Research Journal of

Environmental Sciences

ISSN 1819-3412

Academic
Journals Inc.

www.academicjournals.com

Research Journal of Environmental Sciences 4 (3): 305-316, 2010
ISSN 1819-3412

Comparative Evaluation of Different Post Processing Methods for Numerical Prediction of Temperature Forecasts over Iran

[1]S. Vashani, [2]M. Azadi and [1]S. Hajjam
[1]Department of Meteorology, Science and Research Branch,
Islamic Azad University, Tehran, Iran
[2]Atmospheric Sciences and Meteorological Research Centre, Tehran, Iran

Abstract: In this study we examined the performance of five post-processing methods on WRF model outputs for daily maximum and minimum temperature forecasts in thirty synoptic meteorological stations over Iran. Direct Model Output (DMO) always contains systematic errors which arise mainly from the simplification of the earth topography in the model and deficiencies in the physics of the model. Different methods for post-processing of these outputs are given to remove the systematic errors. The results of the experiments show all methods are successful in removing the systematic errors in the model outputs. Comparing calculated statistical scores like root mean square error, mean absolute error and mean error indicate that Kalman Filtering (KF) and Artificial Neural Network (ANN) methods are better compared to other methods. Due to the importance of specific temperature thresholds in application, we verified the post-processed temperature forecasts for some specific temperature thresholds. The results of some statistical measure such as Proportion Correct (PC), Treat Score (TS) and False Alarm Rate (FAR) showed satisfactory for various thresholds, but better results have been obtained for higher values of maximum temperature and lowest values of minimum temperature.

Key words: Artificial neural network, Kalman filtering, post-processing

INTRODUCTION

Meteorological factors are unmanageable variants that affect the plant's growth directly and indirectly and can impress all effective factors including short, mid and long-term activities in agriculture. For example, orchardists must decide whether or not to protect their orchards each night during the frost season (Baquet *et al*., 1976). Since, the cost of heating an orchard is substantial and since an entire season's harvest is at stake, forecasts of minimum temperature are needed to optimally weight the tradeoffs (Murphy and Winkler, 1979). In order to be able to act sooner to ameliorate the potential impacts of a frost event instead of just reacting to the situation, very accurate and reliable temperature forecasts with enough lead time has always been important for agricultural section. Meteorologists have tried to provide improved forecasts of temperature from a long ago.

Today, Numerical Weather Prediction (NWP), models have become increasingly valued for predicting minimum and maximum temperature. But the forecasts of NWP models for surface temperature are known to have systematic errors partly due to the poor resolution

Corresponding Author: S. Vashani, Department of Meteorology, Science and Research Branch, Islamic Azad University, Tehran, Iran

of the topography and deficiencies in the physical formulation of the model. Statistical post processing methods have been successful in correcting many defects inherent in NWP model forecasts.

Among them, MOS (Glahn and Lowry, 1972) have been most common, with its own strengths and weaknesses. The MOS technique has proven to be a useful tool for the meteorological community, but to establish reliable forecast guidance, it requires a historical data archive for a long period. Studies indicate that at least 2 years of archived data from model runs are needed to derive a useful MOS equation (Jacks *et al*., 1990; Vislocky and Fritch, 1995). Also, the model framework must be stable during and after the data archive period, which means model configurations must be frozen though continuing advances have been made in both spatial resolution and model physics (Kalnay *et al*., 1996). Thus its use in today's rapidly changing model environment is somewhat limited.

In recent years some techniques have been developed that uses short training period (recent 2-3 weeks) to objectively estimate and adjust current forecast errors and yield refined predictions. Using a short training period enable an updating system to respond quickly to changes in error patterns and a longer training period increases statistical stability.

Previous studies (Stensrud and Skindlov, 1996; Mao *et al*., 1999; Eckel and Mass, 2005; Stensrud and Yussouf, 2005; Mccollor and Stull, 2008) have shown how straightforward moving-average and weighted-average methods with short training periods show improvement upon raw point model temperature forecasts. Also, more sophisticated methods of Kalman filtering (Homleid, 1995; Galanis and Anadranistokis, 2002) and artificial neural network (Marzban, 2003) with short training periods have been employed with success to produce forecast guidance.

In this research we compared the performance of the moving average, weighted average, Kalman filtering and artificial neural network techniques for the post processing of the Weather and Research Forecasting (WRF) model output for daily minimum and maximum temperature forecasts over Iran. Given to the extra importance of extreme values and specific thresholds of temperature forecasts for agricultural aims we divided the maximum and minimum temperature to five parts and verified the post processed output for four quintiles separately by using the contingency table. Also, the verification process was repeated for the zero Celsius temperature thresholds.

MATERIALS AND METHODS

Five post processing techniques are used to post-process the daily maximum and minimum temperature forecasts of the WRF model for thirty synoptic meteorological stations over Iran, during the period from 1 November 2008 through 31 April of 2009. The first used method in this paper is a simple Moving Average (MA). In this method forecast error estimates for the previous n-days are weighted equally to estimate the current day error. The second method is the linear weighted average (LIN), in which the previous n-days are weighted linearly so that higher weight is entitled to more recent day's errors in a linear fashion. The third method is again a weighted average with the COS^2 (COS) weighting function. The forth method is a linear Kalman filtering. This method is a two step predictor-corrector method that corrects/updates the estimated forecast error using the last observed error. Let x_k be forecast error at time step k that is to be predicted. The system equation defines the time dependent evolution of x_k: by a persistence of the current bias plus a Gaussian-distributed random term w_k of variance

$$\sigma_w^2 : x_k = x_{k-1} + w_k$$

where w_k represents the random change from t-1 to t and is assumed to be normally distributed with mean zero and variance W. The observable bias at time step k, y_k is assumed to be noisy, with a normally distributed random error term of v_k of variance σ_w^2: $y_k = x_k + v_k$ The objective is to get the best estimate of x_k, which is termed $\hat{x}_k$, by minimizing the expected mean-square error: $P=E(x-\hat{x})^2$.

For updating the estimate x_k, when the observed bias y_k becomes available, the following equation is used:

$$x_k = x_{k-1} + b_k (y_k - x_{k-1})$$

Where, b_k is the Kalman gain and is determined by the following equation:

$$b_k = \frac{p_{k-1} + \sigma_w^2}{p_{k-1} + \sigma_w^2 + \sigma_v^2}$$

And finally for updating the error covariance term the following equation is used:

$$P_k = (P_{k-1} + \sigma_w^2)(1 - b_k)$$

As suggested from previous studies (Roeger *et al.*, 2003; Dle Monache *et al.*, 2006) the value of 0.01 is used for the ratio of r defined by σ_w^2/σ_w^2. For a more complete description of the Kalman filtering and its application in post processing see e.g., Kalnay (1996) and references therein.

The fifth method is a feed-forward Artificial Neural Network (ANN) that is used to calibrate the Direct Model Output (DMO). In this network the neurons from each layer are connected and propagated forward to all the neurons of the following layer. Figure 1 shows a schematic diagram of feed-forward neural network with three layers: input, hidden and output (Hall *et al.*, 1999).

To assess the relative impact of different post-processing methods several common verification statistics are calculated. They include Mean Absolute Error (MAE), Mean Error (ME), mean squared error (RMSE) and skill score based on MAE defined as follows:

$$MAE = \frac{1}{N}\sum_{K=1}^{N} |O - F|$$

$$ME = \frac{1}{N}\sum_{K=1}^{N} O - F$$

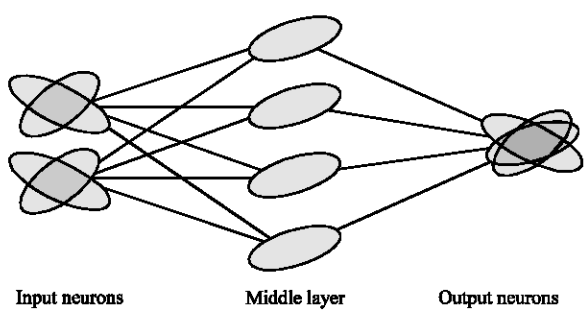

Fig. 1: A schematic of neural network with three layer

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{K=1}^{N}(O-F)^2}$$

$$\mathrm{SKILL\ MAE} = 1 - \frac{\mathrm{MAE}_{PP}}{\mathrm{MAE}_{DMO}}$$

where, N is the number of both observations and forecasts, O_k and F_k the observation and forecast maximum/minimum temperature for the kth day respectively and MAE_{PP} and MAE_{DMO} denote the MAE for the post-processed and direct model output, respectively.

As suggested from previous studies (Eckel and Mass, 2005; Zhong *et al.*, 2005) we considered a window of 14 days training period for all the methods applied here.

Data

Daily WRF forecasts of minimum and maximum surface temperature for 1-5 forecasts days at 15 km grid spacing, initialized at 1200 UTC from the Global Forecast System (GFS) of the National Weather Service's National Centres for Environmental Prediction (NWS/NCEP) are gathered during the period from 1 November 2008 through 31 April of 2009 over Iran. The data consist of observations and bilinear interpolated forecasts of surface temperature at irregularly spaced thirty meteorological synoptic stations located at provincial centres. Figure 2 shows the geographical location of the 30 synoptic stations. The data archive is operated and maintained by the I.R. Iran Meteorological Organization.

Fig. 2: Geographical locations of the thirty synoptic stations in provincial centres over Iran

RESULTS

Using a window of 14 days training period for all the post-processing methods, the statistical measure of ME, MAE, RMSE and SKILL MAE were calculated for daily maximum and minimum temperature forecast periods of days 1-5 and all post-processing methods. Overall, results of all five post-processing methods indicated improvement over the DMO of maximum and minimum daily temperature forecasts, for all forecast days 1-5. Figure 3 shows the calculated ME for DMO and all the post-processing methods for maximum and minimum temperature forecasts. As seen in the Fig. 3a, there is a cold bias, between 3 and 3.5 Celsius in the DMO of daily maximum temperature forecasts. Daily minimum temperature DMO forecast errors ranged from -3.4 to -5 Celsius indicating a warm bias for DMO overnight temperature forecasts (Fig. 3b). All post processing methods have reduced the bias in DMO and ME is close to zero for all forecast days (Fig. 3b).

In terms of mean absolute error (MAE), (Fig. 4a) there is a sharp decrease in MAE especially for short-range daily maximum temperature forecasts for all post processing methods. For maximum temperature forecast DMO the MAE is between 3.1 and 3.8, while after post processing the MAE decreases notably and is between 1.5 and 2 for different methods and forecasts days. Similarly for minimum temperature forecasts the MAE in the DMO ranged from 4.5 to 5.5. That is reduced significantly and ranges from 2.5 to 3 for all post processing methods and forecast days (Fig. 4b).

Calculated Root Mean Squared Error (RMSE) for daily maximum temperature DMO forecasts, shown in Fig. 5a ranged from 3.6 to 5 for days 1-5. All post- processing methods reduced DMO RMSEs ranging from 2.1 to 2.8 Celsius for all forecast days. The daily minimum temperature DMO forecasts error ranged from 4.1 to 5.7 Celsius for days 1-5, which indicates great error dispersion. All post-processing methods reduced the RMSEs from 2.5 to 3.2

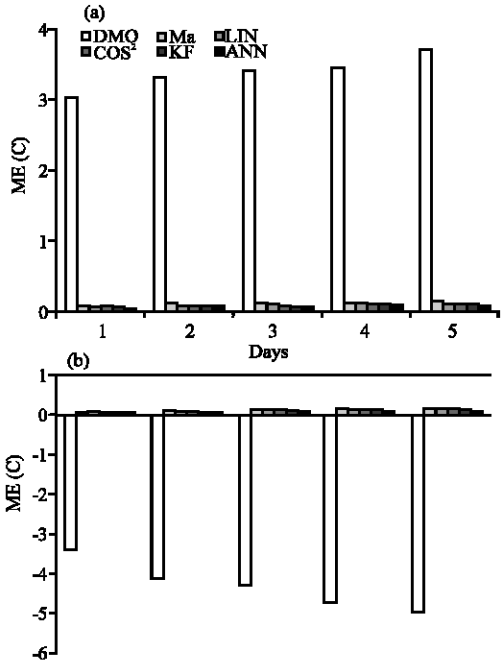

Fig. 3: The mean errors for daily (a) maximum temperature and (b) minimum temperature from the sample for forecast days 1-5

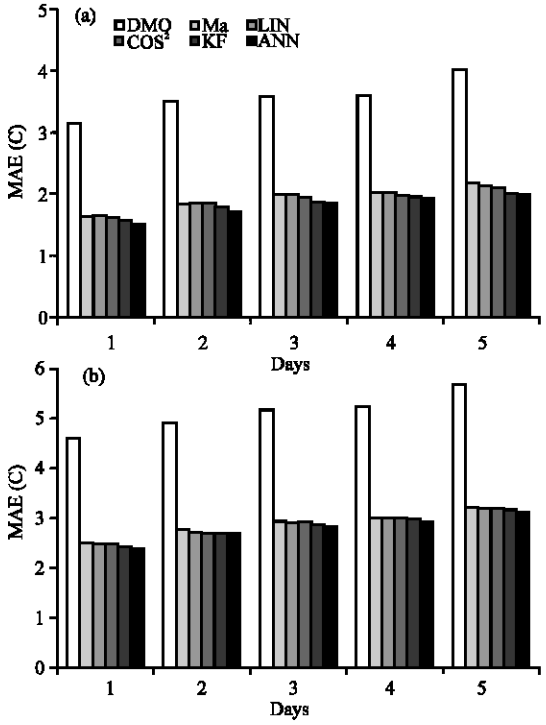

Fig. 4: The same as in Fig. 3 but for MAE

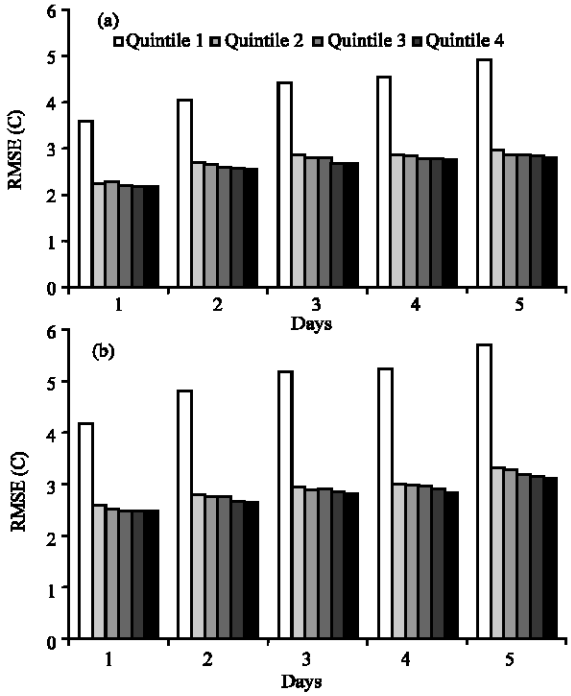

Fig. 5: The same as in Fig. 3 but for RMSE

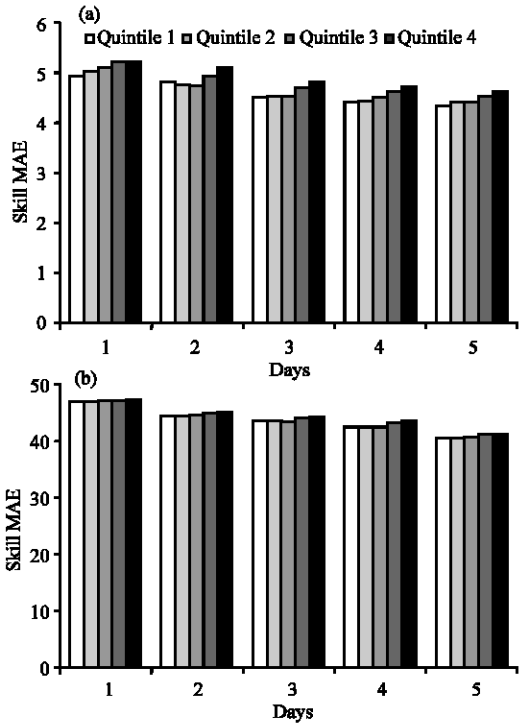

Fig. 6: The same as in Fig.3 but for MAE skill score (relative to DMO)

Celsius for all forecast days. MA, LIN and COS^2 methods show nearly the same reduction in the RMSE over DMO forecasts. KF and ANN methods perform better than other methods while ANN method is the best (Fig. 5b).

To show the skill of different post processing methods over the DMO as the reference forecast, a skill score based on MAE was calculated for both minimum and maximum temperature forecasts (Fig. 6a). The MAE skill score ranges from 0 to 1 with value of zero indicating no improvement skill and a value of one is for perfect forecasting skill. As is seen in Fig. 6a all post processing methods have significant improvement from 45 to 55% over the DMO forecasts. There is no significant difference between post-processing methods but KF method performs best by a slight margin, while ANN technique shows increasing skill for days 1-5.

MAE skill score measured with DMO as the reference forecast is shown in Fig. 6b. MAE skill score indicates that all post processing methods have significant improvement from 40 to 50% over the DMO forecasts. There is no significant difference between post-processing methods but KF and ANN methods perform better by a slight margin, while ANN technique shows increasing skill for days 1-5.

On the whole, the KF and ANN methods showed smaller error and higher skill compared to other methods for both minimum and maximum temperature forecasts.

VERIFICATION

In many cases for agricultural applications, the user is interested to know if the temperature exceeds a particular threshold. In such cases it is suitable to consider the temperature as a discrete quantity by specifying a threshold. It is thus possible to assess the

Table 1: Cell count a is the number of event forecast to occur and did occur. Other cell counts (b, c and d) have similar meanings and n is the total number of forecasts/observations in the sample

Events	Observed	Not observed	Total forecasts
Forecast	a (hits)	b (false alarms)	a+b
Not forecast	c (misses)	d (correct rejections)	c+d
Total observed	a+c	b+d	n = a+b+c+d

quality of the temperature forecasts as if it is a binary event. We present here a brief description of verification procedure for binary events using 2×2 contingency tables, which is discussed in more detailed by Joliffe and Stephenson (2003) and Wilks (2006).

Forecast Verification Using a 2×2 Contingency Table

The joint discrete sample distribution of forecasts and observations is given by a contingency table which is formed by counting different of forecast-observation pairs (Table 1). It is thus defined as follows:

Calculated Quantities in 2×2 Contingency Table

Forecast quality can be assessed using several scalar quantities. Three of them calculated in this study are follows:

Proportion Corrects (PC)

Measures the ratio of correct (hits and correct rejections) to the total number of forecasts expressed as a percentage. It indicates the fraction of correct forecast and is defined as:

$$PC = \frac{(a + d)}{n}$$

It varies between 0 and 1 with value of 1 for a perfect forecast.

Threat Score (TS)

This quantity is similar to PC when correct negatives have been removed from consideration. It shows the fraction of observed and/or forecast events that were correctly predicted and is defines as:

$$TS = \frac{a}{(a + b + c)}$$

False Alarm Ration (FAR)

Measures the ratio of false alarms divided by the total number of events observed. It indicates the fraction of the predicted yes events that did not occur. It is defined as follows:

$$FAR = \frac{b}{(a + b)}$$

PC and TS are positively oriented (the higher the better) and vary between 0 and 1 with value of 1 for a perfect score. FAR also ranges between 0 and 1 but is negatively oriented with values close to zero indicating better forecasts.

VERIFICATION RESULTS

In order to assess the skill of temperature forecast for high and low temperature thresholds, the observed minimum and maximum temperatures for each meteorological station were sorted and divided into five equal parts. The four quintiles were then selected as four temperature thresholds for each station. Due to the fact that 0 Celsius is also an important threshold it was assigned as the fifth threshold for each station. The verification process was then carried out for each threshold separately using contingency tables. Table 2-4 show the calculated statistical scores (PC, TS and FAR) for different post-processing procedures for binary forecast of a specific minimum temperature threshold.

In terms of PC, Table 2 shows that, in general the performances of all methods are better for the first and fourth thresholds. The values of PC for these two thresholds are close to 0.9 for all post-processing methods indicating that in general 90% of the yes/no forecasts for the first and fourth minimum temperature thresholds were correct. The values of PC for the other two thresholds are lower showing a lower skill compared to the first and fourth thresholds.

Table 2: Calculated PC, TS and FAR quantities by using 5 methods for maximum temperature on the first day

Threshold	MA	LIN	COS^2	KF	ANN
PC					
Quintile 1	0.91	0.90	0.91	0.91	0.91
Quintile 2	0.74	0.71	0.76	0.84	0.85
Quintile 3	0.75	0.73	0.74	0.81	0.81
Quintile 4	0.89	0.91	0.89	0.88	0.87
TS					
Quintile 1	0.60	0.62	0.61	0.64	0.64
Quintile 2	0.42	0.40	0.35	0.35	0.37
Quintile 3	0.37	0.36	0.37	0.41	0.38
Quintile 4	0.60	0.65	0.60	0.62	0.61
FAR					
Quintile 1	0.21	0.19	0.20	0.18	0.20
Quintile 2	0.50	0.85	0.41	0.42	0.42
Quintile 3	0.41	0.49	0.50	0.45	0.45
Quintile 4	0.30	0.37	0.31	0.31	0.31

Table 3: Calculated PC, TS and FAR quantities by using 5 methods for minimum temperature on the first day

Threshold	MA	LIN	COS^2	KF	ANN
PC					
Quintile 1	0.75	0.77	0.73	0.84	0.86
Quintile 2	0.57	0.67	0.67	0.78	0.77
Quintile 3	0.71	0.69	0.71	0.72	0.74
Quintile 4	0.81	0.78	0.80	0.86	0.87
TS					
Quintile 1	0.51	0.52	0.50	0.51	0.52
Quintile 2	0.40	0.39	0.42	0.44	0.44
Quintile 3	0.41	0.41	0.39	0.41	0.51
Quintile 4	0.60	0.59	0.61	0.62	0.63
FAR					
Quintile 1	0.21	0.39	0.40	0.31	0.29
Quintile 2	0.50	0.45	0.51	0.41	0.39
Quintile 3	0.41	0.49	0.56	0.43	0.43
Quintile 4	0.26	0.25	0.29	0.25	0.24

Table 4: Calculated PC, TS and FAR quantities by using 5 methods for 0°C on the first day

Quantities	MA	LIN	COS^2	KF	ANN
PC	0.82	0.84	0.85	0.88	0.90
TS	0.72	0.74	0.74	0.75	0.76
FAR	0.10	0.11	0.12	0.10	0.10

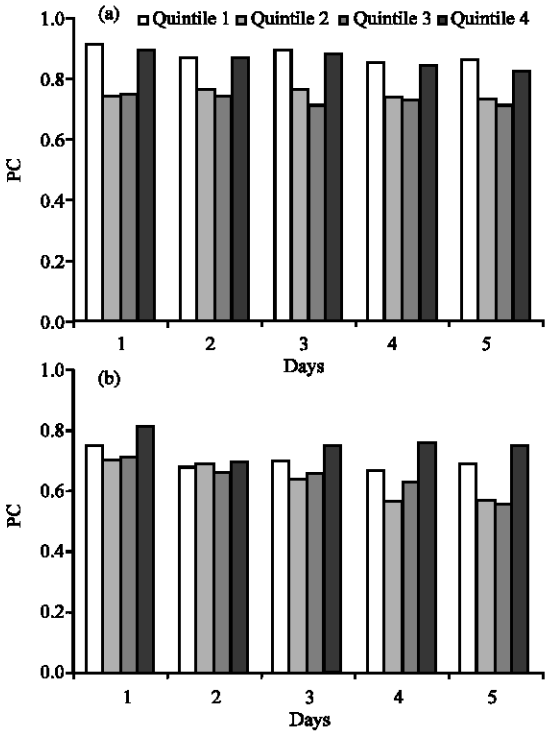

Fig. 7: Proportion correct (PC) for daily (a) maximum temperature and (b) minimum temperature post processed with moving Average technique sample forecast day 1-5. The different Quintiles, from left to right are Quintile 1 to 4. As shown in the legend

60% of the issued/occurred yes forecasts were correctly predicted. The values of TS are close to 0.4 for the other two thresholds. Similar results hold for FAR Table 2 the values of FAR are around 0.2 for the first threshold indicating that in average 20% of the yes forecast for maximum temperature did not occur. There is not much difference between the performances of different post-processing techniques though KF and ANN perform slightly better.

Table 3 shows the same scores as in Table 2 but for minimum temperature threshold forecasts. As seen in Table 3 the results of PC are similar to those for maximum temperature, namely the results for the first and fourth thresholds are slightly better compared to other two thresholds (first and second) and KF and ANN score relatively better than other three methods (MA, LIN and COS^2).

Table 3 shows that the results of TS and FAR are similar to those for maximum temperature for different threshold.

Table 4 shows the calculated TS score for binary temperature forecasts with 0 as the threshold. As is seen in the table, around 90% of the times the below zero Celsius temperature forecast were correct (PC˜ 0.9). Calculated values for TS, presented in Table 4 shows that, around 75% of the forecast/observed below 0 Celsius were correctly predicted (TS˜0.75). Only 10% of the forecast for below zero degree forecast were false alarms (FAR˜0.1). Also, it is seen that KF and ANN perform better than the other three methods for zero degree threshold (Table 4).

As, was seen at the previous figures, the maximum and minimum temperature forecast errors, increased from day 1 to 5 for all the post-processing methods. Figure 7 shows the value of PC score for maximum and minimum temperature forecast using MA method

calculated at different thresholds (four quintiles) for forecast day 1-5. It is seen that the value of PC decreased approximately 10% from days 1 to 5 (Fig. 7). The results are similar for other methods.

DISCUSSION

The atmospheric temperature is one of the most important weather factors that affect the human's activity. Today, meteorological organizations use post-processed model outputs for the forecast of near surface parameters to establish reliable model guidance. Most common methods of MOS (Glahn and Lowry, 1972) and PPM require a long historical data archive for reliable model guidance (Jacks *et al*., 1990; Vislocky and Fritch, 1995). In recent years some techniques have been developed that use short training periods (recent 2-3 weeks) to objectively estimate and adjust current forecast errors and yield refined predictions (Stensrud and Skindlov, 1996; Mao *et al*., 1999; Eckel and Mass, 2005; Stensrud and Yussouf, 2005; McCollor and Stull, 2008). We have compared the performance of five different techniques for the post-processing of daily minimum and maximum temperature forecasts over Iran. The main feature of the post-processing methods used is that they remove the bias in the temperature forecasts without need to long historical data archive, so they can adapt to weather changes quickly and can be used even when the model configuration changes and can be implemented for new stations easily.

CONCLUSION

Results showed that all the methods eliminate the bias and reduce the mean error close to zero. Mean absolute error was also reduced for all forecast days and post-processing methods, though KF and ANN methods gave better results compared to MA, LIN and COS^2 methods. The percentage of successful forecast increased to an acceptable level suitable for operational purposes. In general, KF and ANN methods showed better results compared to other methods and ANN performs slightly better compared to KF for longer forecasts ranges (3 days and above).

In order to see the performance of different methods for high and low temperature thresholds, four thresholds were determined using the four quintiles for observed maximum and minimum temperature. Different statistical scores associated with the contingency tables were calculated for both minimum and maximum temperature forecasts for each threshold. Zero Celsius was also determined as a separate threshold for minimum temperature forecasts. In general better scores (PC, FAR and TS) were obtained for the first and the fourth quintiles. Also maximum temperature forecast showed better results compared to minimum temperature forecasts. Calculated scalars for minimum temperature forecast using zero Celsius as the threshold showed that between 82% for MA and 90% for ANN of the forecasts were correct and the FAR was relatively low around 0.1.

It should be noted, however that all the methods used in this study rely exclusively on statistical properties of the forecast errors and do not use the physics inherent in the relationship between different model variables directly. Such a relationship is accommodated most properly in MOS approach. The main advantage of ANN over other methods used here is that it is possible to use ANN with more than one variable as predictor and thus incorporate the physics for better results (Marzban, 2003).

REFERENCES

Baquet, A.E., A.N. Halter and F.S. Conklin 1976. The value of frost forecasting: A Bayesia appraisal. Am. J. Agric. Econ., 58: 511-520.

Eckel, F. and C.F. Mass, 2005. Aspects of effective mesoscale short-range ensemble forecasting. J. Wea. Forecasting, 20: 328-350.

Galanis, G. and M. Anadranistakis, 2002. A one dimensional kalman filter for the correction of near surface temperature forecasts. J. Meter. Appl., 9: 437-441.

Glahn, H. and R. Lowry, 1972. The use of Model Output Statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11: 1203-1211.

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

Homleid, M., 1995. Diurnal corrections of short-term surface temperature forecasts using the kalman filter. J.Wea. Forecasting, 10: 689-707.

Jacks, E., J.B. Bower, V.J. Dagostaro, J.P. Dallavalle, M.C. Erickson and J.C. Su, 1990. New NGM-based MOS guidance for maximum/minimum temperature, probability of precipitation, cloud amount and sea surface wind. J. Wea. Forecasting, 5: 128-138.

Joliffe, I.T. and D.B. Stephenson, 2003. Forecast Verification: A Practitioners Guide in Atmospheric Science. John Willey and Sons Ltd., New York, USA.

Kalnay, E., G.J. Di Mego, S.J. Lord, H.L. Pan and M. Iredell *et al*., 1996. Recent advances in modeling at the national centers for environmental prediction. Preprints, 11th Conference on Weather Prediction, Norfolk, VA, Amer. Meteor. Soc., J3–J8.

Mao, Q., R.T. Mc-Nider, S.F. Mueller and H.H. Jaung, 1999. An optimal model output calibration algorithm suitable for objective temperature forecasting. J. Wea. Forecasting, 14: 190-202.

Marzban, C., 2003. Neural network for post processing model output: ARPS. Mon. Wea. Rev., 131: 1103-1111.

Mccollor, D. and R. Stull, 2008. Hydro meteorological accuracy enhancement via postprocessing of numerical weather forecasts in complex terrain. Wea. Forecasting, 23: 131-144.

Monache, D.L., T. Nipen, X. Deng, Y. Zhou and R. Stull, 2006. Ozone ensemble forecasts: 2. A kalman filter predictor bias correction. J. Geophys. Res., 111: D05308-D05308.

Murphy, A.H. and R.L. Winkler, 1979. Probabilistic temperature forecasts: The case for an operational program. Bull. Am. Meteorol. Soc., 60: 12-19.

Roeger, R., D. Stull, J. Mcclung, X. Hacker, Deny and H. Modzelewsh, 2003. Verification of mesoscale numerical weather forecasts in mountainous terrain for application to avalanche prediction. J. Wea. Forecasting, 18: 1140-1160.

Stensrud and N. Yussouf, 2005. Bias-corrected short range ensemble forecasts of near surface variables. J. Met. Appl., 12: 217-230.

Stensrud, D.J. and J.A. Shindlov, 1996. Grid point predictions of high temperature from a mesoscale model. J. Wea. Forecasting, 11: 103-110.

Vislocky, R.L. and J.M. Fritch, 1995. Generalized additive models versus linear regression in generating probabilistic MOS forecasts of aviation weather parameters. J. Wea. Forecasting, 10: 669-680.

Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences. 2nd Edn., Academic Press, San Diego, ISBN-10: 0-12-751066-1.

Zhong, S., H.J. In, X. Bian, J. Charney, W. Heilman and B. Potter, 2005. Evaluation of realtime high-resolution MM5 predictions over the Great Lakes region. J. Wea. Forecasting, 20: 63-81.