Mathematical Problems of Computer Science 33, 59–65, 2010.

Statistical Postprocessing of the Output from Numerical Weather Prediction System

Tigran V. Khotsanyan, Vladimir G. Sahakyan and Irina A. Safaryan

Institute for Informatics and Automation Problems of NAS RA

Abstract

This article is devoted to the investigation of relation between really observed meteorological parameters and the forecast of the same parameters produced by a numerical weather prediction system. A statistical correction method of the output from the model of the system is suggested which improves the forecast by coefficient of overlapping, root mean square error and Pearson's correlation.

1. Introduction

Many different methods can be applied for prediction of weather parameters. The numerical weather prediction (NWP) methods are based on the numerical solution of differential equations, which describe the behavior of global weather systems and use the power of computers. The principles of NWP systems functioning are presented by Krishnamurti and Bounoua in [1]. It is desirable that in addition to mean products NWP systems provide timely error correction of the forecast. As noted by Leslie, Fraedrich and Glowacki [2], there are two main kinds of statistical correction schemes. those that correct the numerical model of NWP. and those that simply correct the output from a dynamical model. The statistical techniques for fitting the dynamical model output to real detected data is so called model output statistics (MOS) technique. The technique statistically post-processes NWP forecasts and involves determining a regression relationship between predictors (NWP forecasts, previous observation and climatological information) and predictands (observed weather elements). Hansen and Emanuel [3] presented MOS exploiting for initial condition correction case applying the previous states of system as predictors.

The regression analysis is successful in correction of defects in short-term forecast. For medium and extended range forecasts it can be only used if problems such as the limits of predictability and forecast breakdown are solved. The MOS technique combining with other methods can be used not only for forecasts archive investigation but also for current forecast correction and stochastic modeling to generate ensemble forecasts (see, for example, Clark and Hay [4], Kumar, Maini, Rathore and Singh [5], Silva, Meza and Varas [6]). MOS technique is assigned to statistical web-analysis and correction NWP products and elaboration site-specific guidance. The Meteorological Development Laboratory of the National Weather Serves (NWS) [7] produces guidance for forecasters trough MOS. Description of the guidance for short-term projection can be found by Dallavalle, Erickson and Maloney in [8]. Marzban, Sandgathe and Kalnay [9] compared two postprocessing procedure MOS and Perfect Prog

(PP) in mathematical point of view and suggested a new real-time postprocessor combined both of witch.

In this paper we investigate the quality of the air temperature forecast in Yerevan for period from 01.09.09 to 18.11.09. It is the purpose of this study to demonstrate how raw NWP forecast can be improved owning to statistical postprocessing. Together with traditionally used scores of the skill such as root mean square error (RMSE) and correlation coefficient of Pearson a new score, the coefficient of overlapping (OVL), measuring the amount of agreement of two probability distributions is applied. Statistical properties and possibility of practical applications of OVL are presented by Schmid and Schmidt in [10]. A postprocessor correction method is found that improves the forecast regarding to all of three scores of the skill on the base of regression model. The offered modification of this correction allows to enhance the future forecast. The facility for proposed analysis and error correction statistical procedure is implemented in created and developed for the territory of Armenia NWP web environment presented by Khotsanyan in [11].

2. Measures of forecast skill

Mean square error (MSE) and root of mean square error (RMSE) are used usually as descriptive measures of forecast accuracy. If $X = (x_1, ..., x_N)$ and $Y = (y_1, ..., y_N)$ are vectors of forecasted and actually observed values of some parameter and $\mathbf{e} = (e_1, ..., e_N, e_n = x_n - y_n$ is the forecast errors vector then MSE expressed as

$$\mathrm{MSE} = \frac{1}{N}\sum_{n=1}^{N} e_n^2.$$

and

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}. \tag{1}$$

The assessment category for forecast is realized according [2] in following way. The forecast is very good if RMSE≤ 2.5, good if $2.5 \leq$RMSE≤ 2.9, moderate if $2.9 <$RMSE≤ 3.4 and poor if RMSE> 3.4. We shall also use such assessment in this work though hereinafter it needs revision. The skill of forecast can be explored also on the graph of errors on the time. For a good model, the forecast errors should vary in a horizontal band around zero.

As a score of the forecast skill showing that real and forecasted data have the same trend of the time frequently the Pearson correlation coefficient is applied

$$\mathrm{cor}(X,Y) = \frac{\sum(x_n - \overline{x})(y_n - \overline{y})}{\sqrt{(x_n - \overline{x})^2(y_n - \overline{y})^2}}. \tag{2}$$

The forecast is called correct if X and Y considered as samples from random variables with densitye functions $g_X(x)$ and $g_Y(x)$ coincide on distributions. For these properties checking in many works the scatterplot X versus Y is used (see, for instants [1] and [2]). Departure of points on the scatterplot from straight line y=x indicates incorrect forecasts. We shall use the overlapping coefficient [7] which sampling form can be represented as

$$\mathrm{OVL}(X,Y) = \frac{1}{2}\left(\frac{1}{N_1}\sum_{n=1}^{N_1}\min\left\{1, \frac{\hat{g}_{Y,N_2}(x_n)}{\hat{g}_{X,N_1}(x_n)}\right\} + \frac{1}{N_2}\sum_{n=1}^{N_2}\min\left\{1, \frac{\hat{g}_{X,N_1}(y_m)}{\hat{g}_{Y,N_2}(y_m)}\right\}\right), \tag{3}$$

where $\hat{g}_{X,N_1}(x_n)$ and $\hat{g}_{Y,N_2}(x_n)$ are kernel estimates of corresponding densities functions.

Table 1: Skill scores for the forecast of the air temperature in Yerevan 01.09.09-18.11.09

Forecast horizon (hour)	OVL	Correlation	RSME	Categories on RMSE
00	0.759	0.704	4.25	poor
03	0.800	0.672	3.49	poor
06	0.695	0.0.930	3.22	moderate
09	0.521	0.898	5.40	poor
12	0.533	0.912	5.79	poor
15	0.761	0.908	3.35	moderate
18	0.874	0.876	2.79	good
21	0.903	0.811	2.90	good
24	0.837	0.781	3.12	moderate
27	0.793	0.701	3.61	poor
30	0.739	0.912	5.31	poor
33	0.582	0.890	5.30	poor
36	0.604	0.887	3.93	poor
39	0.733	0.919	3.67	poor
42	0.801	0.905	2.91	moderate
45	0.797	0.838	3.17	moderate
48	0.740	0.835	3.25	moderate
51	0.701	0.745	3.67	poor
54	0.684	0.928	3.34	poor
57	0.589	0.923	5.24	poor
60	0.586	0.877	6.40	poor
63	0.653	0.937	3.85	poor
66	0.696	0.940	2.99	moderate
69	0.700	0.859	3.37	poor
72	0.650	0.878	3.55	poor

Obviously $\mathrm{OVL}(X.Y) = 1$ if and only if the distributions of X and Y are equal and $\mathrm{OVL}(X,Y) = 0$ if and only if supports of the distributions of X and Y have no interior points in common. The overlapping coefficient is a quantitative measure of the skill which is sensitive unlike Pearson's correlation to bias (systematic error) of the forecast. Below results of estimation temperature air of forecast in Yerevan for 80 days over quality measures (1), (2) and (3) are brought. The forecasted data set consists of predicted values of temperature given for three days onward at interval three hours. Thus, each vector $R_t = (r_{t,1}, ..., r_{t,N})$ of actually observed values at hour t is compared with to three vectors of predicted values $F_t = (f_{t,1}, ..., f_{t,N})$, $F_{t+24} = (f_{t+24,1}, ..., f_{t+24,N})$ and $F_{t+48} = (f_{t+48,1}, ..., f_{t+48,N})$ given in the same day, on the eve and two days earlier respectively. Forecast skill scores are shown in Table 1.

It is seen from Table 1, that only for hour horizons 18 and 21 forcasts are good. For these forecasts least values of RMSE and largest values of OVL are obtained. The Pearsons correlation coefficient is highly significant for all hours, but not connected with RMSE and OVL. This fact confirms our a priori suggestion that main source of errors is forecast bias. Thus, the considered forecast needs correction.

3. Statistical postprocessing and error correction

Errors of the NWP model forecast can be corrected by statistical postprocessing of its output. Both of the most popular postprocessing methods (MOS and PP) are based on relating the forecast model to the linear regression model with predictand (weather parameter) R, predictors $V_1, \ldots, V_L$ and errors vector Z of the following form

$$R = c_0 + c_1V_1 + \ldots + c_LV_L + Z. \tag{4}$$

In this study we consider a particular case of model (4) with two predictors which are NWP raw forecast F and real data set observed on the eve R_{-1}, namely

$$R = c_0 + c_1F + c_2R_{-1} + Z. \tag{5}$$

The selection of predictors is determined by the properties of the following forecast models bias correction forecast $F1 = F + (\overline{z} - \overline{f})$, $\overline{z} = \frac{1}{N}\sum_{n=1}^{N} z_n$, $\overline{f} = \frac{1}{N}\sum_{n=1}^{N} f_n$, and persistence forecast $F^{*} = R_{-1}$.

The persistence forecast assumes that the condition at the time of the forecast will not change and is the best by probabilistic approach to prediction (see for details Gneiting. Balabdaoui and Raftery [12]). Since forecast $F1$ reduces RMSE, F^{*} maximizes OVL and the NWP model output F (as it can see in Table 1) significantly correlates with R, we can expect, that their linear function

$$F2 = c_0 + c_1F + c_2R_{-1}, \tag{6}$$

will improve the prediction in average on all three scores of the skill. Coefficients c_0, c_1, c_2, can be estimated according the regression equation (5) from different training samples. We have chosen the whole considered data set (N=80 day) as a training one and apply obtained estimates $\hat{c}_0(N)$, $\hat{c}_1(N)$ and $\hat{c}_2(N)$ to the same sample. Results of comparing forecasts F, $F1$, F^{*} and $F2$ with estimated coefficients are shown on Figures 1, 2, and 3.

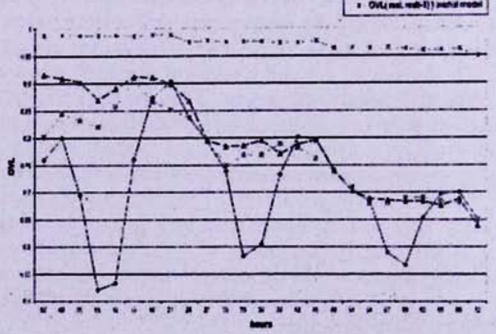

Figure 1. Overlapping coeffcient.

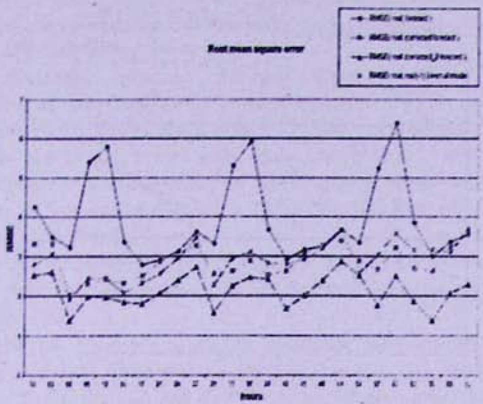

Figure 2. Root mean square error.

As it is expected the forecast $F2$ is better in average than F. $F1$ and F^* on scores RMSE and Pearson's correlation and the closet to the persistence forecast F^* on the score OVL. Thus if the estimates $\hat{c}_0(N)$, $\hat{c}_1(N)$ and $\hat{c}_2(N)$ were known beforehand, then taking them as correction of the raw forecast F in model (6) we would be able to improve significantly the prediction.

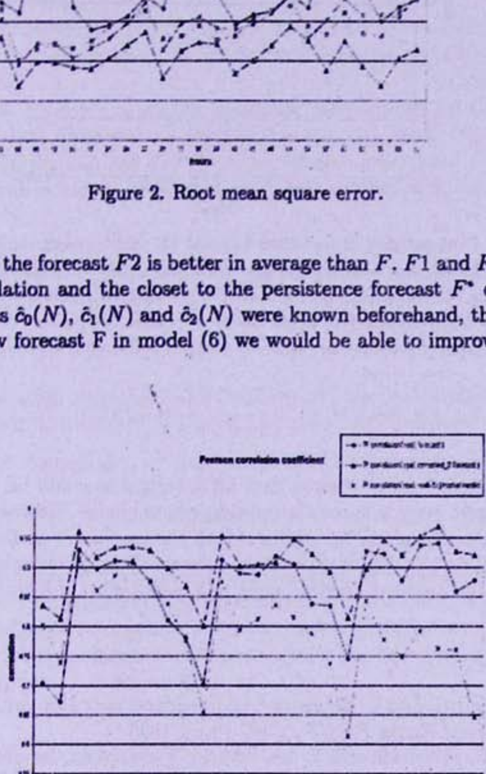

Figure 3. Pearson's correlations.

To develop a real-time postprocessor, we consider the first half of the sample as training and apply obtained estimates of the coefficients $\hat{c}_0(N/2)$, $\hat{c}_1(N/2)$ and $\hat{c}_2(N/2)$ to the second half. This forecast denoted $F3$ is compared with F^* on the RMSE on the Figure 4.

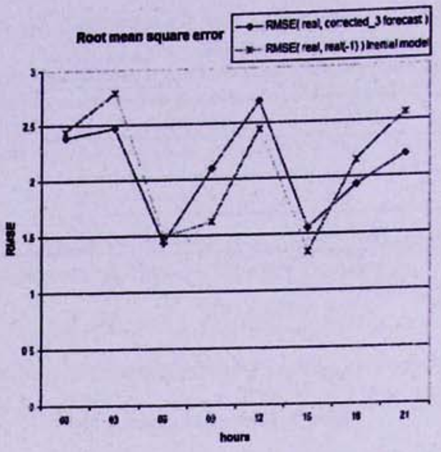

Figure 4. Comparision of real-time forecast F3 with persistence forecast F^{*}.

It can be stated that the coefficients are valid for the second half of sample, since, as it is seen from the Fig. 4, the forecasts $F3$ and F^{*} are approximately equal RMSE skill score.

The model in NWP web site [13] allow chose as predictors powers of forecast and previous real values of the weather parameter. Moreover, the coefficients are not nessesarily regressien model (4).

4. Conclusion

The example considered above showed that MOS-technique could be useful in correction of the raw NWP forecast even in case of simple regression model. Meanwhile the main problem is in exact definition of the period during which the regression coefficients estimated from training sample can be utilized for making the forecast. Further investigation in this direction would be devoted to the mentioned problem and elaboration of a real-time postprocessor.

References

[1] T.N. Krishnamurti, and L. Bounoua,"An Introduction to Numerical Weather Prediction Techniques",*Boca Raton FL US:, CRC Press*, 1996.

[2] L.M. Leslie , K. Fraedrich and T. J. Glowacki "Forecasting the skill of a regional weather prediction model", *Monthy Weather Review*, vol. 117, pp. 550-557, 1989.

[3] Hansen J. A. and K. A. Emanuel, "Forecast 4D-Var: Exploiting Model Output Statistics" *Quarterly Journal of Royal Meteorol. Soc.* vol. 129, pp. 1255-1267, 2003.

[4] M. P. Clark and L.E. Hay, "Use of medium-range numerical weather prediction model output to produce forecasts of streamflow",*Journal of Hydrometeorlogical*, vol. 5, pp. 15-32, 2004.

[5] A. Kumar, P. Maini, L.S. Rathore and S.V. Singh, "Skill of statistical interpretation forecasting system during monsoon season in India", *Atmospheric Science Letters*, vol. 10, p. 13, 2000.

[6] D. O.Silva, F.J. Meza, and E. Varas, "Use of mesoscale model MM5 forecasts as proxies for surface meteorological and agroclimatic variables", *Ciencia e Investigacion Agraria, 36(3)*, pp. 369-380, 2009.

[7] NOAA's National Weather Service. Current MOS Forecast Production, http://www.nws.noaa/mdl/synop/products.php

[8] J. P. Dallavalle, M. C. Erickson, and J. C. Maloney, "Model output statistics (MOS) guidance for short-range projections", *20th Conference of Weather Analysis and Forecasting, Seattle, Amer. Meteor.Soc.*, Preprints, pp.1-11, 2004.

[9] C. Mazbran, S. Sandgathe, E. Kalnay, "MOS, Perfect Prog and Reanalysis", *Mon. Wea. Rev.*,134, pp. 658-663, 2006.

[10] F. Schmid and A. Schmidt, "Nonparametric estimation of coefficient of overlapping - theory and empirical application", *Computational Statistics and Data Analysis*, vol. 50, pp. 1583-1596, 2006.

[11] T.V. Khotsanyan, "Numerical weather prediction environment for the territory of Armenia",*Mathematical Problems of Computer Science* vol. 33, 83–90, 2010.

[12] T. Gneiting, F. Balabdaoui and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness", *Technical Report*, N483, pp. 1-31, Department of Statistics, University of Washington, 2005.

[13] http://www.cluster.am/wrf/stat

[14] D. S. Wilks, T. M. Hamill. "Comparison of ensemble-MOS using GFS reforecast", Mon. Wea. Rev.135, pp. 2380-2389, 2007.

Եղանակի թվային կանխատեսման համակարգի արդյունքների վիճակագրական վերամշակում

Տ. Խոցանյան, Վ. Սահակյան և Ի. Սաֆարյան

Ամփոփում

Աշխատանքը նվիրված է իրականում դիտարկված և կանխատեսված միևնույն օդերևութաբանական պարամետրերի հարաբերության հետազոտմանը: Առաջարկված է կանխատեսված տվյալների ուղղման վիճակագրական եղանակ, որը բարելավում է կանխատեսումը ըստ ծածկվածության գործակցի, միջին քառակուսային սխալի արմատի և Պիրսոնի հարաբերակցության գործակիցը: