Using Neural Network Emulations of Model Physics in Numerical Model Ensembles

Vladimir Krasnopolsky, Michael S. Fox-Rabinovitz, and Alexei Belochitski

Abstract — **In this paper the use of the neural network emulation technique, developed earlier by the authors, is investigated in application to ensembles of general circulation models used for the weather prediction and climate simulation. It is shown that the neural network emulation technique allows us: (1) to introduce fast versions of model physics (or components of model physics) that can speed up calculations of any type of ensemble up to 2 -3 times; (2) to conveniently an naturally introduce perturbations in the model physics (or a component of model physics) and to develop a fast versions of perturbed model physics (or fast perturbed components of model physics), and (3) to make the computation time for the entire ensemble (in the case of short term perturbed physics ensemble introduced in this paper) comparable with the computation time that is needed for a single model run.**

I. Introduction

During the last decade, ensemble techniques demonstrated a significant success in numerical weather prediction (NWP) [1,2] (Palmer et al. 2007, Buizza et al. 2005) and climate simulations [3-6](Broccoli et al. 2003, Murphy et al. 2004, Staniforth et al. 2005, Yoshimori et al. 2005). A traditional ensemble approach widely used in NWP consists of introducing perturbations in initial conditions because NWP problems (specifically, for short- to medium-term weather predictions) are the initial condition problems. Hereafter we will call this kind of ensembles the perturbed initial condition ensemble (PICE).

It was also found that, for both the NWP and climate applications, the spread of PICE forecasts is insufficient to systematically capture reality and perturbing of model physics has been proposed and introduced in some ensemble forecast systems [2,7] (Buiza et al. 1999, 2005). Climate simulation problems are rather boundary condition and right hand side (r.h.s.) forcing problems than initial condition problems. For this kind of problems, an ensemble approach based on perturbation of model physics (or perturbation of forcing) seems to be appropriate. The perturbed physics ensembles are expected to be more effective for climate simulations and projections [5] (Stainforth et al. 2005).

In this paper we investigate different possibilities of using the neural network (NN) emulation technique, introduced in [8,9] Krasnopolsky et al. (2002, 2005) for speeding up calculations of model physics, in combination with ensemble approaches. We discuss two types of perturbed physics ensembles: a long term perturbed physics ensemble (PPE) and a short term perturbed physics ensemble (STPPE). We also show that the NN emulation technique can be efficiently used to create PPE and STPPE. We show that all three aforementioned types of ensembles (PICE, PPE, and STPPE) can significantly benefit in terms of their numerical performance of using NN emulations of model physics; however, STPPE becomes especially efficient (orders of magnitude faster than PICE and PPE) when the NN technique is used to produce the ensemble of perturbed realizations of model physics.

In section 2, we briefly review the PICE and PPE techniques and introduce STPPE approach. In section 3 we describe the NN emulation of model physics technique and introduce fast NN emulation based ensemble versions for PICE, PPE, and STPPE. In section 4, we compare fast STPPE with PPE and PICE using NN emulations for model physics, specifically for long-wave radiation (LWR) of the NCAR CAM. Conclusions are given in section 5.

II. Ensemble Approaches in NWP and Climate Simulations

General circulation models (GCM) used for climate simulations and numerical NWP are complex nonlinear systems, which can be symbolically written as

$$\frac{\partial \psi}{\partial t} + D(\psi, t) = P(\psi, t)$$

and which are composed of many elements: initial conditions, $\psi(0)$, model dynamics, $D(\psi,t)$, model physics, $P(\psi,t) = \sum_k p_k(\psi,t)$ (p_k are parameterizations of physical processes), etc. Here ψ is the atmospheric state vector. Each of these elements as well as boundary conditions can be considered as a specific component that has its own internal (natural) uncertainty. Each of these components may be perturbed within its natural uncertainty to produce an ensemble of model realizations. Each of these ensemble realizations produces a prediction which constitutes an ensemble member.

This work was supported by the NOAA CDEP-CTB Grant NA06OAR4310047. MMAB Contribution No. 265

Vladimir Krasnopolsky is with SAIC at the National Centers for Environmental Prediction, Camp Springs, MD 20746 USA; Michael S. Fox-Rabinovitz, Vladimir Krasnopolsky, and Alexei Belochitski are with Earth System Science Interdisciplinary Center, University of Maryland, MD, USA; (the corresponding author: V. Krasnopolsky, phone: 301-763-8000 x 7262; fax: 301-763-8545; e-mail: Vladimir.Krasnopolsky@noaa.gov).

Formally, an ensemble forecast system may be represented as a set of numerical integrations,

$$\psi_j(T) = \psi_j(0) + \int_0^T [P_j(\psi_j, t) + D(\psi_j, t)]dt \qquad (1)$$

where j = 1,…, N is the number of the ensemble member. All ensemble members are close but different. The ensemble approach allows for integrating the specific information contained in the individual ensemble members into an ensemble that "knows" more/has more information about or represents the predicted climate or weather better than each of the individual ensemble members.

A. Ensembles with Perturbed Initial Conditions

Because NWP model integrations (specifically, for short- to medium-term weather predictions) are based on solving the initial condition problems, a traditional ensemble approach, PICE, widely used in NWP consists of introducing perturbations in initial conditions [10] (Buiza 1997); model physics is not perturbed and P_j are the same in eq. (1) for all ensemble members. Within this approach, each ensemble member run starts from uniquely perturbed initial condition $\psi_j(0)$. After running independently for some prescribed time T, the results of the ensemble member runs are compared with each other and with observations and averaged (see Fig. 1).

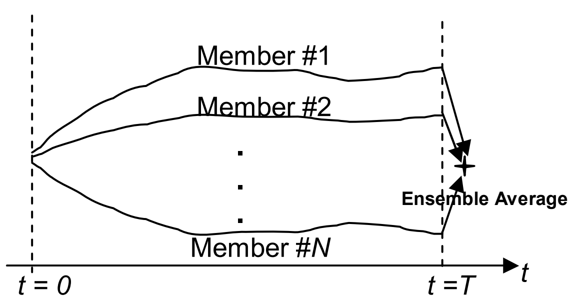

Fig. 1 The PICE and PPE scenario

Usually, the ensemble average describes better an actual weather or climate at the moment t = T than an individual ensemble member. PICEs allow us to observe how small uncertainties in initial conditions develop over model integration time into significant/measurable differences in predicted atmospheric states. For PICEs, the initial time step is the only time step for which an uncertainty is taken into account, i.e. perturbations is introduced in a deterministic NWP model. PICE proved to be an effective tool for NWP; however, it was noticed that the spread of PICE forecast is often insufficient for providing a systematic improvements of NWP (Buizza et al. 2005).

B. Ensembles with Perturbed Physics

It was also shown that, for the NWP and climate problems, a perturbed physics ensemble may provide a larger spread and better results [2,4,7,11] (Buiza et al. 1999, 2005; Stensrud 2000; Murphy et al. 2004). For example, ECMWF operational ensemble forecast system has been already augmented by including perturbed physics ensembles [2,7] (Buiza et al. 1999, 2005).

For climate models, which are not initial condition problems but rather boundary condition and r.h.s forcing problems, an ensemble generation approach based on perturbation of model physics (or perturbation of forcing) appears to be appropriate. Uncertainties in model physics that arise from the fact that the sub-grid effects are taken into account only approximately in model physics parameterizations, which include many uncertain parameters and approximations, have a different nature and spatial and temporal scales than uncertainties in initial conditions. In a sense, model physics parameterizations produce noise/perturbations at each GCM grid point at each time step of its integration. The perturbed physics ensembles (PPE) are shown to be very effective for climate simulations and projections [3-6,11,12] (Kharin and Zwiers 2000, Stensrud et al. 2000, Broccoli et al. 2003, Murphy et al. 2004, Stainforth et al. 2005, Yoshimori et al. 2005). Within this approach, each ensemble member uses a uniquely perturbed version of model physics P_j. PPE can also be used in combination with PICE [5] (Stainforth et al. 2005) as it is shown in eq. (1).

Several different approaches have been used for perturbing model physics:

- Model random errors associated with physical parametrizations are simulated by multiplying the total parametrized tendencies P by a random number r_j sampled from a uniform distribution between 0.5 and 1.5 ($P_j = r_j \times P$) [2,7] (Buiza et al. 1999, 2005).
- One or several model physics parameters controlling key physical characteristics of sub-grid scale atmospheric and surface processes can be perturbed one or several at a time within the scope of their natural uncertainty [4,5](Murphy et al. 2004, Stainforth et al. 2005).
- Different model physics process parameterization schemes can be used to create various versions of perturbed model physics; the different versions are used in different ensemble members [11] (Stensrud et al. 2000).

In section 3 of this paper a new method of generating an ensemble of perturbed physics is introduced that uses NN emulations of model physics [8,9] (Krasnopolsky et al, 2002, 2005) as a tool to create different realizations of model physics.

Usually the same scenario, as that depicted in Fig. 1 for the PICE with perturbed initial conditions, is followed for creating PPE. A particular GCM ensemble member uses a particular version of the perturbed physics, P_j, throughout the entire GCM run, for a long time T. Thus, in PPE different versions of perturbed physics (different realizations of the sub-grid physics) are used for different ensemble members, and each ensemble member exists and evolves over the entire GCM integration period T that is much longer than a characteristic time scale of sub-grid physical processes.

C. Short Term Ensembles with Perturbed Physics

However, using the perturbed physics approach for generating ensembles offers an opportunity to introduce an alternative ensemble approach, namely a new type of ensembles – a short term perturbed physics ensemble (STPPE) that is not possible in the framework of the traditional PICE approach. In the STPPE mode, the ensemble of different realizations (perturbed versions) of model physics is introduced for a time interval comparable with the time scales of the sub-grid processes, namely during one time step (or for some parameterizations for several time steps) of the model integration. Symbolically STPPE can be written as,

$$\psi(T) = \psi(0) + \int_0^T [\sum_{j=1}^{N} P_j(\psi,t) + D(\psi,t)]dt \tag{2}$$

At each time step, an ensemble of different realizations of model physics is generated and averaged. The ensemble average is used to integrate the model for producing the next time step. The STPPE scenario is shown in Fig. 2.

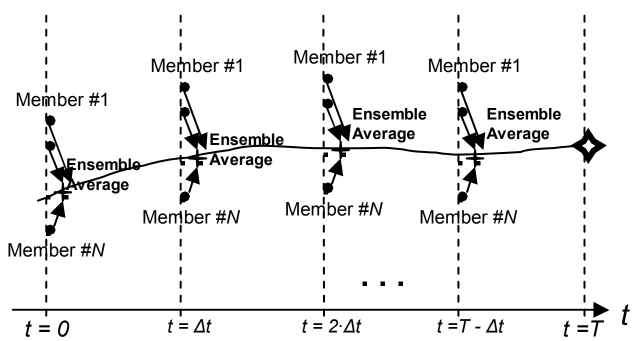

Fig. 2 The STPPE scenario.

The major differences between a PICE or PPE approaches (Fig. 1) and STPPE (Fig. 2) are:

- PICE and PPE consist of N independent model runs; STPPE consists of a single model run.
- In the PICE and PPE approaches, the ensemble averages for climate or weather characteristics are calculated at the end of all N model integrations combining climate or weather characteristics for all single ensemble member runs; within STPPE, the ensemble average is calculated at each integration time step, Δt, for the outputs of the ensemble members composed of perturbed components of model physics. The weather or climate characteristics obtained at the end are the results of this single STPPE run. There is no additional averaging of weather or climate characteristics in this approach.
- SLPPE may be significantly faster then PICE or PPE; if calculations of a perturbed version (or component) of model physics take about $\tau = \frac{1}{m}T$, where $1/m < 1$ is a fraction of T required for calculation of the model physics (or a particular component/parameterization of model physics that is perturbed), and T is the total time required for integration of one PICE member, then the STPPE run takes time

$$T_{STPPE} = [(1-\frac{1}{m}) + \frac{N}{m}] \cdot T \tag{3}$$

whereas PICE or PPE runs take a longer time

$$T_{PICE} = N \cdot T = N \cdot [(1-\frac{1}{m}) + \frac{1}{m}] \cdot T \tag{4}$$

The major technical difficulty in realization of all three discussed ensemble approaches (PICE, PPE, and STPPE) is their time consumption. Both PICE and PPE cost N (N – is the number of ensemble members) times more than a single model run; that is $N \cdot T$, where T is the time required for one GCM run. STPPE costs significantly less time because only model physics is calculated N times. For example, if the calculation of model physics takes 50% of the total model calculation time, STPPE will be about 2 times faster than PICE or PPE runs, assuming that the number of ensemble members is the same, N. If model physics calculation time is reduced the STPPE becomes even more computationally efficient. In the next section, we show that STPPE becomes very efficient (orders of magnitude faster than PICE and PPE) when the neural network (NN) technique is used to produce the ensemble of perturbed realizations of model physics.

III. Neural Network Ensembles with Perturbed Physics

In this section, we discuss NN emulations as a tool for introducing fast versions of model physics and a promising approach for perturbing model physics. The NN emulation technique allows us: (1) to introduce fast versions of model physics (or components of model physics) that can speed up calculations of any type of ensemble up to 2 -3 times; (2) to conveniently an naturally introduce perturbations in the model physics (or a component of model physics) and to develop a fast versions of perturbed model physics (or fast perturbed components of model physics), and (3) to make the computation time for the entire ensemble (STPPE) comparable with the computation time for one single model run. In this section, we use NCAR CAM as a particular example of GCM and its LWR parameterization as a particular example of a model physics component, which can be accelerated and perturbed using NN emulations, to describe the NN perturbed physics ensemble technique.

A. NN emulation technique

The NN emulation technique [8,9,13] (Krasnopolsky et al, 2002, 2005, Krasnopolsky and Fox-Rabinovitz 2006a) is based on the fact that the entire model physics as well as a single parameterizations of model physics may be considered mathematically as a continuous or almost continuous (like a step function) mapping between two vectors X (input vector) and Y (output vector) and symbolically can be written as:

$$Y = M(X); \quad X \in \Re^n, Y \in \Re^m \tag{5}$$

The simplest multi-layer perceptron (MLP) NN is a vector valued NN. It is composed of nonlinear neurons

$$z_j = \phi(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i)$$

Each element of the NN output vector is a linear combination of neurons

$$y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot z_j ;\ q = 1,2,\ldots,m \tag{6}$$

where k is the number of the neurons [14] (see Bishop, 2006 and references therein).

This NN has been used as a generic analytical nonlinear approximation or model for the mapping (5) [15,16] (Funahashi 1989, Krasnopolsky 2007a). Accurate NN emulations for the NCAR CAM long wave and short wave radiation (LWR and SWR) parameterizations have been developed [8,13,17] (Krasnopolsky et al. 2005, Krasnopolsky and Fox-Rabinovitz 2006 a, b). These NN emulations are 20 to 150 times faster that the original CAM SWR and LWR parameterizations correspondingly.

B. *Creating fast perturbed physics using NN emulations*

If we produce N perturbed versions of model physics adding some perturbations to the entire model physics or to one of its components (parameterizations), we can use these N perturbed versions to create PPE members following a traditional scenario of PICE (Fig. 1). These N versions can also be used as members of STPPE following an alternative scenario presented in Fig. 2.

The j^{th} perturbed version of the unperturbed model physics, P, can be written as,

$$P_j = P_j^{NN} = P + \varepsilon_j \tag{7}$$

where P_j^{NN} is a NN emulation number j of the original model physics, P, and ε_j is an emulation error for the NN emulation number j. As we have shown in our previous investigations [8,13,17] (Krasnopolsky et al. 2005, Krasnopolsky and Fox-Rabinovitz 2006 a, b), ε_j can be controlled and changed significantly varying k (the number of hidden neurons) in eq. (6). Not only the value but also the statistical properties of ε_j can be controlled. For example, the systematic components of the emulation errors (biases) can be made negligible (therefore, ε_j are purely random in this case). Thus, ε_j can be made of the order of magnitude of a natural uncertainty of the model physics (or of a particular parameterization) due to unaccounted variability of sub-grid processes (see also discussion in section 4).

Using NN emulations will speed up calculations of all three kinds (PICE, PPE, and STPPE) of ensembles. One PICE or PPE run with N ensemble members using N different NN emulations, each of which is n times faster than the original model physics, as perturbed versions of model physics will take time,

$$T_{PPE}^{NN} = N \cdot [(1 - \frac{1}{m}) + \frac{1}{m \cdot n}] \cdot T \tag{8}$$

Thus, in the case of NCAR CAM, where $m \approx 3/2$ to 2 and $n \approx 10$ to 100, using NNs for PICE or LTPPE will speed up its calculations about two to three times.

The acceleration of calculations of PICE and PPE due to the use of NN emulations of model physics is significant. However, the speed-up will be much more significant in the case of SLPPE. When we use N NN emulations each of which are n times faster than the original model physics, the STPPE run takes time

$$T_{STPPE}^{NN} = [(1 - \frac{1}{m}) + \frac{N}{m \cdot n}] \cdot T \tag{9}$$

It means that STPPE with $N = n$ ensemble members (N different NN emulations of model physics taken as ensemble members) can be run as fast as a single ensemble member of PICE or PPE (see eq. (4)).

Here, the legitimate question to ask is about the efficiency of STPPE. Does it improve the accuracy of a climate simulation to a degree at least comparable with improvements provided by the PICE and LTPPE approaches? This point is discussed in the next section.

IV. Comparisons of Different Ensembles using Perturbed NCAR CAM LWR

For validation of our experiments, we use the NCAR CAM run using the original model physics and the original NCAR CAM initial conditions as a control against which all ensemble members for all three considered types of the ensembles and ensembles themselves are estimated. In other terms, the climate obtained from the 15 years run of NCAR CAM with the original model physics (including original LWR parameterization) and original initial conditions is used below as a control climate. All ensemble members and ensemble averages for different ensembles (PICE, LTPPE and STPPE) are compared with these synthetic "observations". Than to create an ensemble of perturbed physics, we emulated the original LWR parameterization [18] (Collins et al. 2002) with six different NNs which approximate the original LWR parameterization with different limited approximation errors [8,13] (Krasnopolsky et al. 2005, Krasnopolsky and Fox-Rabinovitz 2006 a).

The perturbed LWR parameterizations can be written as,

$$LWR_j^{NN} = LWR + \varepsilon_j \tag{10}$$

where LWR is the original NCAR CAM LWR, LWR_j^{NN} is a NN emulation number j of the original NCAR CAM LWR, and ε_j is an emulation error for the NN emulation number j. Thus, the model physics that includes LWR NN emulation, LWR_j^{NN}, can be considered as perturbed versions of model physics, P_j.

There are many different approaches to creating different NN emulations of the same original parameterization (or NN emulation ensemble) [19] (Krasnopolsky 2007b). We have selected a sufficiently diverse group of six NN emulations mixing two different approaches to create an NN emulation ensemble. Five of these six ensemble members (NN emulations or realizations of LWR) have the same architecture, that is the same number of neurons ($k = 150$ in eq. (6)). However, these NNs are different because:

different initializations for the NN weights have been used to start the NN training; and the NNs have different weights (coefficients) and give slightly different approximations of LWR (i.e. realizations of LWR). The sixth NN emulation ensemble member has a different architecture (k = 90 neurons). In terms of the accuracy of the approximation there is a significant spread between the ensemble members. The approximation rms error varies from 0.28 to 0.40 K/day for the ensemble member NN emulations. It means that by using NN emulations instead of the original LWR parameterization, we introduced on average such a level of perturbation into the LWR model physics.

The distribution of approximation errors (perturbations) is shown in Fig. 3. It is obviously not normal. For the normal distribution with the same mean value and standard deviation, the perturbation values would be very limited; however, because the distribution of ε_j is not normal, there is a small but final probability of larger perturbations. If we compare this perturbations with mean value, μ and standard deviation, σ, of LWR itself (μ = -1.4 K/day and σ = 1.9 K/day), we will see that the majority of perturbations belong to the interval $\mu \pm \sigma$; however, a very small amount of perturbations reaches the magnitude of about $\mu \pm 3\sigma$. Such a distribution of perturbations is in a good agreement with the fact that the parameterizations of model physics on average describe the parametrized processes good enough and the level of errors introduced due to parameterization of sub-grid effects is rather moderate; however, in some cases (e.g., rare or extreme events) the errors may be very significant.

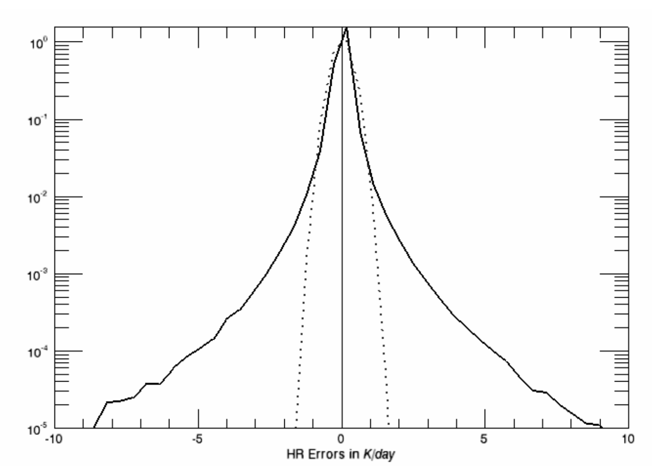

Fig. 3. Probability density function for ε_j. Mean ε_j = 3. 10^{-4} K/day and standard deviation of ε_j is 0.35 K/day. Dashed line shows a normal distribution with the same mean and standard deviation for comparison.

In the case of CAM LWR the NN emulations are about n = 100 times faster than the original LWR parameterization. Since calculation of the original CAM LWR takes about 30% of the model integration time T (m =3 in eqs. (3), (4), (8), and (9)), using LWR NN emulations in PICE and PPE speeds up calculations about 30%, reducing the time required for calculating CAM LWR $n \cdot m$ times. For SHPPE the use of NN emulations provides a much more significant speed up of calculations. Just as an example, a STPPE with N = 100 ensemble members (eq. (8)) runs as fast as a single ensemble member of PICE or PPE (eq. (4)).

Also, to run a PICE that is used for comparison purposes, we created six perturbed initial conditions members by perturbing original initial conditions used for the control run. Then we performed a PICE run (see Fig. 1); six climate simulations have been run with NCAR CAM for 15 years, each with one of these six perturbed initial conditions. Next we performed a LTPPE run (see Fig. 1); six climate simulations have been run with NCAR CAM for 15 years, each with one of the aforementioned six NN emulations (also used as the NN ensemble members for STPPE). The results (climate fields and diagnostics) of each simulation (ensemble member) were compared with the control climate run of NCAR CAM performed with the original LWR and original initial conditions. The climate simulation errors - systematic (bias), rmse, maximum (an extreme positive outlier), and minimum (an extreme negative outlier) - have been calculated for prognostic and diagnostic fields for each ensemble member vs. the control climate. These errors are shown by open circles (for PICE) and asterisks (for LTPPE) at Figs. 4 -6. Then the PICE and LTPPE averages were calculated (shown by filled circles and crosses respectively at Figs. 4 -6).

Next, one STPPE climate run has been performed. For this run, six aforementioned NN emulations were applied and the LWR outputs are calculated as the mean of these six NN emulation outputs, at each time step and at each grid point throughout the entire model integration. The results (mean climate and diagnostics) of these simulations (the SLPPE average/mean) are compared with those of the control run and shown by diamonds at Figs. 4 -9.

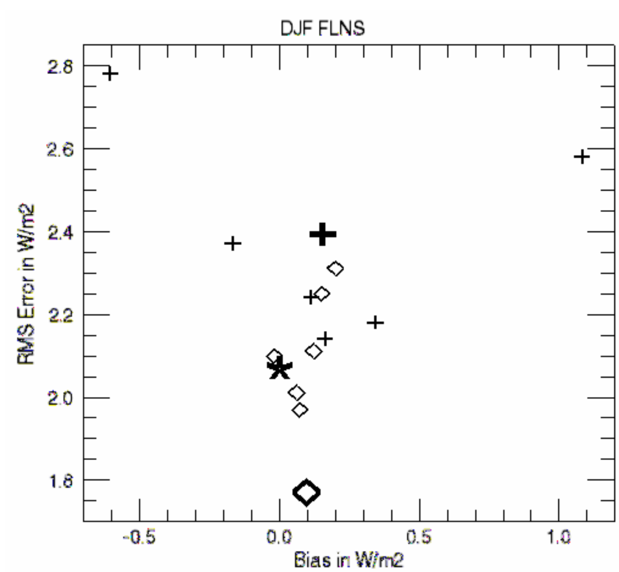

Fig. 4. Mean error (bias) and RMSE for the winter DJF (December through February) surface net LWR flux (FLNS). Diamonds show PICE members, thick large diamond – PICE average; crosses show LTPPE members, thick large cross – LTPPE average; thick large star shows STPPE value.

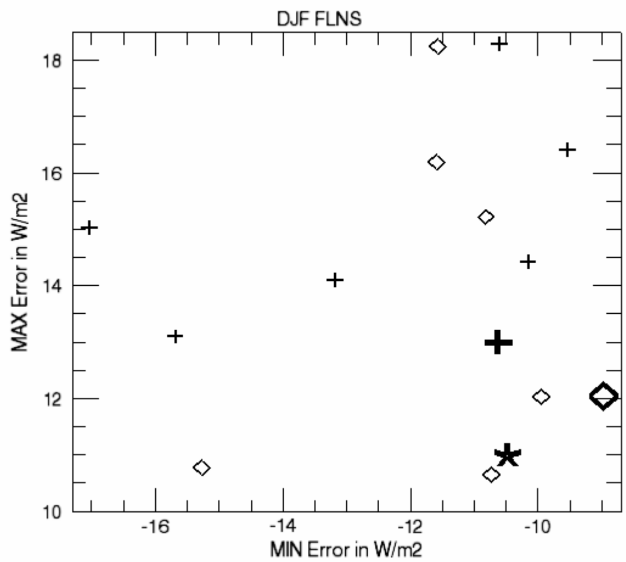

Fig. 5. Min and max errors for FLNS climate. Symbols as in Fig. 4.

Figs. 4 and 5 show the winter DJF (December through February) surface net LWR flux (FLNS) errors, in W/m², as deviations from the control climate. It is noteworthy that min and max errors shown in the right panel of this and the following figures are extreme outliers obtained for the entire 15 years of the model integration. Similarly, Figs. 6 and 7 show the major LWR characteristics, the DJF TOM (top of the model) net LWR flux (FLNT) errors as deviations from the control climate (in W/m²).

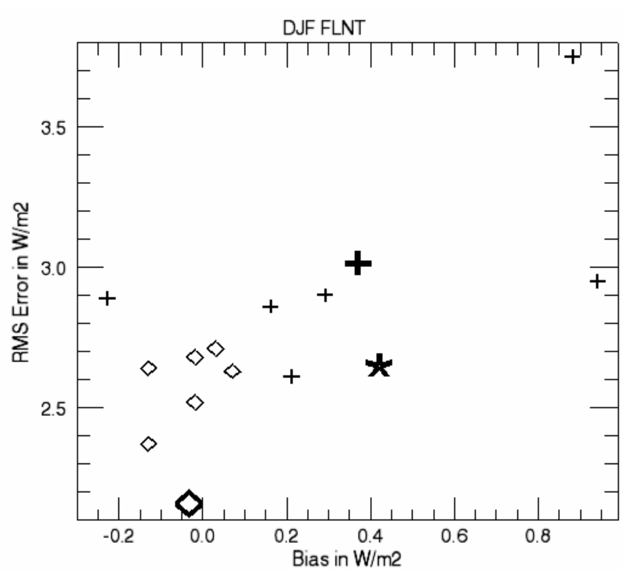

Fig. 6. Mean error (bias) and RMSE for the winter DJF (December through February) top of the model net LWR flux (FLNT). Symbols as in Fig. 4.

Figs. 8 and 9 show errors for DJF pressure at the surface level (PSL) as deviations from the control (in hPa).

The results presented in Figs. 4-9 clearly demonstrate that all three considered ensemble approaches give similar results in terms of improvement of the accuracy of climate simulations. Also they show that LTPPE generates a significantly larger spread of the ensemble members that PICE with random perturbation of initial conditions.

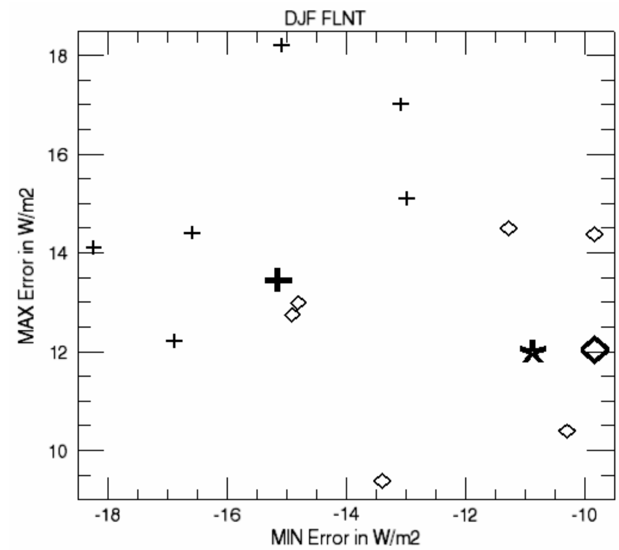

Fig. 7. Min and max errors for FLNT climate. Symbols as in Fig. 4.

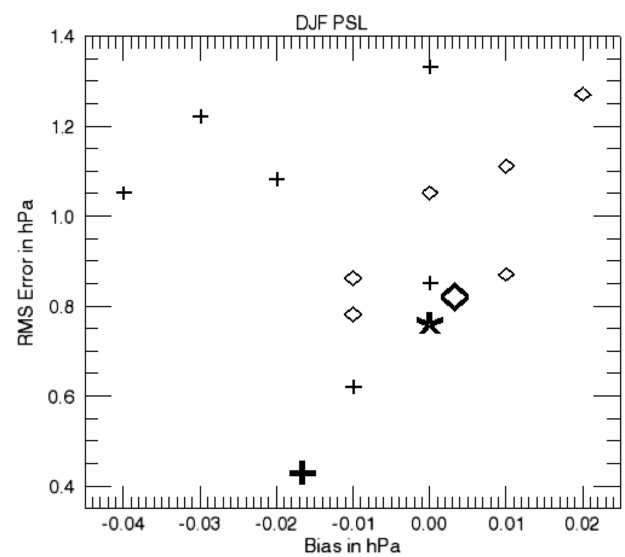

Fig. 8. Mean error (bias) and RMSE for the pressure at the surface level (PSL). Symbols as in Fig. 4.

V. Conclusions and Discussion

In this study we introduced a NN emulation technique as a tool for creating perturbed model physics for using in perturbed physics ensembles. We also introduced a new sort term perturbed physics ensemble (STPPE) approach. It is shown that the neural network emulation technique allows us: (1) to introduce fast versions of model physics (or components of model physics) that can speed up calculations of any type of ensemble up to 2 -3 times; (2) to conveniently an naturally introduce perturbations in the model physics (or a component of model physics) and to develop a fast versions of perturbed model physics (or fast perturbed components of model physics), and (3) to make the computation time for the entire ensemble (in the case of short term perturbed physics ensemble introduced in this paper) comparable with the computation time for one single model run.

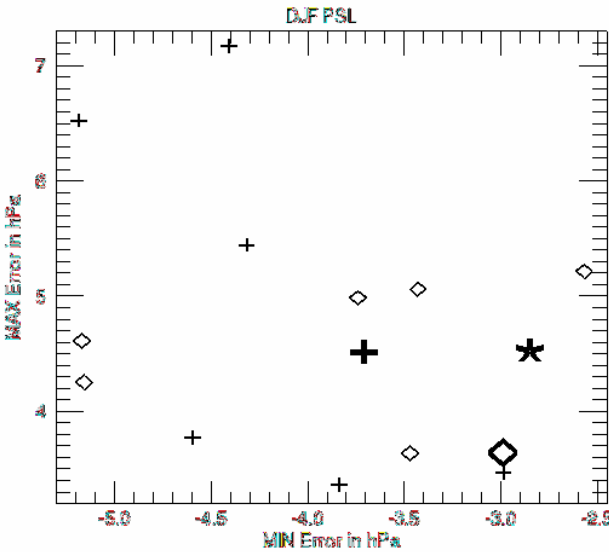

Fig. 7. Min and max errors for PSL climate. Symbols as in Fig. 4.

Preliminary results presented here show that all three ensemble approaches, the perturbed initial conditions ensemble (PICE), the long term perturbed physics ensemble (LTPPE), and STPPE, give similar results: the use of any of these ensembles in the climate simulation significantly reduces the systematic error (bias); it also reduces the random error making it close to that of the best individual ensemble member. The same is true for the extreme (min and max) errors.

All three considered ensembles demonstrate similar improvements of the climate simulation accuracy. Using NN emulations of model physics significantly improves the computational performance of any of investigated ensemble techniques. However, it is important to emphasize that STPPE is significantly faster than PICE and LTPPE. It is 2 N times (12 times for the case of $N = 6$ ensemble members considered in our study) faster than PICE and N times (6 times in our study) faster than LTPPE. Also, our results indicate that LTPPE using NN perturbed physics provides a significantly larger spread of ensemble members than PICE with randomly perturbed initial conditions.

This study is actually a pilot study that introduces and preliminary evaluates NNs as a tool for perturbing model physics and for using it in perturbed physics ensembles. This study also introduces STPPE as a new kind of ensemble approach. Some additional issues should be (and will be) investigated to obtain a more complete picture of advantages and limitations of using this approach:

- In this work we evaluated aforementioned ensemble techniques using the basic statistical metrics like bias, rmse, min and max errors. Various statistical metrics specifically designed for evaluation of ensemble prediction systems (EPS) [1,2] should be applied to perform enhanced quantitative comparison between PICE, LTPPE, and STPPE.
- It was shown that the perturbation ε_j introduced by NN emulation technique can be controlled and changed not only in terms of its value but also in terms of its statistical properties. A broader sample of NN emulations with a broader spread of error statistics should be considered and evaluated.
- In this study we used an unperturbed NCAR CAM run with the original parameterizations of physics as a control run or “synthetic observations”. Similar evaluation should be performed with real observations.
- A climate model, NCAR CAM, was used to evaluate aforementioned ensemble techniques in climate simulation environment. Similar evaluation should be performed in the framework of a numerical weather prediction EPS to evaluate these techniques for their advantages for the weather prediction.
- More realistic perturbation technique like used in [1,2] should be applied to create PICE for comparison with LTPPE and STPPE.
- Some parts of the climate/weather numerical model like convection physics, or full physics (containing boundary layer, land, and ice models), or model chemistry are not so well defined as the model radiation that we perturbed in this study; they introduce larger uncertainties in model calculations. These components may be even better candidates for perturbing them using the NN emulation technique.

REFERENCES

[1] T.N. Palmer, et al.: 2007, The Ensemble Prediction System – Recent and Ongoing Developments, ECMWF Technical Memorandum No. 540

[2] R. Buizza, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems, *Monthly Weather Review*, **133**, pp. 1076–1097

[3] A. J. Broccoli, , K. W. Dixon, T. L. Delworth, T. R. Knutson, and R. J. Stouffer: 2003: Twentieth-century temperature and precipitation trends in ensemble climate simulations including natural and anthropogenic forcing, *J. Geophys. Res.,* **108** (**D24**), 4798-4811, doi:10.1029/2003JD003812.

[4] J.M. Murphy, D.M. Sexton, D.N. Barnett, G.S. Jones, M.J. Webb, M. Collins, D.A. Stainforth: 2004, Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, **430**, 768 - 772

[5] D.A. Stainforth, et al., 2005, Uncertainty in Predictions of the Climate Responses to Rising Levels of Greenhouse Gases, *Nature*, **433**, 403 - 406

[6] M. Yoshimori, T. F. Stocker, Ch. C. Raible, and M. Renold, 2005: Externally Forced and Internal Variability in Ensemble Climate Simulations of the Maunder Minimum, *J. Climate*, **18**, 4253-4270

[7] R. Buizza, , M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, *Quarterly Journal of the Royal Meteorological Society*, **125**, pp. 2887-2908

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

[9] V.M. Krasnopolsky, M.S. Fox-Rabinovitz, and D.V. Chalikov, 2005: "New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Long Wave Radiation in a Climate Model", *Monthly Weather Review*, v. 133, No. 5, pp. 1370-1383.

[10] R. Buizza, 1997: Potential Forecast Skill of Ensemble Prediction and Spread and Skill Distributions of the ECMWF Ensemble Prediction System, *Monthly Weather Review*, **125**, pp. 99–119

[11] D.J. Stensrud, J.-W. Bao, and T.T. Warner, 2000: Using Initial Condition and Model Physics Perturbations in Short-Range Ensemble Simulations of Mesoscale Convective Systems, *Monthly Weather Review*, v. 128, pp. 2077-2107.
[12] V. V. Kharin, F. W. Zwiers, 2000: Changes in the Extremes in an Ensemble of Transient Climate Simulations with a Coupled Atmosphere-Ocean GCM, *J. Climate*, **13**, pp.3760-3788
[13] V.M. Krasnopolsky and M.S. Fox-Rabinovitz, 2006a: "Complex Hybrid Models Combining Deterministic and Machine Learning Components for Numerical Climate Modeling and Weather Prediction", *Neural Networks,* 19, pp. 122-134
[14] Ch. M. Bishop, 2006: *Pattern Recognition and Machine Learning*. Springer, 738 pp.
[15] K. Funahashi, 1989: "On the Approximate Realization of Continuous Mappings by Neural Networks". *Neural Networks*, 2, pp. 183-192
[16] V.M. Krasnopolsky, 2007a, "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*, in press
[17] V.M. Krasnopolsky and M.S. Fox-Rabinovitz, 2006b: "A New Synergetic Paradigm in Environmental Numerical Modeling: Hybrid Models Combining Deterministic and Machine Learning Components", *Ecological Modelling*, v. 191, 5-18
[18] W.D. Collins, J.K. Hackney, and D.P. Edwards, 2002: A new parameterization for infrared emission and absorption by water vapor in the National Center for Atmospheric Research Community Atmosphere Model, *J. Geophys. Res.,* **107** (**D22**), 1-20.
[19] V.M. Krasnopolsky, 2007b: "Reducing Uncertainties in Neural Network Jacobians and Improving Accuracy of Neural Network Emulations with NN Ensemble Approaches", *Neural Networks, 20,* pp. 454-461