Application of Neural Networks to Speed Up Calculations of Radiation Parameterizations in the NOAA Global Ensemble Forecast System and Global Data Assimilation System ensembles

Alexei A. Belochitski[1] and Vladimir Krasnopolsky[2]

[1]IMSG@NOAA/NCEP/EMC, [2]NOAA/NCEP/EMC
Email: Vladimir.Krasnopolsky@noaa.gov

The NOAA Global Ensemble Forecast System (GEFS) and Global Data Assimilation System (GDAS) ensembles have been continually growing in complexity in order to provide forecasters with more accurate probabilistic forecast guidance and improve data assimilation to produce accurate initial conditions. Its demands for HPC resources perpetually outrun the increase in available computational power. For example, the operational GEFS v12 at its peak uses nearly the entirety of the NOAA WCOSS Dell Phase 3.5 supercomputer. Flexible and powerful numerical techniques are required to alleviate growing demands for HPC resources. The Neural network (NN) technique is increasingly being applied in NWP models to achieve this goal. NNs are used to accelerate calculations in NWP models and to develop faster and better constituents (parameterizations) of model physics at NOAA, ECMWF, UKMO, and other weather forecast centers.

The Environmental Modeling Center (EMC) of NCEP has extensive pioneering experience with the successful development of AI/ML applications for global atmospheric modeling, specifically for earlier versions of NCEP GFS and CFS (Krasnopolsky et al. 2010), as well as NCAR CAM. Several of the most time-consuming parts of model physics, including short wave radiation (SWR) and long wave radiation (LWR) parameterizations, have been emulated with high accuracy using an advanced and universal ML NN technique. Table 1 illustrates the performance gains that have been achieved. The speed up does not significantly decrease with increasing vertical resolution of the model.

	NCAR CAM (L=26)		NCEP CFS (L=64)	
	LWR	SWR	RRTMG LWR	RRTMG SWR
Speed up, *times*	150	20	16 (20)	60 (88)

Table 1 Speed up of LWR and SWR parameterizations achieved by using NN emulations in the NCAR CAM and NCEP CFS with different vertical resolutions (L). The presented numbers are for the 2010 versions of the models. Numbers in parentheses are speed-up factors for cloudy-sky conditions.

We showed that a comparable improvement in performance can be achieved in the latest version of GFS which constitutes the core of the GEFS (Belochitski and Krasnopolsky 2021). To assess the potential speed up, in the pilot study we configured a modern version of the GFS to run at the C384L64 (~25 km) configuration, the operational resolution of GEFS v12, and replaced the modern radiative transfer parameterizations (RRTMG) with their shallow NN-based emulators developed for this GFS. Table 2 shows an overall 23% speed up of the model, a result that is to a substantial degree due to an almost an order of magnitude acceleration of radiative transfer calculations.

	Speed up due to using NN emulator of radiation, *times*
Total	1.23
Radiation	9
Dynamical core	1.10

Table 2 Speed up of GFS model components due to the NN emulator of radiative transfer parameterizations in GFS pre-v16 at the GEFS operational resolution of C384L64.

An additional speed up stems from the fact that NN emulations provide a uniform computational efficiency across different atmospheric conditions, while the performance of the original physical parameterizations usually depends significantly on the local atmospheric state. For example, radiative transfer parameterizations take longer to complete their calculations under cloudy conditions, while performance of the NN emulation is invariant with respect to its inputs; therefore, NN emulations provide improved balancing and reduced idle time in an HPC parallel environment, as illustrated by the speed up of the dynamical core in Table 2. Such a speed up will allow us to increase the number of GEFS ensemble members from 30 to 38 and in GDAS from 80 to 98 without an increase in the amount of required computational resources.

Figure 1 illustrates accuracy of the NN emulators over a 10-day forecast. The development methodology for all stages of an NN-based emulation of the model physics components has been worked out and successfully tested. This methodology will be used in the current work with GEFS and GDAS. In a recent pilot study, we demonstrated that a single NN can accurately emulate both the LW and SW radiation parameterizations (RRTMG) in the new FV3 GFS. This development will provide an additional speed up in the calculation of atmospheric radiation.

NN emulators developed using the training software developed at NCEP are FORTRAN modules that are specifically designed for and are completely compatible with the NCEP's developmental and operational HPC environments.

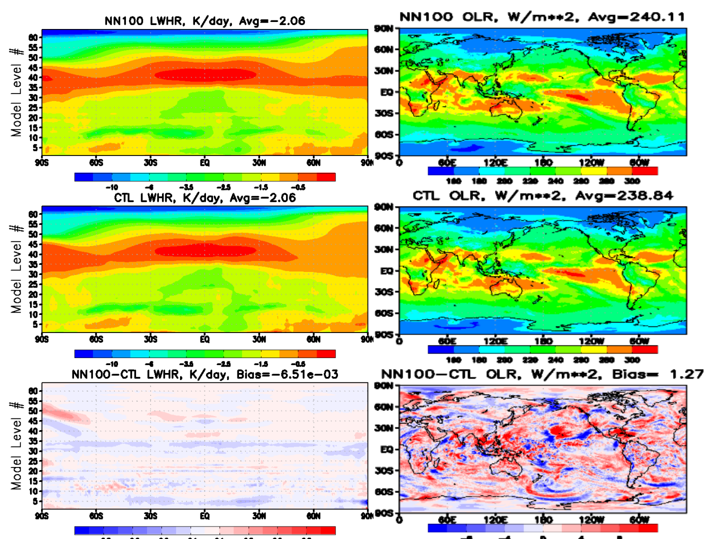

Figure 1 Results of parallel runs of GFS C384L64: time average over 10-day forecast. LWR heating rates (left column), and outgoing long wave radiation flux (right column). Upper row – results produced by NGFS (GFS with NN emulations of radiation parameterizations), medium – by GFS control (CTL) run, and the lower row the difference (NGFS – GFS). Vertical coordinate shows model level number.

References

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Hou, Y. T., Lord, S. J., & Belochitski, A. A. (2010). Accurate and Fast Neural Network Emulations of Model Radiation for the NCEP Coupled Climate Forecast System: Climate Simulations and Seasonal Predictions, *Monthly Weather Review*, 138, 1822-1842. doi: 10.1175/2009MWR3149.1

Belochitski, A. and V. Krasnopolsky (2021). Robustness of neural network emulations of radiative transfer parameterizations in a state-of-the-art general circulation model. *Geosci. Model Dev.*, 14, 7425–7437, 2021. https://doi.org/10.5194/gmd-14-7425-2021