for publication in *Ocean Dynamics*

Predicting significant wave height with artificial neural networks in the South Atlantic Ocean: a hybrid approach

Paula Marangoni Gazineu Marinho Pinto[1*], Ricardo Martins Campos[3,4], Marcos Nicolas Gallo[2] and Carlos Eduardo Parente Ribeiro[1]

[1*]Ocean Engineering Program, COPPE, LIOc, Federal University of Rio de Janeiro, Athos da Silveira Ramos Avenue, Technology Center, I104e, University City, Rio de Janeiro, 21941-611, Rio de Janeiro, Brazil.

[2]Ocean Engineering Program, COPPE, LDSC, Federal University of Rio de Janeiro, Athos da Silveira Ramos Avenue, Technology Center, I100, University City, Rio de Janeiro, 21941-611, Rio de Janeiro, Brazil.

[3]Cooperative Institute for Marine and Atmospheric Studies (CIMAS), University of Miami, 4600 Rickenbacker Causeway, Miami, 33149, Florida, USA.

[4]NOAA, Atlantic Oceanographic and Meteorological Laboratory (AOML), 4301 Rickenbacker Causeway, Miami, 33149, Florida, USA.

*Corresponding author(s). E-mail(s): paulamarangoni95@gmail.com;
Contributing authors: ricardo.campos@noaa.gov; marcosgallo@oceanica.ufrj.br; cparente@globo.com;

Abstract

Accurate simulations of significant wave height (Hs) are extremely important for the safety of navigation, port operations, and oil and gas exploration. Thus, accurate forecasts of Hs are essential for accident

prevention and maintenance of services vital to the economy. Considering the limitations of traditional numerical modeling, such as the typical model underestimation of Hs under severe conditions, forecasting Hs using artificial neural networks is a promising method and a complementary approach. In this study we develop a post-processing model using Long Short-Term Memory (LSTM) algorithm to improve outputs from the numerical model WAVEWATCH III (WW3) at Santos Basin, Brazil. The hybrid scheme is focused on the simulations of 1, 2, 3 and 4-day residues (difference between observations and WW3) using measurements from a local wave buoy moored in deep water. The results of the hybrid model (WW3+LSTM) shows a better performance compared with WW3, being capable of better representing the peak of the events and storms. On average, the gains from using WW3+LSTM reach 3.8% in Correlation Coefficient (CORR), 14.2% in Bias (BIAS), 10.2 % in Root Mean Squared Error (RMSE), and 10.7% in Scatter Index (SI). The hybrid model developed allows high-skill forecasts to be carried out on large domains and through longer horizons.

Keywords: Artificial Neural Network, Long Short-Term Memory, Significant Wave Height, Forecast

1 Introduction

Coastal and offshore activities are often affected by extreme waves that can cause great economic and social losses. For this reason, an accurate prediction of significant wave height (Hs), in particular for port and oil industries, is of vital importance. Currently, several numerical models are used in wave forecasting. However, under severe weather conditions, models usually underestimate the Hs. It occurs mainly because of the underestimation of the wind fields provided as input (Cardone et al, 1996; Caires and Sterl, 2005; Cavaleri, 2009; Campos and Soares, 2016; Kaiser et al, 2022).

Due to the large amount of data assimilated and the calculations performed, these numerical models demand a high computational cost, long processing time (Etemad-Shahidi and Mahjoobi, 2009; Wang et al, 2018) and accurate bathymetry data (Browne et al, 2007). In this context, forecasting Hs using artificial neural networks is a potentially good alternative, especially in scenarios where a fast and accurate forecast is required (Fan et al, 2020).

Several works using artificial neural networks and other machine learning techniques for wave forecasting have already been developed over the years (Deo and Naidu, 1998; Deo et al, 2001; Etemad-Shahidi and Mahjoobi, 2009; Wang et al, 2018; Londhe et al, 2016; Dixit and Londhe, 2016; Fan et al, 2020; Hu et al, 2021; Minuzzi and Farina, 2022; Sun et al, 2022). Londhe et al (2016) use the error between numerical wave model prediction and observations as input to an artificial neural network. The approach that applies this correction

to model prediction improves by 10% and 22% the correlation results compared with observed data for the buoys located on the south coast of India.

Campos et al (2019) simulated the residue given by the difference between the target value and the average ensemble Hs from the Global Wave Ensemble Forecast System - (GWES). The authors used neural networks to simulate this residue in Gulf of Mexico, United States. They observed that this technique presents better results for long-term forecasting (over 4 days).

On the Brazilian coast, Campos and Guedes Soares (2016) developed a hybrid model to reduce the bias of Hs hindcasts. The authors concluded that the neural network model has the best results reducing the final bias from 0.13 to 0.06 meters and the scatter index from 0.13 to 0.03.

In recent years, Long Short-Term Memory (LSTM) neural networks, have been gaining prominence especially in wave prediction (Fan et al, 2020; Hu et al, 2021; Jörges et al, 2021; Minuzzi and Farina, 2022). The study performed by Fan et al (2020), compares Hs prediction using LSTM with other machine learning algorithms. The authors prove that LSTM performs better compared to the other algorithms in the 1 and 6-hour forecasts. In addition, Fan et al (2020) developed a hybrid SWAN-LSTM model (numerical wave model + LSTM) that applied to a buoy was able to improve the standard numerical wave model result by 65%. Minuzzi and Farina (2022) used LSTM to predict Hs in different regions of Brazilian coast and suggested that the neural network could be used as an alternative to numerical models.

Considering the successful results obtained in many studies and the fact that neural networks are an excellent universal mapper (Hornik, 1991), a simulation of the 1, 2, 3 and 4 day residues (difference between Hs observations and Hs of numerical model) was performed using LSTM neural network (Hochreiter and Schmidhuber, 1997). For the present study, a buoy located in the Santos basin was used. The result obtained by LSTM was added to the Hs result of the WAVEWATCH III (WW3) numerical model (Tolman et al, 2014). This new hybrid model (WW3+LSTM) was then analyzed in terms of improvements compared against the initial wave model simulation.

2 Methods

2.1 Study area and data

The Santos basin is located on the southeast coast of Brazil. In this area there are many oil exploration platforms and an intense circulation of ships. The wave climate of the region is characterized by southeast quadrant waves with Hs around 2 meters (Nascimento, 2013). In addition, the region is influenced by South Atlantic Subtropical Anticyclone (SASA), generally associated with a calm weather situation. Severe weather conditions are associated with cyclones and frontal systems. During this condition, waves and more intense southwest quadrant winds are observed (Nascimento, 2013).

The observed data from the Santos buoy (SA) were used in this study. The buoy belongs to the Brazilian National Program (PNBOIA) of the Brazilian

Navy (Fig. 1). For the used period, the Santos buoy was located approximately at latitude 25°16' South and longitude 044°55' West, at a depth of 200 meters. A 3-Meter Axys buoy was used. Its dimension is 3.4 meters in diameter and 1500 Kg. The observations are available with a hourly frequency and the buoy is currently inoperative.

Besides these observed data, the results of the WW3 numerical model were used. The model forecasts of 1, 2, 3 and 4 days were used in order to generate the Hs residual data series. The model has an hourly frequency of availability with a forecast horizon of 10 days. As wind forcing for WW3, the results of the National Centers for Environmental Prediction (NCEP) atmospheric model Global Forecast System - GFS were used. The GFS has a spatial resolution of 0.25° and temporal resolution of 3 hours.

The WW3 version 4.18 (Tolman et al, 2014) was chosen. The modelling was performed with two nested grids. The first with 1° spatial resolution (Global) and the second with 0.25° (South Atlantic). Moreover, the ST4 package (Ardhuin et al, 2010) with Betamax of 1.52 was used.

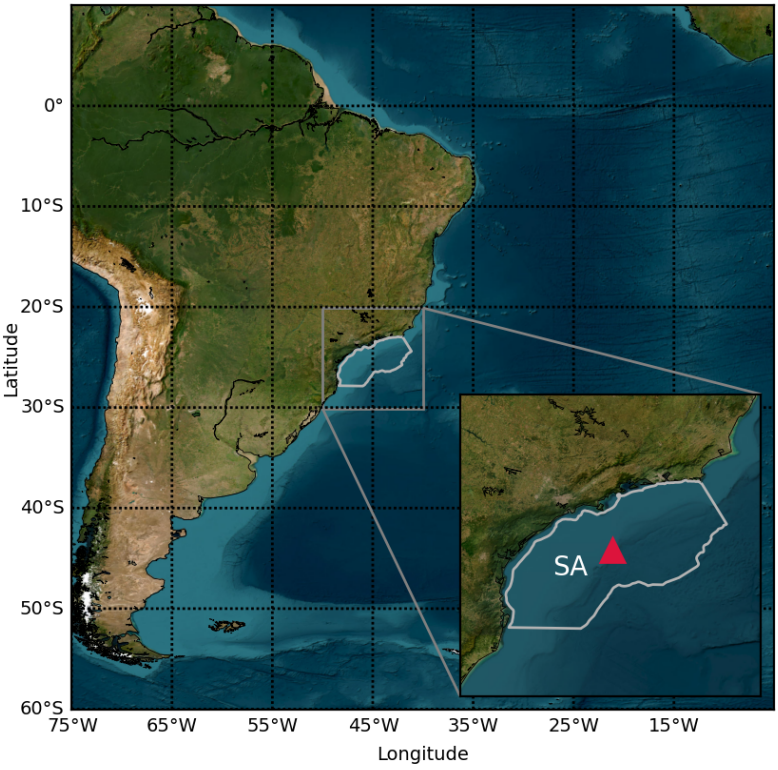

Fig. 1 Location of the study area on the southeast coast of Brazil. The triangle in red within the Santos basin (demarcated area) represents the Santos buoy point, located at 25°16' South and 044°55' West, at a depth of 200 meters.

2.2 Neural Networks

Neural networks were developed by McCulloch and Pitts in 1943 and this technique model how the brain performs a specific task or function of interest (Haykin, 2009). They are an extremely useful tool for mapping complex non-linear relationships, such as those that occur in earth systems (Krasnopolsky, 2013).

Artificial neural networks have a set of processing units (artificial neurons) interconnected by several connectors (artificial synapses) (Silva et al, 2010). It is possible to mathematically define them by the equations 1 and 2.

$$u = \Sigma_{i=1}^{n} w_i * x_i - b \tag{1}$$

$$\tilde{y} = g(u) \tag{2}$$

In the equations, x_i represent the network inputs, w_i are the weights (artificial synapses), b are the bias, u represents the activation potential, Σ are the linear combiner, $g(.)$ are the activation function, and $\tilde{y}$ denotes the output of the neural network.

The simplest neural network, the Multilayer Perceptron (MLP), for example, consists of an input layer, hidden layers and an output layer. Although simple neural networks (with one or two hidden layers) are very effective, for solving more complex problems, Deep Learning usually perform better due to the many hidden layers (Goodfellow et al, 2016).

LSTM networks are one of the multilayer hidden networks used in Deep Learning. They are a type of recurrent neural network that can be used for the prediction of sequential data or time series. They are capable of learning long-term dependencies and are not impaired by very long intervals (Olah, 2015). LSTM has a differentiated architecture (Fig. 2), with a chain structure and instead of neurons, it has memory units called cells.

The process happens as follows: information enters the cells (C_{t-1}) and goes forward to the output (C_t), through a path that is considered the main path of the LSTM. However, along this path, the information is controlled by structures called "gates". They are responsible for determining whether the information will be released (forget gate - Eq. 3), added (input gate - Eq. 4) or read (output gate - Eq. 5) into memory.

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \tag{3}$$

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \tag{4}$$

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \tag{5}$$

Each gate in the network consists of a sigmoid function (σ) (Eq. 6) that has outputs between 0 and 1. Where 0 means that the information will be retained and 1, that the information can follow through.

$$\sigma = \frac{1}{1 + e^{-x}} \tag{6}$$

The information initially enters the forget gate (f_t) through x_t which represents the input of the current moment and h_{t-1}, is the input of the previous result. The output of this gate is multiplied by C_{t-1}. The next step is to check what new information will be added to the memory.

In this step, the input gate (i_t) will decide which information will be updated. Then, a new layer, this time using the hyperbolic tangent activation function (Eq. 7), creates a vector with possible new candidates to be added ($\hat{C}_t$) (Eq. 8).

$$tanh(x) = \frac{sinh(x)}{cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{7}$$

$$\hat{C}_t = tanh(W_c * [h_{t-1}, x_t] + b_c) \tag{8}$$

Thus, the result added in memory is a scalar product between i_t and $\hat{C}_t$. This result is updated in the variable C_t. Where the previous information (C_{t-1}) is multiplied by the forget gate (f_t) and the new information is added in memory (Eq. 9).

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \tag{9}$$

Finally, it is decided which information is going to be the output of the network. In this step, the information passes through the output gate (o_t) described in Eq. 5. The values go to the next memory unit (h_t) from the calculation described in Eq. 10. Thus, the LSTM network only maintains the information it considers most relevant to the learning process. The variables W, x and b, shown in the equations represent, respectively, the values of the weights, input and bias that are learned throughout the backpropagation training.

$$h_t = o_t * tanh(C_t) \tag{10}$$

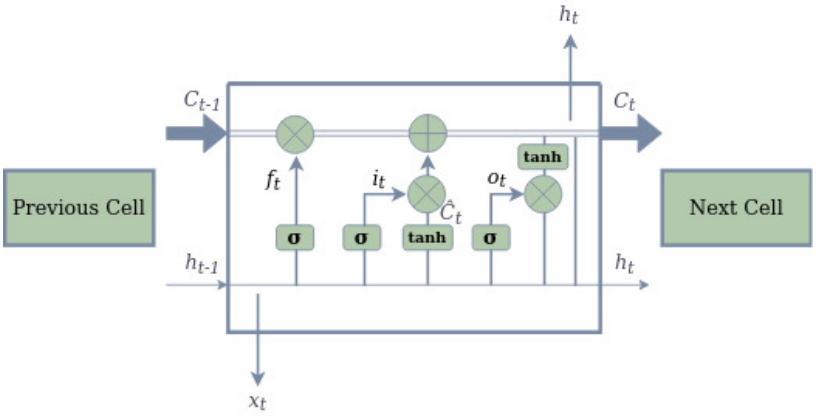

Fig. 2 LSTM network architecture. The figure outlines the processes that occur inside the memory unit (cell). It is possible to observe the input (i_t), forget (f_t) and output (o_t) gates. The LSTM main path starts at the variable C_{t-1} and ends at C_t after passing through the gates.

2.3 Preprocessing and selection of input variables

In data preprocessing, the outliers and missing data were removed from dataset. In addition, the normalization of the input data was performed (placing them between 0 and 1), followed by denormalization after neural network calculations.

In the selection of input variables, Hs observations from the buoy and WW3 model results from 0-23h (Day 1), 24-47h (Day 2), 48-71h (Day 3) and 72-95h (Day 4) were subtracted to constitute the residual data series. Day 1 refers to 1-day simulation, and so on. The period of data was from 2016 to April 2017.

Figure 3 illustrates the residual data series. Observing this figure it is possible to see the high level of noise associated with the residue. Excess noise can cause a risk of overfitting the neural network (Krasnopolsky, 2014) and to avoid this problem a filter (moving average) was applied to the data. In this way, the 3, 6 and 12 hour filters were tested and the best result was obtained with the 3 hour filtering. This filter improved the 1 day forecast by up to 13%, while the 6-hour and 12-hour filtering show an improvement of around 9.7% and 8%, respectively.

The figure also presents the histograms of this data series for 1, 2, 3, and 4 days, respectively. It can be seen in the histograms that the data is concentrated in the lowest residue values, especially the negative residue (model result minus data from buoy). With this behavior, it is possible to note that, in general, there is an underestimation of the numerical model relative to the buoy in all the forecast horizons analyzed.

Table 1 shows statistical metrics of mean, variance and skewness calculated from the residue for each forecast horizon. It can be observed that for all horizons, negative means and positive asymmetry were obtained, characterizing a positive asymmetric distribution. Thus, the results of the table corroborate

what was observed in the histograms (Fig. 3), where the model result tends to underestimate the buoy at lower residue values. Finally, the low variance values observed, increase with the forecast horizon due to the greater dispersion of the data.

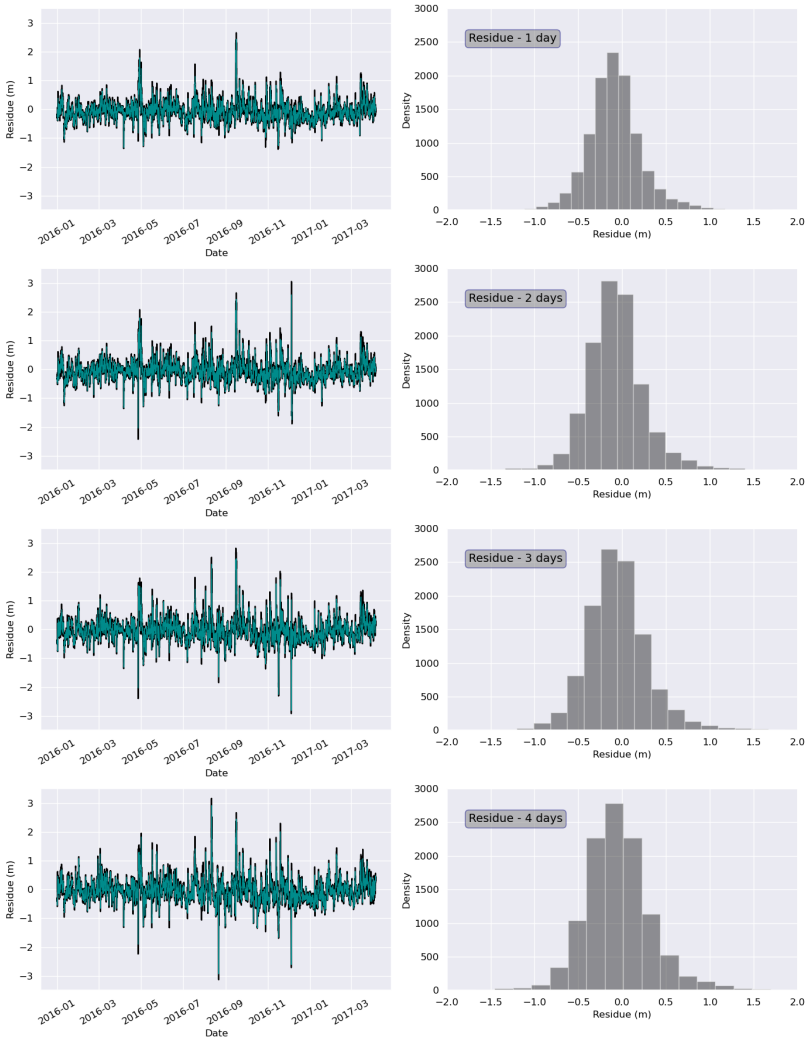

Fig. 3 Residual data series for the 1, 2, 3 and 4 days forecast horizon. The left figure represents the time series of the residue (in black) and the filtered series (in blue). The figure on the right illustrates the histogram for the residue series.

The metocean variables available at the SA buoy were previously analyzed and preprocessed to make it possible to correlate them with the network target variable (residue). Figure 4 illustrates the percentage of missing data per

meteoceanographic variable from Santos buoy. It can be seen that some of these variables have many missing data, which can hinder the correlation analysis. Considering that time series with many periods without data could compromise the neural network training, 30% of missing data was chosen as limit. Thus, variables with missing data above this percentage were excluded from the correlation. Moreover, the discarded variables (above 30% of missing data, such as dew point temperature, humidity and sea surface temperature) did not have significant relation to wave generation and the consequent improvement of the model.

Table 1 Statistics of mean, variance and asymmetry for the 1, 2, 3 and 4 day simulation calculated from the residue.

Statistics	Day 1	Day 2	Day 3	Day 4
Mean (m)	-0.078	-0.069	-0.053	-0.041
Variance (m)	0.107	0.125	0.150	0.168
Skewness	1.027	0.842	0.652	0.699

Fig. 4 Percentage of missing data on meteoceanographic variables available in the Santos buoy. Only data below 30% of missing data were used in correlation wich includes Wind Speed (Wspd), Wind Direction (Wdir), Gust, Air Temperature (Atmp), Pressure (Pres), Significant Wave Height (Hs), Maximum wave height (Wmax), Peak Period (Tp), Mean Wave Direction (Mwd) and Spread. The variables: Dew Point Temperature (Dewp), Humidity (Humi), Sea Surface Temperature (SST), Current Velocity in the cell 1, 2 and 3 (Cvel1, Cvel2, Cvel3) and Current Direction in the cell 1, 2 and 3 (Cdir1, Cdir2, Cdir3) were excluded from the correlation.

The data period used for training and testing the neural network was established based on the availability of the observations and the WW3 model run. The period of 2016 was selected for network training and 2017 from January

to early April for testing. Totaling 80% of the data for training and 20% for testing. Cross-validation of the data was not performed, since for the type of network used (LSTM) the order of events should be preserved.

Figure 5 illustrates the correlation between meteoceanographic variables from the SA buoy (observations) and the 1-day residue of 2016. The correlation results between the variables from the buoy and the 2, 3 and 4 days residue were similar, so the results were not displayed.

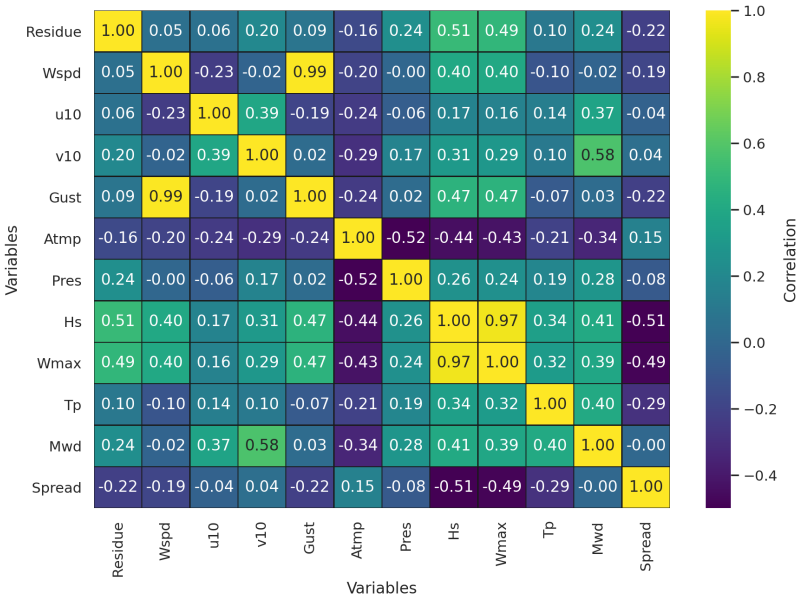

Fig. 5 Correlation between meteoceanographic variables from the SA buoy and the 1-day residue of 2016. The correlation results between the variables from the buoy and the 2, 3 and 4 days residue were similar.

The wind direction variable was decomposed into the zonal (u10) and meridional (v10) components as it is a vector variable. It can be seen in figure 5, that the highest correlation was obtained with Hs. This is because the residue is derived from the significant wave height of the buoy and the model. The other variables show lower correlations with the residue.

In general, since the correlation values obtained are very low, it was decided to use the Hs of 1 hour prior. Also the direction (v10) and wind speed of 3 hours prior as was done by Fan et al (2020) for Hs prediction. Moreover, the variables wind speed and wind direction are the ones that physically have the most influence on wave generation, in addition to the Hs of previous 1 hour (Holthuijsen, 2007; Hashim et al, 2016). The wind is responsible for transmitting energy to the wave, influencing its generation and growth.

The meridional wind direction (v10) was chosen due to its higher correlation with the residue when compared to the zonal direction (u10) (Fig.

5). This demonstrates that the meridional wind direction can have a greater impact on wave generation in the Santos Basin region, forming wind fetches. This correlation result is coherent, since winds from south to north (meridional direction) are characteristic of cyclones and frontal systems coming from polar areas and cyclogenesis regions to the south (Reboita et al, 2010; Campos and Guedes Soares, 2016).

In summary, the neural network inputs are wave heigth (Hs) of 1 hour before, meridional wind direction (v10) and wind speed of 3 hours before and the output is the residue (difference between measured and modeled Hs).

2.4 LSTM Neural Network Configuration

After several tests, the neural network configuration that guaranteed the best performance was obtained with 100 epochs, Adam type optimizer and 300 batch-size that is the frequency at which the error will be calculated and the weights updated. A total of four LSTM layers with 100 neurons each and a dense output layer with one linear neuron were used. After each LSTM layer a dropout layer was used, zeroing 30% of the inputs, in order to minimize the risk of overfitting.

Once these hyperparameters were defined, the LSTM neural network was trained. The training was executed 30 times, considering the random character of the network. This randomness is what guarantees a good learning of the function that is being approximated for the problem. Thus, the final result of the error metrics was the average of these 30 repetitions.

2.5 Error Metrics

To evaluate the results obtained from the network simulation in comparison with the true observed results, the following error metrics were used: Correlation Coefficient (CORR), Bias (BIAS), Root Mean Squared Error (RMSE) and the Scatter Index (SI) (Mentaschi et al, 2013) (Eqs. 11-14). The BIAS and RMSE indicate the systematic error, while the SI indicates the scattered error.

$$CORR = \frac{\Sigma_{i=1}^{n}(p_i - \overline{p}) \times (o_i - \overline{o})}{\sqrt{\Sigma_{i=1}^{n}(p_i - \overline{p})^2} \times \sqrt{\Sigma_{i=1}^{n}(o_i - \overline{o})^2}} \tag{11}$$

$$BIAS = \frac{1}{n}\Sigma_{i=1}^{n}(p_i - o_i) \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(p_i - o_i)^2} \tag{13}$$

$$SI = \sqrt{\frac{\Sigma_{i=1}^{n}[(p_i - \overline{p}) - (o_i - \overline{o})]^2}{\Sigma_{i=1}^{n} o_i^2}} \tag{14}$$

Where n is the total number of samples, i is the instant of time, p_i represent the predicted values and o_i are the observations. The $\overline{o}$ and $\overline{p}$ represent mean values. For the calculation of error metrics, observations were filtered with a 3-hour moving average.

3 Results

3.1 Test Simulations

In this section the results of the test simulations are presented. Figure 6 shows the results of the WW3 model correction from the simulation performed by the LSTM neural network (WW3+LSTM) for each forecast horizon (1, 2, 3 and 4 days, respectively). As well as the comparison with the observations and the WW3 model forecast.

For the 1-day simulation (Fig. 6), it is observed that the WW3+LSTM model perform better the peak of the highest Hs events. In the event on March 15th an improvement of approximately 19.6% is observed at the peak of the event (in the morning) using the WW3+LSTM model. Also on March 20th, there is a gain of 17.1% in the representation of another Hs peak during the evening with the WW3+LSTM model. Then, the underestimation of the WW3 model without the neural network correction is evident in many events. Considering the other forecast horizons, it is noted that both the WW3 model and the WW3+LSTM model slightly attenuate the peak of the higher Hs events as the forecast horizon increases.

Tables 2 and 3 present in error metrics the results of the WW3 and WW3+LSTM models compared to observations. They also indicate the percentage gains of using WW3+LSTM model compared to WW3. It is important to note that gains, when positive, show that the WW3+LSTM model improved the results compared to the original WW3 model. On the other hand, negative gains (or losses), indicate that the WW3+LSTM model showed no improvement in the result.

It is possible to see in the tables that most of the gains are positive. Except for the BIAS of 1-day which has a worse result with the methodology (loss of 7%). The largest gains are observed in the BIAS, above 3-day forecast, where the gains in using the LSTM to correct the forecast reach almost 48% in the 3-day forecast. The other significant gains occurred in RMSE in the 1-day and 2-day forecasts (around 14%).

In general, high correlation values are observed at all forecast horizons (around 0.8). Despite the low gains observed in the CORR, a larger gain is observed in the 1-day forecast (around 6%). Regarding SI, low values and significant gains are observed around 17% in 1-day forecast and around 14% in 2-day forecast. It is also noteworthy that the 4-day horizon showed the smallest gains in most metrics, except for BIAS, where there was an approximate 18% gain.

Regarding the computational costs of neural network, using the parameters chosen in methodology, the costs are extremely low and are restricted to the training process. The simulation was performed using dual-core processors (1.8 GHz) taking approximately two minutes, in Python language. Since the training has been repeated 30 times, the total period takes approximately 1 hour (training) while the prediction, a few seconds.

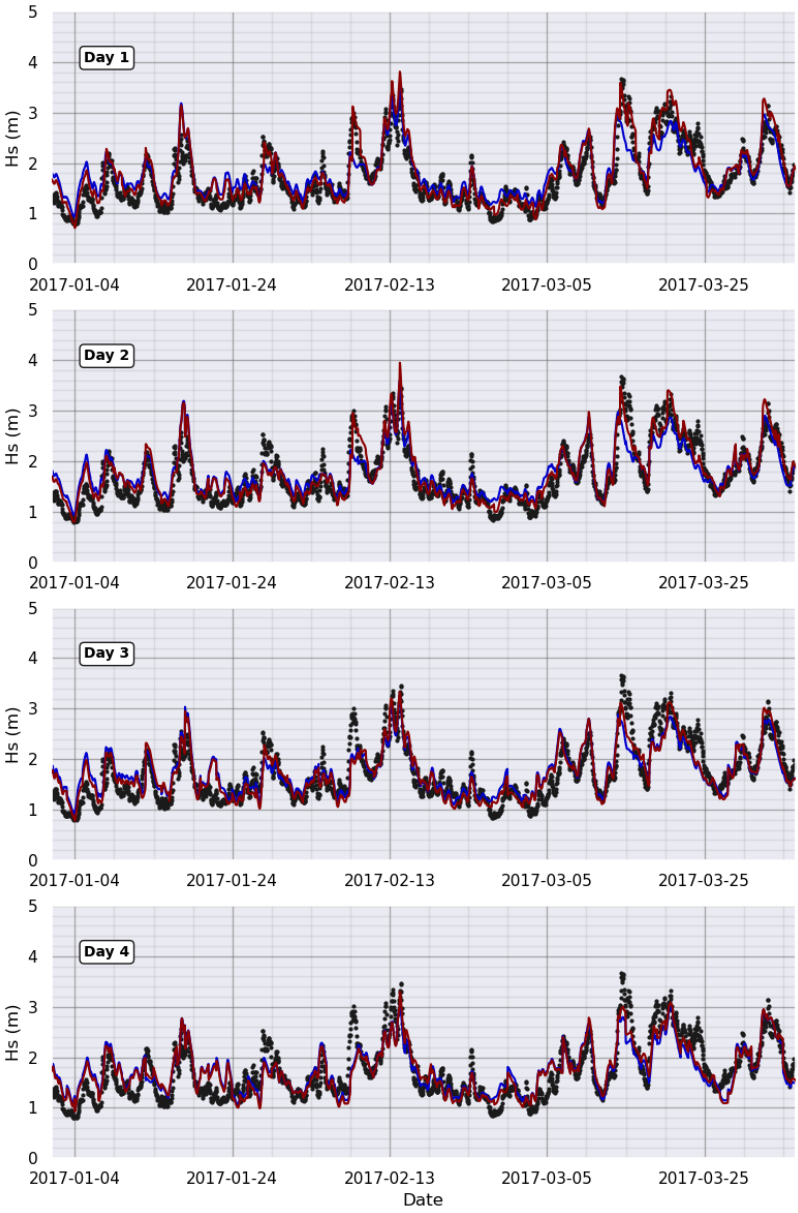

Fig. 6 Comparison of the Hs observations (points in black), with the WW3 model results (solid line in blue) and the model result corrected with the neural network (WW3+LSTM) (solid line in burgundy) for the days of simulation.

Table 2 Hs forecast error metrics for the WW3, WW3+LSTM model and the percentage gains for the 1 and 2-day simulations.

	Day 1			Day 2		
Error Metrics	WW3	WW3+LSTM	Gains(%)	WW3	WW3+LSTM	Gains(%)
CORR	0.875	0.925	5.714	0.864	0.903	4.514
BIAS (m)	-0.083	-0.089	-7.229	-0.067	-0.066	1.493
RMSE (m)	0.290	0.247	14.828	0.296	0.255	13.851
SI	0.153	0.127	16.993	0.159	0.136	14.465

Positive gains - enhanced performance with the WW3+LSTM model.
Negative gains - inferior performance with the WW3+LSTM model.

Table 3 Hs forecast error metrics for the WW3, WW3+LSTM model and the percentage gains for the 3 and 4-day simulations.

	Day 3			Day 4		
Error Metrics	WW3	WW3+LSTM	Gains(%)	WW3	WW3+LSTM	Gains(%)
CORR	0.836	0.858	2.632	0.817	0.836	2.326
BIAS (m)	-0.046	-0.024	47.826	-0.028	-0.023	17.857
RMSE (m)	0.315	0.292	7.302	0.328	0.312	4.878
SI	0.172	0.161	6.395	0.181	0.172	4.972

Positive gains - enhanced performance with the WW3+LSTM model.
Negative gains - inferior performance with the WW3+LSTM model.

3.2 Underestimation Event

Two events occurred in March, 2017 (starting on the 14th and 18th) are highlighted to emphasize the underestimation of WW3 when compared to observations. Figure 7 illustrates the underestimation events for each forecast horizon.

It can be seen that for the March 14th event in the 1-day forecast, the WW3 model failed to reproduce the largest peak of the event, being able to highlight only the first peak. Unlike the WW3+LSTM model, which in addition to reproducing the first peak, also predicts the two following peaks more accurately, as registered in the observations. It can be observed that after the 2-day forecast, the WW3+LSTM model no longer highlights the third peak, showing only the second (which registers the highest Hs).

In the event that started on March 18th, we observe again the underestimation of the WW3 model. In this case, WW3+LSTM model perform better, since is able to highlight the highest peak of the event (around 3.5 meters of Hs).

Overall, the WW3+LSTM model has a much closer forecast to what was observed during this event, improving by about 19.6% the reproduction of the Hs peak on March 14th event and by 17.1% on March 18th event in the 1-day simulation. It is also observed that as the forecast horizon increases, the models

tend to attenuate the peaks. Despite this, it is observed that the WW3+LSTM model is able to indicate the peak of this event even in the 4-day forecast.

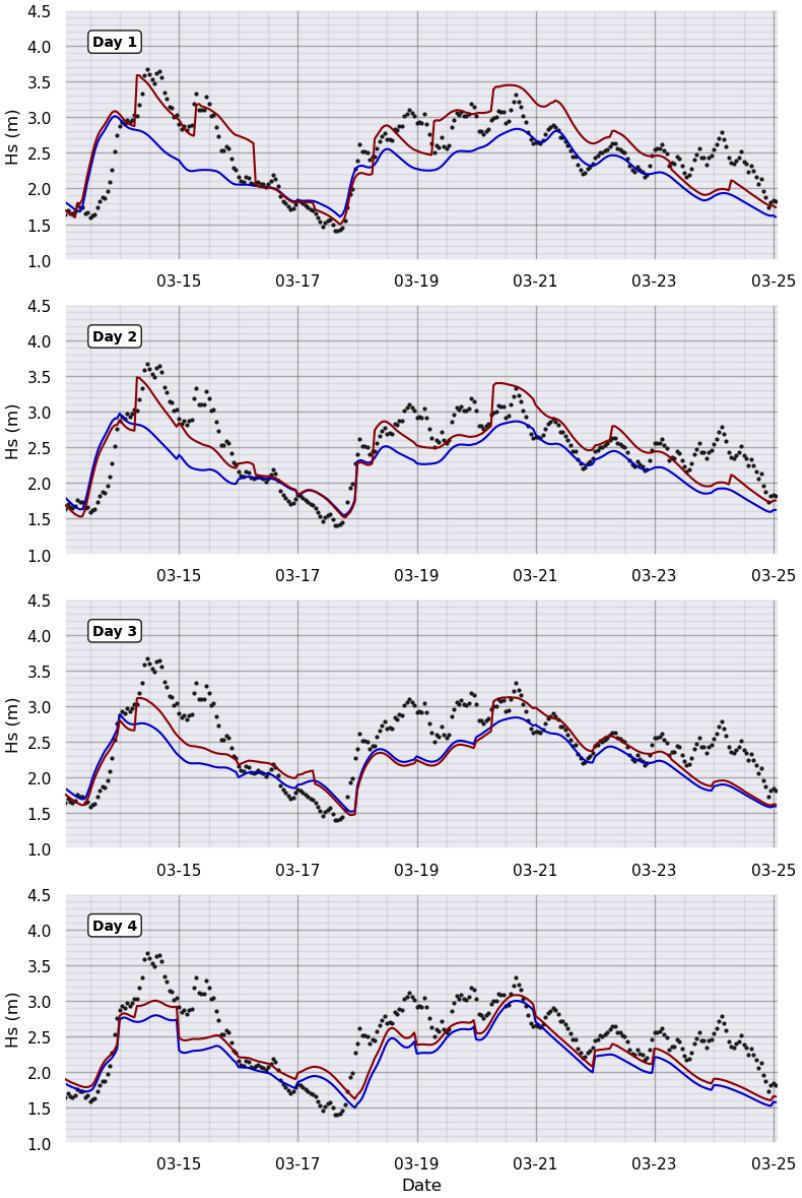

Fig. 7 Underestimation events by WW3 that occurred in March 2017. Hs observations (points in black), WW3 model results (solid line in blue) and model results corrected with the neural network (WW3+LSTM) (solid line in burgundy) for the days of simulation.

3.3 Overestimation Event

Figure 8 shows some events that occurred in January, 2017 that highlight the overestimation of the WW3 model at different forecast horizons (1 to 4 days respectively). In this section the events occured on January 1st and 5th are analyzed. In these events, it is observed that the WW3 and WW3+LSTM models overestimate the observations.

In the January 1st event, a difference of 0.4 m (0.3) is verified between the WW3 model (WW3+LSTM model) and observations. For 1-day simulation, the peak of Hs in January 5th event is 2.02 m in the WW3 model, 1.86 m in the WW3+LSTM model, while the true value (observation) is 1.51 m. Although both models overestimate the observations, the WW3+LSTM results are in better agreement with the buoy measurements. In the other forecast horizons, the same pattern of overestimation could be observed, with emphasis on the fact that in the 4-day horizon there are smaller differences between the WW3 and WW3+LSTM model forecasts.

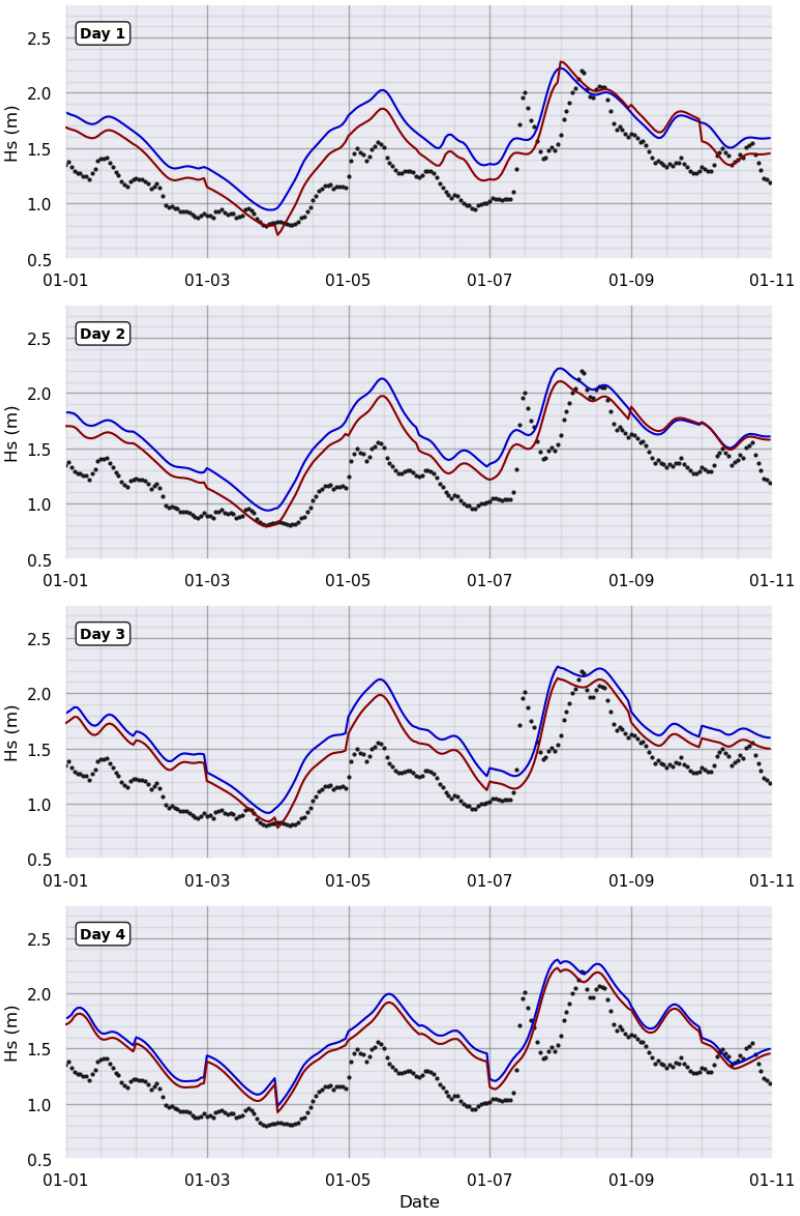

Fig. 8 Overestimation events by WW3 that occurred in January 2017. Hs observations (black dots), WW3 model results (solid line in blue) and model results corrected with the neural network (WW3+LSTM) (solid line in burgundy) for the days of simulation.

4 Discussion

Regarding the choice of input variables, illustrated by the correlation between variables (Fig. 5), it is highlighted that the variables in general have a low correlation with the residue. This is because the random component of the Hs time series is present in it. Thus, the presence of this random component is more evident in the series when the residue is extracted, making its correlation with the other variables considerably lower.

In the results obtained from the test simulations (Figs. 6), it is observed that, in general, the numerical model WW3 represents relatively well the observations - except a few specific extreme events. However, overall, it is observed from the time series that the corrected model (WW3+LSTM) shows a better performance in comparison with the WW3 model, improving especially the peak of extreme wave events. The neural network is able to accurately represent the peak of these events, precisely because it is an universal mapper (Hornik, 1991).

A loss in the accuracy of the statistical metrics is also observed with increasing forecast horizon. This fact can be explained by the chaotic behavior of the coupled atmosphere-ocean-wave system that limits the predictability of this complex system (Lorenz, 1963). Therefore, it is expected that the BIAS, RMSE and SI increase and CORR decrease with increasing horizon as observed in the tables 2 and 3.

Analyzing the tables 2 and 3, it is noted that, as the values are mostly positive, there were only gains in using the WW3+LSTM model for the simulation of Hs. This means that the WW3+LSTM model has an enhanced performance compared to the original WW3 numerical model. This fact supports previous studies (Campos and Guedes Soares, 2016; Londhe et al, 2016) in which hybrid models (numerical model and neural network) overlap in accuracy to conventional numerical models.

Some specific events throughout the time series of the tests are highlighted, where cases of both underestimation and overestimation of the observations by WW3 are observed. In most events the difference between the observations and the model could be corrected by the WW3+LSTM model, as observed in Fig. 7. For the underestimation events that occurred in March, 2017 it is noted that, in general, the WW3+LSTM model satisfactorily reproduces the peak of the events, correcting the underestimation observed in the WW3 model. In the overestimation events that occurred in January, 2017, it is noted that even the corrected WW3+LSTM model are not able to effectively reproduce the observations for the selected events, overestimating them. However, it is worth noting that the overestimation of WW3+LSTM is smaller than that obtained by WW3. Thus, the capacity of the WW3+LSTM model to represent the observations more effectively than the WW3 model stands out, especially in situations of underestimation of the WW3.

According to the results, it is observed that the major contribution of this paper compared to previous studies is the use of LSTM neural networks for prediction of significant wave height using a hybrid approach, combining WW3

results with LSTM post-processing (residue). Due to their unique architecture, as the memory unit for instance, LSTM networks are capable of incorporating information over long periods of time, enabling longer- term forecasts. The present study corroborates with others (Fan et al, 2020; Hu et al, 2021; Minuzzi and Farina, 2022) that LSTM is a suitable neural network for this type of forecast.

The hybrid model developed in the present paper combines the LSTM network with the large scale numerical wave model (WW3), where the output of the neural network is in the form of the residue (difference between measured and modeled Hs). This powerful combination allows forecasts to be carried out on a large scale and with longer time horizons, working as a very effective nonlinear bias correction.

Regarding the South Atlantic Ocean domain, Campos and Soares (2016) developed a hybrid model for forecasting the residue on the Brazilian coast, however the study only corrected the hindcast and was restricted to a simple multilayer perceptron neural network. Recently, Minuzzi and Farina (2022) also used LSTM on the Brazilian coast to performed Hs forecast, but they did not use a hybrid framework including WW3, and they were bounded to 24-hour forecasts whereas the present study, due to the hybrid approach, advanced to forecast leads up to 96 hours.

5 Conclusion

In this work, the impact of using LSTM neural networks to improve the long-term forecast of Hs was evaluated. Overall, it is observed that for the majority of forecast horizons there were gains in using such nonlinear bias correction approach. For example, significant gains of almost 48% in BIAS in the 3-day forecast and about 18% in the 4-day forecast. On average, gains of 3.8% were observed in CORR, 14.2% in BIAS, 10.2% in RMSE, and 10.7% in SI.

Detailed analyses of events showed that WW3+LSTM performs better and is able to obtain results close to the observations. This makes even more evident the advantages of using the correction by the proposed LSTM neural network to increase the accuracy of Hs forecasts. Considering the results obtained, the methodology used in the present study looks very promising - being an efficient and low-cost method for post-processing wave forecasts, with easy implementation. The hybrid model used for forecasting the residue allowed better results that can be replicated on a large scale and in longer forecast time horizons.

Finally, the results suggest that the correction of WW3 model outputs by neural networks (WW3+LSTM) trained with buoy data is able to generate more accurate wave predictions than those made exclusively by WW3. This is mainly due to the fact that neural networks are an universal mapper (Hornik, 1991) and suitable for complex nonlinear problems (Krasnopolsky, 2013), which in the present paper is represented by the residual signal of Hs.

Acknowledgments The authors thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – (Finance Code 001) for the student fellowship of the first author and CNPQ (PQ-2308078/2019-5) for the grant to the third author. The second author is funded by the Cooperative Institute for Marine and Atmospheric Studies (CIMAS), a Cooperative Institute of the University of Miami and the National Oceanic and Atmospheric Administration, cooperative agreement NA20OAR4320472. Authors also thank Oceanographic Instrumentation Laboratory (LIOc) for providing the wave modeling data.

Data Availability The buoy dataset analyzed during the current study is available in the Brazilian National Program repository, [https://www.marinha.mil.br/chm/dados-do-goos-brasil/pnboia-mapa]. The wind dataset (GFS) used in wave modeling is available in the NCEP repository, [https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast]

Declarations

Authors' contributions Paula Marangoni Gazineu Marinho Pinto: conceptualization, methodology, formal analysis and investigation, writing - original draft preparation. Ricardo Martins Campos: conceptualization, methodology, writing - review and editing. Marcos Nicolas Gallo: conceptualization, methodology, writing - review and editing, funding acquisition, supervision. Carlos Eduardo Ribeiro Parente: conceptualization, methodology, funding acquisition, supervision.

Competing interests The authors declare no competing interests.

References

Ardhuin F, Rogers E, Babanin A, et al (2010) Semiempirical dissipation source functions for ocean waves. part i: Definition, calibration, and validation. Journal of Physical Oceanography 40:1917. https://doi.org/10.1175/2010JPO4324.1

Browne M, Castelle B, Strauss D, et al (2007) Near-shore swell estimation from a global wind-wave model: Spectral process, linear, and artificial neural network models. Coastal Engineering 54(5):445 – 460. https://doi.org/https://doi.org/10.1016/j.coastaleng.2006.11.007, URL http://www.sciencedirect.com/science/article/pii/S0378383906001840

Caires S, Sterl A (2005) 100-Year Return Value Estimates for Ocean Wind Speed and Significant Wave Height from the ERA-40 Data. Journal of Climate 18(7):1032–1048. https://doi.org/10.1175/JCLI-3312.1, URL https://doi.org/10.1175/JCLI-3312.1, https://arxiv.org/abs/https://journals.ametsoc.org/jcli/article-pdf/18/7/1032/3796666/jcli-3312_1.pdf

Campos R, Guedes Soares C (2016) Hybrid model to forecast significant wave heights: Proceedings of the 3rd international conference on maritime technology and engineering. MARTECH 2016, Lisbon, Portugal, 4-6 July 2016 pp 1027–1035. https://doi.org/10.1201/b21890-138

Campos RM, Soares CG (2016) Comparison and assessment of three wave hindcasts in the north atlantic ocean. Journal of Operational Oceanography 9(1):26–44. https://doi.org/10.1080/1755876X.2016.1200249, URL https://doi.org/10.1080/1755876X.2016.1200249, https://arxiv.org/abs/https://doi.org/10.1080/1755876X.2016.1200249

Campos RM, Krasnopolsky V, Alves JHGM, et al (2019) Nonlinear Wave Ensemble Averaging in the Gulf of Mexico Using Neural Networks. Journal of Atmospheric and Oceanic Technology 36(1):113–127. https://doi.org/10.1175/JTECH-D-18-0099.1, URL https://doi.org/10.1175/JTECH-D-18-0099.1, https://arxiv.org/abs/https://journals.ametsoc.org/jtech/article-pdf/36/1/113/3416099/jtech-d-18-0099 1.pdf

Cardone VJ, Jensen RE, Resio DT, et al (1996) Evaluation of Contemporary Ocean Wave Models in Rare Extreme Events: The "Halloween Storm" of October 1991 and the "Storm of the Century" of March 1993. Journal of Atmospheric and Oceanic Technology 13(1):198–230. https://doi.org/10.1175/1520-0426(1996)013⟨0198:EOCOWM⟩2.0.CO;2, URL https://doi.org/10.1175/1520-0426(1996)013⟨0198:EOCOWM⟩2.0.CO;2, https://arxiv.org/abs/https://journals.ametsoc.org/jtech/article-pdf/13/1/198/3297784/1520-0426(1996)013 0198_eocowm_2_0_co_2.pdf

Cavaleri L (2009) Wave Modeling—Missing the Peaks. Journal of Physical Oceanography 39(11):2757–2778. https://doi.org/10.1175/2009JPO4067.1, URL https://doi.org/10.1175/2009JPO4067.1, https://arxiv.org/abs/https://journals.ametsoc.org/jpo/article-pdf/39/11/2757/4496396/2009jpo4067_1.pdf

Deo M, Naidu CS (1998) Real time wave forecasting using neural networks. Ocean Engineering 26(3):191 – 203. https://doi.org/https://doi.org/10.1016/S0029-8018(97)10025-7, URL http://www.sciencedirect.com/science/article/pii/S0029801897100257

Deo M, Jha A, Chaphekar A, et al (2001) Neural networks for wave forecasting. Ocean Engineering 28(7):889 – 898. https://doi.org/https://doi.org/10.1016/S0029-8018(00)00027-5, URL http://www.sciencedirect.com/science/article/pii/S0029801800000275

Dixit P, Londhe S (2016) Prediction of extreme wave heights using neuro wavelet technique. Applied Ocean Research 58:241 – 252. https://doi.org/https://doi.org/10.1016/j.apor.2016.04.011, URL http://www.

sciencedirect.com/science/article/pii/S0141118716301067

Etemad-Shahidi A, Mahjoobi J (2009) Comparison between m5 model tree and neural networks for prediction of significant wave height in lake superior. Ocean Engineering 36(15):1175 – 1181. https://doi.org/https://doi.org/10.1016/j.oceaneng.2009.08.008, URL http://www.sciencedirect.com/science/article/pii/S0029801809001905

Fan S, Xiao N, Dong S (2020) A novel model to predict significant wave height based on long short-term memory network. Ocean Engineering 205:107,298. https://doi.org/https://doi.org/10.1016/j.oceaneng.2020.107298, URL http://www.sciencedirect.com/science/article/pii/S0029801820303401

Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT Press, http://www.deeplearningbook.org

Hashim R, Roy C, Motamedi S, et al (2016) Selection of climatic parameters affecting wave height prediction using an enhanced takagi-sugeno-based fuzzy methodology. Renewable and Sustainable Energy Reviews 60:246 – 257. https://doi.org/https://doi.org/10.1016/j.rser.2016.01.098, URL http://www.sciencedirect.com/science/article/pii/S1364032116001283

Haykin SS (2009) Neural networks and learning machines, 3rd edn. Pearson Education, Upper Saddle River, NJ

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Holthuijsen LH (2007) Waves in Oceanic and Coastal Waters. Cambridge University Press, https://doi.org/10.1017/CBO9780511618536

Hornik K (1991) Approximation capabilities of multilayer feedforward networks. Neural Networks 4(2):251–257. https://doi.org/https://doi.org/10.1016/0893-6080(91)90009-T, URL https://www.sciencedirect.com/science/article/pii/089360809190009T

Hu H, van der Westhuysen AJ, Chu P, et al (2021) Predicting lake erie wave heights and periods using xgboost and lstm. Ocean Modelling 164:101,832. https://doi.org/https://doi.org/10.1016/j.ocemod.2021.101832, URL https://www.sciencedirect.com/science/article/pii/S1463500321000846

Jörges C, Berkenbrink C, Stumpe B (2021) Prediction and reconstruction of ocean wave heights based on bathymetric data using lstm neural networks. Ocean Engineering 232:109,046. https://doi.org/https://doi.org/10.1016/j.oceaneng.2021.109046, URL https://www.sciencedirect.com/science/article/pii/S0029801821004819

Kaiser J, Nogueira ICM, Campos RM, et al (2022) Evaluation of wave model performance in the south atlantic ocean: a study about physical parameterization and wind forcing calibration. Ocean Dynamics 72:137–150. https://doi.org/10.1007/s10236-021-01495-4, URL https://doi.org/10.1007/s10236-021-01495-4

Krasnopolsky V (2013) The Application of Neural Networks in the Earth System Sciences, vol 46. Springer Dordrecht, https://doi.org/10.1007/978-94-007-6073-8_6

Krasnopolsky VM (2014) Nn-tsv, ncep neural network training and validation system; brief description of nn background and training software. Office note 478 (National Centers for Environmental Prediction (US)) pp 1–60. https://doi.org/http://dx.doi.org/10.7289/V5QR4V2Z, URL https://repository.library.noaa.gov/view/noaa/6945

Londhe S, Shah S, Dixit P, et al (2016) A coupled numerical and artificial neural network model for improving location specific wave forecast. Applied Ocean Research 59:483 – 491. https://doi.org/https://doi.org/10.1016/j.apor.2016.07.004, URL http://www.sciencedirect.com/science/article/pii/S0141118716300876

Lorenz EN (1963) The predictability of hydrodynamic flow. Trans N Y Acad Sci 25:409–432

Mentaschi L, Besio G, Cassola F, et al (2013) Problems in rmse-based wave model validations. Ocean Modelling 72:53–58. https://doi.org/10.1016/j.ocemod.2013.08.003

Minuzzi FC, Farina L (2022) A deep learning approach to predict significant wave height using long short-term memory. Ocean Modelling p 102151. https://doi.org/https://doi.org/10.1016/j.ocemod.2022.102151, URL https://www.sciencedirect.com/science/article/pii/S1463500322001652

Nascimento MX (2013) Analysis of meteoceanographic conditions in the santos basin region. Master's thesis, Federal University of Rio de Janeiro

Olah C (2015) Understanding lstm networks. URL http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Reboita MS, Gan MA, Rocha RPArd, et al (2010) Precipitation regimes in south america: A literature review. Brazilian Journal of Meteorology 25:185 – 204. https://doi.org/10.1590/S0102-77862010000200004, URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-77862010000200004&nrm=iso

Silva I, Spatti D, Flauzini R (2010) Artificial Neural Networks for Engineering and Applied Sciences. Artliber

Sun D, Huang W, Luo Y, et al (2022) A deep learning-based bias correction method for predicting ocean surface waves in the northwest pacific ocean. Geophysical Research Letters 49(23):e2022GL100,916. https://doi.org/https://doi.org/10.1029/2022GL100916, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GL100916, e2022GL100916 2022GL100916, https://arxiv.org/abs/https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022

Tolman H, Accensi M, Alves JH, et al (2014) User manual and system documentation of WAVEWATCH III version 4.18. U. S. Department of Commerce National Oceanic and Atmospheric Administration National Weather Service National Centers for Environmental Prediction

Wang W, Tang R, Li C, et al (2018) A bp neural network model optimized by mind evolutionary algorithm for predicting the ocean wave heights. Ocean Engineering 162:98 – 107. https://doi.org/https://doi.org/10.1016/j.oceaneng.2018.04.039, URL http://www.sciencedirect.com/science/article/pii/S0029801818305213