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

Office Note 469

Development of neural network convection parameterizations for climate and NWP models using Cloud Resolving Model simulations[1]

Vladimir Krasnopolsky[2],EMC, NCEP/NWS, NOAA
and Earth System Sciences Interdisciplinary Center (ESSIC), University of Maryland (UMD)
Michael Fox-Rabinovitz,ESSIC UMD,
Alexei Belochitski,ESSIC UMD
Philip J.Rasch, Pacific Northwest National Laboratory,
Peter Blossey, University of Washington,
and
YefimKogan, Scripps Institute at University of California at San Diego

November 2, 2011

[1]MMAB Contribution No. 293
[2]e-mail: Vladimir.Krasnopolsky@NOAA.gov

Abstract

A novel approach based on the neural network (NN) technique is formulated and used for development of a NN ensemble stochastic convection parameterization for climate models. This fast parameterization is built based on data from Cloud Resolving Model (CRM) simulations initialized with and driven by the TOGA-COARE data available for the 4-month winter season from November 1992 to February 1993. CRM simulated data were averaged and projected onto the GCM space of atmospheric states to implicitly define a stochastic convection parameterization. This parameterization is emulated using an ensemble of neural networks (NNs). The developed NNs are trained and tested. The inherent uncertainty of the stochastic convection parameterization derived following this approach is estimated. The newly developed NN convection parameterization has been tested in a diagnostic mode of NCAR CAM. It produced reasonable and promising climate simulations for the TOGA-COARE winter. This paper is devoted to discussion of the concept, methodology, initial results,and the major challenges of development of NN convection parameterizations for climate and numerical weather prediction (NWR) models.

1. Introduction

Clouds and convection are among the most important and complex phenomena of the Earth's physical climate system. The processes that control clouds, and through which they interact with other components of the Earth system involve slow and fast fluid motions carrying heat, moisture, momentum and trace constituents, and influence other important physical processes through phase changes of water substances, radiative transfer, chemistry, production and removal of trace constituents, and atmospheric electricity. Clouds are strong regulators of radiant energy, sites for chemical processing, participate in many climate feedbacks, and are potentially very susceptible to anthropogenic change (indirect aerosol effects). They figure prominently in all climate change assessments.

In spite of intense studies for centuries, clouds still provide an intellectual and computational challenge. Because of the vast range of time and space scales involved, researchers typically focus on a particular component of a cloud system, with a narrow range of time and space scales, and prescribe features of the cloud that operate outside of that range. For example, 'box' models treat a small air parcel as an approximately spatially homogeneous medium over length scales of order a meter to explore the evolution of a cloud drop spectrum. Motions of the air as a fluid are prescribed, and the focus is on molecular scale motions (e.g. vapor deposition) and drop scale motions (e.g. drop coagulation). At the other end of the spectrum of representations of clouds is their representation in large scale models, for example, in General Circulation or Global

Climate Models (GCMs). GCMs treat convective clouds very simply. While scientific problems that use these models are among the most computationally intensive applications in the history of scientific exploration, the models employ drastic simplifications in their treatment of many processes important in climate and weather.

Parameterizations must represent the effect of clouds on time and space scales that are well below the resolution of the scales explicitly treated in GCMs, which resolve atmospheric features with space scales of order 100km, and time scales of order 10 minutes. NWP models typically operate at smaller spatial and temporal scales, but most cloud processes are still acting well below these scales. Cloud systems are driven by and interact with the Earth system through a variety of physical processes, and this is the reason that we see the incredible variety of clouds in the atmosphere.

Clouds that exist in situations where buoyant overturning is one of the major controlling component processes are called convective clouds. GCMs frequently treat convective clouds in parameterizations as 'plumes' consisting of updrafts and downdrafts. All of these aspects of clouds must be treated in some fashion in models that attempt to realistically represent atmospheric evolution, and it has become traditional in large scale models to treat many of these regimes and processes (stratiform clouds, convective clouds, microphysics, sub-gridscale cloud fraction, and sub-gridscale cloud overlap) somewhat independently, even though they are intimately connected. For this reason, models produce independent, but connected 'parameterizations' (representations) of each of these processes.

In this paper we focus on improving the representation (parameterization) of convective clouds. Convective clouds are loci for some of the most vigorous vertical motions in the atmosphere, transporting heat, momentum, moisture, and trace constituents from near the surface to 10s of km in altitude in minutes. They are also regions where much of the rain hitting the earth's surface is produced, and the local origin or source of water that appears in stratiform clouds like 'cirrus anvils'.

The representation of cloud processes has long been recognized as a challenge at any of the space and time scales mentioned above, and the problem is particularly difficult for global models. The scientific community realizes that many critical aspects of our ability to represent the atmosphere for climate and NWP problems are being hindered by the representation of clouds, and that we have reached an impasse in our ability to improve these processes. These issues are eloquently discussed by Randall et al. (2003). That paper makes the point that models that explicitly resolve processes at the smaller time and space scales that are relevant to many features of clouds systems (e.g. 10s of meters to less than 10 kilometers, and time scales of seconds to minutes), so called Cloud Resolving Models (CRM) also called Cloud System Resolving Models(CSRMs) or Large Eddy Simulation Models (LES), are usually (but not invariably) able to simulate component aspects and evolution of the cloud systems much more realistically than large scale models. The improvements seen in representing clouds using CSRMs and LES have provided the motivation for the approaches described in section 2.

1.1. Problems seen using traditional approaches in representing convection

In spite of their importance to the Earth system, and the attention that atmospheric modelers have paid to them over the last 30 years, it is clear from a variety of analyses that our current parameterizations of convective processes are inadequate. We list a few of the studies documenting these inadequacies. We stress that the features are present in virtually all global models today, and are also documented in many other studies with many other models.

1. Global atmospheric model used for NWP seem unable to produce the correct balance in the wind, water substances, mass, and temperature field to allow precipitation to be produced realistically when models are started from observed atmospheric initial conditions. This phenomena, termed convective 'precipitation spin up' (see, e.g., Donner and Rasch, 1989) results in substantial imbalance in these atmospheric fields and require explicit adjustments to initial conditions used in weather forecast models to avoid unrealistic transients in precipitation forecasts.

2. Global large scale models used in either forecast or climate mode show problems in predicting the diurnal variation of convection, both over land, and oceans (e.g. Rasch et al., 2006). This problem is less prevalent in CSRM simulations of convection (e.g. Guichard et. al., 2004, Grabowski et al., 2006), which suggest that shallow convection may play a role in moistening the lower troposphere prior to deep convection, and that there may be a need for a time evolving cumulus entrainment rate in the convection.

3. Global Models frequently produce a persistent 'double ITCZ' (Inter-Tropical Convergence Zone) in the western and central tropical Pacific and the Indian Ocean, that

is seen only occasionally in the real world, and only over a very limited region of the pacific (Hack et al, 2006). These biases have been shown to be quite sensitive to the formulation of convection in GCMs (Lin et al, 2006).

4. Although there are still many uncertainties in observational estimates of the burdens of cloud condensate (both liquid and ice), we believe models still have large biases, particularly in convective clouds (see e.g., Rasch and Kristjansson, 1998), and in the partitioning between the convective and stratiform components of tropical cloud systems (see, e.g. Rasch, et. al., 2006; Hack et. al, 2006).

5. Global models frequently produce too much convective precipitation, with a relatively uniform heating profile in the vertical, instead of associated 'stratiform precipitation' driven by convection that has heating aloft, and cooling below (Rasch et al, 2006). These biases produce a far field effect (a teleconnection) that is manifested in errors in the Walker Circulation, in the El Nino Southern Oscillation (ENSO) phenomena, and in stationary wave patterns far from the original location of the heating (e.g. Gill, 1980; Branstator, 1992).

6. Convection is a process that moves tracers rapidly from the surface to the upper troposphere. Numerous comparisons of global models to short and long lived tracers suggest that the balance of deep vertical motions associated with weakly entraining plumes that penetrate deep into the troposphere, and shallower motions associated with stronger lateral entrainment/detrainment events is poorly represented in current atmospheric models (Rasch et al., 2000)

1.2. Current efforts by others to improve the treatment of clouds in large scale models

Because of the problems described above, the research community has been exploring alternate ways of making progress in representing clouds in global models. Among the alternate paradigms for treating clouds in global models are the following:

1. Cloud System Resolving Models (CSRMs). These models, first developed in the 1970s and 1980s (e.g. Krueger, 1988), operate over a limited area (typically continental scale or smaller) at finer spatial and temporal time scales than a global model (although they still require many simplifications). The spatial and temporal average behavior of these models tells us what might be gained if we could resolve many of the phenomena that global models must ignore (e.g. higher resolution fluid dynamic motions that can resolve some updrafts and downdrafts, convective organization, meso-scale circulations, and stratiform and convective components that interact with each other, etc.). It is important to note that CSRMs do not simulate all processes important to cloud systems from first principles. There are still many sets of processes (for example microphysical and turbulent processes) that are still treated crudely, but they do resolve many more phenomena than today's global model parameterizations.These process parameterizations in CSRMs are most often applied locally because the redistribution of mass momentum, energy and moisture by the largest convective/cloud motions is resolved on the global model grid.

The domain average of solutions from these models driven by observationally based estimates of observed meteorological events can be directly compared with the physics

from a single column model of a GCM, which has come to be known as a 'Single Column Model' (SCM). This has proven to be a particularly valuable tool in assessing the state of our understanding of cloud processes. A case in point can be found in the Global Energy and Water Cycle Experiment (GEWEX) Cloud System Study (GCSS) project. This project (see Randall et al, 2003b) was designed to confront CSRMs and SCMs with the observationally based estimates of the driving meteorological forcing, as well as the behavior and response of cloud processes, and the local environment (temperature, moisture, precipitation), to this forcing. GCSS studies have explored the behavior of models (and thus our understanding and representation of the atmospheric processes) of clouds in a variety of regimes, e.g. convective, polar, and boundary layer clouds.

It is an explicit assumption within the GCSS project that CSRMs will provide a more accurate representation of the behavior of clouds than SCMs, and that if that is true then information can be gleaned from the CSRMs that will help in the development of the next generation global model parameterizations. It is reassuring to find that, although not uniformly true, that in most regimes this is the case: CSRMs produce better simulations of deep tropical maritime convection, and summertime mid-latitude continental convection (Randall et al, 2003b; Guichard et al, 2004), and tropical continental convection driven by diurnal variations in heat release in the absence of large scale forcing (Grabowski, 2006). All situations have been simulated more realistically using CSRMs than single column models including parameterized convection. The CSRMs also show a realistic transition from no convection, to shallow, to deep convection as the

day proceeds, with a peak in precipitation that occurs in late afternoon, which SCMs have much more difficulty reproducing.

2. Recently attempts have been made to develop global models that resolve some cloud scale motions (down to horizontal resolutions of about 3km) (Global Cloud Resolving Models or GCRMs, e.g. Miura et al, 2005; Satoh et al, 2005). They are incredibly expensive to run when compared with the cost of a typical general circulation model (see next item), and have generally been used in idealized settings for exploratory experiments (in an water covered planet scenario for example)

3. The term 'super-parameterization', also known as a 'Multiscale Modeling Framework' (MMF), was originally suggested by Grabowski (2001), and subsequently developed by a group at Colorado State University (e.g. Khairoutdinov and Randall, 2001; Randall et al 2003a). This concept refers to the embedding a simplified CSRM into each column of a global model. Because of the simplifications, such a model is substantially less expensive (by a factor of 10^3-10^4) than a GCRM, but it is still enormously more costly (10^2 -10^3) than models using a conventional convective parameterization (Randall et al., 2003a). A typical MMF parameterization might use a two-dimensional CSRM with 64 sub-columns within each GCM column, to produce sub-resolution of about 4km. The simplifications buy a substantial reduction in cost, with a consequent compromise in accuracy of physical representation. There are for example, significant changes in the storm evolution (e.g. vertical velocities) produced in 2-D CSRMs when compared with 3-D formulations, and 2D CSRM simulate too rapid a transition from shallow to deep convection, and too much cloud cover (e.g., Grabowski et

al., 2006). These artifacts will also be present in MMF formulations. Other compromises are discussed more thoroughly in the references.

In this paper we introduce an alternative approach based on neural network (NN) technique – a particular case of statistical learning technique (SLT). This approach allows one to develop a NN convection parameterization, which can be used as a parameterization in GCM and can effectively take into account major sub-grid effects taken into account by other approaches (see above) at a fraction of the computational cost. In Section 2 we introduce our approach, discuss sources of uncertainties in the convection parameterization and estimate these uncertainties. We also describe the NN training and perform an initial evaluation of the developed NN parameterization on an independent training set. In Section 3 we introduce the newly developed NN convection parameterization into NCAR CAM in a diagnostic mode and perform and analyze the results of a climate simulation for the TOGA-COARE 4-month winter period (November 1992 – February 1993). Section 4 contains discussion and Section 5 conclusions.

2. Formulation of our Approach: Development of NN ensemble convection parameterization from CRM data

Our approach is aimed at developing NNs which emulate the behavior of a CSRM run at larger scales (closer to GCM scales) in a variety of regimes and initial conditions. The resulting emulations can be used as a novel, and computationally viable convection parameterizations. If successful, it will produce a parameterization of similar or better quality than the superparameterization or MMF, effectively taking into account subgrid (in terms of GCM scales) effects at a fraction of the computational cost.

As we showed in our previous works (e.g., Krasnopolsky 2007a) any parameterization of model physics can be emulated using multilayer perceptron NNs. This NN is an analytical approximation that uses a family of functions like:

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

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 neurons in (1).

2.1. Design and development of the NN parameterization and the training set

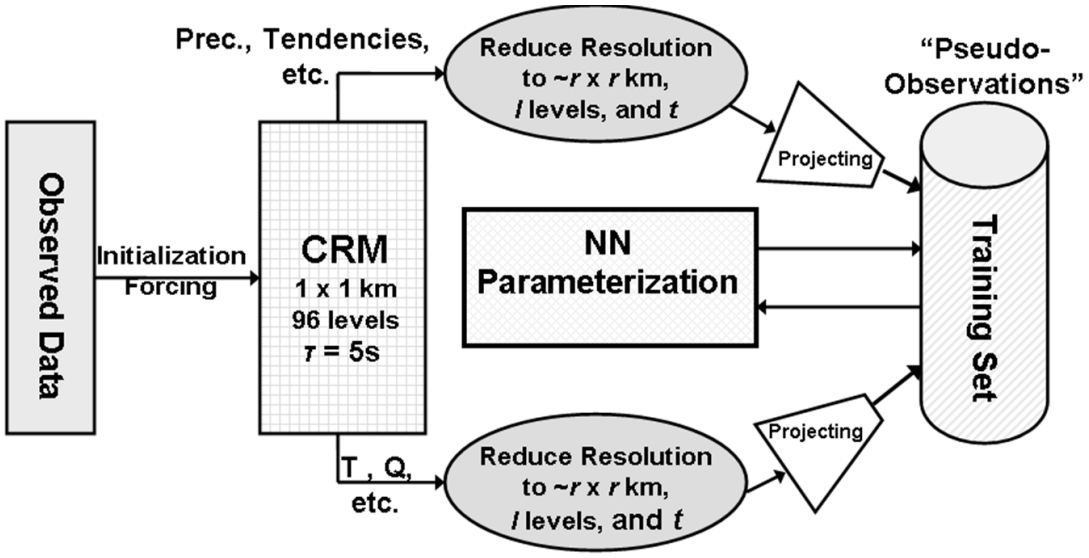


Fig. 1 Development design of a NN convection parameterization.

Fig. 1 summarizes the process of development of the NN convection parameterization. The CRM used in our work is the SAM (System for Atmospheric Modeling)developed and provided by M. Khairoutdinov (Khairoutdinov and Randall, 2003). It has been used for our CRM simulations. SAM uses TOGA-COARE data (ARM or other observations) for initialization and forcing and has the horizontal resolution ρ of about 1 km, 64 or 96 vertical layers, and time integration step of 5 s. We integrate the CRM over the domain of 256 x 256km. The development of an NN parameterization is a multi-step process. These steps are:

1. Simulating CRM data. The SAM is run for the entire TOGA-COARE period (120 days in our current experiments) and the high resolution output of the model is archived.

2. Reducing the resolution of simulated data. The high resolution CRM simulated data are averaged in space and time. The data are averaged to a reduced horizontal resolution of $\rho<r \leq R$, where ρ and R are the CRM and GCM resolutions correspondingly, and are interpolated/averaged to the number of vertical layers $l = L$, where L is the number of vertical layers in GCM.

3. Projecting a CRM space of atmospheric states to a GCM space of atmospheric states. The subset of variablesis selected from the reduced resolution CRM simulated data created at the previous step, and this subset constitutes the NN development dataset. Only variables that can be identified with corresponding GCM variables or can be calculated from or converted to prognostic or diagnostic variables available in GCM, are included in the development set (called "pseudo-observations" in Fig.1; actually they are obtained from the averaged CRM/SAM simulated data). Only these variables are used as inputs and outputs of our NN parameterization. The choice of "inputs" and "outputs" for a NN convection parameterization is very critical, and influences the following discussion. For example, a simple convective parameterization might define "temperature", "water vapor" and the convergence of temperature and water vapor to be "inputs", and produce Q1C and Q2, the apparent heat and moisture tendencies, as the "outputs". These could also be viewed as the inputs and outputs of a CRM, but the outputs Q1C and Q2 clearly depend upon other variables (for example, the condensed water in each CRM column) that are not necessarily considered to be part of either the inputs or outputs of the NN. These variables cannot be included as NN inputs and/or outputs simply because they are not available in GCM. From the point of view of GCM "model reality" these variables

are "hidden" variables responsible for sub-grid scale variability. The acknowledgement of this challenge requires development of the concept of uncertainty and "stochasticity" discussed in the following section. The development set of "pseudo-observations" implicitly represents a stochastic convection parameterization with an uncertainty, which is an inherent feature of such a parameterization (see Section 2.2). The dashed lines in Fig. 1 show that, if it is found to be desirable, the high resolution CRM simulated data and/or even observed data can be added to the development set to enrich sub-grid variability in the development data.

4. CRM vs. GCM mean differences for all variables selected as the NN parameterization inputs and outputs have to be determined. These differences are a result of CAM and CRM being two different models with different temporal and spatial scales and resolutions, with different dynamics and physics; they have different boundary and initial conditions and different forcing.
5. The developed "pseudo-observations" (averaged SAM simulated data) are separated into two sets, one set being used for training and another independent set for testing/validation. Then the NN parameterization is trained using the training set. Due to the inherent uncertainty of pseudo-observations, the NN convection parameterization is implemented as an ensemble of NNs.

The validation procedure for the NN parameterization consists of two steps. First, the trained NN is applied to the test set and error statistics are calculated. Second, the tested NN parameterization is included into GCM to validate its behavior in the model

simulations. This last step is the most important step of our approach. The pseudo-observations used for development of the NN parameterization are not real observations. They represent the "virtual reality" of the averaged SAM simulated data. We use them for development of the NN parameterization to be introduced into CAM. CAM has its own "virtual reality", which may not be in a complete agreement with the averaged SAM simulated data and therefore with the NN parameterization trained on pseudo-observations derived from the averaged SAM simulated data. Thus, special efforts may be required to calibrate/synchronize/make consistent the "virtual realities" of CAM and average SAM simulated data. This issue is investigated in Section 2.5.

Let us now follow the steps of the development process formulated above and discuss the uncertainties introduced at each of these steps into the final NN parameterization. First, we discuss the first three stages of the development, which are performed prior to the use of NN. These stages lead to creating the development data set of "pseudo-observations" that implicitly define the stochastic convection parameterization. Then this parameterization is emulated using an ensemble of NNs.

2.2. Parameterization and its Uncertainties

In this Section we outline the sources of uncertainties, which emerge in the process of the data preparation for development of an NN convection parameterization for GCM based on the data simulated by a CRM. We discuss the major properties/sources of the uncertainties and show that the uncertainty is an inherent part of the data and, therefore,

of the parameterization derived in such a way; and that this parameterization is essentially a stochastic parameterization.

The first three steps of the development process formulated in the previous sections introduce uncertainties in the training data set or "pseudo-observations". The uncertainties are introduced at each of these steps and their sources can be traced step by step.

First, the CRM may be formally considered as a mapping μ that defines the relationship between two vectors: the input vector (x) and the output vector (y) that are composed of CRM variables. At each time step the mapping μ given vector x produces vector y or,

$$y = \mu(x) \tag{2}$$

Here (x, y) are high resolution SAM variables (produced with spatial resolution of 1 km and temporal resolution of 5 s) or SAM simulated data; they are related by SAM and this fact is expressed by eq. (2). The mapping μ is an exact (or deterministic, or physically based) mapping, which means that it is explicitly represented by a complete set ofSAM equations, and that one particular y corresponds to each particular x. The first step in our developmental process consists of simulating the CRM or applying (2) at each time step and continuing with the simulation for a period of time T. The CRM simulation is initialized, and then forced at each time step by large-scale observational data prorated per a time step. However, because the CRM physics (like microphysics) is partially

parameterized, contains a number of simplifications, and the forcings are approximate, CRM is not "perfect" and the CRM simulated data will deviate from the observational data. This difference between the observed reality and the "CRM reality" is the first contribution to the uncertainty of pseudo-observations and the NN parameterization derived from these data.

The second step in our development process consists of averaging high resolution simulated data (x,y) over some area $r(\rho < r < R)$and over a time interval t $(\tau < t < T)$, where ρ = 1km and τ= 5sare the CRM resolution and integration time step respectively; R and T are the GCM resolution and integration time step respectively. As the result, averaged vectors of simulated data $\underline{x}$ and $\underline{y}$ are produced. Here the bar below the symbol means averaging over r and t. Fig. 2 shows the evolution of the spectrum of precipitation rates with respect to changing resolution of the field that we are sampling. It demonstrates the sensitivity of precipitation to resolution changes over the r range from 16 to 256 km. Fig. 3 shows the evolution of the profile of precipitable moisture ("Rain and Snow") with the maximal profile variability for the same range of r.

By changing r and t, we can regulate the amount of sub-grid information (high frequency variability) in pseudo-observations that we want our parameterization (derived from the pseudo-observations) to introduce into CAM. Thus, moving from r = 1 km and t= 5 s (high resolution SAM data) to lower resolution and larger t, we will gradually reduce the sub-grid signal introduced in CAM. Determining the optimal values for r and $t$is one of the most important topics that we have not investigatedyet. However, it can be

investigated only through validation of performance of the developed NN convection parameterizations in CAM.

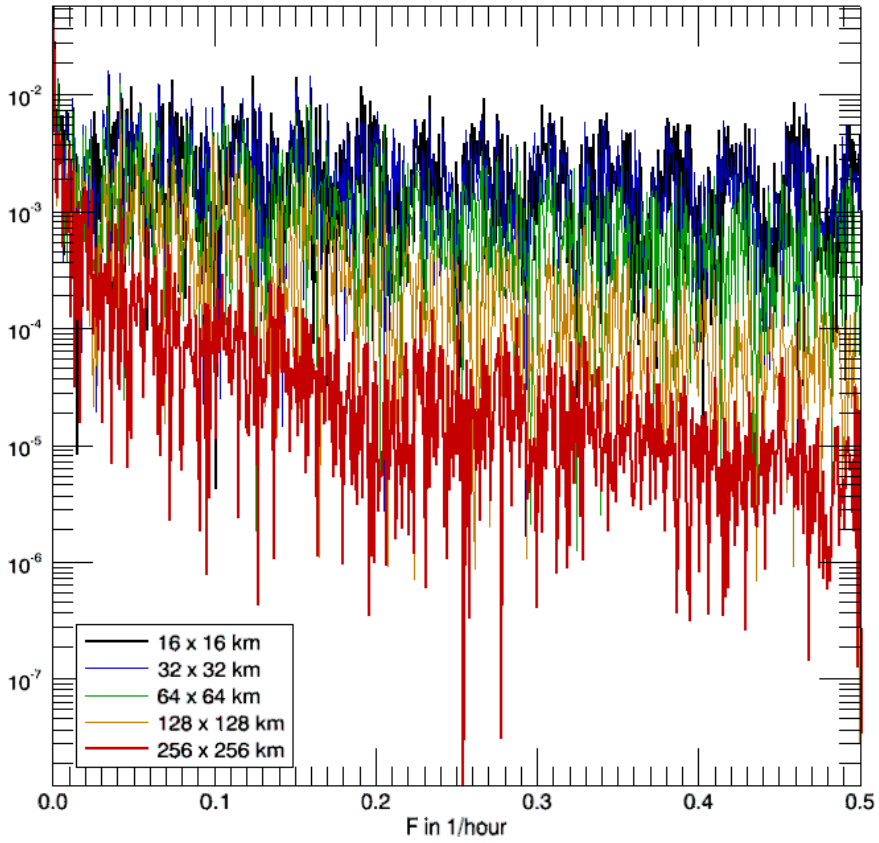


Fig.2 Spectra of horizontally averaged data for precipitation rates for different averaging areas ($r \times r$) with r = 16, 32, 64, 128, and 256 km.

For Fig. 3 a particular r (e.g., r = 16 km) was selected and a time series (120 days in our case) of profiles with 256 (for r = 16 km) profiles for each time step was obtained. Then all but a single profile with the maximum profile variability were removed. This profile is shown in black (for r = 16 km) in Fig. 3. The same has been done for all other considered cases of r (for r = 32 km we have 64 profiles for each time step, etc.).

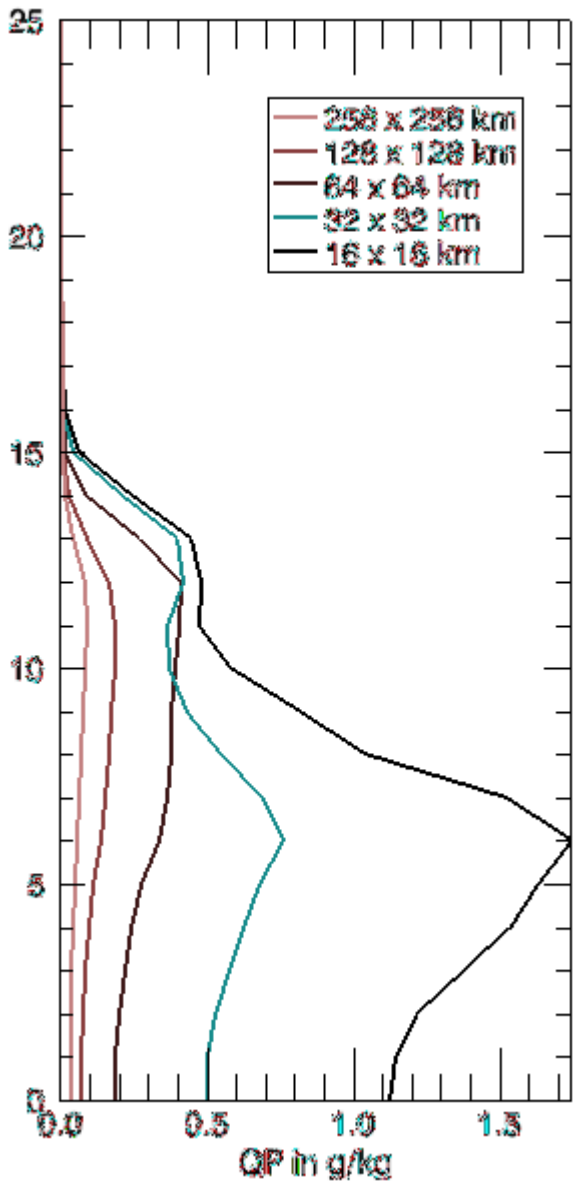


Fig.3 Evolution of the profile of precipitable moisture ("Rain and Snow“) with the maximal profile variability in response to the change of averaging areas ($r \times r$) with r = 16, 32, 64, 128, and 256 km.

As expected, both Figs. 2 and 3 clearly show the reduction of high frequency variability in averaged simulated data $\underline{x}$ and $\underline{y}$ as compared with the original high resolution simulated data x and y. It is important to emphasize that the new variables $\underline{x}$ and $\underline{y}$ are

stochastic variables that are distributed around their mean values with some probability density functions.

Now we assume that there exists a mapping $\underline{\mu}$ between the averaged simulated data $\underline{x}$ and $\underline{y}$, which can be written as,

$$\underline{y} = \underline{\mu}(\underline{x}) + \underline{\varepsilon} \tag{3}$$

where $\underline{\varepsilon}$ is a measure of uncertainty in the mapping introduced by the two previous developmental steps. The uncertainty $\underline{\varepsilon}$ reflects the fact that the mapping (3) is not exact.

The mapping $\underline{\mu}$ is a complex **stochastic mapping** between two **stochastic vector variables** $\underline{x}$ and $\underline{y}$. The "stochasticity" of the mapping $\underline{\mu}$ contributes significantly to the uncertainty $\underline{\varepsilon}$ in (3). The stochastic mapping (3), for each particular value $\underline{x}$, may generate many different values $\underline{y}$ with different probabilities determined by their joint probability density function $\rho(x,y)$. Also one value of $\underline{y}$ can be generated by a stochastic mapping from different values of $\underline{x}$ with different probabilitiesdetermined by their jointprobability density function.

The next step of our developmental process is **projecting**the CRM space of atmospheric states onto the GCM space of atmospheric states. It starts from a transition from averaged CRM variables $\underline{x}$ and $\underline{y}$ to a subset of these variables, X' and Y'. Let us write

$\underline{x}$ and $\underline{y}$ as $\underline{x} = \{X', x'\}$ and $\underline{y} = \{Y', y'\}$. Here we split each vector $\underline{x}$ and $\underline{y}$ into two parts. The **new variables** X' and Y'**include only variables that can be identified with corresponding GCM variables or can be calculated from or converted to prognostic or diagnostic variables available in the GCM**; all other variables x' and $y'$are projected out (averaged out), i. e., can not and will be not used. We can rewrite the mapping (3) before the projection in the new coordinates as,

$$Y' = \underline{\mu}(X', x', y') + \underline{\varepsilon} \tag{4}$$

y' is present in the right hand side of eq. (4) because $y'$and Y' are physically related and correlated. Now let us expand the mapping (4) in the Taylor series at $x' = \overline{x'}$ and $y' = \overline{y'}$ where the upper bar **means averaging over x' and y'.**

$$\begin{aligned} Y' = \underline{\mu}(X', \overline{x'}, \overline{y'}) + \left.\frac{\partial \underline{\mu}(X', x', y')}{\partial x'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot (x' - \overline{x'}) + \left.\frac{\partial \underline{\mu}(X', x', y')}{\partial y'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot (y' - \overline{y'}) + \\ + \left.\frac{\partial^2 \underline{\mu}(X', x', y')}{\partial x' \partial y'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot (x' - \overline{x'}) \cdot (y' - \overline{y'}) + \left.\frac{\partial^2 \underline{\mu}(X', x', y')}{\partial^2 x'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot (x' - \overline{x'})^2 + \ldots + \underline{\varepsilon} \end{aligned} \tag{5}$$

where derivatives in (5) are subsections of the Jacobian (the first order derivatives) and the Hessian (the second order derivatives) matrixes of the mapping (4).

Let us denote the projected variable Y after the averaging over x' and y'as,

$$Y = \overline{\underline{\mu}(X', \overline{x'}, \overline{y'})} = M(X) \tag{6}$$

where the long upper bar means averaging over x' and y'. Let us now estimate an additional uncertainty $\overline{\varepsilon}$ which is introduced by the projection (the upper bar means averaging over x' and y'; the lower bar means averaging over r and t).

The systematic part of this uncertainty can be estimated by combining (5) and (6) and averaging over x' and y',

$$\overline{(Y'-Y)} = \left.\frac{\partial \underline{\mu}(X', x', y')}{\partial x'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot \overline{(x'-\overline{x'})} + \left.\frac{\partial \underline{\mu}(X', x', y')}{\partial y'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot \overline{(y'-\overline{y'})} +$$
$$+ \left.\frac{\partial^2 \underline{\mu}(X', x', y')}{\partial x' \partial y'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot \overline{(x'-\overline{x'})\cdot(y'-\overline{y'})} + \left.\frac{\partial^2 \underline{\mu}(X', x', y')}{\partial^2 x'}\right|_{\substack{x'=\overline{x'} \\ y'=\overline{y'}}} \cdot \overline{(x'-\overline{x'})^2} + \ldots = \tag{7}$$
$$= \alpha \cdot b_{x'} + \beta \cdot b_{y'} + \gamma \cdot c_{x'y'} + \delta \cdot \sigma_{x'}^2 + \eta \cdot \sigma_{y'}^2 + \ldots$$

Here in (7) α, β, γ, δ, and η are coefficients built from the mapping derivatives; $b_{x'}$ and $b_{y'}$ are biases, $\sigma_{x'}^2$, $\sigma_{y'}^2$ are variances of x' and y'; and $c_{x'y'}$ is the correlation coefficient (covariance matrix) between x' and y'. If the averaging is performed over the entire domain of x' and y' (or over the representative training data set), then the biases $b_{x'}$ and $b_{y'}$ are equal to zeros. However, the systematic part of the uncertainty (7) also includes all higher moments of x' and y', which are not zeros; therefore, thesystematic part of the uncertainty $\overline{\varepsilon}$, which introduced by projecting the CRM space of atmospheric states onto

the GCM space of atmospheric states, is, in general, not zero. The random part of the uncertainty $\underline{\varepsilon}$ can be calculated in a similar way and it is not zero as well.

Taking into account the uncertainty $\underline{\varepsilon}$, the mapping (6) between the new vectors of projected variables *X* and *Y* can be written as,

$$Y = M(X) + \bar{\varepsilon} + \underline{\varepsilon} = M(X) + \varepsilon \tag{8}$$

This new mapping (8) is obtained from the mapping (3) by projecting out all omitted CRM variables. The mapping (8) is also a stochastic one and the new component of the uncertainty, the vector $\underline{\varepsilon}$, is due to the projection of the CRM state space onto the GCM state space. This uncertainty emerges due to an unaccounted variability of omitted parametersx' and $y'$of vectors $\bar{X}$ and $\bar{Y}$. After projecting these vectors onto the GCM space, *the projected vectors X and Y do not correspond to any particular values of omitted parameters*x' and y'*; these values are uncertain.* Actually, when we learn mapping *M* from data, the projected vectors *X* and *Y* correspond to the mean values of omitted parameters calculated over the training set. The vector $\underline{\varepsilon}$ contains not only random but also a systematic component. Finally, merging $\bar{\varepsilon}$ and $\underline{\varepsilon}$ into the vector ε we get the mapping (8) that we want to use as a parameterization in GCM.

Thus, *mapping (8) is not an exact mapping as is the mapping (2)*; it is a **stochastic mapping** between two stochastic vector variables *X* and *Y*. Stating this we would like to

emphasize that the *stochastic elements emerge in this problem before introducing the NN emulation technique*. Actually **(8) is a stochastic parameterization**, which inherently contains the uncertainty $\mathcal{E}$**.** The parameterization is *implicitly defined by the training set (X, Y)***. The uncertainty in this case is not a destructive noise; it is an inherent informative part of the stochastic parameterization, which contains important statistical information about sub-grid scale** (in terms of GCM scales) **effects. Actually, the stochastic parameterization is a family of mappings distributed with a distribution function.The range and shape of the distribution function are determined by the uncertainty vector** $\mathcal{E}$**.**

It is noteworthy that the distinction between a deterministic (physically based parameterization represented by a closed set of analytical equations or unambiguous computer code) and stochastic parameterization is not well defined. Actually, physically based parameterizations usually contain the same sources of uncertainty that have been discussed above, such as averaging and projecting some internal or hidden (in terms of GSM) variables. Also they include several additional sources of uncertainties: first, they contain approximations (that can be done in various ways with different results) and second, they include multiple empirical, statistically derived parameters, which introduce significant uncertainty in the parameterization because the parameters are derived from data with uncertainties.

Thus, rigorously speaking, deterministic/physically based parameterizations should be also considered as a family of parameterizations inside the uncertainty corridor (or

stochastic parameterizations). Actually, in a strict sense, any parameterization of physics should be considered as a stochastic one; and it is desirable/necessary to take into account the aforementioned uncertainties when introducing them in GCM using ensembles or some other approaches to obtain better representation of subgrid effects. However, physically based parameterizations used in GCM are usually closed analytical expressions. Theyare usually considered as "exact" parameterizations, at least they are not considered as stochastic parameterizations with uncertainty. Actually such "exact" parameterization is a single member of a family of parameterizations representing a stochastic parameterization; and this single member is obtained by assuming uncertainty to be zero. Stochastic physics, if used, is introduced as an additional procedure (e.g., Palmer and Williams 2008).

2.3. NN Emulation of the Convection Parameterization and Estimation of its Uncertainties

2.3.1. Data

A data set was simulated for the development. It is limited by the length of observational data set needed for driving/forcing SAM simulations. SAM/CRM using the TOGA-COARE forcing was ran for 120 days for the 256 x 256 km domain with 1 km resolution and 96 vertical layers (0 – 28 km). Then it wasaveraged at every hour of model integration to produce a simulation data set with an effective horizontal resolution of 256 km and temporal resolution of 1 hour. Finally, only variables that are available in GCM (NCAR CAM) or can be calculated there have been selected. The final data set consists of 2,800 records (every hour or hourly mean data). The simulation dataset was

partitioned into two parts: a training set consisting of 2,240 records or 80% of data and a test set consistingof 560 records or 20% of data. Namely, first 2,240 records are included in the training set and the last 560 records in the test set.

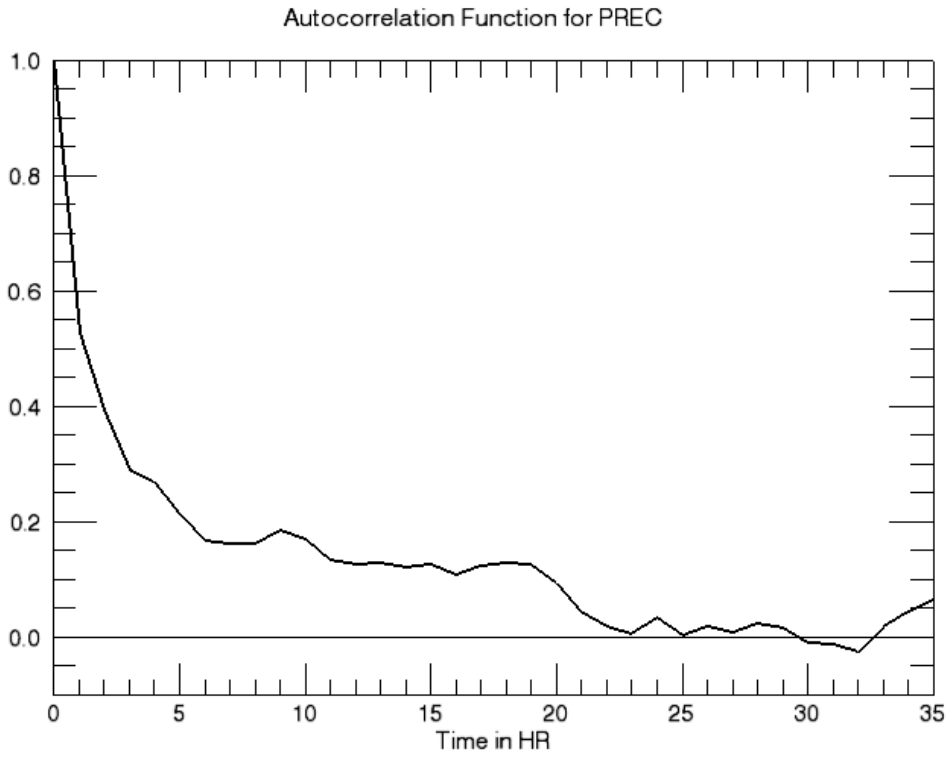


Fig.4 The autocorrelation function for the hourly precipitation data.

To be sure that the one hour time interval sampling is appropriate and our data are not redundant and not strongly correlated, we calculated autocorrelation functions for different parameters. Fig.4 shows the autocorrelation function for the hourly precipitation data. It shows that the data are not significantly correlated, and the data set is not significantly redundant. It also means that our training and test sets are independent.

These two data sets have been used for the NN training and testing/validation. As was noticed in the previous section, these data implicitly represent a stochastic parameterization that inherently contains an uncertainty ε, which is not a useless noise. However, from the point of view of a single NN trained using the data, the data (both component *X* and *Y*of the data) contain a significant level of noise.

Symbolically, the NN emulation of the stochastic parameterization (8) can be written as,

$$Y = M_{NN}(X) + \varepsilon + \varepsilon_{app} \tag{9}$$

whereM_{NN} is a NN emulation of the mapping *M* (8) and ε_{app} is a NN approximation error. Thus, in the case of the stochastic parameterization, the NN emulation task is different from one of emulating a *deterministic* original parameterization in GCM (e. g. Krasnopolsky et al. 2005, 2008a).

In our NN emulation approach applied to a deterministic or "exact" parameterization in GCM, the goal is to emulate the parameterization with a universal mapping (NN) as accurate as possible (Krasnopolsky et al. 2008a, 2010). We can do it because, in this case, we can produce the simulated data using the given "exact" parameterization (which actually is also a mapping) and consider the data as accurate ones (with no noise higher than the round off errors).

In the current work, the situation is completely different**. We do not have an expression (or computer code) for the mapping (8) that we want to emulate with NN (9**). We can only assume that it exists and, in this case, it is a stochastic mapping, which is

embedded into "pseudo-observations". **We cannot simulate "pseudo-observation" that we use for NN training directly.** We derive them using a rather complex data processing ($(x, y) \Rightarrow (\underline{x}, \underline{y}) \Rightarrow (X', Y') \Rightarrow (X, Y)$) from the data (x, y) simulated by a very different exact mapping (SAM) (2). Here we cannot assume that "pseudo-observations" are accurate. T**he uncertainty and stochasticity are the essential conceptual features of our approach here. Without accepting and understanding these features we cannot properly interpret our results and use NN in CAM.**

The case of emulating a stochastic parameterization, which we consider here, is much closer to the task of solving a forward or inverse problem in satellite remote sensing, i.e. to the task of emulating a satellite retrieval algorithm using noisy empirical data (Krasnopolsky 2008b). This important difference should be taken into account when the NN approximation is trained, the approximation error statistics are analyzed and interpreted, and the NN architecture is selected. For example, **in the case of training, the usually used criterion of minimum of the root mean square error should be substituted by the requirement that the root mean square error should not exceed the uncertainty ε or,**

$$\frac{1}{N}\sum_{i=1}^{N}[Y_i - M_{NN}(X_i)]^2 < \varepsilon^2 \tag{10}$$

All NNs that satisfy the condition (10) are equally valid emulations of the stochastic parameterization (8). Actually, each of these NNs emulates a member of the family of

mappings that together represent the stochastic parameterization (8). **Therefore, all NNs satisfying (10) together – the entire ensemble of NNs – represent the stochastic parameterization (8).** It is clear now that any estimate of the magnitude of ε is of a paramount importance for our approach. We will attempt to derive such an estimate in the next sections.

2.3.2. NN architectures,NN trainingand validation

Selecting an emulating NN architecture includes two different aspects and types of decisions: (i) the selection of inputs and outputs and their numbers (*n* and *m* in (1)), which, as we have already mentioned, are determined by the availability of the variables in the GCM, and (ii) the selection of the number of hidden neurons (*k*in (1)) in the emulating NN, which is determined by many factors (the length of the training set, the level of noise in the data, the characteristics of conversions of the training and test errors, etc.).

Table 1 Different NN architectures (combinations of inputs and outputs) investigated in the paper. Tabs is temperature, QV is atmospheric moisture - vapor mixing ratio, W – vertical velocity, U and V are horizontal components of the wind vector, RelH is the relative humidity, Rad – the radiative heating/cooling rates, Q1C – the "apparent heat source", Q2 – the "apparent moist sink", Prec – precipitation rates, and CLD –cloudiness. Numbers in the table show the dimensionality of the corresponding input and output parameters. In:Out stand for inputs and outputs and show their corresponding numbers.

NN Architecture In:Out	**NN Inputs**							**NN Outputs**			
	Tabs	**QV**	**W**	**U**	**V**	**RelH**	**Rad**	**Q1C**	**Q2**	**PREC**	**CLD**
{2} – 47:40	26	21	-	-	-	-	-	21	18	1	-
{3} – 47:59	26	21	-	-	-	-	-	21	18	1	19
{4} – 87:66	26	15	-	23	23	-	-	26	19	1	20
{5} – 58:66	26	15	17	-	-	-	-	26	19	1	20
{6} – 81:66	26	15	17	-	-	23	-	26	19	1	20
{7} – 66:66	26	-	17	-	-	23	-	26	19	1	20
{9} – 84:66	26	15	17	-	-	-	26	26	19	1	20
{11} – 36:55	18	18	-	-	-	-	-	18	18	1	18

Table 1 shows several different (in terms of inputs and outputs) architectures that we have experimented with here. The major inputsfor different NN architectures are the vertical profiles of the following model prognostic and diagnostic fields: Tabs - temperature, QV - water vapor, RelH – relative humidity, U and V – horizontal wind components, W - vertical velocity, and Rad – radiation heating/cooling rates. The major outputsfor all NN architectures (except for the NN architecture {2}), are the vertical profiles (or vectors) of the following model prognostic and diagnostic fields: Q1C (the "apparent heat source"), Q2 (the "apparent moist sink"), PREC (precipitation rates, a scalar),and CLD (cloudiness). Notice that only for the NN architecture {2} the CLD vertical profile was not used as output.

Numbers in Table 1 show how many vertical levels of the corresponding profile have been included as inputs in the NN. Many profiles have zeros, or small constants, or very small values that are almost constant (their standard deviationsare very small) for the entire data set. Zeros and constants should not be included in inputs or outputs because (1) they carry no information about input/output functional dependence and (2) if not removed they introduce additional noise in training. As for small values that are almost constant, these small signals may be in some cases not a noise but very important signals; however, taking into account the level of uncertainty in the problem, information that these small signals may provide is well below the level of uncertainty and is practically useless. Moreover, some of these variables were included in training and no improvement was observed. If they are important, they should be normalized differently or weighted.

Next, the number of hidden neurons (HID) has to be selected. NNs with this number of hidden neurons will be used in comparisons of different architectures. We selected several architectures, varied HID, trained a corresponding NN, and tested it. Figs. 5 and 6 show results of these experiments for two output parameters, Q1C and PREC. The figures show NN errors on training and test sets for HID changing from 1 to 20 for architecture {3} (see Table 1).

It is important to understand that a NN training (a least square minimization) attempts to minimize the total $(\varepsilon + \varepsilon_{app})$ *that is the approximation error and the uncertainty (noise).* ***Because of very different statistical properties of these components, they can be approximately separated and roughly estimated using detailed information about the training and test statistics.*** This issue is discussed in more detail below.

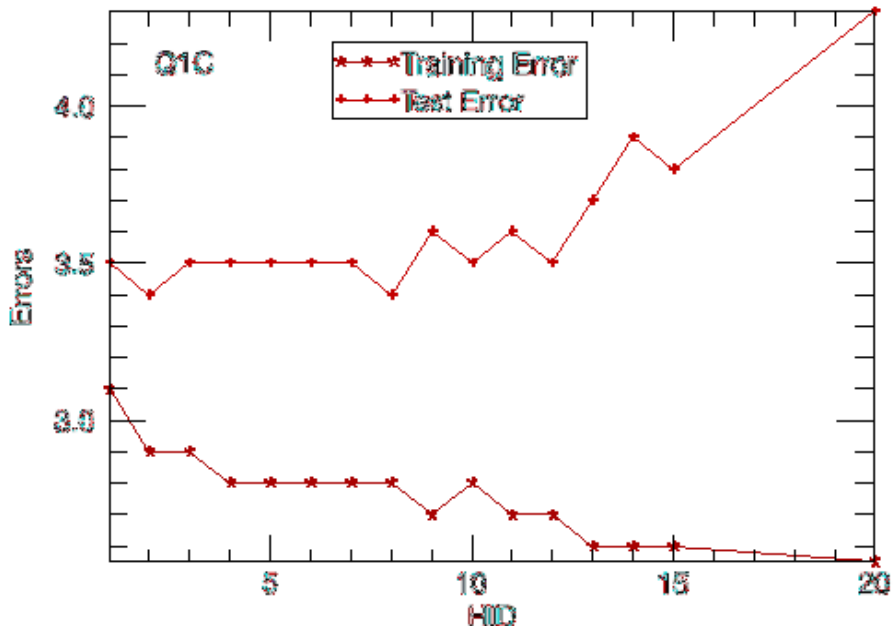


Fig. 5 NN approximation error on training (blue) and test (red) sets for Q1C.

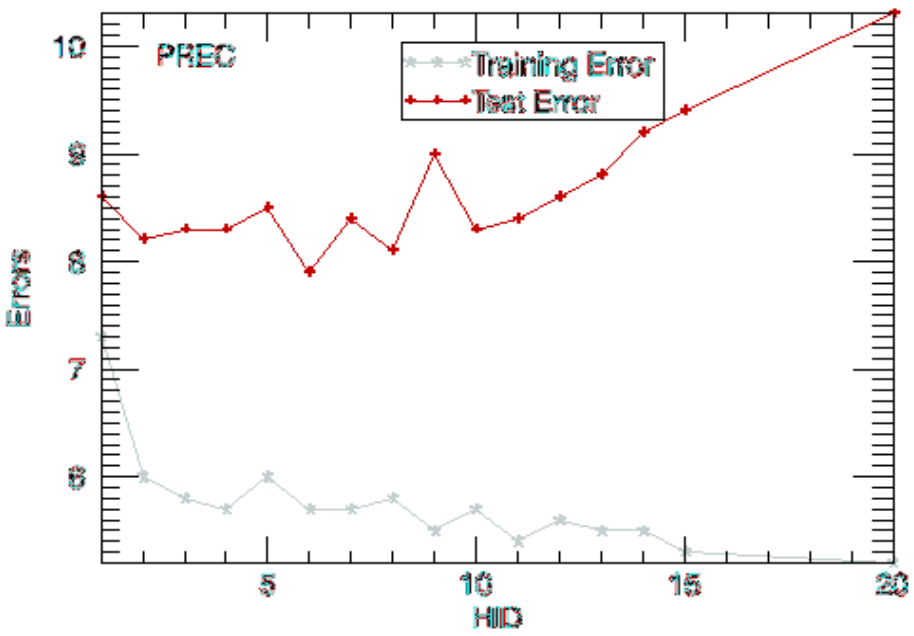


Fig. 6 Same as in Fig. 5 but for precipitation.

Figs. 5 and 6 demonstrate a classic situation that is usually observed when NN is trained using the data with a significant level of noise. The training error, after a sharp initial drop, decreases very slowly. The test error, after an initial drop, stabilizes and then increases. The interpretation of this behavior is well known. After the initial improvement of the approximation of the data due to an increasing flexibility of an approximating NN, a short interval of stability is reached (at HID ~ 3 to 7) when NN fits the signal but filters out the noise. Here the training error keeps decreasing; however, the test error is almost constant. Then with the increase of the flexibility of the approximating NN, it starts fitting the noise (the overfitting occurs). The training error keeps slowly decreasing; however, the test error quickly increases. Table 2 shows the number of fitting parameters (NN weights) in NNs with different HID, which were used for plotting figures 5 and 6. Taking into account that the training set contains 2240 records, it is not surprising that clearly pronounced overfitting is observed at HID > 10.

Table 2. The number of fitting parameters (NN weights) at different values of HID = k (see (1)).

HID	1	2	5	10	15	20
N_C	166	273	594	1129	1667	2199

The above results are presented for the NN architecture {3} (see Table 1); however, other architecture produce similar results. Thus, we can conclude that, for a particular simulation (data set) used, HID = 5 would be a good approximation for the number of hidden neurons in emulating NN. This value is inside of the interval of stability of the test error when the NN emulation fits the mapping (8) but does not fit the noise in the data. For that reason, comparisons of different architectures presented in the following subsection are performed for HID = 5.

2.4. Comparison of different NN architecturesand interpretation of NN training and test results

Table 3 presents comparison of different NN architectures defined in Table 1. NNs presented in the table were trained using the training set (Tr) and tested using an independent test set (Ts); both sets are described above in Section 2.3.1. For each NN and for each NN output variable, three statistics were calculated (bias, RMSE, and correlation coefficient) by comparison of NN generated output variables with the corresponding ones in the training or test set.

Table 3 Accuracy statistics for NNs with different architectures on training (Tr) and independent test (Ts) sets. CC is the correlation coefficient. HID = 5. For each variable and for each statistic the best are shown in red and the second best in green

Data Set	NN Arch.	NN Outputs											
		Q1C(K/day)			**Q2** (K/day)			**Prec**(mm/day)			**CLD** (fractions)		
		Bias	**RMSE**	**CC**	**Bias**	**RMSE**	**CC**	**Bias**	**RMSE**	**CC**	**Bias**	**RMSE**	**CC**
Tr	**{2}**	$3.\ 10^{-3}$	2.8	0.76	$3.\ 10^{-3}$	3.8	0.67	$2.\ 10^{-2}$	5.6	0.87	-	-	-
	{3}	$1.\ 10^{-3}$	2.8	0.75	$2.\ 10^{-2}$	4.0	0.63	$1.\ 10^{-2}$	6.0	0.85	$1.\ 10^{-4}$	0.07	0.91
	{4}	$2.\ 10^{-3}$	2.4	0.78	$2.\ 10^{-2}$	3.7	0.66	$1.\ 10^{-2}$	5.7	0.86	$2.\ 10^{-6}$	0.07	0.92
	{5}	$1.\ 10^{-3}$	2.3	0.81	$2.\ 10^{-3}$	3.7	0.68	$4.\ 10^{-3}$	5.2	0.89	$1.\ 10^{-4}$	0.07	0.92
	{6}	$2.\ 10^{-3}$	2.3	0.80	$1.\ 10^{-3}$	3.8	0.66	$2.\ 10^{-2}$	5.3	0.88	$3.\ 10^{-5}$	0.07	0.91
	{7}	$4.\ 10^{-4}$	2.3	0.80	$1.\ 10^{-3}$	3.8	0.64	$1.\ 10^{-3}$	5.3	0.88	$6.\ 10^{-5}$	0.08	0.89
	{9}	$2.\ 10^{-4}$	2.3	0.81	$3.\ 10^{-4}$	3.7	0.67	$7.\ 10^{-3}$	5.2	0.89	$5.\ 10^{-5}$	0.06	0.93
	{11}	$1.\ 10^{-3}$	3.1	0.73	$4.\ 10^{-3}$	4.0	0.64	$2.\ 10^{-2}$	5.8	0.86	$1.\ 10^{-4}$	0.07	0.90
Ts	**{2}**	-0.3	3.6	0.61	-0.6	4.9	0.46	-2.8	8.8	0.68	-	-	-
	{3}	-0.1	3.5	0.62	0.02	4.7	0.49	-1.1	8.5	0.68	0.03	0.11	0.81
	{4}	-0.6	3.5	0.62	-0.8	5.0	0.44	-5.1	10.6	0.66	**0.01**	0.11	0.81
	{5}	-0.5	**3.0**	0.70	-0.6	4.5	**0.53**	-4.0	8.8	0.73	0.00	0.09	0.86
	{6}	-0.1	2.9	**0.71**	**-0.1**	3.9	0.52	**-1.8**	7.8	**0.74**	**0.01**	**0.08**	0.87
	{7}	**-0.3**	2.9	0.70	**-0.1**	3.9	0.51	-2.6	8.0	**0.74**	**0.01**	**0.08**	**0.88**
	{9}	-0.4	2.9	0.73	-0.5	**4.3**	0.58	-3.3	7.9	0.77	0.00	0.07	0.92
	{11}	-0.7	3.8	0.65	-0.8	4.7	0.51	-4.1	8.6	0.76	0.01	0.10	0.84

All NNs presented in Table 3 have the number of hidden neurons HID = 5. It means that they all are inside the stability interval where NN fits well enough the signal in the training set and is not significantly responsive to noise in the data (see figs. 5 and 6). The *training errors* (Tr) for all output parameters are significantly less sensitive to the selection of the NN architecture and to the selection of HID inside the interval of stability (see figs. 5 and 6) than the test errors (Ts). Thus**, the training errors can be considered as a rough estimate of the noise in the data that is the inherent uncertainty of the stochastic parameterization (8)**.

Following this assumption, **the errors on the test set should be considered as a combination of the uncertainty (an estimate for it is provided by the training error) and an approximation error.** For example, for Q1C, for the architecture {6} the training error is 2.3 K/day and the test error is 2.9 K/day. Thus, assuming that the uncertainty and the approximation error are independent, i.e., that in (9),

$$(\varepsilon + \varepsilon_{app})^2 = \varepsilon^2 + \varepsilon_{app}^2$$

in the test error of 2.9 K/day, only 1.8 K/day can be attributed to the NN approximation error, and the residual 2.3 K/day should be attributed to the uncertainty, ε, of the stochastic mapping (8). If we perform such a correction for all statistics presented in Table 3, we find that, as in the aforementioned example, after the separation of the uncertainty (the training error) the NN approximation errors on the test set, in most of the cases, do not exceed significantly the uncertainty. In our case of NN emulation of a stochastic parameterization, the major criterion for evaluation of the NN emulation (10) (see the end of Section 2.3.1) is the similarity of the approximation error and the uncertainty. Thus, in accordance with this criterion, all NNs presented in Table 3 can be considered as equally valid emulations of the parameterization (4). **These NNs can be considered together as an NN ensemble emulation of the stochastic parameterization (8).**

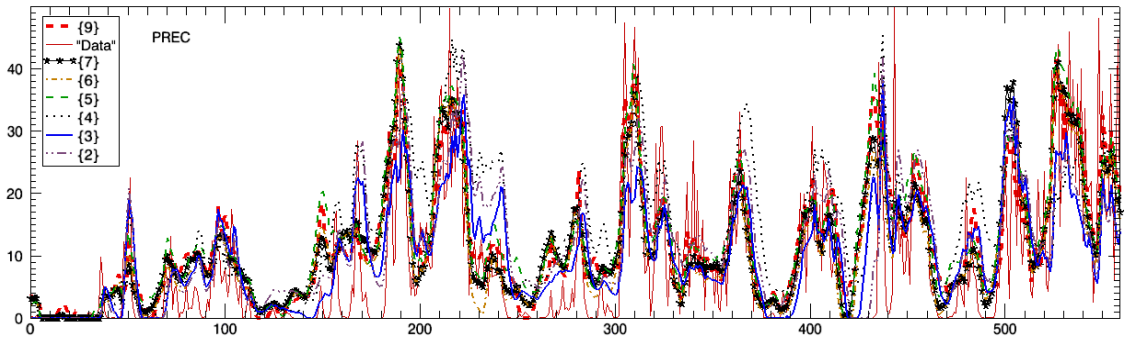


Fig. 7 NN simulations of precipitation on the test set. The legend describes different curves presented in the figure; the number in parentheses indicates the NN architectures described in Table 1.

Figs. 7 to 11 illustrate performance of NNs with different architectures on the independent test set. Fig. 7 demonstrates predictions of precipitation time series produced by different NNs in comparison with "pseudo-observations". These NNs produce an envelope (with a rather measurable spread) which on average gives a very good prediction of precipitation on the test set. The spread of the envelope shows that there is still a measurable difference between NNs with different architectures and some of the members of the envelope (e.g. {9}) give results that are closer to the "pseudo-observations". The magnitude of the spread reflects the uncertainty of the parameterization (8). It is noteworthy that the outliers in this precipitation plot are produced by NNs that donot use the large-scale vertical velocity as their input. It suggests that these neural net architectures are less valid than the ones that use W as an input.

Figures 8, 11, and 12 depict profile statistics for three other outputs of the NN parameterization: Q1C, Q2, and CLD. As in the case of precipitations, different NNs create envelopes with significant spreads for the mean, standard deviation, and RMSE profiles. For the standard deviation, the envelope is shifted with respect to the "pseudo-observation" profile. The variability of the NN standard deviations is less than that of "pseudo-observations" because the uncertainty of the parameterization (8) (its noise) is present in the data; however, each particular NN filters out a part of this noise. It is noteworthy that the RMSE profile includes both the uncertainty of the parameterization and the NN approximation error.

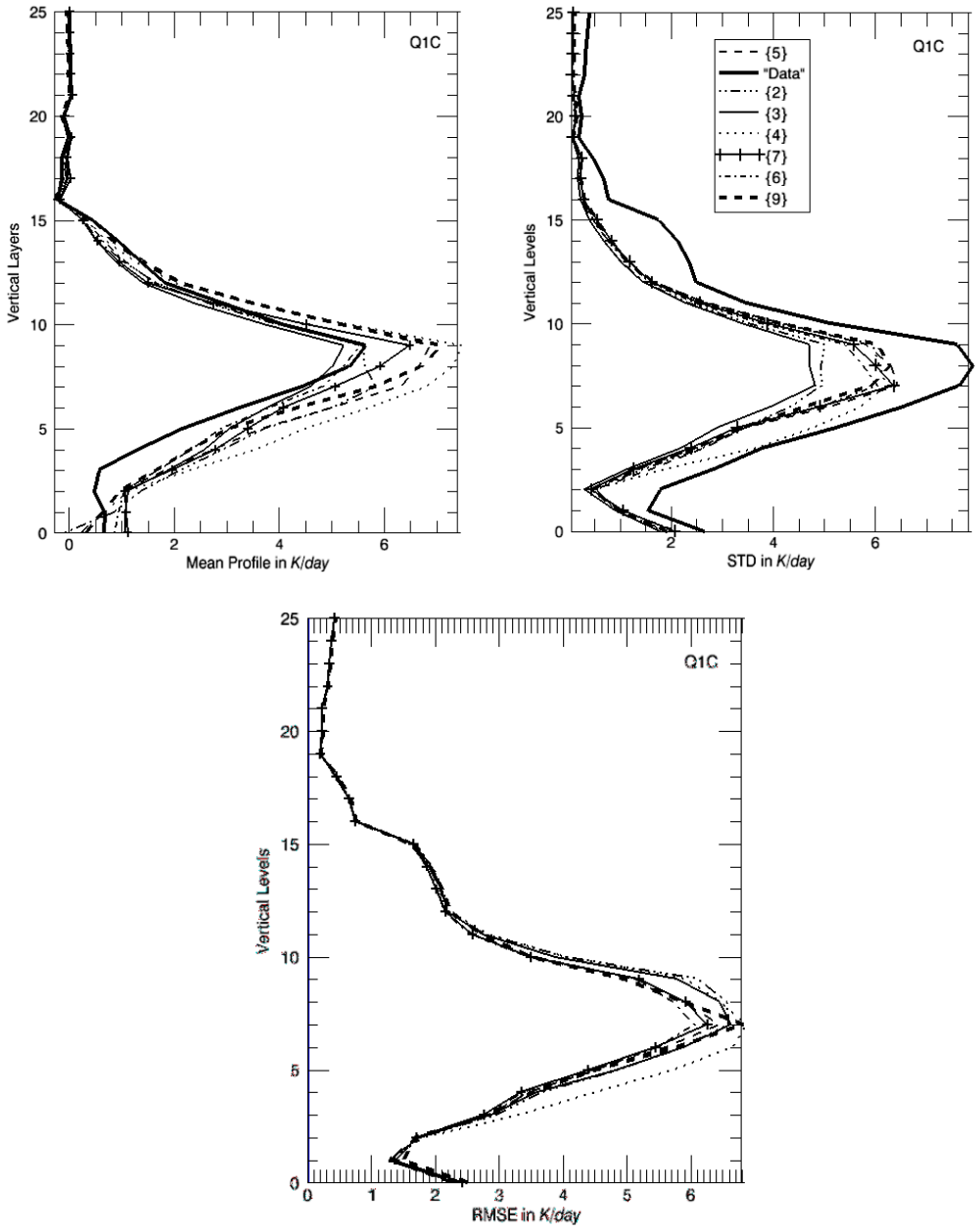


Fig. 8 Q1C (the apparent heat source from convection) statistics on the test set produced by different NNs: mean Q1C profiles (the upper left panel), profiles of Q1C standard deviations (the upper right panel) and RMSE profile(the bottom panel). The legend describes different curves presented in the figure; the number in parentheses indicates the NN architecture described in Table 1.

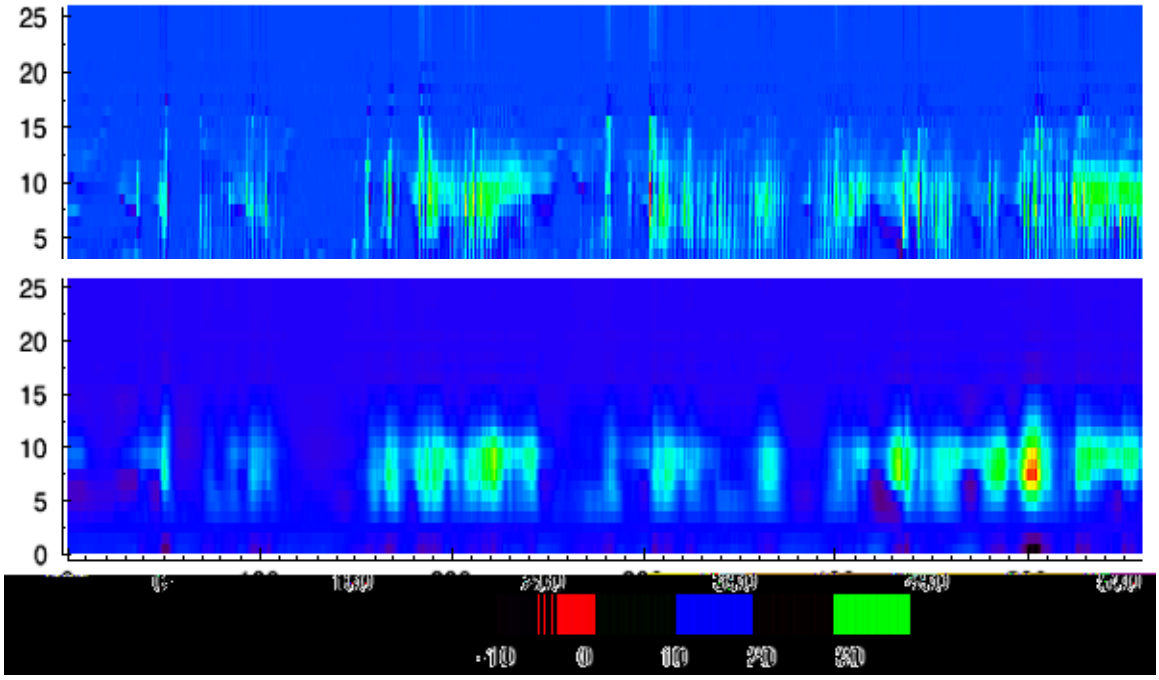


Fig. 9Hovmöller diagrams for Q1C profile time series: SAM “Data” – the upper panel, NN {9} – the lower panel.

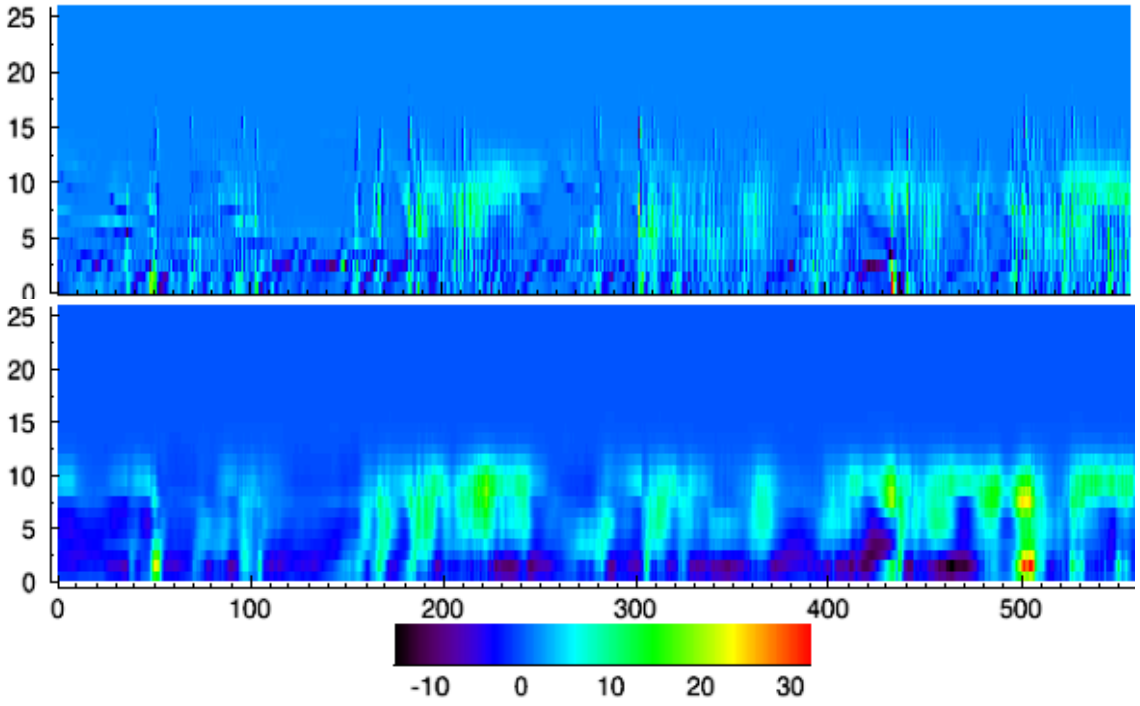


Fig. 10 Same as in Fig. 9 but for Q2 profile time series.

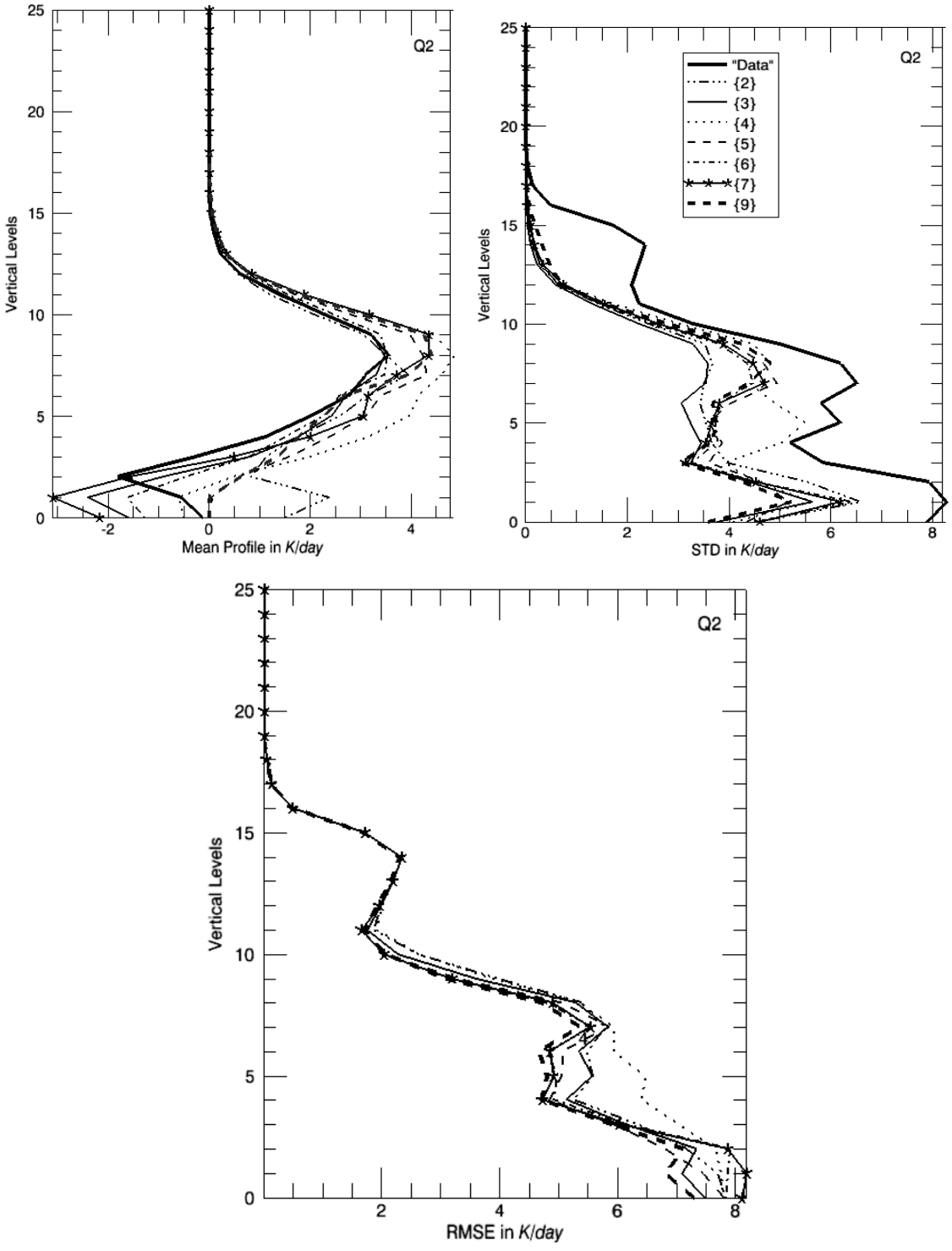


Fig. 11 Same as in Fig. 8 but for Q2.

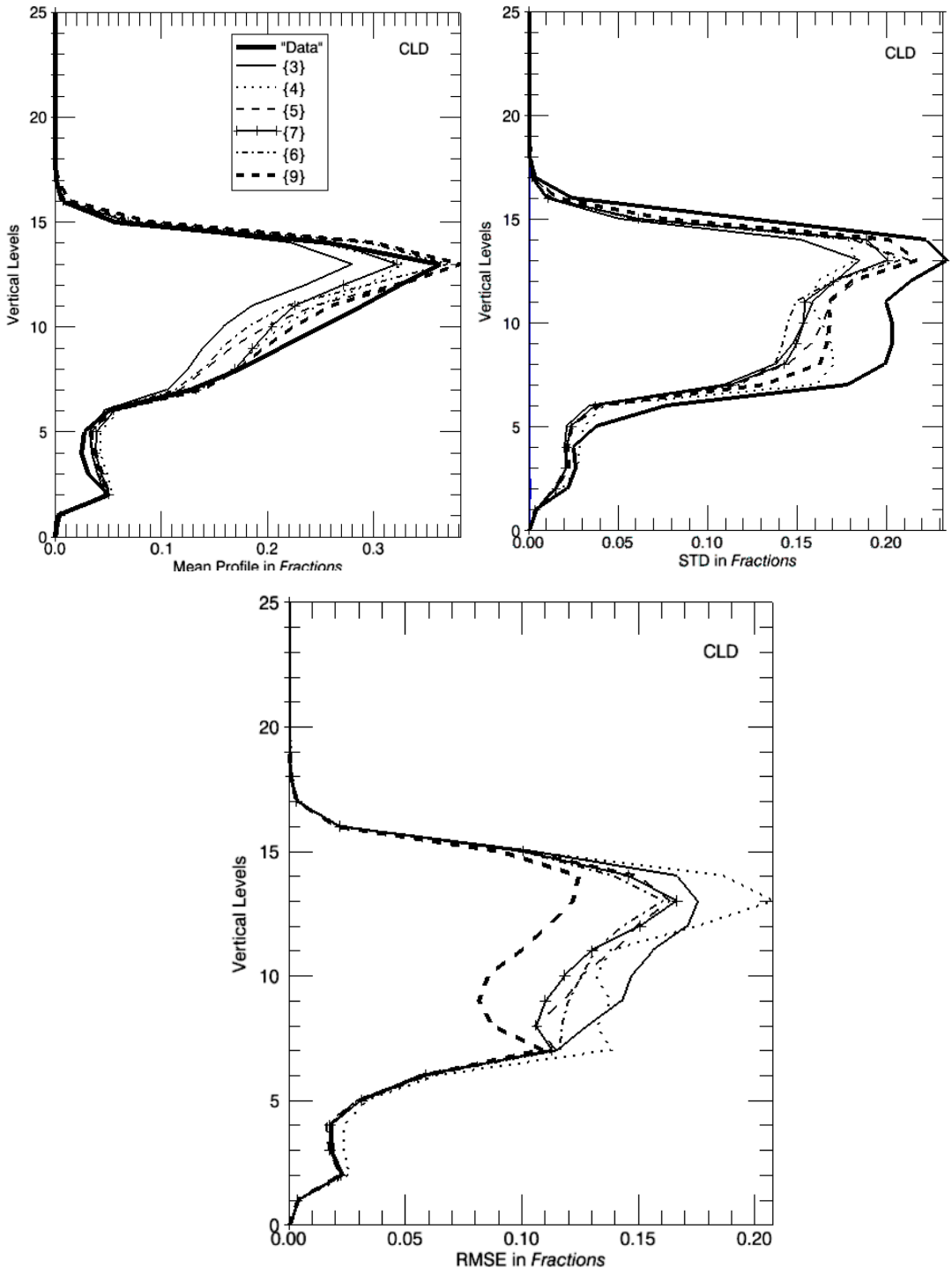


Fig. 12 Same as in Fig. 8 but for CLD

In general, the results show that the statistical structure of "pseudo-observations" and therefore of the parameterization (8) is well represented by envelopes created by different NNs. The spread of these envelopes reflects the magnitude of the uncertainty in the parameterization (8). The differences between the members of the ensemble inside the envelope are small as compared with the uncertainty; however, these differences are significant. They give estimates of the differences between members of the family of parameterizations determined by (8) and implicitly available in "pseudo-observations".

It is noteworthy that in Fig. 11, Q2 for NN{2}does not look physically correct in the upper left panel, namely it is positive in the lower 4 layers. However, NN{3}, which has the same inputs and an additional output – CLD profile, provides a much better physically meaningful Q2. This example illustrates the fact that, in the case of NN approximation, additional output may introduces additional information as well as additional inputs (Krasnopolsky 2007a).

Figs. 9, 10, and 13 shows the Hovmöller diagrams for the time series of Q1C, Q2 and cloudiness (CLD) profiles for a single member of the NN ensemble. The upper panels show "pseudo-observation" and the lower panels show the time series of profiles generated by NN with the architecture {9}. The patterns generated by NN are a bit smoothed, diffused; they are less sharp than the "observed" ones but well recognizable. NN represents the sequence of patterns well and without significant shifts.

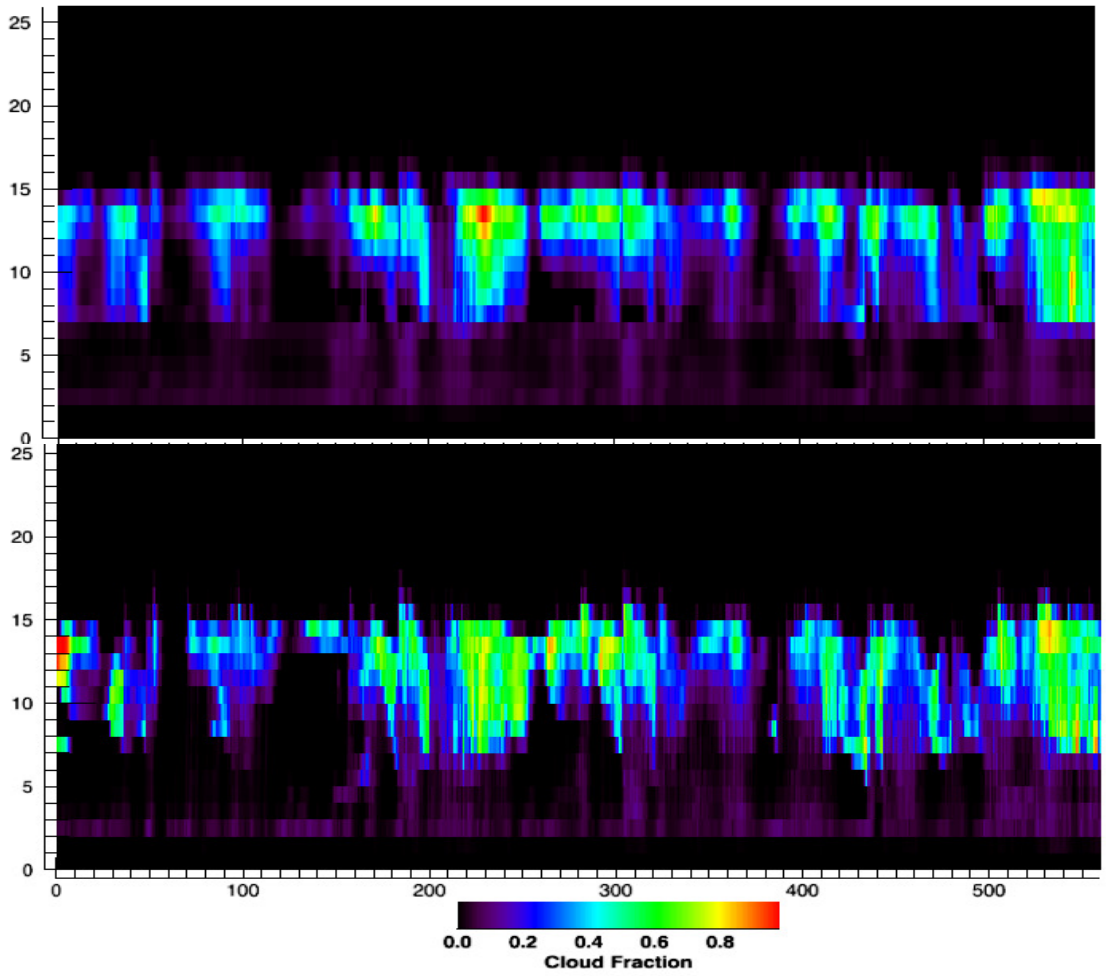


Fig. 13 Same as in Fig. 9 but for CLD profile time series.

We can conclude that Ts errors calculated on the independent test set, which are presented in Table 3 and in figures 7 to 13 cannot be considered as an estimate of the NN accuracy only; they contain a contribution of the parameterization uncertainty $\mathcal{E}$ and should be adjusted as in the example above (see the discussion after Table 3). Another conclusion is that the errors on the test set are more sensitive to the NN architecture; however, the variations of errors for different NN architectures (the spread of the envelope created by different NNs) represent the level of noise in the data or the uncertainty of the stochastic parameterization (8). It means that, in the context of the current application (development of NN emulation for a stochastic convection

parameterization (8)), selecting the best architecture for emulating NN followed by the subsequent use of this "optimal" NN parameterization in GCM, is not the best approach. All NNs presented here (as well as other NNs with the same architecture and the same number of neurons but with different initializations, which we will describe below) can be considered as equally valid emulations of the parameterization (8). They **should be rather considered as members of a particular NN ensemblerealization of the stochastic parameterization (8) represented by a particular data set. Thus, every single NN emulation can be considered as a particular realization of this stochastic parameterization. The NN ensemble parameterization can be used in GCM in several different modes** (Krasnopolsky 2007b, Krasnopolsky et al. 2008c).

We considered above an NN ensemble created by NNs with various architectures. For validation in CAM we excluded the stratospheric levels and used only the tropospheric levels. The TOGA-COARE observations, used for driving SAM, are mostly available for the tropospheric domain, i.e., for the 18 lower levels of CAM. For this reason the architecture {11} (and its ensemble with different initializations), which is similar to the simple architecture 3 but containing only the 18 tropospheric levels, has been trained and used for our experimentation with CAM below (see Section 3).

For testing in NCAR CAM we created an ensemble of ten NNs with the architecture {11} (36 inputs and 55 outputs) with 5 hidden neurons in each, which represents the stochastic parameterization (8) as well. All input and major output variables for the architecture are directly available in CAM for using and comparison. The coefficients (weights) of these

NNs have been initialized with different set of small random numbers before the training. As a result, all these NNs have different weights because the training process converged to different local minima. For these NNs, the errors on training and test sets are similar to those shown in Table 3. They also demonstrate the spread similar to that of for NNs with different architecture shown in Table 3. There exist different methods of creating NN ensembles (Krasnopolsky 2007b) which could be applied also in this but we have chosen a simplest approach.

2.5. Evaluation of CAM vs. SAM differences

In the previous sections we introduced an ensemble of NNs that represented the stochastic convection parameterization (8). These NNs have been derived from the data generated by SAM forced by TOGA-CORE data. We cannot expect that these NNsderived using SAM simulations, would generate output variables similar to those of CAM. These two models create two different "virtual realities" with different temporal and spatial scales and resolutions, with different representations of dynamics and physics; they have different boundary and initial conditions and different forcing. To illustrate and estimate the differences between CAM and averaged SAM simulated data, we ran CAM during the same 120 days TOGA-CORE period and used CAM time series of the convection parameterization inputs and outputs (T, QV, CLD, PREC, etc.) to compare the results with the "pseudo-observations" (averaged SAM data) and estimate CAM vs. SAM differences.

We first made the comparison of the CAM and averaged SAM "virtual realities" in terms of T and QV (the convection parameterization inputs). The significant differences

between CAM and averaged SAM simulated data and derived from them convection parameterization inputs and outputs presented below are not surprising because the models' formulations are different, especially for the stratospheric domain, and SAM is driven by observational meteorological data whereas CAM is simulated in a free mode, with only SST forcing.

Table 4 Bulk statistics for temperature and moisture variables, T and QV, in CAM and in SAM

		Mean	**Standard Deviation**	**Min**	**Max**
T *in K*	CAM	240.6	36.8.	180.7	301.3
	SAM	240.4	35.5	190.5	301.2
QV *in g/kg*	CAM	3.51	5.58	$4.\ 10^{-4}$	19.34
	SAM	3.86	6.08	0.	20.66

Table 4 shows bulk statistics for two time series of profiles of the T and QV in CAM and in averaged SAM simulated data. The bulk statistics are very close for both series with the exception of minimal temperature, which is 10 K cooler in CAM than in the averaged SAM data. However, if we look closer at the data profiles, we will find out that the models' data differ even more significantly. Fig. 14 shows mean temperature profiles for CAM (dashed) and averaged SAM (solid) data. There is a significant temperature differences for the upper 10 layers (above 100 hPa, i.e., throughout the tropical stratosphere); also the variability of the temperature is vertically distributed very differently in the two models.

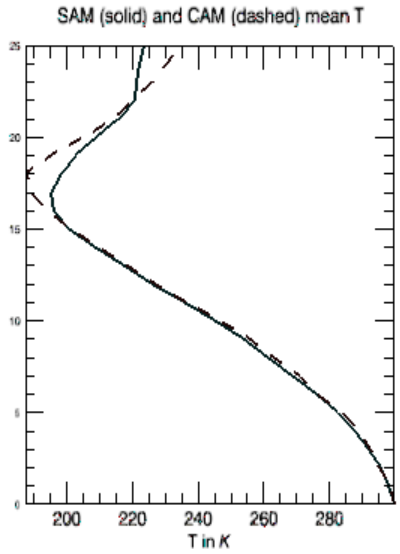


Fig. 14 Mean temperature profiles for averaged SAM (solid) and CAM (dashed) data.

Fig. 15 shows mean moisture (QV) profiles for CAM (dashed) and averaged SAM simulated (solid) data. There are significant moisture differences at the lower layers; also the variability of the moisture is distributed differently in two models. Fig. 16, which depicts the Hovmöller diagrams for QV for the CAM and averaged SAM profile time series. It shows that the patterns are quite different for CAM and SAM (shown in the lower and upper panels, respectively) and also desynchronized in time (what is not surprising due to the fact that the models are driven by different forcing).

Finally, Fig 17 reveals significant systematic differences for T and QV profiles between the two models. Consistently with Figs. 14, 15 and 16, T differences are large, up to 10-12 K by magnitude, for the stratospheric domain, and QV differences are large, up to 3 g/kg by magnitude, for the lower tropospheric domain.

The differences in the stratosphere result from the CRM's forcingsthat do not represent stratospheric processes. The GCM contains a Brewer-Dobson circulation which induces a dynamical cooling that balances the radiative heating in the tropical tropopause layer (between about 14 and 19 km). The forcings for the CRM do not include this, so that the tropopause sinks and the TTL (Tropical Troposphere Layer) warms so that increased radiation to space will balance the energy budget.

The large differences in the stratospheric domain should not affect the NN convection parameterization, which works mostly in the tropospheric domain. This is a justification for using the NN architecture {11} (using only the tropospheric levels; see Table 1) for experimentation with CAM below.

From the bulk statistic presented in Table 4 and from Figs.14 –17 we can conclude that **the CAM simulated and average SAM simulated data (our "pseudo-observations") are located in areas/domains in T – QV space that have similar extents (although min T is shifted by 10 K); however, the shapes of the areas/domains are different.**

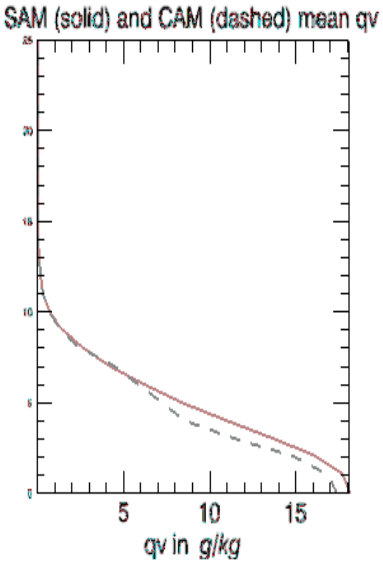


Fig. 15 Mean moisture (qv) profiles for averaged SAM simulated (solid) and CAM simulated (dashed) data.

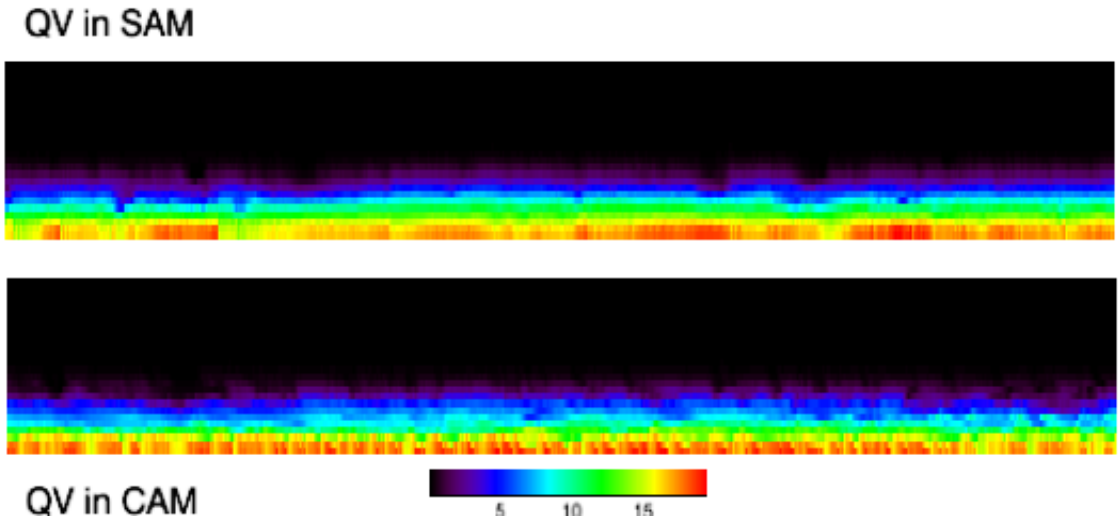


Fig. 16 Hovmöller diagram for the moisture (QV) profile time series in averaged SAM simulated (the upper panel) and in CAM simulated data (the lower panel).

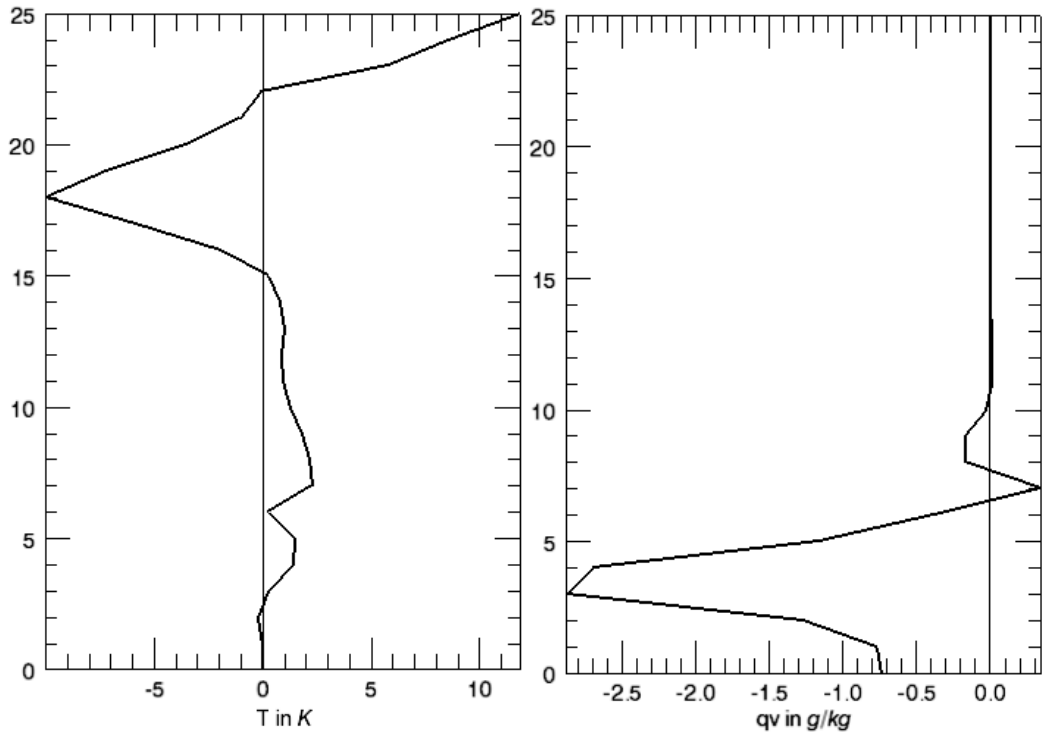


Fig. 17. Differences (CAM simulated – averaged SAM simulated profiles) for T (the left panel) and QV (the right panel).

One interpretation of the SAM-CAM differences shown in Fig. 17are that the CRM has more shallow and mid-level convection than the GCM. This would lead to a drier layer close to the surface. A second interpretation could just be that the model drifts dry at low levels during the run.

In a similar way we investigated the differences between CAM simulated and averaged SAM simulated data for other variables considered as prospective inputs and outputs for convection parameterizations. For example, we evaluated CAM vs. averaged SAM differences for the cloudiness (CLD) and precipitation (PREC) that are outputs of the convection parameterization. The bulk statistics shown in Table 5 demonstrate that CAM and the averaged SAM data are different; there are differences for both CLD and PREC

and variability of these parameters are different. For PREC even the extent of the range is different. It means **that domains for CAM and the averaged SAM data are shifted and their shapes are different.**

Table 5 Bulk statistics for CLD and PREC in CAM simulated and averaged SAM simulated data.

		Mean	**Standard Deviation**	**Min**	**Max**
PREC *in mm/day*	CAM	6.41	7.23.	0.	43.5
	SAM	9.22	11.24	0.	80.8
CLD *in fraction*	CAM	0.159	0.256	0.	1.00
	SAM	0.072	0.154	0.	1.00

The mean CRM precipitation is largely determined by the forcings. That would be the origin of the mean differences shown in Table 5.

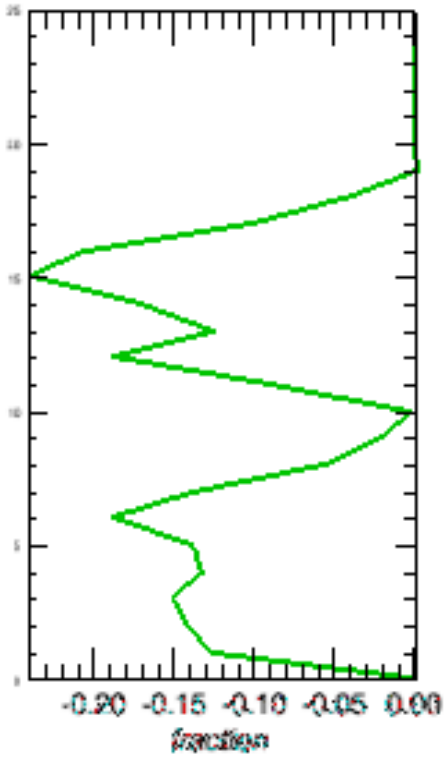


Fig. 18.Averaged SAM vs. CAM differences' profile for CLD.

Fig. 18 shows significant systematic differences for CLD mean profiles between the two models.

In this section we estimated the differences between simulated CAM variables and the corresponding averaged SAM simulated variables. These differences have been estimated as profiles for 3D variables (e.g., T, QV, CLD) or as a number for 2D variables (e.g., PREC) over the one TOGA-COARElocation for the four-months period (November 1992 – February 1993) of available data (the TOGA-COAREwinter, NDJF).

It is noteworthy that the aforementioned differences can be, in principle, used for parameterization calibration aimed at a potential reduction of errors in CAM simulations. Further comments on this issue will be made in the following sections.

3. Validation of NN Convection Parameterization in NCAR CAM

The NN stochastic convection parameterization described in the previous sections is implemented as the ensemble of NNs with the architecture {11},which are trained on the averaged SAM simulated data ("pseudo-observations"). In this section, we introduce the NN stochastic parameterization into CAM. Here our goal is to verify whether the NN ensemble, emulatingthe stochastic convection parameterization (8),provides meaningful/realistic outputs when using CAM inputs. We performed the validation of our stochastic NN convection parameterization in the following experiment.

Over the large Tropical Pacific region (with the area size of 120° x30° and the following coordinates: 150° E <lon<115° W; 15° S <lat< 15° N) we performed the parallel runs with the standard CAM and with the *diagnostic* CAM-NN run (see below) for the TOGA-COARE winter (November 1992 – February 1993, NDJF), the 4-month period for which our NN convection had been actually trained.

For a simple initial testing and validation of the NN stochastic convection parameterization (8) in CAM, we introduced a diagnostic mode of integration. For the diagnostic mode, a standard CAM run is performed but at every grid-point/profile and at each time step the NN ensemble is calculated using the CAM inputs and producing NN ensemble outputs without any feedback into the continuing CAM integration. Hereafter this diagnostic run is called CAM-NN.

For the diagnostic mode of integration, every time the NN convection parameterization is applied, all ten NNs (the NN {11} ensemble members) are evaluated and averages of their outputs are calculated and used as NN ensemble convection parameterization outputs.

This CAM-NNrun using the diagnostic mode is instrumental in the sense that it allows us to produce an initial validation of the developed NN stochastic convection parameterizations. Testing NN parameterizations in a prognostic mode (i.e., feeding back the outputs of the NN convection parameterization into a CAM run) will be done at a

later stage of our development, when "pseudo-observations" will be generated by CRM/SAM driven by CAM simulated data (see discussion in Sections 4 and 5).

3.1Validation of NN convection parameterization using NCAR CAM for the TOGA-COARE location and period (winter of 1992-93, NDJF)

At the first step of our validation, the outputs generated by the ensemble of ten NNs {11} have been compared with CAM data for one grid point at the TOGA-CORE location (- 2° S, 155° E) duringthe TOGA-COARE winter, November 1992 - February 1993. Thus, the CAM data were collocated in space and time with the averaged SAM simulated data.

For CAM-NN we used the CAM generated T and QV as inputs for the ensemble of NN {11}trained on the averaged SAM simulated data.

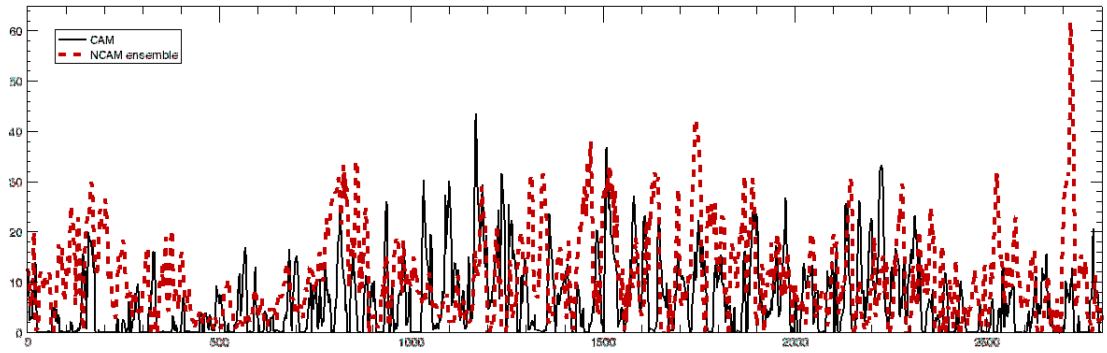


Fig. 19. Precipitation (PREC, in mm/day) time series for CAM (black solid) and CAM-NN (the NN ensemble mean) (red dashed).

Fig. 19 shows PREC time series produced by the original CAM and the NN ensemble mean within CAM-NN. As it was discussed above (see sections 2.4 and 2.5), the time series are not synchronized; however, the scope, the mean, and the frequencies of the time series are quite similar and look reasonable.

Table 6 Bulk statistics for CLD and PREC outputs for SAM, CAM and CAM-NN

		Mean	**Standard Deviation**	**Min**	**Max**
PREC *in mm/day*	SAM	9.22	11.24	0.	80.8
	CAM-NN	8.50	8.14	0.	63.6
	CAM	6.41	7.23.	0.	43.5
CLD *in fraction*	SAM	0.072	0.154	0.	1.00
	CAM-NN	0.104	0.240	0.	1.00
	CAM	0.159	0.256	0.	1.00

Table 6 shows bulk statistics for CLD and PREC variables for CAM, CAM-NN, and SAM. The CAM-NN statistics is rather close to the statistics of SAM and deviate from that of CAM.

Let us stress that we cannot expect full similarity here between CAM-NN and CAM statistics, profiles, and time series. The CAM-NN results are generated by our NN convection parameterization derived from CRM/SAM cloud physics, which is different from the cloud physics implemented in CAM

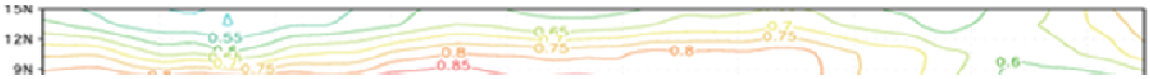
In the next section, we will discuss an extension of our diagnostic tests beyond the TOGA-COARE location, and will analyze the results of the performed parallel CAM and CAM-NN simulations over a large Tropical Pacific region.

3.2Validation of NN convection parameterization forparallel CAM and CAM-NN climate simulations for the TOGA-COARE winter overa Tropical Pacific region

As described above, the developed stochastic NN convection parameterization has been introduced into CAM-NN and run in the aforementioned diagnostic mode, for which CAM inputs have been used for calculating NN convection outputs. In this section, we will compare the parallel CAM-NN and CAM simulations for the 4-months TOGA-COARE period (November 1992 to February 1993) and validate them against the NCEP reanalysis over a large Tropical Pacific region.

Although the NNconvection parameterization has been developed for theTOGA-COARE location (marked by a star in the middle panels of Figs. 20 and 21), we decided to apply the NN convection in the CAM-NNrun for the entire large Tropical Pacific region (with the area size of 120° x30° and the following coordinates: 150° E <lon<115° W; 15° S <lat< 15° N). The stochastic NN convectionparameterization was applied at every grid point and each time step throughout the entire TOGA-COARE winter simulation for the diagnostic CAM-NN run. This is a very tough/hard test for the NNgeneralization ability.

The total cloudiness (CLD) and precipitation (PREC) produced in these parallel climatesimulations for the Tropical Pacific region for the TOGA-COARE period are shown in Figs. 20 and 21.


15N
12N
9N
0.55
0.75
0.85
0.8
0.75
0.8
0.75
0.6

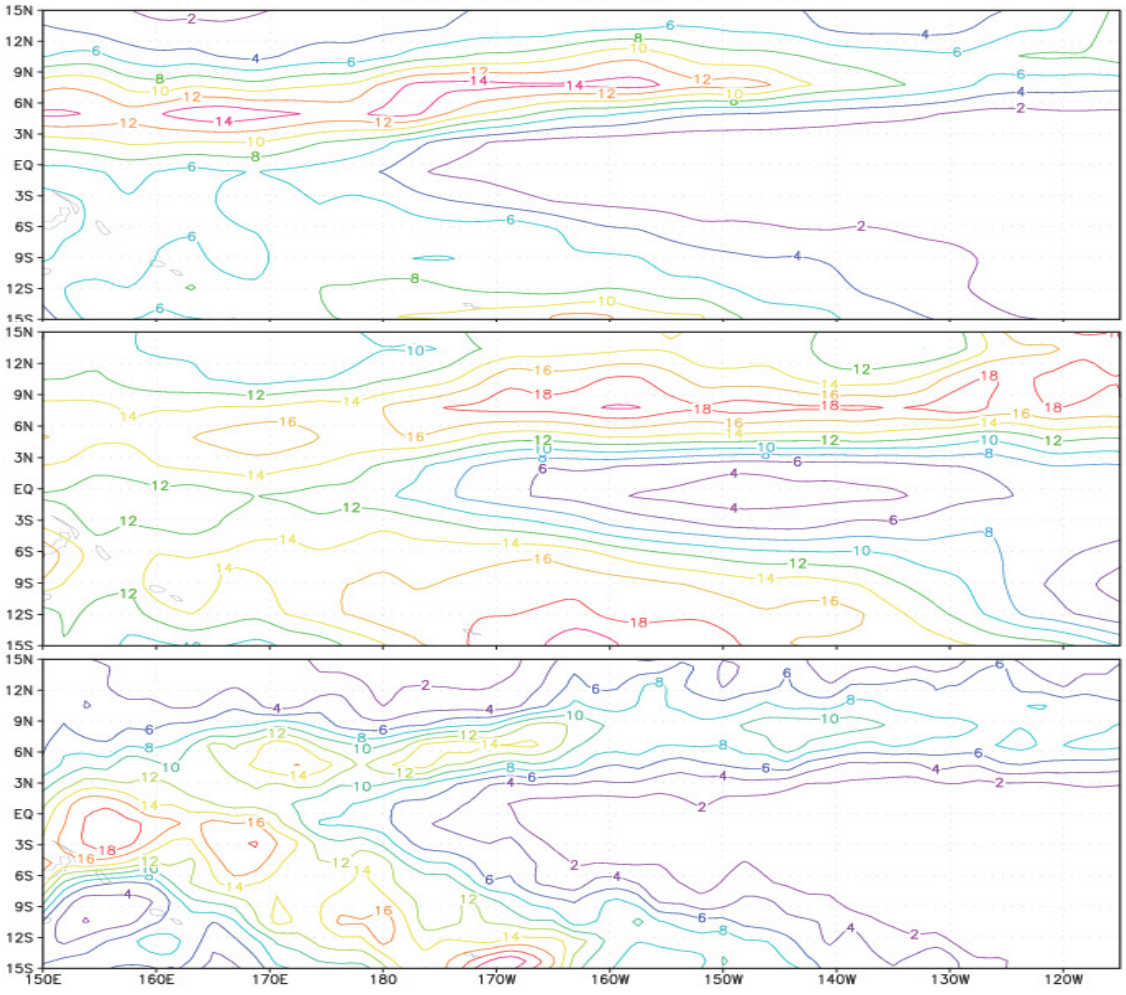


Fig. 21 Same as for Fig. 20 but for total precipitation (PREC, in mm/day). The contour interval is 2 mm/day.

Precipitation patterns for the CAM and CAM-NN runs are overall similar to each other and both are smoother than the precipitation pattern for the NCEP reanalysis. The PREC magnitudes for the CAM-NN run are overall larger than for theCAM run. Compared to the NCEP reanalysis, the PREC magnitudes for the CAM-NN run show an overestimated minimum around the Equator and 150º W and in the areas north of the Equator whereas for the CAM run the PREC magnitudes are underestimated at the Equator east of the aforementioned minimum and in the areas south of the Equator.

The major result is that the regional PREC and CLD distributions presented for the parallel CAM and CAM-NN TOGA-COARE runs show a general pattern consistency and similarity between the CAM-NN and CAM runs and also with the NCEP reanalysis used for validation. The magnitudes for the parallel runs show some systematic differences though.

It is not surprising that there are some differences between the parallel runs and the NCEP reanalysis. When analyzing the results we have to keep in mind that the region is a sparse data area so that the NCEP reanalysis there is significantly (NCEP) model dependent. However, the existing sparse conventional data and satellite data altogether provide valuable information so that the NCEP reanalysis can be definitely used for validation of climate simulations.

We would like to emphasize that at this initial stage of our development of the stochastic NN convection parameterizations, it seems reasonable to compare the CAM and CAM-NN runs in terms of their general consistency between themselves and with the NCEP reanalysis. A detailed climatological analysis of regional and global simulations for all seasons will be done at the next stage of our development. It will be based on using SAM simulations driven by CAM simulated data forcing for developing stochastic NN convection parameterizations for CAM.

Discussion of results presented in this section exemplifies the uncertainty of cloud and precipitation simulations for both parallel CAM and CAM-NN runs. It also underlines the

complexity of analysis and validation of climate model simulations and the limitations of data/information for the tropics that can be reliably used for validation.

In our view, the initial results obtained for the CAM-NN simulation for the TOGA-COARE winter are overall positive and encouraging. They support the validity/soundness of the NN approach for developing stochastic NN convection parameterizations for climate models. Evidently, a significant future effort is needed for practical implementation NN convection parameterizations into climate models.

4. Discussion

In our view, the results presented above in Sections 2 and 3 demonstrate a realistic potential of the presented NN ensemble approach for developing stochastic NN convection parameterizations. Our first attempt in this direction led to meaningful results despite the fact that for our development we used a limited amount of data available over a small area in the Tropical Pacific Ocean (TOGA-COARE) during a limited four month period (the TOGA-CORE winter). We obtained encouraging and physically meaningful results for cloudiness and precipitation not only over this particular TOGA-COARE location and period but also over the extended Tropical Pacific Ocean region.

A careful analysis of results presented in Sections 2 and 3 reveals two major challenges that our approach faces, both related to "pseudo-observations" (averaged CRM/SAM simulated data) used for NN training:

1. Data for initialization and forcing CRM are available only over a few sites (mostly TOGA-COARE and ARM); thus, data simulated by CRM initialized and driven by

observations are not representative enough in terms of different global geographical locations and different weather conditions.

2. There exist significant differences between the averaged data simulated by CRM initialized and driven by observed data (these "pseudo-observations" have been used for training the NN ensemble) and CAM simulated data.

The logical way of dealing with these challenges would be to use for creating "pseudo-observations" the CRM/SAM simulations initialized and driven by CAM simulated data forcing. We consider it to be a major effort of our future developments. Such a setup has several important advantages, which will allow us to meet aforementioned challenges. Specifically, it may allow us to:

1. Use many grid points in the global CAM domain to run CRM/SAM, which will extend the geographical and temporal representativeness of the "pseudo-observation" training set, to create a "global" training set for improving performance of our NN convection parameterization in different locations under diverse weather conditions for all seasons. "Pseudo-observations" generated in such a manner are synchronized with CAM simulations in time.
2. Decrease or hopefully practically eliminate the systematic differences between averaged CRM/SAM simulated data and CAM simulations.

As a result we will be able to develop the NN convection parameterization that is more consistent with other CAM components. Achieving such a consistency is important for

including the NN convection parameterizations in a prognostic mode for CAM-NN simulations.

5. Conclusions

In this paper we introduce a novel approach to development of NN convection parameterizations based on applying the NN ensemble technique. This approach has been conceptually developed and formulated. Several very important notions are introduced which constitute the conceptual skeleton of the approach:

1. Pseudo-observations are the result of averaging and projecting of high dimensional and high resolution CRM data onto the GCM space of atmospheric states. Pseudo-observations contain uncertainty which is a result of averaging and projection of original CRM/SAM simulated**.**
2. Stochastic mapping/parameterization that is implicitly defined by pseudo-observations with uncertainties
3. NN ensemble emulation that is an adequate tool for emulating stochastic mappings/parameterizations.

This approach is used for development of a NN ensemble stochastic convection parameterization for climate models. This fast parameterization is built based on data from CRM simulations initialized and forced/driven with TOGA-COARE data. The CRM/SAM (Khairoutdinov and Randall, 2003) provided by M. Khairoutdinov, has been used for CRM simulations. SAM simulated data were averaged and projected onto the GCM space of atmospheric states to implicitly define a stochastic convection parameterization. Next, these data ("pseudo-observations") were used to emulate the stochastic convection parameterizationusing an ensemble of neural networks (NN)

defined by different choices of input and output variables and different initializations. An ensemble of NNs has been trained and tested. The inherent uncertainty of the stochastic convection parameterization derived from "pseudo-observations"has been estimated.

The comparison of two parallel CAM and CAM-NN climate simulation: for the 4-month TOGA-COARE period revealed significant systematic differences over the TOGA-COARE location between CAM and the averaged SAM simulated variables and between their domains. These systematic differences are much larger in the stratosphere.

An ensemble of NNs trained on averaged SAM simulated data has been introduced into CAM. It produced physically meaningful NN ensemble outputs, which demonstrates the robustness of the NN approach. It led to producing realistic climate distributions of cloudiness and precipitation for the TOGA-COARE period for the Tropical Pacific Ocean region (and to some degree, for the Tropical Indian Ocean region). These results have been obtained for the CAM simulation run in a diagnostic mode using the NN convection parameterization trained over one TOGA-COAR location. Basically, we have shown that climate simulations for CAM and CAM-NN for the TOGA-COARE period, and the NCAR reanalysis are quite similar. Specifically, the CAM-NN simulation for the TOGA-COARE period is physically meaningful. In our opinion, these results demonstrate a realistic potential of the NN ensemble approach for developing stochastic NN convection parameterizations for climate and NWP models.

It is noteworthy that we should not expect a full similarity between CAM-NN and CAM simulated fields and statistics. The CAM-NN results are generated by our NN convection parameterization derived from CRM/SAM cloud physics, which is different from the cloud physics implemented in CAM.

The existing approaches discussed in Section 1.2 require very large increases in the computational cost of current models. Their advantage is transparency, but we believe the computational cost is so high that it makes them impractical for most problems with present day computers. We believe that an ensemble of NNs, trained on the output from more realistic and comprehensive representations of convective processes over a variety of regimes could serve as a practical convective parameterization for global and regional models. The NN ensemble can be used in global models in several different ways (Krasnopolsky et al. 2008c). The results presented here suggest that the NN ensembles are able to accurately represent the behavior of a realistic model of convection and are robust enough to generate physically meaningful results when introduced into a global model. We expect that they will be able to produce a useful parameterization for a variety of application in weather and climate science. The NN convection parameterization provides an opportunity to capture many aspects of the more realistic representation at a fraction of the cost of the alternatives.

Our future plans include:

1. running SAM simulations initialized and forced by CAM simulated data (the most significant step of our future development) to:

   - drive SAM with forcing from a broader range of regimes.

- cover longer time periods, more geographic locations, and more diverse weather conditions

2. performing longer and more representative, in terms of geographic locations, simulations using a MMF, or later GCRM framework.Archiving results from SAM simulations in a global framework to provide a more varied/diverse set of model output to be used as NN training dataset.
3. testing the NN convection parameterization trained using these new data in CAM through performing parallel CAM CAM-NN runs, the later in both diagnostic and prognostic modes.

There are also some consequences to the approach. We are aware of the following issues:

1. We will have abstracted the physical description of the processes driving the system by another conceptual layer. A simple example of this situation can be seen when considering the possible influence of aerosol/cloud interactions. Unless we train the NN using aerosol distribution as 'inputs', and vary those inputs during the training process, we cannot build a parameterization that is sensitive to those processes. Note that many convective parameterizations today are also unable to handle these processes. We must of course have trained the NN parameterization on the 'right' inputs, and have sampled the phenomena of interest over the whole range of parameter space that the NN is expected to perform realistically over. The NN parameterization also needs to predict situations where no convection will occur, and adapt to different convective regimes (like shallow and deep maritime convection, mid-latitude frontal convection, continent mid-latitude summertime convection,

continental deep convection). The parameterization must be able to seamlessly transition from one regime to another, or at the very least recognize where it should not be operating. Several solutions for this problem have been proposed and discussed by Krasnopolsky (2007a, b), and we will further explore them. Again, most current parameterizations have this same constraint.

2. It may be difficult to interpret the model response to variations in the important processes. For example it may be difficult to attribute increased precipitation to a particular process (e.g. a change in accretion processes in rain or snow falling from anvils) because those processes are not explicitly available within the parameterization to monitor. It should be possible to design an NN to output this information.
3. Probably, the NN convection parameterization might be best-suited as a drop-in replacement for only the deep convection parameterization (or perhaps the deep and shallow convection parameterizations). This is based in part on a comment that Martin Miller of ECMWF made once that coarse-resolution superparameterization (or MMF) might be best-suited as a replacement for the deep convection scheme in a GCM, rather than as a replacement for all of the turbulence/cloud schemes.
4. The NN parameterization has one distinct advantage over superparameterization: it can use high-resolution, three-dimensional cloud-resolving simulations for its training set. Superparameterization is constrained to use low-resolution (1km or 4km, typically) simulations that are often two-dimensional.

Acknowledgements. The authors would like to thank Prof.Marat Khairoutdinov (SUNY) for providing SAM and consultations on SAM.This work was supported by DOE grant number: DEFG0208ER64606.

References:

Branstator, G., 1992: The maintenance of low-frequency atmospheric anomalies, J. Atmos. Sci. 49, 1924-1945.

Donner, L. J., and P. J. Rasch, 1989: Cumulus initialization in a global model for numerical weather prediction, Mon. Wea. Rev., 117, 2654-2671.

Gill, A., 1980: Some simple solutions for heat induced tropical circulation, Quart. J. Royal Meteorol. Soc. 106, 447-462.

Grabowski, W. W., 2001: Coupling cloud processes with the large-scale dynamics using the Cloud-Resolving Convection Parameterization (CRCP), J. Atmos. Sci. 58, 978-997.

Grabowski, W. W., 2004: An improved framework for superparameterization, J. Atmos. Sci., 61, 1940–1952

Grabowski, W. W., Bechtold, P., Cheng, A., Forbes, R., Halliwell, C., Khairoutdinov, M., Lang, S., Nasuno, T., Petch, J., Tao, W., Wong, R., Wu, X., Xu, K., 2006: Daytime convective development over land: a model intercomparison based on LBA observations, Quarterly Journal of the Royal Meteorological Society.

Guichard, F., Petch, J. C., Redelsperder, J.-L., Bechtold, P., Chaboureau, J.-P., Cheinet, S., Grabowski, W., Grenier, H., Jones, C. G., Kohler, M., Piriou, J.-M., Tailleax, R. and Tomasini, M., 2004: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models, Q. J. R. Meteorol. Soc., 131, 3139–3172.

Hack, J. J., J. M. Caron, S. G. Yeager, K. W. Oleson, M. M. Holland, J. E. Truesdale, and P. J. Rasch, 2006: Simulation of the Global Hydrological Cycle in the CCSM Community Atmosphere Model Version 3 (CAM3): Mean Features, J. Climate 19, 2199–2221.

Janowiak, J. E., P. A. Arkin, and M. Morrissey, 1994: An examination of the diurnal cycle in oceanic tropical rainfall using satelliteand in situ data. *Mon. Wea. Rev.,* **122,** 2296–2311

Kalnay et al., 1996: The NCEP/NCAR 40-year reanalysis project. *BAMS,* v. 77, No. 3, 437-471.

Khairoutdinov, M. F., and D. A. Randall, 2001: A cloud resolving model as a cloud parameterization in the NCAR Community Climate System Model: Preliminary results. Geophys. Res. Let. 28, 3617-3620.

Khairoutdinov, M. F., and D. A. Randall, 2003: Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *J. Atmos. Sci.,* **60,** 607–625.

Krasnopolsky, V.M., W.H. Gemmill, and L.C. Breaker, 1999: "A multi-parameter empirical ocean algorithm for SSM/I retrievals", Canadian Journal of Remote Sensing, Vol. 25, No. 5, pp. 486-503

Krasnopolsky, V. M., 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, Mon. Wea. Rev., vol. 133, No. 5, pp. 1370-1383.

Krasnopolsky, V. M., 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, 45, RG3009, doi:10.1029/2006RG000200.

Krasnopolsky, V.M., 2007b: "Reducing Uncertainties in Neural Network Jacobians and Improving Accuracy of Neural Network Emulations with NN Ensemble Approaches", Neural Networks, 20, pp. 454–461

Krasnopolsky, V. M., M.S. Fox-Rabinovitz, and A. A. Belochitski, 2008a: "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., 2008b, "Neural Network Applications to Solving Forward and Inverse Problems in Atmospheric and Oceanic Satellite Remote Sensing", a book chapter in "Artificial Intelligence in Environmental Sciences", ed. by S.E. Haupt, A. Pasini, and C. Marzban, Springer, 375 pp

Krasnopolsky,V. M., M.S. Fox-Rabinovitz, and A. A. Belochitski, 2008c: "Ensembles of Numerical Climate and Weather Prediction Models Using Neural Network Emulations of Model Physics", Proc. of the 2008 IEEE World Congress on Computational Intelligence, Hong Kong, June 1-6, 2008, CD-ROM, paper NN0498, pp. 1524-1531

V. M. Krasnopolsky, 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

Krueger, S. K., 1988: Numerical simulation of tropical clouds and their interaction with the subcloud layer, J. Atmos. Sci., 45, 2221–2250.

Lin, J. L., G. N. Kiladis, B. E. Mapes, K. M. Weickmann, K. R. Sperber, W. Lin, M. Wheeler, S. D. Schubert, A. Del Genio, L. J. Donner, S. Emori, J. F. Gueremy, F. Hourdin, P. J. Rasch, E. Roeckner, and J. F. Scinocca, 2006: Tropical Intraseasonal Variability in 14 IPCC AR4 Climate Models, Part I: Convective Signals, J. Climate 19, 2665-2690, doi:10.1175/JCLI3735.1.

Miura, H., H. Tomita, T. Casino, S. Riga, M. Satoh, and T. Matsuno, 2005: A climate sensitivity test using a global cloud resolving model under an aqua planet condition, Geophys. Res. Lett. 32, L19717, doi:10.1029/2005GL023672.

Nesbitt, Stephen W., Edward J. Zipser, 2003: The Diurnal Cycle of Rainfall and Convective Intensity according to Three Years of TRMM Measurements. *J. Climate*, **16**, 1456–1475.

Palmer, T. N., and P. D. Williams, 2008: Introduction. Stochastic physics and climate modeling. Phil. Trans. R. Soc. A (2008) 366, 2421–2427.
doi:10.1098/rsta.2008.0059

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003a: Breaking the cloud-parameterization deadlock, Bull. Amer. Meteor. Soc., 84, 1547-1564.

Randall, D., S. Krueger, C. Bretherton, J. Curry, Duynkerke, M. Moncireff, B. Ryan, D. Starr, M. Miller, W. Rossow, G. Tselioudis, and B. Wielicki, 2003b: Confronting models with data: The GEWEX Cloud Systems Study, Bull. Amer. Meteor. Soc. 84, 455-469, doi:10.1175/BAMS-84-4-455.

Rasch P J, and J E Kristjansson, 1998. A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations. *J. Climate*., 11, 1587-1614.

Rasch, P. J., J. Feichter, K. Law, N. Mahowald, J. Penner, C. Benkovitz, C. Genthon, C. Giannakopoulos, P. Kasibhatla, D. Koch, H. Levy, T. Maki, M. Prather, D. L. Roberts, G.-J. Roelofs, D. Stevenson, Z. Stockwell, S. Taguchi, M. Kritz, M. Chipperfield, D. Baldocchi, P. McMurry, L. Barrie, Y. Balkanski, R. Chatfield, E. Kjellstrom, M. Lawrence, H. N. Lee, J. Lelieveld, K. J. Noone, J. Seinfeld, G. Stenchikov, S. Schwartz, C. Walcek, and D. Williamson, 2000: A comparison of scavenging and deposition processes in global models: Results from the WCRP Cambridge workshop of 1995, Tellus 52, 1025-1056.

Rasch, P. J., M. J. Stevens, L. Ricciardulli, A. Dai, A. Negri, R. Wood, B. A. Boville, B. Eaton, and J. J. Hack, 2006: A Characterization of Tropical Transient Activity in the CAM3 Atmospheric Hydrologic Cycle, J. Climate 19, 2222–2242.

Satoh, M., H. Tomita, H. Miura, S. Iga, and T. Nauno, 2005: Development of a global cloud resolving model – a multi-scale structure of tropical convection, J. of the Earth Simulator 3, 11-19.

Saha S. and co-authors, 2010: The NCEP Climate ForecastSystem Reanalysis, BAMS, 1015-1057.

Sui, C-H., K-M. Lau, Y. N. Takayabu, D. A. Short, 1997: Diurnal Variations in Tropical Oceanic Cumulus Convection during TOGA COARE. *J. Atmos. Sci.*, **54**, 639–655.

Yang, Gui-Ying, Julia Slingo, 2001: The Diurnal Cycle in the Tropics. *Mon. Wea. Rev.*, **129**, 784–801.