An adaptive nonlinear MOS scheme for precipitation forecasts using neural networks

Yuval, William W. Hsieh

Department of Earth and Ocean Sciences,

University of British Columbia,

Vancouver, British Columbia, Canada

Submitted for publication in *Weather and Forecasting*

October 28, 2002

Abstract

A novel neural network (NN) based scheme performs nonlinear Model Output Statistics (MOS) for generating precipitation forecasts from a Numerical Weather Prediction (NWP) model output. Data records from the past few weeks are sufficient for establishing an initial MOS connection which then adapts itself to the ongoing changes and modifications in the NWP model. The technical feasibility of the algorithm is demonstrated in three numerical experiments using the NCEP reanalysis data in the Alaskan panhandle and the coastal region of British Columbia. Its performance is compared to that of a conventional NN–based non–adaptive scheme. When the new adaptive method is employed, the degradation in the precipitation forecast skills due to changes in the NWP model is small, and is much less than the degradation in the performance of the conventional non–adaptive scheme.

1 Introduction

Model Output Statistics (MOS) is a process by which a statistical relationship between the output of a Numerical Weather Prediction (NWP) model and observations is established in order to improve forecasts. This process is most often applied to forecast problems where the variable to be forecasted is not produced by the NWP model, or for downscaling where the spatial resolution of the NWP model is too coarse. Rough terrain, lack of observations, and yet inadequate understanding of various physical processes are additional problems that contribute to reduced predictability of the NWP and call for additional processing by MOS.

A major obstacle in implementing a MOS prediction system is the ongoing modification of NWP models. Improvements in the dynamic and data assimilation schemes, changes in the observation system, and refinements of the temporal and spatial resolution of the numerical solutions all contribute to changes in the NWP model characteristics. The weather system itself is also changing on various time scales. No doubt then that the connection between the NWP output and the weather variables such as precipitation must be changing as well, and that calls for MOS prediction schemes that adapt themselves accordingly. In a recent paper, Wilson and Vallée (2001) described the Updateable MOS (Ross 1987) system of the Meteorological Service of Canada. This system is based on updating the data set from which the linear MOS empirical relationships are developed. A direct estimation and update of parameters of a linear MOS relationship from sequential data can be carried out by Kalman filtering (Grewal 1993). The authors are not aware of an existing technique to continuously update a nonlinear MOS relationship. An interesting alternative for nonlinear MOS forecasts of precipitation was introduced by Xia and Chen (1999). Their Model Output Dynamics scheme is based on the modification of the model vertical velocity given the last rain observations. Factors like topography cannot be incorporated in this method and thus its capability is limited in mountainous areas like British Columbia (BC).

The importance of a short term localized precipitation forecasts for flood control, transportation safety, landslide and avalanche prediction, and for the general public's interest put it in the focus of many research efforts (Ebert 2001; Mao et al. 2000; Hall et al. 1999; Koizumi 1999; Xia and Chen 1999; Kuligowski and Barros 1998a, 1998b; Krzysztofowicz 1998). Prediction of precipitation is notoriously difficult and it is a prime example where NWP models fail very often, calling for establishment of MOS schemes. The relationship between the NWP variables and the true precipitation cannot by any means be assumed linear. Thus, several studies (Hall et al. 1999; Koizumi 1999; Kuligowski and Barros 1998a, 1998b) employed neural networks (NNs) to form the statistical connection. In all these cases the NWP model, and the MOS scheme, were frozen in time.

This paper proposes an NN–based MOS process that adapts itself continuously. At each time point, an NN is trained to form the connection between the 6-h accumulated precipitation at a set of measuring stations and the corresponding NWP forecasts valid for that time. This NN model is a modification of an existing references model trained using all the data, in all the stations during a short period of time sometime before present. The updated model is constructed to best fit the new observed data (using predictors of possibly modified NWP model), keeping the desirable characteristics provided by the reference. The training process is completely automatic and the decisions about the level of modification are made according to a well defined mathematical criterion (Golub et al. 1979).

Providing precipitation forecasts in the coastal region of the North American Pacific Northwest is a challenging problem because of the acute topographical variations, and the large Pacific data void that does not enable proper initializations of NWP models. The current precipitation forecasts for the region are not satisfactory and we hope this work will serve as a basis for the future establishment of a MOS precipitation forecasting system in the observation stations participating in the Emergency Weather Network of BC (currently under construction), using output of the high resolution NWP models run by the Atmospheric Sciences Programme at the University of BC. In this paper, data from the NCEP reanalysis set (Kalnay et al. 1996) are used to demonstrate the technical feasibility of establishing the proposed scheme. The expected differences and difficulties in its application in an operational mode are discussed.

2 The adaptive NN

2.1 The modeling process

A system of N precipitation observation stations is considered. A set of K NWP variables are believed to be associated with the precipitation at each of these stations and are used as predictors. The accumulated observed precipitation between t_{i-1} and t_i in the N observation stations are provided by $\mathbf{y}_i^{\rm obs}$. The corresponding K NWP predictor variables are given in the $K \times N$ matrix $\mathbf{X}_i$, each of its N columns contains the NWP values associated with one observation station at time t_i. The NWP predictors can be variables from grid points in the vicinity of that station, and valid at various time points (e.g. $t_i, t_{i-1}, t_{i-2}, t_{i-3}, \cdots$). The choice of predictors is problem depended and requires some exploration. Additional quantities like local topographical information can also be included. Note that the prediction lead time of the process was not mentioned. It equals the time elapsed between production of the NWP variables at $t_i - \Delta$ and the prediction time t_i.

The number of observation stations N is not constrained to be a constant and its value can change from one time point to another if the number of stations that actually report precipitation vary. This is important for operational systems where more likely than not a few of the stations fail to report at any given time point.

The NN relationship connecting the precipitation values and the NWP model output is described by

$$\mathbf{y}_i^{\mathrm{cal}} = \mathcal{F}[\mathbf{X}_i, \mathbf{w}_i], \tag{1}$$

where $\mathbf{y}_i^{\mathrm{cal}}$ is a vector of current precipitation values calculated by the NN model whose parameters are stored in the vector $\mathbf{w}_i$, using the information provided by $\mathbf{X}_i$. A detailed description of $\mathcal{F}$ will be given in the next subsection.

The values $\mathbf{w}_i$ are the ones that minimize the cost function

$$\phi = ||\mathbf{y}_i^{\mathrm{obs}} - \mathbf{y}_i^{\mathrm{cal}}||^2 + \beta||\mathbf{w}_i - \mathbf{w}_{\mathrm{ref}}||^2 \tag{2}$$

where $\mathbf{w}_{\mathrm{ref}}$ contains the reference NN model parameters, and β is a parameter determining the trade-off between the data fit constraints in the first term of ϕ, and the requirement of the model to be close to the reference one, expressed in the second term. The value of β is chosen automatically by simultaneously minimizing ϕ and the General Cross Validation (GCV) function (Haber and Oldenburg 2000; Yuval 2000). Minimizing the GCV function ensures that the NN model is optimally tuned for the prediction of data points *not used* in the model's development, and for the avoidance of overfitting (Haber and Oldenburg 2000; Golub et al. 1979). A natural candidate for the reference model $\mathbf{w}_{\mathrm{ref}}$ is the previous day's model $\mathbf{w}_{i-1}$. However, we found it more beneficial to use as a reference a model based on the data accumulated over a slightly longer period of time. Using the latter option, it is advisable, although not imperative, to periodically update the reference model as more data is accumulated.

The spatial distribution of the precipitation at the next time point t_{i+1} is predicted by

$$\mathbf{y}_{i+1}^{\mathrm{pre}} = \mathcal{F}[\mathbf{X}_{i+1}, \mathbf{w}_i]. \tag{3}$$

When the observations $\mathbf{y}_{i+1}^{\text{obs}}$ become available, they can be compared to $\mathbf{y}_{i+1}^{\text{pre}}$ in order to evaluate the accuracy by which the NN model predicted the observations. These new observations, and the corresponding NWP values $\mathbf{X}_{i+1}$, are then used to produce the updated NN parameters for the prediction in the next time point.

The process of updating the model elements and predicting future precipitation values is repeated continuously. Small, or no changes to the model parameters are needed if no significant changes have occurred in the system since the last update of the reference model. In that case, a reasonable fit between $\mathbf{y}^{\text{obs}}$ and $\mathbf{y}^{\text{cal}}$ can be achieved by a model close to the reference, and the GCV controlled training is likely to choose a large value for the β in Eq. 2. However, at times of seasonal changes or following modifications in the NWP, the relationship between the NWP output $\mathbf{X}$, and the observed precipitation $\mathbf{y}^{\text{obs}}$ does not remain the same. In this case, the reference model parameters cannot provide an adequate fit between $\mathbf{y}^{\text{obs}}$ and $\mathbf{y}^{\text{cal}}$. The chosen value of β will be appropriately small, enabling a large deviation of the model from the reference to adapt it to the new relationship.

2.2 The neural network model

The relationship between the NWP output variables and the spatial precipitation values is modeled by

$$\mathcal{F}[\mathbf{X}, \mathbf{w}] = \mathcal{F}[\mathbf{X}, \mathbf{W}_1, \mathbf{B}_1, \mathbf{w}_2, \mathbf{b}_2] = \mathbf{w}_2 \left(F \left(\mathbf{W}_1 \mathbf{X} + \mathbf{B}_1 \right) \right) + \mathbf{b}_2, \tag{4}$$

where F is the hyperbolic tangent function, $\mathbf{W}_1$ is an $L \times K$ matrix, $\mathbf{B}_1$ is an $L \times N$ matrix with identical columns, $\mathbf{w}_2$ is a $1 \times L$ row vector, and $\mathbf{b}_2$ is a $1 \times N$ row vector with identical elements. In the NN literature $\mathcal{F}$ is usually called a two-layer feedforward NN. The manipulations of $\mathbf{X}$ by $\mathbf{W}_1$ and $\mathbf{B}_1$ are referred to as the NN's first, or hidden layer, and F is the hidden layer's transfer function. The value of L is called the number of hidden neurons. The larger it is, the more complex is the NN model. It is convenient to store all the elements of $\mathbf{W}_1$, $\mathbf{B}_1$, $\mathbf{w}_2$, and $\mathbf{b}_2$ in one vector of NN model parameters $\mathbf{w}$.

It has been shown by Cybenko (1989), Hornik et al. (1989), and Funahashi (1989), that a two–layer feedforward NN can approximate arbitrarily well any continuous nonlinear function given a set of inputs $\mathbf{X}$ and a sufficient number of hidden neurons L. The NN is thus assured to be able to sufficiently simulate the desired relationship at any given time, no matter how complex and nonlinear this relationship might be. The problem is to find the model parameters that enable the model to capture the actual relationship between the NWP output and the observed values while avoiding fitting to noise (from both measurements and calculation artifacts).

The assumption of our methodology is that, at a given time point, the physical processes governing the relationship between the predictor variables and precipitation is the same in all the stations. The same NN model ($\mathbf{w}_i$ of equation 1) is simulating this relationship, and the predicted spatial differences in the precipitation are due to the differences in the corresponding NWP predictors at the various locations. This assumption certainly does not hold in the very general case. Many different processes like large synoptic troughs, small scale turbulence, and orographic lifting can lead to precipitation, and the relationship between NWP predictors and the precipitation is not necessarily the same in all these cases. For example, it was not surprising to find out that the NN model that works well for predicting large scale synoptic precipitation along the BC coast is not suitable for predictions of precipitation in the interior BC Peace River region where most of the precipitation comes from local summer thunderstorms. Thus the MOS system should include only stations where the precipitation is predominantly generated by similar physical processes. Different systems can be constructed for different regions. In this paper only stations along the BC coast and the Alaska panhandle were included.

3 Data

The data in this study are from the NCEP/NCAR reanalysis project (Kalnay *et al.* 1996) that uses a state of the art global assimilation model to create a comprehensive dataset that is as complete as possible. The output variables are calculated on two different grids, a Gaussian T62 grid with 192×94 points (about $1.9^o \times 1.9^o$), and a $2.5^o \times 2.5^o$ lat–long grid. The six hourly data (four times a day) from 1 January 2000 to 31 December 2001 were used in this paper.

This study considers the data of the coastal Pacific Northwest, from northern Washington state (47.5^oN) to the BC–Yukon border (60.0^oN), including the Alaska panhandle (Fig. 1). The precipitation in this area is given in 31 Gaussian grid points. Sixteen atmospheric variables, believed to be related to the precipitation rate, were chosen as predictors. A deliberate choice was made to use only predictor variables given on the lat-long grid. Their values had to be interpolated to the precipitation grid locations imitating the real life situation where the NWP grid points usually do not coincide with the locations of the precipitation measurement stations. The interpolation was crude—each precipitation grid point was associated with the predictors values in the closest lat-long grid point. This simulates the expected inaccurate interpolation of NWP output values over the rough terrain of the Coast Mountains region.

The predictor variables are: 1000mb, 850mb, and 500mb air temperature (^{o}K), 1000mb, 850mb, and 500mb geopotential height (m), 1000mb, 850mb, 700mb, and 500mb vertical velocity (m/s), 1000mb, 850mb, and 500mb relative humidity (%), 1000mb, 850mb, and 500mb specific humidity (kg/kg). Values of these predictors were available four times a day at 0000, 0600, 1200, and 1800. The six–hourly accumulated precipitation rate (mm/6-h) in between these hours (i.e. 0000-0600, 0600-1200, etc.) at the 31 stations is the predictand. It must be noted that the NCEP precipitation is purely model based so the NN models only simulate the dynamic model that generates it. Neither are the temporal errors in the NWP predictors simulated in this study. This should not affect the conclusions about the effectiveness of the adaptive process as its performance is compared to a non–adaptive one benefiting from the same advantage. For simplicity, only predictors variables valid at the prediction time of the MOS process were used. For an operating MOS system, NWP variables valid for previous time points should also be considered and can help improve consistent temporal errors in the NWP predictors.

4 Results

The suggested methodology is based on updating an NN–based MOS connection established using short data records of many stations simultaneously. Its results are demonstrated, and compared in this section to those achieved by a conventional (e.g. Kuligowski and Barros, 1998a) NN–based MOS connection that is developed individually for each station using long data records, but with no updates to the model. The comparison method is referred henceforth as the benchmark.

For the benchmark, data of one year (1460 time points) were used to develop an individual model at each station connecting the 16 predictor variables, simulating NWP output, to the precipitation predictand. The models were developed using the MATLAB Levenberg-Marquardt routine (Demuth and Beale 1998). The performance of these models was tested in the following year at which the predictors were modified in various ways to simulate the modifications that occur in the NWP output.

A reference model is needed in order to apply the adaptive MOS approach. This reference model was developed using only the data of the last month of the first year. The same modified predictors were used to update the adaptive model and predict the precipitation during the second year. Testing the predictions in this case was carried out by using the model, updated by the new data of a six hour period, to predict the precipitation of the corresponding six hour period in the next day. The frequent updating ensures prompt adaptation to possible changes in the NWP but is not necessary if the changes are known to occur on a less regular basis.

Three numerical experiments were carried out. In the first one, no modifications were applied to the data in the testing year to compare the performance of the MOS schemes in the case where a "frozen" NWP model is used. In practice, NWP models are never "frozen" and two other experiments compared the performance of the two schemes in scenarios where the testing period predictors were modified. In one experiment the modifications were carried out according to $V = V + \nu$ sign(V) |V|, where V is a predictor value, sign(V) $= -1, 0, 1$ for $V < 0, = 0, > 0$, respectively, and $|\cdot|$ denotes taking an absolute value. This results in modification which is linearly proportional to the magnitudes of the predictor values with the parameter ν controlling the amount of modification. The testing year was divided into 10 equal periods with $\nu = 0.01, 0.05, 0.10, 0.05, 0.01, -0.01, -0.05, -0.1, -0.05, -0.01$. Thus the level of the linear modification changed every 146 time points (about five weeks). The second type of modification was carried out according to $V =$ sign(V) $|\mathrm{V}|^{\nu}$ with $\nu = 7/6, 6/5, 5/4, 6/5, 7/6, 6/7, 5/6, 4/5, 5/6, 6/7$ in the ten segments of the testing year. The modified predictor values in this case are proportional to powers of the values, resulting in a nonlinear modification. The predictor series were all normalized prior to the modification so that only the relative magnitudes, not the units of the variables, dictate the amount of modification. Additional experimentation with combined linear and nonlinear predictor modifications, different division of the testing period, and different ranges for the value of ν lead to conclusions similar to those extracted from the results of the three experiments presented below.

Fig. 2 shows scatterplots of precipitation predictions against the observations at all the stations during the period of the testing year using the unmodified data. The predictions in Fig. 2a were produced by models updated every time point according the method proposed in this paper. Fig. 2b shows the corresponding plot for the benchmark where models were trained separately for each station using the full data record of the training year. The reduced scatter in Fig. 2b compared to that in Fig. 2a, and the better corresponding correlation and root mean square error (rmse) scores, given in Table 1, show a clear advantage of the benchmark method. This is not surprising bearing in mind that with no modifications, the data in both training and testing periods in this case are produced by the same frozen NCEP data assimilation model.

The results in Fig. 2 are what we expect to see in the ideal case of a MOS scheme developed for an NWP model that never changes. The benchmark, using training on a much longer data records and tailored model for each station results in superior predictions and should have been chosen for practical use in case of a MOS system using a frozen NWP model. Unfortunately, NWP models invariably change so a more realistic comparison is that of the performance of the two methods using testing predictors data that were somewhat modified.

Fig. 3 shows the scatterplots resulting from the second experiment where the predictors in the testing period were linearly modified. Both plots show more scatter than the plots in Fig. 2 but while Fig. 3a, showing results of the adaptive method, is quite similar to the corresponding plot in Fig. 2a, the scatterplot from the benchmark (Fig. 3b) shows much greater scatter than Fig. 2b. The correlation and rmse skills achieved by the benchmark in this case (Table 1) are significantly worse than those for the non–modified data, and are inferior to those achieved by the adaptive method.

Fig. 4 compares the Probability of Detection (POD), the False Alarm Rate (FAR) and the Threat Score (TS) (Wilks, 1995) of the results achieved by the adaptive method and the benchmark.

These three measures are more suitable than the correlation and rmse skills for ranking predictions of events of interest which are less likely to occur than not. Precipitation forecasts, especially heavy precipitation events are thus better ranked using these measures. The POD is the ratio of predictions of events (i.e. precipitation above certain threshold) that were predicted and materialized, to the number of observations of these events. The FAR is the ratio of predicted events that did not materialize to the total number of predictions of such events. The TS, most common among accuracy measures in precipitation prediction studies, is the ratio of materialized forecasts to the total number of occasions in which an event was forecasted and/or observed. The range of the three measures is $[0, 1]$ with the POD and TS having positive orientation, i.e. a score of unity is best, and the FAR having negative orientation. Readers are advised to consult Wilks (1995) for a more complete discussion of these measures and the contingency table from which they are derived.

The POD values of the adaptive method in Fig. 4 are in general slightly higher than those of the benchmark. With linearly modified predictors, a little degradation is noticed in most of the scores relative to those achieved using non–modified predictors. The better POD values of the the adaptive method, especially at higher precipitation thresholds, signify a better capability in forecasting large precipitation events than the benchmark. This capability is unfortunately, but not surprisingly, accompanied by a higher rate of false alarms when using the non–modified data. However, the false alarm rate of the benchmark degrades much more, and much exceeds that of the the adaptive method when using the linearly modified data. The relatively small degradation in the FAR scores obtained in this case by the adaptive method suggests its possible advantage in reducing false alarms when the NWP outputs are occasionally biased up or down. A similar advantage is demonstrated by inspecting the TS results which take into account both ability to predict events, and to avoid falsely calling for their occurrence. The TS results of the benchmark are slightly higher than those of the adaptive method for the non–modified data, but the degradation resulting from the use of modified predictors pushes the benchmark TS curve well below that of the adaptive method.

Figs. 5 and 6 compare the results achieved in the third experiment where the predictors, simulating NWP output, were nonlinearly modified. Fig. 5a is a scatterplot obtained using the adaptive method. Only a little additional scatter is detected in this plot compared to that in Fig. 2a where non–modified data were used, and the correlation and rmse skills are only slightly degraded. The scatter of the predictions by the benchmark in Fig. 5b is clearly worse in this case compared to the scatter in Fig. 2b, and the corresponding correlation and rmse scores are obviously not as good. The POD, FAR, and TS results shown in Fig. 6 convey similar information to that provided by Fig. 4 in the case of the linearly modified predictors. That is, more accurate predictions were achieved by the benchmark when frozen NWP output was used, but the deleterious effects of alterations in the predictors affected the adaptive MOS method significantly less.

5 Discussion and Conclusions

This paper proposes a method to adapt NN–based MOS prediction of precipitation as new NWP output and observations arrive. Its main advantages are the relatively short length of data record required to establish the MOS connection, and the ability to adapt this connection to changes in the NWP model and/or observations. The method assumes that the physical processes that lead to the precipitation at the different observation locations are similar and that the differences in the MOS connection are only due to differences in the predictors. The performance of the proposed method was demonstrated in three numerical experiments, and was compared to that of a benchmark non–adaptive NN–based MOS scheme. A non–adaptive scheme benefits from the information provided by longer data records (in case they are available) to establish the MOS connection, and can be tailored separately for each location.

The advantage of longer data records was evident in the superior prediction achieved by the benchmark while using data produced by a frozen dynamic model. Modifying the MOS predictors, to simulate changes in the NWP model, resulted in severe degradation of the MOS prediction by the benchmark. The performance of the adaptive scheme was not as severely affected by the modification in the NWP output, suggesting it might be of use for improving operational NWP predictions when frequent modifications in the NWP operation prevent the establishment of a conventional MOS scheme.

The study described in this paper used NCEP reanalysis records of atmospheric variables as the NWP predictors, and the purely model–based NCEP precipitation record as the predictand. Being purely model based, the NCEP precipitation does not necessarily agree with any observed precipitation values. It is rather tuned to agree with related variables like vertical velocity, specific humidity and latent heat flux, that are smoothed over the NCEP analysis grid. Thus the spatial and temporal distribution of the precipitation values are not expected to closely resemble the real ones. Our comparison of the NCEP precipitation records at selected locations where rain gauge observations were available, revealed substantial differences on a value by value basis. However, with the exception of lack of extreme values (above 25 mm/6-h), in the range, irregularity of the series, and the shape of the probability density function, the NCEP precipitation records in the study region are quite similar to the observed ones. The typical difficulties in predicting observed precipitation appear also while predicting the ones produced by the NCEP model.

We thus believe that the use of NCEP precipitation to demonstrate the technical feasibility of establishing an adaptive NN-based MOS scheme for predicting precipitation is justified. Testing on data used for operational forecasts is needed in order to evaluate this method's worth for practical predictions. A main concern is the prediction capability of the NWP output on which the overall performance of the MOS scheme depends. Using NCEP records as simulators of NWP output eliminated that concern in this study but it remains to be tested how well the recently developed high resolution models (1.0 km grid and below) will be capable to capture the small scale, but important, phenomena that are responsible for much of the variability in precipitation in rugged terrain like that of the Pacific Northwest.

To apply the nonlinear adaptive MOS scheme in an operational mode will also require extensive additional exploration to tailor the algorithm to the specific region under consideration. The most important issues to consider are: (a) The division of the locations in the forecast area into groups with similar precipitation patterns. The more homogeneous the groups the better, but from our experience, a minimal number of about 30 stations is needed for each group. (b) An exploration for the best predictors, especially additional ones that convey information not included in the NWP output, like typical local wind patterns, local topography (below the NWP resolution), etc. We believe that properly addressing these issues, and using reasonably accurate NWP outputs, the algorithm presented in this paper is a viable tool for improving precipitation forecasts.

6 Acknowledgments

The need for an adaptive nonlinear MOS scheme was first brought to our attention through discussions with professor Roland Stull. We thank three anonymous reviewers for their helpful comments. This work was supported by research and strategic grants to William Hsieh from the Natural Sciences and Engineering Research Council of Canada.

7 References

Cybenko, G., 1989: Approximation by superpositions of a sigmoidal function. *Math. Control, Signal, Sys.,* **2,** 303-314.

Demuth, H., and M. Beale, 1998: Neural Network Toolbox Version 3.1. The Math Works, Inc.

Ebert, E. E., 2001: Ability of a Poor Man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129,** 2461-2480.

Funahashi, K., 1989: On the approximate realization of continuous mappings by neural networks. *Neural Networks*, **2,** 183-192.

Golub, G. H., M. Heath, and G. Wahba, 1979: Generalized cross–validation as a method for choosing a good ridge parameter. *Technometrics*, **21,** 215-223.

Grewal, M. S., 1993: *Kalman Filtering: Theory and Practice.* Prentice–Hall Information and System Science Series, Prentice–Hall, 381 pp.

Haber, E., and D. W. Oldenburg, 2000: A GCV based method for nonlinear ill–posed problems. *Comput. Geosci.*, **4,** 41-63.

Hall, T., H. E. Brooks, and C. A. Doswell, 1999: Precipitation forecasting using a neural network. *Wea. Forecasting*, **14,** 338-345.

Hornik, K., M. Stinchcombe, and H. White, 1989: Multilayer feedforward networks are universal approximators. *Neural Networks*, **2,** 359-366.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40–year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77,** 437-471.

Koizumi, K., 1999: An objective method to modify numerical model forecasts with newly given weather data using an artificial neural network. *Wea. Forecasting*, **14,** 109-118.

Krzysztofowicz, R., 1998: Probabilistic hydrometeorological forecasts: Toward a new era in operational forecasting. *Bull. Amer. Meteor. Soc.*, **79,** 243-251.

Kuligowski, R. J., and A. P. Barros, 1998a: Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. *Wea. Forecasting*, **13,** 1194-1204.

Kuligowski, R. J., and A. P. Barros, 1998b: Experiments in short–term precipitation forecasting using artificial neural networks. *Mon. Wea. Rev.*, **12,** 470-482.

Mao, Q., S. F. Mueller, and H. H. Juang, 2000: Quantitative precipitation forecasting for the Tennessee and Cumberland River watershed using the NCEP regional spectral model. *Wea. Forecasting*, **15,** 29-45.

Ross, G. H., 1987: An updateable model output statistics scheme. Programme on Short– and Medium Range Weather Prediction, PSMP Report Series, No. 25, World Meteorological Organization, 25-28.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.

Wilson, L. J., and M. Vallée, 2001: The Canadian Updateable Model Output Statistics (UMOS) system: Design and development test. *Wea. Forecasting*, **17,** 206-222.

Xia, J., and A. Chen, 1999: An objective approach for making rainfall forecasts based on numerical model output and the latest observation. *Mon. Wea. Rev.*, **14,** 49-52.

Yuval, 2000: Neural network training for prediction of climatological time series, regularized by minimization of the Generalized Cross Validation function. *Mon. Wea. Rev.*, **128,** 1456-1473.

8 Table captions

Table. 1 The correlation and rmse skills of the results obtained by the adaptive and non–adaptive schemes using the training non–modified data (Test 1), the testing data with linearly modified predictors (Test 2), and the testing data with nonlinearly modified predictors (Test 3).

Table. 2 The slope and intercept values of the least square fit lines in Figs. 1, 2, and 5 (Test 1, Test 2, and Test 3 correspondingly).

9 Figure captions

Fig. 1 Map of the study area. Locations of the NCEP precipitation stations are marked with asterisks.

Fig. 2 Scatterplots of observed and predicted precipitation in the case of the non–modified NWP predictors. The solid line is the perfect one–to–one fit. The dashed line is the least squares fit to the data. The least square parameters are given in Table 2. The predictions are by (a) the adaptive MOS scheme and (b) the benchmark conventional non–adaptive MOS scheme.

Fig. 3 Similar to Fig. 2 but for the case of linearly modified NWP predictors.

Fig 4 Plots of the POD, FAR, and TS scores. X-marks denote results obtained by the adaptive MOS scheme, circles, denote the benchmark. Solid lines are the results for the case of the non–modified NWP predictors, and dashed lines are results for the linearly modified NWP predictors.

Fig 5 Similar to Fig. 2 but for the case of nonlinearly modified NWP predictors.

Fig 6 Similar to Fig. 4 but dashed lines are for the case of the nonlinearly modified NWP predictors.

Table 1: The correlation and rmse skills of the results obtained by the adaptive and non–adaptive schemes using the testing non modified data (Test 1), the testing data with linearly modified predictors (Test 2), and the testing data with nonlinearly modified predictors (Test 3).

	Test 1		Test 2		Test 3	
	corr.	rmse	corr.	rmse	corr.	rmse
Adaptive	0.74	0.73	0.59	1.05	0.69	0.86
Non–Adaptive	0.83	0.56	0.25	2.76	0.52	1.38

Table 2: The slope and intercept values of the least square fit lines in Figs. 1, 2 and 5 (Test 1, Test 2, and Test 3 correspondingly).

	Test 1		Test 2		Test 3	
	Slope	Inter.	Slope	Inter.	Slope	Inter.
Adaptive	0.74	0.57	0.72	0.82	0.78	0.60
Non–Adaptive	0.73	0.31	0.69	1.76	0.83	0.51

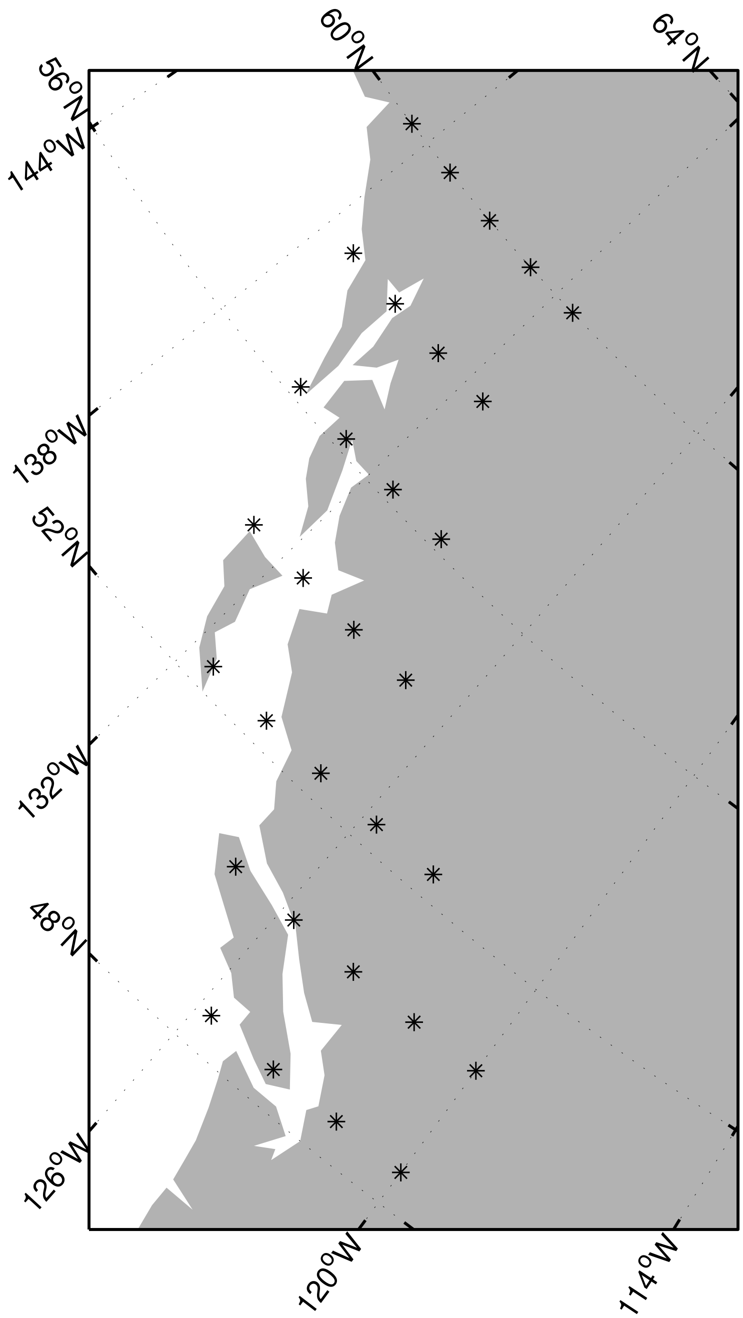

Figure 1: Map of the study area. Locations of NCEP precipitation stations are marked with asterisks.

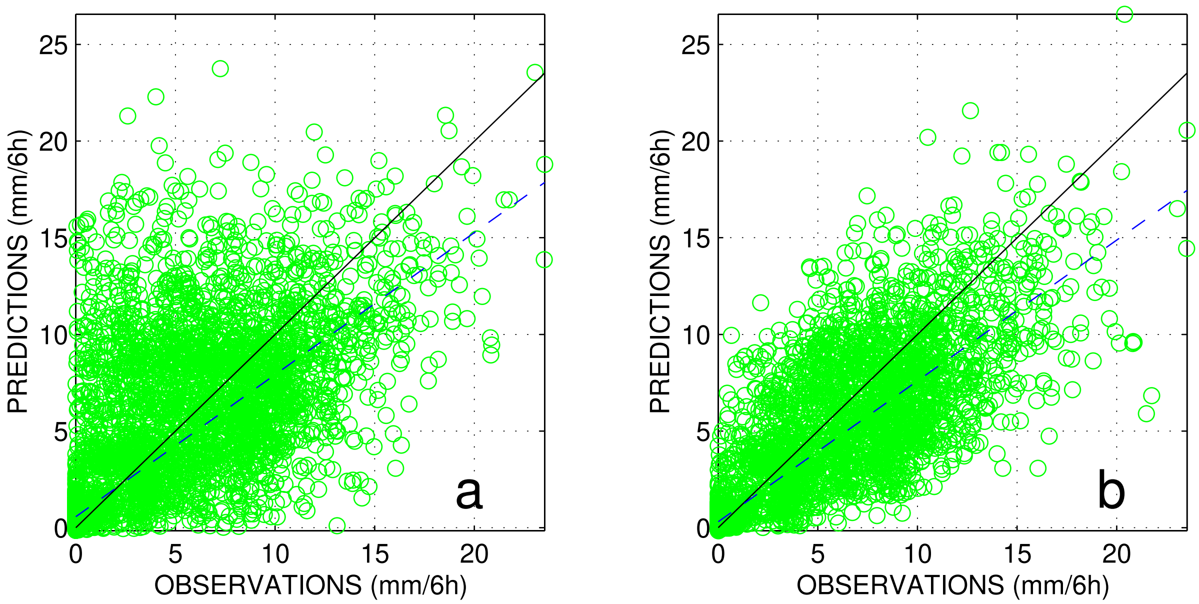

Figure 2: Scatterplots of observed and predicted precipitation in the case of the non–modified NWP predictors. The solid line is the perfect one–to–one fit. The dashed line is the least squares fit to the data. The least square fit parameters are given in Table 2. The predictions are by (a) the adaptive MOS scheme and (b) the benchmark conventional non–adaptive MOS scheme.

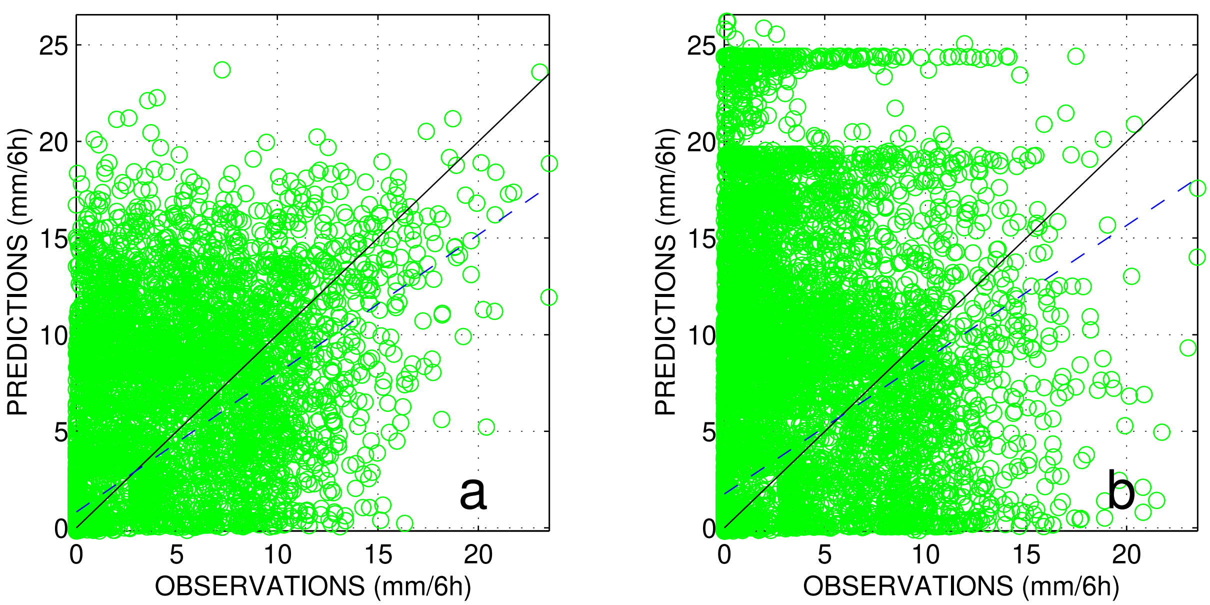

Figure 3: Similar to Fig. 2 but for the case of linearly modified NWP predictors.

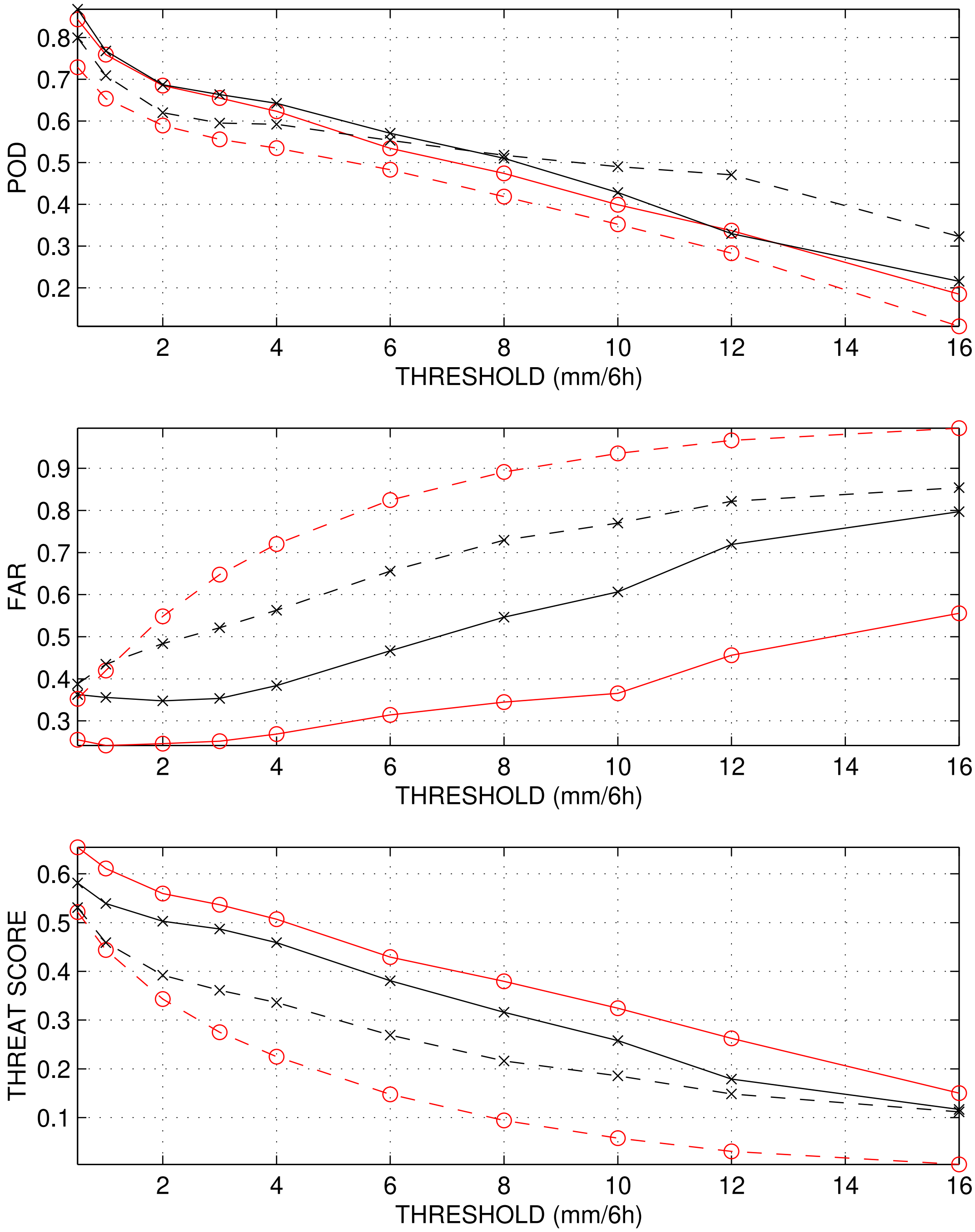

Figure 4: Plots of the POD, FAR, and TS scores. X-marks denote results obtained by the adaptive MOS scheme, circles, denote the benchmark. Solid lines are the results for the case of the non–modified NWP predictors, and dashed lines are results for the linearly modified NWP predictors.

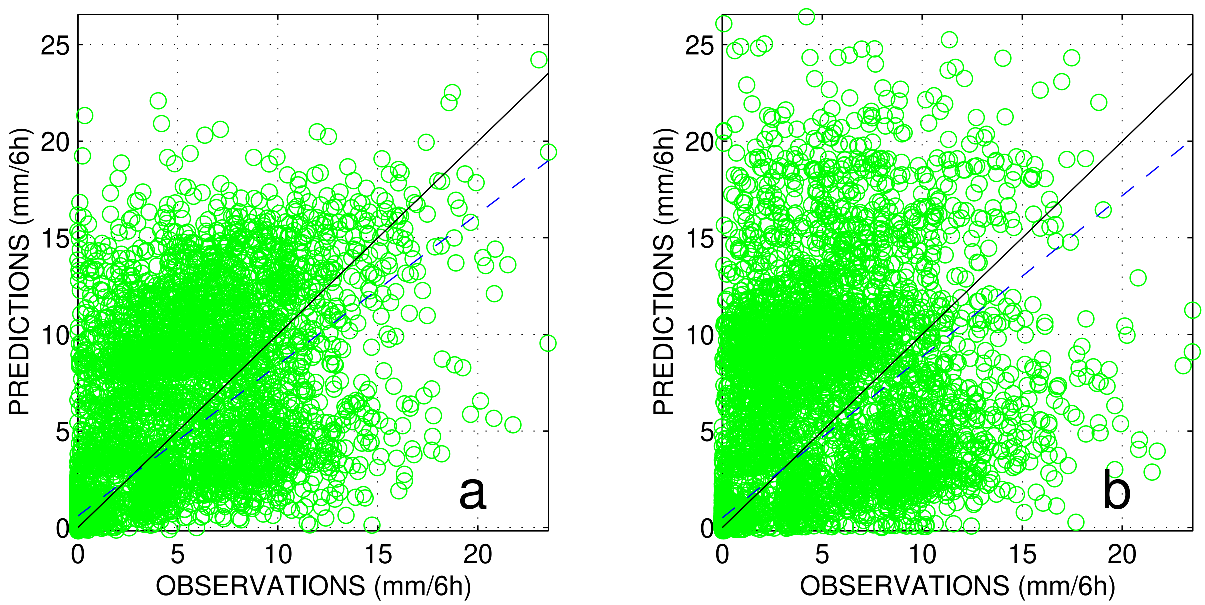

Figure 5: Similar to Fig. 2 but for the case of nonlinearly modified NWP predictors.

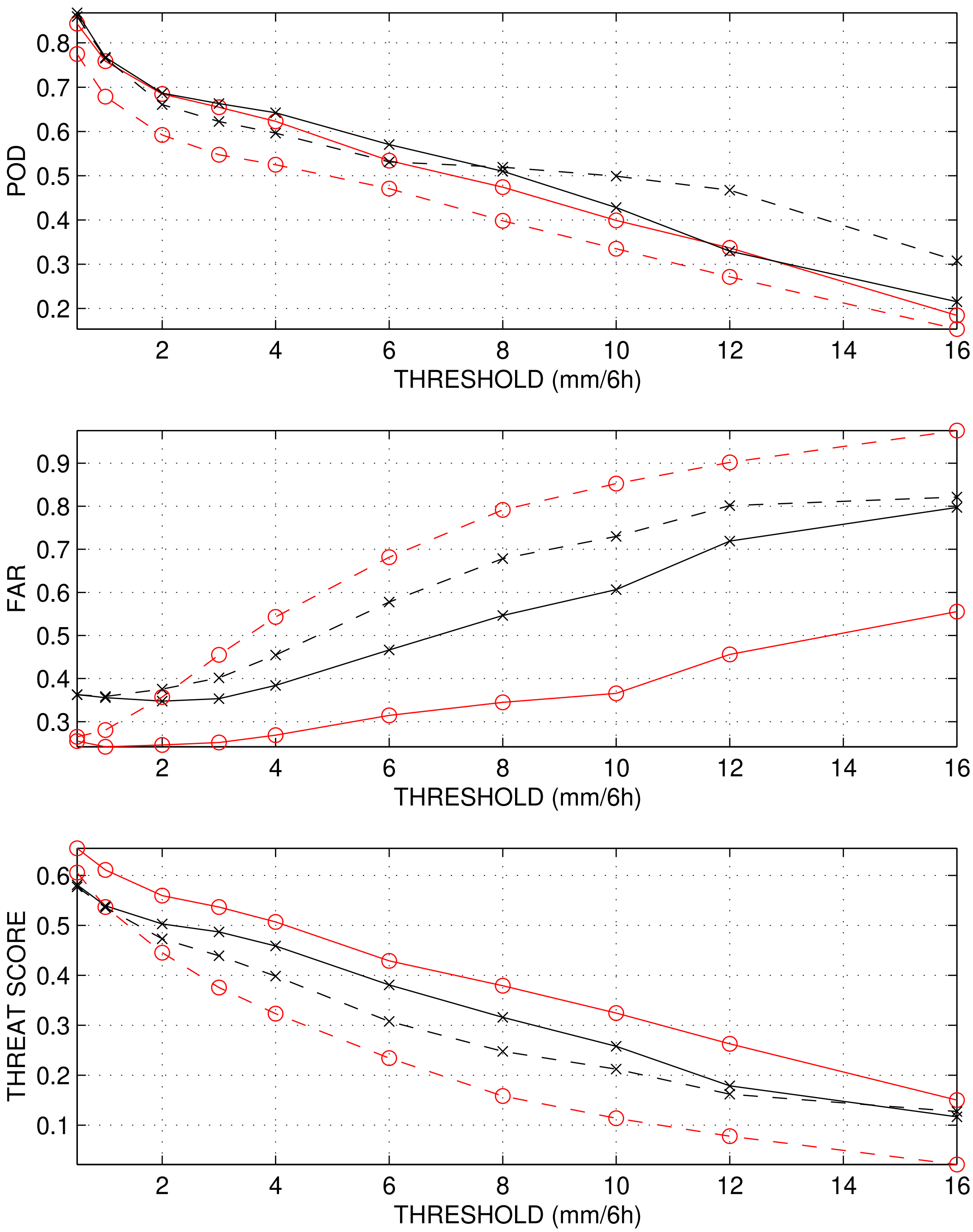

Figure 6: Similar to Fig. 4, but dashed lines are for the case of the nonlinearly modified NWP predictors.