

## New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Longwave Radiation in a Climate Model

VLADIMIR M. KRASNOPOLSKY

*NOAA/NCEP/SAIC, Camp Springs, and Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland*

MICHAEL S. FOX-RABINOVITZ AND DMITRY V. CHALIKOV

*Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland*

(Manuscript received 26 April 2004, in final form 25 October 2004)

### ABSTRACT

A new approach based on a synergetic combination of statistical/machine learning and deterministic modeling within atmospheric models is presented. The approach uses neural networks as a statistical or machine learning technique for an accurate and fast emulation or statistical approximation of model physics parameterizations. It is applied to development of an accurate and fast approximation of an atmospheric longwave radiation parameterization for the NCAR Community Atmospheric Model, which is the most time consuming component of model physics. The developed neural network emulation is two orders of magnitude, 50–80 times, faster than the original parameterization. A comparison of the parallel 10-yr climate simulations performed with the original parameterization and its neural network emulations confirmed that these simulations produce almost identical results. The obtained results show the conceptual and practical possibility of an efficient synergetic combination of deterministic and statistical learning components within an atmospheric climate or forecast model. A developmental framework and practical validation criteria for neural network emulations of model physics components are outlined.

### 1. Introduction

Tremendous developments in numerical modeling and in computing capabilities during the last decades contributed dramatically to the scientific and practical significance of interdisciplinary climate, climate change, and weather prediction. One of the main problems of implementing the best atmospheric, ocean, and chemistry models is the complexity of physical, chemical, and biological processes involved. Parameterizations of model physics are approximate subgrid-scale schemes based on simplified 1D physical process equations and observational data. Still, these parameterizations are so time consuming, even for most powerful modern supercomputers, that different kinds of additional simplifications are usually applied. For example, some of these parameterizations are calculated less frequently than model dynamics (based on solving 3D geophysical fluid dynamics equations). This may negatively affect the accuracy of model physics calculations and its temporal consistency with model dynamics and may lead to a

significant reduction of the accuracy of climate simulations and especially weather predictions.

Calculation of model physics in a typical moderate-resolution general circulation model (GCM), like the National Center for Atmospheric Research (NCAR) Community Atmospheric Model (CAM-2) with T42 ( $\sim 3^\circ$ ) resolution and 26 vertical levels, takes about 70% of the total model computations. Higher uniform and variable model resolutions (e.g., Fox-Rabinovitz et al. 2001, 2002; Duffy et al. 2003) 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 for looking for new alternative numerical algorithms that provide faster and, most importantly, very accurate ways of calculating model physics and chemistry. For example, a traditional statistical technique based on a representation of the input/output relationship as an expansion of hierarchical correlated functions has been investigated in some atmospheric chemistry applications (see Schoendorf et al. 2003, and references therein). However, current climate and weather prediction models are complex and nonlinear and they require a much higher accuracy and better flexibility of approximation

---

*Corresponding author address:* Vladimir Krasnopolsky, 5200 Auth Rd., Camp Springs, MD 20746-4304.  
E-mail: Vladimir.Krasnopolsky@noaa.gov

than those provided by traditional statistical techniques, which are appropriate for simpler applications.

During the last decade, neural network (NN) techniques have found a variety of applications in different fields and, more specifically, in the accurate and fast modeling of atmospheric radiative processes (Krasnopolsky 1997; Chevallier et al. 1998) and in satellite retrieval procedures (Krasnopolsky et al. 1995; Krasnopolsky and Schiller 2003).

Two different NN-based approaches have been developed to speed up calculations of model physics. The first approach developed by Chevallier et al. (1998, 2000) introduces NNs as a convenient tool in the traditional framework of developing new, improved longwave radiation parameterizations. Within his approach, NNs were applied to develop “a new generation of radiative transfer models” (Chevallier et al. 1998). A new NN-based longwave (LW) radiation parameterization (“NeuroFlux”) has been successfully developed for the European Centre for Medium-Range Weather Forecasts (ECMWF) model (Chevallier et al. 1998, 2000). In the NeuroFlux parameterization, the artificial neural network technique was used in conjunction with a classical cloud approximation (the multilayer graybody model). As a result, in this new LW radiation parameterization two NNs were used to compute upward and downward clear-sky parts of LW fluxes, and  $2 \times K$  NNs were used to calculate upward and downward fluxes at each of  $K$  cloud layers. Thus, the developed NeuroFlux is a battery of  $(2 \times K + 2)$  NNs (40 NNs for  $K = 19$ ) (Chevallier et al. 1998). Each NN has different size and is trained separately. NeuroFlux is 8 times faster than the previous parameterization (Chevallier et al. 1998, 2000). NeuroFlux has been used operationally within the ECMWF four-dimensional variational data assimilation (4DVAR) system since October 2003.

A different new approach based on application of NNs has been introduced to ocean models (Krasnopolsky et al. 2000, 2002; Tolman et al. 2004) and as a preliminary study to an atmospheric (Krasnopolsky et al. 2004) model, the NCAR single-column model with the physics identical to that of NCAR CAM-2. This approach introduces an accurate and fast method of calculating the atmospheric physics parameterizations by developing NN *emulations* for *existing* model physics parameterizations. In this approach, the entire parameterization, as a single object (i.e., a continuous or almost continuous mapping), is emulated by an NN. In this case, data used for the NN training are obtained through a GCM simulation with the original parameterization. An NN emulation of a model physics parameterization is a functional imitation of this parameterization, so that the results of model calculations with the original parameterization and with its NN emulation are physically identical. The high quality of NN emulations is achieved because of the high accuracy of approximation of the original components. We prefer

to use the term NN emulation, not NN approximation, to avoid any possible confusion. The term parameterization already means a simplified approximation of physical processes. So, in the context of our approach, the term emulation means a complete functional imitation based on a precise mathematical/statistical approximation (in a classic mathematical sense) of the model physics parameterizations.

The key point is that NN emulation is developed here for the *existing* parameterizations of atmospheric physics. This allows us to preserve the integrity and level of sophistication of the state-of-the-art physical parameterizations of atmospheric processes. Because of the capability of modern machine learning techniques to provide an unprecedented accuracy in the approximation of complex systems like model physics, our NN emulations of model physics parameterizations are practically identical to original physical parameterizations. In other words, the underlying idea of the approach is not developing a new parameterization but rather emulating a parameterization already very carefully tested and validated by its developers offline and then online through experimentation with the entire model. It is achieved by using data for NN training that are simulated by an atmospheric model run with the original parameterization. Using model-simulated data for NN training allows us to achieve an unprecedented accuracy in approximation because simulated data are free of the problems typical in empirical data (problems like high level of observational noise, sparse spatial and temporal coverage, poor representation of extreme events, etc.). In the context of our approach, the accuracy and improved computational performance of NN emulations are always measured against the original parameterization. It is noteworthy that the developed NN emulation has the same inputs and outputs as the original parameterization and is used as its functional substitute in the model.

We would like to clarify that the term “existing” parameterizations means not only the ones currently used in a model but also new, more sophisticated parameterizations that are computationally prohibited in their original form, but will become computationally “affordable” when using their accurate and computationally more efficient NN emulations. Also, by “existing” we mean advanced parameterizations currently under development, like those of cloud physics in cloud-resolving models (also called superparameterizations).

The key objectives or questions of this study are the following: (i) Are these emulations accurate or close enough to the original physical parameterizations so that their use (instead of the original parameterization) allows us to preserve all the richness, integrity, and detailed features of atmospheric physical processes? In other words, is the NN emulation a precise emulation of the original physical parameterization? (ii) Are these emulations fast enough to significantly accelerate

model physics calculations? (iii) Are these statistical/machine learning components able to successfully co-exist or be compatible with the deterministic components of climate models, so that their combination can be efficiently used for accurate and fast climate simulations without any negative impacts on their quality? (iv) Is there a real/productive synergy here? In other words, does this new combination of deterministic and statistical learning approaches lead to new opportunities in climate simulation and weather prediction? An additional objective of this study is to outline a developmental framework and practical criteria for the development and validation of NN emulations for model physics and chemistry components, which could be used as guidelines in the follow-up developments of NN emulations for other components of model physics and chemistry.

More specifically, in this study we apply the NN approach to approximating the LW radiation parameterization in the NCAR CAM-2 (e.g., see the special issue of *Journal of Climate*, 1998, vol. 11, no. 6). Longwave radiation parameterization has been selected for this first study because calculating the LW [and shortwave (SW)] radiation is the most time-consuming part of the atmospheric physics computations. For example, in the NCAR CAM-2 T42, total (LW and SW) radiation takes almost 60% of the time required for the model physics calculations.

The most efficient and convenient way of developing NN emulations for model physics components is developing a single NN for a model physics parameterization. Such an approach was introduced in our preliminary study (Krasnopolsky et al. 2004) wherein we developed NN emulation for the NCAR LW radiation scheme in the framework of the NCAR single-column model. This study showed the feasibility of the approach. In the current study, we extend the developed approach to build an NN emulation for the NCAR CAM-2 LW radiation parameterization. We also performed and analyzed past climate simulations using the NN emulation. A more detailed discussion of its impact on climate simulation, with a comprehensive analysis of climate characteristics and the adjustment of NN emulations to account for climate change, goes beyond the scope of this paper and will be presented later.

It is noteworthy that the initial motivation for this study came from the authors' discussion of the possibility of an effective calculation of model physics on a global *uniform fine* resolution grid for stretched-grid GCMs, instead of calculating model physics on an intermediate, global uniform resolution grid or directly on a stretched grid (e.g., Fox-Rabinovitz et al. 2001, 2002). Calculation of model physics on a global uniform fine resolution grid is feasible only when using the much more computationally efficient NN emulations. These issues will be explored in a separate study.

Neural network emulations of model physics are based on the fact that any parameterization of model

physics can be considered as a continuous or almost continuous (with finite discontinuities) mapping (input vector versus output vector dependence), and NNs (multilayer perceptrons in our case) are a generic tool for approximating such mappings (Cybenko 1989; Funahashi 1989; Hornik 1991; Chen and Chen 1995a,b; Attali and Pagès 1997). Neural networking is an analytical approximation that uses a family of functions like

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

where  $x_i$  and  $y_q$  are components of the input and output vectors respectively,  $a$  and  $b$  are fitting parameters, and  $\phi$  is a so-called activation function (usually 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 (see Ripley 1997 for more details).

In section 2, NN emulation of the NCAR CAM-2 LW radiation parameterization is developed and analyzed in terms of its accuracy and computational performance. In section 3, the use of the developed accurate and fast NN emulation within CAM-2 is validated in terms of its impact on climate simulation. Section 4 contains discussion and conclusions.

## 2. NN emulation of the NCAR CAM longwave atmospheric radiation parameterization

### a. General or background information

The major requirement for developing NN emulations for model physics is obtaining an extremely high accuracy of NN emulations with practically zero biases or systematic approximation errors (i.e., the systematic errors that NN emulation introduces in addition to a bias of the original parameterization). Providing very small additional bias is a necessary condition for assuring that additional errors are not accumulating during long-term climate simulations when using developed NN emulations. The choice of an optimal version of NN emulation is based on the accuracy, not the amount of speedup. All the obtained NN emulations guarantee a very significant speedup anyway.

The T42/26-level NCAR CAM-2 is used in this study. The function of the LW radiation parameterization in atmospheric GCMs is to calculate the heating fluxes and rates produced by LW radiation processes in the atmosphere. The complete description of the NCAR CAM LW radiation parameterization is presented by Collins (2001) and Collins et al. (2002).

The input vectors for the NCAR CAM-2 LW radiation parameterization include 10 profiles [atmospheric

TABLE 1. Statistics estimating the accuracy of HR ( $\text{K day}^{-1}$ ) calculations and computational performance for NCAR CAM-2 LWR using NN emulation vs the original parameterization.  $\text{Bias}_{26}$  and  $\text{RMSE}_{26}$  ( $\text{K day}^{-1}$ ) correspond to the lowest layer. Total mean value for HRs =  $-1.36 \text{ K day}^{-1}$  and standard deviation  $\sigma_{HR} = 1.93 \text{ K day}^{-1}$ . For the lowest level (26th), mean value for HRs =  $-2.22 \text{ K day}^{-1}$  and  $\sigma_{HR} = 5.57 \text{ K day}^{-1}$ . Corresponding statistics ( $\text{K day}^{-1}$ ) for the ECMWF model are shown just as a point of reference.

Model	Bias	RMSE	PRMSE	$\sigma_{\text{PRMSE}}$	$\text{Bias}_{26}$	$\text{RMSE}_{26}$	Performance
NCAR NN90	$-4.0 \times 10^{-4}$	0.33	0.27	0.19	$-6.0 \times 10^{-4}$	0.85	~80 times faster
NCAR NN150	$1.0 \times 10^{-4}$	0.28	0.23	0.16	$-4.0 \times 10^{-3}$	0.79	~50 times faster
NCAR NN200	$5.0 \times 10^{-5}$	0.26	0.21	0.15	$2.0 \times 10^{-3}$	0.71	~35 times faster
ECMWF	0.05	0.1			0.3*	1.4*	~8 times faster

\* These statistics are for the lowest (31st) level in ECMWF model.

temperature, humidity, ozone,  $\text{CO}_2$ ,  $\text{N}_2\text{O}$ ,  $\text{CH}_4$ , two chlorofluorocarbon (CFC) mixing ratios (the annual mean atmospheric mole fractions for halocarbons), pressure, and cloudiness] and one relevant surface characteristic (upward LW flux at the surface). The CAM-2 LW radiation parameterization output vectors consist of the profile of heating rates (HRs), and several radiation fluxes including the outgoing LW radiation flux from the top layer of the model atmosphere [the outgoing LW radiation (OLR)].

The NN emulation of the NCAR CAM-2 LW radiation parameterization has the same inputs [total 220;  $n = 220$  in Eq. (1)] and the same outputs [total 33;  $m = 33$  in Eq. (1)] as the original NCAR CAM-2 LW radiation parameterization. We have developed several NNs, all of which have one hidden layer with 20, 90, 150, 200, 250, or 300 neurons [ $k = 20, 90, 150, 200, 250, 300$  in Eq. (1)]. Varying the number of hidden neurons allows us to demonstrate the accuracy of approximation dependence on this parameter as well as its convergence, and as a result to provide the sufficient accuracy of approximation for the climate model.

The NCAR CAM-2 was run for 2 yr to generate representative datasets. The first year of simulation was divided into two independent parts, each containing input/output vector combinations. The first part was used for training and the second was used for tests (control of overfitting, control of a NN architecture, etc.). The second year of simulation was used to create a validation dataset, completely independent from both training and test datasets. This third dataset was used for validations only. All approximation statistics presented in the rest of this section are calculated using this independent validation dataset.

### b. Bulk approximation error statistics

To ensure a high quality of representation of long-wave radiation processes, the accuracy of the NN emulations have been carefully investigated. Our NN emulations have been validated against the original NCAR LW parameterization. For calculating the error statistics presented in Table 1 and the figures of this section, the original parameterization and its NN emulation have been applied to the validation data. Two sets of

corresponding HR profiles have been generated. The total and level bias (or mean error), total and level RMSE, profile RMSE (PRMSE), and  $\sigma_{\text{PRMSE}}$  presented in Table 1 have been calculated as follows. The outputs of the original parameterization and NN emulation can be represented as  $Y(i, j)$  and  $Y_{\text{NN}}(i, j)$ , correspondingly, where  $i = (\text{lat}, \text{lon})$ ,  $i = 1, \dots, N$  is the horizontal location of a vertical profile,  $N$  is the number of horizontal grid points, and  $j = 1, \dots, L$  is the vertical index, where  $L$  is the number of the vertical levels.

The mean difference,  $B$  (bias or a systematic error of approximation), and the root-mean-square difference (a root-mean-square error of approximation), RMSE, between the original parameterization and its NN emulation, are calculated as follows:

$$B = \frac{1}{NL} \sum_{i=1}^N \sum_{j=1}^L [Y(i, j) - Y_{\text{NN}}(i, j)]$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^L [Y(i, j) - Y_{\text{NN}}(i, j)]^2}{NL}}. \quad (2)$$

These two characteristics [Eqs. (2)] describe the accuracy of the NN emulation integrated over the entire 4D (latitude, longitude, height, and time) dataset. Using a slight modification of Eqs. (2), bias and RMSE for  $m$ th vertical level can be calculated:

$$B_m = \frac{1}{N} \sum_{i=1}^N [Y(i, m) - Y_{\text{NN}}(i, m)]$$

$$\text{RMSE}_m = \sqrt{\frac{\sum_{i=1}^N [Y(i, m) - Y_{\text{NN}}(i, m)]^2}{N}}. \quad (3)$$

The root-mean-square error can also be calculated for each  $i$ th profile:

$$\text{prmse}(i) = \sqrt{\frac{1}{L} \sum_{j=1}^L [Y(i, j) - Y_{\text{NN}}(i, j)]^2}. \quad (4)$$

This error is a function of a horizontal location of the profile. It can be used to calculate mean profile root-

mean-square error, PRMSE, and its standard deviation,  $\sigma_{\text{PRMSE}}$ , which are location independent:

$$\text{PRMSE} = \frac{1}{N} \sum_{i=1}^N \text{prmse}(i)$$

$$\sigma_{\text{PRMSE}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [\text{prmse}(i) - \text{PRMSE}]^2}. \quad (5)$$

Table 1 shows that the profile (PRMSE) error and the RMSE are close but not equal.

Figure 1 illustrates improvements (convergence) in four error statistics when the number of hidden neurons,  $k$ , increases from 20 to 300. These statistics are RMSE (2), RMSE<sub>26</sub> [ $m = 26$  in Eq.(3)], PRMSE, and  $\sigma_{\text{PRMSE}}$  (5). The 26th, or lowest model level errors, are presented because they are the maximum errors for the entire vertical profile. After a sizeable improvement for  $k$  increasing from 20 to 90, the errors practically reach convergence for  $k = 150$ , and even for  $k = 90$ . All our NN emulations have almost zero or negligible systematic errors (biases), which (see Fig. 2) practically do not depend on height and are indistinguishable from each other for the figure scale. The rest of Fig. 2 shows the vertical profiles of RMSEs (2) for six developed NNs. For all NNs with the number of hidden neurons starting at 90, the RMSE (which is a purely random error in the case of a zero bias) for the 10 upper levels does not exceed  $0.2 \text{ K day}^{-1}$ , reaching  $0.4 \text{ K day}^{-1}$  at the 22d level. For the two lowest levels, RMSE is about  $0.6\text{--}0.8$

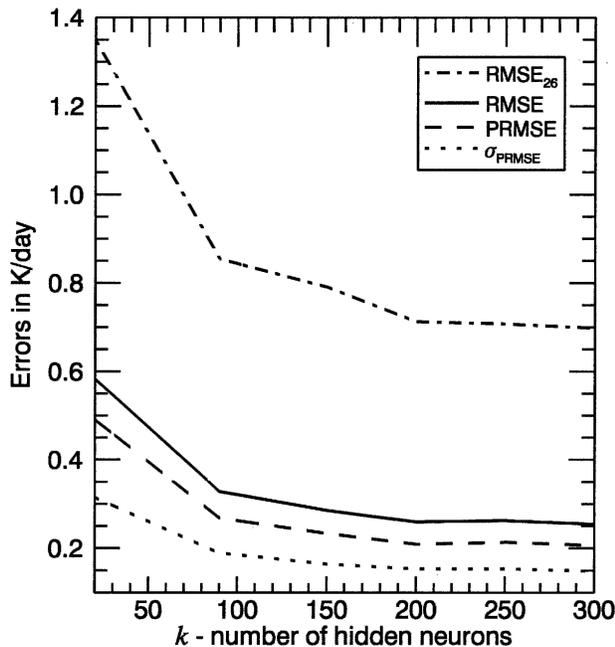


FIG. 1. Convergence of four error statistics when the number of hidden neurons increases from 20 to 300. RMSE [(2)]—solid; PRMSE [(5)]—dashed;  $\sigma_{\text{PRMSE}}$  [(5)]—dotted; and RMSE<sub>26</sub> [(3)]—dotted-dashed lines.

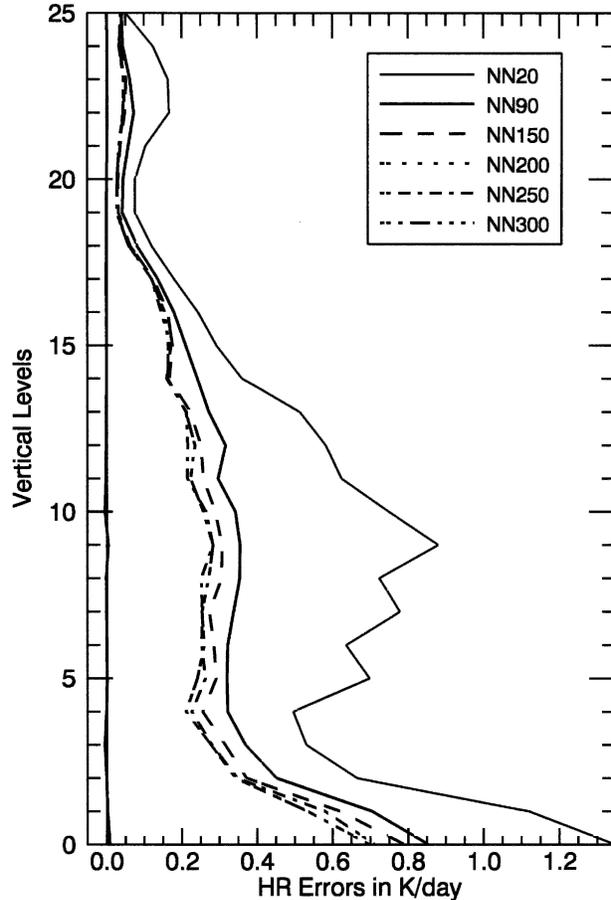


FIG. 2. The vertical profiles of mean approximation errors at each of 26 levels, level biases (the left vertical line), and level RMSEs [(3)], for six developed NNs (NN20—thin solid; NN90—thick solid; NN150—dashed; NN200—dotted; NN250—dashed-dotted; NN300—dashed-double-dotted lines) all in  $\text{K day}^{-1}$ .

$\text{K day}^{-1}$ . Statistics (biases and RMSEs) for the lowest (26th) level are also included in Table 1. The natural variability ( $\sigma$ ) of the HRs is significantly higher (see in the caption of Table 1) at the lowest levels than at the higher ones. Hence, relative errors in the HRs calculated with respect to the natural variability ( $\sigma$ ) are approximately the same as errors at the higher levels.

Hereafter, NN20, NN90, NN150, NN200, NN250, and NN300 stand for NNs with 20, 90, 150, 200, 250, and 300 hidden neurons. Because NN20 is less accurate, and NN250 and NN300 do not provide a significantly better accuracy than NN200, only three NNs, namely NN90, NN150, and NN200, are included in Table 1. Table 1 shows bulk validation statistics for the accuracy of approximation and computational performance for the three best (in terms of accuracy and performance) developed NN emulations. Here, statistics for comparison of the ECMWF operational and NeuroFlux LW radiation parameterizations are also shown just as a reference point. Mean values and standard deviations ( $\sigma_{\text{HR}}$ )

of HRs are presented in the title of Table 1 for a better understanding of relative errors.

In addition to this high approximation accuracy, our NN emulation performs about 80–35 times faster (for NN90, NN150, and NN200, correspondingly) than the original NCAR LW radiation parameterization. Table 1 and Figs. 1 and 2 clearly demonstrate a systematic improvement in approximation accuracy with increasing NN hidden layer size. Table 1 also demonstrates a reciprocal reduction in performance gain, from 80 to 35 times faster, than the original parameterization. This offers an opportunity for the accuracy versus performance trade-off; however, as we mentioned earlier, in this trade-off the key requirement, which allows the successful, synergetic functioning of NN emulation within the model, is to preserve the accuracy and integrity of the description of the corresponding physical process.

All three NNs (including NN90) provide an accurate approximation of the original parameterization. Actually, there is no difference between 1%, 2%, or 3% cost of the original parameterization calculations with NN90, NN150, or NN200 (which are 80, 50, or 35 times faster). Therefore, we would recommend and use the accurate and reliably converged version. Obviously, the final decision on the optimal version of the NN, which has to be implemented into the model, should be made based on testing these NNs in a climate model (see section 3). We would like to stress that the speedup is achieved for the NN emulation of the entire LW radiation scheme that includes calculations of optical properties (emissivity and absorptivity), HRs, and radiative fluxes. Our speedup for the entire LW radiation scheme provides the opportunity of calculating it hourly, that is, as frequently as HRs and radiative flux calculations, at a very limited computational cost.

### c. Detailed evaluation of approximation errors

Both the original parameterization and its NN emulation are complicated multidimensional objects (mappings). In this case, calculating bulk statistics is not sufficient for evaluating the accuracy of the approximation. We evaluated many different statistical metrics of the approximation accuracy, the most important of which are shown in Figs. 1–6. Figure 3 shows the typical level statistics (for level number 20 from the top) based on NN150 emulation. It contains the scatterplot  $Y$  versus  $Y_{\text{NN}}$  (upper-left panel), the distribution of errors or differences  $Y(i, j) - Y_{\text{NN}}(i, j)$  (lower-left panel), and bias and RMSE as functions of HRs (the upper- and lower-right panels, correspondingly). The scatterplot cloud contains hundreds of thousands of points, which are tightly concentrated along the diagonal, with only several points at the upper and lower ends located outside the one sigma interval (marked by dotted lines). The error distribution is strongly peaked about 0 K day<sup>-1</sup>, and it is narrower than the normal distribution

with the same mean and standard deviation (dotted line), which indicates a lesser amount of larger errors than in the case of the normal error distribution. Similar behavior takes place in the distribution of the prmse profile rms errors [see Eq. (4)] shown in Fig. 4. This distribution is also significantly sharper than the normal one with the same mean value and standard deviation.

The two right panels in Fig. 3 show both systematic (bias) and random (RMSE) approximation errors as functions of HRs. An increase in errors at the tails of the HR distribution (shown with the dashed line), where NN was exposed to an insufficient amount of training data, can be clearly seen. More training data should be simulated in this part of the domain to improve the approximation accuracy there.

Figure 5 shows the absolute zonal mean bias (left column) and zonal mean RMSE (right column). Three rows correspond to three different NNs, namely NN90, NN150, and NN200, from top to bottom, correspondingly. When the accuracy of approximation increases (with increasing the number of hidden neurons in the NN), both zonal mean bias and RMSE decrease significantly. Comparing the top, middle, and bottom panels, we see that small areas with bias  $>0.01$  K day<sup>-1</sup> (left column) in the lower part of the atmosphere disappear completely. Also, the small areas of RMSE  $>0.25$  K day<sup>-1</sup> (right column) disappear at the upper levels. In the lower part of the atmosphere, small areas with RMSE  $>1$  K day<sup>-1</sup> (right column) begin to disappear for NN150, and even more so for NN200, and the areas with RMSE  $>0.5$  K day<sup>-1</sup> are confined to just two small spots located in the polar areas.

Figures 6a–c show three typical individual profiles with profile rms errors [prmse; Eq. (4)] close to their mean [prmse; Eq. (5)]. Each of these profiles demonstrates a complicated vertical variability for original and NN emulation profiles that are very close to each other. There is an obvious convergence of the emulating profiles (gray) to the original (black). The prmse for each profile systematically improves when the number of neurons in the NN hidden layer increases from 90 to 200. This convergence, however, is not uniform at some vertical levels. For example, at level 14 in Fig. 6a, NN150 (gray dashed line) is slightly closer to the original profile (black solid line) than NN200 (gray dotted line). It is remarkable that for all individual profiles all the NN emulations are very close to the original profiles. It shows a high uniform convergence and accuracy at the profile and even gridpoint level.

The analysis of approximation errors presented above shows that the NN technique is capable of providing NN emulations with practically zero systematic errors or biases, and small random errors. Moreover, the distributions of errors are usually narrower than the corresponding normal distributions, which indicate a significantly smaller amount of larger errors. An additional analysis shows that larger errors are located in

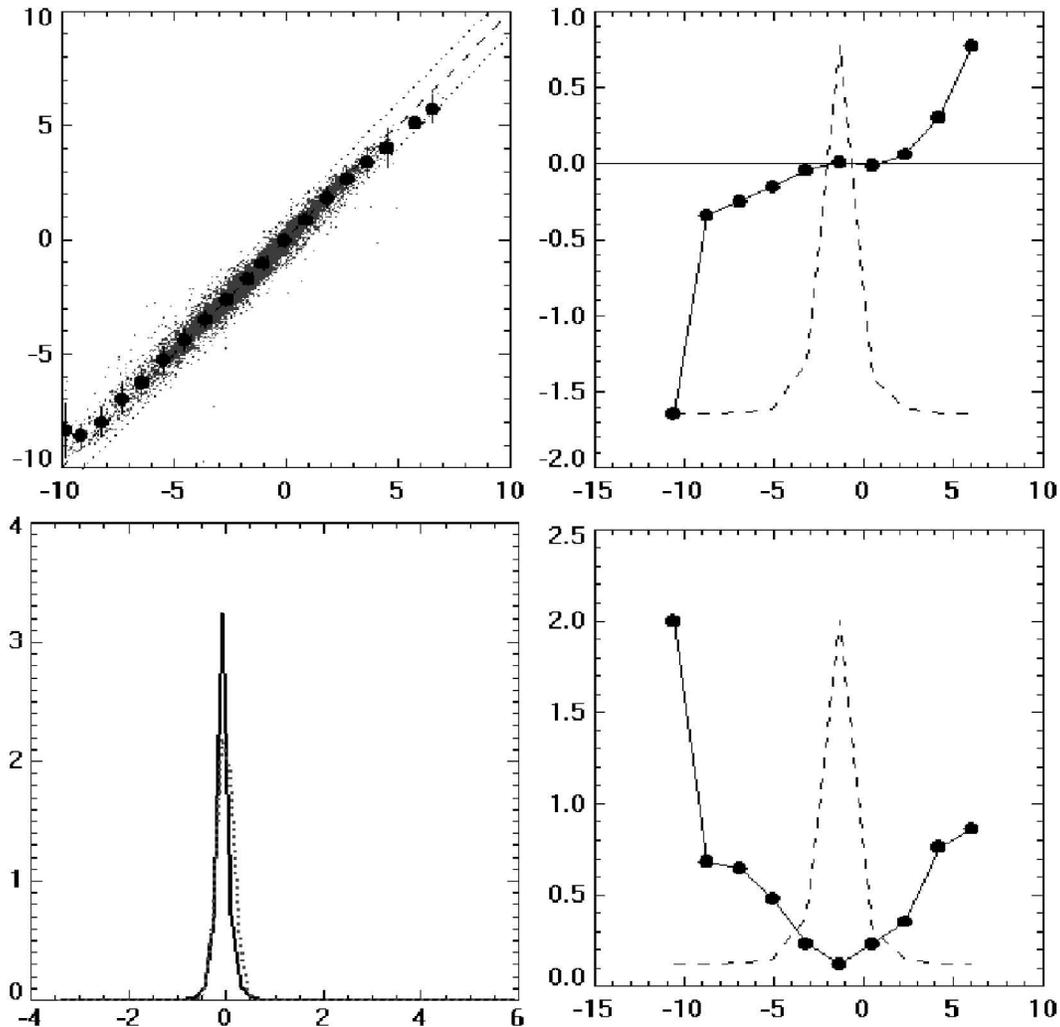


FIG. 3. Typical level statistics (for the 20th from the top level, about 625 hPa). (upper left) The scatterplot  $Y$  (original HRs) vs  $Y_{NN}$  (NN150 HRs) both in  $\text{K day}^{-1}$ . Large dark circles and associated bars show average in the bin and an error bar. (lower left) The distribution of errors or differences  $Y(i, j) - Y_{NN}(i, j)$  and the normal distribution with the same mean and standard deviation (dotted line); horizontal axis corresponds to errors in  $\text{K day}^{-1}$ . (upper right) Bias and (lower right) RMSE ( $\text{K day}^{-1}$ ) as functions of HRs ( $\text{K day}^{-1}$ ) with event distribution (dashed lines).

the areas supported by a smaller amount of training data. In these areas NN is forced to extrapolate. These areas should be enriched by additional simulated data to improve the accuracy of the NN emulation there. Namely, the original parameterization should be run in this subdomain to generate more data.

### 3. Results of climate simulations

In assessing the impact of using an NN emulation of the LW radiation parameterization, the parallel NCAR CAM-2 climate simulations were performed with the original LW radiation parameterization (the control run) and with its NN emulations, which are described in section 2. The climate simulations have been run for 10

yr starting after the training and validation period (see section 2), namely for years 3 through 12. All the comparisons of the control and NN emulation runs presented in this section are done by analyzing the time- (10 yr) mean differences between the results of different runs.

Preservation of time means of prognostic and diagnostic fields is one of the most important/necessary properties in climate simulations. In the climate simulations performed with the original LWR parameterization and its NN emulations, the time-mean surface pressure is almost precisely preserved. For example, for the NN150 run there is a negligible difference of 0.0001% between the NN and control runs (see Table 2). Other time global means, some of which are also

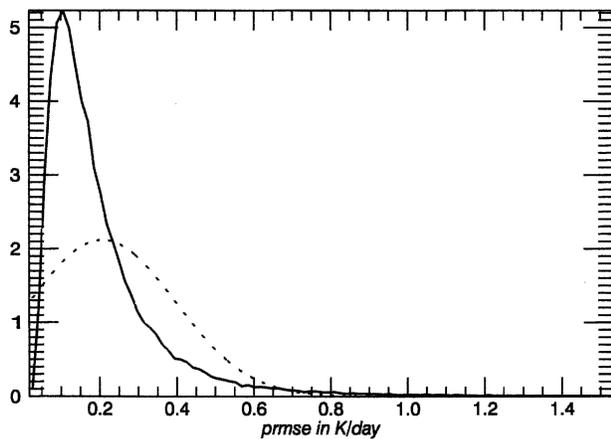


FIG. 4. The distribution of the  $PRMSE [(4)]$ —profile rms errors ( $K \text{ day}^{-1}$ ) and the normal distribution with the same mean and standard deviation (dotted line).

presented in Table 2, show a profound similarity between the simulations in terms of these global time means, with the differences usually within about 0.03%–0.1% and not exceeding 0.1%–0.3%. Other simulations (with NN90 and NN200) show similar results.

Let us now consider some key simulated diagnostic and prognostic fields and their differences, produced in the control and NN emulation runs. The time-mean vertical distributions of zonal means for LW radiation HRs (QRL), potential temperature ( $T$ ), zonal wind ( $U$ ), and specific humidity ( $Q$ ) are presented in Figs. 7–10. The control simulation (with the original LWR parameterization) is shown in (a); (b) and (c) are two simulations, with NN90 and NN150 emulations, correspondingly; (d) and (e) are biases or deviations of these NN90 and NN150 simulations from the control simulation or (b) – (a) and (c) – (a), correspondingly. Therefore, the biases are calculated against the control run. The general assessment of the presented field distributions and their biases is as follows. The field distributions for the control and NN emulation runs are very close, showing a striking similarity to each other that can be seen by the comparing (a) with (b) and (c) for Figs. 7–10. Such a profound pattern similarity is further confirmed and quantified by the small biases shown in (d) and even smaller, almost negligible biases shown in panel (e) of Figs. 7–10.

Now let us discuss the results in more detail. The LWR HRs for the NN90 run (Fig. 7b) show four minor spots located within the 400–500-hPa layer that are not present in the control (Fig. 7a) and the NN150 (Fig. 7c) runs. These differences are more visible in the bias pattern of the NN90 run (Fig. 7d). Actually, small positive and negative biases, in the 0.05–0.2  $K \text{ day}^{-1}$  range, are present in the troposphere. For some spots, within the lower-tropospheric 900–1000-hPa layer, the bias increases to 0.6–1.0  $K \text{ day}^{-1}$  (Fig. 7d). For the NN90 run,

bias can be reduced by extending training datasets for the tails of the distribution as mentioned above in section 2. However, for the NN150 run (Fig. 7e), the bias pattern has cleared up significantly (bias is mostly within 0.01–0.05  $K \text{ day}^{-1}$ ) and contains just a few small positive and negative spots in the lower-tropospheric layer, including those with a maximum magnitude of 0.2–0.4  $K \text{ day}^{-1}$  located near the poles. Bias in the stratospheric domain (above 100 hPa) is significantly smaller than that of the tropospheric domain. For both NN90 and NN150 runs, it is well below 0.05  $K \text{ day}^{-1}$ , especially for the NN150 run (Figs. 7d and 7e), with the exception of polar areas near the top model levels for the NN90 run, where bias reaches 0.1–0.15  $K \text{ day}^{-1}$  within the polar domain (Figs. 7d).

The distributions of temperature, zonal wind, and specific humidity for the control, NN90, and NN150 runs are practically indistinguishable from each other [for Figs. 8–10, cf. (a), (b), and (c)]. The differences between runs can be seen only in bias distributions [cf. (d) and (e) for Figs. 8–10]. Temperature bias for the NN90 run (Fig. 8d) is mostly limited by magnitude to 0.5 K. It increases to a maximum of 1 K for polar domains within the 300–200-hPa layer, and for a few spots at the  $\sim 100$  hPa level. In the middle-stratosphere polar domains, it increases to 1–1.5 K. Bias for the NN150 run (Fig. 8e) is largely reduced everywhere compared to that of the NN90 run (Fig. 8d), mostly to 0.1–0.2 K by magnitude, with the exception of a few very small spots where it approaches 0.5 K.

Zonal wind bias for the entire troposphere and stratosphere domain in the NN90 run (Fig. 9d) is small, mostly below 0.5  $\text{m s}^{-1}$ , and does not exceed 1  $\text{m s}^{-1}$  by magnitude. Bias is slightly larger within the small spot around 25°S at 100 hPa, and it increases to 2.5  $\text{m s}^{-1}$  for the upper model level within the polar domains. For the NN150 run (Fig. 9e), bias is significantly reduced to 0.1–0.2  $\text{m s}^{-1}$ , with a maximum magnitude of only 0.5  $\text{m s}^{-1}$  for a few small spots. It is noteworthy that meridional wind bias for the NN90 run (not shown) is mostly below 0.05  $\text{m s}^{-1}$  and is within 0.1  $\text{m s}^{-1}$  by magnitude. Only in the small areas around the tropical tropopause does bias reach 0.1–0.15  $\text{m s}^{-1}$ , which does not affect the Hadley and Ferrell circulations. Bias for the NN150 run is further reduced to 0.01–0.02  $\text{m s}^{-1}$ , with the exception of a few very small spots where it increases to 0.05  $\text{m s}^{-1}$ .

Specific humidity bias in the NN90 run (Fig. 10d) shows a maximum of 0.3–0.4  $\text{g kg}^{-1}$  around  $\sim 700$  hPa in the equatorial domain. For the NN150 run (Fig. 10e), bias is significantly smaller, mostly 0.01–0.02  $\text{g kg}^{-1}$  by magnitude, with a maximum of 0.05  $\text{g kg}^{-1}$  over a couple of small spots in the subtropical lower troposphere.

The above comparison of biases for NN90 and NN150 runs [see (d) and (e) of Figs. 7–10] confirms that increasing the number of hidden neurons from 90 to 150 leads to a measurable bias reduction that positively

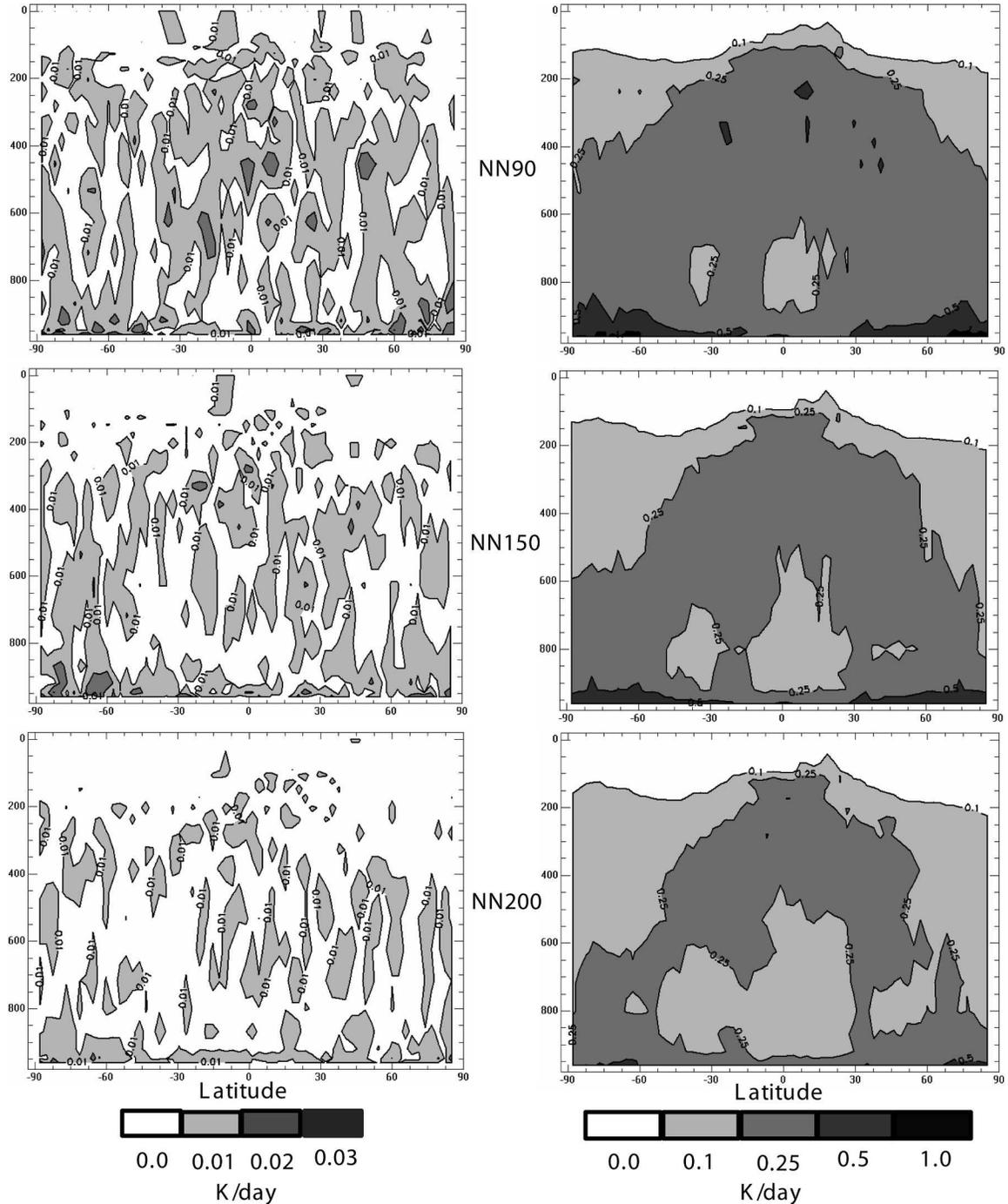


FIG. 5. (left) Absolute zonal mean bias and (right) zonal mean rmse for (top) NN90, (middle) NN150, and (bottom) NN200.

affects the accuracy of the NN150 climate simulation in terms of its profound similarity to the control simulation. Most importantly, biases for both NN90 and NN150 10-yr simulations do not accumulate over time. The LW radiation HRs obtained in these climate simulations maintain the level of approximation accuracy consistent with that obtained in section 2.

#### 4. Discussion and concluding remarks

In this study, we presented a new approach based on a synergetic combination of deterministic modeling and a machine learning technique within an atmospheric model. This approach uses neural networks as a statistical or machine learning technique to develop highly

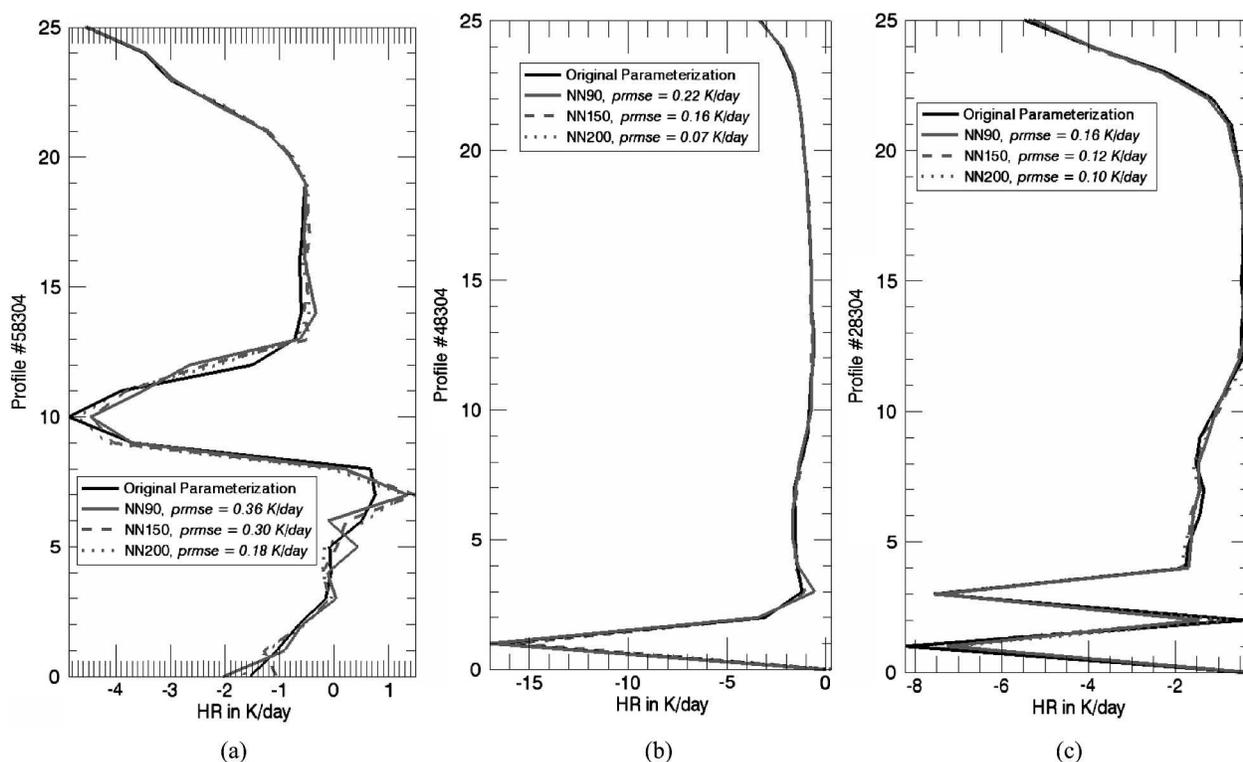


FIG. 6. Typical profiles with  $PRMSE$  [(4)] (a) above, (b) close to, and (c) below  $PRMSE$  [(5)]. Solid black—original profile; solid gray—its NN90 emulation; dashed gray—NN150 emulation; and dotted gray—NN200 emulation.

accurate and fast approximations for model physics components. The approach consists of four major steps:

- 1) Analysis of the structure and complexity of the original parameterization for determining the topology (architecture) of the future NN emulation by specifying all the inputs and outputs and selecting the initial number of neurons in the NN hidden layer.
- 2) Generation of representative datasets for training, testing, and validation. This approach is based on using data simulated during the GCM runs with the original parameterization, which allows us to produce NN emulations that are physically identical to the original parameterizations. For weather prediction applications, the use of blended (simulated, assimilated, and observational) data for NN training could be beneficial. To account for insufficient sampling for some events, it is possible to run the original parameterization offline and generate complementary data to extend sampling. This offline simulation can also be used in the context of adjusting NN emulations for future climate change.
- 3) NN training. Several different versions of NNs, with different number of neurons in a hidden layer, should be trained to determine the optimal size of the hidden layer, which provides the sufficient accuracy of approximation; several initialization procedures and training algorithms should be applied to assure that an optimal minimization is achieved.
- 4) Validation of trained NN emulation consisting of two steps. The first step is validation of the NN approximation against the original parameterization

TABLE 2. Time and global means for mass (mean sea level pressure) and other model diagnostics for the NCAR CAM-2 climate simulations with the original LWR parameterization and its NN emulation (NN150) and their differences (%).

Field	Original LWR parameterization	NN emulation	Difference (%)
Mean sea level pressure (hPa)	1011.480	1011.481	0.0001
Surface temperature (K)	289.003	289.001	0.0007
Total precipitation ( $\text{mm day}^{-1}$ )	2.275	2.273	0.09
Total cloudiness (fractions 0.1 to 1.0)	0.607	0.609	0.3
LWR heating rates ( $\text{K day}^{-1}$ )	-1.698	-1.700	0.1
-OLR ( $\text{W m}^{-2}$ )	234.43	234.63	0.08
Latent heat flux ( $\text{W m}^{-2}$ )	82.84	82.82	0.03

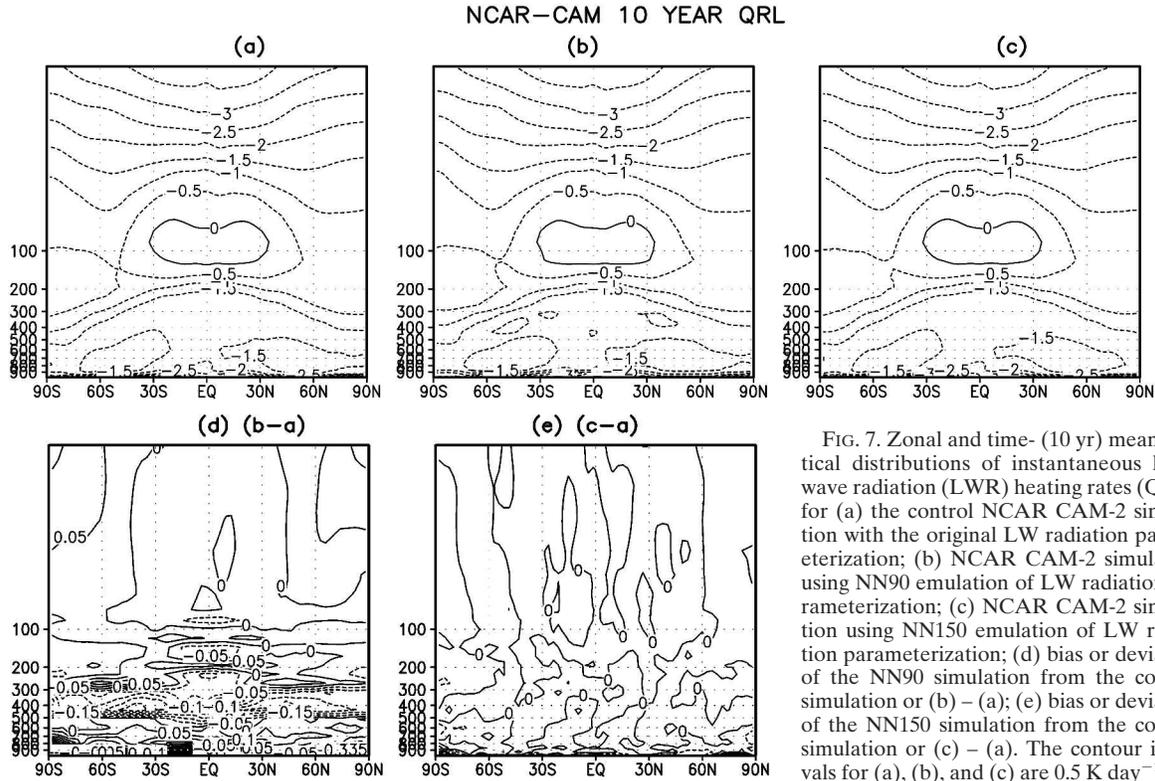


FIG. 7. Zonal and time- (10 yr) mean vertical distributions of instantaneous long-wave radiation (LWR) heating rates (QRL) for (a) the control NCAR CAM-2 simulation with the original LW radiation parameterization; (b) NCAR CAM-2 simulation using NN90 emulation of LW radiation parameterization; (c) NCAR CAM-2 simulation using NN150 emulation of LW radiation parameterization; (d) bias or deviation of the NN90 simulation from the control simulation or (b) – (a); (e) bias or deviation of the NN150 simulation from the control simulation or (c) – (a). The contour intervals for (a), (b), and (c) are  $0.5 \text{ K day}^{-1}$  and for (d) and (e)  $0.05 \text{ K day}^{-1}$ .

using an independent validation dataset. The second validation step consists of performing comprehensive parallel model runs with the original parameterization (the control run) and the developed NN emulations.

In this study, we presented an NN emulation of an atmospheric LW radiation parameterization used in NCAR CAM-2. The LW radiation has been selected as the most time-consuming component of the NCAR model physics. We evaluated the accuracy and computational performance of this NN emulation. The obtained results show the following:

- (i) The conceptual and practical possibility of developing an accurate NN emulation of model physics components, which preserves the integrity and all the detailed features of atmospheric physical processes. The practical possibility of experimentations with NNs with hundreds of inputs, hundreds of neurons in a hidden layer, and tens or more of outputs, using a training dataset with about 100 000 records. The efficient software for a standard workstation without any hardware acceleration has been developed, allowing such experimentations.
- (ii) These accurate NN emulations are very fast (up to 80 times) despite the large size of the corresponding NNs, so the significant speedup of model physics calculations can be achieved without compromising its accuracy.

- (iii) These statistical/machine learning techniques can be successfully combined with deterministic climate model components so that their synergy can be efficiently used for climate simulations without any negative impacts on simulation quality as it was shown by the presented decadal climate simulation.
- (iv) This productive synergy or the new combination of the state-of-the-art deterministic and statistical learning approaches leads to new opportunities in climate simulation and weather prediction. For example, more accurate and more sophisticated atmospheric parameterizations, which may be currently computationally prohibited because they are too time consuming even for most powerful supercomputers available, may exist or be developed in the future. After developing NN emulations for these parameterizations they may become computationally feasible.

The systematic error introduced by NN emulation is negligible and does not accumulate over the model integration in time. The random error for NN emulation is also small, as shown in section 2. The distributions of error are usually narrower than the corresponding normal distributions, which indicates a significantly smaller amount of larger errors. The application of this approach allows us to accelerate the calculation of the LW

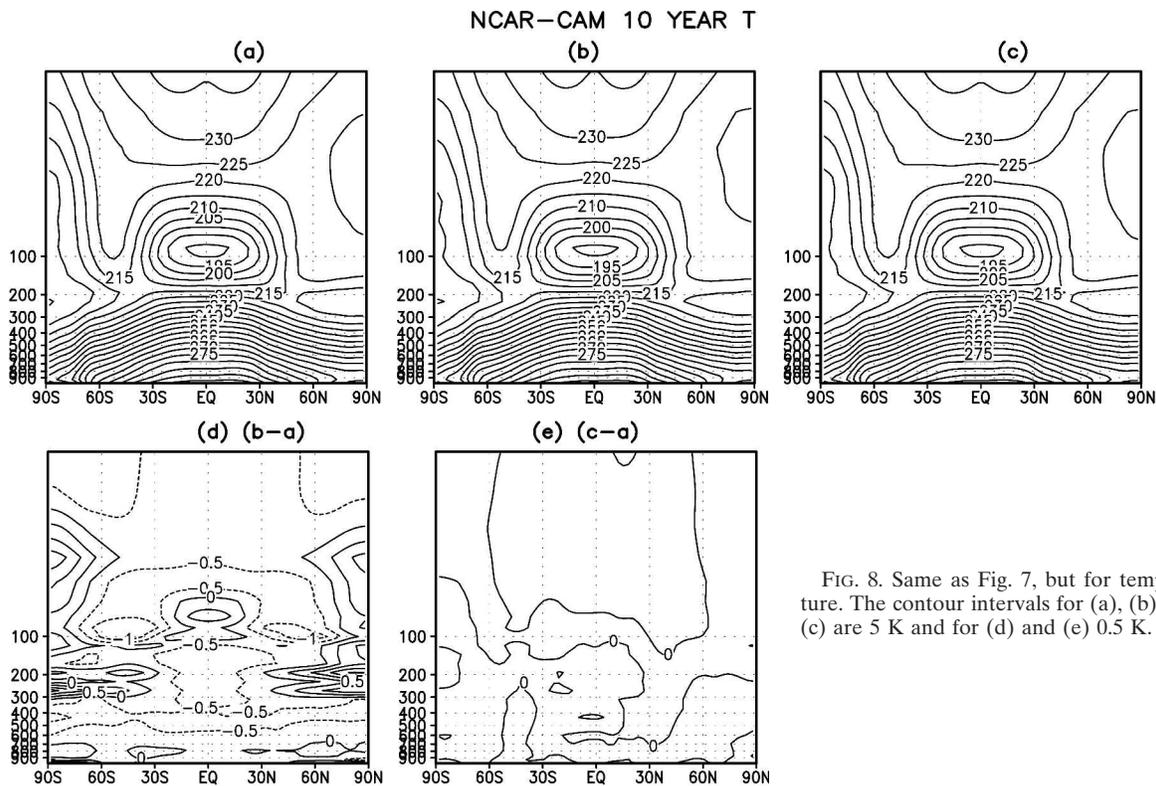


FIG. 8. Same as Fig. 7, but for temperature. The contour intervals for (a), (b), and (c) are 5 K and for (d) and (e) 0.5 K.

radiation parameterization by about 50–80 times (or takes only 1%–2% of the original parameterization computation time) for NN150 and NN90, correspondingly, without compromising the accuracy and integrity of the original longwave radiation parameterization.

The impact of using NN emulation on climate simulation has been assessed by a comparison of some basic climate characteristics of parallel NCAR CAM-2 simulations, calculated with the original LW radiation parameterization and its NN emulations. The differences between the simulations with the original LW radiation parameterization and its NN emulations appear to be very small for simulated fields. The results obtained show that the NN emulation of the considered atmospheric LW radiation parameterization is accurate and provides a significantly improved computational efficiency.

There are several important topics—like adjustments of NN emulations to account for climate change, an explicit evaluation of NN emulation Jacobian, and dealing with giant NNs in the case of higher resolutions—that go beyond the scope of this study. However, these topics should and will be addressed in our follow-up investigations. In this methodological study, we dealt with the past climate. In the follow-up efforts we will consider different options (like, e.g., extensions of the training domain, using recurrent NNs, and adopting some control theory tools), which machine learning techniques can provide, to adjust NN emulations to climate change.

In Krasnopolsky et al. (2002) we demonstrated the acceptable quality of the NN emulation Jacobian for moderate-size NN emulations. For the large-size NN emulations, an explicit investigation of the quality of the NN emulation Jacobian should be conducted. When the Jacobian is a very highly dimensional object, a special study is desirable, like that of Chevallier and Mahfouf (2001). In our case, the high accuracy of approximation and interpolation demonstrated by our NN emulation on an independent validation dataset and during the parallel run of NCAR CAM-2 demonstrates, although indirectly, the practically sufficient accuracy of Jacobians in terms of the considered application.

Increasing the model resolution will result in increasing the size of NNs. Training such NNs will become more and more time consuming. In our previous studies (Krasnopolsky et al. 2002; Tolman et al. 2005), we proposed a solution for this problem. Before applying an NN technique, inputs and outputs are decomposed using EOFs (or another complete basis). Then NN is applied to relate the coefficients of these decompositions. This approach allowed us to reduce the size of input and output vectors (and the size of the NN emulation) by an order of magnitude.

The success of the approach introduced in this paper for approximating the longwave radiation parameterization opens the opportunity for a complete open-minded reexamination of computations for all model physics components. The next logical steps would be

NCAR-CAM 10 YEAR U-WIND

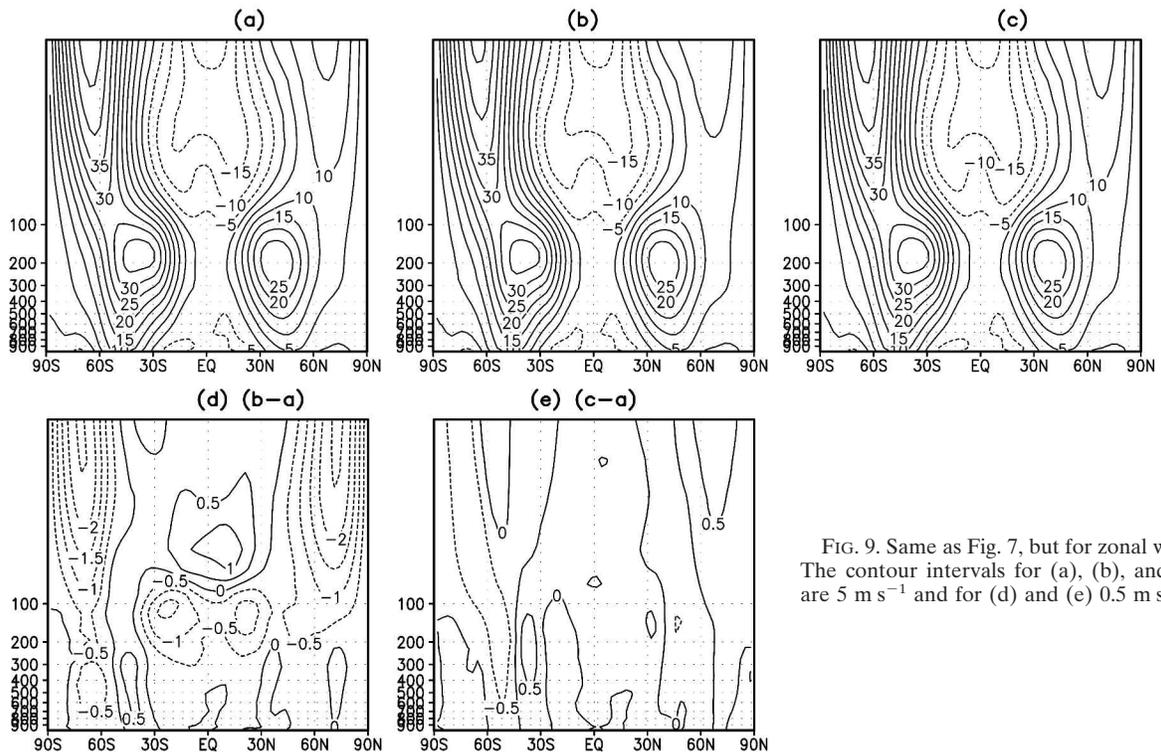


FIG. 9. Same as Fig. 7, but for zonal wind. The contour intervals for (a), (b), and (c) are  $5 \text{ m s}^{-1}$  and for (d) and (e)  $0.5 \text{ m s}^{-1}$ .

NCAR-CAM 10 YEAR Q

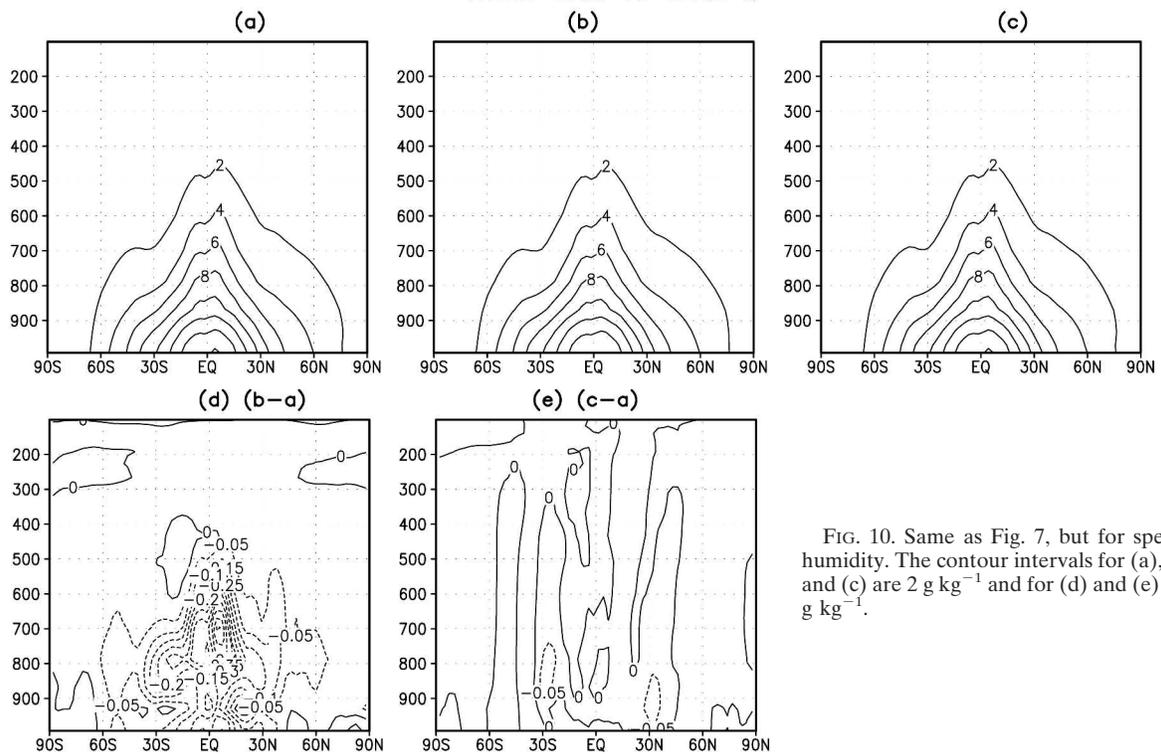


FIG. 10. Same as Fig. 7, but for specific humidity. The contour intervals for (a), (b), and (c) are  $2 \text{ g kg}^{-1}$  and for (d) and (e)  $0.05 \text{ g kg}^{-1}$ .

developing NN emulations for the full atmospheric radiation (including shortwave) block and then cautiously approaching the more nonlinear and, therefore, challenging moisture physics block, including convection, cloudiness, and turbulence. Eventually, NN emulations for all the diabatic forcing components could be introduced. This, in turn, could potentially make an important positive impact on extensive experimentation with the complex models needed to improve climate change and variability assessments, as well as weather prediction. It should be emphasized that the results obtained on the accuracy and efficiency of the NN emulation may facilitate a collaborative effort (with model physics scheme developers) to develop new, more sophisticated parameterizations of model physics (superparameterizations, e.g., cloud physics) that are now computationally prohibitive. This is also true for computational bottlenecks in model dynamics like complicated solvers, iterations, transformations, inversions, etc.

*Acknowledgments.* The authors thank Drs. W. Collins, P. Rasch, J. Tribbia, S. Lord, D. B. Rao, and F. Chevallier for useful discussions and consultations. The authors thank Mr. R. Govindaraju for useful consultations and Mr. A. Belochitsky for programming support. This study is based upon the work supported by the Department of Energy Grant ER63197-0007185.

## REFERENCES

- Attali, J.-G., and G. Pagès, 1997: Approximations of functions by a multilayer perceptron: A new approach. *Neural Networks*, **6**, 1069–1081.
- Chen, T., and H. Chen, 1995a: Approximation capability to functions of several variables, nonlinear functionals and operators by radial basis function neural networks. *Neural Networks*, **6**, 904–910.
- , and —, 1995b: Universal approximation to nonlinear operators by neural networks with arbitrary activation function and its application to dynamical systems. *Neural Networks*, **6**, 911–917.
- Chevallier, F., and J.-F. Mahfouf, 2001: Evaluation of the Jacobians of infrared radiation models for variational data assimilation. *J. Appl. Meteor.*, **40**, 1445–1461.
- , F. Chérut, N. A. Scott, and A. Chédin, 1998: A neural network approach for a fast and accurate computation of longwave radiative budget. *J. Appl. Meteor.*, **37**, 1385–1397.
- , J.-J. Morcrette, F. Chérut, and N. A. Scott, 2000: Use of a neural-network-based longwave radiative transfer scheme in the EMCWF atmospheric model. *Quart. J. Roy. Meteor. Soc.*, **126**, 761–776.
- Collins, W. D., 2001: Parameterization of generalized cloud overlap for radiative calculations in general circulation models. *J. Atmos. Sci.*, **58**, 3224–3242.
- , J. K. Hackney, and D. P. Edwards, 2002: An updated parameterization for infrared emission and absorption by water vapor in the National Center for Atmospheric Research Community Atmosphere Model. *J. Geophys. Res.*, **107**, 4664, doi:10.1029/2001JD001365.
- Cybenko, G., 1989: Approximation by superposition of sigmoidal functions, in mathematics of control. *Signals Syst.*, **2**, 303–314.
- Duffy, P. B., B. Govindasamy, J. P. Iorio, J. Milovich, K. R. Sperber, K. E. Taylor, M. F. Wehner, and S. L. Thompson, 2003: High-resolution simulations of global climate. Part 1: Present climate. *Climate Dyn.*, **21**, 371–390.
- Fox-Rabinovitz, M. S., L. L. Takacs, and R. C. Govindaraju, 2001: A variable resolution stretched grid GCM: Regional climate simulation. *Mon. Wea. Rev.*, **129**, 453–469.
- , —, and —, 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**, 4768, doi:10.1029/2002JD002177.
- Funahashi, K., 1989: On the approximate realization of continuous mappings by neural networks. *Neural Networks*, **2**, 183–192.
- Hornik, K., 1991: Approximation capabilities of multilayer feed-forward network. *Neural Networks*, **4**, 251–257.
- Krasnopolsky, V., 1997: A neural network forward model for direct assimilation of SSM/I brightness temperatures into atmospheric models. *Research Activities in Atmospheric and Oceanic Modeling*, CAS/JSC Working Group on Numerical Experimentation, Rep. 25, WMO Tech. Doc. 792, 1.29–1.30.
- , 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.
- , L. C. Breaker, and W. H. Gemmill, 1995: A neural network as a nonlinear transfer function model for retrieving surface wind speeds from the Special Sensor Microwave Imager. *J. Geophys. Res.*, **100**, 11 033–11 045.
- , D. V. Chalikov, and D. B. Rao, 2000: Application of neural networks for efficient calculation of sea water density or salinity from the UNESCO equation of state. *Proc. Second Conf. on Artificial Intelligence*, Long Beach, CA, Amer. Meteor. Soc., 27–30.
- , —, and H. L. Tolman, 2002: A neural network technique to improve computational efficiency of numerical oceanic models. *Ocean Modell.*, **4**, 363–383.
- , M. S. Fox-Rabinovitz, and D. V. Chalikov, 2004: 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. *Proc. 15th Symp. on Global Change and Climate Variations*, Seattle, Washington, Amer. Meteor. Soc., CD-ROM, P1.20.
- Ripley, B. D., 1997: *Pattern Recognition and Neural Networks*. Cambridge University Press, 403 pp.
- Schoendorf, J., H. Rabitz, and G. Li, 2003: A fast and accurate operational model of ionospheric electron density. *Geophys. Res. Lett.*, **30**, 1492, doi:10.1029/2002GL016649.
- Tolman, H. L., V. M. Krasnopolsky, and D. V. Chalikov, 2005: Neural network approximations for nonlinear interactions in wind wave spectra: Direct mapping for wind seas in deep water. *Ocean Modell.*, **8**, 253–278.