Using Neural Networks for Nonlinear Averaging NCEP Wave Model Ensemble

V. Krasnopolsky[1], R. Campos[2], J.H. Alves[3], S. Penny[2]
[1]NWS/NCEP/EMC, [2]AOSC/UMD, [3]NCEP/EMC/SRG
Email: vladimir.krasnopolsky@noaa.gov; Phone: 301-683-3742

The NCEP Global Wave Ensemble System (GWES) was implemented in 2005 [1] and initially validated by [2]. After upgrades reported in [3], it is now run with four cycles per day, using a spatial grid with 0.5^{o} resolution, with forecast range to 10 days. A total of 20 perturbed members plus a control member compose the GWES, which consists of an implementation of the WAVEWATCH III model [4], forced by winds from NCEP's Global Ensemble Forecast System (GEFS) [5]. A recent assessment and comparison of deterministic and ensemble products using altimeter data is provided in [3]. Their results show that although the general bias of the ensemble system does not show significant improvement over the deterministic global wave, after the fifth forecast day, root mean square errors from the GWES become smaller than the deterministic run. Furthermore, the GWES continuous ranked probability scores (CRPS) systematically outperforms the corresponding deterministic model's mean absolute error (MAE) in all forecast times.

In the current study, we propose an improvement of the quality of output products from the GWES using neural networks (NN), which are initially used to compute nonlinear averages. Currently a conservative ensemble approach is used to calculate the ensemble mean (EM) in the GWES. The ensemble mean for variable *p* is calculated as,

$$EM = \frac{1}{n}\sum_{i=1}^{n} p_i \tag{1}$$

here *n* is the number of ensemble members and p_i is the *i*-th ensemble member.

An improvement upon (1) can be achieved using weighted EM (WEM),

$$WEM = \frac{\sum_{i=1}^{n} W_i \cdot p_i}{\sum_{i=1}^{n} W_i} \tag{2}$$

where W_i are weights subscribed to ensemble members. A priori information can be used to select the weights W_i. In addition, if observational data are available for the variable *p*, eq. (2) can be considered as a linear regression. Solving the linear regression equations (2), W_i and the linear regression EM (LREM) can be found. Eq. (2) assumes a linear relationship between EM and the ensemble members; however, in reality, this relationship may be significantly nonlinear, and we can use a nonlinear statistical tool like NN to derive a relationship between ensemble members and nonlinear EM (NEM),

$$NEM = NN(p_1, p_2, \cdots, p_n) \tag{3}$$

In a previous work [6], we demonstrated that a NN technique can be successfully used for averaging multi-model ensemble for precipitation over the Continental US. We showed that NN provides significantly better results than conservative ensemble or LREM. In fact, NN results are comparable with those obtained by a human meteorologist-analyst. In this pilot study, we apply NN to calculate NEM in GWES.

To start with, we selected a "one buoy location" setup. We used 21 GWES ensemble members from a single grid point near the buoy location at 32.501N and 79.099W (buoy #41004 in the North Atlantic Ocean, water depth of 37 meters); distant only 9.7 km to the nearest model grid point selected. As inputs to our NN we use three model variables related to 5-day forecasts for: significant wave height, H_s, peak period T_p, and wind speed at 10 m height, U_{10}, a total of 63 (3 x 21) inputs. Also, two metavariables: sin and cos of the day of the year were used. Thus, our NN has 65 inputs in total. The NN has three outputs: H_s, T_p, and U_{10}. One year of data for buoy #41004 was used for training the NN outputs. For NN validation we used one year of data collected at buoy #41013, located at a distance of about 100 mi from buoy #41004.

Table 1. Performance of three ensembles (1) to (3) for H_s on independent validation set (buoy #41013). MAE is mean absolute error, SI – scatter index, CC – correlation coefficient, and Max the largest value of H_s in m.

	Bias	**RMSE**	**MAE**	**SI**	**CC**	**Max**
Conservative EM	-0.03	0.445	0.301	0.334	0.759	6.28
LREM	0.28	0.463	0.313	0.348	0.754	4.46
NEM	0.12	0.424	0.29	0.328	0.782	4.3

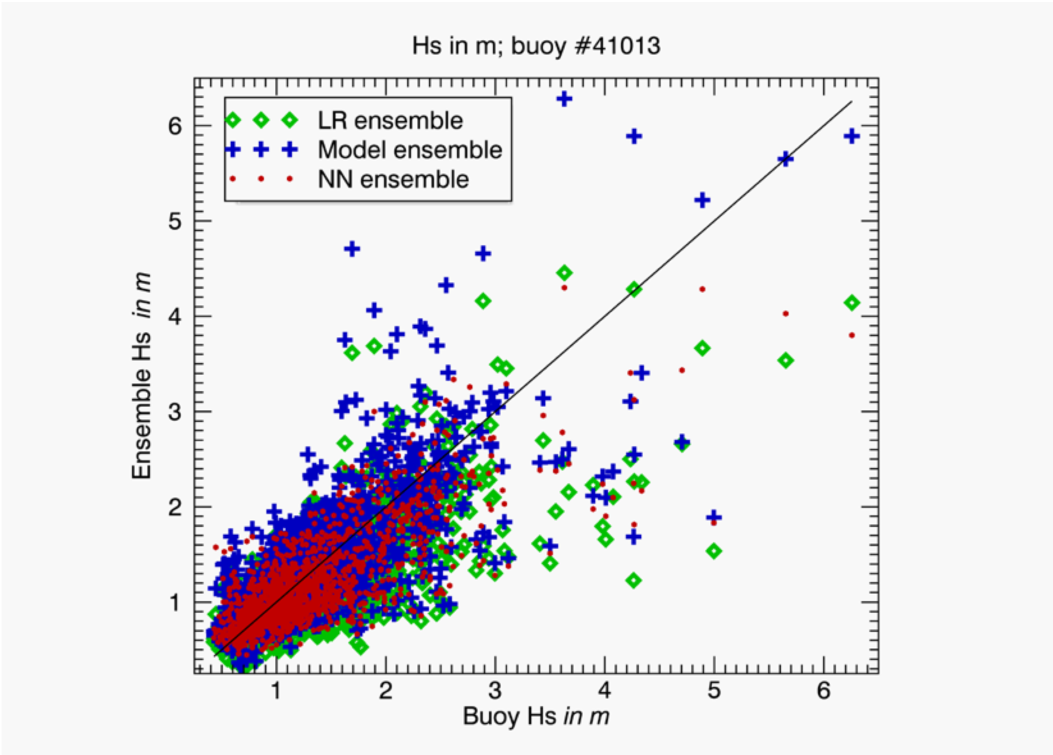

Figure 1 - Scatter plot shows three overlaid scatterplots: blue crosses show wave model ensemble, green diamonds – LREM, and red dots – NEM.

Table 1 shows comparison of three aforementioned ensembles relative to an independent validation set (buoy #41013). NEM outperforms the model ensemble and LR ensemble for all statistics except the max value. Fig. 1 illustrates the reason: there are very few data points with $H_s > 3$ m. These data are not sufficient for NN (and LR) training in the area of high H_s.

As the next step, our investigation will move from one buoy configuration to two regional (Atlantic and Pacific) configurations and, eventually, to a global configuration when one or several NNs will provide NEM over the entire global ocean. In addition, we are going to include the altimeter data in the training process.

References

[1] Chen, H. S., 2006: Ensemble prediction of ocean waves at NCEP. Proc. 28th Ocean Engineering Conf., Taipei, Taiwan, NSYSU, 25–37.

[2] Cao, D., H. S. Chen, and H. L. Tolman, 2007: Verification of ocean wave ensemble forecasts at NCEP. Proc. 10th Int. Workshop on Wave Hindcasting and Forecasting and First Coastal Hazards Symp., Oahu, Hawaii, G1.

[3] Alves, J.H.G.M., Wittman, P., Sestak, M., Schauer,J., Stripling, S., Bernier, N.B., McLean, J., Chao, Y., Chawla, A., Tolman, H., Nelson, G., Klots, S., 2013. The NCEP–FNMOC combined wave ensemble product. Expanding Benefits of Interagency Probabilistic Forecasts to the Oceanic Environment. Bull. American Met. Society, BAMS, December 2013.

[4] WAVEWATCH III® Development Group, 2016: User manual and system documentation of WAVEWATCH III® version 5.16. Tech. Note 329, NOAA/NWS/NCEP/MMAB, College Park, MD, USA, 326 pp. + Appendices.

[5] Hou, D., Z. Toth, Y. Zhu, and Y. Yang, 2008: Impact of a Stochastic Perturbation Scheme on NCEP Global Ensemble Forecast System. Proceedings, 19th AMS conference on Probability and Statistics. New Orleans, LA, 20-24 Jan. 2008.

[6] Krasnopolsky V., and Y. Lin, 2012: "A Neural Network Nonlinear Multimodel Ensemble to Improve Precipitation Forecasts over Continental US", Advances in Meteorology, Volume 2012, Article ID 649450, 11 pages, doi:10.1155/2012/649450