U. S. Department of Commerce
National Oceanic and Atmospheric Administration
National Weather Service
National Centers for Environmental Prediction

Office Note 471

Accurate and Fast Neural Network Emulations of Long and Short Wave Radiation for the NCEP Global Forecast System Model[1]

V. M. Krasnopolsky[1,2] (Vladimir.Krasnopolsky@noaa.gov),

A. A. Belochitski[1,2], Y. T. Hou[1], S. J. Lord[1], and F. Yang[1]

[1]National Centers for Environmental Prediction, NOAA, Camp Spring, MD 20746, USA

[2]Earth System Sciences Interdisciplinary Center, University of Maryland, College Park, MD 20740, USA

May 2012

[1] MMAB Contribution No 297.

Abstract

In this study we used the neural network (NN) emulation approach applied earlier to NCAR Community Atmospheric Model (CAM) and NCEP Climate Forecast System (CFS) radiation to develop NN emulations of the full NCEP Global Forecast System (GFS) model radiation. NN emulations have been developed and tested for the original RRTMG long-wave radiation (LWR) and RRTMG short wave radiation (SWR) parameterizations, which together comprise the full model radiation for the GFS model. The results presented in this note show that the developed NN radiation is very accurate. Also, it is 20 (LWR) and 100 (SWR) times faster than the original LWR and SWR parameterizations. The NN radiation was tested in several parallel 8-day forecasts. During first four days of integration no significant differences between control and NN runs can be observed. The differences observed after four days of integration are small and, after four days of integration, the NN run often demonstrates slightly better performance (higher anomaly correlation, lower bias and RMS errors) than the control run. Comparisons with an additional GFS run using CFS NN radiation demonstrate robustness of the developed NN radiation with respect to changes in the model environment. This is a very important practical result, which shows that the NN radiation does not require frequent updates and may work in the model for many years without retraining. The high speed of NN radiation calculations can be used to: (1) significantly speedup (15-18%) the model integration; (2) calculate radiation more frequently than ones per hour (actually it can be calculated at each integration time step); (3) emulate more advanced and time consuming radiation parameterization (e.g., the newest RRTM-McICA). Also, the NN radiation helps to achieve a significantly better load balance. This study is the first, initial step in evaluating NN radiation in GFS. Further steps will include: (1) more

comprehensive tests in a longer series of 10-day forecasts and (2) evaluation of the NN radiation in parallel runs with more frequent radiation calculations.

1. Introduction

Full (long and short wave) 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 numerical weather prediction (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 more CPU time than the rest of the model components, i.e., 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 (e.g. correlated-k method by Lacis and Oinas (1991)).

To reduce the cost further, calculations are usually made at lower temporal and/or spatial resolutions. Quite drastic reductions in temporal resolution are often made (e.g., radiation calculations are made every one or three hours for the climate and global forecast models at NCEP and UKMO (Manners et al. 2009)). Between radiative transfer calculations major changes may occur in the radiative profiles (caused primarily by two factors: changes in clouds and changes in the angle of incident solar radiation) that are not represented. A reduced horizontal resolution approach (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 ECMWF (Morcrette et al. 2007, 2008). Also a fast neural network (NN) based long wave radiation parameterization NeuroFlux (Chevallier et al. 1998, 2000) is used to speed up radiation calculations for the 4D-Var linearized physics (Janiskova et al., 2002) in ECMWF.

A reduced vertical resolution approach (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. Coˆte´ et al. 1998a, 1998b). 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.

The aforementioned situation served as an important motivation for developing new alternative numerical algorithms that deliver faster calculations of model physics components while carefully preserving their accuracy. An approach to accurate and fast calculation of model radiation, using NN emulations, has been previously proposed, developed and thoroughly tested at NCEP for Climate Forecast System (CFS) (Krasnopolsky et al. 2010) and for the National Center for Atmospheric Research Community Atmospheric Model (Krasnopolsky et al. 2005, 2008). In this study the NN approach has been implemented for the NCEP Global Forecast System (GFS) model. NCEP GFS is a spectral model with the spectral resolution of 574 spectral harmonics and 64 vertical levels (T574L64).

NN emulations of model physics are based on the fact that any parameterization of physics can be considered as a continuous or almost continuous mapping (output vector, *Y*, vs. input vector, *X*, dependence) $Y = M(X)$, and NNs are a generic tool for approximation of such mappings (Krasnopolsky 2007). NN is an analytical approximation,

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

where x_i and y_q are components of the input and output vectors X and Y, respectively, a and b are fitting parameters, and $\phi(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i)$ 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 terms or neurons in (1).

NN emulations approximate the functional dependence between inputs and outputs of a parameterization. They learn this functional dependence during the NN training utilizing a training data set which was simulated using the original parameterization. In Section 2 we describe the data that we used, our NN approach, and the NN validation on independent data set. In Section 3 the results of validation of the radiation in GFS parallel runs are presented. Section 4 contains conclusions and discussion.

2. GFS Radiation and its NN emulations

In this study we developed the NN emulations of the full GFS model radiation. Namely, the NN emulations have been developed and tested for the original long-wave radiation (LWR) and short wave radiation (SWR) parameterizations, which together comprise the full model radiation for the GFS model.

a. NCEP LWR and SWR

The RRTMG-LW is used in the GFS and CFS model as LWR. It is based on the AER's RRTM-LW v2.3 (Mlawer et al., 1997, Iacono et al., 2000). The SWR parameterization used in GFS and CFS is a modified version of AER's RRTMG-SW (v2.3) (Clough et al., 2005). Although both RRTMG-LW and RRTMG-SW are built with fast computation schemes designed for GCM applications, they still represent the most time-consuming part of model physics in the NCEP GFS model. The percentage of the total model computation time used by model physics and by radiation (LWR and SWR) vary largely depending 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 GFS configuration at the T574L64 resolution, with the RRTMG-LW and RRTMG-SW both called every hour, the portion of the radiation computation time is about 15-18% of the total model computation time.

b. Data for training and validation

The GFS model with the original LWR and SWR parameterizations have been used to simulate data for NN training and testing. The data set was composed of 24 ten-day forecasts, each started on the first and fifteenth day of the month over a period of one year (2010). All inputs and outputs from the original LWR and SWR parameterization were saved over the globe eight times

per day (every three hours), resulting in a total of 1920 global data sets generated from the NCEP GFS. About 300 data records have been randomly selected from each global data set. The collected set of about 600,000 input/output radiation vectors was divided into three independent parts, each containing about 200,000 input/output vector combinations (records). The first part of the dataset was used for training, the second for control (control of overfitting, control of a NN architecture, etc.), and the third for validation of the approximation quality of trained NNs only.

c. NN emulation architecture and training

In our previous works (Krasnopolsky et al. 2005, 2008), when developing NN emulations for the radiation parameterizations of NCAR CAM, we followed a straightforward approach in selecting the emulating NN architecture. Inputs and outputs of the emulating NN have been selected to be identical to the inputs and outputs of the radiation parameterization to be emulated. For CAM, which has 26 vertical layers, the LWR emulating NN has 220 inputs and 33 outputs and the SWR emulating NN has 451 inputs and 33 outputs. If we followed the same straightforward approach (the NN emulation has exactly the same inputs and outputs as the original parameterization to be emulated) for NCEP model, which has 64 vertical layers, the LWR emulating NN would have 585 inputs (7 profiles of 64 components each + 2 profiles of 65 components each + 6 scalar variables), and the SWR emulating NN would have 3,277 inputs (49 profiles of 64 components each + 2 profiles of 65 components each + 10 scalar variables). However, even for this straightforward approach, the number of relevant 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 have zero or constant values for the

Table 1. Inputs and outputs of LWR and SWR NNs developed for GFS

LWR NN Inputs			**SWR NN Inputs**	
NN input	**NN input ##**	**Levels**	**NN input ##**	**Levels**
$\cos\left(\frac{month \cdot \pi}{6}\right)$	1	-	1	-
$\sin\left(\frac{month \cdot \pi}{6}\right)$	2	-	2	-
$\cos(Lon)$	3	-	3	-
$\sin(Lon)$	4	-	4	-
Lat	5	-	5	-
Interface pressure	6:28	1,2:2:42,43	6:28	1,2:2:42,43
Interface temperature	29:64	1,2,3:2:61,62:65	29:64	1,2,3:2:61,62:65
Layer H_2O mixing ratio	65:104	1:40	65:104	1:40
Layer O_3 mixing ratio	105:138	31:64	105:138	31:64
Layer total cloud fraction	139:186	1:48	139:185	1:47
Surface emissivity	187	-	-	-
Cos of zenith angle	-	-	186	-
Surface albedo	-	-	187:190	1:4
Total Inputs	**187**		**190**	
LWR NN Outputs			**SWR NN Outputs**	
NN output	**NN output ##**	**Levels**	**NN output ##**	**Levels**
Layer Heating Rates	1:64	1:64	1:64	1:64
Total sky upward flux at toa	65	-	65	-
Clear sky upward flux at toa	66	-	66	-
Total sky downward flux at toa			67	-
Total sky upward flux at sfc	67	-	68	-
Clear sky upward flux at sfc	68	-	69	-
Total sky downward flux at sfc	69	-	70	-
Clear sky downward flux at sfc	70	-	71	-
Total sky downward uv-b flux at sfc	-	-	72	-
Clear sky downward uv-b flux at sfc	-	-	73	-
Total Outputs	**70**		**73**	

upper (e.g., water vapor) or lower (e.g., ozone) 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 should not be used for NN training. Constant inputs (zero or nonzero) do not contribute to the functional input/output relationship and should not be used as inputs and/or outputs for NN emulations. Moreover, if they were used, they would introduce an additional noise (an approximation error). Thus, 92 such constant inputs have been removed (see Table 1), which, in addition, significantly reduced the emulating NN dimensionalities.

Also, some input profiles contain a lot of redundancy that, if properly identified, can be used to reduce the input dimensionality (Krasnopolsky et al. 2009). Some profiles depend on the vertical coordinate very smoothly. For example, autocorrelation functions (ACF) for vertical profiles of several model variables are shown in Fig. 1. ACF of a profile shows the correlation between adjacent components of the profile (between values of the corresponding variables at the adjacent model levels). Slowly decreasing ACF (like those for pressure and temperature shown in black and red in Fig. 1) shows that the adjacent components of the profile are highly correlated and that redundant information is introduced if all of them are used as inputs for the emulating NN. For such profiles a sampling can be applied to reduce the redundancy and dimensionality of the NN input vector. For these profiles every other or even every third level can be selected as NN input. For some other profiles (e.g., cloud fraction shown in pink in Fig. 1) the corresponding ACFs decrease very quickly, which means that the redundancy for these variables is insignificant and the sampling should not be applied. In the case of LWR and SWR NN emulation, for the pressure and temperature profiles we applied samplings shown in Table 1.

This procedure allowed us to eliminate 50 redundant NN inputs without any significant reduction in the approximation accuracy (see Table 2).

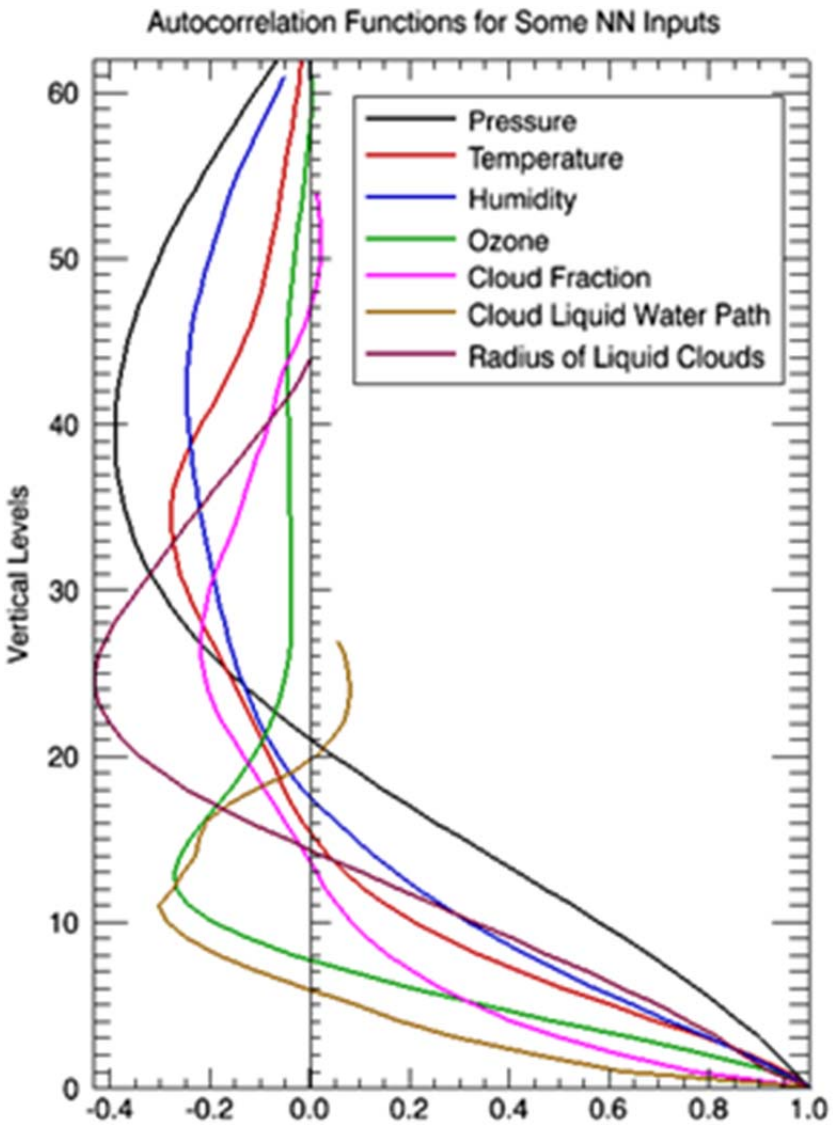

Fig. 1 Autocorrelation function for several NN input profiles. The horizontal axis shows the correlation and the vertical axis – the lag in vertical levels. Curves have different length for different input parameter profiles because the profiles have different number of nonzero components.

In addition, for SWR, 2688 inputs describing the optical depth, 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 $\tau = \frac{2 \cdot \pi}{T} \cdot q$, where q is the month of the year, and $T = 12$. Such a substitution is possible because, in NCEP CFS and GFS aerosol model, aerosols are calculated using the specific humidity profiles and 3-D lookup tables composed of climatological monthly data, different for different months of the year. It means that in terms of functional input/output dependences, the aerosol characteristics are the functions of *lat*, *lon*, τ, and the 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. Thus, *the SWR emulating NN emulates the aerosol model and SWR.*

In the current work, we generalized this approach to reduce even more the size of the SWR and LWR emulating NNs. Both SWR and LWR use, in addition to the cloud fraction profile, four other cloud characteristic profiles: layer cloud liquid water path (LCLWP), layer mean effective radius for liquid droplet (LMERLD), layer cloud ice water path (LCIWP), and mean effective radius for ice cloud (MERIC). These four profiles are calculated in microphysics block using models (equations) that use the specific humidity and atmospheric temperature profiles (Moorthi et al. 2001). Since the profiles of the specific humidity and atmospheric temperature have been already included as inputs in NNs emulating SWR and LWR, and the four aforementioned profiles are correlated with the cloud fraction profile, the emulating NNs are capable of emulating the part of microphysics that calculate these four additional profiles also. Thus, we excluded these four profiles from SWR and LWR NN inputs (totally 256 inputs have been eliminated). As a result, the developed LWR emulating NN emulates actually, in addition to

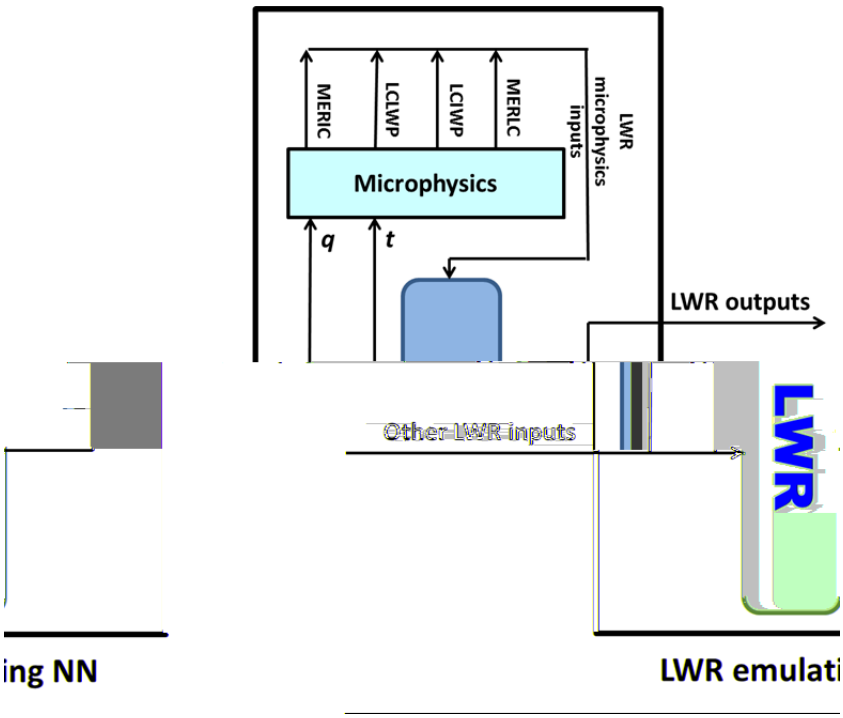

Fig. 2 The LWR emulating NN, which emulates LWR and a part of microphysics that calculates LCLWP, LMERLD, LCIWP, and MERIC.

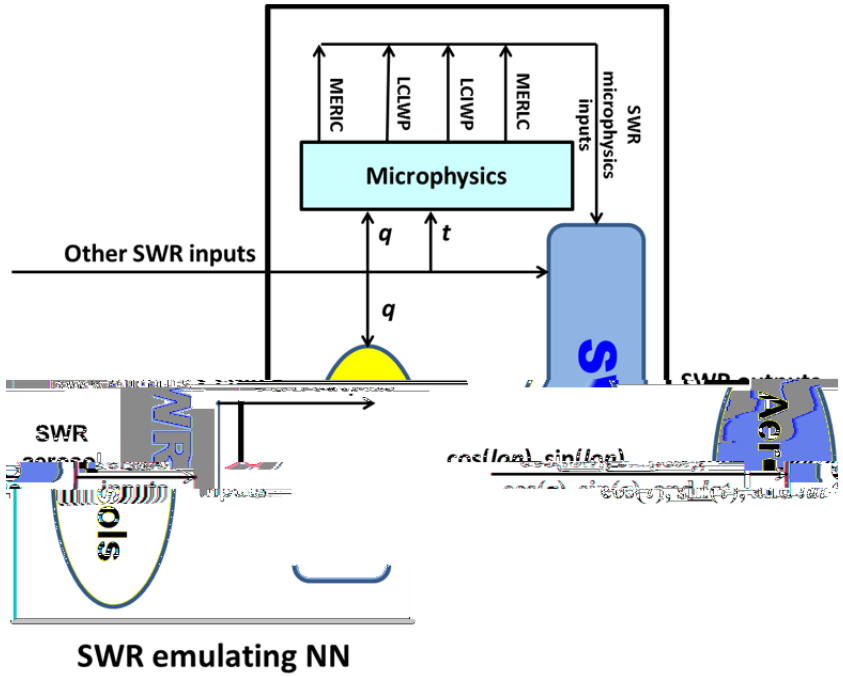

Fig. 3 The SWR emulating NN, which emulates SWR, a part of microphysics that calculates LCLWP, LMERLD, LCIWP, and MERIC, and the aerosol model.

LWR parameterization, also the cloud microphysics calculations of LCLWP, LMERLD, LCIWP, and MERIC (see Fig. 2). The developed SWR NN emulation emulates actually, in addition to SWR parameterization, also LCLWP, LMERLD, LCIWP, MERIC cloud microphysics calculations, and aerosol model (see Fig. 3). As the result, all these inputs of the original LW and SW radiation parameterizations are not included as inputs of the NN emulations of SW and LW radiation parameterizations and, therefore, are not presented in Table 1. Thus, the NN emulations developed for GFS radiations have the number of inputs 187 (LWR) and 190 (SWR) and the number of outputs 70 and 73 respectively (see Table 1). It is noteworthy, that our

Table 2. Statistics estimating the accuracy of heating rates (in K/day) calculations and the computational performance for NCEP GFS (T574L64) LWR and SWR using NN emulation vs. the original parameterization. For comparison, NCEP CFS (T126L64) LWR and SWR statistics are also shown. Total statistics show the mean error or bias, RMSE (A1), PRMSE, and σ_{PRMSE} (A4) for the entire 3-D HR fields. Layer (for the top and bottom layers) statistics show the mean error and RMSE (A2) for one horizontal layer (the top or bottom layer). The NN complexity NC (2) and average speedup η are also shown.

	Statistics	GFS RRTMG-LW	CFS RRTMG-LW	GFS RRTMG-SW	CFS RRTMG-SW
Total Error Statistics	**Bias,** *eq. (A1)*	$8. \cdot 10^{-3}$	$2. \cdot 10^{-3}$	$-7. \cdot 10^{-3}$	$5. \cdot 10^{-3}$
	RMSE, *eq. (A1)*	0.52	0.49	0.26	0.20
	PRMSE, *eq. (A1)*	0.38	0.39	0.18	0.16
	σ_{PRMSE}**,** *eq. (1)*	0.36	0.31	0.19	0.12
Bottom Layer Error Statistics	**Bias,** *eq. (1)*	$2. \cdot 10^{-2}$	$-1. \cdot 10^{-2}$	$-3. \cdot 10^{-2}$	$9. \cdot 10^{-3}$
	RMSE, *eq. (1)*	0.55	0.64	0.20	0.22
Top Layer Error Statistics	**Bias,** *eq. (1)*	$5. \cdot 10^{-2}$	$-9 \cdot 10^{-3}$	$-1. \cdot 10^{-3}$	$1. \cdot 10^{-2}$
	RMSE, *eq. (1)*	0.13	0.18	0.13	0.21
NN Complexity	N_C *See eq. (1)*	25,870	33,294	26,473	45,173
Speedup, η	***Times***	20	16	100	80

SWR and LWR NN emulations represent the cloudy-sky radiation as well as clear-sky radiation using the same eq. (1) (with weights different for SWR and for LWR); thus, the calculation of radiation under cloudy condition takes the same time as for the clear sky.

The NN emulation complexity, N_C, presented in Table 2 is calculated as,

$$N_C = k \cdot (n + m + 1) + m \quad (2)$$

where n and m are the numbers of inputs and outputs respectively, and k is the number of neurons in the emulating NN (1).

A number of NNs has been trained using the training set described above. The developed NN emulations use from 60 to 150 neurons in one hidden layer and have the same inputs and outputs presented in Table 1. Then bulk validation statistics for the accuracy of approximation and computational performance for the developed NNs emulations have been estimated using an independent data set. From the set of trained NNs, the LWR and SWR NNs with 100 hidden neurons have been selected for testing in GFS. The accuracy of the selected NN emulations has been estimated against the original GFS radiation parameterizations; the statistics are presented in Table 2. For these NN emulations, bias is negligible (about 10^{-3} K/day) and RMSE is limited (about 0.3 – 0.5 K/day). Obtaining very small NN emulation biases is important to ensure non-accumulating errors in the course of model integrations using NN emulations. The NN emulations developed for GFS are less complex and faster than those developed previously for CFS (Krasnopolsky et al 2010). However, as Table 2shows, they are as accurate as the CFS LWR and SWR NNs. The developed highly accurate NN emulations for LWR and SWR, in terms of code-by-code comparison at each model time step when LWR and SWR are calculated,

are about 20 and 100 times faster than the original/control NCEP GFS LWR and SWR respectively.

3. Validation in GFS

As the next step, the developed LWR and SWR NN emulations were validated in GFS model integrations. The LWR and SWR emulations with 100 neurons have been selected for an initial validation because they seem to be acceptable in terms of both their accuracy and minimal complexity (see the previous section). Several 8-day forecasts have been run using the GFS model (T574L64, 2011 version). Here we present results for three runs performed from August 1 to August 8, 2010. This period was selected for a validation run because there were several atmospheric events during this time:

1. Tropical storm Colin moved through the western Atlantic to Bermuda from August 2 to 8.

2. In the Eastern Pacific tropical storm Estelle moved from the coast of Mexico to about 400 miles southwest of Baja California form Aug 6 to 10.

3. Tropical storm Domeng occurred in the Western Pacific to the north east of the Philippines from August 3 to 5.

Three GFS (T574L64) runs have been performed in parallel:

1. control run labeled as PRNNCTL (black) in the figures, which uses the original radiation codes;
2. run labeled as PRNNFULLGFS (red) that uses LWR NN and SWR NN developed for GFS using GFS (T574L64) simulated data (see above); and
3. run labeled as PRNNFULLCFS (green) that uses LWR NN and SWR NN developed for CFS (Krasnopolsky et al 2010) using CFS data simulated by an old (T126L64, 2006) version of atmospheric model (see below).

It is noteworthy that NCEP CFS used for development and validation of CFS NN radiation (Krasnopolsky et al. 2010) incorporated the following: the NCEP GFS (version of 2006) atmospheric model with 126 spectral components and 64 vertical levels (T126L64) coupled with the 40-level interactive MOM4 ocean model, the interactive Noah land model with four soil levels with improved treatment of snow and frozen soil, an interactive sea ice model with fractional ice cover and depth allowed, a sub-grid scale mountain blocking, and a seasonal climatological aerosol treatment. The CFS NN radiation has been developed based on a data set accumulated from a 17-year CFS run. The GFS NN radiation presented in this study has been developed based on 24 ten-day forecasts produced by the high resolution T574L64 (version of 2011) GFS (uncoupled atmospheric part of CFS), which, in addition to much higher spectral resolution, incorporates many changes in physics and other model components as compared with the version of CFS, for which the CFS NN radiation has been developed. Thus, the CFS NN radiation has not been trained and validated for a significantly different model environment of the current version of GFS. In the third GFS run presented below we use this old CFS NN radiation in the new T574L64 GFS. Comparing results of this run with the control run and with the GFS run using new GFS NN radiation developed for the current version of GFS allows us to evaluate robustness of the NN emulation approach with respect to the changes in the model.

Figures 4 to 12 present various statistics (anomaly correlations, biases, and RMS errors) routinely used for evaluation and comparison of GFS runs. The comparisons of anomaly correlations, biases, and RMS errors have been performed for instantaneous model prognostic and diagnostic fields produced at each day of the 8-day forecasts. The NN radiation and control runs are very close in terms of calculated statistics.

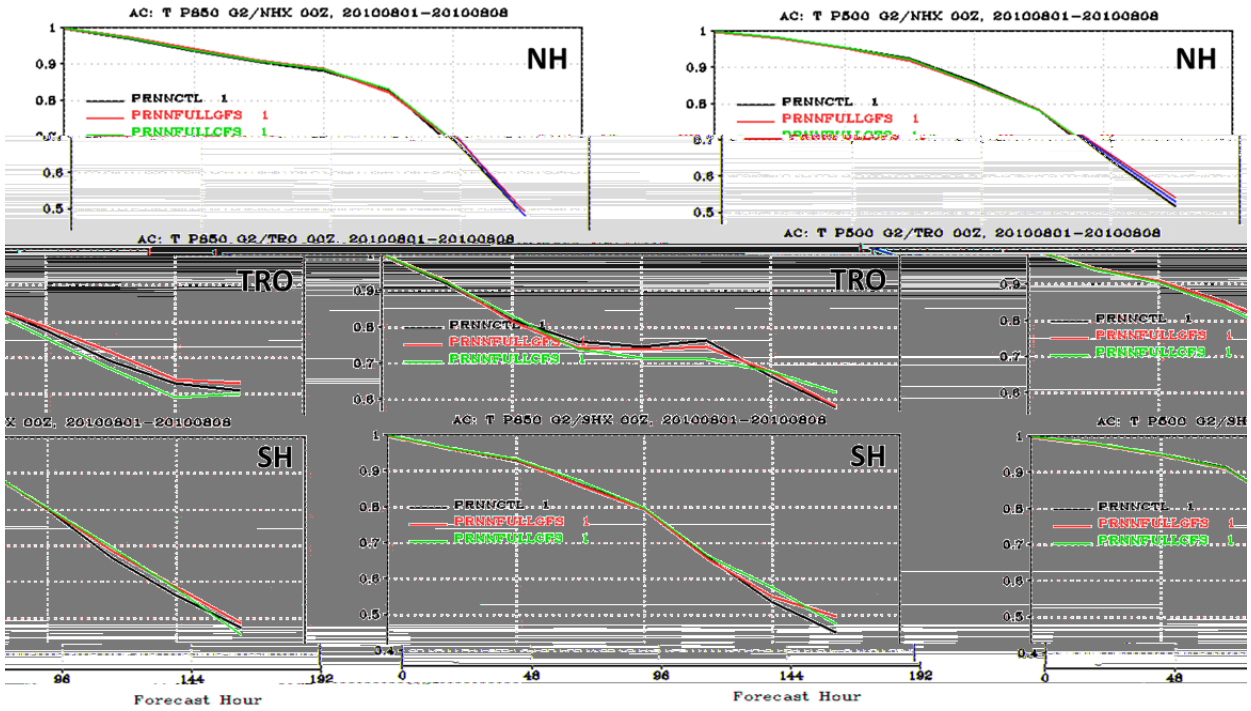

Fig. 4 Anomaly correlation at 850 mb (left column) and 500 mb (right column) for the northern hemisphere (upper row), tropics (medium row), and southern hemisphere (lower row) calculated for temperature fields. Three GFS (T574L64) runs are shown: black line – control run with the original LWR and SWR (PRNNCTL); green line – run with NN SWR and LWR developed for CFS (PRNNCFS); and red line – run with NN SWR and LWR developed for the current version of GFS (PRNNGFS).

For example, Fig. 4 shows the anomaly correlation calculated at 850 mb and 500 mb (left and right column respectively) for the temperature field. The upper and lower rows show results for the northern and southern hemispheres correspondingly. The middle row shows results for the tropics. Fig. 5 shows the anomaly correlation calculated for the geopotential height field for the 500 mb level and Fig. 6 shows the anomaly correlation calculated for the surface pressure field for the northern hemisphere (upper row), tropics (medium row), and southern hemisphere (lower row) correspondingly.

The differences between NN runs and control run increase from day one to day eight remaining small. As could be expected, the *PRNNGFS* run (red) closer follows the control run (black) than

the *PRNNCFS* run (green). The CFS LWR and SWR NNs have been transplanted into new significantly different version of GFS. Relatively small difference between the PRNNGFS (red) and PRNNCFS (green) runs demonstrates the robustness of the NN emulation approach with respect to changes in the model environment.

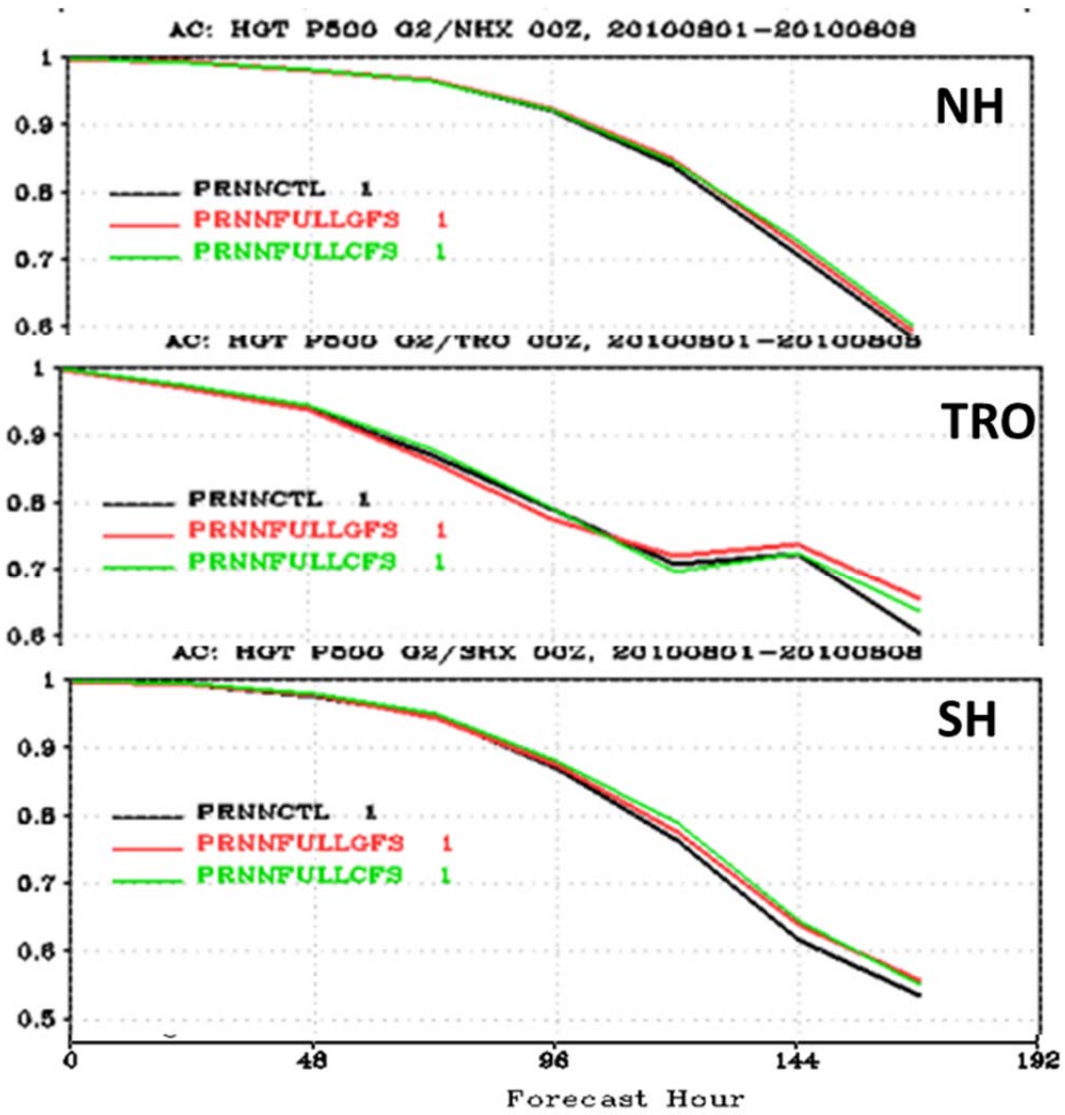

Fig. 5 Anomaly correlation at 500 mb for the northern hemisphere (upper row), tropics (medium row), and southern hemisphere (lower row) calculated for geopotential height fields. Three GFS (T574L64) runs are shown: black line – control run with the original LWR and SWR (PRNNCTL); green line – run with NN SWR and LWR developed for CFS (PRNNCFS); and red line – run with NN SWR and LWR developed for the current version of GFS (PRNNGFS).

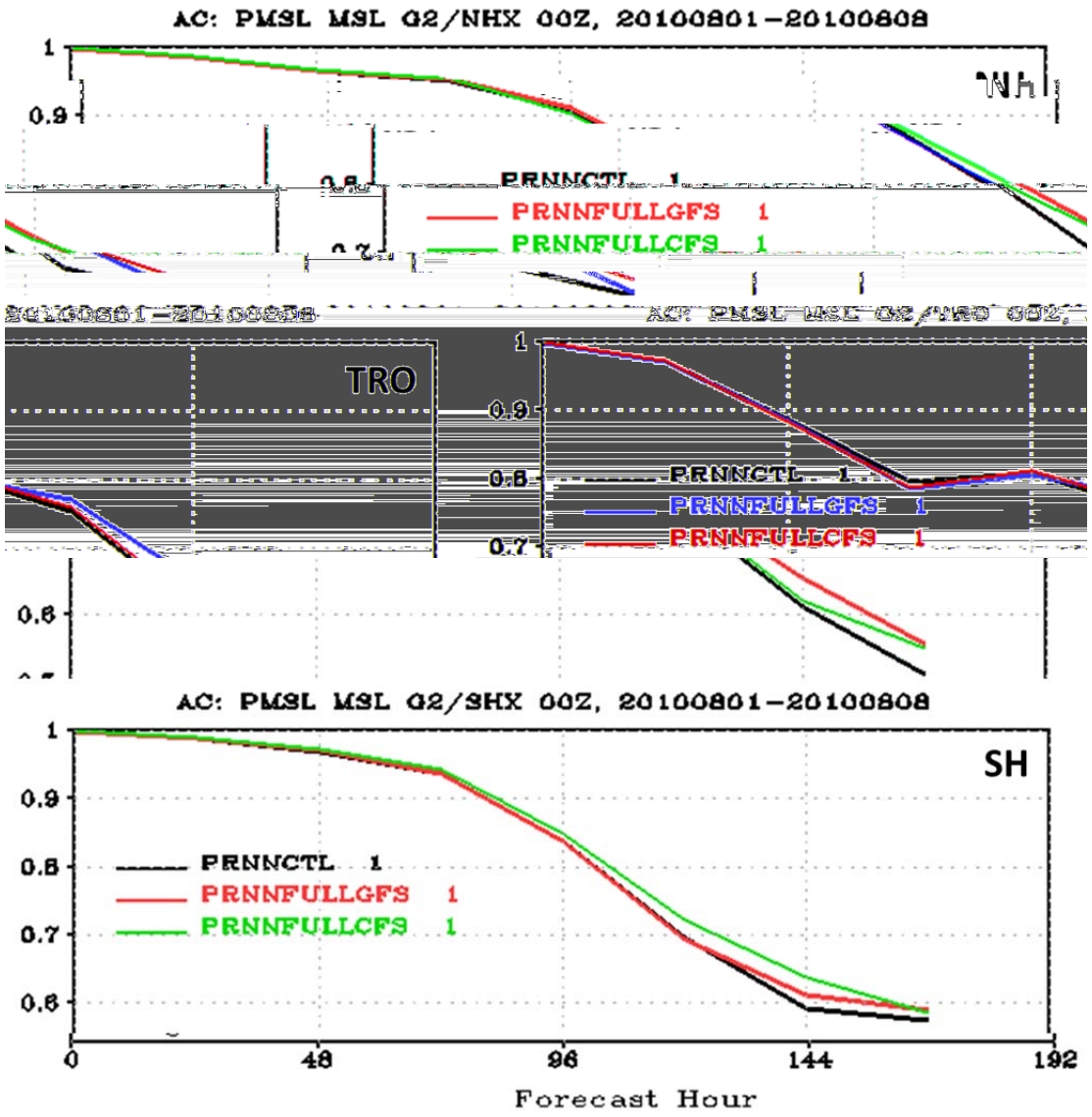

Fig.6 Anomaly correlation calculated for the surface layer pressure fields. See capture to Fig. 5.

It is noteworthy that after 80 to 100 hours of integration, the NN runs demonstrate slightly higher values of AC (see Figs. 4, 5, and 6) than the control run with the original radiation parameterizations. This interesting result should be further investigated in a longer series of 10-day forecasts. If confirmed, this result may indicate that NN emulations, in addition it their major function to speed up the radiation calculation, also effectively suppress high frequency noise present in the model.

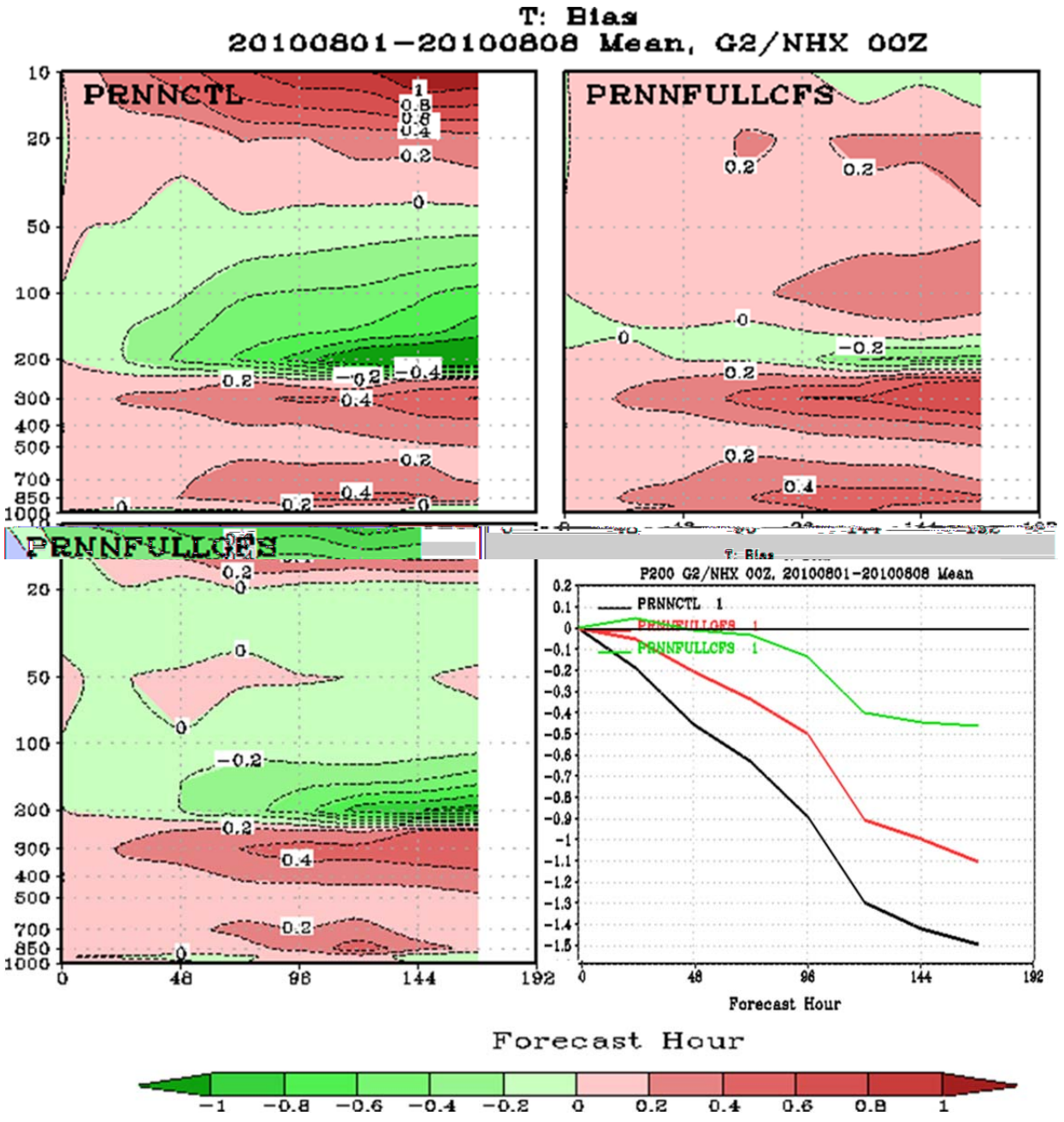

Fig.7 Temperature bias calculated for the northern hemisphere as a function of the forecast time (horizontal axis) and height in mb (vertical axis) for the control run, PRNNCTL (upper left), for PRNNCFS (upper right), and for PRNNGFS (lower left). Lower right panel shows the bias at 200 mb level: black line – control run (PRNNCTL); green line – PRNNCFS run; and red line – PRNNGFS run.

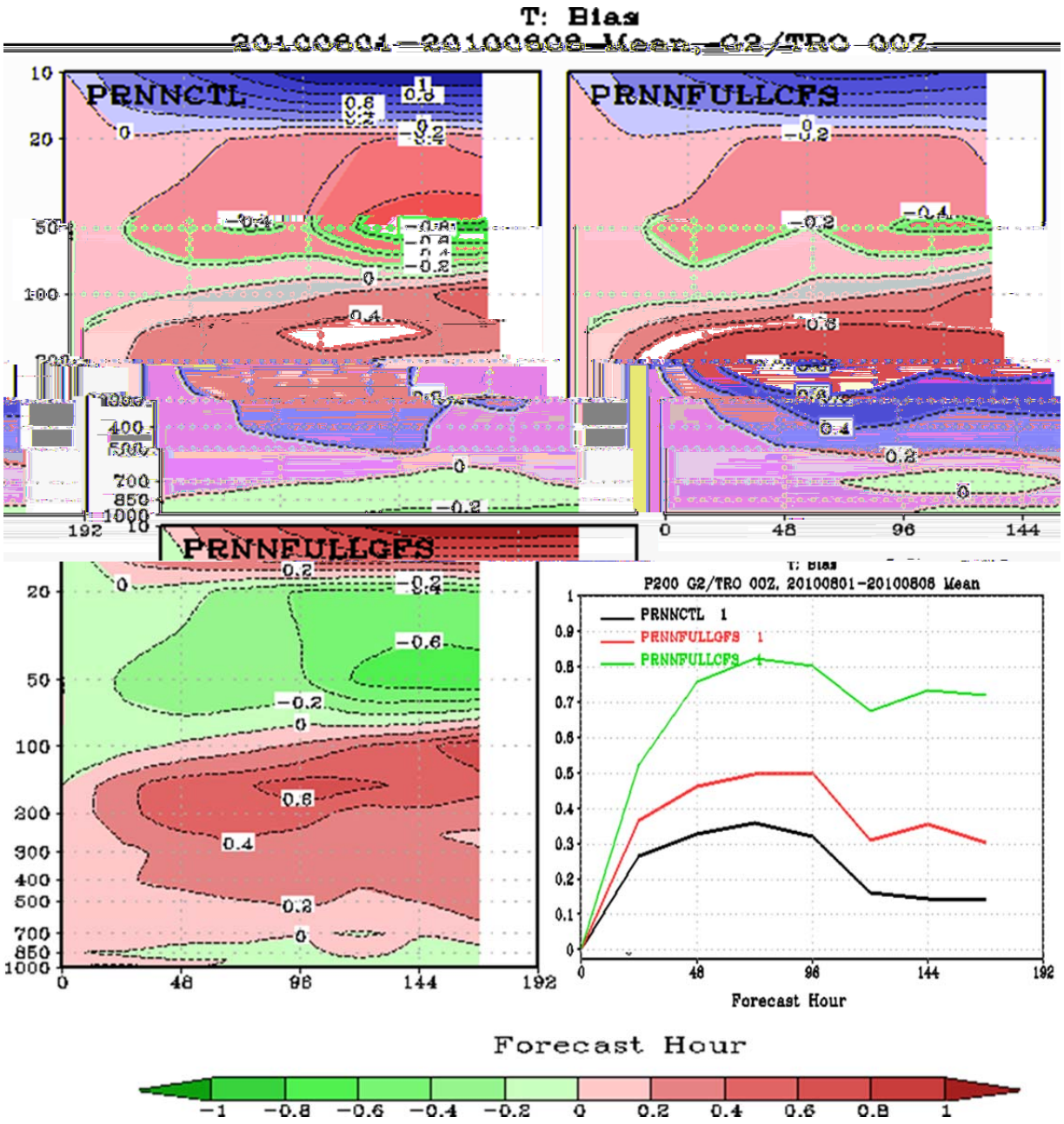

Fig.8 The same as in Fig. 7 calculated for the tropics.

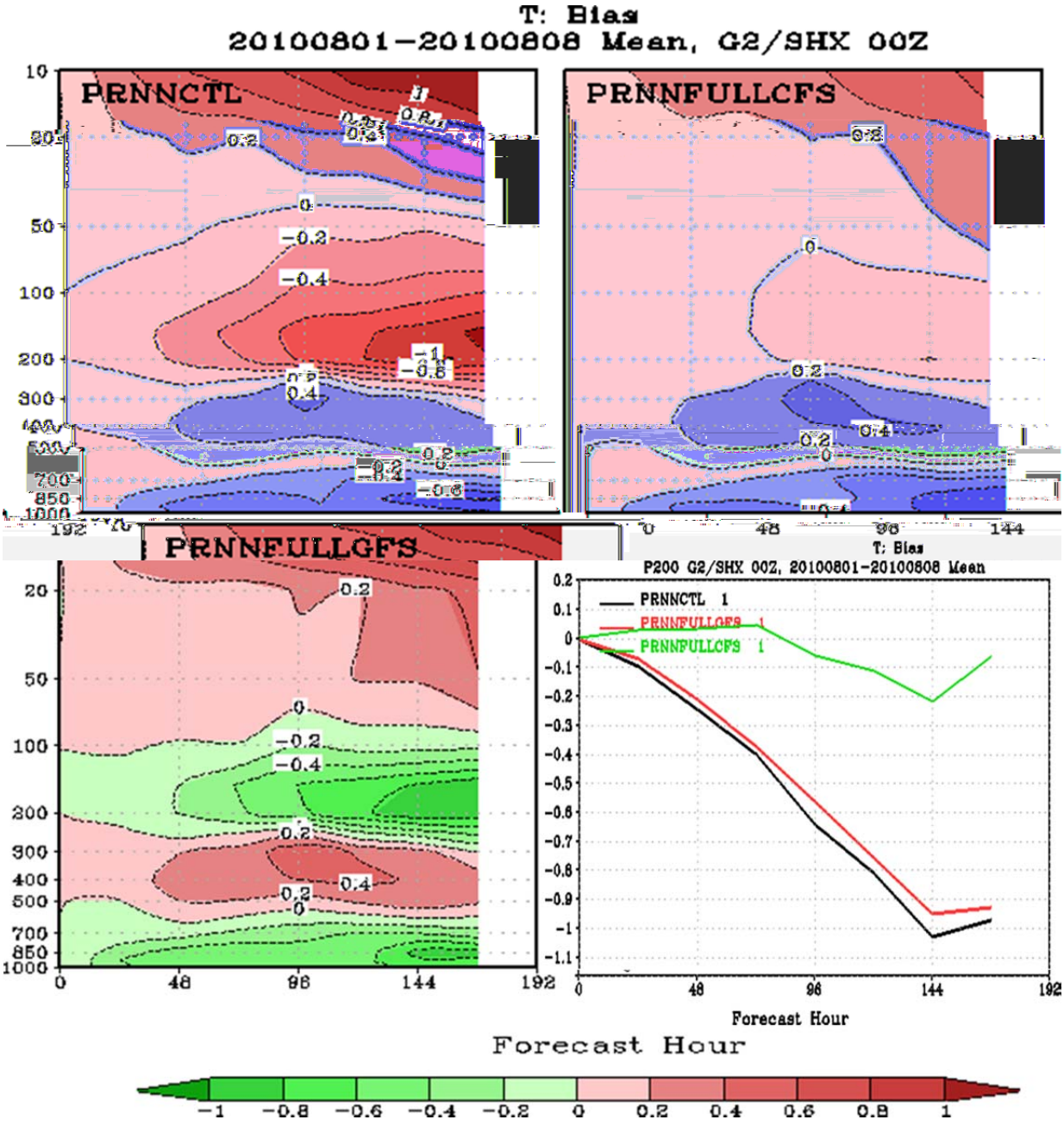

Fig.9 The same as in Fig. 7 calculated for the southern hemisphere.

Figs. 7 to 9 show the temperature bias calculated for the northern hemisphere (Fig. 7), for the tropics (Fig. 8), and the southern hemisphere (Fig. 9) as functions of the forecast time (horizontal axis) and height in *mb* (vertical axis). The upper left panel shows bias for the control run, the upper right – for *PRNNCFS* run, and the lower left – for *PRNNGFS* run. The lower right panel shows the time series of the three other panels at 200 mb.

Figs. 10 to 12 show the vector wind RMS errors and differences calculated for the northern hemisphere (Fig. 10), for the tropics (Fig. 11), and the southern hemisphere (Fig. 12) as functions of the forecast time (horizontal axis) and height in *mb* (vertical axis). The upper left panel shows the RMS error for the control run, the upper right – the RMS differences for (*PRNNCFS* – control) run, and the lower left – for (*PRNNGFS* – control) run. The lower right panel shows the time-series of the RMS errors for the three runs at 200 mb.

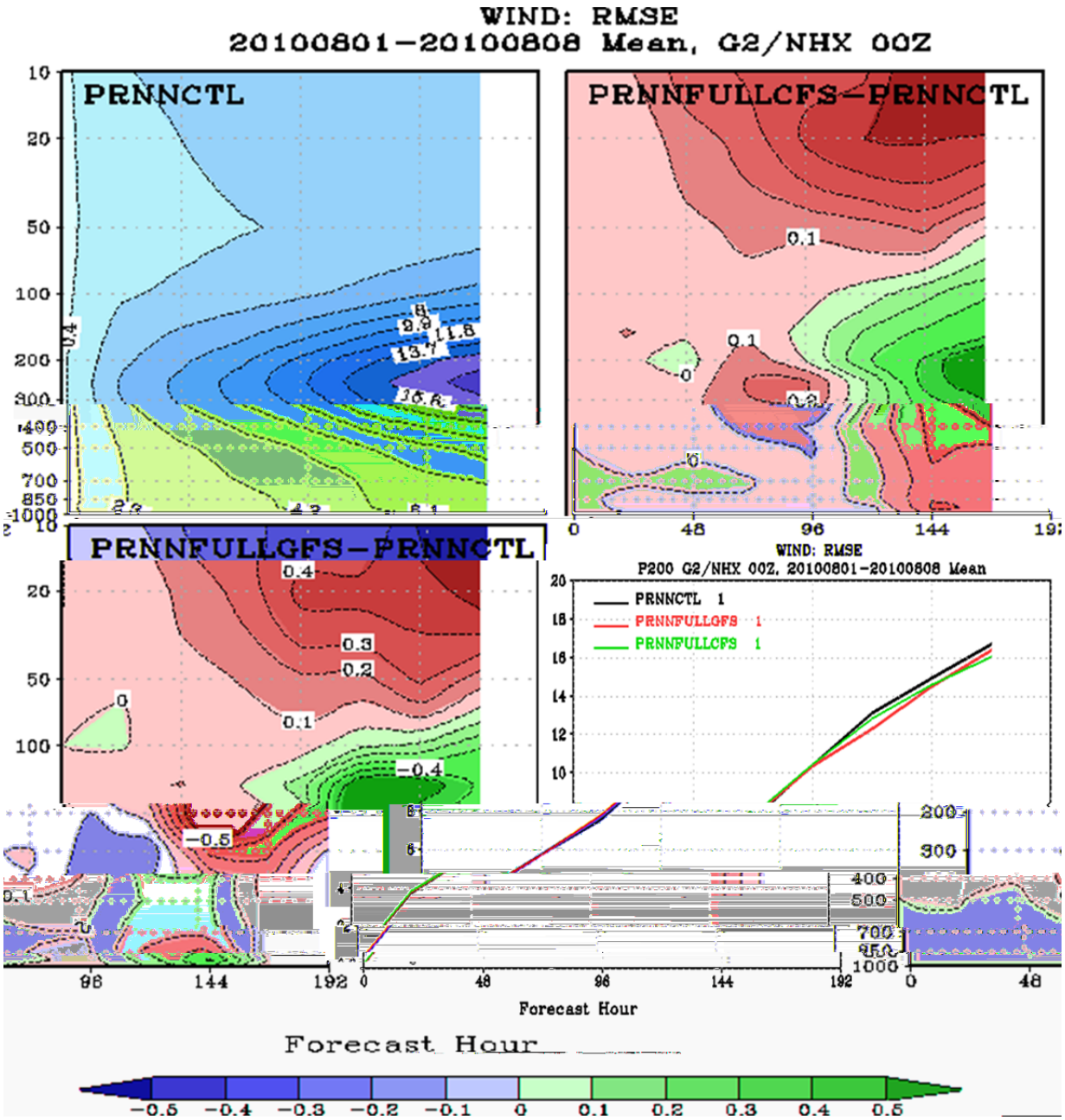

Fig.10 Vector wind RMSE calculated for the northern hemisphere as a function of the forecast time (horizontal axis) and height in mb (vertical axis) for the control run, PRNNCTL (upper left), for (PRNNCFS – PRNNCTL) (upper right), and for (PRNNGFS – PRNNCTL) (lower left). Lower right shows the RMSE at 200 mb level: black line – control run (PRNNCTL); green line – PRNNCFS run; and red line – PRNNGFS run.

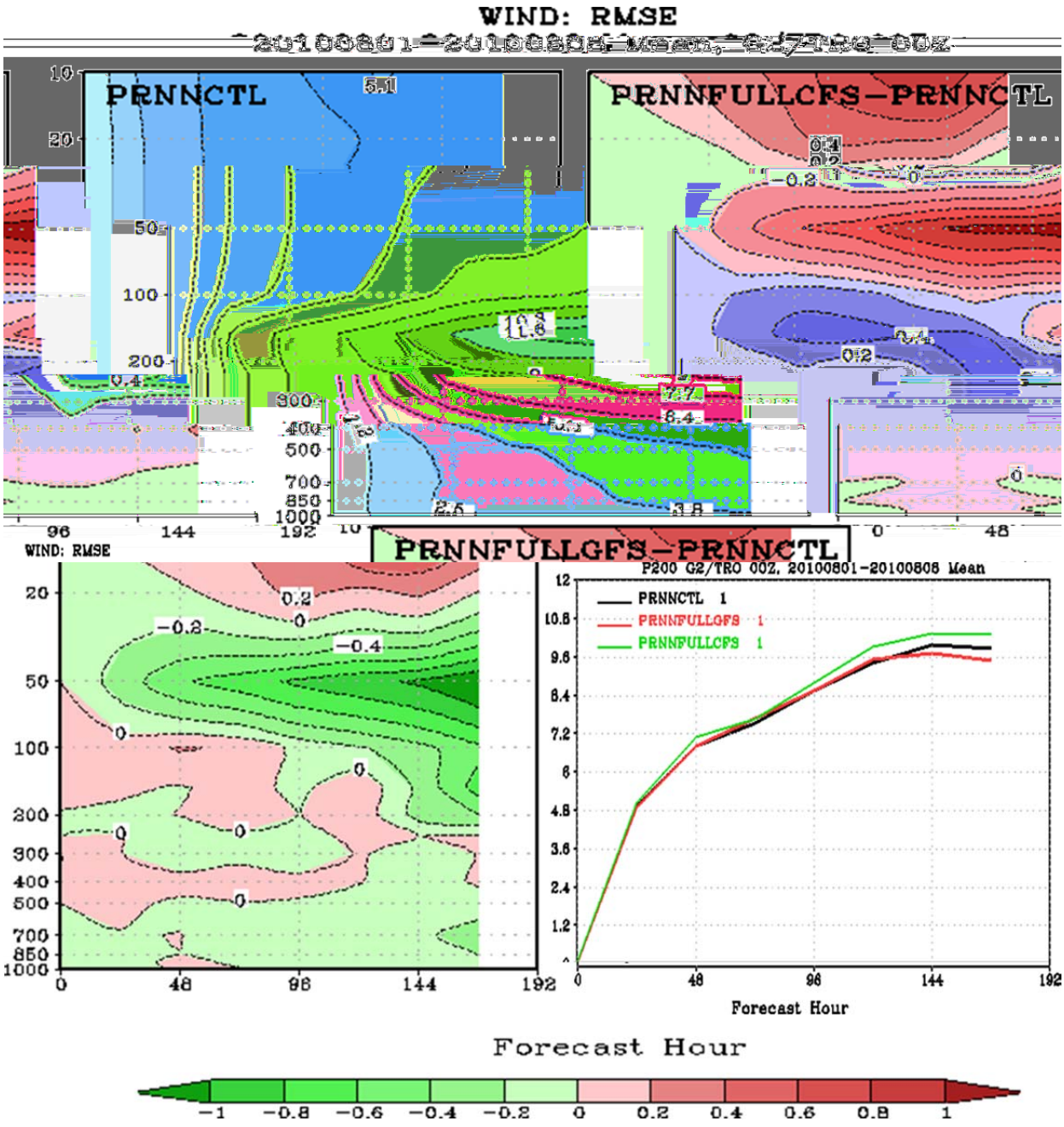

Fig.11 The same as in Fig. 10 calculated for the tropics.

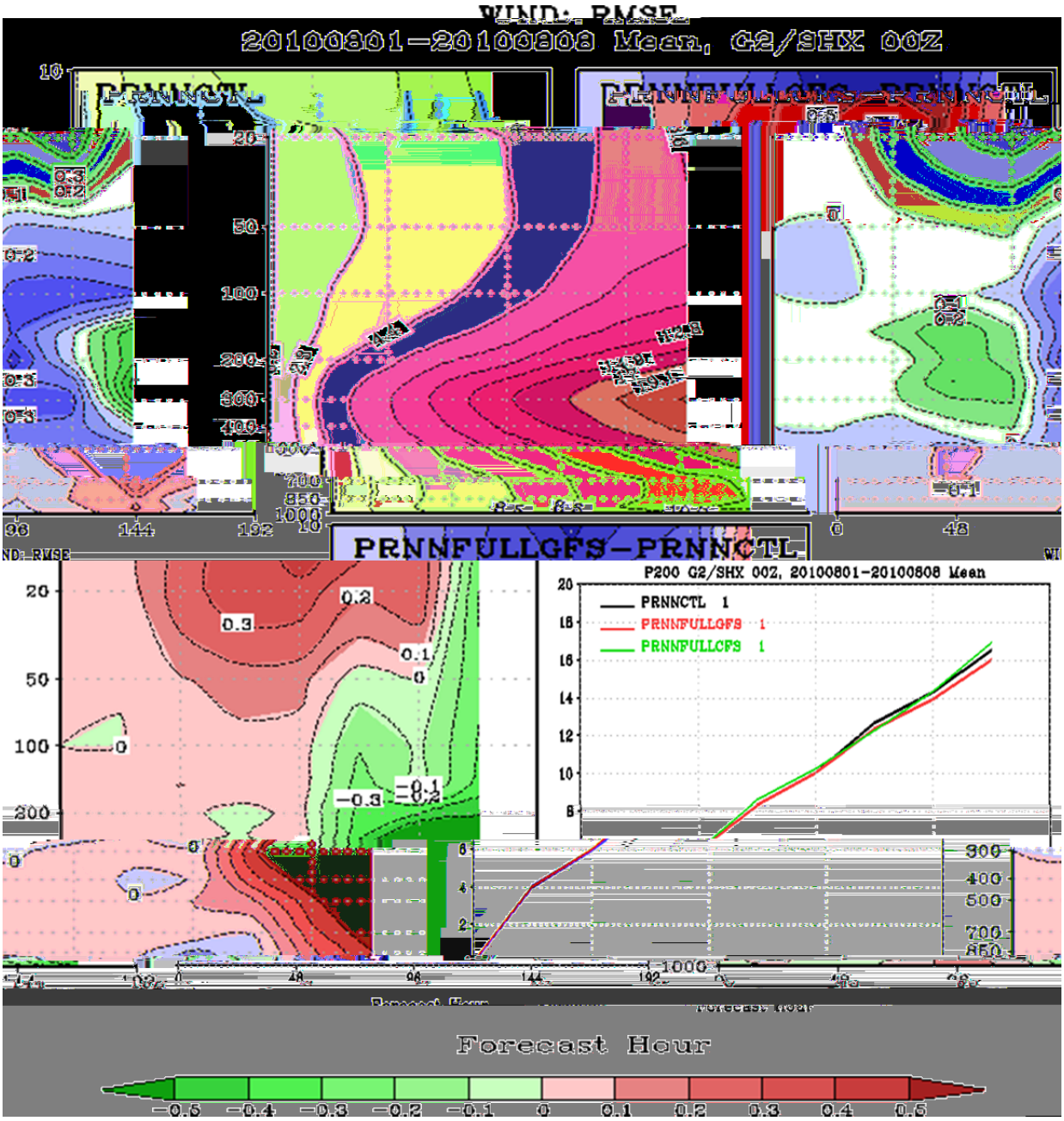

Fig.12 The same as in Fig. 7 calculated for the southern hemisphere.

4. Discussion and Conclusions

In this note we presented a new version of NN radiation developed for the latest version of NCEP GFS. The accuracy of the developed NN radiation has been evaluated using the independent validation data set (see Section 2) and in 8 day forecasts (see Section 3). In Section 3 we showed the results for one of performed 8 day forecasts. The results for other forecasts are very similar.

The results presented in this note show that **the developed NN radiation is very accurate**; the *PRNNGFS* run closely follows the control run, *PRNNCTL*. The differences between *PRNNGFS* and control runs are small; they increase slowly with the forecast time (Figs. 4 to 12); however, in many cases the NN radiation run *PRNNGFS* demonstrates slightly better results (higher anomaly correlation, lower bias and RMSE) at larger forecast times.

The additional *PRNNCFS* run allowed us to evaluate the robustness of the NN radiation and the NN emulation approach in general with respect to the changes in the model. The comparison of three runs (control, *PRNNGFS,* and *PRNNCFS*) demonstrated small differences between them, which shows **high level of robustness of the developed NN radiation with respect to changes in the model environment**. It shows that the developed NN radiation (CFS NN radiation) survived the transplantation from an old version of coupled model (CFS) to the newest version of uncoupled GFS. It also survived about 5 years of constant model evolution resulted in many changes in other than radiation physics parts of the model. After all these changes, the CFS NN radiation still produces reasonable results**. This is a very important practical result, which shows that the NN radiation does not require frequent updates and may work in the model**

for many years without retraining. Of course, when the original radiation parameterization or the vertical resolution of the model is changed, the NN radiation has to be retrained.

In addition to high accuracy, **the developed NN GFS radiation is very fast**. The high speed of NN radiation calculations can be used in several different ways:

1. The original radiation LW and SW parameterization can be simply substituted by the NN radiation and the NN radiation can be calculated in the model at the same frequency (ones per hour). This is the least efficient use of the NN radiation, which nevertheless provides a significant speedup of the total model integration of about 15-18%.
2. The GFS NN LWR is 20 times faster and NN SWR is 100 times faster than the original parameterization. This very significant speed up can be used to calculate radiation more frequently than ones per hour (actually it can be calculated at each integration time step). Such a run will take as much time as the current run (with ones per hour frequency of radiation calculations). In this case, the model run with the NN radiation would be many times faster than that with original parameterization calculated with the same frequency. Also in this case, in addition to a significant speedup, improvements in the quality of the forecast could be expected due to improvements in the radiation-cloud interaction.
3. The developed NN emulation approach can be used to emulate more advanced and time consuming radiation parameterization, which currently cannot be afforded in GFS. For example, NN emulation could be developed for the newest RRTM-McICA (Monte Carlo independent column approximation) radiation, the most sophisticated but slowest version of RRTM radiation. In this case, use of the NN emulation approach could lead to improvements in model radiation physics and in the quality of the forecast.

It is noteworthy that in addition to the speedup of radiation calculations, the use of the NN radiation provides an additional significant advantage as compared to the use of the original parameterizations, namely it helps to achieve a significantly better load balance (Krasnopolsky et al 2010). The radiatve transfer calculations take different time under different cloud conditions because of the different complexity of cloud-radiation interaction. For a more complex cloud-radiation interaction (deep convection) the calculation of the original LWR and SWR parameterizations takes ~22% and ~57% more time respectively than for clear sky conditions. Obviously, the time of the NN radiation calculations does not depend on the cloud conditions.

This study is the first, initial step in evaluating NN radiation in GFS. Further steps will include:

1. More comprehensive tests in a longer series of 10-day forecasts.
2. Evaluation of the NN radiation in parallel runs with more frequent radiation calculations.

Also refinement of NN emulations for the GFS model based on longer training set, implementation of the concept of a compound parameterization including a quality control procedure**,** and the NN ensemble approach (Krasnopolsky et al. 2007) will be introduced.

Appendix: 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 4-D (latitude, longitude, height, and time) data set and are calculated as follows:

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{L} [Y(i,j) - Y_{NN}(i,j)]^2}{N \times L}} \qquad \text{(A1)}$$

where $Y(i,j)$ and $Y_{NN}(i,j)$ are outputs from the original parameterization and its NN emulation, respectively, where i = *(latitude, longitude), i=1,…, 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 the vertical levels.

Using a minor modification of Eq. (A1), the bias and RMSE for the m^{th} vertical level of the model can be calculated:

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

$$RMSE_m = \sqrt{\frac{\sum_{i=1}^{N} [Y(i,m) - Y_{NN}(i,m)]^2}{N}} \qquad \text{(A2)}$$

The root mean square error profile can be calculated as a combination of the level $RMSE_m$ for $m = 1,\ldots, L$.

The root mean square error has been calculated for each i^{th} vertical profile:

$$prmse(i) = \sqrt{\frac{1}{L}\sum_{j=1}^{L}[Y(i,j) - Y_{NN}(i,j)]^2} \tag{A3}$$

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

$$\begin{aligned} PRMSE &= \frac{1}{N}\sum_{i=1}^{N} prmse(i) \\ \sigma_{PRMSE} &= \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}[prmse(i) - PRMSE]^2} \end{aligned} \tag{A4}$$

The statistics (A4) and (A1) both describe the accuracy of the NN emulation integrated over the entire 4-D data set. However, because of a different order of integration it reveals different and complementary information about the accuracy of the NN emulations.

Acknowledgements. The authors would like to thank Drs. B. Ferrier and H. Juang for the thorough relieving of the manuscript and helpful comments.

REFERENCES

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

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

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. Transfer, 91, 233-244, doi:10.1016/j.jqsrt.2004.05.058.

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

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

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, & F. Chevallier, 2002: Linearized radiation and cloud schemes in the ECMWF model: development and evaluation. Quarterly Journal of the Royal Meteorological Society, 128, 1505–1528.

Krasnopolsky, V.M., M.S. Fox-Rabinovitz, and D.V. Chalikov, 2005: “Fast and Accurate Neural Network Approximation of Long Wave Radiation in a Climate Model”, Monthly Weather Review, vol. 133, No. 5, pp. 1370-1383.

Krasnopolsky, V.M., 2007: “Neural Network Emulations for Complex Multidimensional Geophysical Mappings: Applications of Neural Network Techniques to Atmospheric and Oceanic Satellite Retrievals and Numerical Modeling”, Reviews of Geophysics, 45, RG3009, doi:10.1029/2006RG000200.

Krasnopolsky, V.M., M.S. Fox-Rabinovitz, A. A. Belochitski, 2008: "Decadal climate simulations using accurate and fast neural network emulation of full, long- and short wave, radiation.", *Monthly Weather Review,* 136*,* 3683–3695,

doi: 10.1175/2008MWR2385.1.

Krasnopolsky, V.M., S. J. Lord, S. Moorthi, and T. Spindler, 2009, "How to Deal with Inhomogeneous Outputs and High Dimensionality of Neural Network Emulations of Model Physics in Numerical Climate and Weather Prediction Models", Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, USA, June 14-19, 2009, pp.1668-1673

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

Lacis, A. A., V. Oinas, 1991: A description of the correlated k-distribution method for modeling non-gray gaseous absorption, thermal emission and multiple scattering in vertically inhomogeneous atmospheres. *J.Geophys. Res.* **96**: 9027.9063.

Manners, J., J.-C. Thelen, J. Petch, P. Hill & J.M. Edwards, 2009: Two fast radiative transfer methods to improve the temporal sampling of clouds in NWP and climate models, *Q. J. R. Meteorol. Soc.* **135**: 457 – 468; doi: 10.1002/qj.385

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., P. Bechtold, A. Beljaars, A. Benedetti, A. Bonet, F. Doblas-Reyes, J. Hague, M. Hamrud, J. Haseler, J.W. Kaiser, M. Leutbecher, G. Mozdzynski, M. Razinger, D. Salmond, S. Serrar, M. Suttie, A. Tompkins, A. Untch and A. Weisheimer, 2007: Recent advances in radiation transfer parameterizations, ECMWF Technical Memorandum No. 539, October 18, 2007

Morcrette, J.-J., G. Mozdzynski and M. Leutbecher, 2008: A reduced radiation grid for the ECMWF Integrated Forecasting System, *Monthly Weather Review*, 136, 4760-4772, doi: 10.1175/2008MWR2590.1

Moorthi S., Hua-Lu Pan and Peter Caplan, 2001: "Changes to the 2001 NCEP Operational MRF/AVN Global Analysis/Forecast System", Technical Procedures Bulletin Series No. 484, National Weather Service, Office of Meteorology.