Decadal Climate Simulations Using Accurate and Fast Neural Network Emulation of Full, Longwave and Shortwave, Radiation*

VLADIMIR M. KRASNOPOLSKY

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

MICHAEL S. FOX-RABINOVITZ AND ALEXEI A. BELOCHITSKI

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

(Manuscript received 19 September 2007, in final form 18 February 2008)

ABSTRACT

An approach to calculating model physics using neural network emulations, previously proposed and developed by the authors, has been implemented in this study for both longwave and shortwave radiation parameterizations, or to the *full model radiation*, the most time-consuming component of model physics. The developed highly accurate neural network emulations of the NCAR Community Atmospheric Model (CAM) longwave and shortwave radiation parameterizations are 150 and 20 times as fast as the original/control longwave and shortwave radiation parameterizations, respectively. The full neural network model radiation was used for a decadal climate model simulation with the NCAR CAM. A detailed comparison of parallel decadal climate simulations performed with the original NCAR model radiation parameterizations and with their neural network emulations is presented. Almost identical results have been obtained for the parallel decadal simulations. This opens the opportunity of using efficient neural network emulations for the full model radiation for decadal and longer climate simulations as well as for weather prediction.

1. Introduction

One of the main problems in development and implementation of state-of-the-art numerical climate and weather prediction models is the complexity of physical processes involved. Some of the model physics parameterizations, such as radiation, are time consuming even for most powerful modern supercomputers, and because of that are calculated less frequently than other model physics components and model dynamics. This may negatively affect the accuracy of a model's physics calculation and its temporal consistency, which may, in turn, reduce the accuracy of climate simulations and weather predictions.

The calculation of model physics in the general circulation model (GCM) used in this study, the National Center for Atmospheric Research (NCAR) Community Atmospheric Model (CAM), with the T42 (~3°) horizontal resolution and 26 vertical levels (T42L26), takes about 70% of the total model computations. Evidently, this percentage is model dependent but full model radiation is the most time-consuming component of GCMs (e.g., Morcrette et al. 2007a,b). Such a situation is an important motivation for looking for new alternative numerical algorithms that provide faster calculations of model physics—in particular, of model radiation—while carefully preserving their accuracy. We used the NCAR CAM with the climatological sea surface temperature forcing for the study as a test bed available to the community for experimentation. It is noteworthy that currently we are developing NN emulations for radiation for a coupled ocean–atmosphere model, namely for the National Centers for Environmental Prediction (NCEP) Climate Forecasting System (CFS) model. The preliminary results discussed in Krasnopolsky et al. (2008a) show that neural network

* National Centers for Environmental Prediction Marine Modeling and Analysis Branch Contribution Number 258.

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

DOI: 10.1175/2008MWR2385.1

(NN) emulations for radiation produce positive results for the coupled model, comparable with those presented here for CAM forced by climatological SST.

During the last decade new emerging NN techniques have found a variety of applications in different fields and, more specifically, to accurate and fast modeling of atmospheric radiative processes (Krasnopolsky 1997; Chevallier et al. 1998) and for satellite retrieval procedures (e.g., Krasnopolsky 1997; Krasnopolsky and Schiller 2003). The NN techniques have been successfully applied to development of a new longwave radiation parameterization ("NeuroFlux") for the European Centre for Medium-Range Weather Forecasts (ECMWF) model (Chevallier et al. 1998, 2000). NeuroFlux, which is 8 times as fast as the previous parameterization, consists of a battery of about 40 neural networks. NeuroFlux has been used operationally within the ECMWF four-dimensional variational data assimilation (4DVAR) system since October 2003.

A new approach based on the application of a statistical learning technique—NNs—to *emulation* of model physics parameterizations has been introduced to ocean models (Krasnopolsky et al. 2002) and recently applied by the authors to the NCAR CAM2 (version 2) atmospheric physics parameterizations (Krasnopolsky et al. 2005). A new type of GCM, the hybrid GCM (HGCM), has also been introduced (Krasnopolsky and Fox-Rabinovitz 2006a,b). HGCM is based on a synergetic combination of statistical learning and deterministic model components. The approach uses a particular kind of statistical or machine learning technique, NNs, for accurate and fast *emulation* of model physics components. The term "emulation" means a complete, accurate, and robust functional imitation of the input–output relationship or mapping that exists between input and output vectors of model physics parameterizations (Krasnopolsky 2007a).

The NN emulations are developed for the *existing* (i.e., original) parameterizations of atmospheric physics. This allows us to preserve the integrity and the level of sophistication of the state-of-the-art physical parameterizations of atmospheric processes. Because of the capability of modern statistical learning techniques to provide an unprecedented accuracy for emulations of complex, multidimensional, and multiscale systems like model physics, our NN emulations of model physics parameterizations are practically identical to the original physical parameterizations in terms of the functional input–output relationship. They are usually significantly (from one to five orders of magnitude) faster than the original parameterizations (Krasnopolsky et al. 2002, 2005). In other words, *the underlying idea of the approach is not to develop a new parameterization, but rather to emulate a parameterization already carefully tested and validated by its developers.* This is achieved by using data for NN training simulated by an atmospheric model run with the original parameterization. Using model-simulated data for NN training allows us to achieve an unprecedented accuracy of the approximation because simulated data are free of the problems typical for empirical data (a high level of observational noise, sparse spatial and temporal coverage, poor representation of extreme events, etc.). The accuracy and speedup of NN emulations are always measured against the original parameterization. The developed NN emulation has the same inputs and outputs as the original parameterization, which allows us to use it as a functional substitute for the original parameterization.

The key objective of this study is validating the efficiency of developed NN emulations for the NCAR CAM full-radiation block in terms of a close similarity of decadal climate simulations using the original radiation parameterizations (the control simulation) and their NN emulations (the NN run).

In the study, we apply the NN approach to approximating both the longwave radiation (LWR) and shortwave radiation (SWR) parameterizations in the NCAR CAM [e.g., see the special issue of the *Journal of Climate* (1998, Vol. 11, No. 6)]. Calculation of the LWR and SWR or the full/total model radiation is the most time-consuming part of the atmospheric physics calculations. For example, the NCAR CAM T42 total radiation (LWR and SWR) takes ~70% of the time required for calculation of model physics. Because calculation of model physics takes about 70% of the total model computations, the calculation of T42L26 full radiation takes ~50% of the total model calculation time. A description of NN emulations for LWR in the NCAR CAM and results of their use in 10-yr climate simulations with the model are provided in Krasnopolsky et al. (2005). Also, the NN emulations for LWR and SWR are discussed in Krasnopolsky and Fox-Rabinovitz (2006a,b). In this study, we will concentrate on a discussion of the results of a decadal (50 yr) NCAR CAM climate simulation using NN emulations simultaneously for both LWR and SWR (i.e., for their combined application as the full model radiation). The model simulation using the combined NN emulations is compared with or actually validated against the control model simulation using the original LWR and SWR parameterizations.

In section 2, the NN approach and developed NN emulations for NCAR CAM LWR and SWR are briefly described in terms of their design, accuracy, and computational performance. In section 3, the results of the two parallel decadal model simulations, one using

TABLE 1. Statistics estimating the accuracy of HRs (K day^{-1}) calculations and computational performance for NCAR CAM2 LWR and SWR using NN emulation vs the original parameterization.

	Bias (K day^{-1})	PRMSE (K day^{-1})	σ_{PRMSE} (K day^{-1})	Performance (times as fast)
LWR NN 50	3.0×10^{-4}	0.28	0.20	150
SWR NN 55	-4.0×10^{-3}	0.15	0.12	20

the combined LWR and SWR NN emulations for full model radiation and the other using the original model radiation (the control) are compared in terms of closeness of their spatial and temporal variability characteristics. Section 4 contains the conclusions.

2. NN emulations for the NCAR CAM radiation

a. Background information on NN emulations for LWR and SWR

The NN emulations of model physics are based on the following considerations. Any parameterization of model physics can be formulated as a continuous or almost continuous mapping (input vector versus output vector dependence) and can be symbolically written as

$$\mathbf{Y} = M(\mathbf{X}); \quad \mathbf{X} \in \Re^n, \quad \mathbf{Y} \in \Re^m, \tag{1}$$

where M denotes the mapping, n is the dimensionality of the input space (the number of NN inputs), and m is the dimensionality of the output space (the number of NN outputs). The NNs (multilayer perceptrons, in our case) are a generic tool for approximation of such mappings (Funahashi 1989; Hornik 1991).

The NN is an analytical approximation that uses a family of functions like

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

where x_i and y_q are components of the input and output vectors $\mathbf{X}$ and $\mathbf{Y}$, respectively; a and b are fitting parameters; and $\phi(b_{j0} + \Sigma_{i=1}^{n} b_{ji}x_i)$ is a "neuron." The activation function ϕ is 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. Definitions of NN terminology can be found in many places, for example, in the recent book by Bishop (2006) and in Krasnopolsky (2007a); however, Eq. (2) is sufficient to understand the subject of this paper.

The major goals for developing NN emulation for model physics are to obtain an extremely high accuracy for NN emulation with practically zero biases or systematic errors. This is a necessary condition for obtaining nonaccumulating errors during long-term climate simulations with developed NN emulations. The choice of an optimal version of NN emulation is based on accuracy, not on a speedup of computation. All the NN emulations obtained provide a very significant speedup anyway. The most efficient and convenient way of developing NN emulations for model physics components is to develop a single NN for a model physics parameterization. Such an approach has been introduced and discussed in Krasnopolsky et al. (2005) and Krasnopolsky and Fox-Rabinovitz (2006b).

The LWR and SWR parameterizations together compose the full model radiation. The LWR and SWR parameterizations or the *full model radiation* have been emulated using NNs in the NCAR CAM2. The function of the radiation (LWR and SWR) parameterizations in atmospheric GCMs is to calculate radiation fluxes and heating rates produced by the LWR and SWR atmospheric processes. The complete description of NCAR CAM atmospheric LWR and SWR parameterizations is presented by Collins (2001) and Collins et al. (2002). A very general and schematic outline of these parameterizations, illustrating the complexity that makes them a computational bottleneck in the NCAR CAM physics, is given in Krasnopolsky and Fox-Rabinovitz (2006b).

The input vectors for the NCAR CAM2 LWR parameterization include 10 profiles [i.e., atmospheric temperature, humidity, ozone, CO_2, N_2O, CH_4, two CFC mixing ratios (the annual mean atmospheric mole fractions for the halocarbons), pressure, cloud emissivity, and cloud cover] and one relevant surface characteristic (i.e., the upward LWR flux at the surface). The LWR 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 LWR or OLR).

The NN emulation of the LWR parameterization has exactly the same inputs [total 220 inputs; $n = 220$ in Eq. (1)] and the same outputs [total 33 outputs; $m = 33$ in Eq. (1)] as the original LWR parameterization. Krasnopolsky et al. (2005) have developed several NNs, all of which have one hidden layer with from 20 to 300 neurons [$k = 20$–300 in Eq. (2)]. Varying the number of

TABLE 2. Time (40-yr) and global means and their differences for model prognostic and diagnostic fields for the NCAR CAM2 control climate simulation with the original LWR and SWR, and for the parallel simulation with NN emulations for the full radiation using NN 50 (LWR) and NN 55 (SWR). Sea level pressure (SLP), 2-m temperature (T2M), 200-hPa zonal wind (U-200), total precipitation rate (TPR), total cloud amount (TCA), high-level cloud amount (HLCA), low-level cloud amount (LLCA), midlevel cloud amount (MLCA), total gridbox cloud liquid water path (TGCLWP), total gridbox cloud ice water path (TGCIWP), top-of-model net longwave flux (TOMNLW), top-of-model net shortwave flux (TOMNSW), and top-of-model longwave cloud forcing (TOMLWC).

Field	Control	NN full radiation	Mean diff	RMS diff	Min diff	Max diff
SLP (hPa)	1011.48	1011.50	0.02	0.52	−2.04	1.57
T2M (K)	287.37	287.27	−0.1	0.26	−1.64	0.78
U-200 (m s^{-1})	16.21	16.29	0.08	0.86	−2.31	3.95
TPR (mm day^{-1})	2.86	2.89	0.03	0.2	−1.84	1.19
TCA (%)	60.71	61.12	0.41	1.42	−7.50	5.76
HLCA (%)	43.05	43.29	0.24	1.63	−7.52	8.01
LLCA (%)	31.67	31.93	0.26	1.06	−5.20	4.78
MLCA (%)	19.11	19.14	0.03	0.81	−4.86	4.39
TGCLWP (g m^{-2})	60.23	60.59	0.36	3.02	−19.43	14.95
TGCIWP (g m^{-2})	8.82	8.83	0.01	0.39	−1.69	1.45
TOMNLW (W m^{-2})	234.48	234.54	0.06	2.32	−8.37	11.56
TOMNSW (W m^{-2})	234.91	234.17	−0.74	2.17	−12.44	18.94
TOMLWC (W m^{-2})	29.33	29.07	−0.26	2.45	−15.59	7.64

hidden neurons allows us to demonstrate the dependence of the accuracy of approximation on this parameter as well as its convergence, and as a result, provide a sufficient accuracy of approximation for the climate model.

The input vectors for the SWR parameterization include 21 vertical profiles (i.e., specific humidity, ozone concentration, pressure, cloud cover, layer liquid water path, liquid effective drop size, ice effective drop size, fractional ice content within cloud, aerosol mass mixing ratios, etc.), the solar zenith angle, and the surface albedo for four different bands. The SWR parameterization output vectors consist of a vertical profile of HRs and several radiation fluxes. The NN emulations of the SWR parameterization have 173 inputs and 33 outputs. We have developed several NNs, all of which have one hidden layer with from 50 to 200 neurons [k = 50–200 in Eq. (2)].

It is noteworthy that the number of NN inputs is less than the number of input profiles multiplied by the number of the vertical layers plus the number of relevant single level characteristics. Many input variables (e.g., all gases) have zero or constant values for upper vertical layers. These constant inputs were not used for NN training to improve the accuracy of the approximation. Constant inputs (zero or nonzero) do not contribute to the functional input–output relationship and should not be used for the development of NN emula-

TABLE 3. Time (40-yr) and global means for model prognostic and diagnostic fields for the NCAR CAM2 control climate simulation with the original LWR and SWR, and the simulation with NN emulations for the full radiation using NN 50 (LWR) and NN 55 (SWR), and their mean (bias), rms, min, and max errors vs observations* or reanalysis. Sea level pressure (SLP), 2-m temperature (T2M), 200-hPa zonal wind (U-200), total precipitation rate (TPR), surface LW downwelling flux (SLWDF), surface SW downwelling flux (SSWDF), surface net LW flux (SNLWF), surface net SW flux (SNSWF), all fluxes are in watts per meter squared. In each cell two errors, separated by a slash, are presented: NN full radiation vs observation and control vs observations.

Variable	Bias	RMS error	Min error	Max error
SLP (hPa)	−0.12/−0.14	3.40/3.55	−29.44/−31.58	8.12/8.02
T2M (K)	−0.32/−0.28	3.12/3.11	−26.30/−25.34	24.05/24.33
U-200 (m s^{-1})	1.01/0.93	4.56/4.73	−11.71/−11.42	15.03/15.36
TPR (mm day^{-1})	0.16/0.17	1.20/1.15	−5.41/−5.21	6.87/9.02
SLWDF (W m^{-2})	−4.52/−4.66	13.35/12.68	−67.04/−66.72	45.23/43.32
SSWDF (W m^{-2})	2.04/2.08	18.23/19.03	−53.14/−61.55	72.35/68.77
SNLWF (W m^{-2})	14.84/14.70	20.61/20.62	−16.89/−15.62	85.23/87.78
SNSWF (W m^{-2})	−0.84/1.80	16.54/16.3	−53.21/−55.7	70.49/71.43

* LW and SW observational data: International Satellite Cloud Climatology Project (ISCCP) (July 1983–December 2000). Other observational data: sea level pressure (SLP) and U-200: NCEP–National Center for Atmospheric Research (NCAR) reanalysis (1979–98; Kalnay et al. 1996; Kistler et al. 2001); T2M: Legates and Willmott (1920–80; Legates and Willmott 1990); TPR: Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; 1979–98; Xie and Arkin 1997).

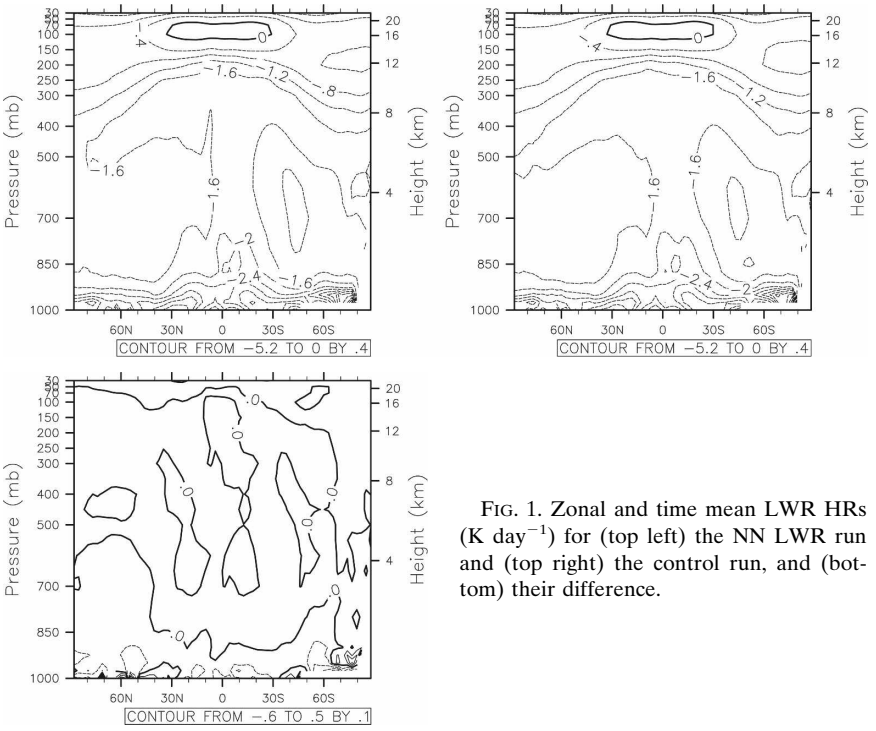

FIG. 1. Zonal and time mean LWR HRs (K day^{-1}) for (top left) the NN LWR run and (top right) the control run, and (bottom) their difference.

tions. Moreover, if they were used, they would introduce an additional noise (an approximation error).

The NCAR CAM2 (T42L26) was run for 2 yr to generate representative datasets. The representative dataset adequately samples the atmospheric state variability. 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

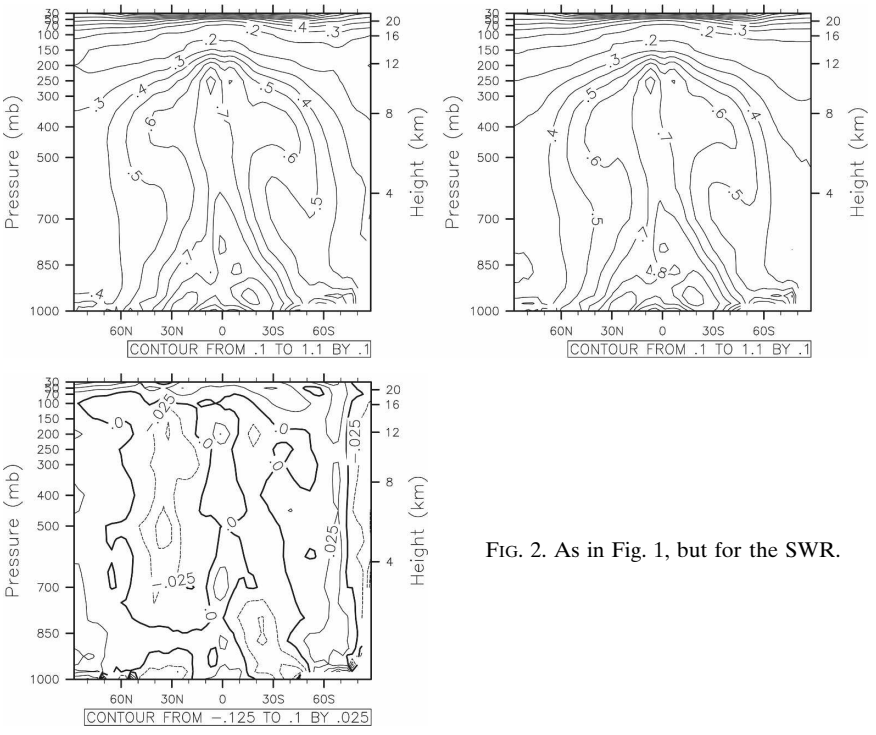

FIG. 2. As in Fig. 1, but for the SWR.

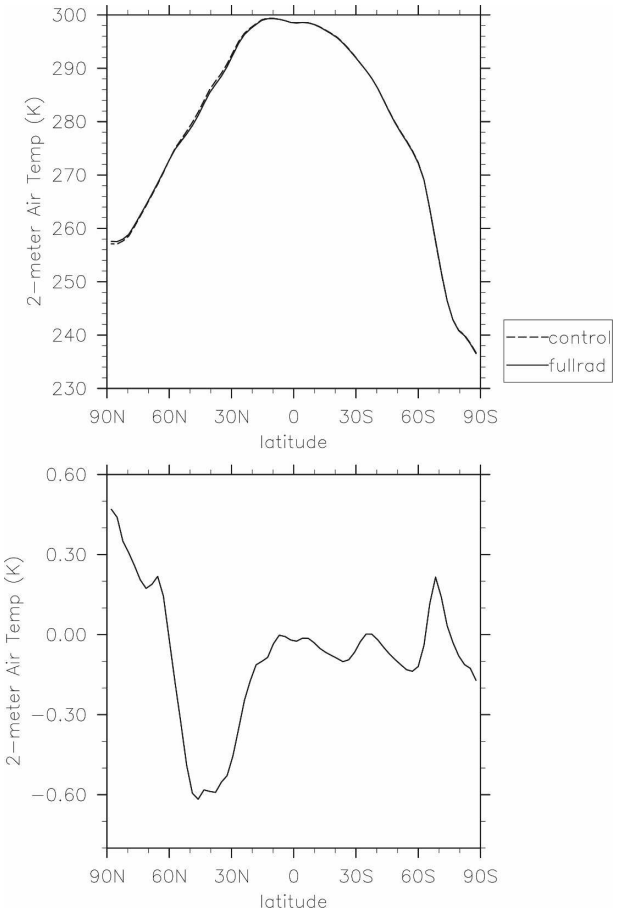

FIG. 3. Zonal and time mean 2-m temperature (K) for (top) the full-radiation NN and control runs and (bottom) their difference.

the second for tests (i.e., control of overfitting, control of NN architecture, etc.). The second year of simulation was used to create a validation dataset completely independent of both the training and test datasets. The third part or the validation set was used for validations only. All approximation statistics presented in this section are calculated using this independent validation dataset. The accuracy of the NN run (i.e., biases and rmse) is calculated against the control run.

b. Bulk approximation error statistics

To ensure a high quality of representation of the LWR and SWR radiation processes, the accuracy of the NN emulations has been carefully investigated. Our NN emulations have been validated against the original NCAR CAM LWR and SWR parameterizations. To calculate the error statistics presented in Table 1 and the following figures of this section, the original parameterizations and their NN emulations have been calculated using a validation dataset. Two sets of the corresponding HR profiles have been generated for both LWR and SWR. Total and level bias (or mean error), total and level RMSE, profile RMSE or PRMSE, and σ_{PRMSE} have been calculated (Krasnopolsky et al. 2005; Krasnopolsky and Fox-Rabinovitz 2006a). Some of these statistics presented in Table 1 have been calculated as follows. The outputs of the original parameterization and the NN emulations can be represented as $Y(i, j)$ and $Y_{\mathrm{NN}}(i, j)$, respectively, where $i = (\mathrm{lat}, \mathrm{lon})$, $i = 1, \ldots, N$ is the horizontal location of a vertical profile, N is the number of horizontal grid points, and $j = 1, \ldots, L$ is the vertical index, where L is the number of vertical levels.

The mean difference, B (i.e., bias or a systematic error of approximation), between the original parameterization and its NN emulation, is calculated as follows:

$$B = \frac{1}{NL}\sum_{i=1}^{N}\sum_{j=1}^{L}[Y(i,j) - Y_{\mathrm{NN}}(i,j)]. \quad (3)$$

The root-mean-square error has been calculated for each ith profile:

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

This error can be used to calculate the mean profile root-mean-square error, PRMSE, and its standard deviation, σ_{PRMSE}:

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

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

Table 1 shows bulk validation statistics for the accuracy of approximation and computational performance for the best (in terms of accuracy and performance) developed NN emulations: NN 50 [$k = 50$ hidden neurons in Eq. (2)] for the LWR emulation and NN 55 [$k = 55$ hidden neurons in Eq. (2)] for the SWR emulation.

The error profiles for LWR and SWR are shown in Fig. 4 of Krasnopolsky (2007a). The NN emulations developed for LWR and SWR are highly accurate. They have practically zero bias and a quite small PRMSE. Zonal mean differences between the NN emulation and the original parameterization for radiative fluxes at the top of the atmosphere and at the surface have also been produced. The differences appear to be uniformly small for all latitudes, mostly within ± 0.5 W m^{-2} and do not exceed ± 1 W m^{-2}.

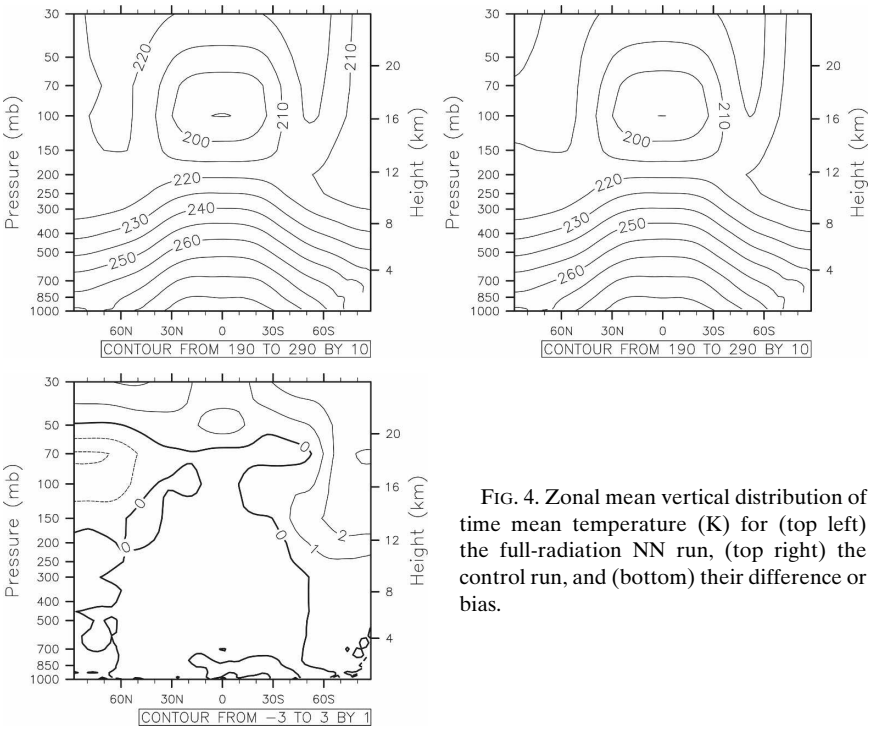

FIG. 4. Zonal mean vertical distribution of time mean temperature (K) for (top left) the full-radiation NN run, (top right) the control run, and (bottom) their difference or bias.

The NN emulations using 50 neurons in the hidden layer provide, if run *separately* at every model physics time step (1 h), a speedup of roughly 150 times for LWR and 20 times for SWR as compared with the original LWR and SWR, respectively. It is noteworthy that the main reason for the smaller performance gain for NN SWR versus NN LWR is that the original CAM SWR parameterization is simpler and about 10 times faster than the original CAM LWR.

Using NN emulations *simultaneously* for LWR and SWR or for the full model radiation results in an overall significant, 13-fold acceleration of calculations for the entire/full model radiation block. It is worth clarifying, for a better understanding of the overall speedup, that for the usual control run the original LWR (including time-consuming optical properties calculations) is calculated less frequently, only every 12 h or twice a day, and only computationally inexpensive heating rates and radiative fluxes are calculated every hour. Notice that all other inputs, including cloud cover, which is represented by a vertical profile of cloud fraction, are updated hourly. For the model run using NN emulations, LWR (including both optical properties and heating rates and radiative fluxes) is calculated more frequently, every hour, consistent with SWR and other model physics calculations. We also performed an additional costly control run with the original LWR calculated every hour, as it is done in the LWR NN run, for a limited period (10 yr). The results of the two control runs appeared to be very close. The difference between them is significantly less than the difference between each of them and the LWR NN run. Because of that we decided to validate the 40-yr full-radiation NN run against the usual control run.

3. Validation of parallel decadal model simulations

The comparisons between diagnostic and prognostic fields for the relatively short parallel model runs, one using the original LWR or SWR (the control run) and the other one using NN emulations for the LWR or SWR parameterizations, are presented in Krasnopolsky et al. (2005) and in Krasnopolsky and Fox-Rabinovitz (2006a,b). They show that the parallel runs produce the results that are close to each other. Therefore, both components of radiation, LWR and SWR, can be successfully emulated using the NN approach.

These results opened the opportunity to use both NN emulations, for LWR and SWR (or full model radiation) simultaneously, in a multidecadal simulation using NCAR CAM (T42L26), the results of which are discussed below. The results of multidecadal climate simulations performed with NN emulations for both LWR and SWR (i.e., for the full model radiation) have been validated against the parallel control NCAR CAM simulation using the original LWR and SWR. Below we

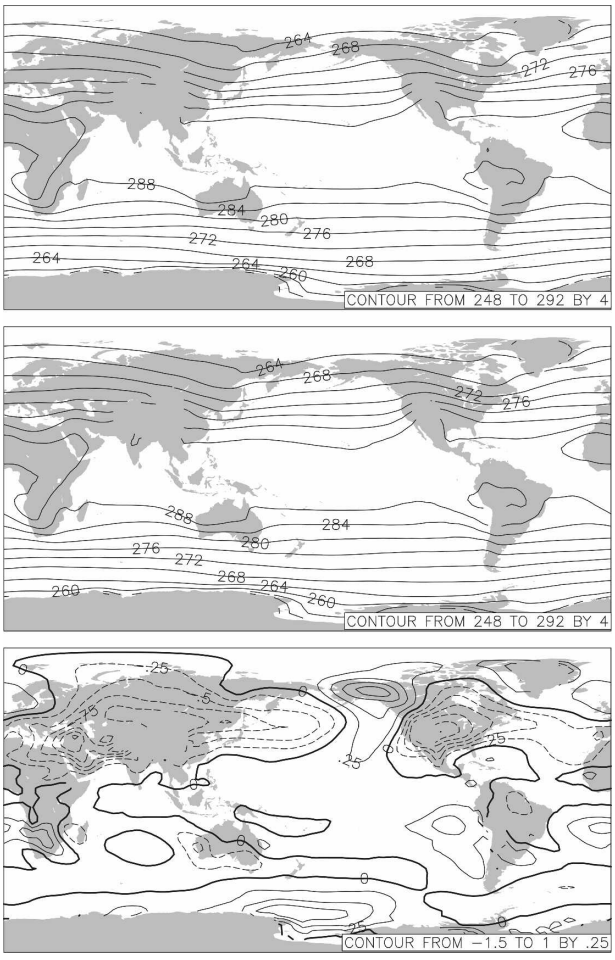

FIG. 5. Time mean temperature at 850 hPa (K) for (top) the full-radiation NN run and (middle) the control run, and (bottom) their difference.

estimate closeness of the results for these parallel 50-yr climate simulations. Note that the first 10 yr of simulations are not included in the validation to avoid the impact of spinup effects, so that years 11–50 are used for the validation. The spinup is done for the original NCAR CAM; it is not related to the use of NN emulations. We will analyze below the differences between the parallel runs in terms of time and spatial (global) means as well as temporal characteristics.

Table 2 presents comparisons between the parallel control and NN emulation runs in terms of the time (40 yr) and global mean characteristics and the differences between the results of the parallel runs. Basically, the differences, in terms of their mean, rms, minimum, and maximum characteristics, between the parallel runs, are small. More specifically, there are small mean differences (bias), 0.02 hPa and −0.1 K, in sea level pressure and 2-m temperature, respectively, between the NN and control runs. For these fields, rmse, minimum, and maximum differences are also small. Other time and global mean differences presented in Table 2, including such sensitive fields as total precipitation, total cloud amount, cloud amounts for high, low, and midlevel clouds, total gridbox cloud liquid and ice water paths, top-of-model net longwave flux, top-of-model net shortwave flux, and cloud forcing, also show a close similarity, in terms of all presented difference characteristics, between the parallel simulations. These differences are within typical observational and reanalysis errors/uncertainties. Note that minimum and maximum differences in Table 2 (and in Table 3) are not averaged in space and time but rather are obtained from monthly mean gridpoint values for the entire 40-yr simulations.

To further analyze the closeness of the parallel control and NN emulation runs, Table 3 presents their mean (bias), rms, minimum, and maximum errors (i.e., validation results for the runs) versus observations or reanalysis. Such a validation allows us to verify to what degree the model simulation with the NN emulations for full radiation can deviate from the control simulation in terms of model simulation versus observation errors [i.e., whether the deviation between the runs is within the general uncertainty (or overall error levels) for typical decadal model simulations such as those of the Atmospheric Model Intercomparison Project (AMIP; Gates et al. 1999)]. Both the prognostic (i.e., sea level pressure and zonal wind at 200 hPa) and diagnostic (i.e., 2-m temperature, precipitation, long- and shortwave radiation characteristics) fields presented in Table 3 show similar errors for the parallel NN and control runs including sensitive minimum and maximum errors. The differences between the errors for the parallel runs are smaller than the errors themselves for each run. Also, the differences are smaller than those of typical single or multimodel ensemble integrations (e.g., Gates et al. 1999). In other words, the differences between the errors for the parallel runs are well within the general uncertainty of climate model simulations.

Let us discuss the differences between the parallel simulations in terms of spatial and temporal characteristics. Zonal and time mean heating (or cooling) rates for LWR and SWR are presented in Figs. 1 and 2, respectively. The HR patterns (the top panels) are very similar and their differences (the bottom panels) are small. It confirms that the NN emulations for LWR and SWR are very close to their original parameterizations throughout the model simulations. It is noteworthy that the HR differences in SWR and especially in LWR are a bit larger near the surface because HRs are larger there (Figs. 1 and 2). For the zonal means it is not easy to distinguish between the ocean and land. However, the differences seem to be larger over the mountainous

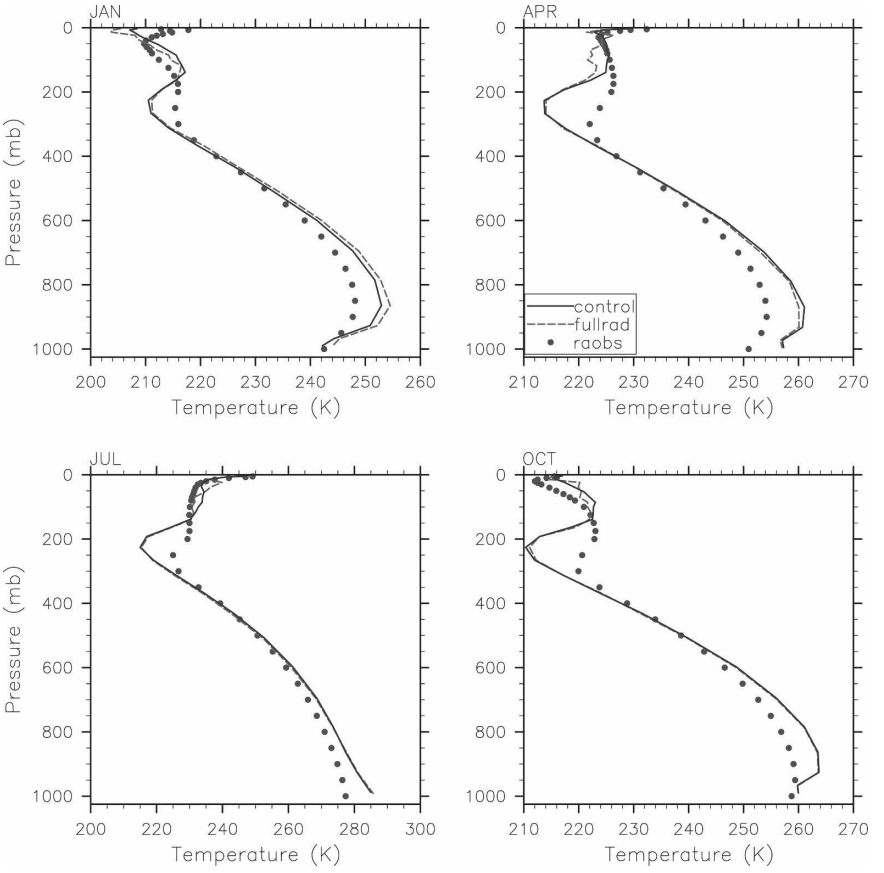

FIG. 6. Vertical profile of time mean temperature (K) at the Resolute, Canada, station for the full-radiation NN run (the dashed line), the control run (the solid line), and observations (the dotted line): (top left) January, (top right) April, (bottom left) July, and (bottom right) October.

Antarctica region (60°–90°S) as well as over the Northern Hemisphere midlatitudes where the major mountains are located (such as those in Europe, Asia, and North America).

Figure 3 shows a very close similarity in zonal and time mean 2-m temperature for the parallel simulations (the top panel) where their differences are within the −0.6- to 0.5-K range (the bottom panel).

The zonal and time mean vertical distributions of temperature for the parallel runs (Fig. 4) are close to each other and their difference or mean bias is practically zero, with minimum and maximal biases within approximately 2–2.5 K by magnitude. This larger zonal bias occurs in the stratosphere mostly over the southern polar domain. However, it is comparable with typical observational and/or reanalysis errors/uncertainties (just as a reference) and also comparable with the differences between the NCEP and ECMWF reanalyses.

Close similarities have also been obtained for the results of parallel runs in terms of time mean spatial fields such as 850-hPa temperature presented in Fig. 5. The horizontal fields presented in the top and middle panels are close to each other. For the difference field (the bottom panel), bias is very small (−0.06 K), RMSE is small (0.34 K), and minimum and maximum values (∼−1.6 and ∼0.9 K) are well within observational or reanalysis errors/uncertainties.

In addition to global distributions such as shown in Fig. 5 it is important to assess the differences between the parallel simulations at a local (station) level, an example of which is presented in Fig. 6. The vertical distributions of time mean temperature are very close for both runs at the local level as well.

Now we compare the results of the parallel simulations in terms of temporal characteristics. Figure 7 shows the winter–summer differences for time mean temperature at 850 hPa. Their patterns are very similar and the minimum and maximum values are very close. The global mean time series for monthly mean tem-

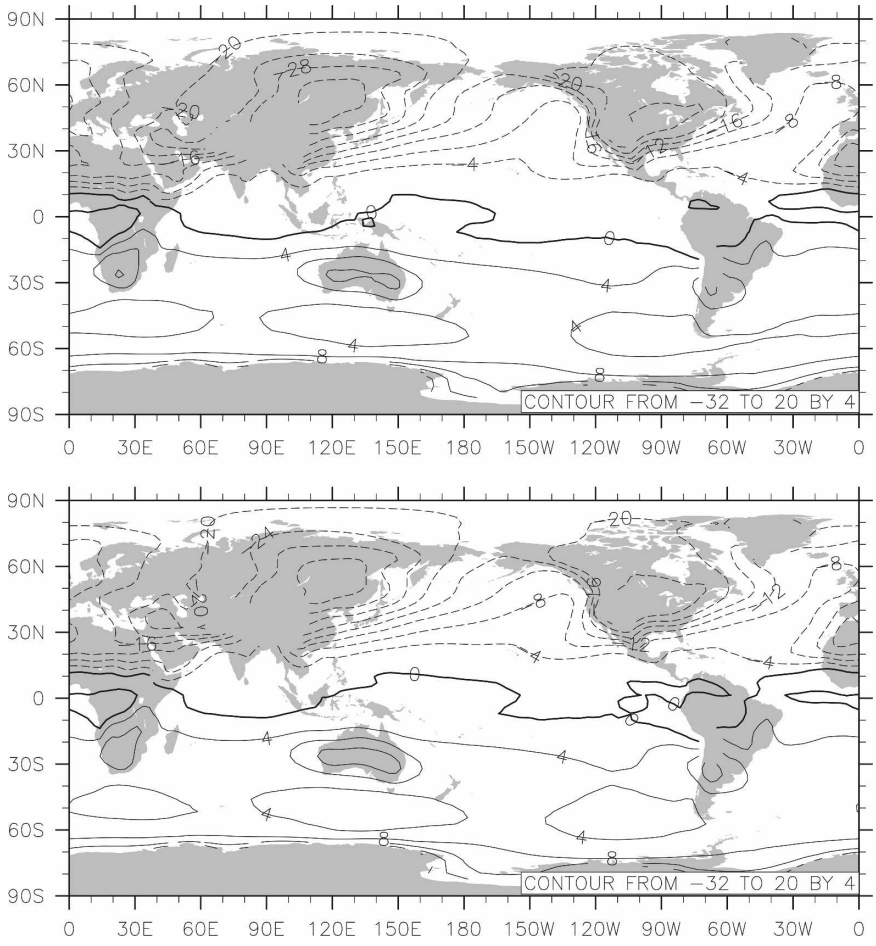

FIG. 7. Winter–summer difference for time mean temperature at 850 hPa (K) for (top) the full-radiation NN run and (bottom) the control run.

perature at 850 hPa presented in Fig. 8 are very similar throughout the entire decadal simulations for the parallel runs, with maximum differences not exceeding 0.3–0.5 K. These maximum differences occur in January and July and are well below the observation and reanalysis errors/uncertainties.

The annual cycle for global and time (40 yr) mean temperature at 850 hPa is presented in Fig. 9. It shows very small differences between the runs, with the maximum within 0.2 K for January. The precipitation annual cycles shown in Fig. 10 are very close for both runs (the top panels) and their differences or bias (the bottom panel) is quite small. It is noteworthy that there still is a coherent signal in Fig. 10; namely, the ITCZ precipitation maximum is slightly stronger in the NN run, and the subsidence regions north and south of the ITCZ are slightly drier. These minor features may be related to the prescribed SSTs (which strongly regulate the tropical convection) used for model simulations. Close similarity has also been obtained for other model prognostic and diagnostic fields in term of their spatial and temporal characteristics.

The results obtained confirm the profound similarity in parallel climate simulations, which justifies the possibility of using efficient neural network emulations of full model radiation for decadal and longer climate simulations as well as for weather prediction models. The methodology developed can be applied to other LWR and SWR schemes used in a variety of models, process studies, and other applications.

4. Conclusions

In this study, we presented an approach based on a synergetic combination of deterministic modeling based on physical (i.e., first principle) equations and statistical learning (i.e., NN emulation) components within an atmospheric model. The statistical learning approach was used to develop highly accurate and fast NN emulations for model physics components. Here we

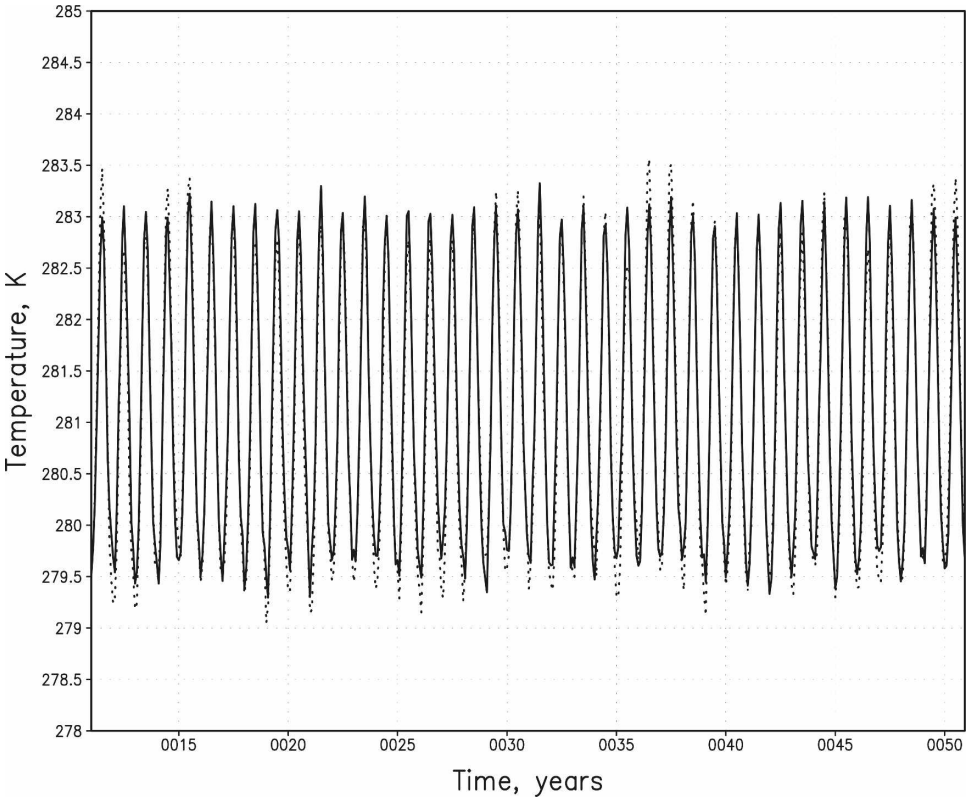

FIG. 8. Global mean time series for monthly mean temperature at 850 hPa (K) for the full-radiation NN run (the dotted line) and the control run (the solid line).

presented an NN emulation of the full atmospheric radiation (i.e., for long- and shortwave radiation parameterizations used in numerical climate and weather prediction models).

This study has shown the practical possibility of using highly efficient NN emulations for the full model radiation block for decadal (50 yr) climate simulations. A high accuracy and increased speed of NN emulations for the NCAR CAM full radiation (LWR and SWR) has been achieved. The systematic errors introduced by NN emulations of full model radiation are very small and do not accumulate during the decadal model simulation. The random errors of NN emulations are also small as is shown in section 2. Almost identical results have been obtained for the parallel 50-yr climate simulations as shown in section 3. These results show the potential of developing efficient NN emulations for model physics components and the entire model physics.

The NN emulation approach presented here is very robust. It was applied to both LWR and SWR parameterizations in different models with different dynamical cores and with different resolutions. For example, in addition to the NCAR CAM applications presented here, this approach was applied to the National Aero-

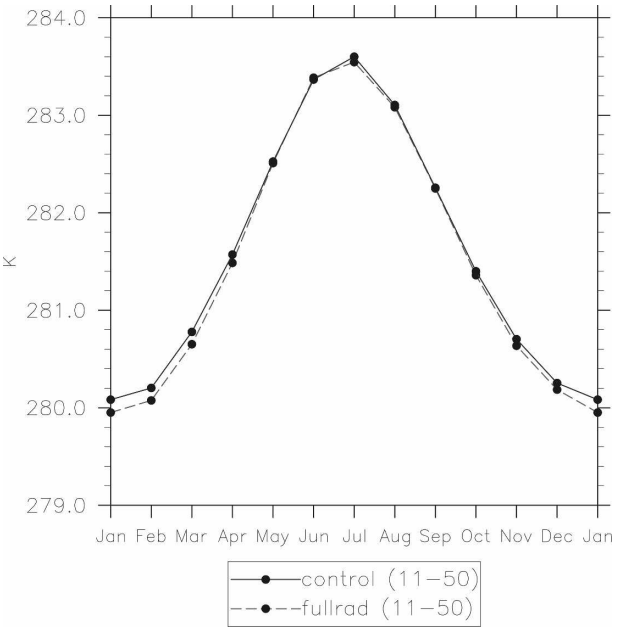

FIG. 9. Annual cycle for global and time (40 yr) mean temperature at 850 hPa (K) for the full-radiation NN run (the dashed line) and the control run (the solid line).

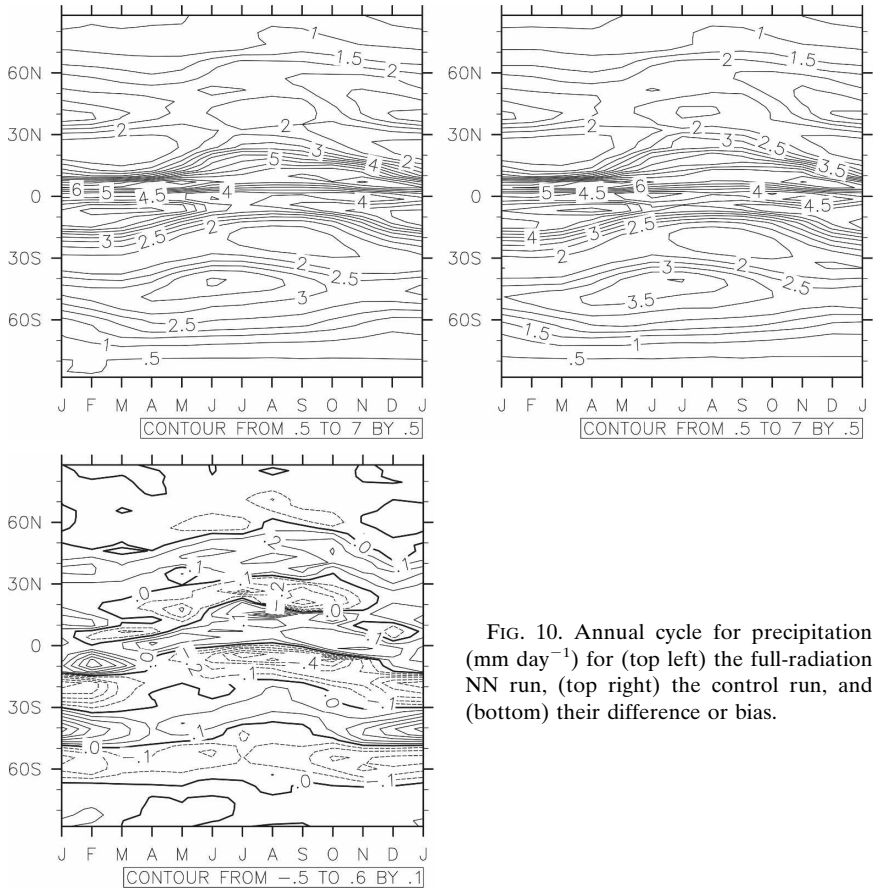

FIG. 10. Annual cycle for precipitation (mm day^{-1}) for (top left) the full-radiation NN run, (top right) the control run, and (bottom) their difference or bias.

nautics and Space Administration's (NASA's) Seasonal-to-Interannual Prediction Project (NSIPP) model and to the NCEP CFS model, using NN emulations of the LWR parameterizations (Krasnopolsky and Fox-Rabinovitz 2006a; Krasnopolsky et al. 2008a). In all these cases, the systematic errors introduced by NN emulations are negligible and the random errors are very small, similar to errors presented in this paper. The computational speedups achieved in all these cases are also similar (about two orders of magnitude).

Applying the NN emulation approach, which allows such a significant speedup with preservation of the accuracy and functional integrity of the model physics, may present some challenges that can be resolved using the tremendous flexibility of statistical learning techniques and of the NN technique in particular. Because NN emulations are statistical approximations, there exists a small probability of larger approximation errors or outliers. The major reason for larger errors is high dimensionality n of the input space of the mapping in (1), which reaches several hundreds for NCAR CAM and approaches 1000 for models with higher vertical resolution. It is impossible to sample uniformly a domain in such a high-dimensional space. Far corners of the domain may remain underrepresented in the training set. During the NN run, if input vectors, which belong to these underrepresented far corners, are encountered, they may cause larger errors in the NN outputs. These larger errors can be successfully controlled using a compound parameterization technique with a quality control procedure for removing larger errors (Krasnopolsky 2007a; Krasnopolsky et al. 2008b) or using the NN ensemble approach with NN emulations (Fox-Rabinovitz et al. 2006). The compound parameterization technique can also be used as a method of enriching the training dataset by inclusion of underrepresented atmospheric states (Krasnopolsky et al. 2008b). The NN emulation technique is sensitive to the resolution of the model used, especially to vertical resolution, which determines the NN emulation architecture (i.e., the number of inputs and outputs). Every time the vertical resolution of the model is changed, the NN emulation needs to be retrained. It is noteworthy that NN retraining can be done routinely and takes limited time and effort once the practical framework for a specific model is developed.

In some applications of the developed NN emulation (e.g., in a data assimilation system or for an error and sensitivity analysis) not only the NN emulation but also its first derivatives (i.e., NN Jacobian) are used. High accuracy of a NN emulation does not automatically guarantee the accuracy of the NN Jacobian. An approach that allows us to calculate accurately the NN Jacobian was developed by Krasnopolsky (2007b).

As mentioned above, the NN emulations described in this study have been developed only for the *existing* model parameterizations. Extension of the NN approach to developing *new parameterizations* goes beyond the scope of this study and could be done as a collaborative effort with parameterization developers interested in implementation of more sophisticated and realistic model physics, which are now computationally prohibitive. Also, it is noteworthy that the NN emulation technique can be applied to accelerate calculations of the model chemistry.

Acknowledgments. The authors thank Drs. W. Collins, P. Rasch, J. Tribbia (NCAR), F. Baer, and D. Chalikov (University of Maryland) for useful discussions and consultations. This study is based upon the work supported by the NOAA/CDEP/CTB Grant NA06OAR4310047.

REFERENCES

Bishop, Ch. M., 2006: *Pattern Recognition and Machine Learning*. Springer, 738 pp.

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. *J. Appl. Meteor.,* **37,** 1385–1397.

——, 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. *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.

Fox-Rabinovitz, M. S., V. Krasnopolsky, and A. Belochitski, 2006: Ensemble of neural network emulations for climate model physics: The impact on climate simulations. *Proc. 2006 Int. Joint Conf. on Neural Networks,* Vancouver, BC, Canada, IEEE, 9321–9326.

Funahashi, K., 1989: On the approximate realization of continuous mappings by neural networks. *Neural Networks,* **2,** 183–192.

Gates, W. L., and Coauthors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.,* **80,** 29–55.

Hornik, K., 1991: Approximation capabilities of multilayer feedforward network. *Neural Networks,* **4,** 251–257.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Krasnopolsky, V. M., 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/TD-792, 1.29–1.30.

——, 2007a: Neural network emulations for complex multidimensional geophysical mappings: Applications of neural network techniques to atmospheric and oceanic satellite retrievals and numerical modeling. *Rev. Geophys.,* **45,** RG3009, doi:10.1029/2006RG000200.

——, 2007b: Reducing uncertainties in neural network Jacobians and improving accuracy of neural network emulations with NN ensemble approaches. *Neural Networks,* **20,** 454–461.

——, 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.

——, and M. S. Fox-Rabinovitz, 2006a: A new synergetic paradigm in environmental numerical modeling: Hybrid models combining deterministic and machine learning components. *Ecol. Modell.,* **191,** 5–18.

——, and ——, 2006b: Complex hybrid models combining deterministic and machine learning components for numerical climate modeling and weather prediction. *Neural Networks,* **19,** 122–134.

——, D. V. Chalikov, 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, 2005: Fast and accurate neural network approximation of longwave radiation in a climate model. *Mon. Wea. Rev.,* **133,** 1370–1383.

——, ——, Y. T. Hou, S. J. Lord, and A. A. Belochitski, 2008a: accurate and fast neural network emulations of longwave radiation for the NCEP Climate Forecast System Model. Preprints, *20th Conf. on Climate Variability and Change,* New Orleans, LA, Amer. Meteor. Soc., P3.10.

——, ——, H. L. Tolman, and A. A. Belochitski, 2008b: Neural network approach for robust and fast calculation of physical processes in numerical environmental models: Compound parameterization with a quality control of larger errors. *Neural Networks,* **21,** 535–543, doi:10.1016/j.neunet.2007.12.019.

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.,* **41,** 11–21.

Morcrette, J.-J., G. Mozdzynski, and M. Leutbecher, 2007a: A reduced radiation grid for the ECMWF Integrated Forecasting System. ECMWF Tech. Memo. 538, 22 pp.

——, and Coauthors, 2007b: Recent advances in radiation transfer parametrizations. ECMWF Tech. Memo. 539, 52 pp.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558.