By collocating three years (2016–2018) of ASCAT measurements and WaveWatch III sea state hindcasts at a global scale, huge amount data points (>8 million) were employed to train the multi-hidden-layer deep learning model, which has been established to map the inputs of thirteen sea state related ASCAT observables into the wave heights. The ASCAT significant wave height estimates were validated against hindcast dataset independent on training, showing good consistency in terms of root mean square error of 0.5 m under moderate sea condition (1.0–5.0 m). Additionally, reasonable agreement is also found between ASCAT derived wave heights and buoy observations from National Data Buoy Center for the proposed algorithm. Results are further discussed with respect to sea state maturity, radar incidence angle along with the limitations of the model. Our work demonstrates the capability of scatterometers for monitoring sea state, thus would advance the use of scatterometers, which were originally designed for winds, in studies of ocean waves.

Keywords: Advanced Scatterometer (ASCAT); significant wave height; WaveWatch III; deep learning; multi-hidden-layer neural network



Citation: Wang, H.; Yang, J.; Zhu, J.; Ren, L.; Liu, Y.; Li, W.; Chen, C. Estimation of Significant Wave Heights from ASCAT Scatterometer Data via Deep Learning Network. *Remote Sens.* **2021**, *13*, 195. <https://doi.org/10.3390/rs13020195>

Received: 23 November 2020

Accepted: 5 January 2021

Published: 8 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The knowledge of ocean surface wave is important for various scientific and operational studies. Significant wave height (SWH), which is traditionally defined as the average of the 1/3 largest in a record of ocean wave heights, is the most valuable and commonly used parameter for ocean wave applications, e.g., ocean engineering, marine navigation and wave power evaluation. For decades, the SWH observations from space have mostly relied on altimeters at a global scale. The space-borne altimeters provide robust SWH with the precise <0.5 m. However, their nadir measurements along track suffer from a sparse spatial sampling (see e.g., [1,2]). Another globally SWH data source is from synthetic aperture radars (SARs) operating in wave mode, which could produce reliable estimates only for long waves (swells) [3] or total SWH by empirical algorithms (i.e., CWAVE-like models see [4–7]). Additionally, SAR wave mode measurements are even less dense than altimeters (e.g., 100 km sampling along the Sentinel-1 orbit).

Polar-orbiting scatterometers, the real-aperture radar instruments, are monitoring oceans for large regions with high temporal revisit compared to altimeters and SARs. For instance, the Advanced Scatterometer (ASCAT) onboard Meteorological Operational satellite (MetOp-A) of European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), which was launched in 2006 and is still in orbit presently, achieves global ocean coverage in 1.5 days approximately. Unfortunately, the primary oceanic objective of scatterometers is to provide information on the winds [8] rather than waves. Thus, estimating SWH from the scatterometers with wide swath and rapid revisit is an attractive alternative compared to other existing wave sensors from space.

In fact, the long ocean waves indeed affect the normalized radar cross section (NRCS, or σ^0) observed by the C-band radars using real (scatterometer) or synthetic aperture (SAR) [9–11]. Moreover, it is evident that sea state is a factor impacting wind speed retrieval from radars in C-band [12]. The main mechanism for this is that the presence of large-scale waves changes the wind stress [13,14], and impact small-scale roughness, consequently, changes the radar backscattering over the ocean. In this context, although they are typical ocean wind sensors, theoretically, estimating sea state information is feasible from the wide-swath scatterometers.

However, it remains a challenge. Nowadays, the operational scatterometers wind retrieval is based on the relationship between radar backscattering and the wind vectors, which has been well described by the empirical geophysical model function (GMF) (e.g., the continuously updated CMOD family for C-band, see [15,16]). In contrast, the relation between the scatterometer NRCS and SWH is complex and subtle compared to the known wind-NRCS GMF, so that no GMF explicitly exists relating waves to NRCS at present.

In the literature, the only efforts aiming at SWH estimation from scatterometers were made by Guo et al. These studies focused on retrieving SWH from ERS-1/2 [17] and QuikSCAT [18] sensors through the simple artificial neural network, which could be considered as a nonlinear data driven technique to interconnect the complicated relationships between inputs and outputs. In recent years, the technology of artificial neural network has dramatically progressed and been widely used for geophysical retrieving ocean wind/waves from microwave sensors, for example, estimating SWH [6] and ocean winds [19,20] from C-band SARs. Especially, as the growing availability of big Earth data, deep learning techniques (by employing larger and deeper neural networks) have emerged and been increasingly applied in remote sensing field [21–23]. However, the pioneering works [17,18] conducted a decade ago utilized a shallow architecture of three-layer neural networks which are based on a small dataset of co-locations between scatterometer and the geographical limited buoys.

This paper proposes a novel approach for estimating SWH from MetOp-A/ASCAT based on deep learning neural network by training enormous matchups of scatterometer and numerical wave model hindcast at a global scale. The remainder of the paper is organized as follows. The data sources and methods for collecting matchups are introduced in Section 2, followed by the analysis on ASCAT-buoy data in Section 3 to investigate the possibility of estimating SWH from ASCAT variables. In Section 4, we develop the deep learning neural network model for inferring SWH from ASCAT. The validation results with respect to the deep learning SWH approach for ASCAT are presented in Section 5, followed by the discussion in Section 6. Finally, conclusions and perspectives are given in the last section.

2. Data Sets

2.1. ASCAT Data

MetOp-A/ASCAT is a vertically polarized C-band (5.255 GHz) scatterometer with two sets of three antennae. The fixed fan-beam antennae are oriented at 45° (fore-beam), 90° (mid-beam), and 135° (aft-beam) with respect to the satellite flight direction, resulting in two swaths of 550 km on both sides separated by a gap of about 360 km [24]. Here,

we focus on the ASCAT data over a three-year period from January 2016 to December 2018. Two types of MetOp-A/ASCAT products are used: Level-1B data containing NRCS measurements released by EUMETSAT, and the Level-2 wind vector products which were generated based on Level-1B data by Royal Netherlands Meteorological Institute (KNMI) using GMF of CMOD5n [15]. Both Level-1B and Level-2 ASCAT products are with grid spacing of 25 km and approximately 50 km resolution.

For each 25×25 km wind vector cell (WVC) of ASCAT, the following variables are employed for the development of deep learning based SWH algorithm:

- (1) Triplet of NRCS (σ_{fore}^0 , σ_{mid}^0 , and σ_{aft}^0 , denoting fore-beam, mid-beam and aft-beam, respectively);
- (2) Triplet of incidence angles, (θ_{mid} , θ_{fore} , and θ_{aft} , ranging of $25\text{--}53^\circ$ and $34\text{--}64^\circ$ for middle beam and side beams, respectively);
- (3) Triplet of radar backscattering variability factor (K_p) defined as

$$K_p = \frac{\sqrt{\text{var}(\sigma^0)}}{\sigma^0} \quad (1)$$

This value can be regarded as a metric of the uncertainty in the mean backscatter (σ^0) resulted from sensor speckle noise, data processing, and also spatial heterogeneities of the ocean surface [25,26];

- (4) Ocean wind vectors, including 10-m wind speed (U_{10}) and triplet of cosine values of wind direction relative to the radar beams ($\cos \varphi_{fore}$, $\cos \varphi_{mid}$ and $\cos \varphi_{aft}$).

2.2. Numerical Wave Model Hindcast

The significant wave height hindcast used here are from the database of Integrated Ocean Waves for Geophysical and other Applications (IOWAGA) project of Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER). The wave hindcasts were performed using the numerical ocean wave model of WaveWatch III (WW3) with the parameterization described in [27,28]. The spatial and temporal sampling of WW3 is 0.5° and three hours, respectively.

IOWAGA WW3 is forced by the European Centre for Medium-range Weather Forecasts (ECMWF) winds (T1279, since January 2010), which is of approximately 64 km effective resolution [29], and the corresponding resolution of IOWAGA WW3 used here is around 64 km. Hence, datasets of WW3 (~ 64 km resolution) and 25 km spacing ASCAT (~ 50 km resolution) are roughly of the same scale (note that 12.5 km spacing ASCAT is of approximately 28 km resolution, much smaller than WW3), making sampling artefacts less of an issue here for training our deeply learned algorithm.

2.3. ASCAT-WW3 Matchups: Training, Validation and Test Data

In this study, the deep learning dataset was built by col-locating the ASCAT measurements and WW3 SWH hindcasts within the spatio-temporal criteria of 0.1° and 0.5 h. Here, we did not adopt the col-location methodology by linearly interpolating (in space and time) the model grided data to every points of satellite data, since it would modify the model outputs and bring additional uncertainties.

Following the procedure used previously e.g., [25], ASCAT data at high latitudes ($>50^\circ$) and less than 100 km offshore distance are excluded, in order to avoid sea ice and land/island contaminations, respectively. In addition, the following quality control procedures are applied.

- (1) Data flagged as suspicious wind retrievals in ASCAT Level-2 products, with rejection percentage of 12.7% approximately.
- (2) Apart from the wind and sea state, oceanic rainfall also affects the C-band radar backscattering (e.g., for scatterometers [30] and SARs [31]). Thus, we used Integrated Multi satellite Retrievals for Global Precipitation Measurement (Imerg) [32] late

product (version 6) with global grid of 0.1° and 0.5 h to reject rainy conditions (rain rate >0 mm/h). Rate of rejection according to this quality control is around 9.4%.

This yields a database containing more than 16 million data pairs. The entire matchups are then randomly shuffled and split into three groups: training (50%), validation (20%), and test (30%) set. The training and validation data set are used for the developing of the deep learning model, while the formal is for directly tuning the parameters (weights and biases) of the model and the latter is for the cross-validating and determining the hyperparameters (see Section 4 for details). In contrast, the test data, which are never seen by deep learning model during the tuning procedure (i.e., independent on the model training), are remained for the evaluation. The locations of ASCAT-WW3 col-locations (training set totaling of 8,102,567 points) are shown in Figure 1, representing a global spatial distribution.

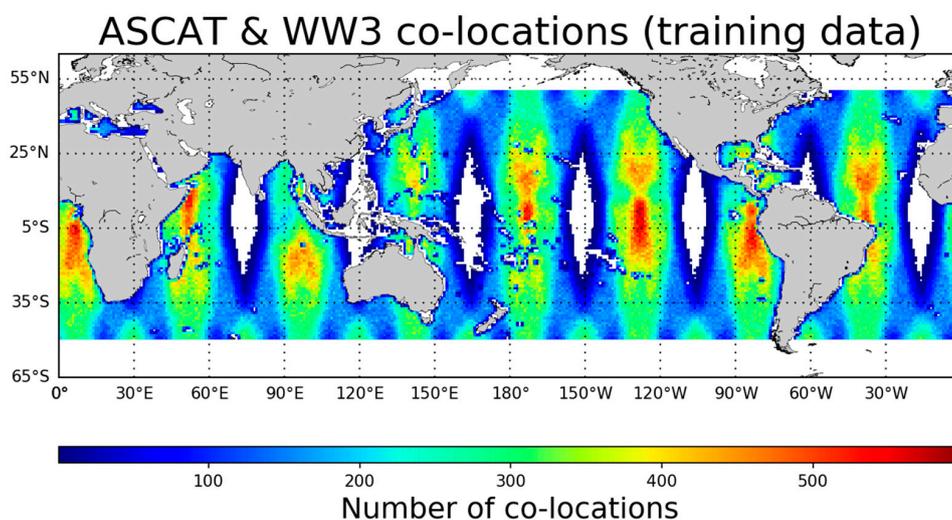


Figure 1. Number of Advanced Scatterometer (ASCAT)-WW3 matchups (training dataset) on $1^\circ \times 1^\circ$ grid.

2.4. ASCAT-Buoy Co-Locations

We also carried out additional co-location against buoy in situ. Hourly SWH records were collected from buoy network of National Data Buoy Center (NDBC). As depicted in Figure 2, 39 NDBC buoys used here are moored off Hawaii Islands (red box), Gulf of Mexico (green box) and the North America coasts. In this study, we collocated the ASCAT data with the buoy in situ by limiting the distance within $25/\sqrt{2}$ km and time separation less than 30 min.

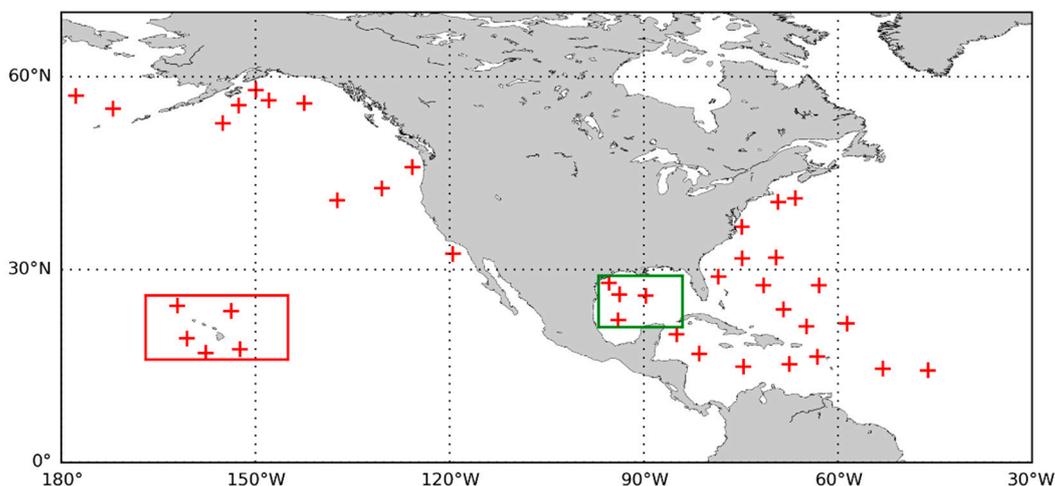


Figure 2. Locations of National Data Buoy Center (NDBC) buoys.

It should be emphasized here that, the ASCAT data that are co-located with buoys were excluded from the training data (ASCAT-WW3) to ensure the independence of the evaluating using buoys on the modeling of deep learning algorithm.

3. Data Analysis Based on ASCAT/Buoy Co-Locations

3.1. Wind–Wave Relationship

Figure 3 shows the dependency of significant wave height on wind speed measured from buoy. Here, buoy wind speeds at different heights have been converted to the equivalent neutral 10-m winds (U_{10} using the coupled ocean atmosphere response experiment bulk algorithm [33]).

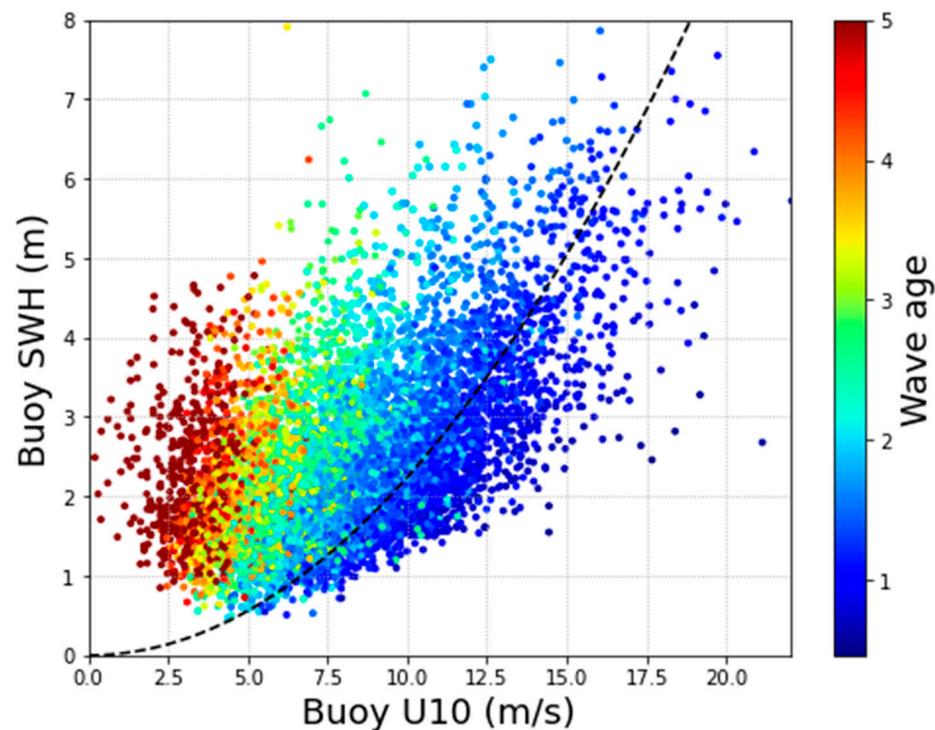


Figure 3. Scatter diagram of 10-m wind speed and significant wave height (SWH) from ASCAT-buoy co-located data set. The color denotes wave age, and dashed curve presents the wind–wave relation in fully developed sea state from Pierson–Moskowitz spectrum according to Equation (2).

It is clear from Figure 3 that the connection of wind speed and SWH is not good enough so that it is difficult to predict SWH only from wind speed via simply regression approach. For instance, under different sea condition (developing or fully developed wind wave, swell, and mixed sea), wind–wave relations differ. This is illustrated in Figure 3 with colors denoting the values of wave age (ratio of peak phase speed of wave to wind speed), which characterize the sea state maturity.

Regarding fully developed sea state, which corresponds to unlimited wind fetch and wave age of approximately 1.2, the wind–wave connection could be modelled by well-known Pierson–Moskowitz spectrum [34]:

$$SWH_{pm} \approx 0.22 \frac{U_{10}^2}{g} \quad (2)$$

where g denotes the gravity acceleration. The curve of this relationship is plotted as dashed line in Figure 3. Using this assumption, the SWH prediction from wind speed has accuracy of 1.25 m against in situ in terms of root mean square error (RMSE) defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{j=1}^N (x_i - y_i)^2} \quad (3)$$

If we employ the advanced wave spectrum model proposed by Elfouhaily [35], in which the wave age is taken into account, the RMSE of 1.01 m is obtained for Elfouhaily SWH prediction against buoy.

Although wind and wave are highly coupled with each other, from our study, the prediction only using wind speed (even take sea state maturity, which is unknown in ASCAT, and the geophysical law into consideration) could not reach a good performance. We need to seek the possibility of using ASCAT NRCS as additional inputs in our deep learning approach.

3.2. ASCAT Wind Speed Accuracy and Sea State Impact

Since we plan to use ASCAT-derived wind speed instead of the surface truth, it is necessary to explore its accuracy and particularly the SWH impact on scatterometer derived U_{10} in advance. In the literature, assessments of ASCAT wind speed retrievals against buoy in situ show reliable accuracy (e.g., [36]). On the basis of our dataset, comparison provides good consistency with RMSE of 0.85 m/s and bias of 0.14 m/s for ASCAT wind speed retrievals, which is in agreement with independent studies (e.g., RMSE of 1.10 m/s in [37]).

In order to investigate the impact factors in ASCAT wind speed retrievals, we computed the residuals between ASCAT and buoy wind speed. Then, the correlation coefficients between the U_{10} residuals and various buoy measured parameters including SWH were calculated and listed in Table 1. The COR is expressed as:

$$\text{COR} = \frac{\sum_{i=1}^N (x_i - \langle x_i \rangle) - (y_i - \langle y_i \rangle)}{\sqrt{\sum_{i=1}^N (x_i - \langle x_i \rangle)^2 \sum_{i=1}^N (y_i - \langle y_i \rangle)^2}} \quad (4)$$

Table 1. Correlation coefficients between U_{10} residuals (ASCAT minus buoy) and various oceanic/atmospheric parameters.

Parameter	Correlation Coefficient
Buoy U_{10}	−0.24
Buoy SWH	0.12
Air temperature at sea surface	−0.21
Sea surface temperature	−0.18
Air-sea temperature difference	−0.20
Air pressure	−0.07

Our analysis indicates that SWH and ASCAT retrieved U_{10} error are related (COR of 0.12), which is consistent with the findings from previous studies (e.g., [9]). In fact, Table 1 shows that SWH is not the only and even not the most significant impact factor for wind speed residual. As addressed by Stopa et al. [11], this could be explained by the fact that merging of the three antennae measured NRCS (at different azimuthal and incidence angles) in wind retrieval processing mitigates the sea state impacts.

Therefore, from this experiment, it is possible to use ASCAT-derived wind speed in our SWH deep learning model instead of the surface truth of U_{10} .

3.3. Influence of Sea State on ASCAT NRCS

The first studies on sea state impact on backscattered radar cross section were conducted in the 1980s [38,39]. From previous works, it is evident that sea state (ocean wave height and/or the sea maturity) influences on the NRCS from altimeters (e.g., [39,40]) and scatterometers (e.g., [9]). This property has been taken into account and were used to develop two-parameter wind retrieval algorithms for nadir-viewing altimeters [40,41].

Here, we investigate the wave height influence on ASCAT NRCS by computing correlation coefficients between NRCS residual and buoy observed SWH at each wind speed bins from 3.5 ± 0.1 to 11.5 ± 0.1 m/s. NRCS residuals were computed as the difference between ASCAT observations and predictions using GMF of CMOD5n [15]. As illustrated in Figure 4, the correlations are stronger under low wind conditions (corresponding large value of wave age, or swell dominated sea state). In addition, we carried out the analysis using winds derived from ASCAT (blue curve in Figure 4) or buoy “surface truth” (red in Figure 4) as inputs to CMOD5n. Although the correlations are lower when using ASCAT winds than buoy in situ, the same dependency trend could be clearly seen from Figure 4. It is worth noting that we obtained the same results (not shown) by using another GMF of CMOD5h [42].

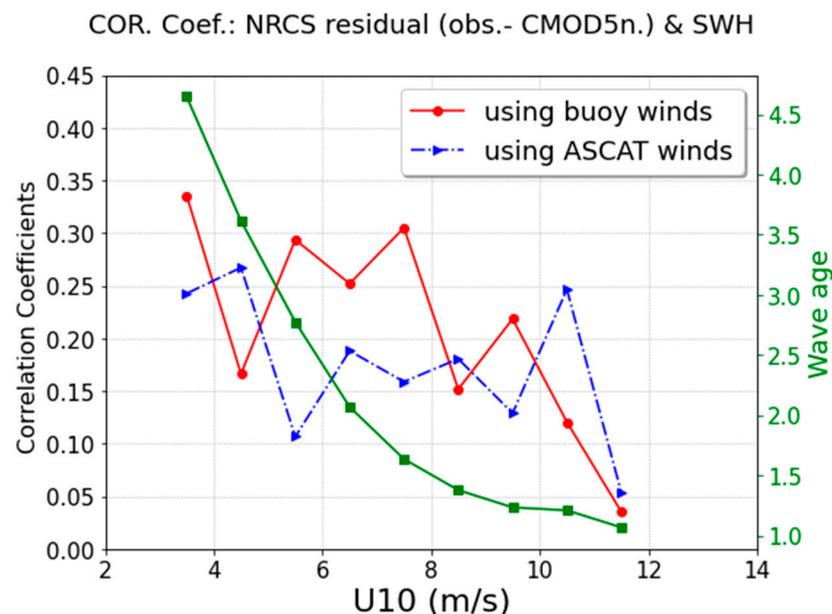


Figure 4. Correlation coefficients between normalized radar cross section (NRCS) residual (ASCAT measured minus geophysical model function (GMF) of CMOD5n predicted) and SWH for various wind speed bins. Red and blue symbols denote the inputs to CMOD5n from winds derived from ASCAT or measured by the buoy. Green squares present the corresponding wave age for each wind speed.

To summarize, from our analysis based on the ASCAT-buoy data, we expect that combining radar backscattering and wind speed from ASCAT could work in deeply learned SWH prediction.

4. Development of Deep Learning Model

4.1. Establishment of Deep Learning Network

As illustrated in Figure 5, the proposed neural network contains input layer representing ASCAT observables, several hidden layers and the rightmost output layer with one node corresponding to significant wave height estimates. The inputs here are 13 ASCAT observables: 4×3 measurements (NRCSs, incidence angles, azimuthal angles and backscattering variability factor) from fore, mid and aft-beams along with ASCAT inferred wind speed. The selection of these observables will be discussed in Section 4.2.

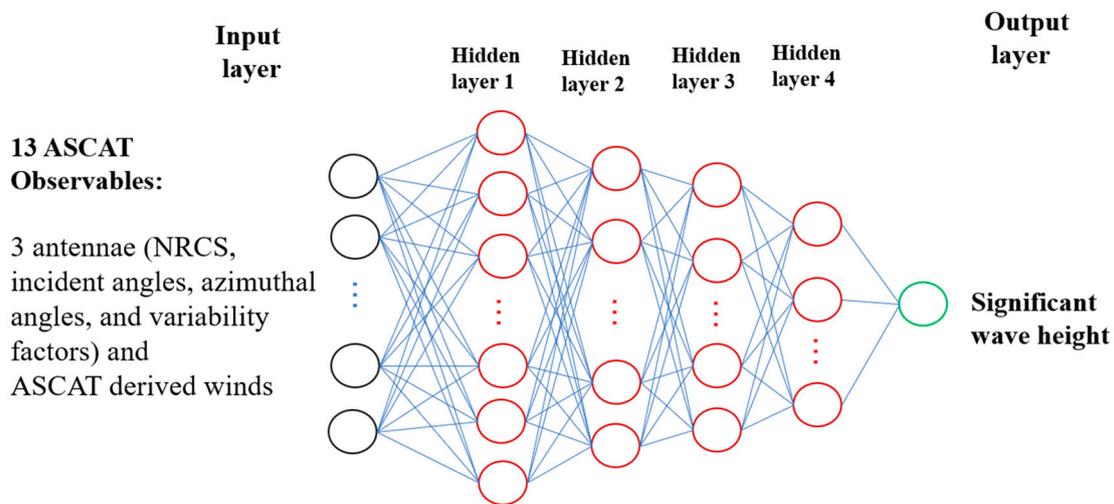


Figure 5. Architecture of the deep neural network.

In this multi-hidden layer network, for each neuron (h_i), the computation is a linear combination of nodes in the previous layers (h_{i-1}), followed by a nonlinear activation function f . The expression for this unit processing is:

$$h_i = f(\text{dot}(W_i, h_{i-1}) + b_i) \quad (5)$$

where W_i (weights) and b_i (bias) are the training parameters corresponding to this node, and the entire network is constructed by these layer-wise operations from the input to output layer.

Here, to tune the optimal training parameters (weights and biases) using the training dataset, we use the standard back-propagation algorithm to minimize the loss function:

$$J = \sum_{n=1}^m (y_n - Y_n)^2 \quad (6)$$

where y_n are the output value (i.e., SWH estimated from the model), and Y_n are the “true” data (i.e., SWH from WW3).

To optimize the model hyperparameters, grid search tests were carried out and the ones that maximize the model scores with respect to the validation dataset [43] was selected. Finally, the model proposed here has 4 hidden layers with neurons number of 512, 256, 128, and 64, respectively, by employing Mish [44] as the nonlinear activation function.

In addition, in order to obtain a more robust multi-hidden layer network, the following techniques are used.

- (1) The optimizer to train neural network. Nesterov Adaptive Moment Estimation (Nadam) [45,46] with a batch size of 512 has been selected among several other existing optimizers (i.e., stochastic gradient descent [47] or Adam [48]). This is because the loss function minimizing process converges much faster when using Nadam after a series of experiments for our model. We follow the recommendation of Nadam parameters [46]: $\beta_1 = 0.9$, $\beta_2 = 0.999$.
- (2) A learning rate schedule, called “ReduceLRonPlateau”, is employed. This simple trick decreases the learning rate by a factor of 10 once performance regarding validation dataset has stopped improving. Consequently, the model could efficiently benefit from this reducing strategy once learning stagnates.
- (3) For denser and deeper networks, the model complexity often makes the training process continue for too long, and this results in an overfitted model which fails to generalize. In our experiment, we invoke a “early stopping” trick, which monitors loss

function on validation dataset and quits training when there is no further improving for 10 continuous epochs.

In our study, the algorithm is implemented based on the Python deep learning library of Keras (version 2.2.4) [49] with Tensorflow backend [50].

4.2. Feature Selection for Deep Learning Network

In this section, we aim at quantifying the importance of the input features and seeking the optimal in the constructed deep learning neural network by comparing model scores based on validation dataset.

Recall the description in Section 2.1. For a certain ASCAT WVC, addition to wind speed U_{10} , possible inputs could consist in two sets of observation regarding at least three beams: (1) NRCS σ^0 along with associated geometric parameters θ and $\cos \varphi$, and (2) backscattering variability factor K_p . Hence, there are a total of 14 permutations as listed in Table 2, together with their performances using the metric RMSE.

Table 2. Performance (root mean square error, unit: m) of deep learning model against validation set using different variables from ASCAT.

Radar Beam Permutations		$\sigma^0 + \theta + \cos\varphi + U_{10}$	$\sigma^0 + \theta + \cos\varphi + U_{10} + K_p$
One beam	Fore	0.7367	0.7261
	Mid	0.7153	0.7030
	Aft	0.7319	0.7263
Two beams	Aft+Fore	0.6168	0.6087
	Mid+Aft	0.6384	0.6335
	Mid+Fore	0.6365	0.6305
Three beams	Fore+Mid+Aft	0.5841	0.5746

Two trends are revealed from Table 2.

- (1) Performance is improving due to diversity of antenna.

The accuracy of the estimates is considerably improved by RMSE of approximately 0.15 m when using the combination of three beams as opposed to only one. This result indicates that jointly use of the ASCAT triplet is critically important for deriving sea states from scatterometer, whereas observations from only one beam were considered when Guo et al. developed their neural network for ERS-1/2 [17].

There are two aspects of antenna diversity, i.e., incidence angle and antenna-looking azimuth, impacting the model skill. Regarding incidence angle, if only one antenna observation is used, the best is found when applying the mid beam. The possible reason is that wave impact on scatterometer NRCS decreases with increasing incidence angle [10], noting that mid beams cover the range 25° to 55° , and the side beams 35° to 65° , and in terms of antenna-looking azimuth, as shown in two beam permutations, it is apparent that the model using combination of two side beams outperforms the combination of mid-side beams. These findings implicate that the wave-led tilt modulation of NRCS is dependent on both incident and azimuthal angle.

- (2) Adding K_p works for performance improvement.

A slightly improved verification score is obtained after incorporating K_p regardless of beam permutations, reflecting the relative importance of K_p for SWH estimation. In fact, although K_p is commonly regarded as metric of scatterometer measurement noise, it is evident that K_p contains geophysical information and has been demonstrated in ice-type classification [25]. Besides, in analogy to K_p , there is a parameter called normalized variance $cvar$ in SAR community for describing the image I_s homogeneity. $cvar$ is expressed by

$$cvar = \text{var}\left(\frac{I_s - \langle I_s \rangle}{\langle I_s \rangle}\right) \quad (7)$$

and it has been proved that, in addition to SAR NRCS, the SWH is related to the image homogeneity *cvar* [4,6,7].

Additionally, very little improvement (no shown) has been found when incorporating NRCS predictions using the GMF of CMOD5n [15] or CMOD7 [16]. This is likely due to the fact that the neural networks, which are known to be good at modeling nonlinear interactions between the input variables, have already taken the wind-NRCS GMF into account implicitly.

4.3. Primary Comparison against Baseline

Here, we define a baseline for SWH estimation solely from wind speed. Using Pierson–Moskowitz spectrum in assumption of fully developed sea [34], a RMSE of 1.33 m is obtained using validation data for this simple solution. Since Pierson–Moskowitz model is only applicable to the fully developed sea state, another baseline could be defined by the more mature model from Elfouhaily spectrum [35], by using inverse wave age computed from WW3 outputs. We obtained a better result with RMSE of 0.95 m by Elfouhaily model.

From Table 2, performances of all neural networks using different variable combinations (the largest RMSE = 0.7367 m) outperform the baselines of Pierson–Moskowitz or Elfouhaily. This indicates the effectiveness of the proposed deep learning model.

To end, optimal combination of the inputs has thirteen variables: four sets of triplets for three ASCAT beams (NRCS, incident angles, azimuthal angles, and backscattering variability factors) and the ASCAT derived wind speed. Thus, the neural network trained using these inputs features will be evaluated and discussed hereafter.

5. Performance Verification

In this section, the performance of the established deep learning model with aforementioned optimal input variables combination ($\sigma^0 + \theta + \cos \varphi + U_{10} + K_p$ for three beams and U_{10}) is assessed using WW3 hindcast and buoy in situ. The statistics of bias, RMSE, the Scatter Index (SI), and correlation coefficient (COR) used in the evaluation:

$$\text{bias} = \frac{1}{N} \sum_{i=1}^N (x_i - y_i) \quad (8)$$

$$\text{SI} = \frac{1}{\langle x_i \rangle} \sqrt{\frac{1}{N} \sum_{i=1}^N [(y_i - \langle y_i \rangle) - (x_i - \langle x_i \rangle)]^2} \times 100\% \quad (9)$$

5.1. Comparison against WW3

Figure 6 shows the comparison of significant wave heights from ASCAT against WW3 based on the collocation independent on the deeply learning process of neural network (i.e., test dataset, see Section 2.4). From the entire dataset, the performance of 0.59 m RMSE, 0.85 correlation coefficient, and 23.12% SI is presented. If we exclude the outliers that had a SWH difference between ASCAT and WW3 larger than three standard deviations, with the percentage of rejection on the order of 1%, the RMSE and SI could be reduced to 0.54 m and of 21%, respectively.

Wave height data are binned in 1 m intervals from WW3 hindcast as overlaid in the scatter plots in Figure 6, with the error bar representing the standard deviation for each bin. At low sea state (<1 m), one could see a remarkable overestimation of ASCAT estimates from deep learning. The explanation for this low-SWH insensitiveness of our approach is probably the fact that the radar backscatter signals are weak and noisy due to the nondominant contribution of Bragg scattering in this circumstance (note that low sea states are commonly associated to light winds).

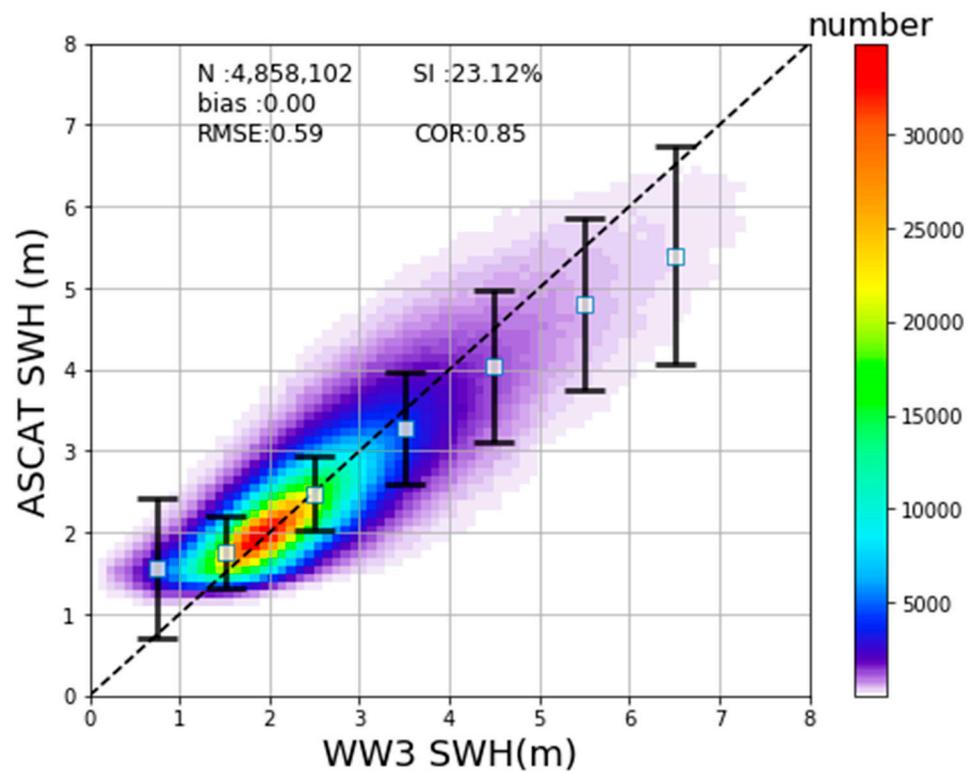


Figure 6. Significant wave heights estimated from ASCAT via the deep learning network versus WW3 hindcast data (from test dataset). Colors denote the data point numbers within $0.1 \text{ m} \times 0.1 \text{ m}$ bins.

Besides, an underestimation (bias of -0.67 m) and the relatively large RMSE (0.94 m) could be found for high waves ($>5 \text{ m}$). Likewise, degradation performance of scatterometer derived winds is also found at high winds region (e.g., [51]). In addition, number of matchup for high sea state is limited (account for 3% only when $\text{SWH} > 5 \text{ m}$), which may be responsible for the imperfect modelling skill by deep learning in this regime.

Hence, from this comparison, against WW3 hindcast at global scale, the best performance (RMSE of 0.5 m , SI of 20% and correlation coefficient of 0.80) could be achieved under moderate sea condition ranging from 1.0 to 5.0 m (93% for total test dataset). In terms of the wind speed, this favorable sea state (1.0 – 5.0 m) corresponds to the range from 3.2 m/s (0.1th percentile) to 16.8 m/s (99.9th percentile) in this study. These scores are close to the SWH retrieval requirement of altimeters.

5.2. Buoy Comparison

Figure 7 shows the comparison of the SWH retrieval from MetOp-A/ASCAT and NDBC buoy in situ. One can see a reasonable SWH estimates from ASCAT with respect to buoy. The results provide larger RMSE and SI (0.7 m and 28%) compared to the WW3 verification although the similar correlation coefficient of 0.78.

The degraded scores regarding SI and RMSE against buoys than numerical modeling is consistent with the published studies (see SAR SWH empirical algorithms with WW3 [4,6]). For instance, WW3 trained neural network for SAR radar images yields SWH retrievals of 0.58 m and 0.72 m RMSE compared to WW3 and buoy data, respectively, as reported by Stopa et al. [6], which is quite close to our results here.

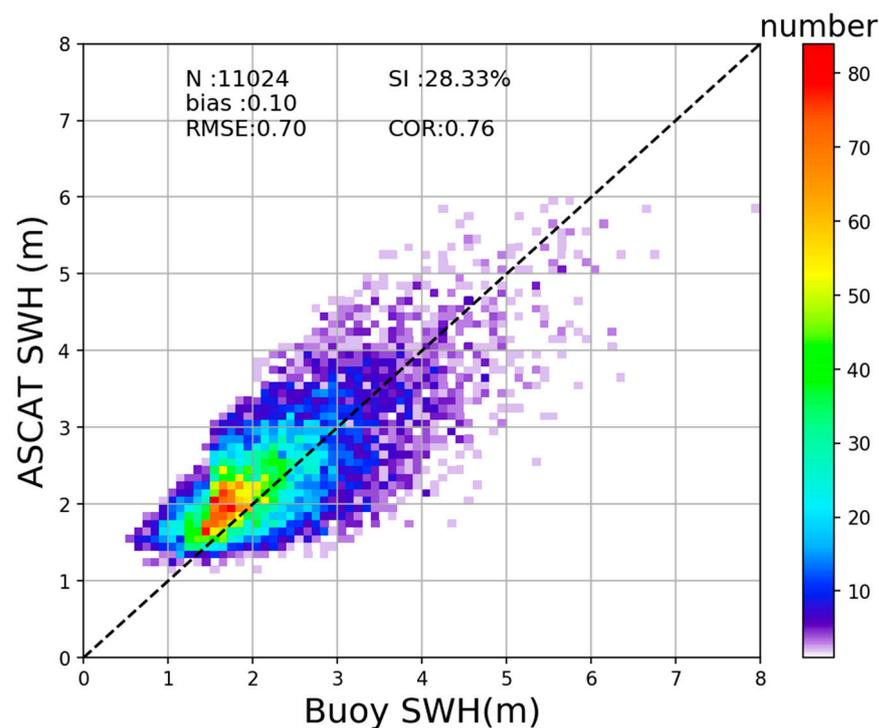


Figure 7. ASCAT SWH deeply learned estimates versus buoy measurements.

The reasons could be various:

- (1) Modelling predictions versus in situ observations. Although IOWAGA WW3 has been proved to be a reliable database [28], predictions are actually numerical instead of in situ and may result in discrepancies. For instance, on the basis of our dataset, RMSE of 0.32 m is found for WW3 against buoy SWH.
- (2) WW3 outputs are distributed almost globally (Figure 1), but the geographic coverage of NDBC buoys is regional (Figure 2). Particularly, the average observed SWH is 2.07 m from buoys while WW3 SWH hindcasts have a mean value of 2.52 m in this study. This means that buoy observations are skewed toward low sea states, where our proposed model are unfavorable. (See the regionality analysis in Section 6.2.)
- (3) Besides, the point observations from buoy could not be regarded as “truth” data free of error, especially within the 50 km radar footprint. Thus, the representativeness error of buoy measurements may be also responsible due to the sampling artefacts, as documented in triple collocation analysis (e.g., [52,53]).

Although these could be the possible explanations of the larger errors compared against buoys found here, it remains to do further studies. One feasible remedy for this could be the further cross-calibration between ASCAT derived SWH and buoy data (many more matchups needed) to refine our model (e.g., see corrections for 10-years SAR sea state products [54]). However, this is beyond the scope of the paper.

6. Discussion

In this section, we explored the applicability of the proposed deep learning model under various conditions and regions. Since the huge number of data points, influence investigations were still based on the WW3 test data.

6.1. Wave Maturity Influence

Mixed conditions of local generated wind sea and remotely propagated swell are common in the ocean. Although sophisticated sea state classification method has been

proposed [55], here we simply use parameter wave age to describe the maturity of the sea state. This mostly used notation is expressed as

$$\beta = \frac{gT_p}{2\pi U_{10}} \quad (10)$$

T_p represent the WW3 hindcast wave period at spectral peak. According to Pierson–Moskowitz [34], the sea state can be classified as swell when the wave age $\beta > 1.2$. In general, lower values of wave age correspond to younger wave maturity.

The variation of RMSE (red) as a function of the wave age along with bias (blue) for our proposed deep learning model is presented in Figure 8. It is shown that the sea state maturity indeed affects the ASCAT SWH deep learning model significantly. In other words, proposed deep learning model performs better for swell (larger wave age) than wind sea (smaller wave age). This behavior is consistent with sea state influence on ASCAT-CMOD5n NRCS residual addressed in Section 3.3.

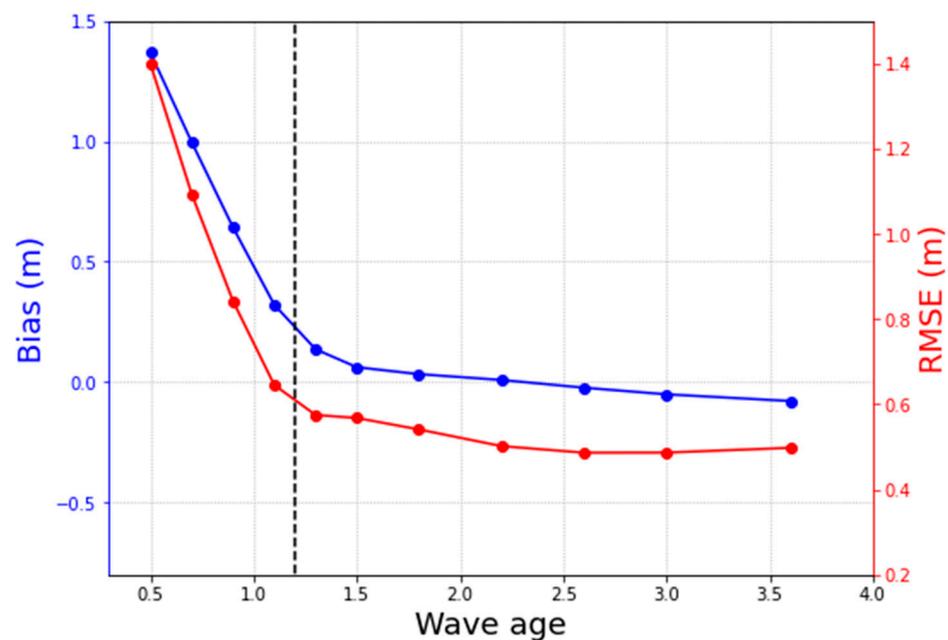


Figure 8. Performances of the deep learning derived ASCAT SWH against WW3 hindcast with respect to wave age. Blue and red symbols represent bias and RMSE, dashed line is for wave age = 1.2.

6.2. Regionality Analysis

In Figure 9, we show the geographical distribution of the errors regarding our deep learning prediction. The global maps of bias and RMSE indicate that the proposed approach performs better in some regions (e.g., eastern tropical of the Pacific, the Atlantic and the Indian Oceans) while larger errors occur in the areas such as northwest Pacific, the northwest Atlantic, and the Mediterranean Sea. Interestingly, the favorable regions are consistent with swell pools found by [56]. Figure 9c also shows the locations of poor (good) performance are in line with small (large) values of wave age mainly corresponding to the wind seas (swells).

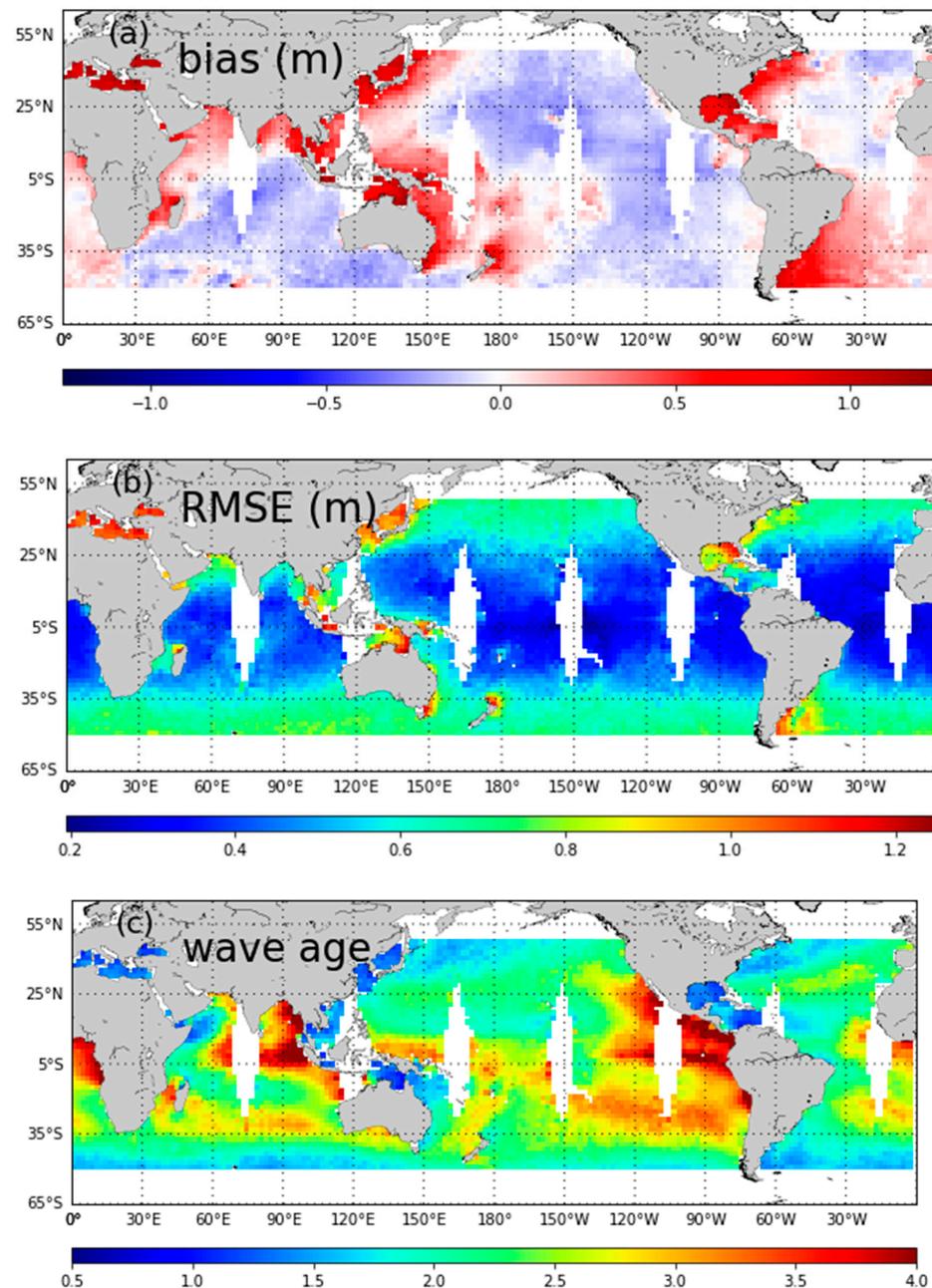


Figure 9. Map of (a) bias and (b) RMSE of deep learning ASCAT SWH estimates against WW3 hindcast and (c) wave age computed from WW3. All maps are on the grid of $2^\circ \times 2^\circ$.

Specifically, the regionality could be highlighted by two regions: Gulf of Mexico (97°W – 84°W , 21°N – 29°N , green box in Figure 2) and Hawaii Islands (164°W – 151°W , 16°N – 26°N , blue box in Figure 2), which are dominated by wind sea and the swell, respectively. In the Gulf of Mexico, statistic of bias (RMSE) is 0.80 m (0.88 m) and 0.76 m (0.84 m) against WW3 and buoys (#42001, 42002, 42019, and 42055), respectively. For the region of Hawaii Islands, a better performance could be achieved: bias = -0.2 m, RMSE = 0.53 m against ww3, and bias = -0.1 m, RMSE = 0.56 m against buoy (#51000, 51002, 51003, 51004, and 51101), respectively.

6.3. Incidence Angle Influence

As a fixed-beam scatterometer, incidence angle of ASCAT corresponds to the position of WVC cross satellite track. In order to investigate the dependent of error characteristics on

the incidence angle, we here employed the cross-track index (CTI, from 0 to 41) indicating the location of ASCAT WVC cross the swath.

Figure 10 presents the performances of the ASCAT SWH deep learning model with respect to the CTI or corresponding incidence angle of mid/side-beams, showing almost constant values in terms of all the metrics: scatter index (blue) and correlation coefficient (red) in Figure 10a, and bias and RMSE in Figure 10b.

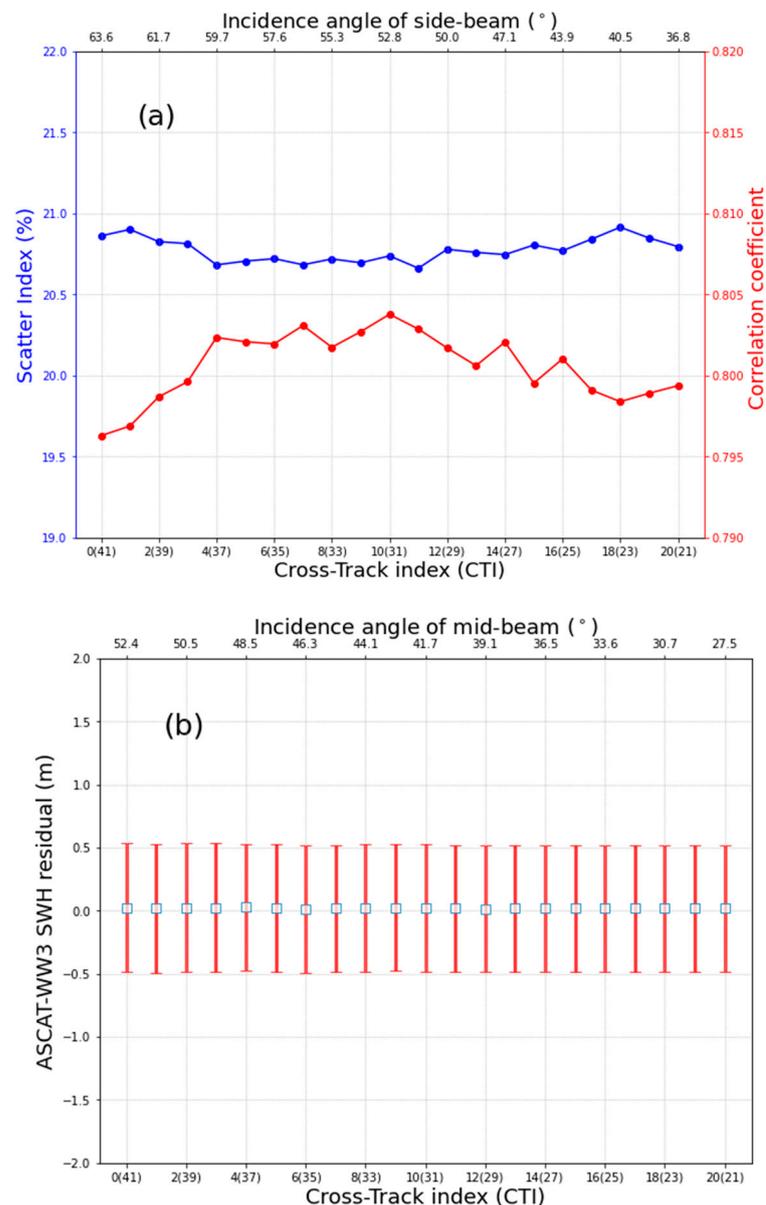


Figure 10. Performances of the ASCAT derived SWH against WW3 (test dataset) with respect to cross-track index and incidence angle of (a) side-beams and (b) mid-beam. (a) Scatter index (blue) and correlation coefficient (red). (b) Bias (box) and RMSE (error bar).

From previous studies, it was found that the sea state impact of wind speed residual decreases with increasing incidence angle for C-band scatterometer of ASCAT [11] and Sentinel-1 SAR [12]. Surprisingly, there is almost no evidence of any radar incidental dependency on SWH residual of our deep learning algorithm here.

Additionally, we also conducted the experiments against geophysical and instrument parameters like NRCS and K_p , and similarly found no apparent dependencies on SWH errors. These could be resulted from the incorporation of these parameters (for instance,

from low as 25° to high as 64° for incidence angle) into the model, and thereby bring adjustment into the SWH estimates due to the deeply learning.

7. Conclusions and Perspectives

In this paper, we have developed and implemented a deep learning neural network to estimate significant wave height from ASCAT scatterometer by training a big matchup database of satellite observation and wave model hindcast.

The proposed neural network was devised to include one input layer (ASCAT data), four hidden layers, and one output layer (SWH estimates). Thirteen variables derived from L1b/2 ASCAT products were taken as inputs: 4 sets of triplets (NRCS, incident angles, azimuthal angles, and variability factors) for three antennae, and the ASCAT derived wind speed. The coefficients of deep learning algorithm were determined by tuning of the neural network using 8+ million pairs of ASCAT-WW3 col-locations.

The statistical assessment against WW3 SWH hindcast (ranging from 0.5–5 m) shows the RMSE of 0.5 m, correlation coefficients of 0.80 and SI of 20%, respectively. The deep learning model tends to overestimate low wave heights and underestimate high wave heights, and larger errors are found in wind–sea dominated regions. Future research should focus on improving the proposed deep learning model under these unfavorable conditions.

Results indicate that the proposed data driven approach is reasonable for wave heights estimation directly from the scatterometers observations, although these remote sensors with frequent revisit and wide-swath coverage are routinely only adopted for monitoring ocean winds. Particularly, the SWH estimates from scatterometer through our model could provide denser spatio-temporal sampling than existing space-borne sensors so far.

The deep learning algorithm could also be trained for other scatterometers, especially recently launched sensors (e.g., HY-2B [51]). SWH derived from multi-scatterometers would be combined to produce maps of global wave climate and furthermore towards long-term climate records [57]. In addition, application to finer 12.5 km spacing ASCAT data, or even the ultra-high resolution scatterometer products [58] would provide wave field observation with fine spatial resolution suitable for coastal engineering, provided the sampling artefacts are tackled.

The proposed deep learning algorithm is encouraging for sea state estimation from wide swath scatterometers but still leave room for improvements. In the future, the external parameter of sea state maturity (wave age) could be used for developing correction while the position information (longitude and latitude) could be employed as QC flags for improving the retrieved SWH. In addition, further investigations of calibration and cross-validation against buoy in situ will also be dedicated.

Author Contributions: Conceptualization, H.W. and J.Y.; funding acquisition, J.Z.; investigation, H.W., L.R. and W.L.; methodology, H.W., J.Y. and L.R.; project administration, J.Y.; resources, J.Z.; Validation, Y.L., W.L. and C.C.; writing—original draft, H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2016YFC1401007.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. These include IOWAGA WaveWatch-III model hindcast from <ftp.ifremer.fr>, NDBC buoy observations from <http://www.ndbc.noaa.gov>, Metop-A/ASCAT L1 & L2 products from <https://archive.eumetsat.int> and <http://projects.knmi.nl/scatterometer>.

Acknowledgments: The authors are grateful for two anonymous reviewers for giving constructive comments and suggestions to help improve this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Queffeuilou, P. Long-term validation of wave height measurements from altimeters. *Mar. Geod.* **2004**, *27*, 495–510. [[CrossRef](#)]
2. Durrant, T.H.; Greenslade, D.J.M.; Simmonds, I. Validation of Jason-1 and Envisat remotely sensed wave heights. *J. Atmos. Ocean. Technol.* **2009**, *26*, 123–134. [[CrossRef](#)]
3. Chapron, B.; Johnsen, H.; Garello, R. Wave and wind retrieval from SAR images of the ocean. *Ann. Telecommun.* **2001**, *56*, 682–699.
4. Schulz-Stellenfleth, J.; Konig, T.; Lehner, S. An empirical approach for the retrieval of integral ocean wave parameters from synthetic aperture radar data. *J. Geophys. Res.* **2007**, *112*, 1–14. [[CrossRef](#)]
5. Li, X.M.; Lehner, S.; Bruns, T. Ocean wave integral parameter measurements using Envisat ASAR wave mode data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 155–174. [[CrossRef](#)]
6. Stopa, J.E.; Mouche, A.A. Significant wave heights from Sentinel-1 SAR: Validation and applications. *J. Geophys. Res.* **2017**, *122*, 1827–1848. [[CrossRef](#)]
7. Wang, H.; Wang, J.; Yang, J.S.; Ren, L.; Zhu, J.H.; Yuan, X.Z.; Xie, C.H. Empirical algorithm for significant wave height retrieval from wave mode data provided by the Chinese satellite Gaofen-3. *Remote Sens.* **2018**, *10*, 363. [[CrossRef](#)]
8. Vogelzang, J.; Stoffelen, A. Scatterometer wind vector products for application in meteorology and oceanography. *J. Sea Res.* **2012**, *74*, 16–25. [[CrossRef](#)]
9. Karaev, V.; Panfilova, M.; Guo, J. Influence of the type of sea waves on the backscattered radar cross section at medium incidence angles. *Izv. Atmos. Ocean Phys.* **2016**, *52*, 904–910. [[CrossRef](#)]
10. Quilfen, Y.; Chapron, B.; Collard, F.; Vandemark, D. Relationship between ERS scatterometer measurement and integrated wind and wave parameters. *J. Atmos. Ocean. Technol.* **2004**, *21*, 368–373. [[CrossRef](#)]
11. Stopa, J.E.; Mouche, A.A.; Chapron, B.; Collard, F. Sea state impacts on wind speed retrievals from C-band radars. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2017**, *10*, 2147–2155. [[CrossRef](#)]
12. Li, H.; Mouche, A.A.; Stopa, J.E. Impact of sea state on wind retrieval from Sentinel-1 wave mode data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 559–566. [[CrossRef](#)]
13. Chen, S.; Qiao, F.; Jiang, W.; Guo, J.; Dai, D. Impact of surface waves on wind stress under low to moderate wind conditions. *J. Phys. Oceanogr.* **2019**, *49*, 2017–2028. [[CrossRef](#)]
14. Vincent, C.L.; Thomson, J.; Graber, H.C.; Collins, C.O. Impact of swell on the wind-sea and resulting modulation of stress. *Prog. Oceanogr.* **2019**, *178*, 102164. [[CrossRef](#)]
15. Hersbach, H. Comparison of C-band scatterometer CMOD5.N equivalent neutral winds with ECMWF. *J. Atmos. Ocean. Technol.* **2010**, *27*, 721–736. [[CrossRef](#)]
16. Stoffelen, A.; Verspeek, J.; Vogelzang, J.; Verhoef, A. The CMOD7 geophysical model function for ASCAT and ERS wind retrievals. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2123–2134. [[CrossRef](#)]
17. Guo, J.; He, Y.; Perrie, W.; Shen, H.; Chu, X. A new model to estimate significant wave heights with ERS-1/2 scatterometer data. *Chin. J. Oceanol. Limnol.* **2009**, *27*, 112–116. [[CrossRef](#)]
18. Guo, J.; He, Y.; Chu, X.; Cui, L.; Liu, G. Wave parameters retrieved from QuikSCAT data. *Can. J. Remote Sens.* **2009**, *35*, 345–351. [[CrossRef](#)]
19. Shao, W.; Zhu, S.; Zhang, X.; Gou, S.; Jiao, C.; Yuan, X.; Zhao, L. Intelligent wind retrieval from Chinese gaofen-3 SAR imagery in quad polarization. *J. Atmos. Ocean. Technol.* **2019**, *36*, 2121–2138. [[CrossRef](#)]
20. Li, X.-M.; Qin, T.; Wu, K. Retrieval of sea surface wind speed from spaceborne SAR over the Arctic Marginal Ice Zone with a neural network. *Remote Sens.* **2020**, *12*, 3291. [[CrossRef](#)]
21. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [[CrossRef](#)]
22. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [[CrossRef](#)] [[PubMed](#)]
23. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* **2020**, *241*, 111716. [[CrossRef](#)]
24. European Organization for the Exploitation of Meteorological Satellites. ASCAT Product Guide. 2015. Available online: www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_DMT_151558&RevisionSelectionMethod=LatestReleased&Rendition=Web (accessed on 29 October 2020).
25. Anderson, C.; Bonekamp, H.; Duff, C.; Figa-Saldana, J.; Wilson, J.J.W. Analysis of ASCAT ocean backscatter measurement noise. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 2449–2457. [[CrossRef](#)]
26. Lin, W.M.; Portabella, M.; Stoffelen, A.; Vogelzang, J.; Verhoef, A. ASCAT wind quality under high subcell wind variability conditions. *J. Geophys. Res.* **2015**, *120*, 5804–5819. [[CrossRef](#)]
27. Ardhuin, F.; Rogers, E.; Babanin, A.V.; Filipot, J.-F.; Magne, R.; Roland, A.; van der Westhuysen, A.; Queffeuilou, P.; Lefevre, J.-M.; Aouf, L.; et al. Semiempirical dissipation source functions for ocean waves. Part I: Definition, calibration, and validation. *J. Phys. Oceanogr.* **2010**, *40*, 1917–1941. [[CrossRef](#)]
28. Rasle, N.; Ardhuin, F. A global wave parameter database for geophysical applications. Part 2: Model validation with improved source term parameterization. *Ocean Model.* **2013**, *70*, 174–188. [[CrossRef](#)]

29. Abdalla, S.; Isaksen, L.; Janssen, P.; Wedi, N. Effective spectral resolution of ECMWF atmospheric forecast models. *ECMWF News Lett.* **2013**, *137*, 19–22. Available online: <https://www.ecmwf.int/sites/default/files/elibrary/2013/17358-effective-spectral-resolution-ecmwf-atmospheric-forecast-models.pdf> (accessed on 29 October 2020).
30. Lin, W.; Portabella, M.; Stoffelen, A.; Verhoef, A.; Turiel, A. ASCAT wind quality control near rain. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4165–4177. [[CrossRef](#)]
31. Liu, X.; Zheng, Q.; Liu, R.; Wang, D.; Duncan, J.H.; Huang, S.-J. A study of radar backscattering from water surface in response to rainfall. *J. Geophys. Res.* **2016**, *121*, 1546–1562. [[CrossRef](#)]
32. Sungmin, O.; Foelsche, U.; Kirchengast, G.; Fuchsberger, J.; Tan, J.; Petersen, W.A. Evaluation of GPM IMERG early, late, and final rainfall estimates using WegenerNet gauge data in Southeastern Austria. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6559–6572.
33. Fairall, C.W.; Bradley, E.F.; Hare, J.E.; Grachev, A.A.; Edson, J.B. Bulk Parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. *J. Clim.* **2003**, *16*, 571–591. [[CrossRef](#)]
34. Pierson, W.J.; Moskowitz, L. A proposed spectral form for fully developed wind seas based on the similarity theory of S.A. Kitaigorodskii. *J. Geophys. Res.* **1964**, *69*, 5181–5190. [[CrossRef](#)]
35. Elfouhaily, T.; Chapron, B.; Katsaros, K.; Vandemark, D. A unified directional spectrum for long and short wind-driven waves. *J. Geophys. Res.* **1997**, *102*, 15781–15796. [[CrossRef](#)]
36. Verspeek, J.; Stoffelen, A.; Portabella, M.; Bonekamp, H.; Anderson, C.; Saldaña, J.F. Validation and calibration of ASCAT using CMOD5.n. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 386–395. [[CrossRef](#)]
37. Wang, Z.; Stoffelen, A.; Zou, J.; Lin, W.; Verhoef, A.; Zhang, Y.; He, Y.; Lin, M. Validation of new sea surface wind products from Scatterometers Onboard the HY-2B and MetOp-C satellites. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 4387–4394. [[CrossRef](#)]
38. Glazman, R.E. Wind-fetch dependence of Seasat scatterometer measurements. *Int. J. Remote Sens.* **1987**, *8*, 1641–1647. [[CrossRef](#)]
39. Glazman, R.E.; Pilorz, S.H. Effects of sea maturity on satellite altimeter measurements. *J. Geophys. Res.* **1990**, *95*, 2857–2870. [[CrossRef](#)]
40. Karaev, V.Y.; Kanevsky, M.B.; Balandina, G.N.; Cotton, P.D.; Challenor, P.G.; Gommenginger, C.P.; Srokosz, M.A. On the problem of the near ocean surface wind speed retrieval by radar altimeter: Two-parameter algorithm. *Int. J. Remote Sens.* **2002**, *23*, 3263–3283. [[CrossRef](#)]
41. Gourrion, J.; Vandemark, D.; Bailey, S.; Chapron, B.; Gommenginger, G.P.; Challenor, P.G.; Srokosz, M.A. A two-parameter wind speed algorithm for Ku-band altimeters. *J. Atmos. Ocean. Technol.* **2002**, *19*, 2030–2048. [[CrossRef](#)]
42. Soisuvann, S.; Jelenak, Z.; Chang, P.S.; Alsweiss, S.O.; Zhu, Q. CMOD5.H—A high wind geophysical model function for C-Band vertically polarized satellite scatterometer measurements. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3744–3760. [[CrossRef](#)]
43. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
44. Misra, D. Mish: A self-regularized non-monotonic neural activation function. *arXiv* **2019**, arXiv:1908.08681.
45. Dozat, T. Incorporating Nesterov momentum into Adam. In Proceedings of the 4th International Conference on Learning Representations, Workshop Track, San Juan, PR, USA, 2–4 May 2016.
46. Sutskever, I.; Martens, J.; Dahl, G.E.; Hinton, G.E. On the importance of initialization and momentum in deep learning. *Proc. Mach. Learn. Res.* **2013**, *28*, 1139–1147.
47. LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1989**, *1*, 541–551. [[CrossRef](#)]
48. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
49. Chollet, F. Keras. 2015. Available online: <https://github.com/keras-team/keras> (accessed on 29 October 2020).
50. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. TensorFlow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
51. Wang, H.; Zhu, J.; Lin, M.; Zhang, Y.; Chang, Y. Evaluating Chinese HY-2B HSCAT ocean wind products using buoys and other scatterometers. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 923–927. [[CrossRef](#)]
52. Portabella, M.; Stoffelen, A. On scatterometer ocean stress. *J. Atmos. Ocean. Technol.* **2009**, *26*, 368–382. [[CrossRef](#)]
53. Janssen, P.A.E.M.; Abdalla, S.; Hersbach, H.; Bidlot, J.-R. Error estimation of buoy, satellite, and model wave height data. *J. Atmos. Ocean. Technol.* **2007**, *24*, 1665–1677. [[CrossRef](#)]
54. Li, X.-M.; Huang, B. A global sea state dataset from spaceborne synthetic aperture radar wave mode data. *Sci. Data* **2020**, *7*. [[CrossRef](#)]
55. Karaev, V.Y.; Meshkov, E.M.; Chu, X. Features of sea-wave classification in problems of remote sensing. *Izv. Atmos. Ocean. Phys.* **2013**, *49*, 919–929. [[CrossRef](#)]
56. Chen, G.; Chapron, B.; Ezraty, R.; Vandemark, D. A global view of swell and wind sea climate in the ocean by satellite altimeter and scatterometer. *J. Atmos. Ocean. Technol.* **2002**, *19*, 1849–1859. [[CrossRef](#)]
57. Verhoef, A.; Vogelzang, J.; Verspeek, J.; Stoffelen, A. Long-term scatterometer wind climate data records. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2186–2194. [[CrossRef](#)]
58. Hutchings, N.; Kilpatrick, T.; Long, D.G. Ultrahigh resolution scatterometer winds near Hawaii. *Remote Sens.* **2020**, *12*, 564. [[CrossRef](#)]