Ensemble of Neural Network Emulations for Model Physics: The Impact on Climate Simulation

Vladimir M. Krasnopolsky[1,2], Phone: 301-763-8000 x 7262, Fax: 301-763-8545,
Email: Vladimir.Krasnopolsky@noaa.gov
Michael S. Fox-Rabinovitz[1], and A. Belochitski[1]
[1]Earth System Science Interdisciplinary Center, University of Maryland, USA
[2]SAIC at NCEP/NOAA, USA

1. Introduction

In our previous studies, we developed NN emulations of climate model physics for the NCAR CAM-2 [1, 2]. Specifically, a number of the NN emulations of the original long wave radiation (LWR) have been individually trained, with slightly different approximation and interpolation accuracies. In this study, we investigate the ability of NN ensembles, created from these NN emulations, to provide a better approximation and interpolation than their individual ensemble members, and to assess the NN ensemble impact on climate simulation.

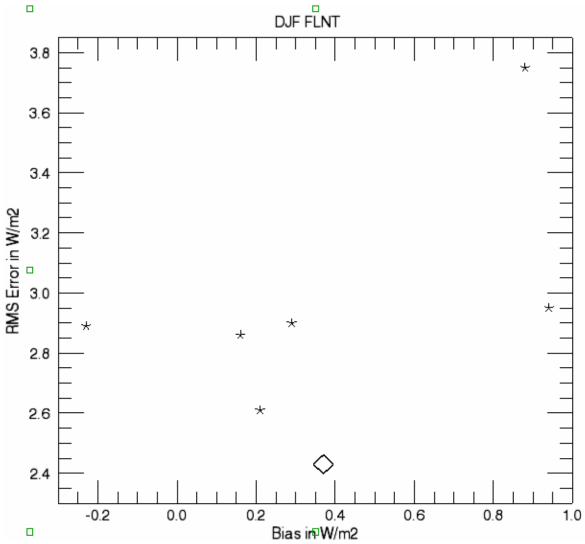
As a nonlinear model or nonlinear approximation, the NN approximation problem allows for multiple solutions or for multiple NN emulations of the same LWR. For example, the original LWR, used in NCAR CAM, can be approximated with NNs with different numbers of hidden neurons, with different weights (resulting from the NN training with different initializations), different partitions of the training set, etc. The availability of multiple NN emulations, providing complimentary information about the original parameterization, opens an attractive opportunity of introducing a NN ensemble approach. It allows for integrating the complimentary information, contained in the individual ensemble members, into an ensemble that “knows” more about or represents the original LWR better than each of the individual ensemble members (a particular NN emulation). Moreover, the NN ensemble, when it is used in a climate model instead of a single NN emulation of the original LWR, may provide a better accuracy of the climate simulation.

An ensemble of NNs consists of a set of members, which are individually trained NNs. Their outputs are combined when applied to a new input data to improve the interpolation ability. We used a conservative ensemble [3] with a simple (with equal weights for all members) averaging of the member outputs providing the ensemble statistics.

2. NN ensemble for NCAR CAM LWR

The NN ensemble approach is an extension of our development of NN emulations we worked on for the last several years. Our approach to developing NN emulations for NCAR CAM LWR is described in detail in [1]. NCAR CAM-2 is integrated for two years to generate representative data sets for training, validation and test of NN emulations. All NN emulations and ensembles of the NN emulations are tested against the control which within our framework is obviously the original NCAR CAM LWR. Mean difference *B* (bias, or a systematic error) and the standard deviation (SD) of the difference between the original LWR and its NN emulation (random error), maximum and minimum errors, are calculated.

The final and the most important test is performed by estimating the accuracy of decadal climate simulation runs with single NN ensemble members and with the NN ensemble vs. the control climate run with the original NCAR CAM LWR. We have selected a sufficiently diverse group of six NN emulations from the NN emulations we have already trained. Four of six members have the same architecture and were trained with perturbed initial conditions for the NN weights. Two others have different number of hidden neurons. These six NN emulations constitute the NN ensemble. Climate simulations have been run with NCAR CAM for 25 years with each of these six ensemble members (each of the six LWR NN emulations). The results



(climate fields and diagnostics) of each simulation are compared with the control climate run of NCAR CAM performed with the original LWR. The climate simulation errors (systematic, random, maximum, and minimum) have been calculated for each ensemble member. For the net LWR flux at the top of the model atmosphere (FLNT or OLR in W/m^2, for DJF (Dec-Jan-Feb)), bias and RMS errors for each ensemble member are shown as stars in the figure on the left. Then the NN ensemble climate run has been performed. For this run, six NN emulations are applied and the LWR outputs are calculated as the mean of these six NN emulation outputs, at each time step and at each grid point throughout the model integration. The use of the NN ensemble in climate simulation significantly reduces the systematic error (bias); it also reduces the RMS error to the value smaller than that of the best individual ensemble member. Bias and RMS error for the ensemble are shown as diamond in the figure. A significant reduction is obtained for the extreme errors as well.

3. Conclusions

A new NN ensemble approach is presented. It is applied to NN emulations of the LWR parameterization in NCAR CAM-2. It is shown that practically all individual NN emulations of LWR that we have trained in the process of developing an optimal NN LWR emulation, can be used within the NN ensemble approach for climate simulation. Using the NN ensemble results in a significant reduction of climate simulation errors, namely the systematic and random errors, the magnitudes of the extreme errors or outliers and, in general, the number of large errors.

References

[1] V.M. Krasnopolsky, M.S. Fox-Rabinovitz, and D.V. Chalikov, "New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Long Wave Radiation in a Climate Model", *Monthly Weather Review*, v. 133, No. 5, pp. 1370-1383, 2005.

[2] V.M. Krasnopolsky and M.S. Fox-Rabinovitz , "A New Synergetic Paradigm in Environmental Numerical Modeling: Hybrid Models Combining Deterministic and Machine Learning Components", *Ecological Modelling*, v. 191, pp. 5-18, 2006.

[3] S.V. Barai and Y. Reich, “Ensemble modeling or selecting the best model: Many could be better than one”, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing,* 13, pp. 377-386, 1999