Nonlinear Wave Ensemble Averaging using Neural Networks

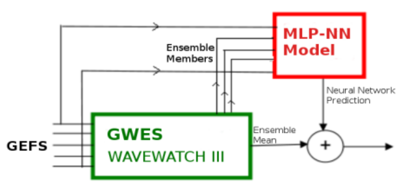
Vladimir Krasnopolsky[1], Ricardo Martins Campos[2], Jose-Henrique Alves[3], Stephen Penny[2]

[1]EMC/NCEP / NOAA Center for Weather and Climate Prediction
[2]Dept.of Atmospheric & Oceanic Science / University of Maryland
[3]SRG/EMC/NCEP / NOAA Center for Weather and Climate Prediction

e-mail: **vladimir.krasnopolsky@noaa.gov**

1. Introduction

An experiment applying neural networks (NNs) to nonlinear wave ensemble averaging is conducted in the Gulf of Mexico (GOM). It is an approach that expands the traditional arithmetic ensemble mean (EM) to a nonlinear mapping that better captures the differences among the ensemble members. The NNs have the members of the Global Wave Ensemble System (GWES) as input, and NN outputs are trained using six NDBC buoys (42001, 42002, 42003, 42039, 42055, 42360). The variables selected for the study are 10-m wind speed (U10), significant wave height (Hs), and peak period (Tp) for the year of 2016. The first experiments testing NN architectures for GWES nonlinear wave ensemble averaging at single locations were reported by Campos et al. (2017). It was found that the best NN model was composed of two layers with 11 neurons at the intermediate layer, using a basis function of the hyperbolic tangent, sequential training, and the log function applied to time series of significant wave height. Equation 1 shows the final strategy of the NN simulation, where the simple arithmetic EM is first calculated and then the NN model is applied to model the residue (difference between the target value and EM). This method can focus the NN simulation on the nonlinear part, instead of simulating the whole signal with linear and nonlinear components (Krasnopolsky, 2013). It builds a more robust model that provides reliable ensemble averages at different metocean conditions and sea severities.



$$NEM = EM + NN_r(p_1, p_2, \cdots, p_n) \tag{1}$$

2. Neural Network training and sensitivity tests

At each particular grid point, the inputs of the NN consisted of 63 ensemble members (20 ensembles plus one control member, for U10, Hs, and Tp) plus the sine and cosine of time (to capture seasonality effects). To introduce space in the NN model, latitude and longitude were included for a total of 67 inputs. In order to determine the complexity of the NN model required to obtain the optimal training, architectures with 12 different numbers of neurons, 8 different filtering windows (time-domain), and 100 seeds for the random initialization were studied. We constructed different NNs for specific forecast days, from Day 0 to Day 10. The number of neurons and filtering windows (hours) using the moving average method are, respectively, $N\ [\,2, 5, 10, 15, 20, 25, 30, 35, 40, 50, 80, 200]$ and $FiltW\ [\,0, 24, 48, 96, 144, 192, 288, 480]$ hours. Two thirds of the available data was selected for training and 1/3 for the test set, using a cross-validation scheme with 3 cycles.

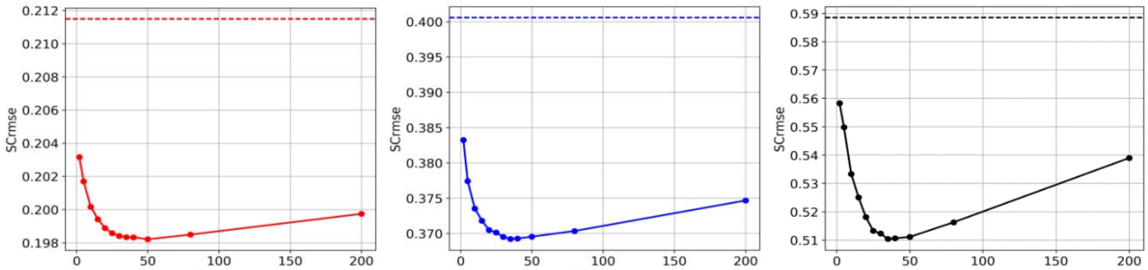


Figure 1 – Scatter component of the RMSE ($SCrmse$) obtained for the NN training tests on forecast Day 0 (red), Day 5 (blue), and Day 10 (black) for significant wave height Hs. Results involving different initializations and filtering windows were averaged to analyze the sensitivity to the number of neurons only. The solid line is the NN model results while the dashed line is the result for the ensemble mean (EM); shown to compare their performances. Points at the plots represent the number of neurons equal to 2, 5, 10, 15, 20, 25, 30, 35, 40, 50, 80, and 200.

Results obtained show that the bias is not very sensitive to the number of neurons so a few neurons are sufficient to improve the bias; however, the scatter error is highly sensitive to the number of neurons (Figure 1). The scatter component of the RMSE ($SCrmse$, Mentaschi et al., 2013) and the correlation coefficient (CC) are continuously improved by higher number of neurons; however, when the number of neurons approaches 40 to 50 neurons the results start deteriorating. Regarding the filtering window, optimum results were found between 48 to 192 hours; which is mainly because the moving average filtering removes high frequency noise from the signal which helps the NN to minimize the scatter error. The best NN configuration for each forecast day was independently selected for the final simulations and assessments. The first row of Figure 2 shows two significant improvements. The first is the benefit of the NCEP ensemble approach of GWES, indicating that the arithmetic ensemble mean has better skill than the control run. The second is the further improvement provided by the nonlinear ensemble averaging using NNs. Both are more evident at longer forecast times.

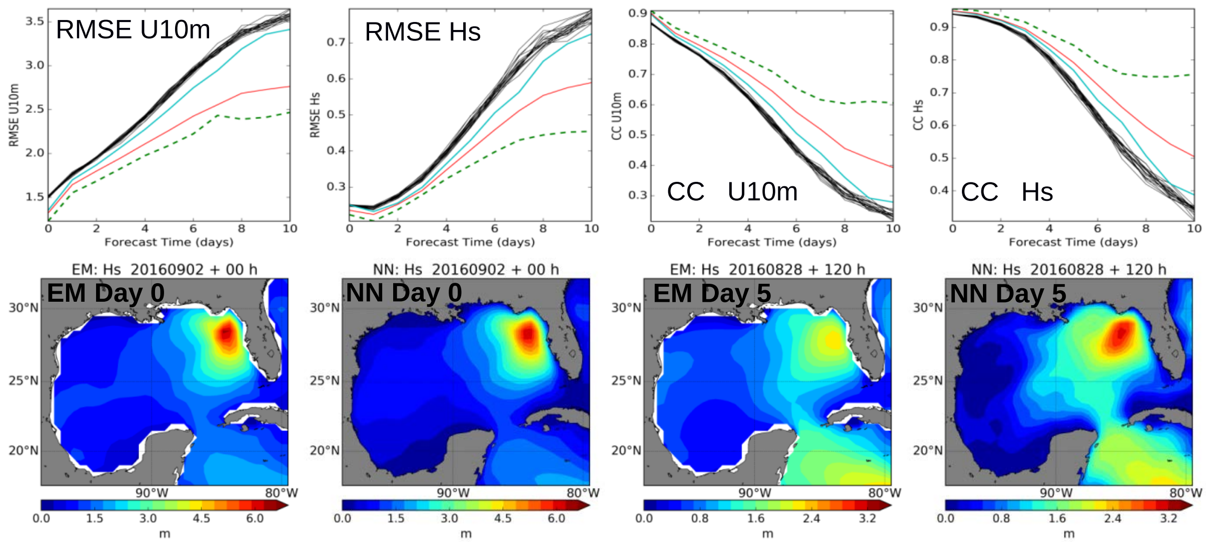


Figure 2 – RMSE and CC as a function of forecast time for U10 and Hs, at the top row, where the black curves are the GWES members, cyan is the deterministic (control) run, red the arithmetic ensemble mean (EM), and green the nonlinear ensemble averaging using NNs. The bottom row shows the maps of Hs related to Sep/02/2016 00Z (Hurricane Hermine), with different forecast times.

3. Conclusions

Our results show that using NN models demonstrates their main advantages at longer forecasts times. Although the NNs do not deal with physical aspects, the improvement of the NN experiment is not restricted to one variable, but the U10, Hs, and Tp have all benefited. It was verified that the conservative ensemble approach (EM) is excellent in reducing the scatter component of the errors but it does not improve the systematic bias, as was expected. The NN experiment described was able to reduce the systematic errors as well as the scatter error, proving to be a useful tool not only for bias correction but also to significantly improve the whole forecast. Error metrics applied to all variables and forecast ranges indicated that, apart from the bias of U10 in the Day-10 simulations, all the results using NNs proved to be better than the arithmetic mean (EM). The GOM maps of Figure 2 present an example for Hurricane Hermine, when the NNs better captured the extreme event than EM, especially at Day 5.

References

Campos, R.M., Krasnopolsky, V., Alves, J.-H., Penny, S., 2017, Improving NCEP's Probabilistic Wave Height Forecasts Using Neural Networks: A Pilot Study Using Buoy Data. National Oceanic and Atmospheric Administration - Office Note 490, http://doi.org/10.7289/V5/ON-NCEP-490

Krasnopolsky, V., 2013. "The Application of Neural Networks in the Earth System Sciences. Neural Network Emulations for Complex Multidimensional Mappings", Springer, 200 pp.

Mentaschi, L., Besio, G., Cassola, F., Mazzino, A., 2013. Problems in RMSE-based wave model validations. Ocean Modelling, 72, 53–58.