Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 11, n. 1, 2025.

Trabalho apresentado no XLIII CNMAC, Centro de Convenções do Armação Resort - Porto de Galinhas - PE, 2024

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Prediction of Significant Wave Heights by an Ensemble of Neural Networks

Felipe C. Minuzzi[1]
UFSM, Santa Maria, RS
Leandro Farina[2]
UFRGS, Porto Alegre, RS

Abstract Due to the chaotic behaviour of the differential equations which model the problem of predicting ocean waves variables, a well know strategy to overcome the difficulties is basically to run several simulations, by for instance, varying the initial condition, and averaging the result of each of these, creating an ensemble. Moreover, in the last few years, considering the amount of available data and the computational power increase, machine learning algorithms have been applied as surrogate to traditional numerical models, yielding comparative or better results. In this work, we present a methodology to create an ensemble of different artificial neural networks architectures, namely, MLP, RNN, LSTM, CNN and a hybrid CNN-LSTM, which aims to predict significant wave height on five different locations in the Brazilian coast. The networks are trained using NOAA's numerical reforecast data and target the residual between observational data and the numerical model output. A new strategy to create the training and target datasets is demonstrated. Results show that our framework is capable of producing high efficient forecast, with an average accuracy of 80%, and a increasingly reduction of computational cost.

Keywords Ocean Waves, Numerical Simulation, Neural Networks, Fluid Dynamics

1 Introduction

Numerical simulations of both weather and ocean parameters rely on the evolution of nonlinear dynamical systems that have a high sensitivity on initial conditions. Considering that errors in the observations and analysis are present, and therefore in the initial conditions, the concept of a unique deterministic solution of the governing equations becomes fragile [6, 11]. To circumvent this drawback, one can use an ensemble of simulations with different initial conditions, to represent the uncertainty of the data, and generates different solutions in which its average can provide a better understanding of the medium range behaviour of the system.

Albeit mathematical-physical models can be solved using traditional numerical solvers, the amount of available quality data prompt the use of machine learning algorithms as an low-cost alternative, achieving better performance in a computational time that is incredibly reduced. In this sense, artificial neural networks (ANNs) are one of the most promising tools for numerical simulations and act as an important alternative to problems with random patters such as those found in ocean modelling [10, 12].

Benefiting from both the advantages of using ensemble and artificial neural networks, we aim to provide in this work a new methodology to forecast significant wave height H_s on five different locations in Brazilian coast. We build five different architectures of artificial neural networks in

[1]felipe.minuzzi@ufsm.br
[2]farina@alum.mit.edu

which each predicts the residual between the observed value and H_s output from a numerical model. The prediction of a variable residual instead of its actual value has an advantage when using neural networks within ensemble predictions, as the final result will be added to the ensemble mean (EM), in the training process, the parameters of the activation function will only update the EM deviates from the target value [2, 3]. The final result is calculated by averaging H_s from the different neural network architectuPROGRAMAÇÃO GERALres, which is reconstructed by adding a forecast H_s with the neural network residual.

2 Ensemble methodology

Artificial neural networks (ANNs) are a kind of machine software that is designed to model the way in which the brain performs a particular task, and is able to learn and generalize huge sets of data [9]. From a mathematical standpoint, they can be considered as multiple nonlinear regression methods able to capture hidden complex nonlinear relationships between input and output variables [14].

In its simplest form, know as the perceptron, the structure of an ANN is based on a unit, or neuron (y_k), which receives a linear combination of weighted input and bias, i.e [9],

$$y_k = \phi\left(\sum_{j=1}^{m} \omega_{kj} x_j + b_k\right) \tag{1}$$

where $\omega_{kj}x_j$ for each j consists of the multiplication of the synaptic weight $\omega_{kj}x_j$ and the data x_j and b_k indicates the bias, which has the effect of increasing or lowering the net input of the activation function ϕ. As we aim to make our network accountable for non-linear dependencies, the activation functions need to be also non-linear, such as the log sigmoid or the hyperbolic tangent sigmoid functions. Nevertheless, this choice is user-defined and may depended on the application.

Several artificial neural network (ANN) architectures, based on layers of neurons, are possible. Different approaches to how information circulates throughout the network are also possible. In this work, we construct five different architectures of neural networks, namely, multilayer perceptron (MLP), recurrent neural network (RNN), long short-term memory (LSTM), convolutional neural network (CNN) and a hybrid CNN-LSTM, and average the results of each of these, to construct an ensemble of neural networks. We invite the reader to the references [7, 9] to a full description of the artificial neural networks used in this work.

As the use of neural networks to predict a residual has already been discussed and applied with satisfactory results [2, 3], we propose a different methodology to construct the datasets that will be used to train each of the neural networks mentioned in the previous section. The target, i.e., the variable that will be predicted is the residual of the significant wave height H_s, calculated as the difference between the real observed value and the forecast output of a numerical model. We consider the net residual, and not the absolute values, to account for negative values.

As the output of the neural networks consist of residuals, we reconstruct H_s by adding these residuals to a numerical prediction of H_s for the respective forecast horizon. We opted for this framework because allows us to generate an operational forecast that can be used on a daily basis. Afterwards, an average of the five H_s results is calculated which yields the final result of the algorithm's prediction.

3 Creating the training datasets

Figure 1 shows a schematic of the training methodology developed. In this framework, we build a feature dataset, in which each column have a numerical time series forecast for a specific

DOI: 10.5540/03.2025.011.01.0459

lead time. First column contains data from 3-hour lead forecasts of H_s, column two, 6-hour lead forecasts, and so on until the $n-$th column. Each row of this dataset represents a date and time, and its length define the size of the training phase. For the target dataset, each element (i, j) will be the residual between the numerical model in the features dataset at position (i, j) and the real measured data obtained from the buoy at that date and time. In this sense, the predictions of the neural network will be of one row and n columns, in the time position immediately following the last row of the target dataset. Therefore, the network will have predicted the residual, and considering each of the lead times of the columns, as it was constructed in the features dataset.

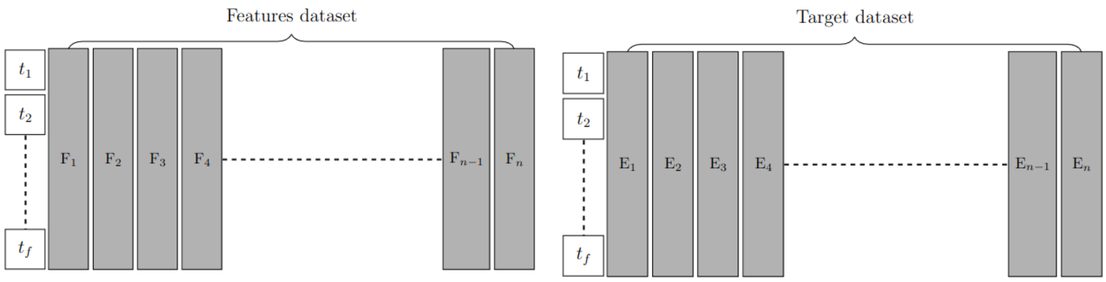

Figure 1: Schematic of the methodology for training applied in this work. Here, $t_1, \ldots, t_f$ represents the time series. In the target dataset, E_i represents the different between the numerical model forecasts from columns F_i and buoy data. The F_i columns are numerical model forecast on a specific lead time (column $i = 1$, lead time is 3hrs, $i = 2$, lead time is 6hrs, and so on until the $n-$th column). Source: from the authors.

In the training phase of every neural network, a cross-validation scheme was implemented, where 80% of the data is selected for the training and 20% for validation. This strategy is an excellent framework to avoid overfitting of a model, i.e., a model that yields a good accuracy to the validation set (seen data) and a bad result to unseen data. Since we are training with time series data, the order of events is important, which can be a problem when using cross-validation. To circumvent this issue, we perform a cross-validation on a rolling basis, where the training dataset is divided into smaller batches of data, and the cross-validation is applied to these batches. We train in a subset of data and then forecast the later data points of the batch to check accuracy. The same forecasted data points are then included as part of the next batch of training. This strategy also avoid excess in the memory usage of the training phase. To define the batch size, several simulations were performed, and a optimal value of twelve data points was obtained.

The Python library TensorFlow [1] and its Keras API [5] are used in this work to implement the neural networks. The model is compiled using the mean absolute error as loss function which is optimized by the Adam algorithm. The networks are build with six hidden layers (the hybrid CNN-LSTM has six hidden layers for each of the architectures) and the hyperbolic tangent is used as activation function, to account for negative values of the residual. A similar structure of simulation had already been used with satisfying results [12].

4 Data and area of study

The objective of our work is to forecast significant wave height H_s. The framework described in the previous section is used to predict the residue between numerical and observational H_s data, which later is reconstructed by adding these residuals to a numerical prediction of H_s. We use NOAA Wave Ensemble Reforecast data [4] as input, which is a 20-year global wave reforecast generated by the WAVEWATCH III model, forced by NOAA's Global Ensemble Forecast System (GEFSv12) [8]. The wave ensemble was run with one cycle per day, spatial resolution of $0.25^o \times 0.25^o$ and temporal resolution of three hours. The forecast range is sixteen days, which is also the same range in which we perform the forecast using the neural networks in this work.

Data from five buoys are also used for this study. All of them are located in the Brazilian coast, ranging from longitude 49° 86'W to 38° 25'W and latitudes 31° 33'S to 3° 12'S. These buoys belong to the National Program of Buoys (PNBOIA) of the Brazilian Navy, which aims to collect oceanographic and meteorological data of the Atlantic Ocean [13]. We interpolated data for missing points in these datasets. The training (and consequently the prediction) period is also determined to be the one with the least missing points. Table 1 presents the longitudes and latitudes of the five buoys. One limitation of our work can be inferred from Table 1, which shows the depth, in meters, of the buoys locations. Depending on the position, these can be considered coastal, which is not the goal of the NOAA Wave Ensemble, designed for deep waters.

The prediction period varies for each buoy, since some of the buoys used in this work are on maintenance and do not have real time data. We gathered this information for each buoy in Table 1. The training period is from 2013 until each buoy's prediction starting date. The results are shown for every three hours, the same temporal resolution of the reforecast data.

Table 1: Geo-spatial latitude and longitude location of the five buoys used in this work, period of prediction, water depth, WMO identification number and city of location in Brazil.

	Longitude	Latitude	Period of prediction	Depth (m)	WMO	City/State location
Buoy 1	49° 86' W	31° 33' S	20/02/2019 – 08/03/2019	200	31053	Rio Grande/RS
Buoy 2	47° 15' W	27° 24' S	30/10/2018 – 15/11/2018	200	31231	Itajaí/SC
Buoy 3	42° 44' W	25° 30' S	28/04/2018 – 14/05/2019	2164	31374	Santos/SP
Buoy 4	34° 33' W	8° 09' S	31/10/2015 – 16/11/2015	200	31229	Recife/PE
Buoy 5	38° 25' W	3° 12' S	08/04/2018 – 24/04/2018	200	31229	Fortaleza/RN

5 Results and discussion

In this section, we present the results of the prediction carried out with the ensemble of artificial neural networks that was described above (referred as NN ensemble in what follows). The residual that is the target of each simulation is added to a numerical forecast from NOAA Wave Ensemble Reforecast. To analyse the accuracy of our results, we evaluate the performance of our proposed model with three metrics: mean absolute percentage error (MAPE), mean absolute error (MAE) and the root mean squared error (RMSE). All the metrics are calculated against buoy data observations.

Figures 2 and 3 illustrate the comparison results between observed data, NOAA reforecast numerical model and this work ensemble of neural networks. As we can see, there is no quantitative improvement in the MAPE metric if we compare the numerical model and the neural networks ensemble. Buoys locations at Santos (Fig 2, middle) and Recife and Rio Grande (Fig. 3, upper and bottom) show the greatest discrepancy; the first and the second with a better accuracy for the NN ensemble while the third with a quantitative similar result to the numerical model..

Figure 2 highlights how the NN ensemble fails to predict H_s peaks in buoy location at Itajaí. The reason might be because we are training the models and reconstructing H_s with data from NOAA numerical simulation and since global wave models are known to not represent extreme events very well, the pattern is also learned by the neural networks. The poor representation of peaks is also seen in other buoy locations, and this could be addressed as one of the drawbacks of the proposed methodology. It is well known that, in the numerical forecast of ocean waves, peaks, storms and extremes events are difficult to predict. The training of the neural networks are based on the NOAA numerical results, which can explain the limitation. Also, data imbalance,

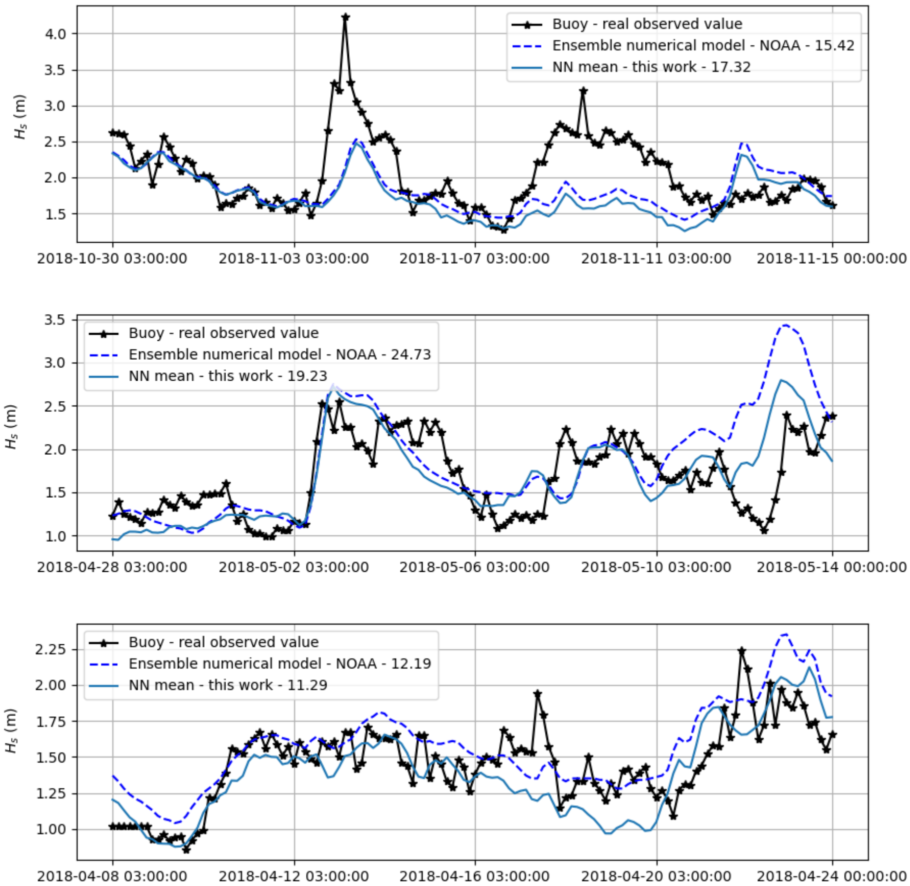

Figure 2: Comparison between real observed data, NOAA reforecast numerical simulation and this work ensemble methodology. Numbers in the legend refers to MAPE metric. Upper: Itajaí, Middle: Santos, Bottom: Fortaleza. Source: from the authors. Source: from the authors.

since peaks and storms represent a small portion of the training dataset, as well as the use of a numerical global model, are others problems that prevents from getting better results.

One can also see that the ensemble in fact learn and predict a residual that is variable according with the initial error, as the graphs show, and although in the beginning of the prediction both numerical and NN have the same behaviour, later on the prediction period the lines get apart from each other, specially where the numerical model is known to lose accuracy, the ensemble of neural networks maintains it. The results show also the same pattern of balance in the error if one looks at the metrics MAE and RMSE, as can be seen in Tab. 2. However, the cost of simulation for the ensemble is vastly reduced compared to the numerical simulation, which can be seen as an advantage. For each neural network architecture, our algorithm took approximately 32 minutes for training and the prediction of a single time value took 2.62×10^{-6} seconds. Thus, the sixteen days predictions period (128 steps) took 3.35×10^{-4} seconds. The simulations were performed in a machine with Intel Xeon processor with 20 cores, 128 Gb of RAM memory, with a GeForce RTX 2080 Ti graphics card. We parallelize all the training and prediction step, so the results for each architecture are given in the same time.

DOI: 10.5540/03.2025.011.01.0459

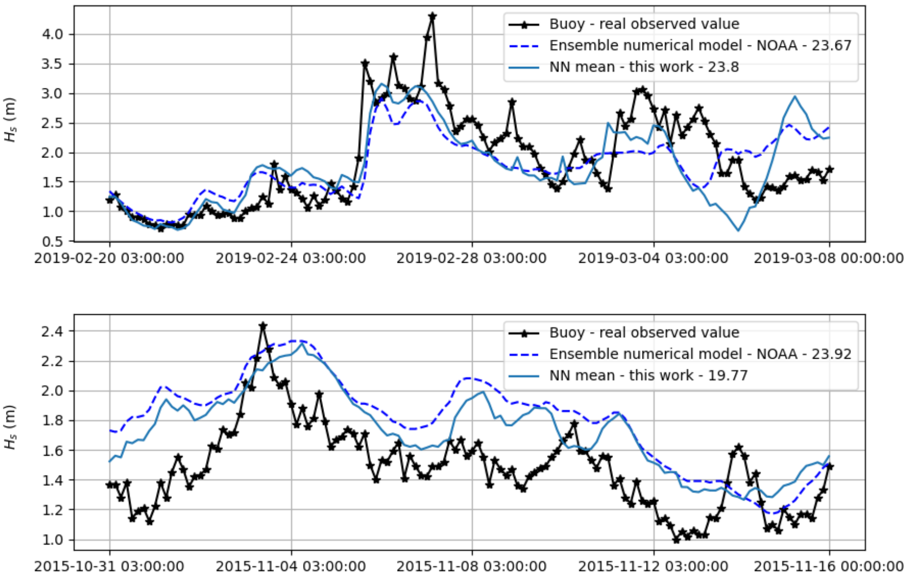

Figure 3: Comparison between real observed data, NOAA numerical simulation and this work ensemble methodology. Numbers in the legend refers to MAPE metric. Upper: Rio Grande, Bottom: Recife. Source: from the authors.

Table 2: Error metrics for each buoy locations. Comparison between ensemble neural network and NOAA numerical model against real observational data.

Error metrics	Itajaí		Santos		Fortaleza		Rio Grande		Recife	
	NN	NOAA	NN	NOAA	NN	NOAA	NN	NOAA	NN	NOAA
MAPE (%)	17.32	15.42	19.23	24.73	11.29	12.19	23.8	23.67	19.77	23.92
MAE (m)	0.40	0.40	0.30	0.39	0.17	0.17	0.44	0.44	0.28	0.34
RMSE (m)	0.56	0.56	0.38	0.55	0.21	0.22	0.58	0.57	0.31	0.38

6 Conclusions

We propose in this work a surrogate methodology to traditional numerical models creating an ensemble of different architectures of artificial neural networks. The results shows that our framework have a good accuracy with metrics that are comparable or, for some cases, superior than the NOAA numerical model. Also, the neural networks ensemble does not reproduce the behaviour of loosing accuracy as the lead time forecast increase, a well known drawback of numerical models. Comparing our result with the historical error of NOAA numerical data for each lead time, we also see an improvement in the performance. The difference in the results between each of the neural network architecture also shows that the strategy of using an ensemble was appropriate. Another major contribution of the present work is that it is the first one to use NOAA Wave Ensemble reforecast data, a large dataset that carries real information on the decay of skill as a function of the forecast lead time, which allows a better discussion about prediction.

Although our model gives highly accurate predictions, there are some limitations in the results, such as the forecast of peaks. From the neural networks perspective, the architectures that behaved poorly in the simulations should be removed from the set to improve the ensemble results. Since there is not a pattern on which architecture is worst for each location, we want to show in this work that the ensemble methodology can improve, if the right networks are chosen. Besides, as mentioned in the text, we considered numerical simulations from a global wave model, in coastal locations that are not suitable for these. All these issues will be addressed in future works.

Acknowledgements

F. C. Minuzzi thanks the financial support from FAPERGS, under the term No. 23/2551-0000796-7. We also acknowledge NOAA and PNBOIA for providing the necessary data.

References

[1] Martín A. et al. **TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems**. Software available from www.tensorflow.org. 2015. URL: `https://www.tensorflow.org/`.

[2] R. M. Campos et al. "Improving NCEP's global-scale wave ensemble averages using neural networks". In: **Ocean Modelling** (2020), p. 101617.

[3] R. M. Campos et al. "Nonlinear wave ensemble averaging in the Gulf of Mexico using neural networks". In: **Journal of Atmospheric and Oceanic Technology** 36.1 (2019), pp. 113–127.

[4] R.M. Campos et al. "Description and Validation of a new 20-year global wave ensemble reforecast data." In: **3rd International Workshop on Waves, Storm Surges, and Coastal Hazards. October 1-6, 2023. University of Notre Dame, Indiana, USA.** 2023. URL: `https://github.com/NOAA-EMC/gefswaves%5C_reforecast`.

[5] F. Chollet et al. **Keras**. 2015. URL: `https://github.com/fchollet/keras`.

[6] L. Farina. "On ensemble prediction of ocean waves". In: **Tellus A: Dynamic Meteorology and Oceanography** 54.2 (2002), pp. 148–158.

[7] I. Goodfellow, Y. Bengio, and A. Courville. **Deep learning**. Vol. 1. MIT press Cambridge, 2016.

[8] T.M. Hamill et al. "The Reanalysis for the Global Ensemble Forecast System, Version 12." In: **Monthly Weather Review.** 150-1 (2022), pp. 59–79. URL: `https://registry.opendata.aws/noaa-gefs.`.

[9] S. S. Haykin. **Neural networks and learning machines**. New York: Prentice Hall, 2009.

[10] N. P. Juan and V. N. Valdecantos. "Review of the application of Artificial Neural Networks in ocean engineering". In: **Ocean Engineering** 259 (2022), p. 111947.

[11] G. J. Komen et al. **Dynamics and modelling of ocean waves**. Cambridge University Press, 1996.

[12] F. C Minuzzi and L. Farina. "A deep learning approach to predict significant wave height using long short-term memory". In: **Ocean Modelling** 181 (2023), p. 102151.

[13] Brazilian Navy. **PNBOIA - National Program of Buoys**. Accessed on 12-04-2021. 2020. URL: `https://www.marinha.mil.br/chm/dados-do-goos-brasil/pnboia`.

[14] D. J. Peres et al. "Significant wave height record extension by neural networks and reanalysis wind data". In: **Ocean Modelling** 94 (2015), pp. 128–140.

