Research Article

A Neural Network Nonlinear Multimodel Ensemble to Improve Precipitation Forecasts over Continental US

Vladimir M. Krasnopolsky^{1,2} and Ying Lin¹

¹ National Centers for Environmental Prediction, NOAA, College Park, MD 20740, USA

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

Correspondence should be addressed to Vladimir M. Krasnopolsky, vladimir.krasnopolsky@noaa.gov

Received 26 May 2012; Accepted 31 July 2012

Academic Editor: Hann-Ming Henry Juang

Copyright © 2012 V. M. Krasnopolsky and Y. Lin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel multimodel ensemble approach based on learning from data using the neural network (NN) technique is formulated and applied for improving 24-hour precipitation forecasts over the continental US. The developed nonlinear approach allowed us to account for nonlinear correlation between ensemble members and to produce “optimal” forecast represented by a nonlinear NN ensemble mean. The NN approach is compared with the conservative multi-model ensemble, with multiple linear regression ensemble approaches, and with results obtained by human forecasters. The NN multi-model ensemble improves upon conservative multi-model ensemble and multiple linear regression ensemble, it (1) significantly reduces high bias at low precipitation level, (2) significantly reduces low bias at high precipitation level, and (3) sharpens features making them closer to the observed ones. The NN multi-model ensemble performs at least as well as human forecasters supplied with the same information. The developed approach is a generic approach that can be applied to other multi-model ensemble fields as well as to single model ensembles.

1. Introduction

For numerical weather prediction (NWP) models, the rainfall is one of the most difficult fields to predict accurately. Detailed knowledge of the atmospheric moisture and vertical motion fields is critical for forecasting the location and amount of rainfall, but these are difficult quantities to predict and observe accurately. Precipitations are determined by cloud dynamics and microphysical processes involved. Clouds and convection are among the most important and complex phenomena of the atmospheric 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 [1]. These processes are highly variable in time and space.

The temporal scales of involved processes range from days to several seconds (some microphysics events), and the

spatial scales change from thousands of kilometers (cyclone phenomena) to tens of micrometers (size of water droplets in the cloud). In such a situation a single NWP model cannot adequately represent the cloud dynamics and microphysical processes involved in rainfall generation because majority of these processes occur on a subgrid scale, which means that they have time and space scales that are well below the resolved scales explicitly treated in NWP models. Usually NWP models have spatial resolution, R , from several (regional models) to several tens (global models) kilometers. It means that the model does not resolve any processes that occur inside the grid cell of the size $R \times R$ km. Such processes are called subgrid processes. Therefore, NWP models must resort to parameterizations that treat subgrid processes (e.g., convective clouds) in a very simplified way to effectively take into account subgrid processes (cloud-related processes), which determine very important parameters like the amount of precipitation.

Using parameterized convective physics introduces uncertainties in quantitative precipitation forecasts (QPFs) due to at least two reasons: (1) usually there exist various

different approaches to develop parameterizations; therefore, various parameterizations exist and are used in different models that produce different QPFs and (2) a particular model using a particular parameterization produces QPF that is determined by the large-scale (described by the model) conditions independently of the fine-scale situations (unresolved by the model). Fine-scale scenarios may vary significantly depending on a particular location and time and lead to the actual amount of precipitation different for each particular location and time and different from QPF predicted by the model. In addition to the uncertainty in QPFs due to aforementioned limitations of the forecast models, the uncertainties in QPFs can arise as a result of errors in the observations due to shortcomings in observing systems. Thus, QPF should be treated as a stochastic variable with a significant uncertainty and with the statistical characteristics that may depend on time and space location.

To compensate for shortcomings in observing systems and model physics and to reduce uncertainty in QPFs, there has been a trend in recent years toward ensemble forecasting, which consists in the realization of the number of model integrations, that is, different perturbed runs of the same NWP model or different NWP models. Ensemble prediction systems (EPSs) that use perturbed initial conditions have been extensively tested and used in operations at the European Centre for Medium-Range Weather Forecasts (ECMWF) and the US National Centers for Environmental Prediction (NCEP) [2, 3]. Using this strategy, one can estimate the probability of various events and possibly also the uncertainty associated with a particular forecast. The ensemble average has repeatedly been shown to give a more accurate forecast than a single realization of the forecast model [4, 5, 9]. Drawbacks with the single-model EPSs are (1) this technique is very computationally expensive and lower-resolution versions of the models are generally employed to reduce the computational expenses that reduce the quality of the forecast and (2) assuming that errors result primarily from uncertainties in the initial conditions, any biases present in the model itself will also be present in the ensemble and may require calibration. The recent introduction of “stochastic” or “perturbed” model physics attempts to account for uncertainties in the model subgrid-scale processes [2, 6, 7].

Multimodel ensemble (MME) is another approach that has been taken to address aforementioned issues. Ebert [8] exhaustively investigated advantages and problems of MME approach using an MME composed of seven operational NWP global and regional models. In the case of MME, the ensemble is composed of output from different high-resolution NWP models (often run in different operational weather prediction centers), rather than a reduced resolution single model with perturbed initial conditions run by EPSs. Unlike EPSs that use singular vectors or breeding modes to generate optimal perturbations to the initial conditions, MME samples the uncertainty in the initial conditions via different observational data, data assimilation systems, and initialization methods used by different operational centers. MME also samples the uncertainty in model formulation due to the differences in model dynamics, the variety of model

physical parameterizations, numerics, and resolutions. As a result, MME can be considered as an approach, in which all components of NWP system are perturbed not just initial conditions or model physics. On the other hand, as any approach, MMEs have some downsides. For example, there is the inherent lack of control over the range of the intermodel differences, given that the model outputs are used “as is.” Also, models included in MME require continuous monitoring because the quality of some of the models may deteriorate with time. Despite these limitations, many authors [8–10] demonstrated superior performance of MME for QPFs.

In this paper, we introduced a new nonlinear MME based on learning from data using a neural network (NN) ensemble technique. NNs and NN ensembles have been successfully applied in other climate and meteorological applications [11–13]. The purpose of this study is to examine improvements that a nonlinear NN-based MME may introduce over a regular (linear) MME for the case of precipitation forecast. The next section describes the forecast and verifying data that we used in the study. Section 3 reviews linear methods of combining ensemble members and calculating the ensemble prediction and introduces a nonlinear NN-based MME. Section 4 contains results and their discussion. The paper finishes with conclusions.

2. Forecast and Verification Data

Twenty-four hour precipitation forecasts over the continental US (ConUS) are available from eight operational models, including NCEP’s own mesoscale and global models (NAM and GFS), the regional and global models from the Canadian Meteorological Centre (CMC and CMCLB), global models from the Deutscher Wetterdienst (DWD), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Japan Meteorological Agency (JMA), and the UK Met Office (UKMO).

Also NCEP Climate Prediction Center (CPC) 1/8 degree daily gauge precipitation analysis is available. The CPC’s analysis is used for the training of NNs and for the verification of model predictions as the “ground truth”.

All gridded data fields were interpolated to the same grid, the 40 km Lambert-conformal AWIPS Grid 212 that encompasses ConUS and has 23,865 grid points per field.

Comparisons of model QPFs with the CPC analysis indicate that all models demonstrate similar behavior at lower levels of precipitation they are slightly wetter than the CPC analysis and at the higher levels (>50–60 mm/day) they are dryer than the CPC analysis (for detailed discussion, see [14]). Moreover, locations of highs and lows and details of precipitation features are different in the precipitation fields produced by different models. The reasons of these differences are discussed in Section 1. The model results (24 h forecast) for three models (NAM, GFS, and ECMWF) together with the CPC verification analysis are shown in Figure 1 for October 24, 2010. This figure illustrates the problem and demonstrates a typical disagreement between model predictions.

FIGURE 1: The model results for three models (NAM, GFS, and ECMWF) together with the CPC verification analysis, all for 24 h period ending at 12 Z on October 24, 2010. Red and blue ellipses show high and low precipitation areas, respectively. Color bars on the left side of the figures show the color codes for the amount of precipitation from 0 to 175 mm/day. Different shades of green correspond to the amount of precipitation from 1 to 10 mm/day and different shades of red from 25 to 75 mm/day. The figure illustrates the differences in model forecasts, especially for high and low precipitations.

Predictions of all aforementioned eight models over the first six months of 2010 are shown in Figure 2 as a scatter plot, which presents 24 h QPF plotted versus the CPC analysis. More than 4×10^6 events are presented there. The figure demonstrates a tremendous spread in the models' results. The uncertainty of the forecast is especially large at higher levels of the precipitation amount. The same information is shown as binned scatter plots of all eight-model 24 h predictions versus CPC verification analysis in Figure 3. The models create an envelope with the spread increasing with the increase of the precipitation rate. All models have increasingly low bias at high levels of precipitations.

Figures 2 and 3 illustrate very well the aforementioned problems and demonstrate the stochastic nature of the system under consideration. The conservative ensemble mean (thick solid line with asterisks), which is defined in the next section (see (1)), goes in the middle of the envelop, not improving upon the situation. In the next sections, we investigate possibilities of improving the Multimodel ensemble technique for 24 h precipitation forecast using linear (multiple linear regression) and nonlinear (NN) techniques to improve upon the conservative linear ensemble (1).

FIGURE 2: Scatter plot showing 24 h precipitation forecasts obtained by eight models over the first six months of 2010 versus corresponding CPC analysis. More than 4×10^6 events are presented.

FIGURE 3: Binned scatter plot for eight models (ensemble members) and conservative ensemble EMEM (1).

3. Method

In MME as well as in EPS based on a single model, the final product is a combination of the ensemble members. At a particular time in a particular location for an ensemble with N ensemble members, N predictors, P_i , $i = 1, \dots, N$, are available for a particular variable P . To produce an ensemble prediction, ensemble members should be combined in a predictand. The simplest and most common combination of the ensemble members is the arithmetic ensemble mean (EM), which is calculated as a simple average of ensemble members (a.k.a. conservative ensemble):

$$EM = \frac{1}{N} \sum_{i=1}^N P_i, \quad (1)$$

where N is the total number of ensemble members and P_i is the i th ensemble member generated by the model number i . This approach to combining ensemble members has two major advantages: (i) it does not require any additional information; therefore, (ii) the unique result (1), EM, can always be calculated. The major disadvantage is that there is no guaranty that (1) makes the best use of information carried by the set of predictors.

More sophisticated approaches use the weighted ensemble mean (WEM):

$$WEM = \sum_{i=1}^N W_i \cdot P_i, \quad (2)$$

$$\sum_{i=1}^N W_i = 1,$$

where ensemble members are subscribed with weights, W_i , that are usually based on some ad hoc considerations. For example, if from the past experience it is known that some models give better prediction than others, they can be subscribed with higher weights in (2).

Also, multiple linear regression technique has been applied [15, 16] to determine optimal weights, W_i , for combining the ensemble members. This approach can be used only if a training dataset is available to learn the regression coefficients from data; a significant improvement was demonstrated using weighted ensemble mean over the simple ensemble mean.

If training data are available, (2) can be generalized and other predictors, x_i , $i = 1, \dots, m$, can be included in the linear regression

$$WEM = \sum_{i=0}^m a_i \cdot x_i + \sum_{i=m+1}^{N+m} a_i \cdot P_{i-m}. \quad (3)$$

The aforementioned approaches (both simple and weighted mean) implicitly assume a linear dependence between ensemble members and the best predicted value (the amount of precipitations in our case). However, in many cases, predictors are significantly correlated. In the case considered in this paper, it happens because QPFs produced by different NWP models for the same time and location are similar and correlated. Linear regression technique becomes numerically ill conditioned when dealing with correlated predictors and may require additional numerical efforts. Also, in some cases the assumption about linear dependence may be incorrect *per se*. For example, for longer forecast times when bifurcation of the ensemble forecasts may occur, it can lead to misleading results. Also for fields (like precipitation fields) with high gradients and sharp, localized features, the assumption of linearity may lead to significant problems in MME predictions (see more detailed discussion in the following sections). In such cases, the dependence between the ensemble members and the best predicted value may be a complex nonlinear one.

In this study, we relaxed the linearity assumption and allowed for an arbitrary nonlinear dependence between the MME members and the best predicted value, MME, as

$$MME = f(X), \quad (4)$$

where the vector $X = \{x, P\}$, $P = \{P_i\}_{i=1, \dots, N}$, is a vector of the ensemble member predictors and $x = \{x_i\}_{i=1, \dots, m}$ is a vector of additional predictors, which may accommodate time and location dependencies, and so forth.

A neural network (NN) technique is used to approximate this arbitrary nonlinear dependence (4) using a training set composed of past data to learn NN weights from data. The NN technique is used because NN is a universal approximator that can approximate any continuous or almost continuous dependence given a representative data set for training [17, 18]. The nonlinear NN ensemble mean (NNEM), which we introduce here, is defined following

[16, 19]; it is an analytical multilayer perceptron that can be written as

$$\text{NNEM} = a_0 + \sum_{j=1}^k a_j \cdot \phi \left(b_{j0} + \sum_{i=1}^n b_{ji} \cdot X_i \right), \quad (5)$$

where X_i are components of the input vector X (the same as in (4)) composed of the same N inputs (ensemble members) as those used for EM and WEM equations (1) and (2) plus optional additional input parameters (see (4) and Section 4), n is the number of inputs ($n \geq N$), a and b are fitting parameters (weights), and $\phi(b_{j0} + \sum_{i=1}^n b_{ji} \cdot X_i)$ is a so-called “neuron.” For the activation function, ϕ , we use a hyperbolic tangent, and k is the number of neurons in (5).

It is noteworthy to repeat that expression (5) is capable of approximating any nonlinear relationship between non-stochastic variables. However, the training set that is used for training NN (5) is composed of inputs and outputs with uncertainties. The inputs, X , contain vectors of QPFs, P , predicted by NWP models, and the outputs contain observed precipitation (CPC verification analysis). Both inputs and outputs contain significant uncertainties (see Section 1) and are stochastic variables. It means that the nonlinear function f (4) is also a stochastic function because it describes a relationship between two stochastic variables.

Actually, the stochastic function is a family of functions, each of which describes a relationship between two considered variables inside a corridor determined by the uncertainties of these variables with a probability determined by a joint probability distribution of these variables. As a result, a single NN (5) cannot provide an adequate approximation for the stochastic function (4). However, the NN technique is rich and flexible enough to solve this problem. It was suggested [20] that an ensemble of NNs can be used as an adequate tool for approximating stochastic functions (and mappings). Thus, in this paper we produce multiple NNs (an NN ensemble) to approximate the stochastic function (4), f . Each NN ensemble member is represented by (5). Finally, QPF is calculated as NN ensemble average

$$\text{MNNEM} = \frac{1}{q} \sum_{i=1}^q \text{NNEM}_i, \quad (6)$$

where q is the number of NNs (5) in the NN ensemble, and each NNEM_i is one of q NNEMs (5). There are many different method of creating NN ensembles. In this study, we used an ensemble of NNs (5) that have different weights a and b corresponding to different local minima of the error function minimized during the NN training. These minima have been found using different initializations of NN weights a and b .

Using NN ensemble for calculating MME mean, MNNEM, (6) has an additional advantage. It allows us to calculate the scatter of the NN ensemble members as the standard deviation of NNEMs

$$\sigma = \sqrt{\frac{1}{q} \sum_{i=1}^q (\text{NNEM}_i - \text{MNNEM})^2}, \quad (7)$$

σ may be helpful for estimating the uncertainty of MME forecast.

Because of the model problems described previously, the research community has been exploring various ways of making better precipitation forecasts. Among the approaches investigated at NCEP and in this paper, we consider an eight-member MME, which is averaged in three different ways calculating (i) a conservative ensemble, EM (1), (ii) WEM based on multiple linear regression (8), and (iii) a nonlinear NN ensemble MNNME (6), based on NNEM (5). Different aforementioned MME techniques have been applied for calculating 24-hour precipitation forecast over the ConUS territory [14].

First, we introduced and investigated an improved linear technique. We defined WEM (3) as a multiple linear regression in the following way [21]:

$$\text{WEM} = a_1 \cdot \text{cjd} + a_2 \cdot \text{sjd} + a_3 \cdot \text{lat} + a_4 \cdot \text{lon} + \sum_{i=1}^8 a_{i+4} \cdot P_i, \quad (8)$$

where $\{a_i\}_{i=1,\dots,12}$ are regression parameters, $\text{cjd} = \cos((\pi/183) \cdot \text{jday})$, $\text{sjd} = \sin((\pi/183) \cdot \text{jday})$, jday is the Julian day, lat is the latitude, lon is the longitude, and P_i are the ensemble members in a particular grid point of ConUS grid with the coordinates lat and lon and at a particular time (jday). Thus, the multiple linear regression (8) has 12 input parameters in total.

Each MME NN member (NNEM) is defined as in (5) where the input vector X is composed of the same $n = 12$ inputs as those used for WEM (8). The first two parameters account for the nonstationarity of the environment in the form of cyclostationary behavior (the annual cycle). Because we used only one year of the past data for development, we took into account only these basic cyclostationary changes; however, more general time dependence could be introduced if necessary.

$k = 7$ hidden neurons in (5) were selected after multiple trials to avoid overfitting [21]. Both WEM and NNEM have one output, 24 h precipitation forecast. The same CPC analysis corresponding to the time of the forecast was used to train outputs in both cases. It is noteworthy that the regression parameters for WEM and NN weights for NNEM do not depend on the location and time and are the same for all grid points and times. After WEM and NNEM are trained, they are used with the same set of regression coefficients (or weights for NN) in any grid point of the ConUS grid at any time. Thus, the results depend on time and location only through input parameters.

The amount of precipitation is an unbalanced data set since a good percent of grids points are zeros, while the rest of them have various values that are greater than zero. The distribution of the amount of precipitation is very asymmetric and far from normal distribution. Actually, the amount of precipitation has a log-normal distribution, which means that the distribution of the logarithm of the amount of precipitation is close to normal (see Figure 4). Because we minimize the mean square error function for

FIGURE 4: Probability density function for logarithm of the amount of precipitation (prec) (solid line). Dashed line shows the normal distribution with the same mean and variance.

training, the unbalanced data set with asymmetric nonnormal distribution can significantly deteriorate the accuracy of the training process. To alleviate the problem, we used the logarithm of the amount of precipitation as the output of NNs (5), which we trained to balance the data set and to work with the data distributed almost normally (see Figure 4).

4. Results and Discussion

The WEM and NNEMs have been developed using 2009 data (more than 310,000 in/out records [21]). They have been validated on independent data for the first half of 2010, for example, the results shown in Figures 2 to 5 have been calculated using these independent validation data. For the case studies presented in Figures 6 to 8 and for calculating statistics shown in Figure 9, completely independent data that cover the period from October 2010 to July 2011 have been used.

Figure 3 shows the binned scatter plot for the amount of precipitation over the ConUS territory during the first six months of 2010. It shows the eight available models together with EM (1) results versus CPC analysis. As can be seen in Figure 3, the conservative ensemble EM (1) goes inside (in the middle of) the envelope created by the models. In general, EM provides a better placement of precipitation areas; however, it does not improve the situation significantly. Moreover, EM (1) smoothes, diffuses features, and reduces spatial gradients; it has high bias for low level of precipitations (large areas of false low precipitations) and low bias at high level of precipitations (highs are smoothed out and reduced). These problems are illustrated in Figures 5, 6, 7, and 8. They motivated us to search for improved techniques including nonlinear NN ensemble. Our

validation showed that, for precipitation fields, WEM (8) does not significantly improve upon the regular Multimodel ensemble EM (1). In Figure 5, these two ensemble means, EM and WEM, are shown by thick solid and dashed black lines correspondingly. As can be seen from Figures 3 and 5, all models, EM, and WEM are slightly wetter than the CPC analysis at lower precipitation amounts and significantly drier than the CPC analysis at higher precipitation amounts. The linear ensembles, EM and WEM, do not change the situation significantly (see Figures 5(a) and 5(b)). Also the multiple linear regression ensemble, WEM, does not introduce any significant improvement upon EM.

On the other hand, there is a significant difference between linear ensemble averaging techniques (1) and (8) and the nonlinear one (5). EM (1) is always unique. WEM (8) always provides the unique solution for a given training set. Nonlinear ensemble averaging, NN ensemble mean NNEM (5) in particular, can provide multiple solutions for a given training set. For accurate training data (nonstochastic function (4) with no uncertainty), different solutions have different approximation errors, and the best solution with the smallest approximation error can be selected. For training data with the high level of uncertainty (noise), like our data shown in Figures 2 and 3, multiple solutions have almost the same approximation accuracy close to the uncertainty of the data. It means that all these solutions provide equally valid nonlinear averaging of the MME.

In terms of the NN approach, we trained an ensemble of ten NNs (5) with the same architecture ($n = 12$ inputs, one output and $k = 7$ hidden neurons) but different initialization values for weights a and b (see (5)). All ten NNs were initialized with different small random numbers using the initialization procedure developed in [22]. The training of these NNs, which is a nonlinear minimization of an error function, leads to ten different local minima of the error function with approximately the same value of the approximation error. However, because these ten NNs have different weights a and b (see (5)), they produce very different results in the areas where the uncertainty of the data is higher (higher levels of precipitations). Increasing the number of the ensemble members beyond ten does not lead to significant improvements in results.

The results of the application of different MME averaging NNs (NN ensemble members) to the validation data set are shown in Figure 5. It shows binned scatter plots for EM (1), WEM (8), and ten NNEMs (5) ($NNEM_i, i = 1, \dots, q$ and $q = 10$). Figure 5(a) displays the whole interval of precipitation values from 0 to 145 mm/day and Figure 5(b) magnifies the lower precipitation area from 0 to 50 mm/day.

All ten NNEMs are in a good agreement at the lower levels of precipitation. They diverge significantly at the higher levels of precipitation. Their large spread reflects the uncertainty in the data that is the uncertainty of MME, that is, the differences in predicting higher levels of precipitation by the different members of the MME (see Figure 2).

It is noteworthy that in the training and validation data sets less than 0.5% of the data records have precipitation values greater than 50 mm/day and only a few records have precipitation values greater than 100 mm/day.

FIGURE 5: Binned scatter plots for EM (1) (black solid), WEM (8) (black dashed), ten NNEMs (5) ($NNEM_i, i = 1, \dots, q$ and $q = 10$, all blue), and MNNEM (red) that is defined by (6). (b) It shows the magnified lower precipitation area of (a).

FIGURE 6: Comparison of three 24 h forecasts: EM (b), MNNEM (c), and HPC (d) versus CPC analysis, all for 24 h period ending at 12 Z on October 24, 2010. Red ellipses show high-precipitation areas and blue ellipses show low. Color bars on the left side of the figures show the color codes for the amount of precipitation from 0 to 175 mm/day. Different shades of green correspond to the amount of precipitation from 1 to 10 mm/day and different shades of red from 25 to 75 mm/day.

FIGURE 7: The same as in Figure 6 but for December 6, 2010.

FIGURE 8: The same as in Figure 6 but for January 4, 2011.

FIGURE 9: ETS score (a) and BIAS score (b) for the period of eight months from November 15, 2010 to July 15, 2011. Five different 24 h MME forecasts are presented: EM (1): solid red, WEM (8): dashed pink, HPC forecast: dashed blue, MNNEM (5): dashed light blue, and one of NNEM (one member of the NN ensemble (5)) forecasts: dashed brown.

To improve statistical significance of nonlinear NN MME averaging (especially at higher precipitation values), we consider the ten aforementioned NNs as an ensemble of averaging NNs and calculate the NN ensemble mean MNNEM using (6). It is shown in Figure 5 by a red solid line. MNNEM produces a significant improvement relative to EM and WEM at higher levels of precipitations (Figure 5(a)); it significantly reduces the low bias at higher precipitation levels (35 mm/day and higher). It also improves results at low-precipitation levels, significantly reducing high bias at lower precipitation levels (from 0 to 10 mm/day). However, at medium precipitation levels from ~ 12 to 30 mm/day, MNNEM and the majority of NN ensemble members have lower bias than EM and WEM, which can be seen in Figure 5(b). Thus, the nonlinear NN ensemble averaging approach is flexible enough to negotiate the wetness at lower amounts of precipitations and improve upon the dryness bias at the higher amounts.

Figures 6 to 8 demonstrate three case studies that show advantages of nonlinear NN ensemble forecast, MNNEM, as compared with the conservative ensemble forecast, EM. Here, we do not show WEM (8) results because visually they are not distinguishable from EM results. The CPC analysis for the time corresponding to the forecast is used for verification. Also, manual 24 h forecast produced at the Hydrometeorological Prediction Center (HPC) by human forecasters is also presented for comparison. To produce the HPC forecast, forecasters use the model forecasts as well as all available observations and satellite data (including sequential satellite images) [23].

Figures 6 to 8 demonstrate that the nonlinear NN averaging of MME improves positioning of precipitation features inside the precipitation fields. It removes significant areas of false low level precipitations produced by a standard EM (1) technique. It sharpens the features and enhances precipitation fronts and maximums. The MNNEM technique provides a forecast that is comparable with the HPC forecast while using much less resources and time.

In conclusion of the discussion, the statistical results that characterize the accuracy of positioning precipitation features are shown in Figure 9. The statistics cover the period

of eight months from November 15, 2010 to July 15, 2011. As it was mentioned previously, NNs have been trained on 2009 data.

The Equitable Threat Score (ETS) [24] measures that fraction of observed events that are correctly predicted, adjusted for correct predictions that are due to random chance. Possible ETS ranges from $-1/3$ to 1 (perfect forecast would have a score of 1 for every precipitation threshold). Bias score is simply the ratio of areal coverage of forecast versus observed precipitation exceeding a given threshold. An ideal forecast would have a bias score of 1 at every threshold.

Summarizing, the MNNEM forecast is comparable with the HPC forecast and significantly better than EM at the threshold values less than 0.1 inch/day and more than 1 inch/day, which is in a good agreement with the statistics presented in Figure 5.

5. Conclusions

In this paper, we introduce a nonlinear NN Multimodel ensemble approach to improve 24 h Multimodel ensemble precipitation forecast. This straightforward application of NNs to the problem produced promising results. We showed that our NN MME improves upon simple linear ensemble; it

- (1) significantly reduces high bias at low-precipitation level;
- (2) significantly reduces low bias at high-precipitation level;
- (3) sharpens features making them closer to observed ones.

It is noteworthy that the NN Multimodel ensemble forecast works at least as well as the forecast produced by human forecasters. NN forecast is produced without using any additional information that is available to the forecasters, and it is less time and resource consuming.

It is also noteworthy that the NN technique is flexible enough to accommodate the time and space dependence of the environment in which NN works through additional

time and location dependent inputs. The NN technique also allowed us to take into account the stochastic nature of the problem. We successfully used NN ensemble technique to approximate a stochastic function (mapping), which performs an averaging of MME taking into account significant uncertainty in the data used for NN training.

Any ensemble averaging technique (linear or nonlinear) based on using past data, including the NN approach introduced in this paper, requires additional maintenance efforts as compared with the simplest conservative ensemble (1). The MME ensemble members, the NWP models, are evolving complex systems. Their prediction quality changes with time. Accordingly, the function f (4) changes with time. Therefore, the quality of its approximation (5), (8), or any other should be permanently monitored, and if necessary, the approximations should be periodically adjusted (retrained). For the approach introduced in this paper, the NN ensemble trained on 2009 data still worked well in 2011. Hence, if the NN ensemble requires a retraining, it should be performed once per M ($M > 2$) years.

In this study, we implemented NN in a simple way, supplying it with the same information which the linear Multimodel ensemble approaches use. The flexibility of the NN approach allows us to introduce in future studies more sophisticated NN approaches. For example, we are planning to introduce information available to a human analyst (and HPC forecast itself) as additional inputs to our NNs. We are also planning to implement a field-wise approach taking inputs from several neighboring grid points, and so forth. Also, several previous forecasts and/or CPC analyses can be used as additional inputs.

The nonlinear NN averaging approach that we developed in this paper is a generic approach. Although here we applied it to precipitation fields, it is clear that it can be applied to other fields as well. Also here we applied this approach to calculate the MME mean; it can be applied as well to calculate nonlinear ensemble mean in a single model EPS.

Acknowledgment

The authors would like to thank Dr. Stephen Lord for initiating this work and for stimulating discussions and the reviewers for very helpful and constructive comments. MMAB Contribution no. 296.

References

- [1] P. J. Rasch, M. J. Stevens, L. Ricciardulli et al., "A characterization of tropical transient activity in the CAM3 atmospheric hydrologic cycle," *Journal of Climate*, vol. 19, no. 11, pp. 2222–2242, 2006.
- [2] R. Buizza, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, "A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems," *Monthly Weather Review*, vol. 133, no. 5, pp. 1076–1097, 2005.
- [3] T. N. Palmer, R. Buizza, M. Leutbecher et al., "The ensemble prediction system recent and ongoing developments," ECMWF Technical Memorandum 540, 2007.
- [4] Z. Zhang and T. N. Krishnamurti, "Ensemble forecasting of hurricane tracks," *Bulletin of the American Meteorological Society*, vol. 78, no. 12, pp. 2785–2795, 1997.
- [5] R. Buizza and T. N. Palmer, "Impact of ensemble size on ensemble prediction," *Monthly Weather Review*, vol. 126, no. 9, pp. 2503–2518, 1998.
- [6] R. Buizza, M. Miller, and T. N. Palmer, "Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System," *Quarterly Journal of the Royal Meteorological Society*, vol. 125, no. 560, pp. 2887–2908, 1999.
- [7] V. M. Krasnopolsky, M. S. Fox-Rabinovitz, and A. A. Belochitski, "Ensembles of numerical climate and weather prediction models using neural network emulations of model physics," in *Proceedings of the IEEE World Congress on Computational Intelligence*, pp. 1524–1531, Hong Kong, June 2008, CD-ROM, Paper NN0498.
- [8] E. E. Ebert, "Ability of a poor man's ensemble to predict the probability and distribution of precipitation," *Monthly Weather Review*, vol. 129, no. 10, pp. 2461–2480, 2001.
- [9] J. Du, S. L. Mullen, and F. Sanders, "Short-range ensemble forecasting of quantitative precipitation," *Monthly Weather Review*, vol. 125, no. 10, pp. 2427–2459, 1997.
- [10] M. S. Speer and L. M. Leslie, "An example of the utility of ensemble rainfall forecasting," *Australian Meteorological Magazine*, vol. 46, pp. 75–78, 1997.
- [11] V. M. Krasnopolsky, "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*, vol. 45, no. 3, Article ID RG3009, 2007.
- [12] A. J. Cannon and P. H. Whitfield, "Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models," *Journal of Hydrology*, vol. 259, no. 1–4, pp. 136–151, 2002.
- [13] Y. B. Dibikey and P. Coulibaly, "Temporal neural networks for downscaling climate variability and extremes," *Neural Networks*, vol. 19, no. 2, pp. 135–144, 2006.
- [14] Y. Lin and V. M. Krasnopolsky, "Simple- and Modified-Poor Man's QPF Ensembles, and a Neural Network Approach," in *Proceedings of the 91th Annual AMS Meeting, 24th Conference on Weather and Forecasting/20th Conference on Numerical Weather Prediction*, 2011, Paper 6A.2.
- [15] T. N. Krishnamurti, C. M. Kishtawal, T. E. LaRow et al., "Improved weather and seasonal climate forecasts from multimodel superensemble," *Science*, vol. 285, no. 5433, pp. 1548–1550, 1999.
- [16] T. N. Krishnamurti, C. M. Kishtawal, Z. Zhang et al., "Multimodel ensemble forecasts for weather and seasonal climate," *Journal of Climate*, vol. 13, no. 23, pp. 4196–4216, 2000.
- [17] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [18] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [19] V. M. Krasnopolsky, "Reducing uncertainties in neural network Jacobians and improving accuracy of neural network emulations with NN ensemble approaches," *Neural Networks*, vol. 20, no. 4, pp. 454–461, 2007.
- [20] V. Krasnopolsky, M. Fox-Rabinovitz, A. Belochitski, P. J. Rasch, P. Blossy, and Y. Kogan, "Development of neural network convection parameterizations for climate and NWP models using Cloud Resolving Model simulations," NCEP

Office Note, No.469, pp. 70, 2011, <http://www.emc.ncep.noaa.gov/officenotes/newernotes/on469.pdf>.

- [21] V. M. Krasnopolsky and Y. Lin, "Application of the neural network technique to develop a nonlinear multi-model ensemble for precipitation over ConUS," in *Proceedings of the 91th Annual AMS Meeting, 9th Conference on Artificial Intelligence and its Application to the Environmental Sciences*, 2011, Paper 1.1.
- [22] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '90)*, pp. 21–26, June 1990.
- [23] D. R. Novak, C. Bailey, K. Brill et al., "Human improvement to numerical weather prediction at the Hydrometeorological Prediction Center," in *91th Annual AMS Meeting, AMS 24th Conference on Weather and Forecasting/20th Conference on Numerical Weather Prediction*, 2011, Paper No 440.
- [24] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*, Academic Press, 2nd edition, 2006.