Accurate and Fast Neural Network Emulations of Model Radiation for the NCEP Coupled Climate Forecast System: Climate Simulations and Seasonal Predictions*

V. M. KRASNOPOLSKY

NOAA/NCEP, Science Applications International Corporation, Camp Springs, and Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

M. S. FOX-RABINOVITZ

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

Y. T. HOU AND S. J. LORD

NOAA/NCEP, Camp Springs, Maryland

A. A. BELOCHITSKI

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

(Manuscript received 14 July 2009, in final form 2 December 2009)

ABSTRACT

The approach to accurate and fast-calculating model physics using neural network emulations was previously developed by the authors for both longwave and shortwave radiation parameterizations or the full model radiation, which is the most time-consuming component of model physics. It was successfully tested for a moderate-resolution uncoupled NCAR Community Atmospheric Model (CAM) that is driven by climatological SST for a decadal climate simulation mode. In this study, the approach has been further developed and implemented into the NCEP coupled Climate Forecast System (CFS) with significantly higher resolution and time-dependent CO_2. The higher complexity of NCEP CFS required further adjustments to the neural network emulation methodology. Validation of the approach for the NCEP CFS has been performed through a decadal climate simulation and seasonal predictions. The developed highly-accurate neural network emulations of longwave and shortwave radiation parameterizations are, on average, 16 and 60 times faster than the original/control longwave and shortwave radiation parameterizations, respectively. The authors present a detailed comparison of parallel decadal climate simulations and seasonal predictions performed with the original NCEP model radiation parameterizations and with their neural network emulations. The differences between the parallel runs are overall within or less than the observation errors and uncertainties of reanalysis. Moreover, the differences (both in terms of bias and RMSE) are of a similar magnitude as the model's internal variability. These results justify the practical use of efficient neural network emulations of full model radiation for climate simulations and seasonal predictions.

1. Introduction

Calculation of model physics in a GCM usually takes a significant part 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. 2007, 2008; Manners et al. 2009). In both climate modeling and NWP, the calculation of radiative transfer is necessarily a trade-off between accuracy and computational efficiency. There exist very accurate methods such as line-by-line procedures that could be employed ideally to calculate radiative fluxes for every grid point at every time step. If the radiation transfer were to be computed for every grid point and at all time steps, it would generally require

* National Centers for Environmental Prediction Marine Modeling and Analysis Branch Contribution Number 277.

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

DOI: 10.1175/2009MWR3149.1

as much CPU time or more than the rest of the model components, that is, model dynamics and other physical parameterizations (Morcrette et al. 2008). Therefore, a number of simplifications are usually made to reduce this cost to manageable levels. For example, in the majority of modern radiative schemes, the correlated-*k* method (Lacis and Oinas 1991) is typically used to reduce the integration over wavelength by effectively binning wavelengths with similar absorption coefficients (*k* terms). This simplification reduces greatly the number of monochromatic radiative transfer calculations required. The number of *k* terms can be adjusted, which provides a trade-off between the accuracy and efficiency required for a given application; however, the correlated-*k* methods cannot be made sufficiently computationally efficient to allow calculations for every grid point at every time step.

To reduce the cost further, calculations are usually made at lower temporal and/or spatial resolutions. Drastic reductions in temporal resolution are often made [e.g., radiation calculations are made every 1 or 3 h for the climate and global forecast models at the National Centers for Environmental Prediction (NCEP) and the Met Office (UKMO; Manners et al. 2009)]. Between radiative transfer calculations, major changes may occur in the radiative profiles (caused primarily by changes in clouds and in the angle of incident solar radiation) that are not represented. A reduced horizontal resolution approach, in which the radiative calculations are performed on a coarser grid with a following interpolation of the results to an original finer grid, is used to speed up radiation calculations at the European Centre for Medium-Range Weather Forecasts (ECMWF; Morcrette et al. 2007, 2008). A reduced vertical resolution approach in which the full radiation is calculated at every other vertical level and interpolated on the intermediate levels, is used in the Canadian operational Global Environmental Multiscale model (e.g., Côté et al. 1998a,b). Such approaches reduce horizontal or vertical variability of radiation fields. Thus, these approaches may reduce the accuracy of a model's radiation calculation and its spatial or/and temporal consistency with other parts of model physics and with model dynamics, which may, in turn, affect negatively the accuracy of climate simulations and weather predictions.

Such a situation is an important motivation for developing new alternative numerical algorithms that provide faster calculations of model physics while carefully preserving their accuracy. Two techniques have been proposed to improve temporal and spatial resolution of radiation calculations: 1) a technique that improves interpolation of the radiative calculations from the coarse grid to the fine grid (Morcrette et al. 2008) or improves radiative calculations between the time steps for which full radiative calculations are performed (Venema et al. 2007; Manners et al. 2009), and 2) a technique that introduces either new and fast radiation parameterizations (Chevallier et al. 1998, 2000) or accurate and fast emulations of existing radiation schemes and parameterizations (Krasnopolsky 1997; Krasnopolsky et al. 2005a, 2008a) that can be used in a model at each grid point and at each time step instead of original slow radiative calculations.

A fast neural network (NN)-based longwave radiation parameterization NeuroFlux (Chevallier et al. 1998, 2000) has been developed and tested in the ECMWF model. The NeuroFlux approach has a limited application (Krasnopolsky et al. 2005b), because it has been developed for a particular formulation (Washington and Williamson 1977) of the longwave radiation physics only. Also, because of NeuroFlux's suboptimal design (Krasnopolsky et al. 2005b), at vertical resolution of 60 layers and more, accuracy and rapidity of NeuroFlux cannot be achieved simultaneously (Morcrette et al. 2008). Consequently, the NeuroFlux is used at ECMWF only for the four-dimensional variational data assimilation linearized physics (Janiskova et al. 2002) for which the accuracy requirements are less stringent.

In our previous studies (Krasnopolsky et al. 2005a, 2008a; Krasnopolsky and Fox-Rabinovitz 2006a,b) we demonstrated that the neural network emulation approach can be successfully used to speed up significantly (by 1–2 orders of magnitude) the calculations of model radiation while providing a sufficient accuracy of decadal (50 yr) climate simulations. We also demonstrated that this approach is a generic one; namely, it can be used not only for emulating any formulation of the longwave radiation physics but also for emulating any formulation of shortwave radiation physics.

In the previous studies (Krasnopolsky et al. 2005a, 2008a), we used a moderate-resolution National Center for Atmospheric Research (NCAR) Community Atmospheric Model (CAM), coupled with a land model, with the T42 (approximately 3°) horizontal resolution and 26 vertical levels (T42L26). In that study, CAM was driven by the climatological sea surface temperature (SST) forcing (with no ocean model coupled).

In this study we apply the NN emulation approach to the higher complexity NCEP Climate Forecasting System (CFS; Saha et al. 2006), which required further development of the neural network emulation methodology. We demonstrate that the NN emulation approach for model radiation can be successfully applied to the significantly higher resolution coupled ocean–atmosphere–land–ice model with time-dependent CO_2. The atmospheric part of CFS has spectral T126 horizontal resolution and 64 vertical levels (T126L64); it is

coupled with the 40-level interactive modular ocean model (MOM4), with a state-of-the-art 3D land model, and with an ice model.

In section 2, the coupled NCEP CFS is described briefly. In section 3, the extended NN emulation methodology and developed NN emulations for the NCEP CFS longwave radiation (LWR) and shortwave radiation (SWR) are briefly described in terms of their design, accuracy, and computational performance. In section 4, the results of the parallel decadal model simulations and seasonal predictions, one using both LWR and SWR NN emulations for calculation of full model radiation (the NN run) and one using the original model radiation (the control run) are compared in terms of similarity of their spatial and temporal variability characteristics. Also, the differences between both parallel runs are compared with the model's internal variability. Section 5 contains conclusions.

2. The NCEP Climate Forecast System

The operational NCEP CFS is described in detail in Saha et al. (2006) and the references therein. The coupled NCEP CFS version (being tested for the operational use) used in our study incorporates the following: the NCEP Global Forecast System (GFS) 64-level atmospheric model, the 40-level interactive MOM4, the interactive Noah land model containing four soil levels with improved treatment of snow and frozen soil, an interactive sea ice model with fractional ice cover and depth allowed, a subgrid-scale mountain blocking, a new seasonal climatological aerosol treatment, a historical CO_2 database from global observations collected by the World Meteorological Organization, a variable solar constant database, and a historical stratospheric volcanic aerosol distributions (Sato et al. 1993).

The NCEP GFS model is a mature, state-of-the-art spectral atmospheric GCM (AGCM) used in operational medium-range weather forecasts. The operational GFS version has a variable horizontal spectral resolution of up to T382 or approximately 38 km. The hybrid sigma–pressure coordinate and a conservative finite-difference scheme are used in the vertical domain. The operational model is run with 64-layer vertical resolution between the surface and 0.27 hPa (approximately 60 km). The current version is implemented in the Message Passing Interface parallel environment. The GFS incorporates parameterizations of a variety of physical processes important in the troposphere and stratosphere and implemented in a ''plug compatibility'' format that facilitates model development. The radiation components contain a GCM version (v2.3) of the rapid radiative transfer model (RRTM) for LWR (hereafter RRTMG-LW) developed at AER, Inc. (e.g., Mlawer et al. 1997; Iacono et al. 2000) and an SWR based on Chou's parameterization scheme (Hou et al. 2002; Chou and Suarez 1999). In the new coupled CFS used in this study the SWR of the operational GFS has been replaced by v2.3 of the AER's RRTM SWR (hereafter RRTMG-SW; e.g., Clough et al. 2005) to improve the accuracy of SWR calculation.

3. NN emulations for the NCEP CFS radiation

a. Background information on NCEP CFS LWR and SWR

LWR and SWR parameterizations in an atmospheric model calculate radiation fluxes and heating rates in the earth–atmospheric system. The RRTMG-LW in the new CFS model employs a computationally efficient correlated-*k* method for radiative transfer calculations. It contains 16 spectral bands with various number of quadrature points (*g* points) in each of the bands that sums up to a total of 140 *g* points (e.g., Mlawer et al. 1997; Iacono et al. 2000). Active gas absorbers include H_2O, O_3, CO_2, CH_4, N_2O, O_2, and four types of halocarbons. A maximum-random cloud overlapping scheme is used for cloudy-sky radiative transfer, and a climatological aerosol scheme provides the global distribution of aerosol optical depth. In this study, a 1-h frequency of radiation calculation is applied to both SWR and LWR. In the current version of the LWR parameterization, the level of atmospheric CO_2 and its time dependence is presented by its global mean value that increased from 350 to 380 ppmv during the period of model integration used in this study (1990–2006).

Beside the RRTMG-LW, which is a faster member of the RRTM LWR family, we have also experimented with another version of the RRTM LWR (hereafter RRTMF-LW) in this study. The RRTMF-LW is based on AER's RRTM-LW v3.0. It uses a full 16 *g* points in each of the 16 spectral bands that add to a total of 256 versus the reduced total of 140 in the faster RRTMG-LW. Unlike the diffusivity approach (one zenith angle of approximately 53°) in the faster RRTMG-LW, the RRTMF-LW uses multiangle radiance integration over a hemisphere to yield better accuracy (we set it at three angles in the study). As a result, the RRTMF-LW is approximately 5 times slower than the RRTMG-LW in exchange for improved accuracy (Mlawer et al. 1997).

The SWR parameterization used in the new CFS is a modified version of AER's RRTMG-SW (v2.3; Clough et al. 2005). It contains 14 spectral bands with various numbers of *g* points in each of the bands to a total of 112. RRTMG-SW uses a fast two-stream radiative transfer scheme and supports sophisticated absorption and scattering processes by clouds, aerosols, and absorbing gases (H_2O, O_3, CO_2, CH_4, N_2O, and O_2). Thus, in the current

version of the SWR parameterization the level of atmospheric CO_2 and its time dependence is presented by the entire 3D CO_2 field that changes with time in accordance with the change of the mean CO_2 level.

Although both RRTMG-LW and RRTMG-SW are built with fast computation schemes designed for GCM applications, they still represent the most time-consuming physics in the NCEP CFS model. The percentage of the total model computation time used by model physics and by LWR and SWR vary depending largely on the model horizontal and vertical resolution, the time step, the frequency of radiative calculations, and the computing environment (e.g., the number of processors and threads). For example, in the new CFS configuration at the T126L64 resolution, with the new RRTMG-LW and RRTMG-SW both called every 1 h, the portion of the radiation computation time is approximately 57% of the total AGCM model computation time.

b. *Background information on the NN emulation approach*

As we showed in our previous works (e.g., Krasnopolsky et al. 2005a; Krasnopolsky 2007a) any parameterization of model physics can be emulated using NNs. NN is an analytical approximation that uses a family of functions such as:

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

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, $\phi(b_{j0} + \sum_{i=1}^{n} b_{ji}x_i)$ is a "neuron," ϕ is the activation function and is usually a hyperbolic tangent, n is the number of inputs, m is the number of outputs, and k is the number of neurons. Definitions of NN terminology can be found in many places [e.g., Bishop 2006 and Krasnopolsky 2007a]; however, (1) is sufficient to understand the subject of this paper. The numerical complexity of NN in (1) can be well approximated by the number of NN weights and fitting parameters in (1) (Krasnopolsky 2007a):

$$N_C = k(n + m + 1) + m. \tag{2}$$

The time T_{NN} required to estimate the NN in (1) is directly proportional to the NN numerical complexity N_C, that is, $T_{\mathrm{NN}} = cN_C$, with the coefficient of proportionality c depending mainly on the hardware and software environment of the computer used.

Obviously, N_C increases linearly with the increase of model vertical resolution (the number of the vertical layers L), because both n and m depend linearly on L. As a result, T_{NN} increases linearly with the increase of model vertical resolution. The time required to estimate the original parameterization T_O also increases with the increase of vertical resolution. For the original parameterization, the dependence of the calculation time on vertical resolution is strongly conditioned by the numerical scheme implemented. For example, the calculation time increases for RRTMG-LW approximately linearly with the increase of L versus an L^2 relationship for most of other types of LWR used in many models.

Thus, the dependence of the speedup η provided by an NN emulation on model vertical resolution is determined by the ratio T_O:T_{NN}. Therefore, the change of η with the increase of model vertical resolution will strongly depend on the physical complexity of the original parameterization and on the numerical scheme implemented (see also the discussion at the end of section 3f).

The major goal for developing NN emulations for model physics is to obtain a sufficiently high accuracy for NN emulation with practically zero biases or systematic errors (calculated against the original model physics), which is a necessary condition for obtaining nonaccumulating errors during long-term climate simulations that use developed NN emulations. The choice of an optimal version of NN emulation is based on the accuracy, not on a speedup of computation. All of the obtained NN emulations provide a 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, discussed, and applied in our research (e.g., Krasnopolsky et al. 2005a,b, 2008a).

c. *NN emulations for full model radiation*

The LWR and SWR parameterizations together comprise the full model radiation. The full model radiation for the NCEP CFS has been emulated using two NNs, one for LWR and one for SWR.

The input and output vectors for NNs, emulating the LWR or SWR parameterizations, include the same parameters as those of the input and output vectors for the original LWR or SWR parameterizations, respectively. For the LWR NN emulation, these input parameters are the following nine profiles: atmospheric pressure, temperature, specific humidity, ozone mixing ratio, total cloud fraction, cloud liquid water path, mean effective radius for liquid cloud, cloud ice water path, and mean effective radius for ice cloud. The time-dependent CO_2 is presented, as in the original LWR parameterization, by its changing with time global mean. The LWR parameterization (and LWR NN emulation) output vectors consist of the profile of heating rates (HRs) and five radiation

fluxes: the total sky outgoing LW radiation flux from the top layer of the model atmosphere (the outgoing LWR or OLR), the clear-sky upward flux at the top of the model atmosphere, the total sky upward flux at the surface, the total sky downward flux at the surface, and the clear-sky downward flux at the surface.

The NN emulation of the LWR parameterization includes all nonconstant inputs of the original LWR [$n = 556$ in (1)]. It has the same outputs [$m = 69$ in (1)] as the original LWR parameterization. We have developed several NNs, all of which have the same aforementioned inputs and outputs, with k changing from 50 to 200 in (1). Varying k in (1) allows us to demonstrate the dependence of the accuracy of approximation on this parameter as well as its convergence, and as a result, to provide a sufficient accuracy of approximation for the model (e.g., Krasnopolsky et al. 2005a).

The input vectors for the SWR parameterization include 55 vertical profiles: atmospheric pressure, temperature, specific humidity, ozone, CO_2, N_2O, O_2, and CH_4 volume mixing ratios, total cloud fraction, cloud liquid water path, mean effective radius for liquid cloud, cloud ice water path, mean effective radius for ice cloud, and 3 profiles (optical depth, single scattering albedo, and asymmetry parameter) for each of 14 different species of aerosols. The input vectors include also the solar zenith angle, the solar constant, and the surface albedo for four different bands. The SWR parameterization output vectors consist of one vertical profile of HRs and nine radiation fluxes: three fluxes at the top layer of the model atmosphere (the total sky outgoing SW radiation flux, the total sky downward flux, the clear-sky upward flux), four radiation fluxes at the surface (the total sky upward and downward fluxes and the clear-sky upward and downward fluxes), and the downward (the total and clear sky) fluxes in the UV-B spectral band.

The NN emulations of the SWR parameterization have 562 inputs and 73 outputs. We have developed several NNs, with k changing from 50 to 200 in (1). It is noteworthy that, as in the case of the NN emulation of LWR, the number of NN inputs is less than the number of input profiles multiplied by the number of vertical layers plus the number of relevant single-level characteristics. Many input variables (e.g., almost all gases) have zero or constant values for the upper vertical layers, and for some gases the entire volume mixing ratio profile is a constant (obtained from climatological data).

To improve the accuracy of the approximation, these constant inputs were not used for NN training. Constant inputs (zero or nonzero) do not contribute to the functional input/output relationship and should not be used for development of NN emulations. Moreover, if they were used, they would introduce an additional noise (an approximation error). In addition, for SWR, 2688 inputs describing the optical depth, the single scattering albedo, and asymmetry parameters of 14 aerosol species were substituted by five inputs: $\cos(\tau)$, $\sin(\tau)$, cos(lon), sin(lon), and lat, where lon is the longitude, lat is the latitude, and $\tau = 2\pi q T^{-1}$, where q is the month of the year, and $T = 12$ is evidently the number of months in the year. Such a substitution is possible, because in NCEP CFS aerosols are calculated using the specific humidity profiles and 3D lookup tables composed of climatological monthly data, different for different months of the year. It means that the aerosol inputs are actually highly correlated, and, in terms of functional input/output dependences, the aerosol characteristics are the functions of lat, lon, τ, and the profile of specific humidity only. Since the profile of the specific humidity has been already included in NN SWR inputs, only the five aforementioned additional variables have to be included to allow NN to completely emulate the contribution of aerosols into SWR.

We would like to emphasize that not including the constant profiles as inputs into the NN emulation or reducing the number of highly correlated profiles (as in the case of aerosols) does not in any way diminish the accuracy of the NN emulation (see sections 3e and 4). All changing/nonconstant inputs are included into the NN emulation. All constant inputs are included in the original parameterization when NCEP CFS has been used to generate the NN training datasets (see section 3d), and the NN emulation receives information about them from the training data during the NN training process. From the practical point of view, if the values of constant inputs are changed or are made variable in the future, the NN emulation inputs will be adjusted correspondingly and NN will be retrained.

It was mentioned above that radiation parameterizations have two kinds of outputs: the heating rates and the radiative fluxes. They are not completely independent; there is a balance relationship between them (see appendix A). The outputs of original radiation parameterizations satisfy the balance relationship with high accuracy because the relationship is explicitly (or implicitly) included into the parameterizations. Obviously, the outputs of the NN emulations satisfy it only approximately but with high accuracy. A balancing procedure has been developed to balance NN outputs exactly. Table 1 shows that the balancing procedure does not practically affect the overall accuracy of LWR NN and only marginally improves the overall accuracy of SWR NN.

d. Generating datasets for NN training and validation

The NCEP CFS (T126L64) has been run for 17 yr to generate representative training datasets. The representative dataset samples the atmospheric state variability

TABLE 1. Statistics estimating the accuracy of HR (K day^{-1}) calculations and the computational performance for NCEP CFS (T126L64) LWR and SWR using NN emulation vs the original parameterization. For comparison, NCAR CAM (T42L26) LWR and SWR statistics are also shown. Total statistics show B and RMSE using (B1) and PRMSE and σ_{PRMSE} using (B4) for the entire 3D HR fields. Top- and bottom-layer statistics show B and RMSE using (B2) for one horizontal layer (top or bottom layer). Also, the changes in statistics due to applying the balancing procedure (see appendix A) are shown for RRTMG-LW and -SW NN emulations. The N_C in (2) and average η are shown.

	LWR				SWR		
		NCEP CFS				NCEP CFS	
Statistics	NCAR CAM	RRTMG*	Change due to balancing	RRTMF*	NCAR CAM	RRTMG*	Change due to balancing
Total error							
B	3.0×10^{-4}	2.0×10^{-3}	6.0×10^{-4}	7.0×10^{-4}	-4.0×10^{-3}	5.0×10^{-3}	-3.0×10^{-3}
RMSE	0.34	0.49	1.0×10^{-4}	0.42	0.19	0.20	-5.0×10^{-3}
PRMSE	0.28	0.39	3.0×10^{-4}	0.30	0.15	0.16	-5.0×10^{-3}
σ_{PRMSE}	0.2	0.31	1.0×10^{-4}	0.30	0.12	0.12	1.0×10^{-3}
Bottom-layer error							
B_m	-2.0×10^{-3}	-1.0×10^{-2}	-6.0×10^{-4}	6.0×10^{-3}	-5.0×10^{-3}	9.0×10^{-3}	-8.0×10^{-3}
RMSE$_m$	0.86	0.64	1.0×10^{-5}	0.67	0.43	0.22	-0.01
Top-layer error							
B_m	-1.0×10^{-3}	-9.0×10^{-3}	6.0×10^{-4}	2.0×10^{-3}	2.0×10^{-3}	1.3×10^{-2}	4.0×10^{-3}
RMSE$_m$	0.06	0.1782	4.0×10^{-3}	0.09	0.17	0.21	1.0×10^{-3}
N_C	12 733	33 294	—	93 969	11 418	45 173	—
η**	150	16	—	21	20	60	—

* RRTMG and RRTMF are different versions of the radiation code developed by AER, Inc. (see section 2 and references therein).
** Here, η shows the average (over a global dataset) speedup or how many times NN emulation is faster than the original parameterization in a sequential single processor code by code comparison.

adequately; that is, it represents all possible states produced by the model as fully as possible (including the states introduced because of time-dependent CO_2 concentration). All inputs and outputs of original LWR and SWR parameterizations have been saved for two days per month, that is, for one day at the beginning and one day in the middle of the month, every 3 h (8 times per day) to cover the annual and diurnal cycles. From each of the three-hourly global dataset, 300 events (the set of input and output profiles) have been randomly selected. The obtained dataset has been divided into three independent parts, each containing approximately 200 000 input/output vector combinations. The first part has been used for training, the second part for tests (control of overfitting, control of NN architecture, etc.), and the third part (an independent validation dataset) has been used for validation of trained NN only. All approximation statistics presented in this section are calculated using this independent validation dataset. The accuracy of the NN emulation; that is, mean errors (or biases), RMSE, etc., are calculated against the control (the original parameterization).

e. Bulk approximation error statistics

To ensure a high quality of representation of the LWR and SWR processes, the accuracy of their NN emulations has been carefully investigated. The NN emulations have been validated against the original NCEP CFS LWR and SWR parameterizations. To calculate the error statistics presented in Tables 1 and 2 and in Fig. 1, the original parameterizations and their NN emulations have been applied to the validation dataset. Two sets of the corresponding HR profiles have been generated for both LWR and SWR. Total and level mean differences (or mean errors), total and level RMSE, profile RMSE (PRMSE), and σ_{PRMSE} have been calculated (see appendix B and Krasnopolsky 2007a for the definition of these statistics).

Table 1 shows bulk validation statistics for the accuracy of approximation of HR and the average computational performance or the speedup (see section 3f). For comparison, the information on the NN emulations for NCAR CAM LWR and SWR (Krasnopolsky et al. 2008a) is also presented in Table 1. Total statistics show the mean difference, RMSE [see (B1) in appendix B], PRMSE, and σ_{PRMSE} in (B4) for the entire 3D HR fields. Also, layer statistics in (B2) for the top and bottom atmospheric layers are included to illustrate the accuracy of NN emulations in the areas of the increased nonlinearity (Morcrette et al. 2008). Although the two models as well as their embedded radiation parameterizations are different, comparisons between NCAR CAM with 26 vertical layers and NCEP CFS with 64 vertical layers allow us to observe a general dependence of the NN accuracy on the model vertical resolution (see also error profiles shown in Fig. 1).

As can be concluded from Table 1 and Fig. 1, NN emulations for both LWR and SWR handle very well the nonlinearity at the top of the atmosphere where mean

TABLE 2. Comparison of calculation time (seconds per 3000 profiles) and η (as in Table 1) for RRTMG-LW and -SW original parameterizations and NN emulations under different cloud conditions. The calculations were performed using a single processor of IBM Power 6 supercomputer.

	RRTMG-LW			RRTMG-SW		
Parameterization	Clear sky	Three-layer clouds	Deep convection	Clear sky	Three-layer clouds	Deep convection
Original	9.6	10.1	11.7	33.8	42.8	52.9
NN	0.6	0.6	0.6	0.6	0.6	0.6
η	16	16.8	19.5	56	71	88

differences and RMSEs are very small, with RMSEs being even smaller than the total RMSE. At the bottom layer, the nonlinearity does not cause significant increases in mean differences; the RMSEs increase approximately 2 times but remain sufficiently small compared with the total RMSE.

It is noteworthy that the approximation errors are identified as being sufficiently small if they are of such a small magnitude that they have almost negligible impacts on model behavior as demonstrated in section 4 for NCEP CFS and by Krasnopolsky et al. (2008a) for NCAR CAM. Only validation of NN emulations in parallel model runs allows us to make final conclusions about the sufficient smallness of the approximation errors.

In terms of the presented accuracy statistics, there are practically no differences between NCAR CAM with 26 vertical layers and NCEP CFS with 64 vertical layers. As shown in Fig. 1, the entire vertical distributions of errors for both LWR and SWR are similar for these two models. Thus, the accuracy of our NN emulation approach does not depend significantly on vertical resolution of the model. It does depend on the vertical location of the atmospheric layer. The layer RMSE increases near the surface for both models.

f. Estimation of speedup

The N_C in (2) and average η are also shown in Table 1. These characteristics complement our discussion on the

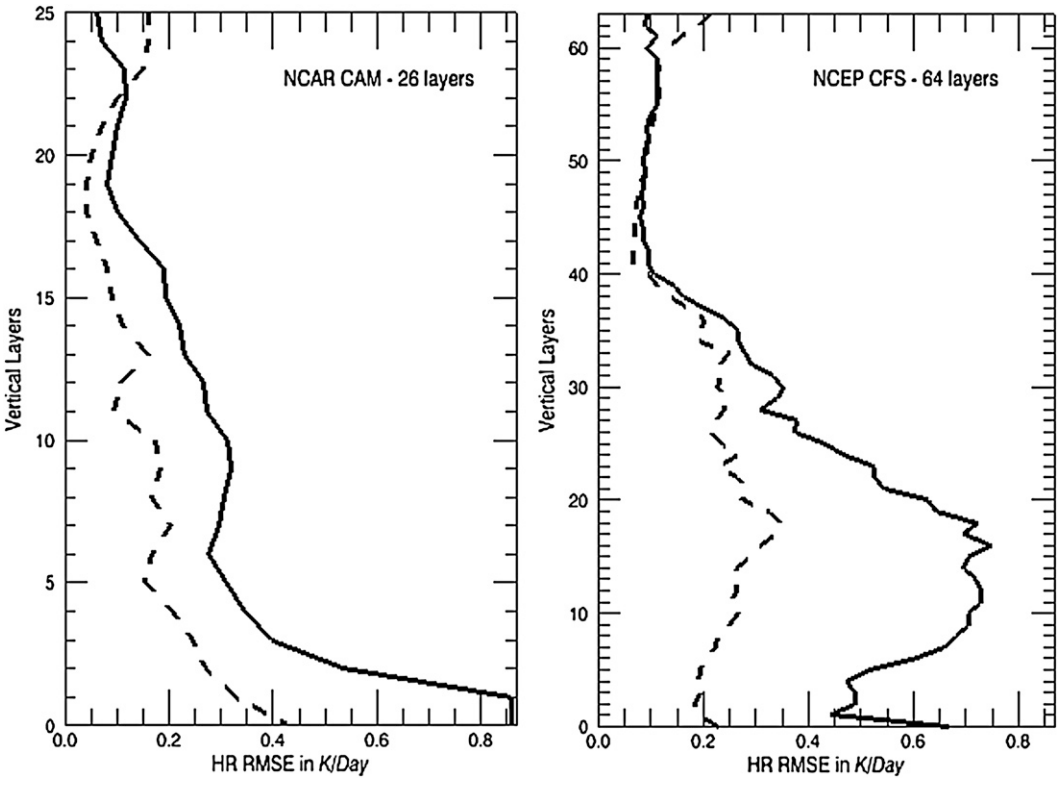

FIG. 1. Vertical distributions of NN emulation RMSEs for (left) NCAR CAM (26 vertical layers) and (right) NCEP CFS (64 vertical layers). The solid line is LWR, and the dashed line is SWR. The errors and their vertical distributions are similar for both models.

dependence of the speedup on vertical resolution (see section 3b). For the LWR parameterization, we see a significant decrease of the speedup for NCEP CFS with 64 vertical layers versus NCAR CAM with 26 vertical layers, although the LWR NN emulation for NCEP CFS is still 16 times faster than the original parameterization. For the SWR parameterization the opposite tendency is observed; that is, the speedup for NCEP CFS SWR NN is more than 3 times higher than that of NCAR CAM SWR NN.

The seemingly contradictory speedups for LWR and SWR emulations can be explained (as was mentioned in section 3b) by the interplay of the two main contributing factors: the physical complexity of the radiation calculation itself (the number of treated species, spectral bands, parameterization schemes, etc.), and the dependence of the particular numerical scheme implemented in the radiative transfer on the number of vertical model layers. The results presented in Table 1 illustrate the fact that the numerical scheme implemented in the NCEP CFS RRTMG-LW parameterization is significantly more efficient (linear with respect to L) than that of the original NCAR CAM LWR parameterization (quadratic with respect to L). Thus, a smaller speedup factor is produced by the NN emulation for NCEP CFS LWR than for NCAR CAM LWR.

The NCEP CFS's RRTMG-SW includes more spectral bands and g points and uses a more complex treatment for a larger variety of absorbing/scattering species; thus NN emulation shows a larger η than for NCAR CAM SWR. In any case, our NN emulation approach is significantly less dependent (in terms of both the accuracy and speedup) on the increase of vertical resolution than the NN-based LWR parameterization NeuroFlux for which at vertical resolution of 60 layers and more, both accuracy and speedup could not be achieved simultaneously (Morcrette et al. 2008). For our NN emulation approach with the model with 64 vertical layers, the desired accuracy of the NN emulation could be achieved simultaneously with a significant speedup of approximately 16 times for the LWR and of approximately 60 times for the SWR parameterizations.

The radiative transfer calculations take different amounts of time under different cloud conditions because of the different complexity of cloud–radiation interaction. We performed more detailed estimations of speedup separately for three different types of cloudiness: clear sky, three layer clouds, and a more complex cloud condition of deep convection. 3000 profiles have been used for each test. The results for the calculation time and speedup are presented in Table 2. For a more complex cloud–radiation interaction (deep convection) the calculation of the original LWR and SWR parameterizations takes approximately 22% and approximately 57% more time, respectively, than for clear-sky conditions. Obviously, the time of the NN radiation calculations does not depend on the cloud conditions. Thus, the speedup is significantly higher for the more complex cloud–radiation interaction.

As Table 2 shows, the average speedup presented in Table 1 is closer to the minimal speedup obtained under clear-sky conditions. The results presented in Tables 1 and 2 were obtained in a sequential code-by-code comparison and represent adequately the situation when the model is run on a single processor; however, if we compare the control model run using the original parameterizations with the NN run when both runs use multiple processors and threads, the actual speedup will be significantly higher and closer to the maximum value shown in Table 2, because it will be determined by the slowest calculation, which is the deep convection condition. Radiation in the control run for all other cloud conditions is calculated faster, but the next integration time step will not start before the radiation calculations for the deep convection condition are completed; however, the time of the NN radiation calculations in the NN run does not depend on the cloud conditions. Thus, in the case of parallel calculations, utilizing multiple processors and threads, in addition to a significant speedup, the use of the NN emulations in the model provides an additional advantage; namely, it helps to achieve a significantly better load balance.

Using NN emulations simultaneously for full model radiation results in an overall significant—approximately 20%–25%—speedup of NCEP CFS climate simulations and seasonal predictions when both LWR and SWR are calculated every 1 h. The η provided by NN emulations (Table 1) can be also used for more frequent calculations of model radiation.

4. Validation of parallel decadal model simulations and seasonal predictions

In this section we present comparisons between two parallel 17-yr NCEP CFS model runs: one using the original LWR and SWR parameterizations (the control run) and one using their NN emulations. Also, the differences between parallel runs are compared with the model's internal variability. Both spatial and temporal characteristics of prognostic and diagnostic fields are compared for the parallel runs.

a. Measures for assessment of the impact of using NN emulations of full model radiation

We show below the differences between the two parallel runs. To evaluate the NN induced changes, we compare them with such commonly used measures as

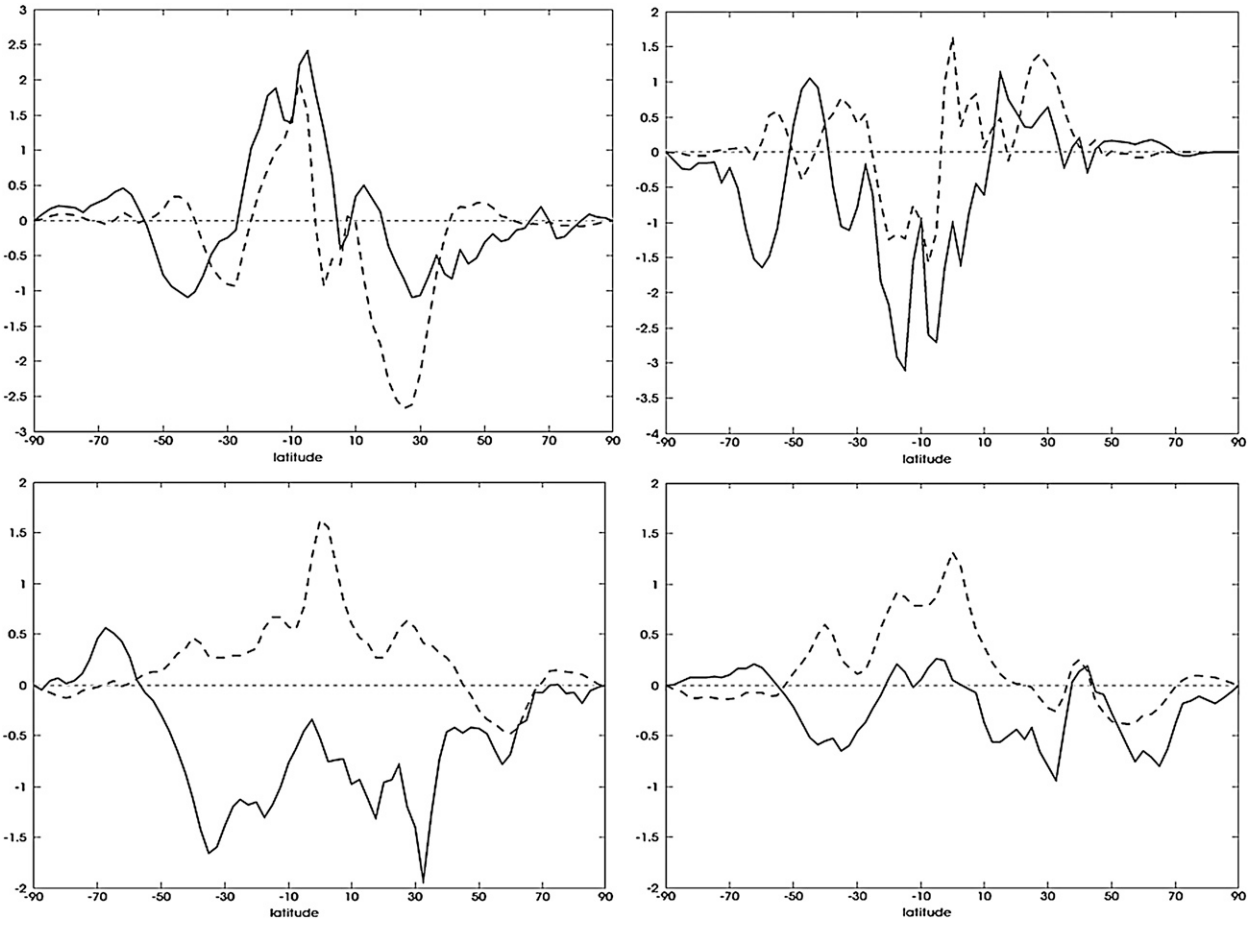

FIG. 2. (top) Zonal and time-mean top-of-atmosphere upward (left) LWR and (right) SWR flux (W m^{-2}) for DJF. The solid line is NN − CTL (the difference between the NN and a control run), and the dashed line is CTL1 − CTL (the difference between two control runs) presented for comparison. (bottom) Zonal and time-annual mean (left) downward and (right) upward surface LWR flux (W m^{-2}). The fluxes' differences are multiplied by cos(lat) to equalize the areas.

observation errors or uncertainties of reanalysis. We show that the differences are smaller than these quantitative measures.

To emphasize how small the changes introduced by the use of NN emulations are, we also find it appropriate to use a measure derived from the model itself; namely, the model's internal variability. Because a GCM is an essentially nonlinear system, it may produce something like a butterfly effect, that is, a significant reaction or response even to small perturbations in the model or in the model computational environment (e.g., routine changes in computer hardware, operational system, compilers, libraries, etc.). Any change, even infinitesimal, in model formulation, initial conditions, or computational environment makes two model integrations diverge, with the effect that after the deterministic predictability is lost (which takes just weeks for the atmosphere, although longer for the ocean), the timing and location of weather patterns becomes essentially independent for the two integrations. Hence the two control model runs produced with the aforementioned small changes provide, in essence, two independent samples of the model's climatology, and their difference represents the model's internal variability.

Thus, we can state that the approximation error of NN emulation is negligible and, therefore, the NN's accuracy is sufficient for the use in the model if the differences or changes introduced in the model results by using the NN emulation are of the same order of magnitude as the aforementioned model's internal variability.

To estimate the model's internal variability, we produced two control runs with the original NCEP CFS model configuration, that is, without NNs. The first run was performed before and the second run after the routine changes (introduced quasi-regularly by system administrators) of the version of the FORTRAN compiler and libraries. Small differences between these two runs, which are similar to those due to changes in a computer

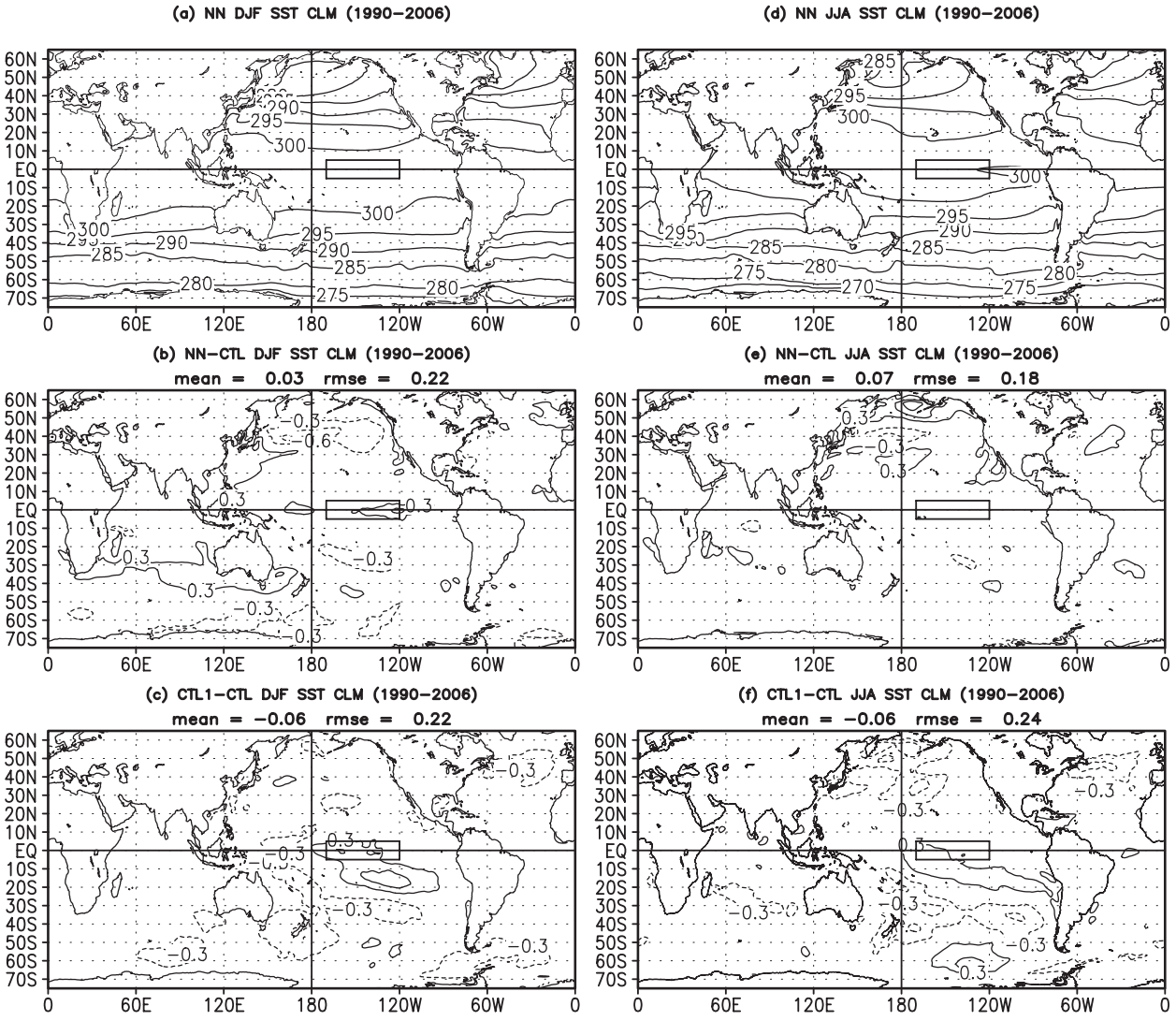

FIG. 3. The 17-yr (1990–2006) time-mean SST distributions and mean differences for (left) DJF and (right) JJA. (top) Full radiation NN runs, (middle) NN − CTL, and (bottom) CTL1 − CTL shown for comparison. The contour intervals for SST distributions are 5 K and for SST differences are 0.3 K. Numbers above the figures in the middle and lower rows show global mean difference and RMSE.

operation system and/or in hardware, are shown below together with the differences between the parallel NN and control runs for comparison purposes as an additional measure of the NN emulation accuracy. Presenting the model's internal variability helps us to better evaluate the differences in climate simulations and seasonal predictions caused by using NN emulations for model radiation and to emphasize how small these differences are.

b. Comparison of parallel runs

1) CLIMATE SIMULATIONS

The results of 17-yr (1990–2006) climate simulations performed with NN emulations for the full model radiation have been validated against the parallel control NCEP CFS simulation using the original LWR and SWR. We analyze the differences between the parallel runs in terms of spatial (global) means as well as temporal characteristics.

Let us discuss first the differences between the parallel simulations in terms of spatial and temporal radiation characteristics. The differences between the NN radiation and control runs and the differences between two control runs for zonal and time-mean LWR and SWR fluxes are presented in Fig. 2. The upper row of Fig. 2 shows the differences for zonal and time-mean top of atmosphere upward longwave (left panel) and shortwave (right panel) fluxes (W m^{-2}) for the NH winter (December–February; DJF). The lower row of Fig. 2 shows the differences for zonal and time-mean downward

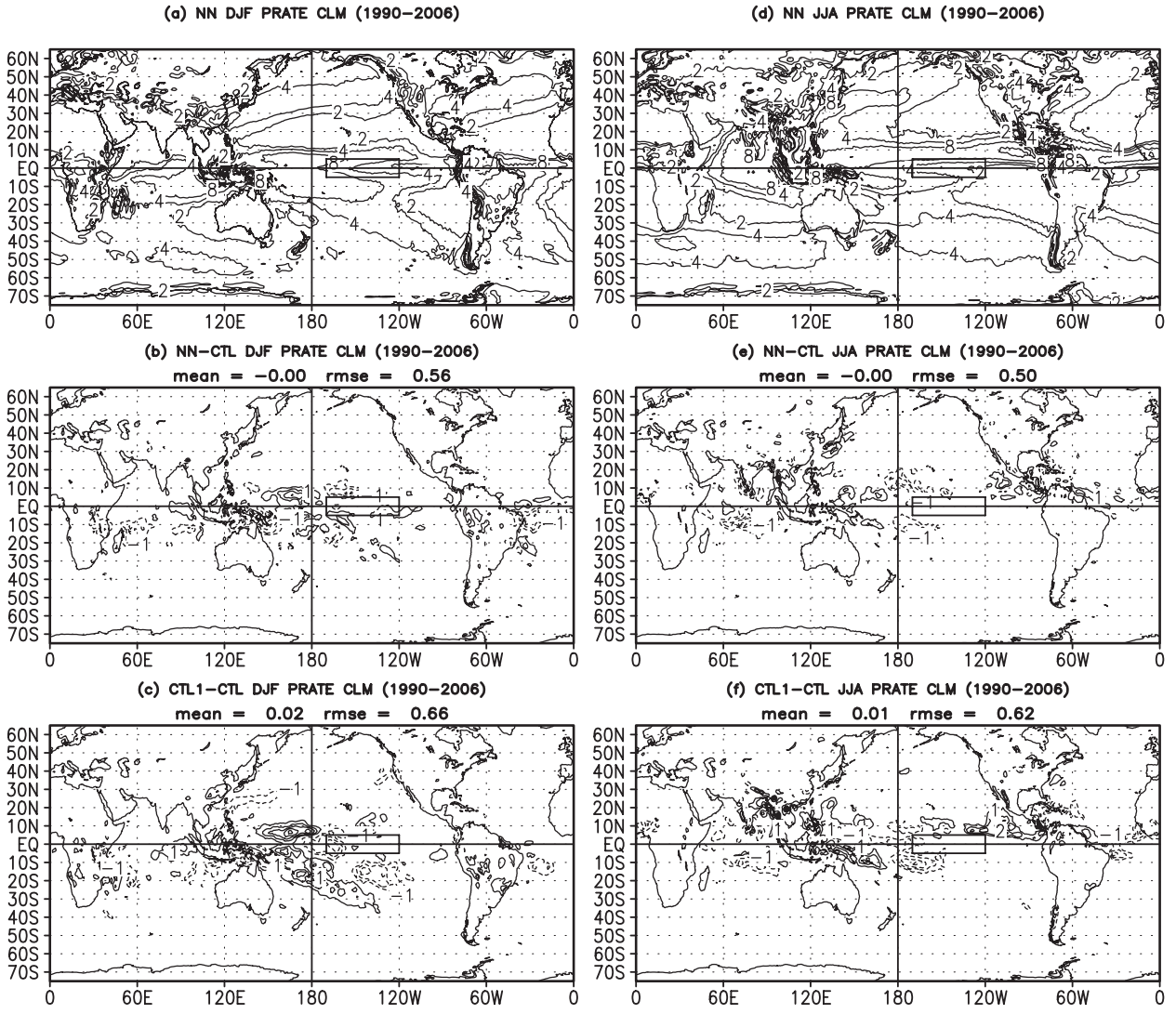

FIG. 4. As in Fig. 3, but for PRATE. (top) The contour levels for PRATE fields are 2, 4, 8, 16, and 32 mm day^{-1}. (middle),(bottom) The contour interval for PRATE differences is 1 mm day^{-1} with 0 mm day^{-1} contour removed for clarity.

(left panel) and upward (right panel) surface longwave fluxes (W m^{-2}). For the fluxes presented in Fig. 2, both the differences between the NN radiation and control runs and the differences between two control runs are small and similar by magnitude. They do not exceed 2–3 W m^{-2}; that is, they are within observational errors and uncertainties of reanalysis (e.g., Kalnay et al. 1996; Kistler et al. 2001). The similarity of the differences by magnitude means that both the differences between the NN radiation and control runs are comparable with the model's internal variability. The HR differences are also very close in magnitude to, but do not exceed, the model's internal variability described in section 4a.

Let us discuss now prognostic and diagnostic fields, such as SST, precipitation, different types of clouds, and time series that are potentially sensitive to changes in the model resulting from using NN emulations. Close similarities have also been obtained for these results of the parallel runs in terms of time-mean spatial fields presented in Figs. 3–7, which have the same design. In each of these five figures, the left column shows results for DJF and the right column for NH summer (June–August; JJA). The top panels [labeled (a) and (d)] show fields produced by the full radiation NN run, the middle panels [(b) and (e)] show mean errors or the difference between the full radiation NN run and the control run (CTL) NN − CTL, and the bottom panels [(c) and (f)] show the differences between two control runs (i.e., model's internal variability) CTL1 − CTL presented for comparison. Notice that spatial (global) and time-mean errors and

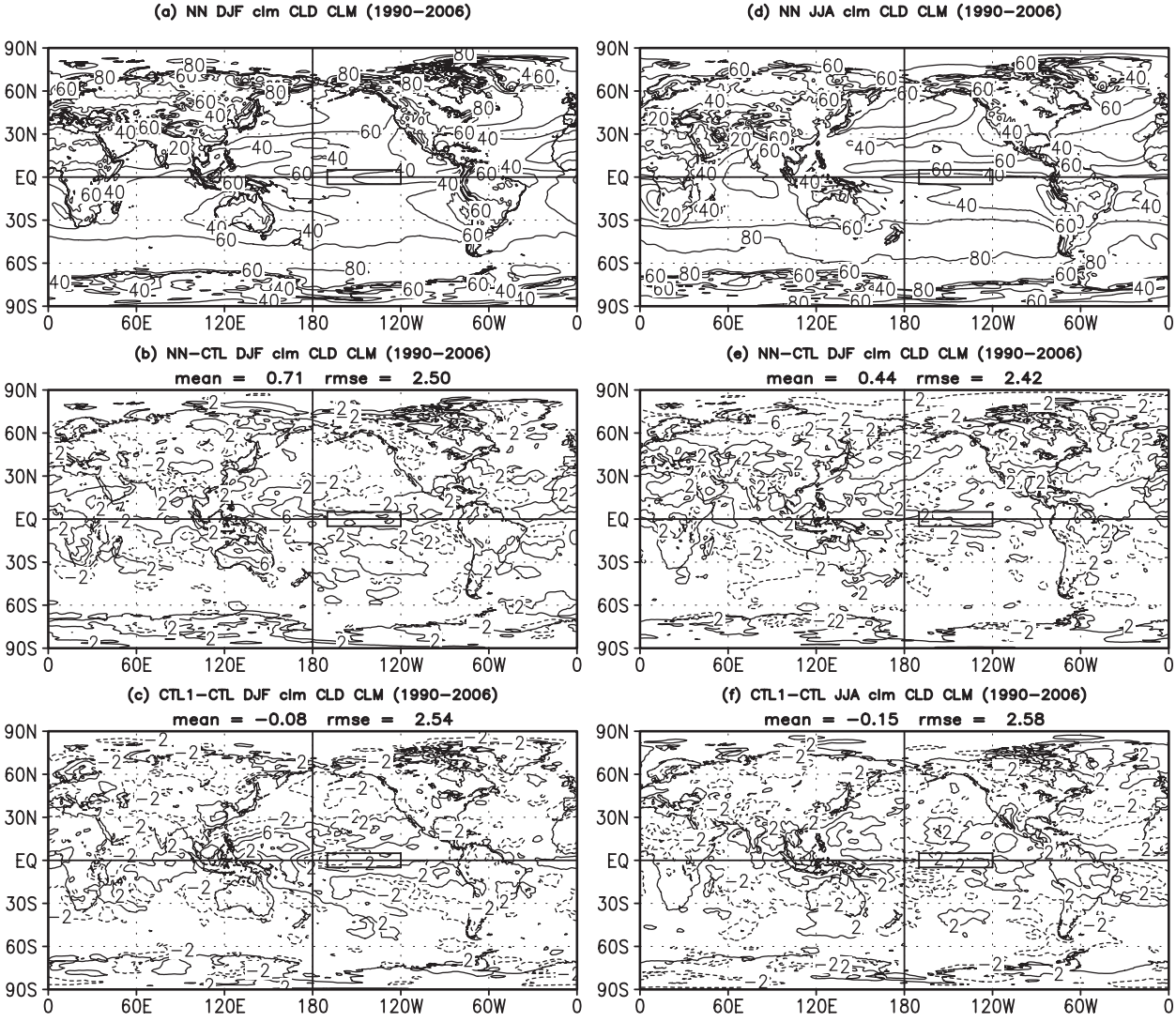

FIG. 5. As in Fig. 4, but for clm CLD. The contour interval (top) for clm CLD fields is 20% and (middle),(bottom) for differences is 4% with 0% contour removed for clarity.

RSMEs are shown in the panel titles for NN − CTL and CTL1 − CTL.

The 17-yr (1990–2006) time-mean SST distributions and differences for the full radiation NN run versus the control run and the differences between two control runs (model's internal variability) are presented for NH summer and winter in Fig. 3. The SST mean differences and RMSEs for NN − CTL are very small; they are not larger than those of CTL1 − CTL. The time- and global mean errors are near zero and RMSEs are just a small fraction of 1 K. The results for other two seasons (NH spring and fall) are similar.

Figure 4 show the 17-yr (1990–2006) time–mean distributions and differences for total precipitation (PRATE) for the parallel full radiation NN and control runs for NH summer and winter, respectively. The PRATE mean differences are quite limited and occur mostly in the tropics; they are also very close in magnitude (as well as RMSEs) and pattern to the model's internal variability. The results for other seasons are similar.

Figures 5–7 show comparisons for the parallel full radiation NN and control runs for different types of clouds. They present the 17-yr (1990–2006) time-mean distributions and differences of total cloud (clm CLD; Fig. 5), convective precipitation clouds (cvl CLD; Fig. 6), and boundary layer clouds (bcl CLD; Fig. 7) for NH summer and winter. Clouds are very sensitive to any changes in the model and, therefore, provide a suitable and sensitive estimate of the accuracy of NN emulations.

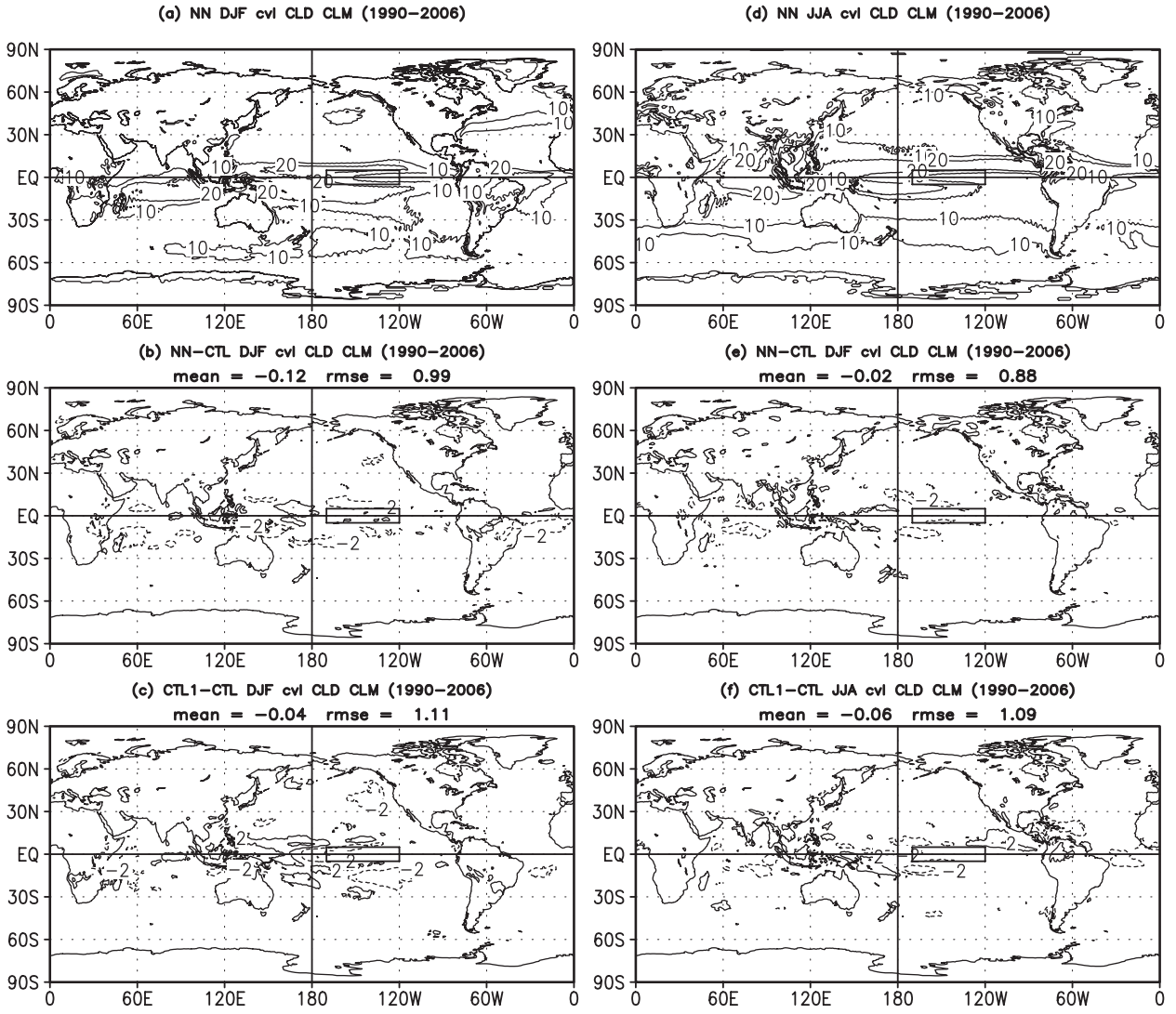

FIG. 6. As in Fig. 5, but for cvl CLD. The contour interval (top) for cvl CLD fields is 10% and (middle),(bottom) for differences is 4%.

For all types of clouds shown in Figs. 5–7, the cloud patterns and differences for parallel total radiation NN and control runs are very close for both seasons presented. The situation is similar for other seasons (NH spring and fall) and types of clouds, such as low-, mid-, and upper-level clouds. The mean differences are very small and occur mostly in the tropics. They have the same magnitude (as well as RMSE) and pattern as the differences between two control runs or model's internal variability shown for comparison. For all presented clouds, the time- and global mean errors are approximately 0%–1%, and RMSEs are approximately 1%–2.5%.

Let us compare now the results of the parallel NN and control runs in terms of temporal characteristics. The global mean time series for monthly means of the total precipitable water (PWAT), with the seasonal cycle subtracted, are presented in Fig. 8. The figure shows the 17-yr (1990–2006) time series for the parallel full radiation NN run (the dashed–dotted line) and for two control runs described in section 4a (the solid line for CTL and the dotted line for CTL1). The time series for PWAT presented in Fig. 8 for the parallel full radiation NN and the CTL show an overall similarity for the entire 17-yr period. The differences between two control runs are similar but marginally larger.

Figure 9 shows the 17-yr time series for the Niño-3.4 index for the reanalysis from the NCEP climate data assimilation system (CDAS) and for the parallel full radiation NN and the two control runs described in section 4a. The Niño-3.4 index is calculated over the small area in the equatorial Pacific Ocean shown by the black rectangle in Figs. 3–7. The time series for the Niño-3.4 index are

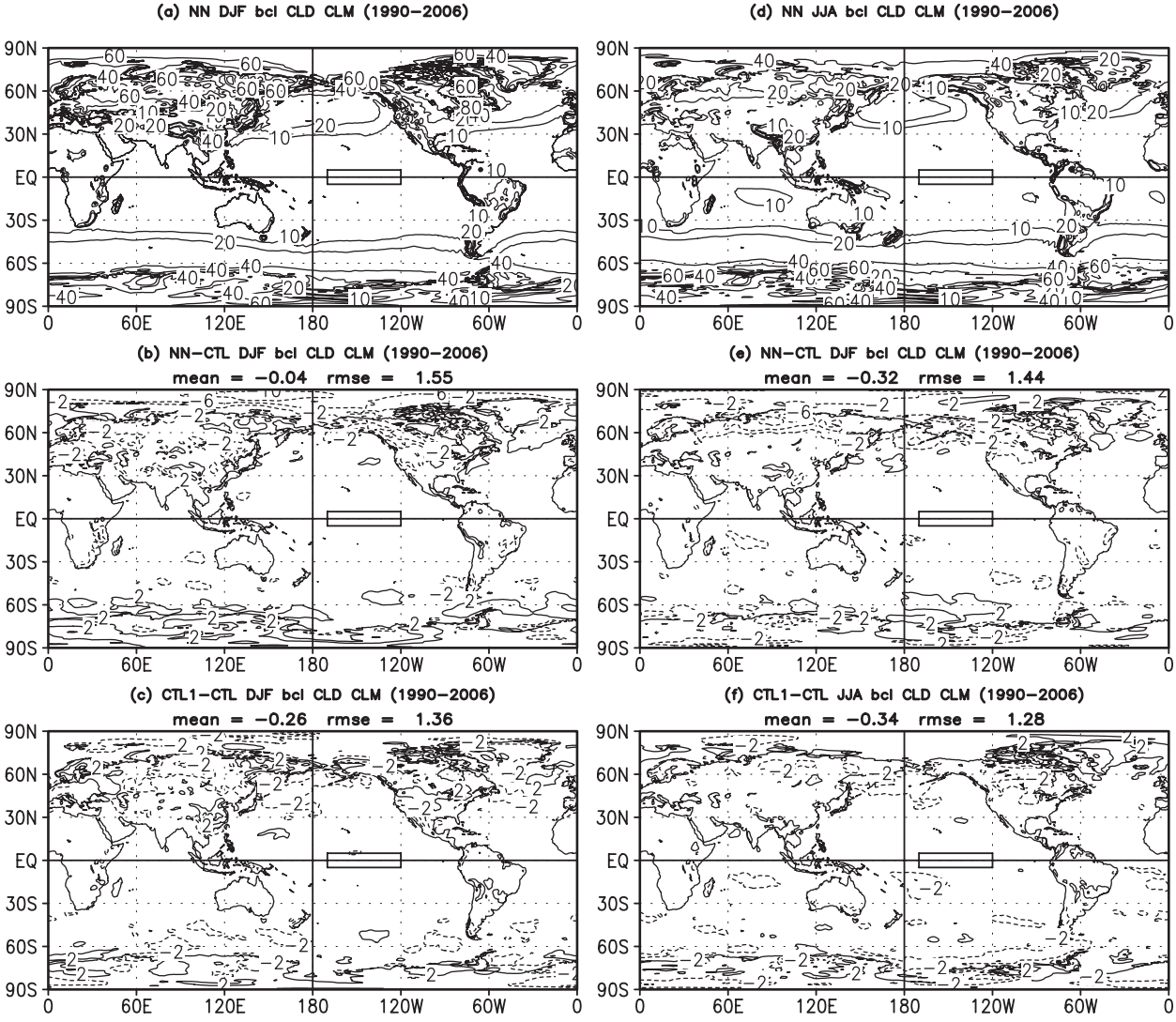

FIG. 7. As in Fig. 5, but for bcl CLD. (top) The contour levels for bcl CLD fields are 10%, 20%, 40%, 60%, 80%, and 100%. (middle),(bottom) The contour interval for bcl CLD differences is 4%.

affected by a quite limited SST anomaly sampling for the relatively small area and are very sensitive to any changes in the model or in its computational environment as can be seen from Fig. 9. The explanation for the different details of the Niño-3.4 time series is that timing and magnitude of El Niño–Southern Oscillation (ENSO) events is chaotic and subject to different phases of internal variability in the different runs. As can be seen from the standard deviation values included in Fig. 9, the model overestimates the ENSO variability compared with CDAS. The overall dissimilarity of the indices or their deviation from CDAS is not larger than that of the two control runs from CDAS and from each other.

Figure 10 shows the 17-yr time series for global mean temperature at 850 hPa for the parallel full radiation NN and the two control runs. All three time series are similar; the differences do not exceed 0.5 K. The small differences between the full radiation NN and control runs are of the same magnitude as those of between two control runs.

The time-mean-simulated products presented in Figs. 3–7 as well as other model simulated products show that mean differences and RMSEs for the full radiation NN run are small; that is, they are overall within the observational errors or uncertainties of reanalysis and are of a similar magnitude as the model's internal variability.

Close similarity has also been obtained for other model prognostic and diagnostic fields in term of their spatial and temporal characteristics. Summarizing, from the obtained validation results, we can conclude that the differences between decadal climate simulations produced by the parallel full radiation NN and control runs

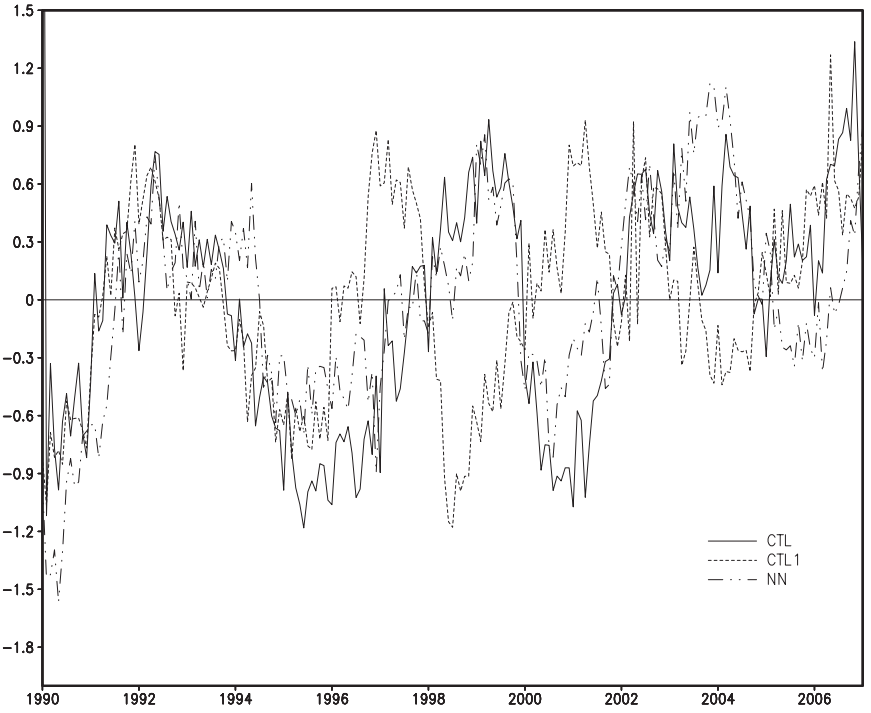

FIG. 8. The 17-yr (1990–2006) time series of PWAT (kg m^{-2}), with the seasonal cycle subtracted, for the full radiation NN run (dashed–dotted line), CTL (solid line), and CTL1 (dashed line).

are overall within or less than the observation errors and uncertainties of reanalysis (e.g., Kalnay et al. 1996). Moreover, these differences (both in terms of mean differences and RMSEs) are of a similar magnitude as the model's internal variability or the differences between two control runs, which are regularly introduced in climate models by routine changes in computer environment (like changes in hardware, operational system, and/or compilers).

2) SEASONAL PREDICTIONS

We performed similar validation for seasonal predictions for 1990. Basically, the results are similar to those presented above. Figure 11 shows NN − CTL and CTL1 − CTL for seasonal predictions of SST, clm CLD, PRATE, and cvl CLD.

All the patterns for the control and NN runs (not shown) are quite similar. The differences between seasonal predictions produced by the parallel full radiation NN and control runs are slightly larger than the differences for climate simulations shown in section 4b(1). It is mostly due to a significantly shorter simulation period used for seasonal predictions; however, the differences are still comparable with the observation errors and uncertainties of reanalysis. The differences do not increase significantly from season one to season four. For the seasonal predictions, NN − CTL is close in magnitude and does not exceed CTL1 − CTL or model's internal variability described in section 4a. The time- and global mean differences and RMSEs (shown in panel titles) are also quite small for NN − CTL and comparable with those of CTL1 − CTL.

The examples of seasonal predictions show overall reasonable results. We realize that for practical implementation seasonal predictions should be produced in an ensemble mode (typically, including several tens of ensemble members), to reduce the impacts of internal variability and to estimate forecast uncertainty; however, this kind of testing is supposed to be done by an implementation group and goes beyond the scope of this study.

5. Discussion and conclusions

In this study, the NN emulation approach (Krasnopolsky et al. 2005a,b, 2008a) has been further developed and tested in the state-of-the-art, high-resolution, coupled NCEP CFS. The developed highly accurate neural network emulations of longwave (RRTMG-LW and RRTMF-LW) and shortwave (RRTMG-SW) radiation parameterizations are on average 16, 20, and 60 times faster than the original/control longwave and shortwave radiation parameterizations, respectively (see Table 1). The results were obtained in a sequential code-by-code comparison and adequately represent the situation when

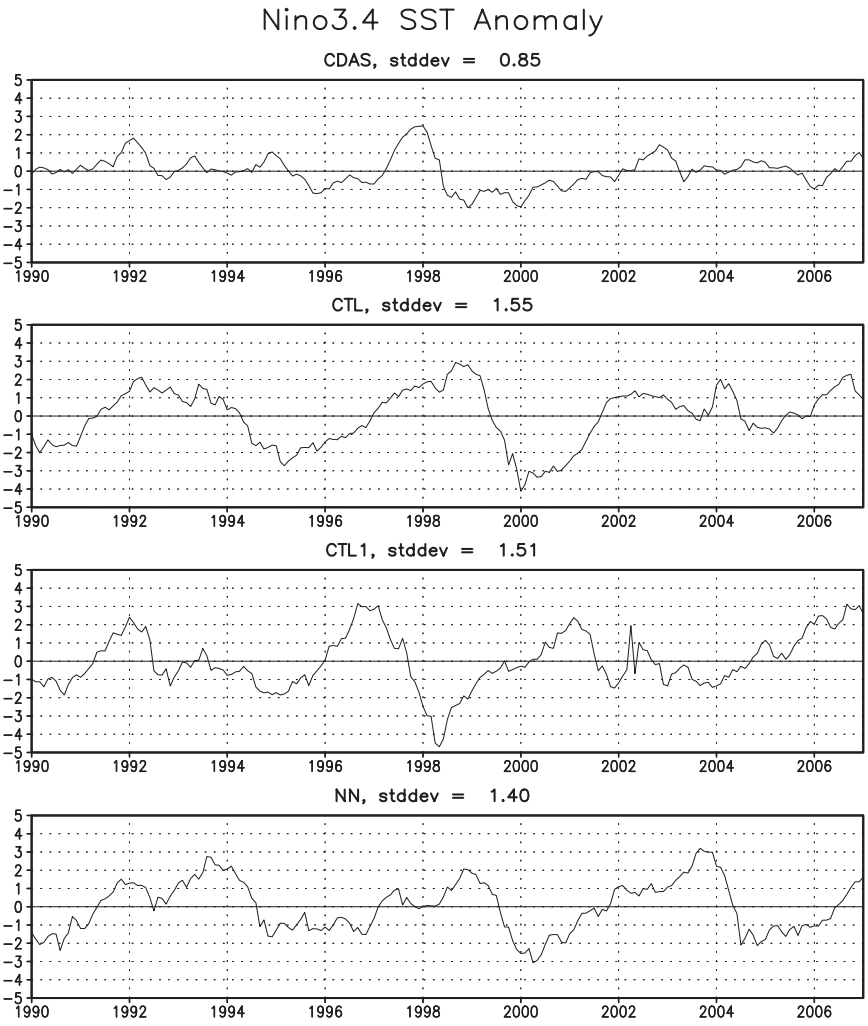

FIG. 9. The 17-yr (1990–2006) time series of the Niño-3.4 index for (top to bottom) CDAS, CTL, CTL1, and parallel full radiation NN. The Niño-3.4 index is calculated over the area in the Pacific Ocean shown by a rectangle in Figs. 3–7.

the model is run on a single processor; however, if we compare the control model run using the original parameterizations with the NN run when both runs use multiple processors and threads, the actual speedup will be significantly higher. In the case of parallel calculations utilizing multiple processors and threads, in addition to a significant speedup, the use of the NN emulations in the model provides an additional advantage; namely, it helps to achieve a significantly better load balance.

The use of the full NN model radiation provides: 1) an overall speedup of approximately 20%–25% for climate simulations and seasonal predictions and 2) an opportunity to increase significantly the frequency of radiation calculations (e.g., to calculate model radiation at every model dynamic time step) without increasing the total model calculation time.

The full NN model radiation was used for the 17-yr climate model simulation and seasonal prediction with the NCEP CFS that has T126 spectral horizontal and high vertical (64 layers) resolutions. We demonstrated the profound similarity for the parallel climate simulations, produced with NN emulations and original radiation (the control run). We have also shown that the model mean differences and RMSEs associated with NN emulations are quite comparable with internal variability of the model for 17-yr integrations and seasonal predictions, which justifies the possibility of using computationally efficient neural network emulations of full model radiation for decadal and longer climate simulations as well as seasonal predictions.

Comparisons with similar results (Krasnopolsky et al. 2008a) obtained for NCAR CAM presented in section 3f

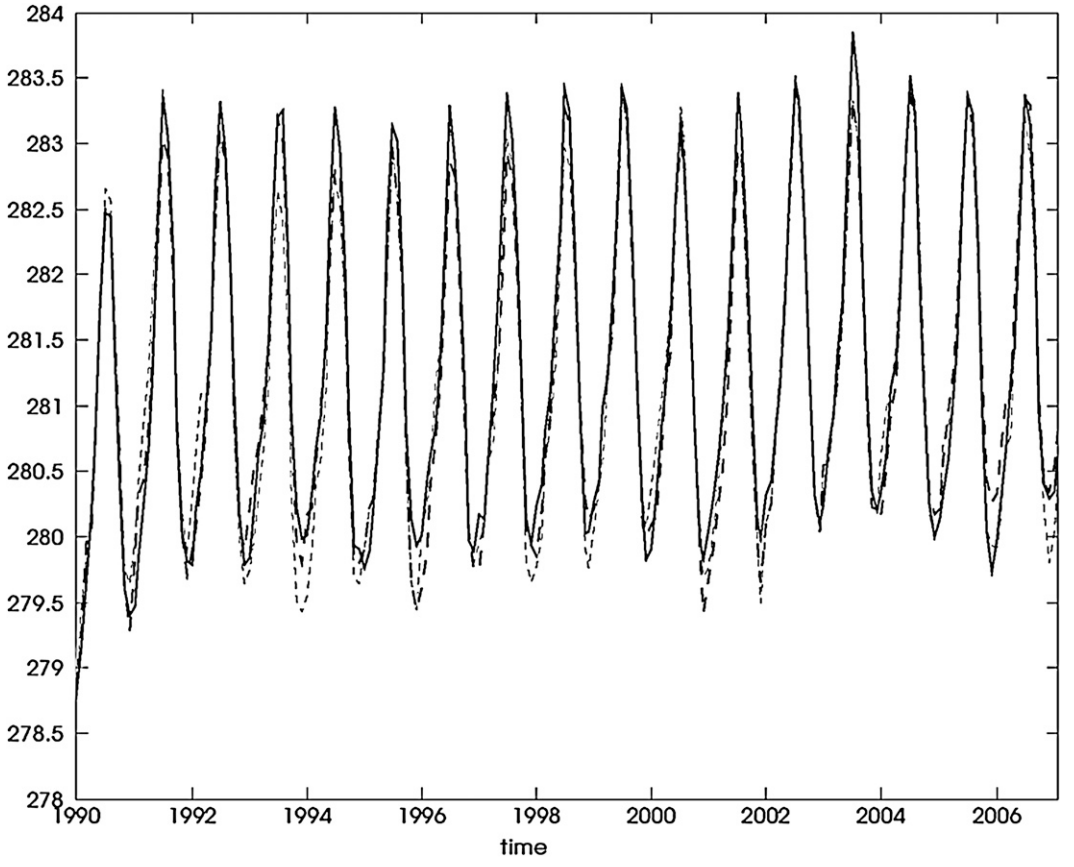

FIG. 10. The 17-yr (1990–2006) time series of global mean 850-hPa temperature (K) for the parallel full radiation NN (solid line), CTL (long-dashed line), and CTL1 (short-dashed line).

and Table 1 show that our NN emulation approach works for the high-resolution (T126L64) NCEP CFS as well as for the lower resolution (T42L26) NCAR CAM. The NN emulation approach has already been applied to both LWR and SWR parameterizations and tested in different models with different dynamical cores and with different resolutions (Krasnopolsky and Fox-Rabinovitz a,b; Krasnopolsky et al. 2008a). It is significantly less dependent (in terms of both the accuracy and speedup of calculations) on the increase of vertical resolution than the NN approach introduced by Chevallier et al. (1998) for developing a NN-based LWR parameterization NeuroFlux. At vertical resolution of 60 layers and more, both accuracy and rapidity of NeuroFlux cannot be achieved simultaneously (Morcrette et al. 2008). As we demonstrated in this study, our NN emulation approach can achieve simultaneously both the desired high accuracy and significant speedup at vertical resolution of 60 layers and more.

Applying the NN emulation approach, which allows us to achieve such a significant speedup with preservation of the accuracy and functional integrity of model physics, may create 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 obtaining larger errors is high dimensionality n of the input space of the mapping [(1)], which reaches several hundreds for NCEP CFS and may reach thousands for future models with significantly higher vertical resolution. It is difficult 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 belonging to these underrepresented far corners of the domain 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) and/or using the NN ensemble approach with NN emulations (Krasnopolsky 2007b).

The compound parameterization technique can also be used as a method of enriching the training dataset

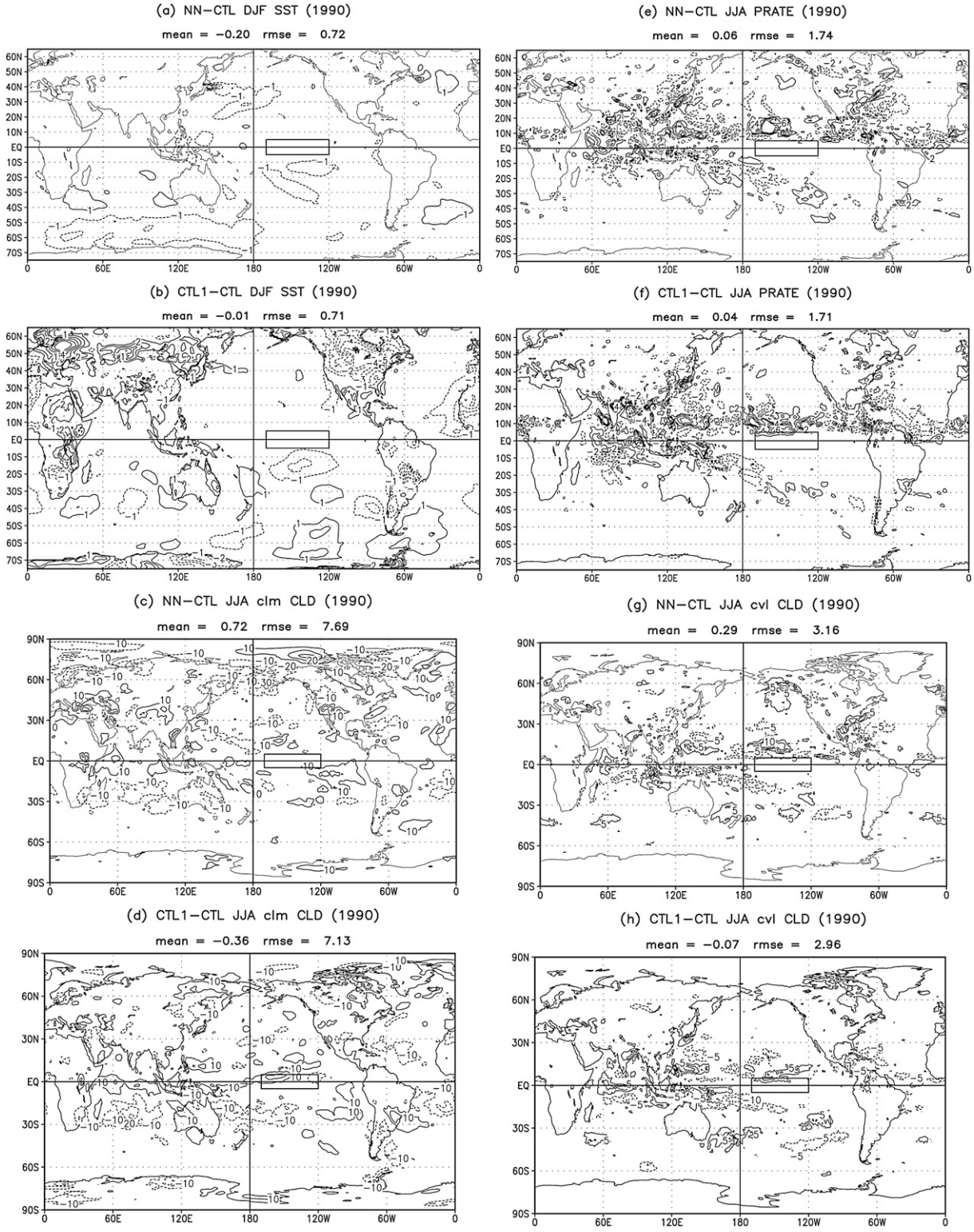

FIG. 11. The NN − CTL and CTL1 − CTL for seasonal predictions for 1990 for: (a),(b) DJF SST, (c),(d) JJA clm CLD, (e),(f) JJA PRATE, and (g),(h) JJA cvl CLD. The contour interval for SST fields is 1 K, for clm CLD is 10%, for PRATE is 2 mm day^{-1}, and for cvl CLD is 5%.

by inclusion of underrepresented atmospheric states (Krasnopolsky et al. 2008b). Taking into account the ability of NN to be adjusted sequentially; that is, using one record of the training set per time, the compound parameterization may be used as a constituent of a dynamically adjustable NN emulation. Such a dynamically adjustable NN emulation will be retrained online every time when recognizable changes happen in the environment and an unusual atmospheric state is pinpointed by the compound parameterization. For example, the NN emulations presented in this paper have been trained for the CO_2 level changes in the range of several tens of percent. Dynamically adjustable NN emulations will be able to perform better and under significantly higher changes of the CO_2 level.

Because model vertical resolution determines the NN emulation architecture, that is, the number of inputs and outputs, every time the vertical resolution of the model is changed (which is usually done quite rarely), the NN emulation needs to be retrained. It is noteworthy that NN retraining can be done routinely and takes a very limited time and effort once the practical framework for a specific model is developed.

In some applications of the developed NN emulation (in a data assimilation system or for an error and sensitivity analysis) not only NN emulation but also its first derivatives (NN Jacobian) are used. High accuracy of 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 model chemistry and other components.

Acknowledgments. The authors thank Drs. H.-L. Pan, S. Saha, S. Moorthi, and M. Iredell for a valuable help with practical use of NCEP CFS and for useful discussions and consultations. We also thank Drs. S. Moorthi and G. White for reading and commenting on the manuscript. This study is based upon the work supported by the NOAA/CDEP/CTB Grant NA06OAR4310047.

APPENDIX A

Balancing NN Radiation Outputs

There exist an integral relationship that relates pressure, heating rates, and fluxes. For example, this relationship for the imbalance ε can be written as,

$$\varepsilon = \frac{\sum_{k=1}^{L} \alpha_k h_k}{\sum_{k=1}^{L} \alpha_k} + \frac{\Phi}{\sum_{k=1}^{L} \alpha_k} = 0, \quad \Phi = \begin{cases} F_{\text{tup}} - F_{\text{sup}} + F_{\text{sdn}} & \text{for LWR} \\ F_{\text{tup}} - F_{\text{tdn}} - F_{\text{sup}} + F_{\text{sdn}} & \text{for SWR} \end{cases}, \tag{A1}$$

where $\alpha_k = (p_k - p_{k-1})G^{-1}$, p_k is pressure at a vertical level k, G is a constant, F_{tup} is the total sky outgoing LWR or SWR flux at the top of the atmosphere, F_{tdn} is the total sky downward SWR flux at the top of the atmosphere, F_{sup} is the total sky upward LWR or SWR flux at the surface, and F_{sdn} is the total sky downward LWR or SWR flux at the surface.

The outputs of original radiation parameterizations satisfy the relationship in (A1) with high accuracy, because these relationships are explicitly (or implicitly) included into the parameterizations. The outputs of the NN emulations will obviously satisfy (A1) only approximately; that is, in this case $\varepsilon \neq 0$; however, ε is small. For example, for the RRTMG-LW NN emulation presented in Table 1, mean value for ε is 6.5×10^{-4} K day^{-1}. A correction can be introduced for the heating rates via $\tilde{h}_k = h_k + \varepsilon$. The correction makes $\tilde{h}_k$ satisfy the relationship in (A1). This correction is very small and, as the results presented in Table 1 show, this balancing procedure does not practically affect the overall accuracy of LWR NN and marginally improves the overall accuracy of SWR NN.

APPENDIX B

Error Statistics Used in this Study

The statistics presented in Table 1 have been calculated as follows. The mean difference B (bias or systematic error of approximation) and the root-mean-square difference RMSE (a root-mean-square error of approximation)

between the original parameterization and its NN emulation describe the accuracy of the NN emulation integrated over the entire 4D (latitude, longitude, height, and time) dataset and are calculated as follows:

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

$$\mathrm{RMSE} = \left\{ (NL)^{-1} \sum_{i=1}^{N} \sum_{j=1}^{L} [Y(i, j) - Y_{\mathrm{NN}}(i, j)]^2 \right\}^{1/2}, \quad \text{(B1)}$$

where $Y(i, j)$ and $Y_{\mathrm{NN}}(i, j)$ are outputs from the original parameterization and its NN emulation, respectively; $i = 1, \ldots, N$ is the horizontal location of a vertical profile on a latitude–longitude grid; N is the number of horizontal grid points; $j = 1, \ldots, L$ is the vertical index; and L is the number of the vertical levels.

Using a minor modification of (B1), the bias and RMSE for the mth vertical level of the model can be calculated:

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

$$\mathrm{RMSE}_m = \left\{ N^{-1} \sum_{i=1}^{N} [Y(i, m) - Y_{\mathrm{NN}}(i, m)]^2 \right\}^{1/2}. \quad \text{(B2)}$$

The root-mean-square error profile shown in Fig. 1 was calculated as a combination of the level RMSE_m for $m = 1, \ldots, L$.

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

$$\mathrm{PRMSE}(i) = \left\{ L^{-1} \sum_{j=1}^{L} [Y(i, j) - Y_{\mathrm{NN}}(i, j)]^2 \right\}^{1/2}. \quad \text{(B3)}$$

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

$$\mathrm{PRMSE} = N^{-1} \sum_{i=1}^{N} \mathrm{PRMSE}(i)$$

$$\sigma_{\mathrm{PRMSE}} = \left\{ (N - 1)^{-1} \sum_{i=1}^{N} [\mathrm{PRMSE}(i) - \mathrm{PRMSE}]^2 \right\}^{1/2}. \quad \text{(B4)}$$

The statistics in (B4) and (B1) both describe the accuracy of the NN emulation integrated over the entire 4D dataset; however, because of a different order of integration it reveals different and complementary information about the accuracy of the NN emulations.

REFERENCES

Bishop, C. 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.

Chou, M.-D., and M. Suarez, 1999: A solar radiation parameterization for atmospheric studies. Tech. Rep. NASA/TM-1999-104606, Vol. 15, 40 pp.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, 2005: Atmospheric radiative transfermodeling: A summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transf.,* **91,** 233–244.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998a: The operational CMC-MRB global environmental multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.,* **126,** 1373–1395.

——, J.-G. Desmarais, S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998b: The operational CMC-MRB global environmental multiscale (GEM) model. Part II: Mesoscale results. *Mon. Wea. Rev.,* **126,** 1397–1418.

Hou, Y.-T., S. Moorthi, and K. A. Campana, 2002: Parameterization of solar radiation transfer in the NCEP models. NCEP/EMC Tech. Memo. 441, 46 pp.

Iacono, M. J., E. J. Mlawer, S. A. Clough, and J.-J. Morcrette, 2000: Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR community climate model, CCM3. *J. Geophys. Res.,* **105** (D11), 14 873–14 890.

Janiskova, M., J.-F. Mahfouf, J.-J. Morcrette, and F. Chevallier, 2002: Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation. *Quart. J. Roy. Meteor. Soc.,* **128,** 1505–1528.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–472.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–268.

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 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.

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

——, ——, and ——, 2005b: Reply. *Mon. Wea. Rev.,* **133,** 3724–3729.

——, ——, and A. A. Belochitski, 2008a: Decadal climate simulations using accurate and fast neural network emulation of full, longwave, and shortwave radiation. *Mon. Wea. Rev.,* **136,** 3683–3695.

——, ——, 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.

Lacis, A. A., and V. Oinas, 1991: A description of the correlated *k*-distribution method for modeling nongray gaseous absorption, thermal emission and multiple scattering in vertically inhomogeneous atmospheres. *J. Geophys. Res.,* **96D,** 9027–9063.

Manners, J., J.-C. Thelen, J. Petch, P. Hill, and J. M. Edwards, 2009: Two fast radiative transfer methods to improve the temporal sampling of clouds in NWP and climate models. *Quart. J. Roy. Meteor. Soc.,* **135,** 457–468.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.,* **102** (D14), 16 663–16 682.

Morcrette, J.-J., and Coauthors, 2007: Recent advances in radiation transfer parameterizations, Tech. Memo. 539, ECMWF, 52 pp.

——, G. Mozdzynski, and M. Leutbecher, 2008: A reduced radiation grid for the ECMWF Integrated Forecasting System. *Mon. Wea. Rev.,* **136,** 4760–4772.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate,* **19,** 3483–3517.

Sato, M., J. E. Hansen, M. P. McCormick, and J. B. Pollack, 1993: Stratospheric aerosol optical depths, 1850-1990. *J. Geophys. Res.,* **98D,** 22 987–22 994.

Venema, V., A. Schomburg, F. Ament, and C. Simmer, 2007: Two adaptive radiative transfer schemes for numerical weather prediction models. *Atmos. Chem. Phys.,* **7,** 5659–5674.

Washington, W. M., and D. L. Williamson, 1977: A description of NCAR GCM's in general circulation models of the atmospheres. *Methods in Computational Physics,* J. Chang, Ed., Academic Press, 111–172.