P1.20 USING NEURAL NETWORKS FOR FAST AND ACCURATE APPROXIMATION OF THE LONG WAVE RADIATION PARAMETERIZATION IN THE NCAR COMMUNITY ATMOSPHERIC MODEL: EVALUATION OF COMPUTATIONAL PERFORMANCE AND ACCURACY OF APPROXIMATION

Vladimir M. Krasnopolsky*,
Earth System Science Interdisciplinary Center, University of Maryland and SAIC at NCEP/NOAA
Michael S. Fox-Rabinovitz,
Earth System Science Interdisciplinary Center, University of Maryland, and
NASA/Goddard Space Flight Center, Global Modeling and Assimilation Office,
and Dmitry V. Chalikov
Earth System Science Interdisciplinary Center, University of Maryland

1. APPROACH

Tremendous developments in numerical modeling and in computing capabilities during the last decades have contributed dramatically to scientific and practical significance of interdisciplinary climate, climate change and weather prediction numerical modeling. One of the main problems of development and implementation of these high-quality high-resolution atmospheric and oceanic models is the complexity of physical, chemical and biological processes involved. Parameterizations of model physics, adjusted to model resolution and computer resources, are approximate schemes based on simplified 1-D first principles equations and empirical data. The parameterizations are so time consuming, even for most powerful modern supercomputers, that they have to be calculated less frequently than model dynamics (based on solving 3-D geophysical fluid dynamics equations). This negatively affects the accuracy of the model physics calculations and the temporal consistency and may lead to a significant reduction of the accuracy of climate simulations and weather predictions. For example, calculation of a model physics package in a typical moderate (a few degrees) resolution GCM (General Circulation Model), such as the NCAR (National Center for Atmospheric Research) CAM-2 (Community Atmospheric Model) takes approximately 80% of the total model computations. Higher and variable model resolutions [*e.g. Fox-Rabinovitz et al. 2002*] and more frequent model physics calculations, desirable for temporal consistency with model dynamics, would increase the percentage to more than 90%.

Such a situation is an important motivation to look for alternative, faster ways of calculating model physics and chemistry. One of the alternative approaches is based on the idea of using fast and accurate statistical techniques for approximating atmospheric physics and chemistry parameterizations. For example, a traditional statistical technique based on representation of input/output relationship as an expansion of hierarchical correlated functions [*Rabitz et al., 1998; Rabitz and Alis, 1999*], has been successfully used in some atmospheric chemistry applications (see [*Schoendorf et al., 2003*] and references there).

* *Corresponding author address*: Vladimir M. Krasnopolsky, SAIC at NCEP/NOAA, MD 20746, Email: Vladimir.Krasnopolsky@noaa.gov

During the last decade a new emerging approach based on neural network (NN) approximations has found applications in a large variety of in different fields; specifically, for fast and accurate approximation of model physics processes [*Krasnopolsky and Chevallier 2001,2003*] and for satellite retrieval procedures [*Krasnopolsky, 1997; Krasnopolsky and H. Schiller, 2003*]. Recently, the NN approach has been used to develop a fast (8 times faster than the original parameterization) and accurate long-wave (LW) radiation parameterization for the ECMWF model [*Chevallier et al. 1998, 2000*]. The NN approach has been also used for approximations of model physics in ocean numerical models [*Krasnopolsky et al., 2002*] where acceleration of calculation from 10 to 10^4 times has been achieved, as compared with original parameterizations. In this study, we apply the NN approach to approximating the LW radiation parameterization in NCAR CAM [*e.g., Journal of Climate, 1998*]. Calculation of the LW radiation is the most time consuming part of the atmospheric physics calculations. It takes about 70% of time required for calculation of model physics and, therefore, about 60% of the total model calculation time.

There are two major approaches in developing NNs for model physical processes, following either the physical or the computational structure the scheme. Following a physical structure is a logical or preferable approach in the case of developing new or modified parameterizations using mostly observational and possibly other data. Such an approach has been used by Chevalier [1998] (see also [*Krasnopolsky and Chevalier 2001, 2003*]) to develop the highly efficient NN based LW radiation parameterization (NeuroFlux) for the ECMWF model. Because of a strong coupling of clouds and radiation in the parameterization scheme, layer-by-layer NNs were developed that resulted in a battery of 40 NNs describing the LW radiation scheme. However, when developing NN approximations for already existing parameterizations, it is logical and more efficient to follow the computational (rather than physical) structure of a model physics scheme. Practically it means that the computational timing should be assessed and "bottlenecks" (the most time consuming parts of a parameterization) need to be determined. The NN approximation(s) then are applied to speed-up these parts. However, the most efficient and convenient way is developing only one NN approximation for a model physics parameterization. Such an approach has been introduced in this study. One NN approximation has been developed for the NCAR CAM-2 LW radiation parameterization with well-structured inputs and outputs.

NN approximations can be used in model physics because any parameterization can be considered as a continuous or almost continuous mapping (input vector vs. output vector dependence). A NN is a generic tool to approximate such mappings and is an analytical approximation that uses a family of functions like:

$$y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot f(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i); \quad q = 1,2,\ldots,m \qquad (1)$$

where x_i and y_q are components of the input and output vectors respectively, a and b are fitting parameters, and f is a so called activation function (usually it is a hyperbolic tangent), n and m are the numbers of inputs and outputs respectively, and k is the number of neurons in the hidden layer (for more details see appendix in [*Krasnopolsky et al., 2002*]).

2. NN APPROXIMATION FOR THE NCAR CAM LONG WAVE ATMOSPHERIC RADIATION

The function of the LW radiation parameterization in atmospheric GCMs is to calculate heat fluxes caused by LW radiation processes in the atmosphere, from which model-layer heating rates are obtained. The complete description of the NCAR CAM-2 atmospheric LW radiation parameterization is presented in [*Collins 2001, 2002*]. Since the calculations of cloudiness are completely separate from the calculations of radiation effects, we are able to approximate the entire LW radiation parameterization with only one NN, with cloudiness being a part of the input vector.

The NN developed for approximation of the NCAR CAM-2 LW radiation parameterization has 101 inputs (n = 101 in eq. (1)), which include six profiles (atmospheric temperature, humidity, ozone concentration, path length for CO_2, path length for H_2O, and cloudiness) and two relevant surface characteristics (surface pressure and upward LW flux on a surface). This NN has 19 outputs (m = 19 in eq. (1)): a profile of the heat rates (HRs) $\{q_k\}_{k=1,\ldots,18}$ and downward LW flux to the surface. The NN has one hidden layer with 90 neurons (k = 90 in eq. (1)) that provide sufficient accuracy of approximation. An important diagnostic parameter, the outgoing LW radiation (OLR) at the top of the atmosphere, is calculated using the NN output HRs $\{q_k\}_{k=1,\ldots,18}$,

$$OLR = G^{-1} \sum_{k=1}^{L} q_k (P_k - P_{k+1}) \qquad (2)$$

where G is a constant and P_k is the atmospheric pressure at the level k, and $L = 18$ (18 vertical levels was used in this study).

A representative data set covering the entire year of 2002, consisting of about 100,000 input/output combinations, has been generated using a single column model with the physics identical to that of NCAR CAM-2. The eighteen level single column model is run for one time step to generate the input/output data set using the NCAR/NCEP reanalysis initial conditions for the first day of each month of the year (every six hours). This dataset is divided into three parts, each containing approximately 33,000 input/output combinations. The first part is used for NN training, the second one is used for tests (control of overfitting, control of a NN architecture, etc.), and the third part is used for validations only.

Table 1*. Accuracy and Computational Performance of LW NN Approximation for NCAR CAM-2 and the ECMWF Model vs. their Corresponding Original Parameterizations*

Parameter	Model	Bias	RMSE	Mean	σ	Performance
HR (*K/d*)	ECMWF	0.2	0.45			8 times faster
	NCAR	2. 10^{-4}	0.05	-1.43	1.76	65 times faster
OLR (*Wt/m 2*)	ECMWF	0.8	1.9			
	NCAR	0.01	0.9	240.5	46.9	

Table 1 shows a bulk validation statistics for the accuracy of the NN approximation and its computational performance. The accuracy and performance of the ECMWF NeuroFlux approximation is also shown for comparison. The NN approximation has been evaluated against the original parameterization. In order to calculate error statistics presented in Table 1, both the original parameterization and its NN approximation have been applied to the validation data. Two sets of the corresponding HR profiles and OLRs have been generated. Bias (or mean error) and RMSE presented in Table 1 have been calculated as the mean differences between these two sets of HRs and OLRs. Mean value and standard deviation (s) of HRs and OLRs are presented for a better understanding of relative errors. The ECMWF results for the NeuroFlux are also shown for comparison. Our NN approximation has very high accuracy with an almost negligible systematic error (bias). Most importantly, it performs 65 times faster than the original parameterization. This speed-up is achieved for NN approximation of the entire LW radiation scheme and includes calculations of optical properties (emissivity and absorptivity), as well as HRs and radiative fluxes. During a normal execution of CAM-2 the relatively expensive emissivity and absorptivity calculations are made every 12 hours, while HRs and radiative fluxes (about 1% of the computing time needed for the entire

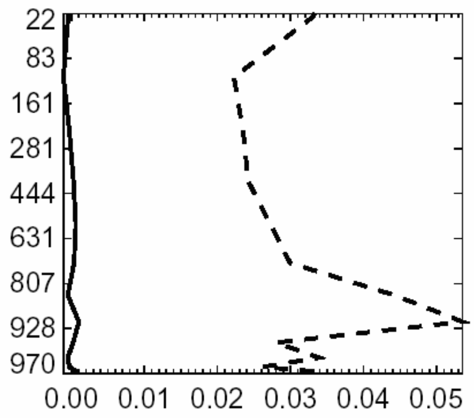

Fig. 1

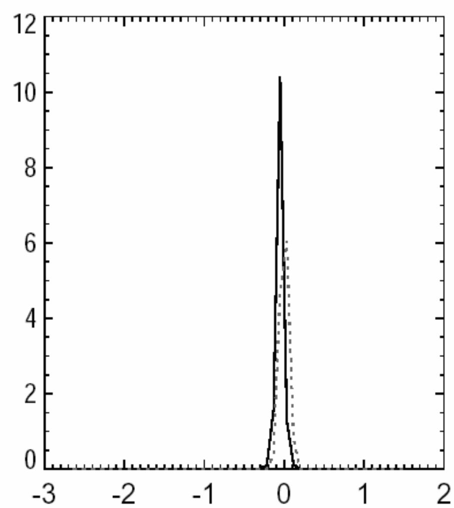

Fig.1 *(the upper pannel) Profiles of bias (solid line) and RMSE (dashed line) for HRs. The vertical axis units are in mb. At each vertical level, both bias and RMSE are normalized using the standard deviation of the HRs at the same vertical level. (the lower panel) Distribution of errors for HRs (solid line). The horizontal axis shows HRs in Kelvin per day. The normal distribution with the same mean value and standard deviation (dotted line) is presented for comparison.*

LW radiation scheme) are calculated every hour. Our speed-up for the entire LW radiation scheme provides the opportunity of calculating optical properties every hour, i.e. as frequently as calculations of HRs and radiative fluxes.

Since both the original parameterization and the NN approximation are very complicated multidimensional objects (mappings), calculating bulk statistics is not sufficient for evaluating the accuracy of the approximation. We evaluated many different statistical metrics of the approximation accuracy.

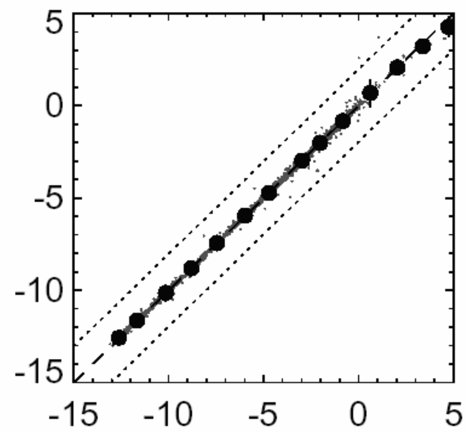

Fig. 2

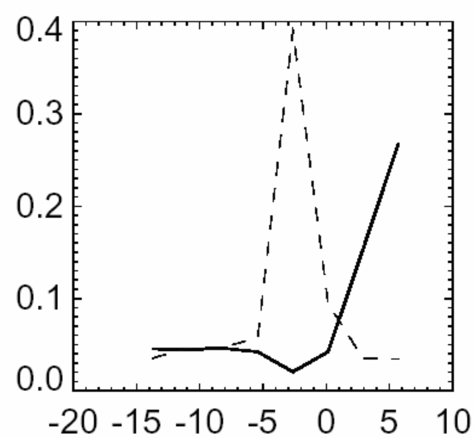

Fig.2. *The fourth vertical level. The upper panel shows a scatter plot (small black dots) and binned scatter plot (large black dots) for HRs generated by the original parameterization (horizontal axis) and its NN approximation (vertical axis). Two dotted lines show one standard deviation for the HRs at this vertical level. The lower panel shows RMSE as a function of HRs. The dashed line shows the distribution of occurrences.*

Fig. 1 (the upper panel) shows vertical profiles for bias and RMSE in units of standard deviations for each particular level. The bias is practically zero (see Table 1). The RMSEs are also very small; they do not exceed 5-6% of the standard deviations at the corresponding level. Fig. 1 (the lower panel) shows the distribution of errors in approximating HRs. The distribution is strongly peaked about 0. K/d. It is very close to the normal distribution with the same mean and standard deviation. Fig. 2 demonstrates other characteristics of the approximation accuracy statistics. The figures correspond to the fourth (from the ground) vertical layer of the NCAR CAM-2, which

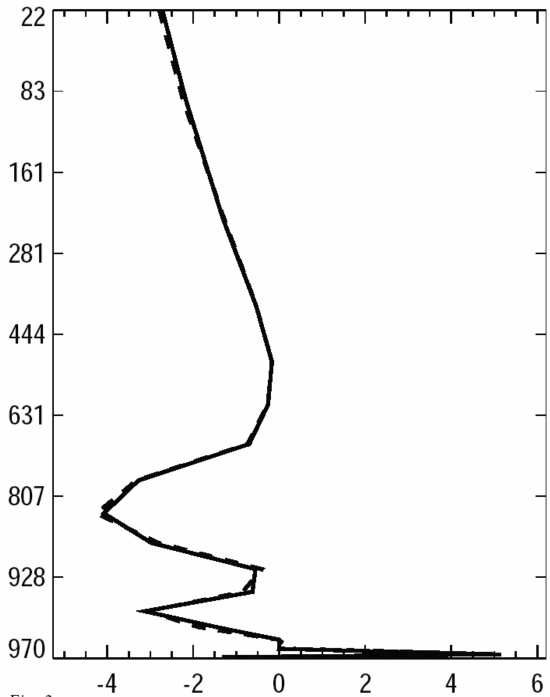

Fig. 3

***Fig. 3**. An instantaneous HR profile (solid) with a complicated cloudiness and its NN approximation (dashed). The vertical axis units are in mb. The horizontal axis shows HRs in Kelvin per day.*

is heavily affected by cloudiness. The upper panel shows the scatter plot for HRs and the lower panel shows the approximation RMSE as a function of HRs; it also shows the distribution of occurrences. The approximation errors are small and random (the systematic error or bias is very small) in the areas well supported by data. In the areas where there are little data (tails of the distribution) errors increase. In these areas the NN is forced to extrapolate. These areas should be enriched by simulated data to improve the accuracy of the NN approximation there, namely, the original parameterization should be run in this sub-domain to generate more data.

Fig. 3 shows an instantaneous HR profile and its NN approximation. The profile demonstrates a very complicated vertical distribution including a significant amount of cloudiness. It is noteworthy that the NN approximation is still very close to the original HRs even for complicated cloudiness profiles.

The NN approximation of the NCAR CAM-2 LW radiation parameterization shows performance and accuracy comparable with highly efficient ECMWF NeroFlux (see Table 1). The NN approximation of the NCAR CAM-2 LW radiation parameterization has a simpler architecture. A simpler NN architecture could be introduced because of a clear separation between calculations of cloudiness and radiative processes in the original NCAR CAM-2 LW radiation parameterization. The original ECMWF radiation parameterization did not have this more straightforward modular structure.

3. CONCLUSIONS

In the study, we evaluated the accuracy and performance of a NN approximation developed for the LW radiation parameterization in NCAR CAM. We selected the LW radiation because it is the most time consuming component of the model physics. Application of this approach allows calculation of LW radiation 65 times faster than the original parameterization without compromising the accuracy of approximation. The systematic error introduced by the NN approximation is negligible. The random error is also very small and does not exceed several percent of the natural variability of the radiation parameters.

Acknowledgments. The authors would like to thank Drs. W. Collins, P. Rasch, and J. Tribbia for useful discussions and consultations. This study is based upon work supported by the National Science Foundation under Grant No. 0105839, by the Department of Energy Grant ER63197-0007185, and by the NASA Grant No. 622-24-28.

REFERENCES

Chevallier, F., F. Chéruy, N. A. Scott, and A. Chédin, 1998: "A neural network approach for a fast and accurate computation of longwave radiative budget", *Journal of Applied Meteorology*, *37*, 1385-1397

Chevallier, F., J.-J. Morcrette, F. Chéruy, and N. A. Scott, 2000: "Use of a neural-network-based longwave radiative transfer scheme in the EMCWF atmospheric model", *Quarterly Journal of Royal Meteorological Society*, *126*, 761-776

Collins, W.D., 2001: "Parameterization of generalized cloud overlap for radiative calculations in general circulation models", *J. Atmos. Sci., 58*, pp. 3224-3242

Collins, W.D., J.K. Hackney, and D.P. Edwards, 2002: "A new parameterization for infrared emission and absorption by water vapor in the National Center for Atmospheric Research Community Atmosphere Model*", J. Geophys. Res., 107* (**D22**)

Fox-Rabinovitz, M. S., L. L. Takacs, and R. C. Govindaraju, 2002: "A variable-resolution stretched-grid general circulation model and data assimilation system with multiple areas of interest: Studying the anomalous regional climate events of 1998", J. Geophys. Res., 107(D24), 4768, doi:10.1029/2002JD002177

Journal of Climate, 1998: Vol. 11, No. 6 (the special issue).

Krasnopolsky, V.M., and F. Chevallier, 2001: Some neural network applications in environmental sciences. *ECMWF Technical Memorandum No. 359*

Krasnopolsky, V.M., and F. Chevallier, 2003: "Some neural network applications in environmental sciences. Part II: Advancing computational efficiency of environmental numerical models", *Neural Networks*, *16*, 335-348

Krasnopolsky, V.M. and H. Schiller, 2003: "Some neural network applications in environmental sciences. Part I: Forward and inverse problems in satellite remote sensing", *Neural Networks*, *16*, 321-334

Krasnopolsky, V.M., D.V. Chalikov, and H.L. Tolman, 2002: "A neural network technique to improve computational efficiency of numerical oceanic models", *Ocean Modelling*, *4*, 363-383

Krasnopolsky V., 1997: "Neural works for standard and variational satellite retrievals", *Technical Note, OMB contribution No. 148*, NCEP/NOAA

Krasnopolsky, V.M., et al., 2000: "Application of neural networks for efficient calculation of sea water density or salinity from the UNESCO equation of state", *Proceedings, Second Conference on Artificial Intelligence*, AMS, Long Beach, CA, 9-14 January, 2000, pp. 27-30

Rabitz, H., O.F. Alis, J. Shorter, and K. Shim, 1998: Efficient input-output model representations, *Comput. Phys. Comm.*, *115*, 1-10;

Rabitz, H., and O.F. Alis, 1999: General foundation of high dimensional model representations, *J. Math. Chem., 25*, 2-3

Schoendorf, J., H. Rabitz, and G. Li, 2003: A fast and accurate operational model of ionospheric electron density, *Geophys. Res. Lett., 30*, 1492-1495