Revisit of the direct assimilation of scatterometer 'sigma0'

Technical report on

The development of the ANN ASCAT σ^0 Observation Operator (WP 1)

Saleh Abdalla, Sean Healy, Giovanna De Chiara, and Matthew Chantry

European Centre for Medium-Range Weather Forecasts (ECMWF)
Shinfield Park, Reading, Berkshire, RG2 9AX, UK

Initial draft: July 2023
Final draft: November 2023

Table of Contents

1 Background

Precise knowledge of ocean surface roughness is important for a large range of applications, such as the computation of air/sea heat fluxes, assimilation in Numerical Weather Prediction (NWP) systems, for forcing ocean models and surface-wave models, as well as for climate studies.

At ECMWF scatterometer winds have been assimilated into the Integrated Forecasting System (IFS) since 1996, beginning with ERS-1 and ERS-2 Scatterometer data. Currently Metop-B ASCAT [5][15] and Metop-C ASCAT wind products are assimilated together with HY-2B observations. The importance of scatterometer observations and their positive impact on NWP and wave models has long been recognized [7][17][18]. In particular, C-band scatterometers, like the ones onboard the ERS and the Metop satellite series, provide information in almost all-weather conditions. These observations are therefore important, often unique, for the analysis of extreme events (usually characterized by cloud and rainfall) such as tropical cyclones (TCs) and extra-tropical storms. It has also been confirmed that ASCAT observations have a substantial impact on the ocean variables in coupled ocean-atmosphere assimilation systems [8][20].

The initial study 'Characterisation of ocean surface roughness in NWP' was performed by ECMWF under contract EUM/CO/18/4600002207/Sli (see [6]). This project was aimed at exploring how to increase the value of scatterometer observations in NWP, by assessing the assimilation of new geophysical variables closer to what the scatterometer actually measures over the ocean. The assimilation of surface stress and stress equivalent winds were tested. But the experiments performed with these new geophysical variables did not show any clear benefit on the quality of the analysis and forecast.

One difficulty with the stress assimilation may be related to the mapping of the ambiguous winds vectors provided by the geophysical model function (GMF) (currently the CMOD5.n [16] is used in operations) to stress space. This additional processing step is unlikely to change the real information content of the observation. An alternative approach to be further investigated is the direct assimilation of backscatter coefficient "sigma0" ($\boldsymbol{\sigma}^0$; also known as the normalized radar cross section, NRCS) triplets, thereby removing the GMF from the observation processing chain.

Several efforts were dedicated to retrieve wind vector from $\boldsymbol{\sigma}^0$ (e.g., [16] [19] [21]) and use the retrieved wind vector in data assimilation (e.g. [17] [18] [21]). The direct assimilation of $\boldsymbol{\sigma}^0$ was discussed in the 1990's and was largely discarded at the time (e.g., [21]), irrespective of the promising results that reported otherwise (e.g. [19] [23]). The two main reasons for not pursuing the direct assimilation of $\boldsymbol{\sigma}^0$ appear to be handling the wind direction ambiguity of the wind - $\boldsymbol{\sigma}^0$ relationship, and the non-linearity associated with this transform. Subsequently, there appears to have been very little work on the direct assimilation of scatterometer $\boldsymbol{\sigma}^0$ values. In fact, the GMF tangent-linear and adjoint codes required to assimilate $\boldsymbol{\sigma}^0$ are not currently available.

It is time to revisit these results. The 4D-Var assimilation approach naturally uses prior information to constrain ill-posed retrieval problems. Ill-posed problems are ambiguous by definition, because multiple atmospheric states can produce the same measured value. In addition, recent progress tackling non-linear problems in the "All Sky" assimilation of radiances [12] might suggest that some non-linearities considered problematic in the 1990's might now be overcome. However, this suggestion needs to be tested.

Preliminary results obtained at ECMWF with a $\boldsymbol{\sigma}^0$ forward operator based on a neural network have been presented with promising results during the final meeting of the initial study and during the ECMWF-EUMETSAT Science bilateral, on 26/04/2022. ECMWF has trained an artificial neural network (ANN) to forward model ASCAT-B scatterometer $\boldsymbol{\sigma}^0$ triplets. This work used ECMWF 10-m neutral equivalent wind vectors and measurement geometry parameters ("features" in ANN terminology) to simulate operational

ASCAT-B $\boldsymbol{\sigma}^0$ triplet values ("targets"), using "trusted" measurements that have passed operational quality control and have been assimilated. The work was performed using the Keras/TensorFlow [3][1] software packages (in python). The early results are encouraging, and equally importantly the ANN approach seems to be very flexible, so that the impact of other variables on the fit to $\boldsymbol{\sigma}^0$ can also be tested. This could include atmospheric air density, sea-state information, rain-rate.

We note that an ANN approach was used more than 20 years ago to develop scatterometer forward models [4]. Furthermore, ANN is also being tested for scatterometer soil moisture applications [2].

2 Used Data

ASCAT backscatter data were retrieved from the ECMWF operational Observational Data Base (ODB) archived on ECMWF Meteorological Archival and Retrieval System (MARS) [14]. Only observations that were assimilated actively are used for training since training requires the best available data. Assimilated observations are usually of good quality. Scatterometer data including backscatter observations with their corresponding geometry (incidence and azimuthal angles) and model wind vectors from both background (first-guess, FG) and analysis (AN) and other related information are stored in ODB during data assimilation. Observations from ASCAT-B were mainly used for training purposes. Validation was carried out using observations from both ASCAT-B (but during different time windows) and ASCAT-C.

Several periods (mainly from 2020 and 2021) were used depending on the availability of the data. The final training, however, was based on ASCAT-B observations covering the whole year spanning from 1 August 2020 till 31 July 2021.

For other variables, IFS short forecasts (FC) from step 1 till step 12 (hours) from the ECMWF operational suite were used. Fields of those variable were retrieved from MARS and were collocated with the ASCAT data from ODB. Closest grid point and time were selected to reduce the impact of interpolation. Both atmospheric and ocean-wave data were used to examine their impact on the learning ANN process.

3 Artificial Neural Network (ANN) Model

An Artificial Neural Network (ANN) model of ASCAT backscatter ($\boldsymbol{\sigma}^0$) has been developed using Keras/TensorFlow ([3][1]) package. Model wind vector (in addition to other model variables that will be specified later) and the geometry of the scatterometer beam with respect to the wind vector are the input variables ("features"). ASCAT $\boldsymbol{\sigma}^0$ is the ANN model output "target".

ANN training is equivalent to the traditional CMOD training (e.g. [16]) but with more flexibility and without the need of the precise knowledge regarding the underlying physics. Once enough existing features-target data are available, the ANN can extract the relation (~ physical laws) between the pair. In fact, ANN only figures out the "pattern" in the data. Therefore, the found "relation" cannot be written in a closed form. However, ANN model parameters are produced and stored for later usage for verification and implementation of the model.

The following ANN configuration is used:

- A Sequential ANN model is used since it is most appropriate for a plain stack of layers where each layer has exactly one input tensor and one output tensor (https://keras.io/guides/sequential_model/).
- The input layer is composed of, at least, four input nodes. There is one node for each wind vector component, one node for each of the scatterometer incidence and azimuthal angles. Other nodes are added for each additional physical variable to be used.
- Two identical hidden dense layers are implemented. The "dense layer" is a term used to describe a layer of nodes with each node receiving input from all the nodes of the previous layer.
- Each layer is composed of 64 nodes.
- The activation function transforms the input to the output (also called activation) at the node level. Linear activation functions are easy to implement and execute. However, to learn complex mappings, nonlinear activation functions are required. Two activation functions were tested:
    - The rectified linear (ReLU) activation function which is a piecewise linear function that will output the input directly if its argument is positive, otherwise, it will output zero. Its derivative is 1 for positive arguments and zero otherwise. These characteristics make ReLU the favoured activation function for many types of neural networks.
    - The hyperbolic tangent (tanh) activation function, which is an S-shaped function symmetric around 0 and saturates at the values of 1 and -1 for high positive and negative values, respectively. This saturation for high argument values is a challenge for the learning process. tanh activation function is much less efficient compared to the ReLU function. However, this function is the recommended one for data assimilation.
- The output layer is composed of a single node representing $\boldsymbol{\sigma}^0$.
- The loss function, which is the equivalent term for the cost function in data assimilation, is the sum of the errors to be minimised. ANN model training is achieved through an optimisation of the loss function. The loss is represented by the mean square error (MSE), which is the average of the squared differences between the predicted and actual values.
- The optimisation algorithm (optimiser) is selected as the stochastic gradient descent (SGD) which is a very efficient optimiser.

Summary of ANN model configuration is listed in Table 1.

Several deviations from the above configurations were considered. Networks that include 2 hidden layers with 16, 32, 64 and 128 nodes as well as 3 layers with 64 and 32 nodes were tested. The impact was marginal and, therefore, for the sake of efficiency and accuracy we decided to use the above configuration in terms of layers and nodes.

Table 1: Summary of ANN model configuration.

Model Elements	Description
Model type	Sequential
Number of layers	1 input, 1 output and 2 hidden (3 hidden layers configuration was tested)
Type of hidden layers	Dense
Number of nodes	64 for each hidden layer (other number of nodes were tested)
Activation function	The following were used separately: - rectified linear (ReLU) - hyperbolic tangent (tanh)
Input layer	At least 4 nodes: - 2 nodes for model wind vector; - 2 nodes for scatterometer geometry; and - 1 or more nodes for each additional variable (if used)
Output layer	1 node for $\boldsymbol{\sigma}^0$
Loss function	Optimisation of mean square error (MSE)
Optimiser	Stochastic gradient descent (SGD)

4 ANN Model Training and Validation

4.1 ANN Observation Operator Based on Wind Vector

ASCAT-B observations collocated with the IFS model wind vectors were used for the ANN training. The final training was based on data covering the whole year spanning from 1 August 2020 till 31 July 2021. First, only IFS model first guess (FG) wind vector was used for training and validation. The trained ANN model (ANN_{FG}) was used to estimate $\boldsymbol{\sigma}^0$ values corresponding to IFS model FG wind vector data covering the 9-month period from 1 August 2021 to 30 April 2022. Later this period was extended till the end of March 2023, i.e. a total of 20 months and the change in the statistics was marginal.

The scatter plot showing the correlation between validation results of $\boldsymbol{\sigma}^0$ values estimated using ANN_{FG} and FG winds and those measured by ASCAT-B for the 9-month validation period are shown in Figure 1. With most of the data pairs being narrowly scattered around the symmetric 45-degree line, the agreement is rather excellent. However, a small number of collocations show higher scatter. Another important observation in Figure 1 is the inability of ANN_{FG} model to produce $\boldsymbol{\sigma}^0$ values below -31.5 dB. Such low values are associated with very light wind. This is in fact in agreement with available experimental evidence (e.g. [10]) that winds with speeds below a certain threshold (below around 2.0 m/s) is unable to produce any perturbation to the water surface. Therefore, any perturbations, which allow the radar signal to be scattered back to the scatterometer antennas, under very light winds are due to other variables like ocean swell. This is also clear in the variation of $\boldsymbol{\sigma}^0$ with respect to the wind speed shown in Figure 2.

Time series of the monthly standard deviation of the difference (SDD), mean absolute error (MAE), bias and correlation coefficient (C.C.) between $\mathrm{ANN_{FG}}$ estimated $\boldsymbol{\sigma}^0$ (based on FG wind input) and measured $\boldsymbol{\sigma}^0$ by ASCAT-B and ASCAT-C are shown in Figure 3. The monthly SDD, MAE and C.C. values are virtually constant for the whole training (in case of ASCAT-B) and validation periods (August 2020 till March 2023). The results from ASCAT-B and ASCAT-C are also the same with very minor advantage for ASCAT-C (smaller SDD values and higher C.C. values during few months). The differences are ignorable. The validation against the ASCAT-B data set that was used for training does not give any advantage for that data set. The minor advantage of ASCAT-C results is still valid.

However, $\boldsymbol{\sigma}^0$ estimates by $\mathrm{ANN_{FG}}$ using FG wind show a small bias as can be seen from Figure 3. Bias with respect to ASCAT-B $\boldsymbol{\sigma}^0$ is zero during the training period, as expected. However, the bias shows an increasing trend afterwards. $\boldsymbol{\sigma}^0$ bias with respect to ASCAT-C shows a rather constant shift of about 0.07 dB compared to bias against ASCAT-B. Otherwise, the bias with respect to ASCAT-C follows the same pattern as that of ASCAT-B. This implies that ASCAT-C $\boldsymbol{\sigma}^0$ is higher than that of ASCAT-B. The increasing trend could be coming from the scatterometers or from the IFS model forecast.

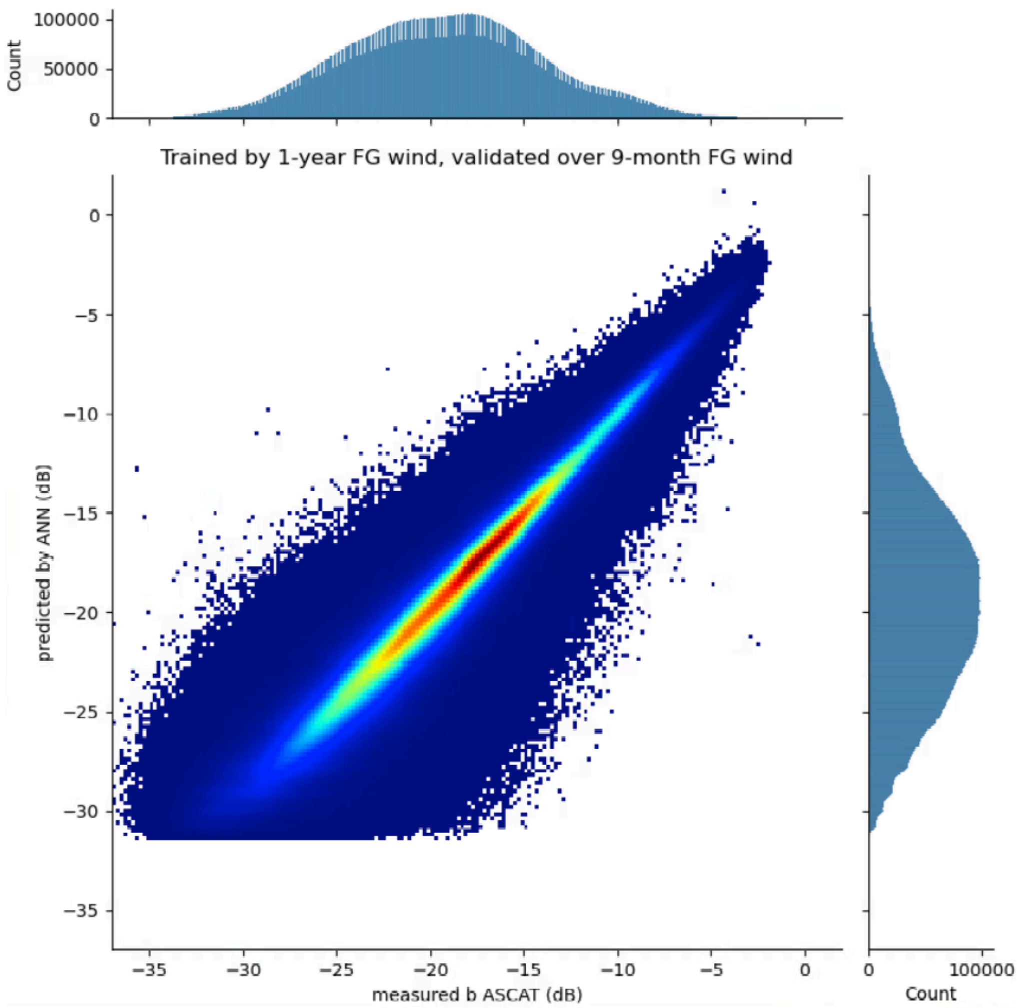

Figure 1: Validation of 9-month (1 August 2021 – 30 April 2022) of the $\mathrm{ANN_{FG}}$ model $\boldsymbol{\sigma}^0$ estimates using IFS model FG wind vectors against measured $\boldsymbol{\sigma}^0$ by ASCAT-B.

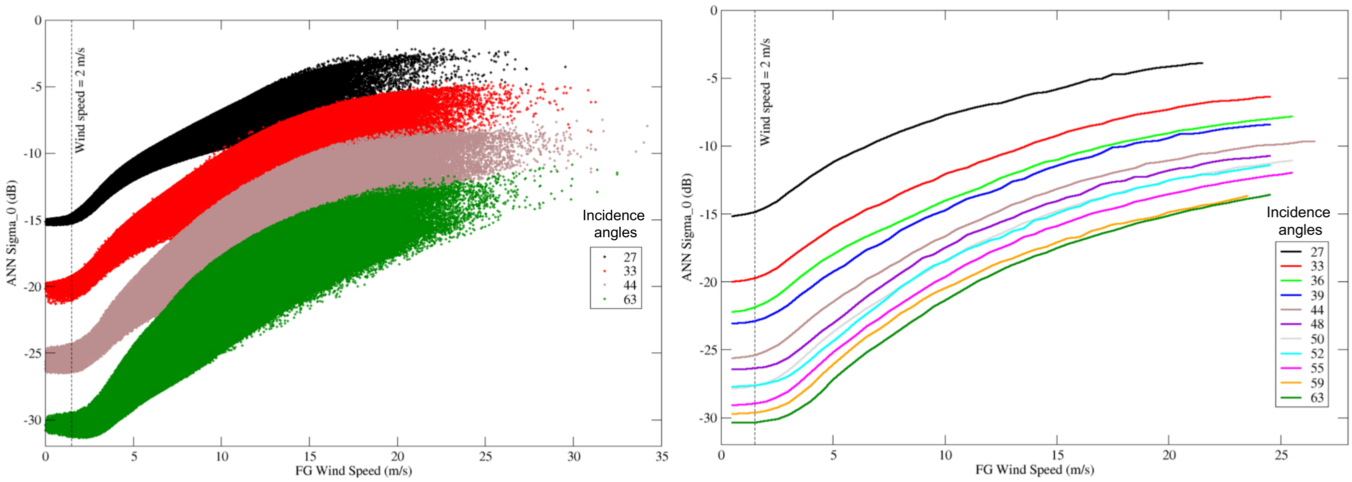

Figure 2: Relation between ANN sigma_0 and wind speed for different incidence angles (training using FG wind only). The left-hand panel shows the data clouds for 4 incidence angles. The right-hand panel shows the centreline of the data clouds of 11 incidence angles.

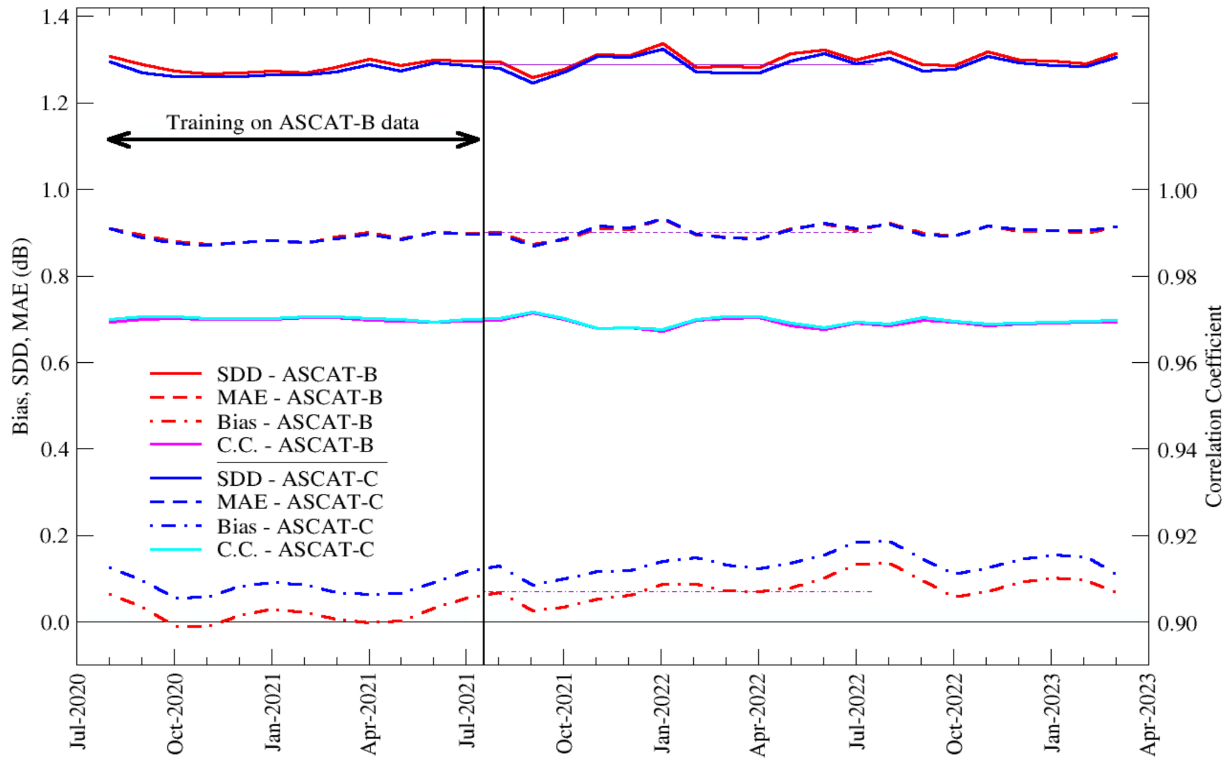

Figure 3: Time series of monthly standard deviation of the difference (SDD), mean absolute error (MAE), bias and correlation coefficient (C.C.) between ANN estimated $\boldsymbol{\sigma}^0$ *(based on FG wind) and measured* $\boldsymbol{\sigma}^0$ *by ASCAT-B and ASCAT-C. Training was carried out using ASCAT-B data during the period shown by the two-sided arrow.*

$\boldsymbol{\sigma}^0$ bias (between ANN_{FG} estimates and the scatterometers) in Figure 3 shows a seasonal variation with peaks in January and August and troughs in April and October.

The geographical distribution of bias of ANN_{FG} model $\boldsymbol{\sigma}^0$ estimates using IFS model FG wind vectors with respect to the measured $\boldsymbol{\sigma}^0$ by ASCAT-B during the 1-year period from 1 August 2021 to 31 July 2022 is shown in Figure 4. The small bias (below 0.1 dB) dominates the map. However, there are higher bias values in the Tropics. It is possible to relate high biases with major ocean currents like: the Pacific South Equatorial Current, the Pacific Equatorial Counter-current, the Guinea Current, Kuroshio Current, Agulhas Current and, to less extent, the Gulf Stream.

The geographical distribution of SDD and scatter index (SI; which is the SDD normalised by the mean value) of ANN_{FG} model $\boldsymbol{\sigma}^0$ estimates using IFS model FG wind vectors with respect to the measured $\boldsymbol{\sigma}^0$ by ASCAT-B during the 1-year period from 1 August 2021 to 31 July 2022 are shown in Figure 5 and Figure 6, respectively. The SDD is predominantly below 1.5 dB. That corresponds to SI value of 8%. There are, however, some areas with high SDD and SI values like around the Indonesian Islands and the Pacific coast of Central America.

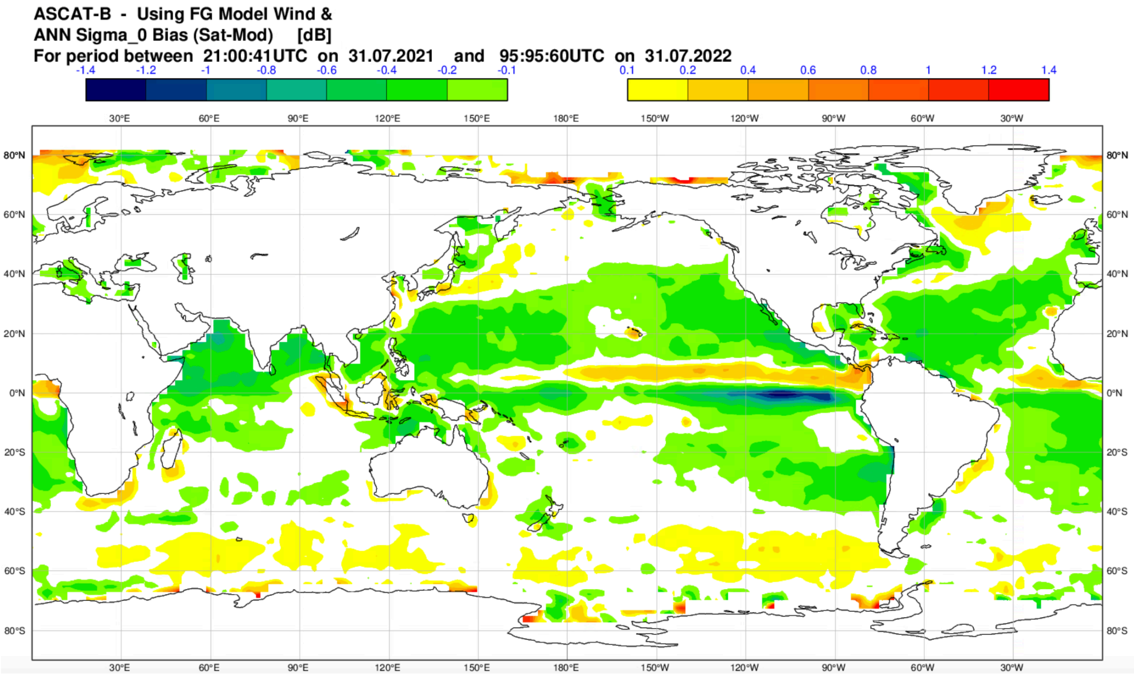

Figure 4: The geographical distribution of bias between ANN_{FG} model $\boldsymbol{\sigma}^0$ estimates using IFS model FG wind vectors and the measured $\boldsymbol{\sigma}^0$ by ASCAT-B during the 1-year period from 1 August 2021 to 31 July 2022. The unit is dB.

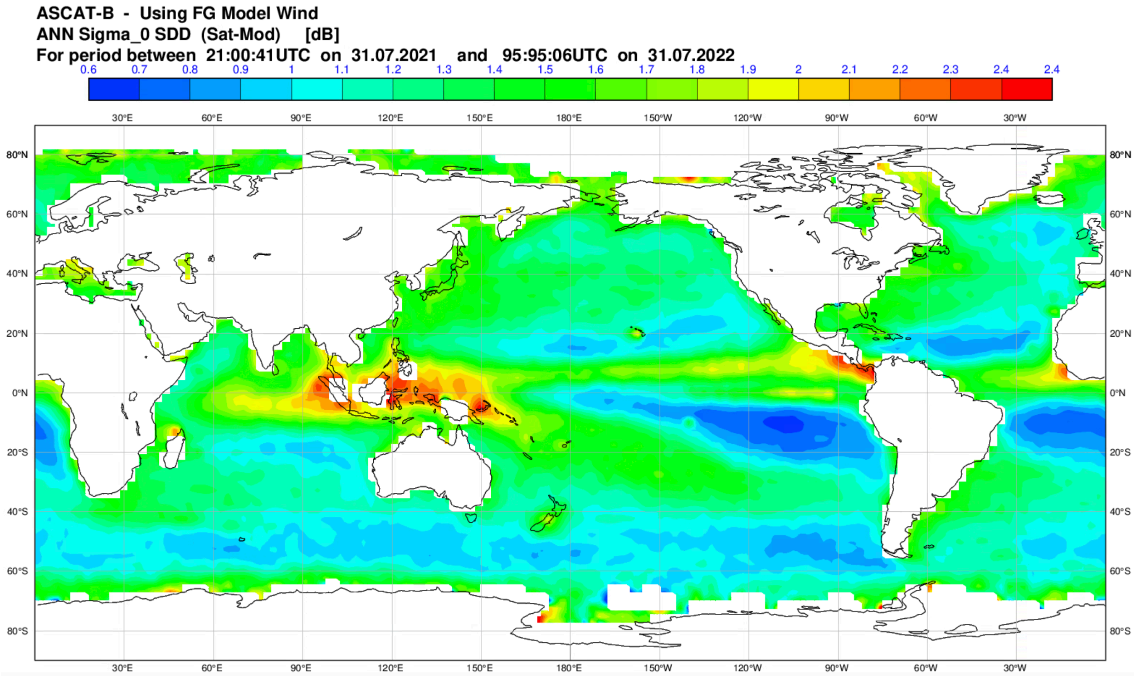

Figure 5: Same as for Figure 4 but for the standard deviation of the difference (SDD). The unit is dB.

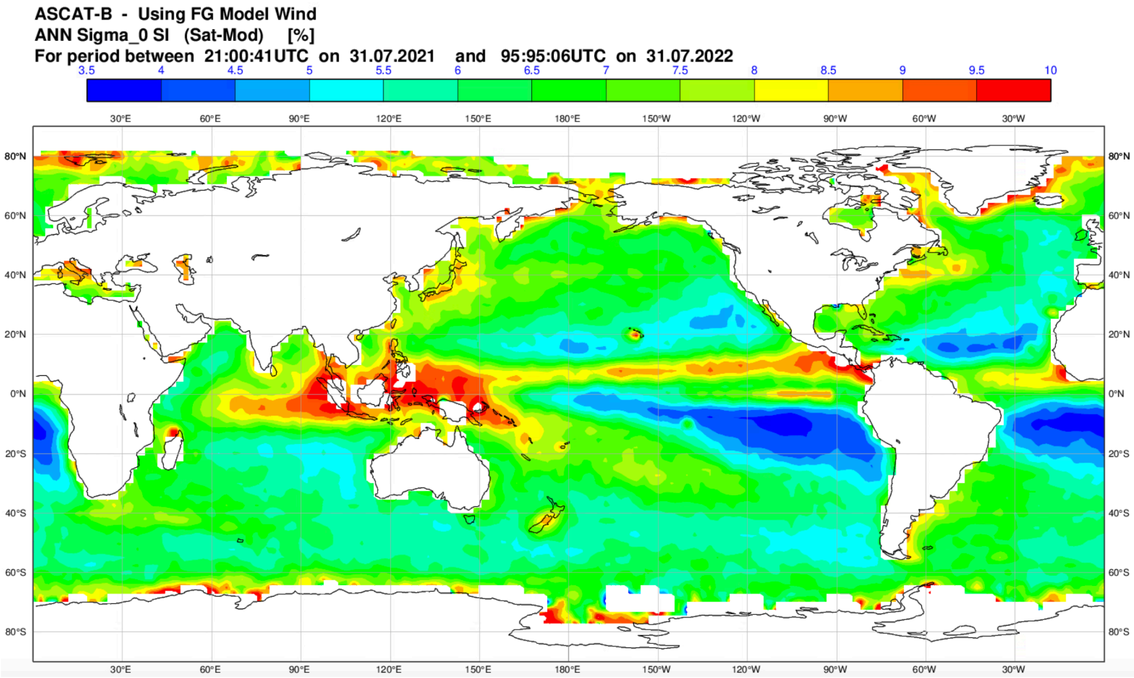

Figure 6: Same as for Figure 4 but for the scatter index (SI), which is the SDD normalised by the mean value. The unit is "percentage".

4.2 Use of Analysis versus First-Guess Wind Vector

Model analysis (AN) winds are in theory the best winds available from a numerical weather prediction (NWP) system like the ECMWF IFS. These AN model winds are the result of data assimilation (i.e. merging observations with the model) process. Currently, IFS assimilates ASCAT-B and ASCAT-C scatterometer winds after inversion. Therefore, the use of AN winds is usually not preferred to be used for training. This is the main reason behind starting with the use of the model first-guess (FG) winds, which are the results of model short forecasts that are needed to represent the model state (background) in the data assimilation system. Model FG data are not impacted by the observations within the same time window They are, however, impacted by data from previous time windows. This is not a real issue unless there is a "systematic error".

The ANN training that was done so far was based on model FG wind (ANN_{FG}) and can be represented schematically using Figure 7. Alternatively, one can FG wind in training by AN wind (ANN_{AN}) as shown in Figure 8. For the validation, it is possible to estimate $\boldsymbol{\sigma}^0$ values using either ANN model (ANN_{FG} or ANN_{AN}) and either wind type: FG wind as shown in Figure 9 or AN wind as shown in Figure 10. The statistics from all four combinations are listed in Table 2 for FG winds and in Table 3 for AN winds.

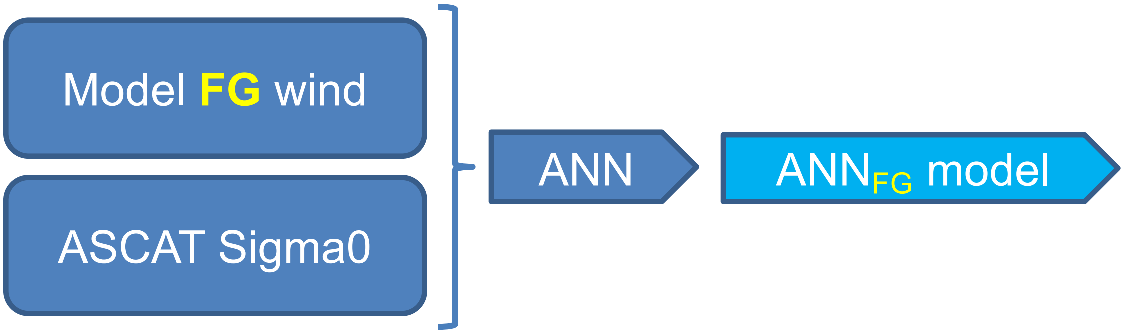

Figure 7: Schematic diagram showing the ANN_{FG} training.

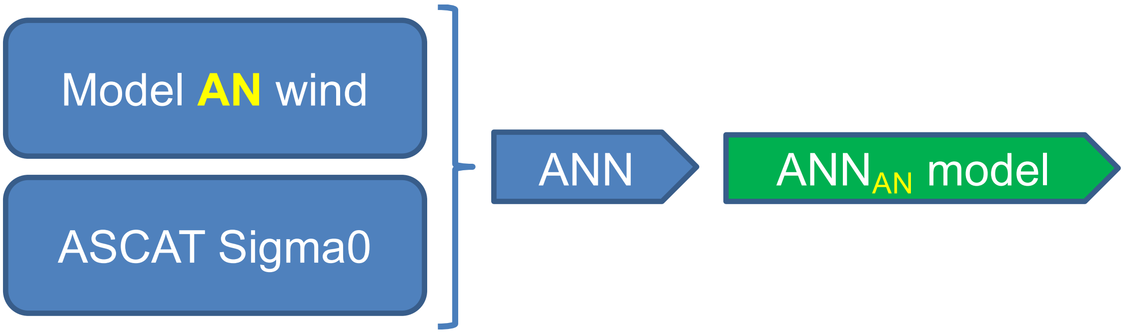

Figure 8: Schematic diagram showing the ANN_{AN} training.

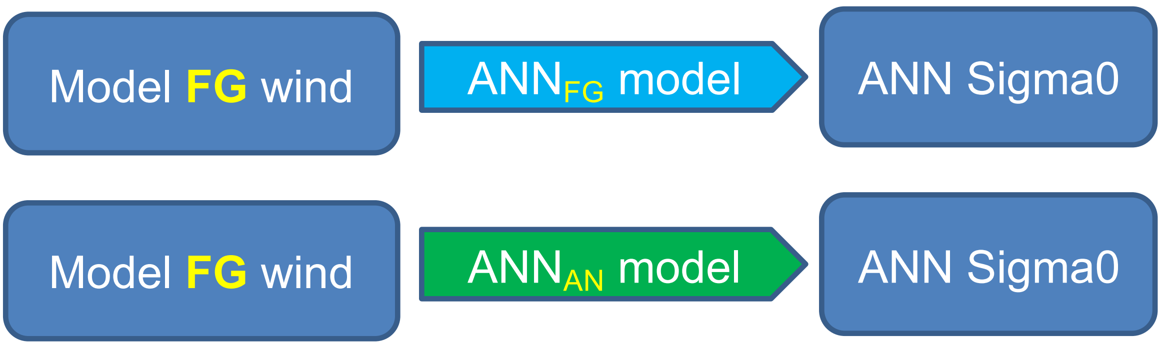

Figure 9: Schematic diagram showing the use of model FG wind to validate both ANN trained models; namely: ANN_{FG} and ANN_{AN}.

It is clear from Table 2 and Table 3 that both models perform equally well for any given wind type. In fact, the wind type used in the $\boldsymbol{\sigma}^0$ estimation dictates the performance. AN winds produce better $\boldsymbol{\sigma}^0$ values than FG winds irrespective of wind used in the model training. However, ANN_{AN} model produces slightly better $\boldsymbol{\sigma}^0$ values when used with AN wind.

Table 2: Verification of ANN models trained using FG (ANN_{FG}) and AN (ANN_{AN}) winds when used with FG winds.

N = 1266843	ANN model trained using FG wind (ANN_{FG})	ANN model trained using AN wind (ANN_{AN})
Bias	0.023	0.048
SDD	1.314	1.316
R	0.9682	0.9681

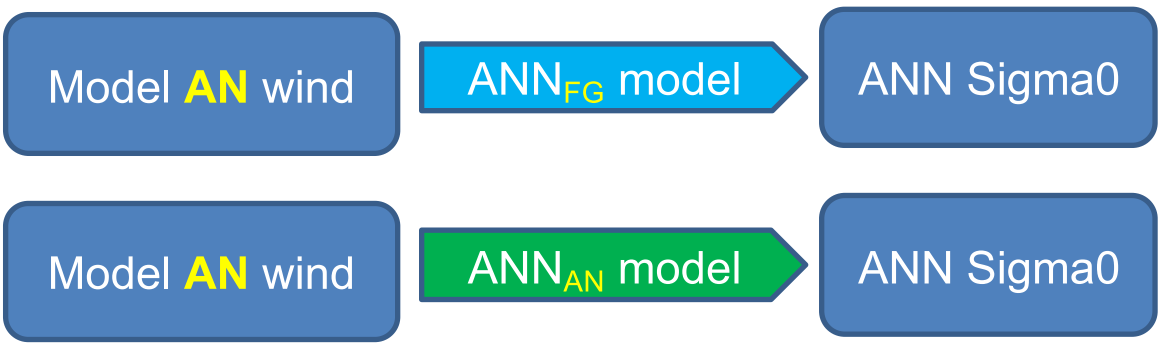

Figure 10: Same as Figure 9 but the use of model AN wind instead of model FG wind for validation.

Table 3: Same as Table 2 but when used with AN winds.

N = 1266843	ANN model trained using FG wind (ANN_{FG})	ANN model trained using AN wind (ANN_{AN})
Bias	-0.012	0.005
SDD	1.012	0.980
R	0.9813	0.9824

4.3 Impact of Seasonality

In order to get a feeling of the impact of training duration and its timing with respect to the validation period on the performance of the ANN model, we divided ASCAT-B $\boldsymbol{\sigma}^0$ data collocated with the IFS model FG wind covering a period of 8 months into two equal parts: the first 4-month period between 1 August and 30 November 2020 was reserved for training while the following 4-month period from 1 December 2020 to 31 March 2021 was reserved for validation. (This was done at an early stage when there was not much data available).

Various durations and timings were selected from the training period to train ANN_{FG} $\boldsymbol{\sigma}^0$ models. In total 10 models were trained. The model durations and timings are schematically shown in the timeline in the left-hand panel of Figure 11 and Figure 12 (same timeline in both figures). All models were validated against the whole validation data set from the second 4-month period as shown in the same timeline. The SDD and correlation coefficient for each model are shown in the right-hand panels of Figure 11 and Figure 12, respectively.

It is clear that the best results are achieved with training using 4 months of data followed closely by training using the 3 months just before the start of the validation period (the top two horizontal bars in Figure 11 and Figure 12). The third best is the ANN model trained using the 3 months with a month gap from the validation start (third horizontal bar in the figures) followed closely with the model trained using the closest two months to the validation period (fourth bar in the figures). The performance of the models based on the other two-month periods degrades as the gap to the validation period increases. The two models based on monthly data closet to the validation period perform better than the model based on the furthest 2-month period. The models based on monthly data show better performance when they are closer to the validation period. The only exception is the model based on data from October 2020.

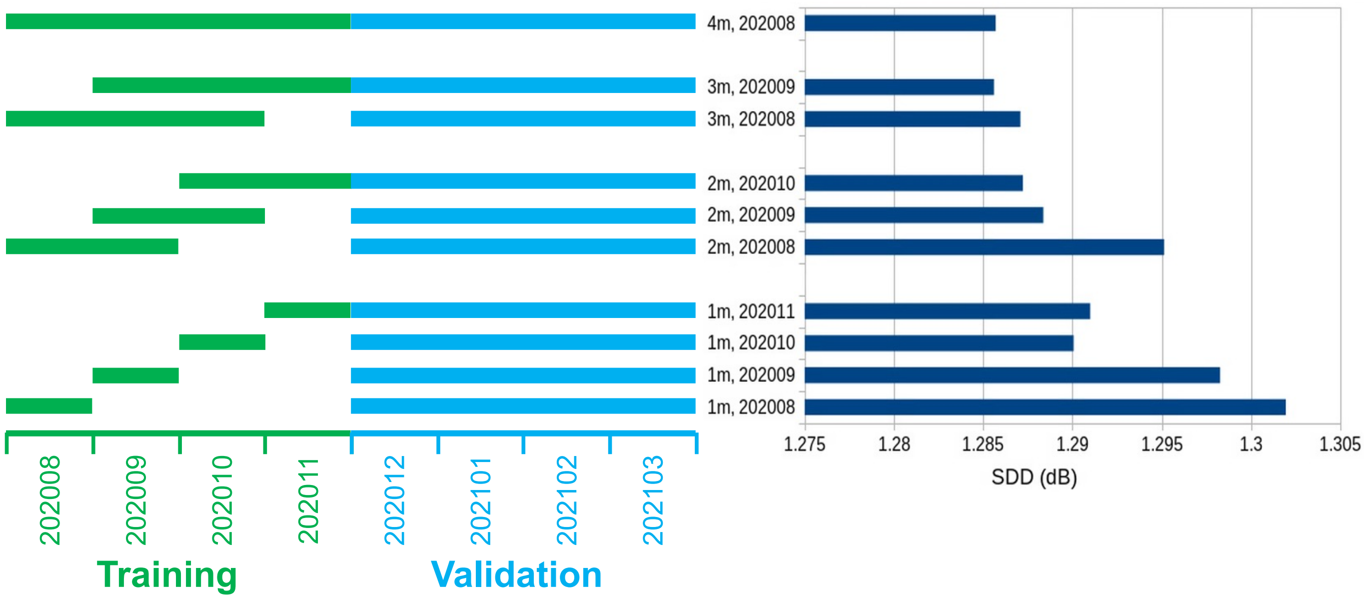

Figure 11: Impact of training duration and its timing with respect to validation period on SDD. Validation of ANN_{FG} model trained using datasets (FG wind and ASCAT-B $\boldsymbol{\sigma}^0$) for various durations within 4-month period between 1 August and 30 November 2020 against ASCAT-B data covering the following 4-month period between 1 December 2020 to 31 March 2021. The labels next to the horizontal bars, represent the duration of training in months and the month at which the training duration commenced. Left-hand panel shows the corresponding timeline of the training and validation periods.

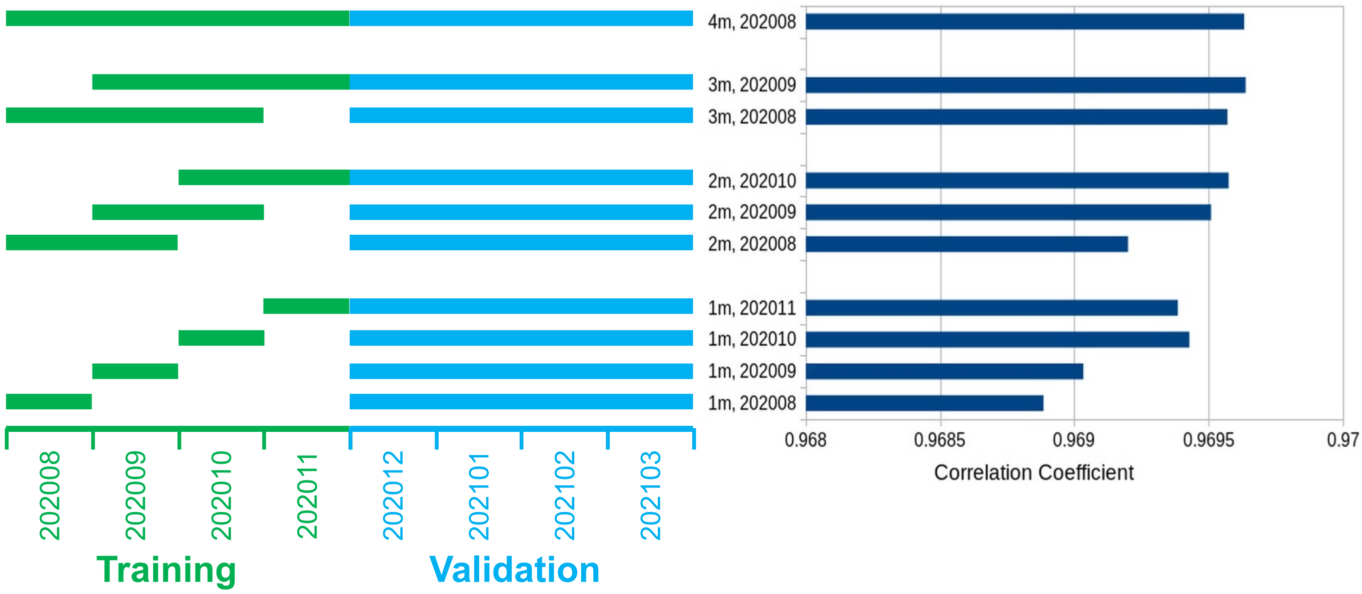

Figure 12: Same as Figure 11 but for impact on the correlation coefficient. The timeline on the left-hand panel is the same as that in Figure 11.

Although the duration of the training data seems like an important factor to produce better performance, it is in fact how much the data are closer to the validation period which is more important. It is important to recall from Figure 3 that a model based on a full year of data behaves equally good at all periods of time including those which are more than a year away. One can conclude that a year of data hopefully covering a wide range of weather conditions is needed for training a successful ANN model.

4.4 Impact of Including Air Density

Air density is one of the variables which influence the impact of wind speed on the water surface. This impact was studied and included in CMOD-7 [22]. The air density as computed from IFS was included as a feature (input) to the ANN model. Validation of $\boldsymbol{\sigma}^0$ values estimated by the ANN model against ASCAT-B $\boldsymbol{\sigma}^0$ data shows marginal impact on SDD and correlation coefficient. However, upon comparing the geographical distribution of the difference between ANN estimated $\boldsymbol{\sigma}^0$ and the ASCAT-B measurements shown in Figure 13 to those produced using FG wind alone in Figure 4, it is possible to see that the differences are getting smaller with the inclusion of air density.

4.5 Impact of Sea-State Variables

Since scatterometers detect the surface capillary gravity waves, other processes that contribute to (or hinder) the growth of such waves are needed to be considered while developing a scatterometer $\boldsymbol{\sigma}^0$ observation operator. The first candidate is the ocean waves. Although ocean waves have scales higher than the scale of the gravity-capillary waves, they do interact with each other. ECMWF runs an ocean wave model called ECWAM with IFS. The model predicts a wide range of sea-state variables. A list of some of those variables are listed in Table 4. The full list with description is available from [11]. The model short (up to 12 hours) forecast (FC) sea-state variables are used with the FG wind as well as model AN sea-state variables are used with AN wind.

Several combinations of sea-state variables were used for training and validation of the trained models. The training period runs from 1 August to 30 November 2020 while the validation period runs from 1 December 2020 to 31 March 2021. IFS FG model winds and ASCAT-B $\boldsymbol{\sigma}^0$ measurements are used for both training and validation. The SDD, the percentage reduction in SDD and the correlation coefficient values for each model are shown in Figure 14, Figure 15 and Figure 16, respectively. The three plots suggest that the inclusion of significant wave height (SWH), mean zero-crossing wave period (also known as mean period based on second moment, MP2), normalized energy flux into waves (PHIAW), and normalized energy flux into ocean (PHIOC) to FG wind produces the best performance with a reduction of slightly less than 8% in SDD. The best performance for adding just two sea-state variables is for the inclusion of SWH and MP2 with about 5% reduction in SDD. For a single variable, the best sea-state variable to include is PHIOC with a reduction of about 2% in SDD.

The scatter plot comparing $\boldsymbol{\sigma}^0$ from the ANN model that was trained using FG wind and the FC sea-state variables of significant wave height (SWH), mean zero-crossing wave period (MP2), normalized energy flux into waves (PHIAW), and normalized energy flux into ocean (PHIOC) to ASCAT-B $\boldsymbol{\sigma}^0$ measurements is shown in Figure 17. The ANN model is now able to produce $\boldsymbol{\sigma}^0$ values below 31.5 dB. This is quite an improvement compared to ANN models with wind-only training (see Figure 1).

The geographical distribution of bias between $\boldsymbol{\sigma}^0$ estimated by few ANN models that were trained using 1-year of data (1 August 2020 – 31 July 2021) and the measured $\boldsymbol{\sigma}^0$ by ASCAT-B during the 1-year period from 1 August 2021 to 31 July 2022 are shown in Figure 18, Figure 19 and Figure 20. The improvement that each sea-state variable combination on the bias clear.

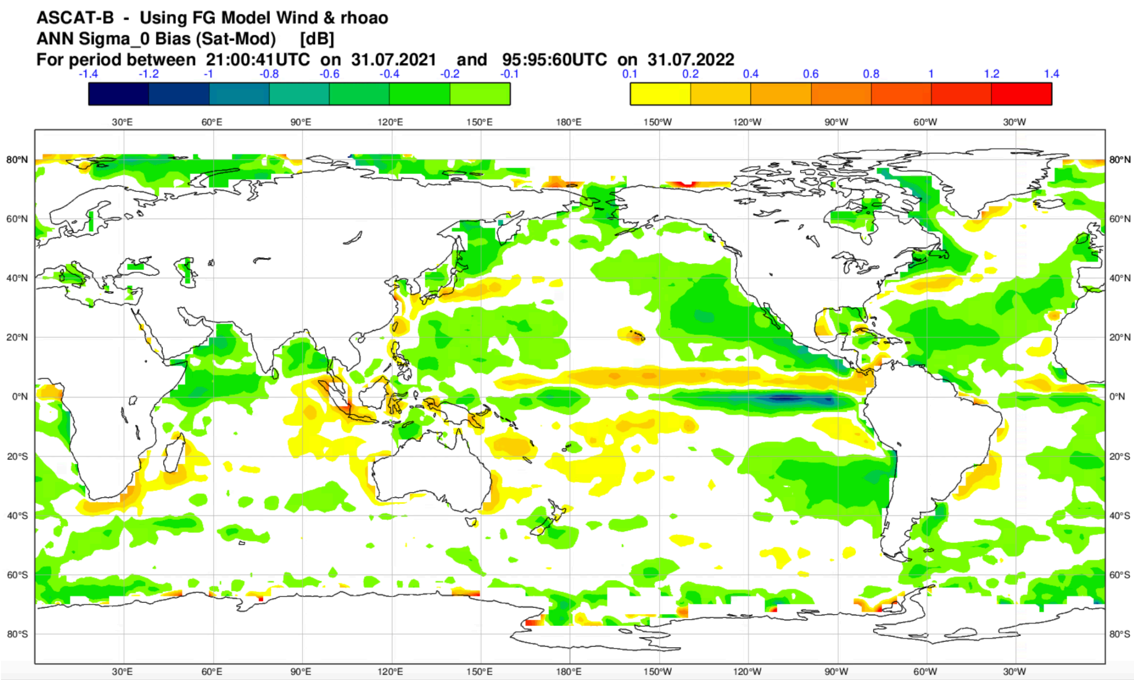

Figure 13: The geographical distribution of bias between $\boldsymbol{\sigma}^0$ *estimated by an ANN model that was trained using FG wind and air density and the measured* $\boldsymbol{\sigma}^0$ *by ASCAT-B during the 1-year period from 1 August 2021 to 31 July 2022. The unit is dB.*

Table 4: The used sea-state variables (sorted based on the short name).

Variable name	Short name	Units
Benjamin-Feir index	bfi	dimensionless
Coefficient of drag with waves	cdww	dimensionless
Mean wave period based on first moment	mp1	s
Mean zero-crossing wave period	mp2	s
Mean wave period	mwp	s
Mean wave direction of first swell partition	mwd1	degree
Mean wave direction of second swell partition	mwd2	degree
Mean wave direction of third swell partition	mwd3	degree
Mean wave period of first swell partition	mwp1	s
Mean wave period of second swell partition	mwp2	s
Mean wave period of third swell partition	mwp3	s
Mean square slope of waves	msqs	dimensionless
Normalized energy flux into waves	phiaw	dimensionless
Normalized energy flux into ocean	phioc	dimensionless
Peak wave period	pp1d	s
Air density over the oceans	rhoao	kg m^{-3}
Significant height of combined wind waves & swell	swh	m
Significant wave height of first swell partition	swh1	m
Significant wave height of second swell partition	swh2	m
Significant wave height of third swell partition	swh3	m
Normalized stress into ocean	tauoc	dimensionless
U-component stokes drift	ust	m s^{-1}
V-component stokes drift	vst	m s^{-1}
Wave spectral directional width	wdw	dimensionless
Wave energy flux magnitude	wefxm	W m^{-1}
Wave spectral kurtosis	wsk	dimensionless
Wave spectral peakedness	wsp	dimensionless
Wave spectral skewness	wss	dimensionless
Free convective velocity over the oceans	wstar	m s^{-1}

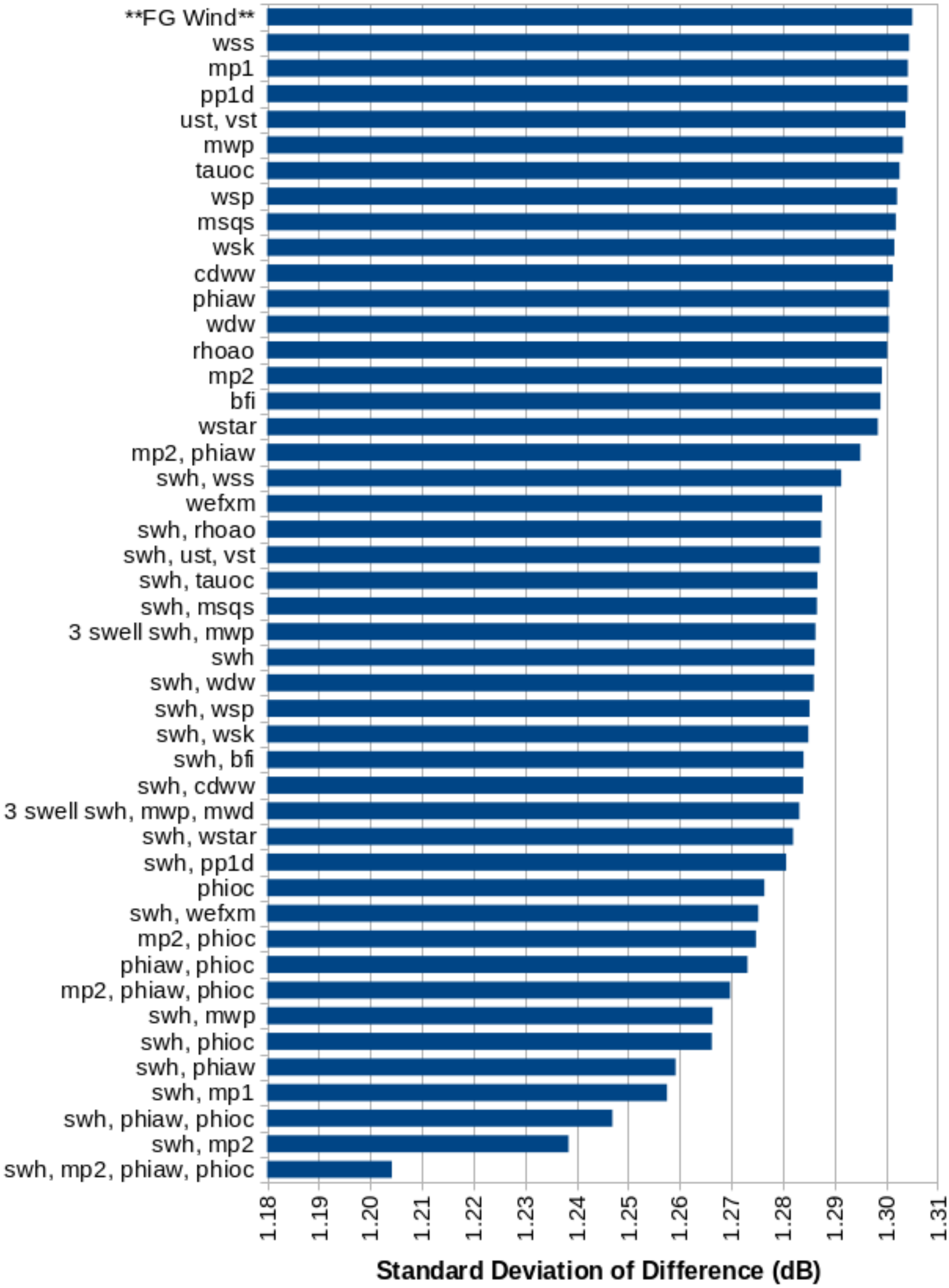

Figure 14: Impact of adding sea-state variables to FG winds for training and validating the trained models against ASCAT-B **σ**⁰ *measurements in terms of SDD. Training period is between 1 August and 30 November 2020 while the validation period is between 1 December 2020 to 31 March 2021. The labels next to the horizontal bars represent the FC sea-state variables used for that model. The label "**FG wind**" refers to the reference FG-wind only model. A list of variable full names is listed in Table 4.*

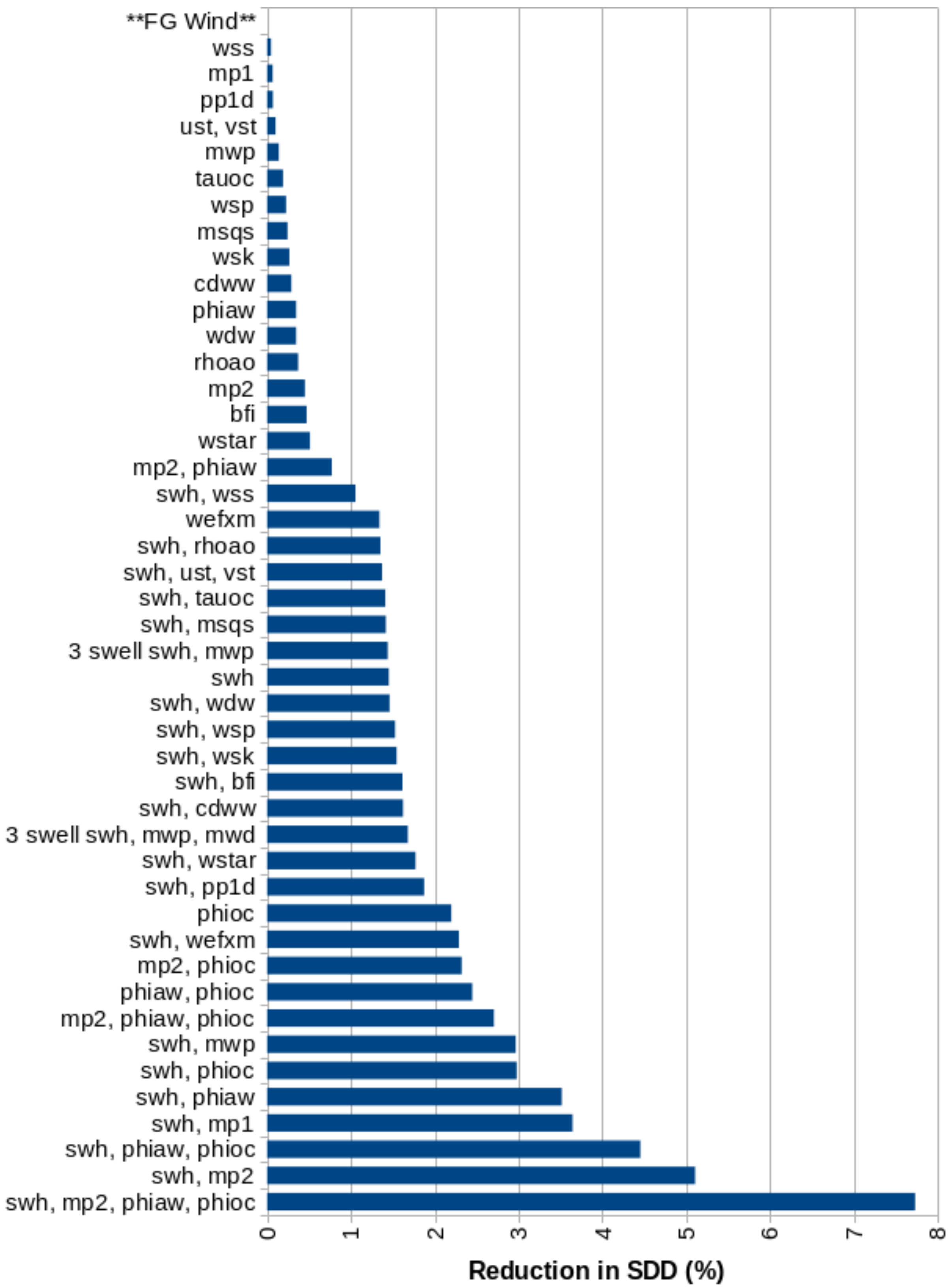

Figure 15: As Figure 14 but the impact is in terms of percentage reduction in SDD.

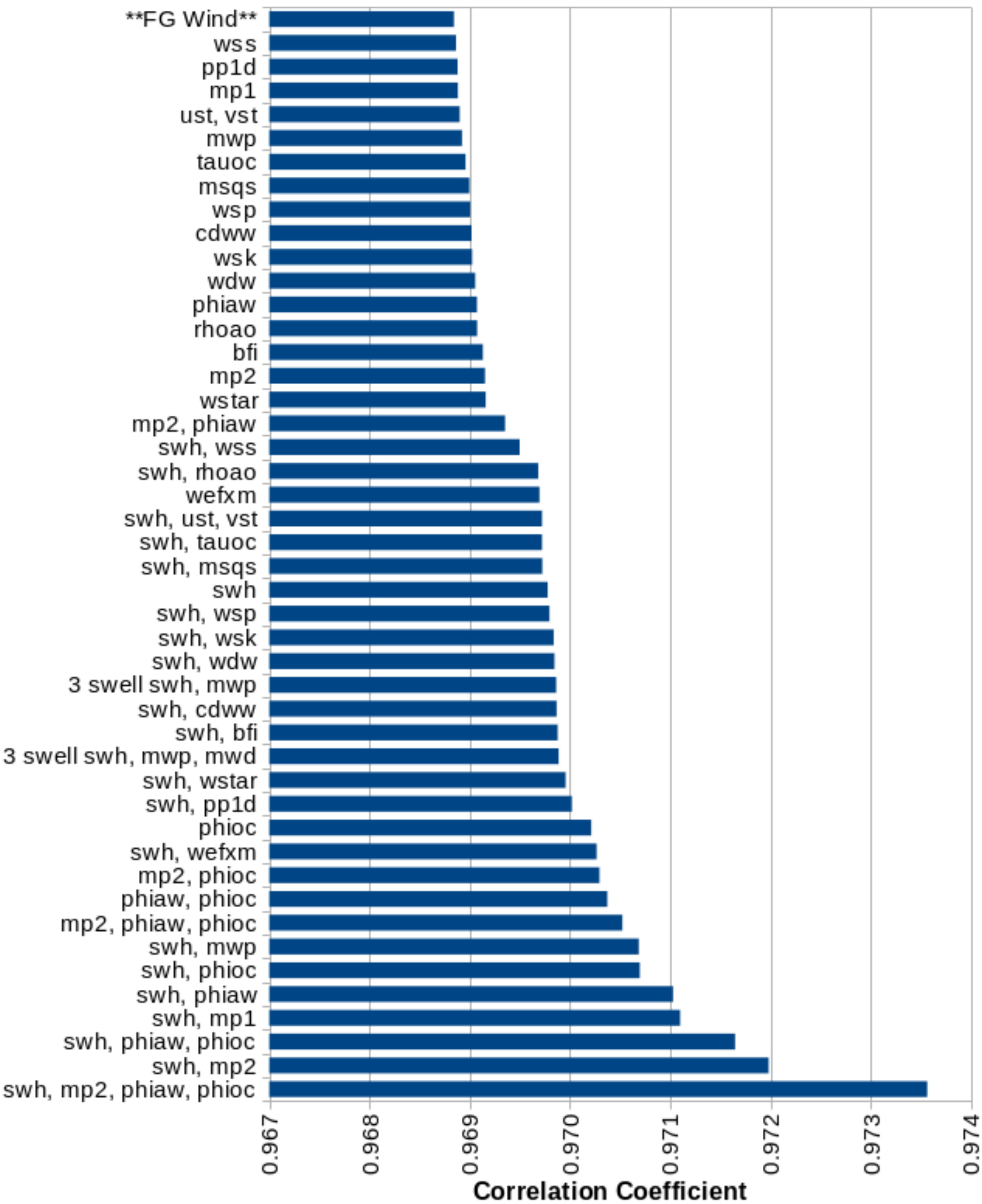

Figure 16: As Figure 14 but the impact is in terms of the correlation coefficient.

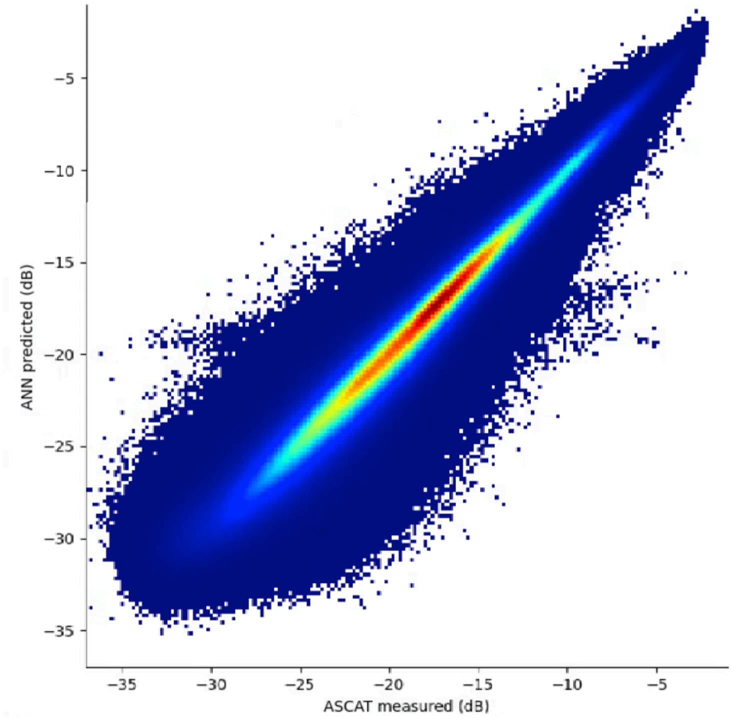

Figure 17: Comparison of $\boldsymbol{\sigma}^0$ *estimated using an ANN model that trained using FG wind and the FC sea-state variables of significant wave height (SWH), mean zero-crossing wave period (MP2), normalized energy flux into waves (PHIAW), and normalized energy flux into ocean (PHIOC) against ASCAT-B* $\boldsymbol{\sigma}^0$ *measurements over a 9-month (1 August 2021 – 30 April 2022) period. The ANN model was trained on 1 year of FG wind vector and 4 FC sea-state variables covering the period from 1 August 2020 to 31 July 2021.*

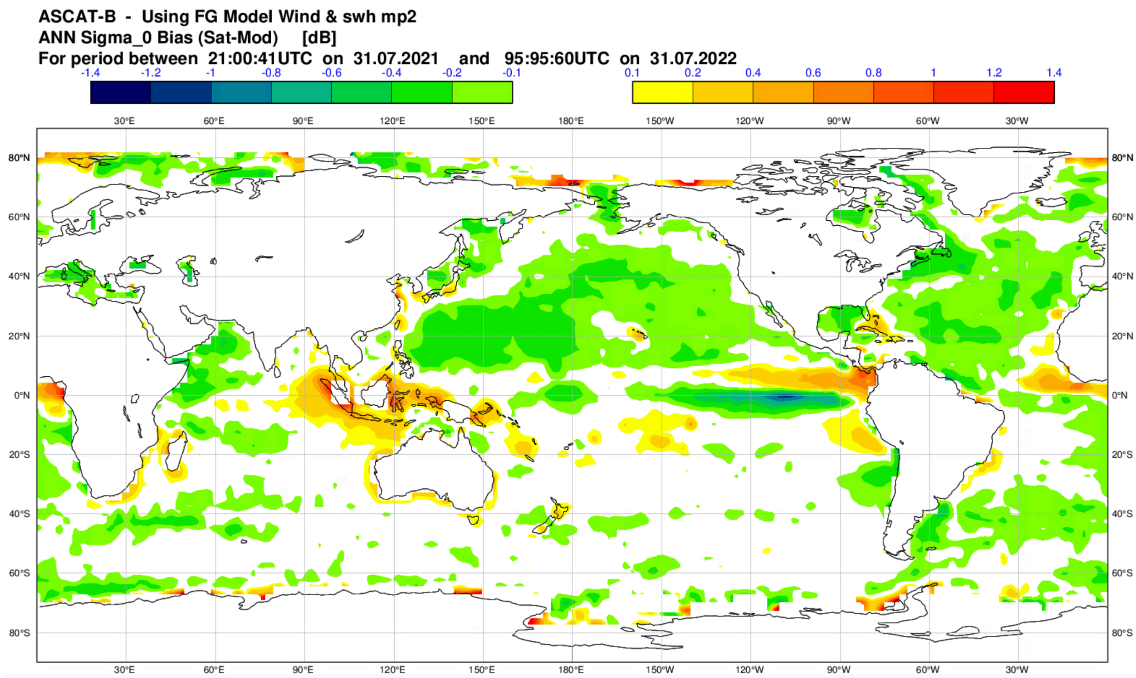

Figure 18: The geographical distribution of bias between $\boldsymbol{\sigma}^0$ *estimated by an ANN model that was trained using FG wind and FC sea-state variables of significant wave height and mean zero-crossing wave period and the measured* $\boldsymbol{\sigma}^0$ *by ASCAT-B during the 1-year period from 1 August 2021 to 31 July 2022. The unit is dB.*

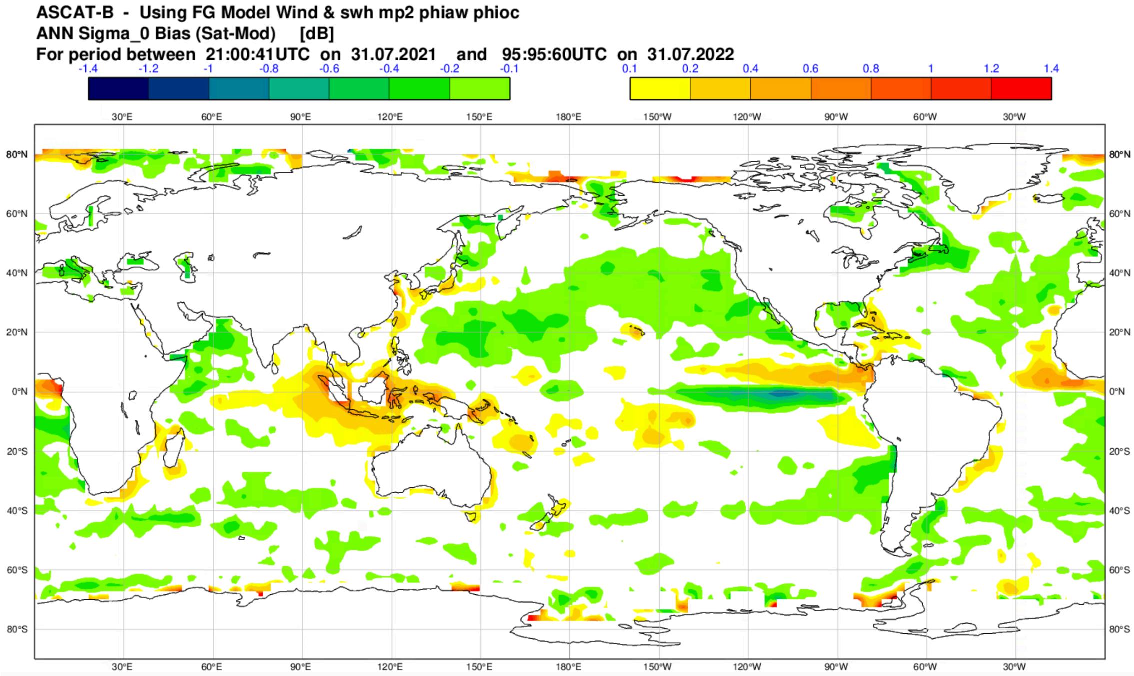

Figure 19: Same as Figure 18 but for the ANN model that was trained using FC energy flux into waves and energy flux into ocean on top of those used in the model of Figure 18 (i.e. FG wind and significant wave height and mean zero-crossing wave period).

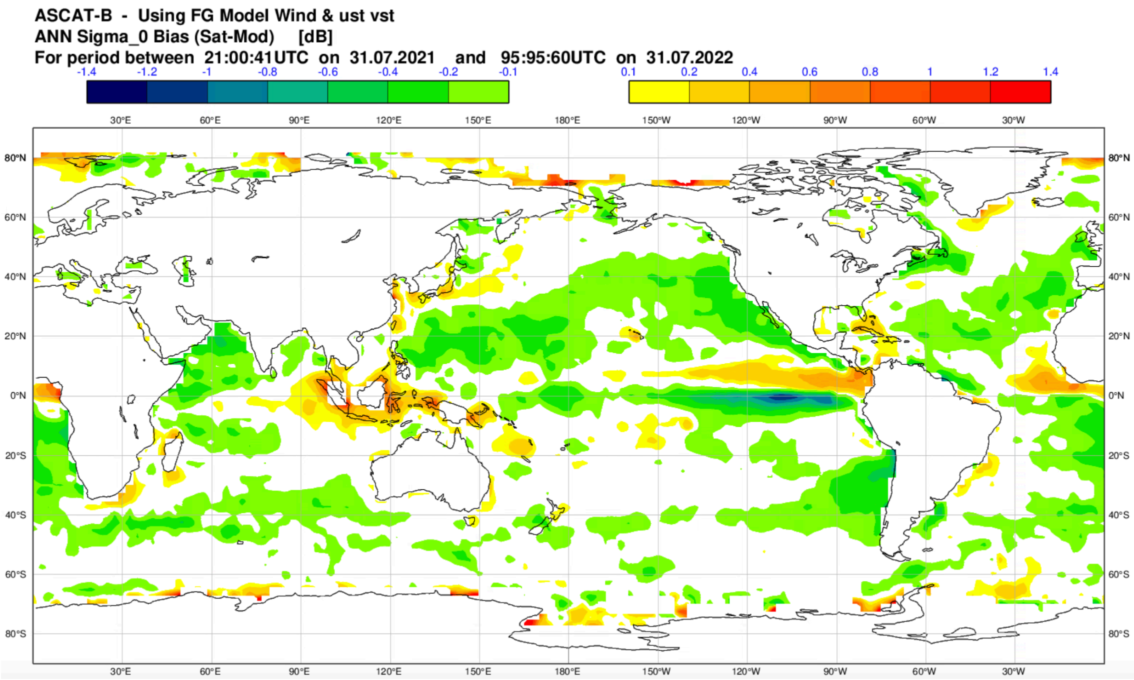

Figure 20: Same as Figure 18 but for the ANN model that was trained using FG wind and FC Stokes' drift vector.

4.6 Impact of Atmospheric and Oceanic Variables

The development of surface capillary gravity waves, which impact the backscatter of the SCAT signals, is impacted by several atmospheric and oceanic processes apart from wind strength. A (nonexclusive) list of some of these processes is provided in Table 5. The ECMWF IFS model provides global estimates of them. The oceanic variables considered here are sea surface temperature (sst), sea water practical salinity (so) and current velocity (eastward and northward sea water velocity, ocu and ocv, respectively). The remaining variables in Table 5, such as surface stress, surface roughness, precipitation, are all atmospheric variables. The full list with brief description is available from [12]. The model short forecast (FC up to 12 and 24 hours atmospheric and oceanic variables, in respective order) are used in addition to the FG wind. Note that the ANN modelling also involves the ASCAT incidence and azimuth angles for all cases even if they are not mentioned. As before, ASCAT-B $\boldsymbol{\sigma}^0$ values collocated with the IFS model variables covering the period from 1 August 2020 to 31 July 2021 (1 year) are used for training the ANN models. The data covering the following year (from 1 August 2021 to 31 July 2022) are used for the validation of the models.

Table 5: The used atmospheric and oceanic variables (sorted based on the short name).

Variable name	**Short name**	**Units**
Charnock	chnk	dimensionless
Convective precipitation	cp	m
Forecast surface roughness	fsr	m
Instantaneous 10-metre wind gust	i10fg	m s^{-1}
Instantaneous eastward turbulent surface stress	iews	N m^{-2}
Instantaneous northward turbulent surface stress	inss	N m^{-2}
Large-scale precipitation	lsp	m
Mean sea level pressure	msl	Pa
Eastward sea water velocity	ocu	m s^{-1}
Northward sea water velocity	ocv	m s^{-1}
Sea water practical salinity	so	psu
Sea surface temperature	sst	K
Total cloud cover	tcc	(0.-.1)
Total column rain water	tcrw	kg m^{-2}
Total column snow water	tcsw	kg m^{-2}
Total column vertically-integrated water vapour	tcwv	kg m^{-2}
Total precipitation	tp	m
Friction velocity	zust	m s^{-1}

The impact of using various combinations of atmospheric and oceanic variables on fitting the measured ASCAT-B $\boldsymbol{\sigma}^0$ values for the validation period compared to the use of wind velocity components only can be seen in Figure 21, Figure 22 and Figure 23. The impact in terms of standard deviation of the difference (SDD) between $\boldsymbol{\sigma}^0$ values from the ANN model and the ASCAT-B measurements is shown in Figure 21. The percentage of reduction of SDD compared to the computation of ANN $\boldsymbol{\sigma}^0$ using FG wind only is shown in Figure 22. The correlation coefficients of ANN model trained using various combinations of atmospheric and oceanic variables when compared against the measured ASCAT-B $\boldsymbol{\sigma}^0$ are shown in Figure 23.

The use of extra atmospheric variables improves the performance of the ANN model by up to ~10% reduction of SDD. This is the case of using IFS model forecast surface stress vector (iews and inss), surface roughness (fsr), magnitude of friction velocity (zust), instantaneous 10-m wind gust (i10fg) and large-scale precipitation (lsp) in addition to the usual IFS model FG wind vectors.

For better visualisation the impact, the combinations are grouped into three groups, from lower to top in Figure 21, Figure 22 and Figure 23:

- FG wind velocity in addition to others,
- FC surface stress in addition to others,
- Both FG wind velocity and FC surface stress in addition to others.

In general, the combinations in the group “FG wind & others” (lower) show lower impact than the other two groups. The combinations of the group “FC surface stress & others” (middle) show more impact than members of the “FG wind & others” group. The best results are obtained when both FG wind and FC stress are used in combination of other variables.

In general, the more variables included in the ANN model, the better the results are. However, including a member of each of the following group of variables improves the performance:

- FG (10-m) wind
- Surface stress, friction velocity, gust
- Surface roughness, Charnock parameter
- Precipitation (large scale)

It seems that each of those groups includes intrinsic relevant pattern or patterns that complement the patterns in the other groups. The other atmospheric variables, not included in the groups above, introduce minor improvements. It was surprising that adding the oceanic variables, especially the surface current, show almost no impact. Adding the water salinity tends degrade the results slightly. This may be a consequence of the relatively poor ocean model resolution except for the tropical areas. This will be tested later by training using data from that area only.

Figure 24 shows the scatter plot of ANN $\boldsymbol{\sigma}^0$ from FG wind only against ASCAT-B $\boldsymbol{\sigma}^0$ over one year (1 August 2021 to 31 July 2022). Figure 25 shows a similar scatter plot for ANN $\boldsymbol{\sigma}^0$ from FG wind in addition to surface stress (iews, inss). Figure 26 shows the same but with adding surface roughness (fsr), magnitude of friction velocity (zust), 10-m wind gust (i10fg) and large-scale precipitation (lsp) to wind and stress vectors. The improvement can be realised by comparing Figure 26 (wind & stress & roughness & precipitation) to Figure 24 (wind only) and even to Figure 25 (wind & stress). The histograms of the three ANN $\boldsymbol{\sigma}^0$ cases: “wind only”; “wind & stress”; and “wind & stress & roughness & precipitation” are compared to the histogram of ASCAT-B $\boldsymbol{\sigma}^0$ and shown in Figure 27. Adding extra atmospheric variables to the wind produces more small ANN $\boldsymbol{\sigma}^0$ values (below 31.5 dB) as can be seen in Figure 27, especially the lower panel. Nevertheless, those small $\boldsymbol{\sigma}^0$ values are less represented in the ANN models which suggests that there is still room for improvement.

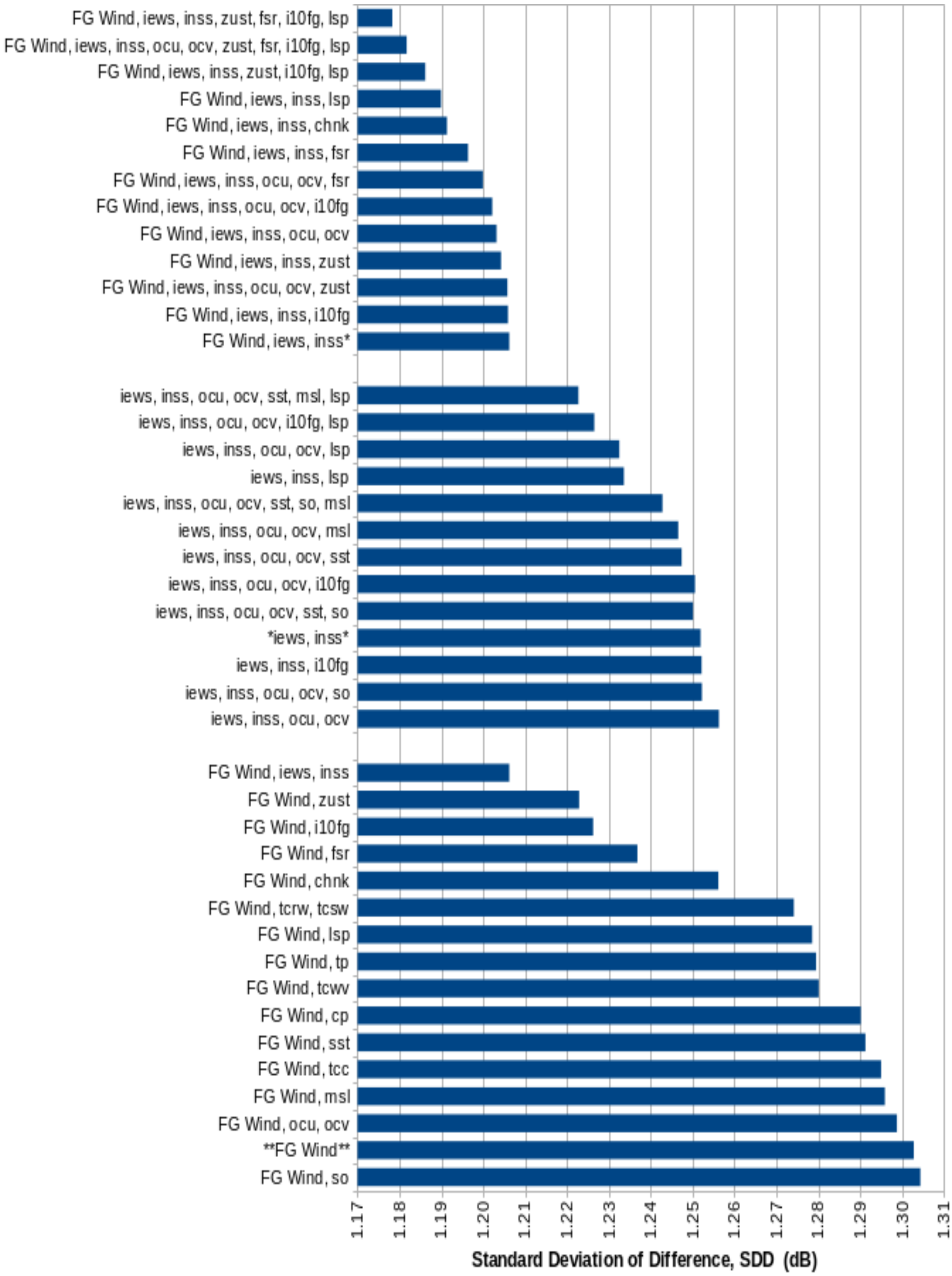

Figure 21: Impact of using relevant atmospheric and oceanic variables (including FG winds) for training and validating the trained models against ASCAT-B σ^0 *measurements in terms of SDD. Training period is between 1 August 2020 and 31 July 2021 while the validation period is between 1 August 2021 to 31 July 2022. The labels next to the horizontal bars represent the combination of atmospheric and oceanic variables used for that model. The label "**FG wind**" refers to the reference FG-wind only model. A list of variable full names is listed in Table 5.*

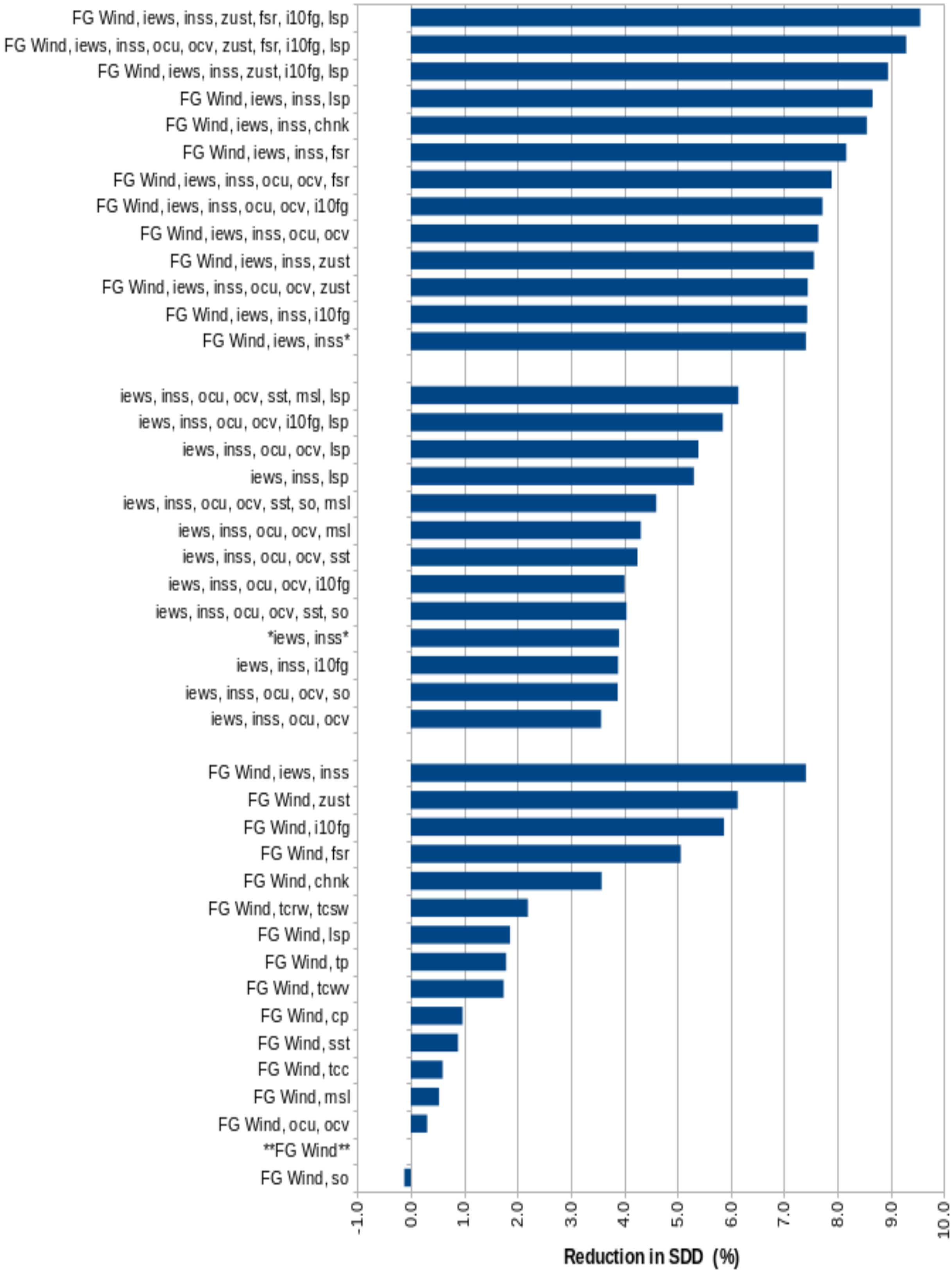

Figure 22: As Figure 21 but the impact is in terms of percentage reduction in SDD.

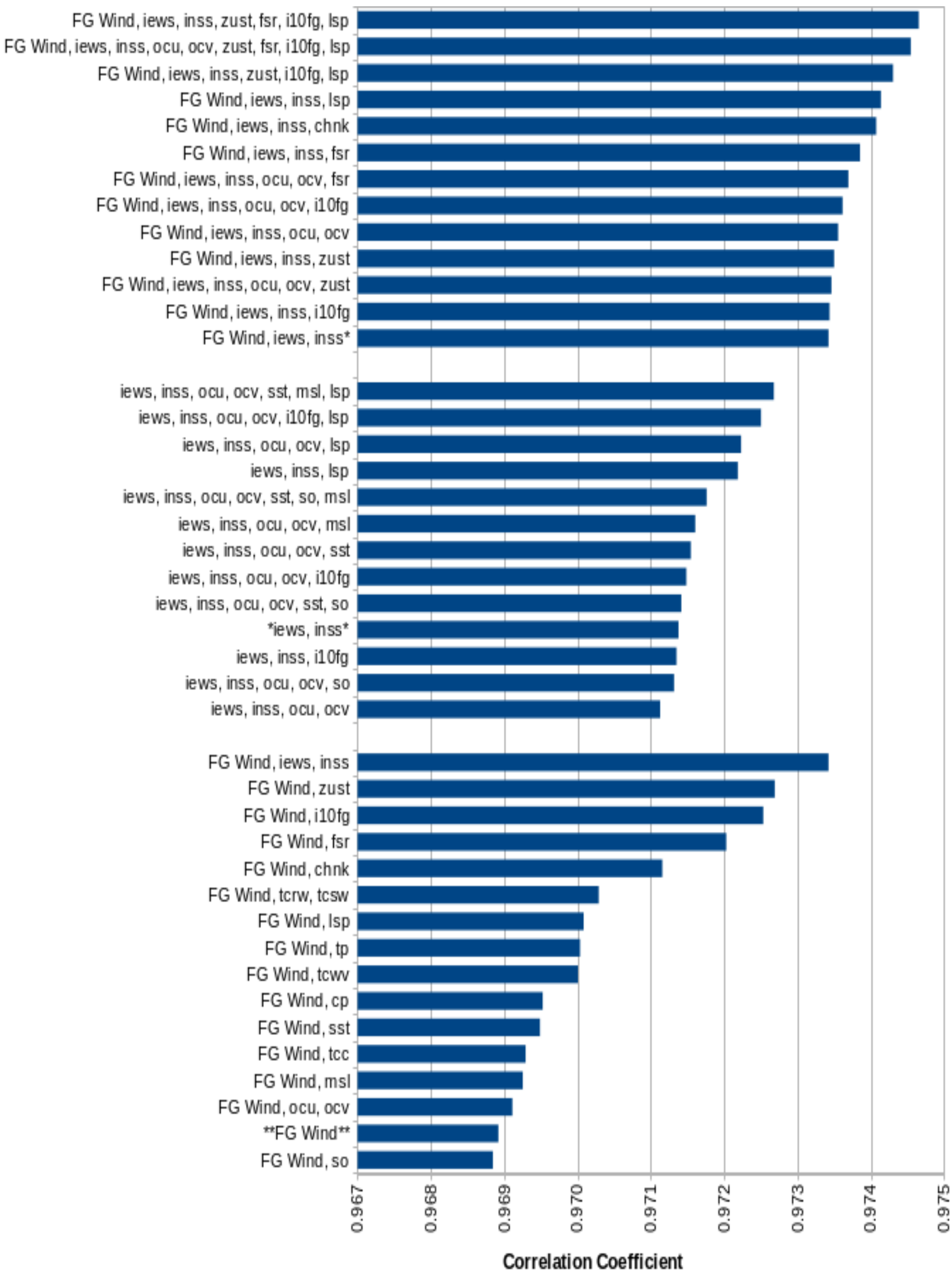

Figure 23: As Figure 21 but the impact is in terms of the correlation coefficient.

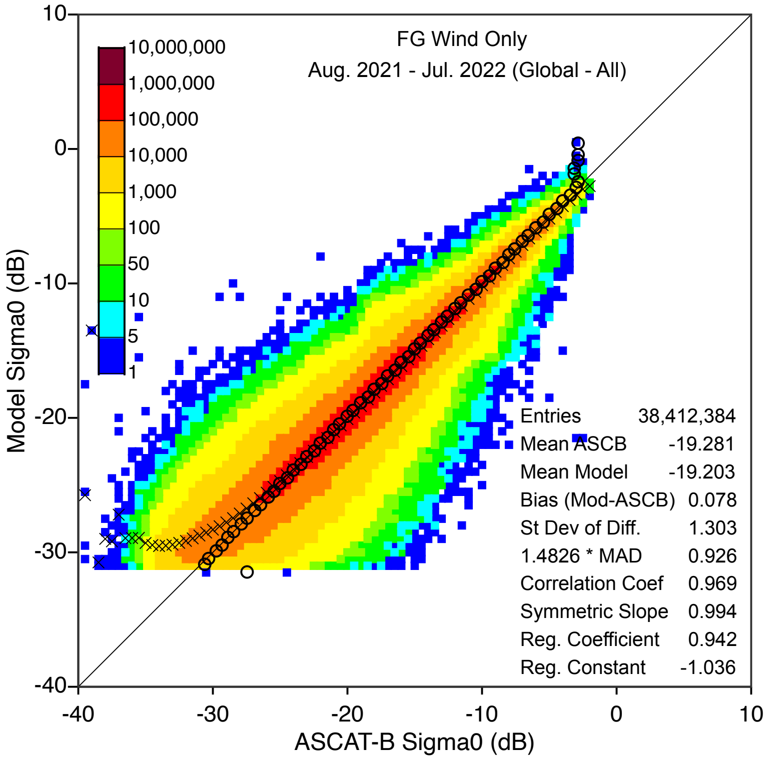

Figure 24: Validation of 12-month (1 August 2021 – 31 July 2022) of the ANN$_{FG}$ model $\boldsymbol{\sigma}^0$ *estimates using IFS model FG wind vectors against measured* $\boldsymbol{\sigma}^0$ *by ASCAT-B. Model training period is between 1 August 2020 and 31 July 2021.*

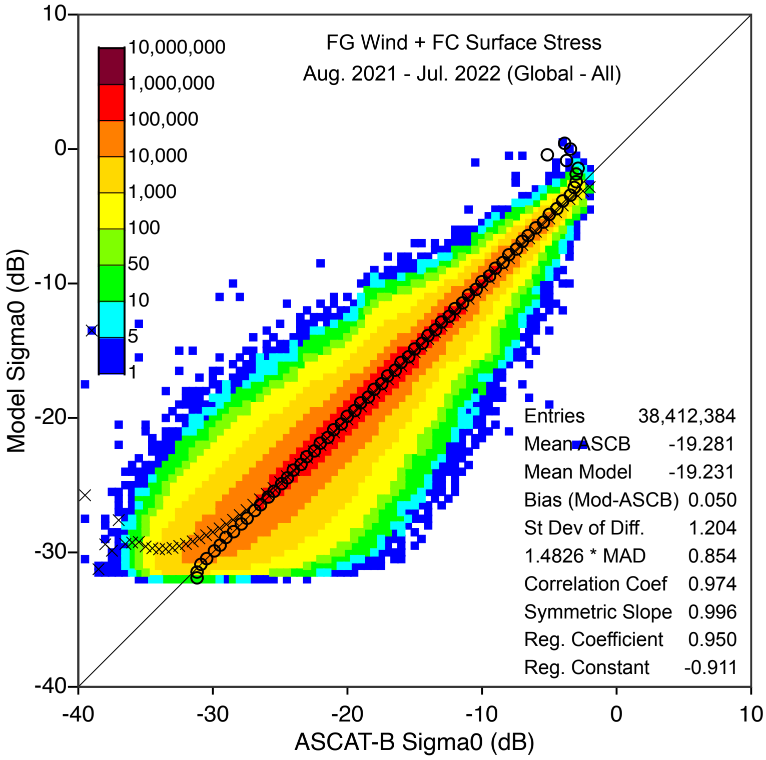

Figure 25: Same as Figure 24 but with the ANN$_{FG}$ model $\boldsymbol{\sigma}^0$ *estimates involves IFS model forecast instantaneous surface stress vector (iews and inss) in addition to IFS model FG wind vectors.*

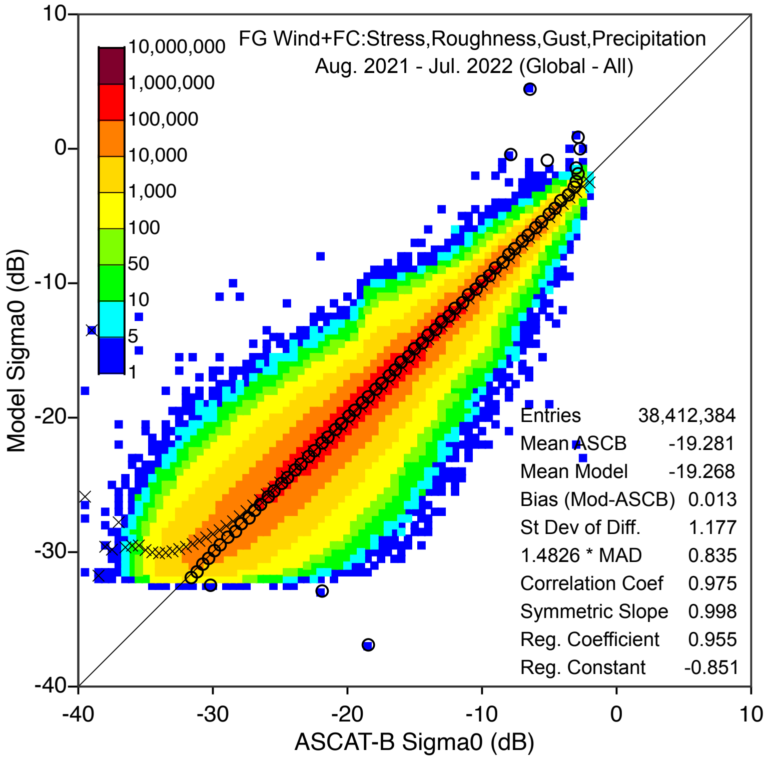

Figure 26: Same as Figure 24 but with the ANN$_{FG}$ model $\boldsymbol{\sigma}^0$ estimates involves IFS model forecast instantaneous surface stress vector (iews and inss) as in Figure 25, surface roughness (fsr), magnitude of friction velocity (zust), instantaneous 10-m wind gust (i10fg) and large-scale precipitation (lsp) in addition to IFS model FG wind vectors.

The geographical distribution of bias between $\boldsymbol{\sigma}^0$ estimated by three ANN models of "wind only" (Figure 24); "wind & stress" (Figure 25); and "wind & stress & roughness & precipitation" (Figure 26) from one side and ASCAT-B $\boldsymbol{\sigma}^0$ measurements, on the other side, over the validation year (1 August 2021 to 31 July 2022) are shown in Figure 28, Figure 29 and Figure 30, respectively. The reduction in bias when more atmospheric variables are added can be clearly seen.

5 Discussion

5.1 Accuracy of the ANN Models

The value of the standard deviation of the difference (SDD) between $\boldsymbol{\sigma}^0$ estimated by ANN models presented earlier and the measured $\boldsymbol{\sigma}^0$ varied roughly between 1.0 and 1.3 dB. This seems to be a high value compared to the expected error in $\boldsymbol{\sigma}^0$ measured by ASCAT family. Potential contributors to the rather high SDD values are the ANN modelling and the wind used for training and validation.

The same IFS FG wind vector used for the validation of the various ANN models was used for various CMOD geophysical model functions. Note that only wind is used (i.e. the impact of air density was not considered here). The resulted SDD values compared to ASCAT-B $\boldsymbol{\sigma}^0$ measurements are shown in Table 6. It is clear that the high SDD values are not due to the ANN training.

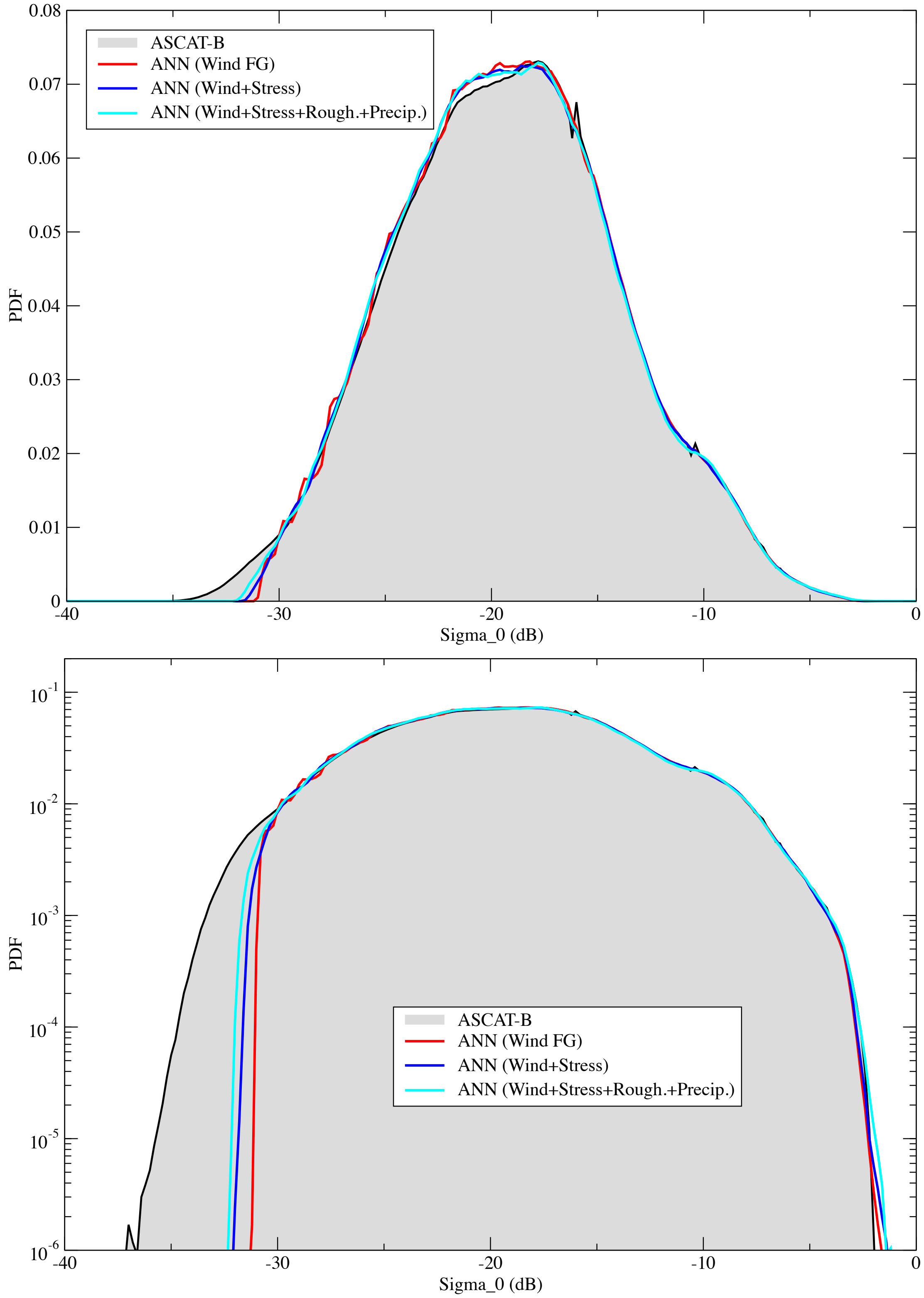

Figure 27: Histograms of ASCAT-B **σ**⁰ *measurements (black line and grey-shaded area) for the 1-year validation period (from 1 August 2021 to 31 July 2022). The corresponding histograms from the ANN models that use FG wind vector, and some other atmospheric variables like surface stress vector, surface roughness, … etc. Upper panel shows histograms on a linearly scaled y-axis while the lower panel shows the histograms on a logarithmically scaled y-axis.*

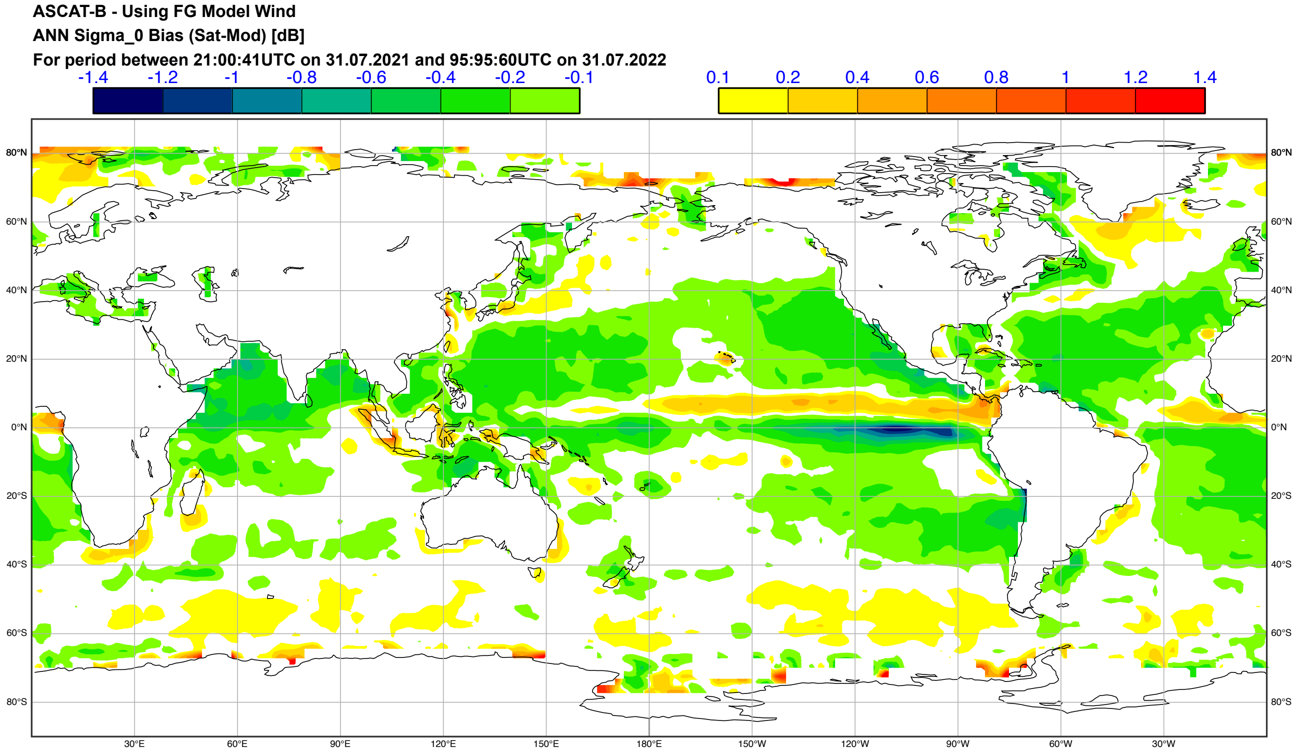

Figure 28: The geographical distribution of bias between ANN_{FG} model $\boldsymbol{\sigma}^0$ estimates using IFS model FG wind vectors and the measured $\boldsymbol{\sigma}^0$ by ASCAT-B during the 1-year period from 1 August 2021 to 31 July 2022. The unit is dB. Minor differences from Figure 4 are due to slightly different number of collocations.

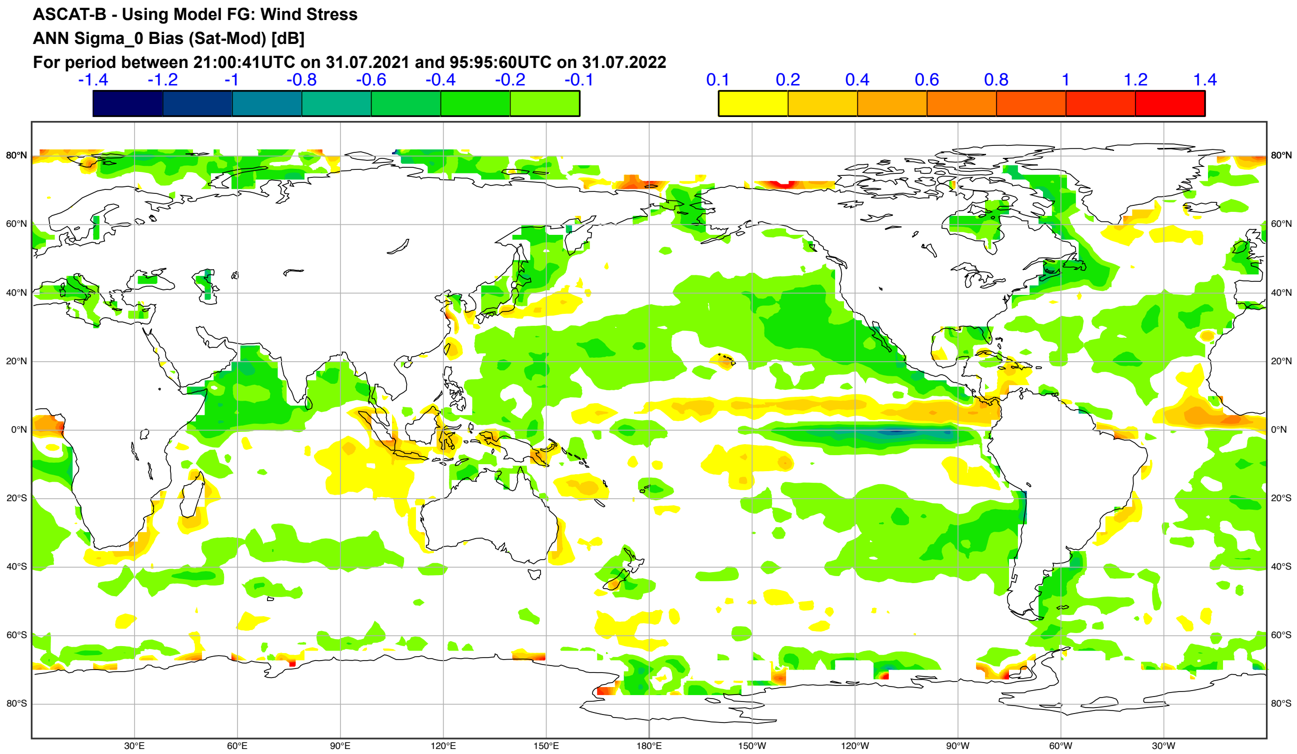

Figure 29: Same as Figure 28 but for the ANN model that was trained using FC instantaneous surface stress vector (iews and inss) on top of those used in the model of Figure 28 (i.e. FG wind vector only).

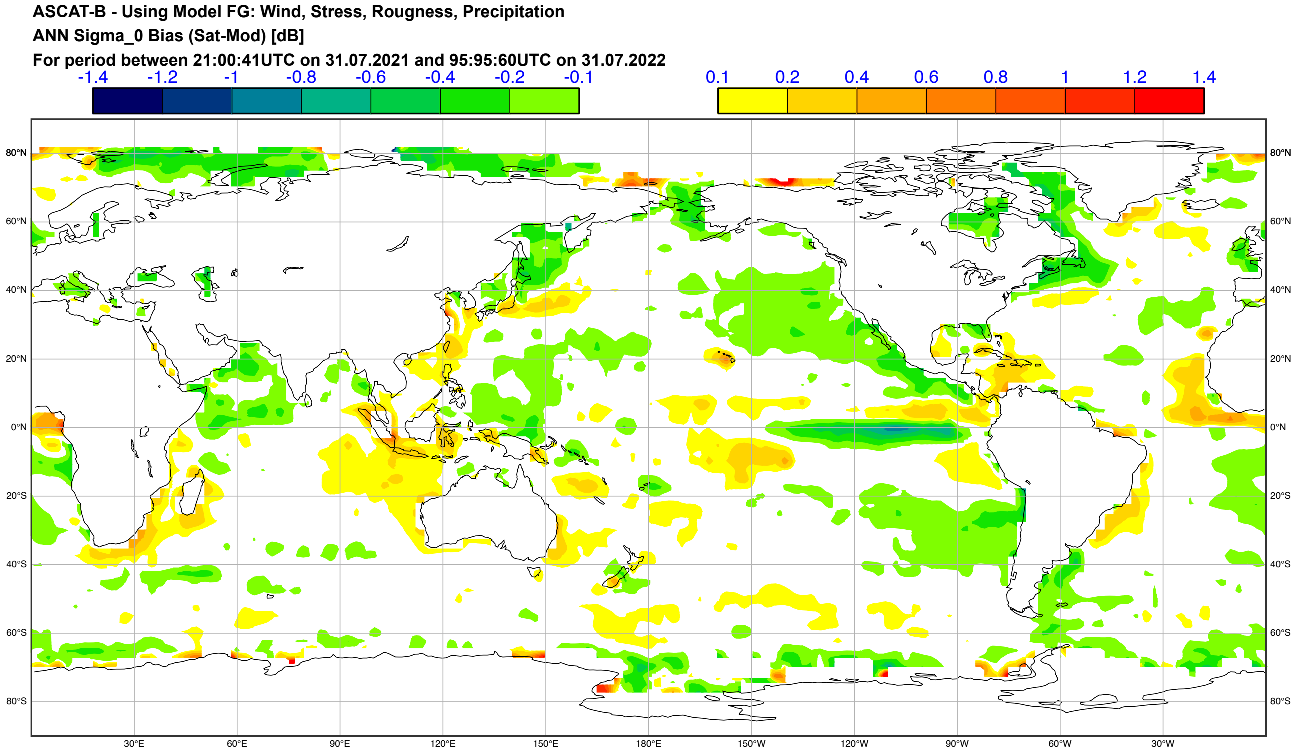

Figure 30: Same as Figure 29 but for the ANN model that was trained using FC surface roughness (fsr), magnitude of friction velocity (zust), instantaneous 10-m wind gust (i10fg) and large-scale precipitation (lsp) on top of those used in the model of Figure 29 (i.e. FG wind vector and FC instantaneous surface stress vector).

Table 6: Standard deviation of the difference (SDD) of $\boldsymbol{\sigma}^0$ *model estimates (using IFS model FG wind vectors) compared to* $\boldsymbol{\sigma}^0$ *measurements from ASAC-B.*

Model	**SDD (dB)**
$\mathrm{ANN_{FG}}$	1.3299
CMOD-5n	1.4689
CMOD-6	1.4327
CMOD-7	1.4621

Furthermore, a simple simulation was carried out by perturbing the model wind speed using Gaussian random noise with standard deviation (SD) of 1.0 m/s (to represent the model forecast error). The wind speed without perturbation is assumed to the truth in this case. CMOD-5n, CMOD-6 and CMOD-7 GMF's were used to compute $\boldsymbol{\sigma}^0$ from both the unperturbed wind speed (the truth in this simulation test) and the perturbed wind speed. The SDD value from each mode (comparing $\boldsymbol{\sigma}^0$ values from the perturbed wind speed to the wind speed without perturbation) are shown in Table 7. Therefore, 1.0 m/s wind error leads to more than 1.2 dB error in $\boldsymbol{\sigma}^0$ according to various CMOD GMF's. Since the wind speed error of IFS model around 1.0 m/s, SDD in the ANN simulation of 1.3 dB is something to be expected.

Table 7: Standard deviation of the difference (SDD) of $\boldsymbol{\sigma}^0$ *model estimates from perturbed wind speeds compared to* $\boldsymbol{\sigma}^0$ *estimates from unperturbed wind speeds.*

Model	SDD (dB)
CMOD-5n	1.226
CMOD-6	1.226
CMOD-7	1.319

Based on the above, one can estimate the wind speed error for model FG to be smaller than 1 m/s and for the model AN to be smaller than 0.8 m/s (I have perturbed the wind speed data using noise with SD of 0.8).

5.2 Comparison Against CMOD5N and CMOD7 GMF's

The previous sub-section, in specific Table 6, provided a gross comparison between ANN_{FG} model and various CMOD GMF's. ANN_{FG} model is slightly better than CMOD models.

Figure 31 shows a comparison between $\boldsymbol{\sigma}^0$ as estimated from ANN_{FG}, CMOD-5n and CMOD-7 against $\boldsymbol{\sigma}^0$ as measured by ASCAT-B for wind vectors aligned with the scatterometer beams (within 2°) for incidence angles of 27° and 59°. In general, the agreement against the ASCAT-B measurements and among various models is quite good. However, at low values of wind speed and $\boldsymbol{\sigma}^0$, contrary to CMOD-5n and CMOD-7 $\boldsymbol{\sigma}^0$ estimates, ANN $\boldsymbol{\sigma}^0$ estimates saturate below ~ 1.5 m/s. There is quite a support for this behaviour in the ASCAT-B data. Furthermore, this behaviour agrees with the available experimental evidence (e.g. [10]) that winds with speeds below a certain threshold (below around 2.0 m/s) is unable to produce any perturbation to the water surface. Therefore, the saturation value of $\boldsymbol{\sigma}^0$ could be a result of noise.

5.3 Comparison Against PARMIO Model

We obtained the code for the Passive and Active Reference Microwave to Infrared Ocean (PARMIO) model [9]. After having it running, several model configurations were changed in order to produce results that are comparable to the scatterometer data. Figure 33 shows the variation of $\boldsymbol{\sigma}^0$ estimated by PARMIO model for C-band with respect to the wind speed for various incidence angles. The comparison between $\boldsymbol{\sigma}^0$ estimated by ANN^{FG} and that by PARMIO model for several incidence angles is shown in Figure 34. The comparison does not look very good. It seems that having PARMIO model working properly needs more work and experience that what we have available for this project. Furthermore, the outcome of this comparison will not contribute to the aims of this project. Therefore, we decided not to pursue this effort any further.

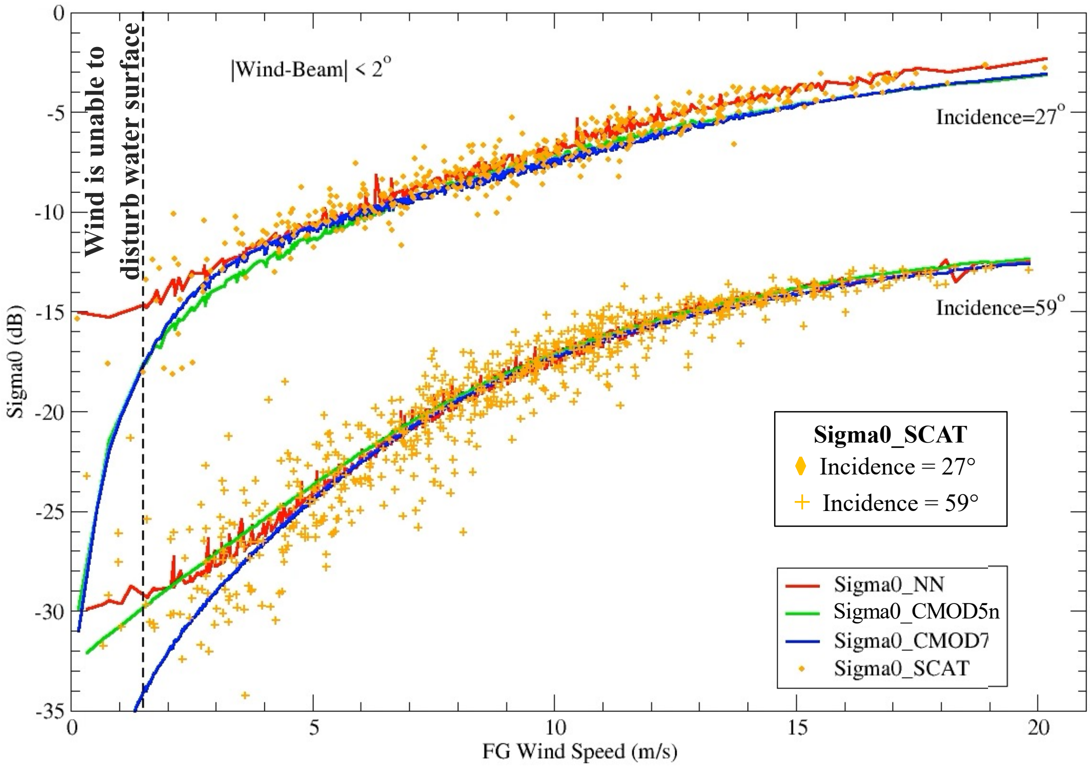

Figure 31: Comparison between **σ**^{0} *as estimated from ANN$_{FG}$ (Sigma0_NN in the plot), CMOD-5n and CMOD-7 against* **σ**^{0} *as measured by ASCAT-B (Sigma0_SCAT in the plot) for wind vectors aligned with the scatterometer beams (within 2°) for incidence angles of 27° and 59°.*

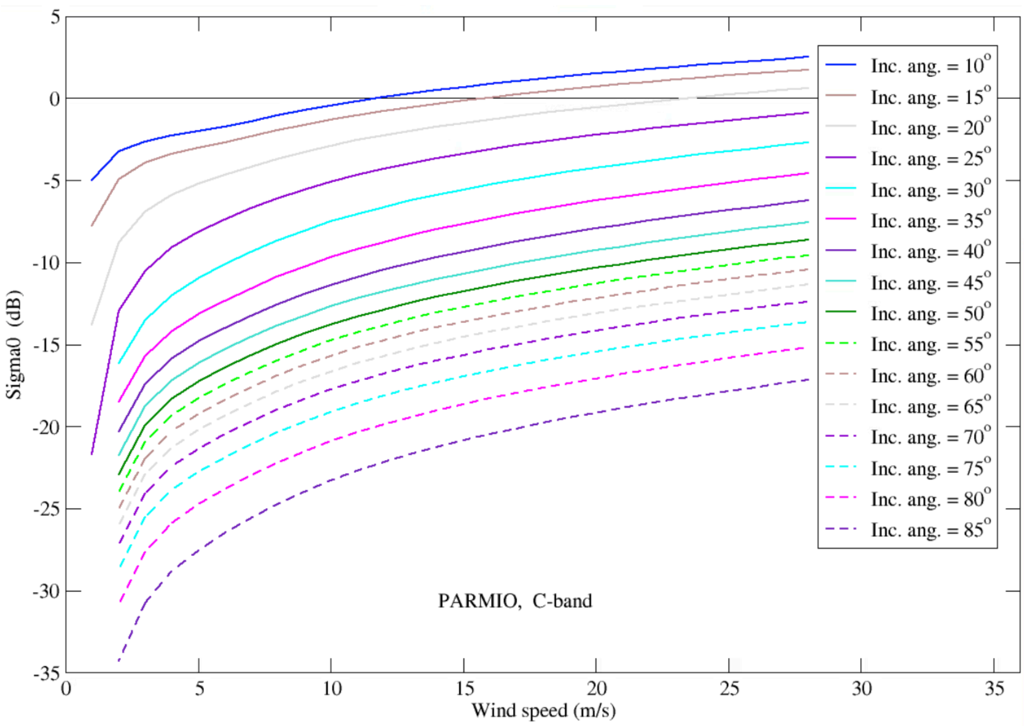

Figure 32: The variation of **σ**^{0} *estimated by PARMIO model for C-band with respect to the wind speed for various incidence angles.*

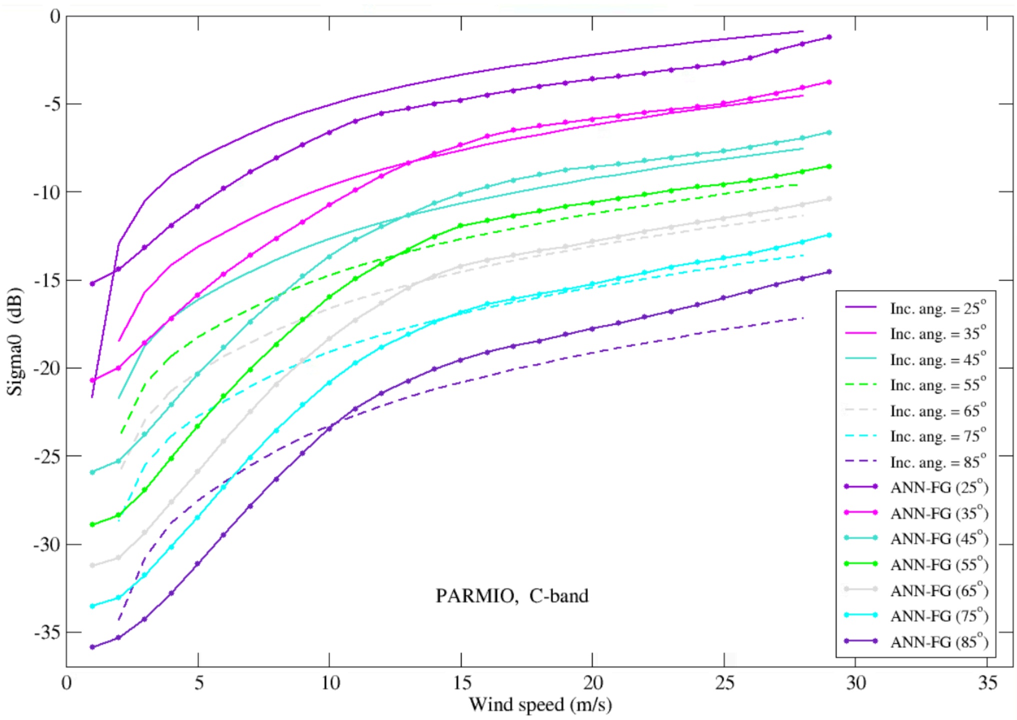

Figure 33: The comparison between ANN$_{FG}$ and PARMIO $\boldsymbol{\sigma}^0$ *estimates for various incidence angles.*

6 Conclusions

Artificial neural networks (ANN) is an attractive tool that can be used to develop numerical models based on the data without the need for a deep understanding of the detailed physical laws. Its flexibility to include extra features (input variables) with minimal effort is another strength.

A database consisting of collocations of ASCAT-B and ASCAT-C backscatter coefficients, $\boldsymbol{\sigma}^0$, and the corresponding ECMWF IFS model wind vectors and other related variables covering a period of more than 30 months was constructed for training and validation of the developed ANN models.

With $\boldsymbol{\sigma}^0$ as the target, ANN models based on first-guess (FG) surface wind vectors, analysis (AN) surface wind vectors and combinations of wind vectors and other variables as features were trained and validated. In general, 1-year of ASCAT-B data (together with their model collocations) were reserved for training. The rest of the data as well as the whole ASCAT-C data were used for validation. The 1-year of training data was needed to cover most of the atmospheric conditions.

ANN models, irrespective of being trained against FG or AN wind, were successful in predicting $\boldsymbol{\sigma}^0$ values for the data which they have not seen before. The compare very well with CMOD family of geophysical model functions (GMF's). ANN models based on wind vector only were not able to produce any $\boldsymbol{\sigma}^0$ values below 31.5 dB. This turned out to be the regime of wind speed below ~1.5 m/s, which is known to be around the value of the threshold wind speed to produce any water surface gravity-capillary waves with scale capable of backscattering the radar signal. ANN models based on wind vector and sea-state variables could produce $\boldsymbol{\sigma}^0$ values below 31.5 m/s. This is an indication that such backscatter values are not produced by wind but other processes like the ocean swell.

The value of the standard deviation of the difference (SDD) between $\boldsymbol{\sigma}^0$ estimated by ANN models presented earlier and the measured $\boldsymbol{\sigma}^0$ varied roughly between 1.0 and 1.3 dB. This rather high value is due to the accuracy of the IFS model wind.

7 Further Work

The performance of the trained ANN models will be investigated further especially under extreme conditions like tropical cyclones.

Training using oceanic variables in the tropics could be done to check whether a positive impact on the ANN modelling of $\boldsymbol{\sigma}^0$.

8 Appendix

The geographical distribution of the IFS mean surface wind over one year from 1 August 2021 to 31 July 2022 is shown in Figure 34 for reference.

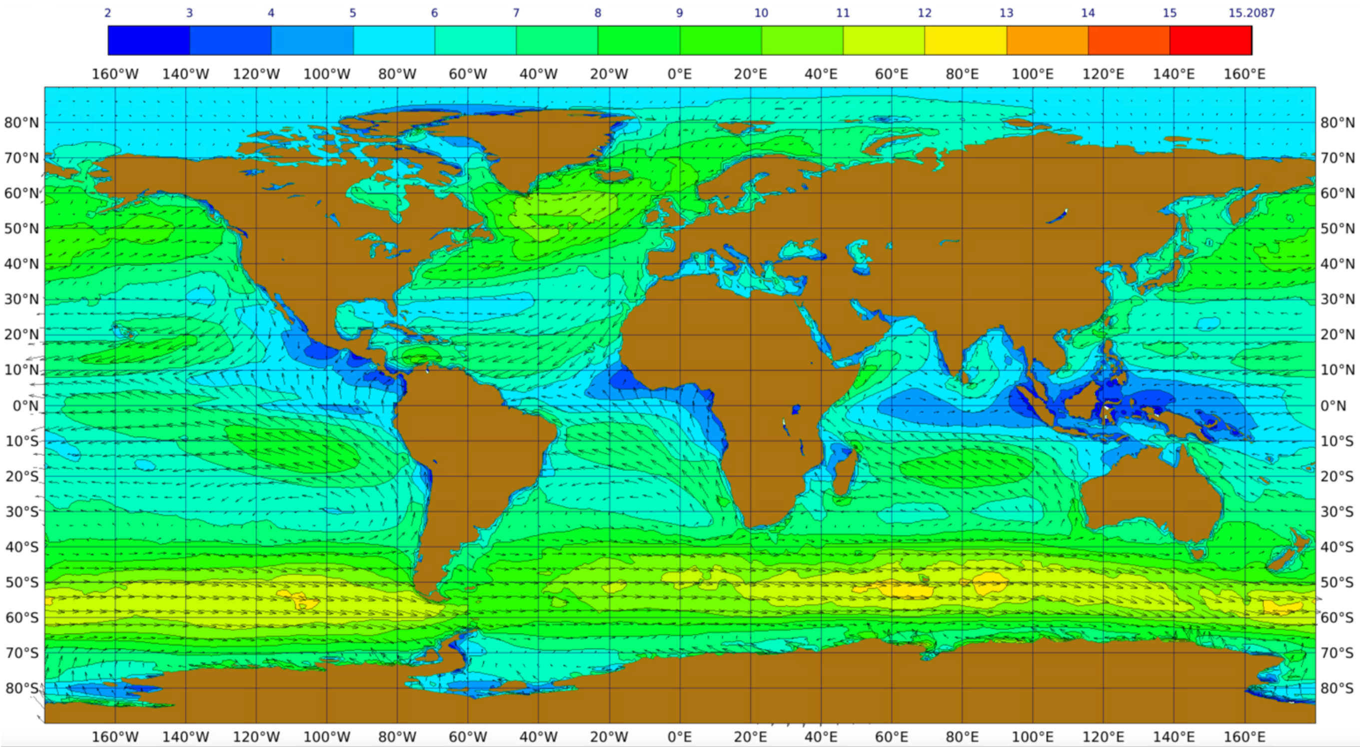

Figure 34: The geographical distribution of the IFS mean surface wind over one year (1 August 2021 – 31 July 2022).

9 References

1 Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Schuster, M., Monga, R., Moore, S., Murray, D., Olah, C., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X., TensorFlow: Large-scale machine learning on heterogeneous systems (2015), Software available from tensorflow.org

2 Aires, F, Weston, P, de Rosnay, P, Fairbairn, D. (2021), Statistical approaches to assimilate ASCAT soil moisture information – I. Methodologies and first assessment. Q J R Meteorol Soc. 1823–1852. https://doi.org/10.1002/qj.3997

3 Chollet, F. et al. (2015), Keras. GitHub. Retrieved Sep 22, 2022 from https://github.com/fchollet/keras

4 Cornford, D., Nabney, I. T., and Ramage, G. (2001), Improved neural network scatterometer forward models, J. Geophys. Res., 106(C10), 22331– 22338, doi:10.1029/2000JC000417

5 De Chiara, G. (2013), "Preparation for the assimilation of ASCAT-B Scatterometer winds," ECMWF Research Department Memorandum, RD13-294.

6 De Chiara, G., S. Abdalla, L. Isaksen, S. Healy (2022), "Characterization of ocean surface roughness in NWP", Task 4 report, Contract No. EUM/CO/18/4600002207/SLi

7 De Chiara, G., S. English, P. A. E. M. Janssen and J.-R. Bidlot (2016), "ASCAT ocean surface wind assessment," ECMWF Technical Memorandum 776.

8 De Chiara G., L. Isaksen, S. English (2018), "ECMWF Support for EPS/ASCAT – Ocean Wind Assessment", Final Report Contract No. EUM/CO/15/4600001497/JF.

9 Dinnat, E., English, S., Prigent, C., Kilic, L., Anguelova, M., Newman, S., Meissner, T., Boutin, J., Stoffelen, A., Yueh, S., Johnson, B., Weng, F., and Jimenez, C. (2023), "PARMIO: A Reference Quality Model for Ocean Surface Emissivity and Backscatter from the Microwave to the Infrared". Bull. Amer. Meteor. Soc., 104, E742–E748. https://doi.org/10.1175/BAMS-D-23-0023.1

10 Donelan, M. A., and Plant, W. J. (2009), "A threshold for wind-wave growth", J. Geophys. Res., 114, C07012. doi:10.1029/2008JC005238

11 ECMWF (2023a). "IFS Documentation CY48R1 - Part VII: ECMWF Wave Model", ECMWF web page. doi: 10.21957/cd1936d846 , Retrieved Jul 11, 2023. https://www.ecmwf.int/en/elibrary/81373-ifs-documentation-cy48r1-part-vii-ecmwf-wave-model

12 ECMWF (2023b). "Parameter Database", ECMWF web page, Retrieved Nov 16, 2023, https://codes.ecmwf.int/grib/param-db/

13 Geer, A. J., Lonitz, K., Weston, P., et al. (2018), "All-sky satellite data assimilation at operational weather forecasting centres". Q J R Meteorol Soc.; 144: 1191– 1217. https://doi.org/10.1002/qj.3202

14 ECMWF (2020). "Data Handling System", ECMWF web page. Retrieved Sep 22, 2022 from https://www.ecmwf.int/en/computing/our-facilities/data-handling-system

15 Hersbach, H. (2007), "Preparation for assimilation of surface-wind data from ASCAT at ECMWF," ECMWF Research Department Memorandum R60.9/HH/075.

16 Hersbach, H. (2008), "CMOD5-N: A C-band geophysical model function for equivalent neutral wind," ECMWF Technical Memorandum 554.

17 Isaksen, L. and Stoffelen, A. (2000), "ERS scatterometer wind data impact on ECMWF's tropical cyclone forecasts", IEEE Transactions on Geoscience and Remote Sensing, 38(4), 1885-1892.

18 Isaksen, L. and Janssen, P. A. E. M. (2004), "Impact of ERS scatterometer winds in ECMWF's assimilation system", Q. J. R. Meteorol. Soc., 130, 1793-1814.

19 Janssen, P. A. E. M., Wallbrink, H., Calkoen, C. J., van Halsema, D., Oost, W. A., and Snoeij, P. (1998), "VIERS-1 scatterometer model", J. Geophys. Res., 103(C4), 7807– 7831. doi:10.1029/97JC02911

20 Laloyaux, P., M. Balmaseda, D. Dee, K. Mogensen, P. A. E. M. Janssen (2016), "CERA: A coupled data assimilation system for climate reanalysis", Q. J. R. Meteorol. Soc., 142, 65-78.

21 Stoffelen, A. and Anderson, D. (1997), "Ambiguity removal and assimilation of scatterometer data". Q. J. R. Meteorol. Soc., 123: 491-518. https://doi.org/10.1002/qj.49712353812

22 Stoffelen, A., J. A. Verspeek, J. Vogelzang and A. Verhoef (2017), "The CMOD7 Geophysical Model Function for ASCAT and ERS Wind Retrievals", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 5, pp. 2123-2134. doi: 10.1109/JSTARS.2017.2681806

23 Thépaut, J., Hoffman, R. N., & Courtier, P. (1993), "Interactions of Dynamics and Observations in a Four-Dimensional Variational Assimilation", *Monthly Weather Review*, *121*(12), 3393-3414. Retrieved Sep 16, 2022, from https://journals.ametsoc.org/view/journals/mwre/121/12/1520-0493_1993_121_3393_iodaoi_2_0_co_2.xml