J12.15

LONG-RANGE FORECASTING BY EEOF EXTRAPOLATION

Constantin Mares * and Ileana Mares
National Institute of Meteorology and Hydrology, Bucharest, Romania

ABSTRACT

This paper presents an optimum combination of two robust statistical techniques that can be used to improve the skill of long-range weather forecasts. The first method uses decomposition and analysis based on EEOF (Extended Empirical Orthogonal Functions), with a 3-month data window, for temperature and precipitation fields in Romania. Using Rule N to select the significant components led to 3 modes for temperature and to 9 modes for precipitation. In linear extrapolation, an AR model is used to produce forecast the time series of the EEOF components. The parameters of this model are determined by a method consistent with the maximum entropy method, which is why this model is named AR-MEM. In order to select model order, 7 criteria are tested, some of which are efficient, while the others are consistent. Model parameters are determined from observational data over the period 1950 – 1990. The Heidke skill score is computed using independent data (1991-1997). The best results have been obtained for the temperature field filtered by the first 3 EEOF modes, for the meteorological stations situated in the central part of Romania. For precipitation, the forecast based on the EEOF 1 component with one-step ahead, led to skill scores worse than those obtained using persistence in all cases.

1. INTRODUCTION

While many different types of Empirical Orthogonal Functions (EOF) techniques are available (Kim and Wu 1999), Extended Empirical Orthogonal Functions (EEOF) as Weare and Nasstrom (1982) has been chosen for the present study, because it is best suited to forecasting. The conventional EOF identifies the main *patterns* of variability, which are coherent in space. In EEOF, those patterns coherent both in space and in time are identified (Wang et al. 1995). In such analyses, the field is studied in *m* successive moments, that is a mobile *window* is inserted of length *m*. In a traditional EOF analysis, *m* = 1. The way in which the *window* is selected depends on the aim of the analysis. Vautard and Ghil (1989) discussed the *power* of the EEOF method for identifying physical oscillations, as well as for analyzing, filtering and forecasting time series. One disadvantage of the EEOF analysis is that it is ineffective when the spatial coherence is small. Another disadvantage is a large computational memory requirement. Among the relatively recent applications of EEOF is Tangang et al. (1998), where they forecast ENSO events using an extrapolation of the first 7 EEOF components of the sea level pressure field, and the SSTA (Sea Surface Temperature Anomalies) with the help of a neural network model.

* *Corresponding author address*: Constantin Mares, National Institute of Meteorology and Hydrology, Bucuresti-Ploiesti , 97, Bucharest 013686, Romania, e-mail: constantin.mares@meteo.inmh.ro

Compagnucci et al. (2001) studied the 1000 hPa geopotential height field using Principal Sequence Pattern Analysis (PSPA), which is an extension of Principal Components Analysis set in T-mode (Richman 1986). Since the purpose of their paper is to obtain spatial patterns and their time evolution, the PSPA methodology in T-mode proved to be much more efficient than the EEOF analysis in S-mode.

In the present study extrapolation is performed linearly for each principal component (PC) series of the significant EEOF, by means of an autoregressive model (AR).

The rest of this paper is organized as follows. Section 2 describes the data and methods. Section 3 presents the results of extrapolating the first 3 significant principal components series for temperatures, and the first principal component series of EEOF for precipitation, using an AR-MEM model. Summary and concluding remarks follow in Section 4.

2. DATA AND METHOD

The meteorological time series (1950-1997) analyzed in this paper include mean monthly temperature values from 31 stations, and precipitation amounts from 33 stations in Romania, relatively uniformly distributed across the country. In the first stage decomposition and analysis based on EEOF with a 3-month data window, for temperature and precipitation fields in Romania are applied. After that, a linear prediction model is applied to the temperature and precipitation fields for a spatial encompassing Romania, for each component of EEOF and for separately for time series from each stations. The parameters of this model are determined by a method consistent with the maximum entropy method (Burg, 1975), which is why this model is named AR-MEM However, this technique differs from the one used by Barnett and Hasselmann (1979), because the orthogonal basis used to represent meteorological fields, is not in EOF, but in EEOF. Because climate variables are, in general, cyclostationary, the annual cycle has been eliminated from the data by normalizing, with respect to the average and variance of each month. The number of predictors is reduced as the number of significant components. The statistical significance of the retained components is based on Rule N (Preisendorfer, 1988).

In order to determine the grid point predictability for temperature and precipitation, (i.e. at stations), Mares and Mares (1997), utilize a single algorithm (FPE) for selecting the order of an AR-MEM model. In the present paper, new criteria, namely CAT (Parzen, 1977), AIC (Akaike, 1970), AICW (Wilks, 1995), BIC (Schwarz, 1978), HQ (Hannan and Quinn, 1979) and AIC_C (Hurvich and Tsai, 1989) are applied. The use of these criteria as well as filtering the initial field by the significant EEOF improves the forecast skill.

The results on the criteria performance of the simulated time series are presented in Mares and Mares (2003).

The model parameters have been determined from the 1950-1990 observation, and the verifications, by means of the Heidke score for persistence, have been achieved using independent data for the period 1991-1997.

3. EEOF EXTRAPOLATIONS

3.1. Temperature field

The AR-MEM parameters, for the 1950-1990 period have been determined for the first 3 EEOF components. Tests have been performed for the 1991-1997 period (which contain a maximum of 82 values because of the extended EOF with a 3 month window). For the first PC of EEOF the following orders for AR-MEM have been obtained: 25 using FPE, CAT and AIC, 18 using AICW, 9 using BIC, 28 using HQ and 17 using AIC_C. In order to test the model orders on independent data, the correlation coefficients between forecast and observed values have been used. The root-mean-square-error (RMSE) was also considered in this test, because this is a measure sensitive to magnitude errors, while the correlation coefficient is immune to magnitude errors. Comparisons between patterns as well as magnitudes can be made with RMSE. The model orders and correlation coefficients (R) between forecast and observational series for first principal component EEOF are presented in the upper part of Table 1. It is clear that all the criteria have almost the same performance, considering the values of the correlation coefficients and the RMSE. BIC is an exception, it produces larger errors. When the observed and models series are standardized, the RMSE is always less than 1.

Oscillatory components can disturb forecasts based on an AR model, which is why, the existence of oscillatory components need to be determined. These oscillatory components are revealed using a periodogram (Vialar 1968) as well as a power spectrum built with AR-MEM models (Burg 1978; Ulrych and Bishop 1975).

A period of 26 months was significant at the 95% level for the first PC of EEOF. In this case, the periodic series corresponding to 26 months was extracted from the initial series X_1 of EEOF1. The same procedure is used for the residual series, X_2. Once the model orders have been selected by the 7 criteria, then 82 values are extrapolated one step ahead and, finally, the performance of the 7 criteria are tested. The results are presented in the middle part of Table 1, corresponding to the X_2 series. The HQ criterion, with order 18, is sensitive to the nature of the series. Together with AICW and AIC_C, it produces the best results. These results for the residual series X_2 are weaker than for the X_1 series, but the periodic part of the series has not been added yet.

The Heidke score has been chosen for this study in order to determine if the forecast values using an AR-MEM model are better than some other reference forecasts. The Heidke skill score (HSS) is calculated for the persistence forecast as well as for the forecasts based on an AR-MEM model. The improvement of the forecast over persistence is given by the following ratio: **(accuracy of forecast – accuracy of persistence) /(1- accuracy of persistence)**. When based upon a 2 x 2 classical contingency table (Zhang and Casey 2000), with a, b, c and d the elements of the table (as defined in Wilks 1995), the Heidke skill score is defined as: $\mathbf{2(ad-bc)/[(a+c)(c+d)+(a+b)(b+d)]}$. A perfect forecast produces a HSS of 1, while a forecast that is no better than the reference produces a HSS of 0, and forecasts that are worse than the reference have negative HSS values. In this study, the values analyzed using the HSS have been classified in two categories: above and below normal, depending on the mean value of an 82 element time series, which is used in the verification procedure. The periodic time series was then added to forecast series and, in this way, the X_3 series is obtained. For the X_3 series, correlation coefficients as well as HSS compared to persistence are presented in the lower part of Table 1. An improvement of the initial forecast is apparent and this improvement is due to the HQ, AICW and AIC_C criteria for optimum model orders equal to 18, 18 and to 17, respectively. The HSS based on a persistence reference is 0.390 for the first PC of EEOF. In Table 1, HSS represents skill scores in comparison with persistence.

Analyzing Table 1, it can be seen that all the criteria are more useful for the first PC of EEOF temperature, forecast by an AR-MEM model, than forecasts based on persistence. The BIC criterion has the least skill. Based on the HSS, the AR-MEM model is much better than the forecast based on persistence. The FPE, CAT, AIC and AICW criteria indicated the same order of 26 for the second principal component of EEOF, and, same order of 23 for the third principal component The HQ criterion also had the same order (28) in both cases, and BIC and AIC_C had the same order (10) for EEOF2 and order 6 for EEOF3. These orders together with the correlation coefficients between predicted and observed series, and RMSE are presented in Table 2.

TABLE 1. Order of the models selected on 7 criteria, correlation coefficient (R) between forecast and observed values and RMSE for the first PC of EEOF for three cases: predicted initial series X_1, predicted residual series X_2 and final series X_3 obtained by the sum of X_1 and periodic time series. The HSS in comparison with persistence is also indicated for X_3.

EEOF 1	FPE-CAT-AIC	AICW	BIC	HQ	AIC_C
Series X_1					
Model order	25	18	9	28	17
R	0.868	0.866	0.849	0.868	0.867
RMSE	2.877	2.900	3.077	2.881	2.840
Series X_2					
Model order	25	**18**	9	**18**	**17**
R	0.852	0.857	0.819	0.857	0.856
RMSE	2.894	2.842	3.192	2.842	2.853
Series X_3					
R	0.867	**0.872**	0.837	**0.872**	**0.871**
HSS	0.442	0.362	0.280	0.362	0.362

TABLE 2. Performance of AR-MEM models depending on the selected order for temperature PC of EEOF 2 and 3.

Component		FPE-CAT-AIC-AICW	BIC-AIC_c	HQ
E	Model order	26	10	28
E	R	0.652	0.623	0.657
O	RMSE	4.122	4.213	4.099
F	HSS	0.342	0.318	0.342
2				
E	Order	23	6	28
E	R	0.816	0.789	0.814
O	RMSE	2.827	3.009	2.841
F	HSS	0.700	0.700	0.700
3				

For EEOF2, the best criterion is HQ, based on the correlation coefficient and the associated RMSE. For the PC of EEOF3, the best results correspond to FPE, CAT, AIC and AICW criteria, while the weakest to the BIC criterion. Note that the correlation coefficients are higher for EEOF3 than for EEOF2, which are also shown by the associated RMSE values. The HSS for persistence (which is the reference forecast) is negative in both time series corresponding to EEOF2 and EEOF3 and has the following values: -0.074 and –0.467, respectively. In comparison, the performance of AR-MEM model is very good for EEOF3, with an HSS equal to 0.7 for all the orders indicated by the 7 criteria. The skill in this case was higher than for the first PC of EEOF where persistence exists and for which an AR-MEM model is not as useful. The PC of EEOF2 has a HSS equal to 0.342 for the forecasts based on an AR-MEM model for the criteria similar to FPE and for HQ. An HSS is equal to 0.318 resulted for the forecasts performed using the order indicated by BIC and $AIC_{C.}$

The temperature at a station should be predicted with skill using an AR-MEM model and a filtering procedure. In this case, a new time series is constructed, by means of the first three PCs together with the first three EEOF eigenvectors. Then the efficiency of the station forecast was tested using the observed field. For this purpose, correlation coefficients between forecast and observed series, as well as, RMSE associated are used. The FPE, CAT and AIC criteria give same values for the reconstructed series, while the other four criteria yield different orders. For lags 0 and 1 the results are good, the correlation coefficients have a high level of significance corresponding to the 82 values, and the RMSE values are small. For lag 2, the performance of the model decreases rapidly, due to the small length of the window, as expected from von Storch and Frankignoul (1998). They showed that when the lead-time becomes longer than the window length, a sudden drop in skill occurs. For this reason, only the results for lag 0 and 1 will be discussed. For lag 0, the highest correlation coefficient is given by the HQ criterion for 18 of the 31 stations, and by FPE-CAT and AIC criteria is best for the rest of the stations. The highest correlation coefficient (0.954), with RMSE equal to 0.330, was obtained for lag 1 at the Sibiu meteorological station, in the central part of Romania. For Sibiu station, the behavior of the predicted time series (filtered by the three EEOF modes) with one step ahead compared to the observed series is shown in Fig. 1. For lag 1 at Sibiu the HQ criterion produces, a high HSS = 0.762 compared to persistence, which has HSS = - 0.024.

Correlation coefficients, together with RMSE values, obtained using the HQ criterion for all Romanian stations, are presented in Figs. 2 and 3. From these figures, it is clear that the model performs worst in the mountain areas. The best skill, highest correlation coefficient and lowest RMSE, appears in the Southwestern sub-Carpathian area. Several other stations in and around the Carpathian ring also have good skill.

The high variability in temperature over the Romanian territory, due to the mountains, represents a problem element for the methods used in the present study.

Figs 4 and 5 present the correlation coefficients and RMSE values, after having eliminated the stations with the largest variability (Vf. Omu, Fagaras, and Miercurea Ciuc). Figs. 11 and 12 show the capability of the model to improve the monthly temperature forecast in the central parts of the analyzed domain, compared to the boundary zones. The model performs worse in the mountain and in valley stations, where can be a large variability in the analyzed fields.

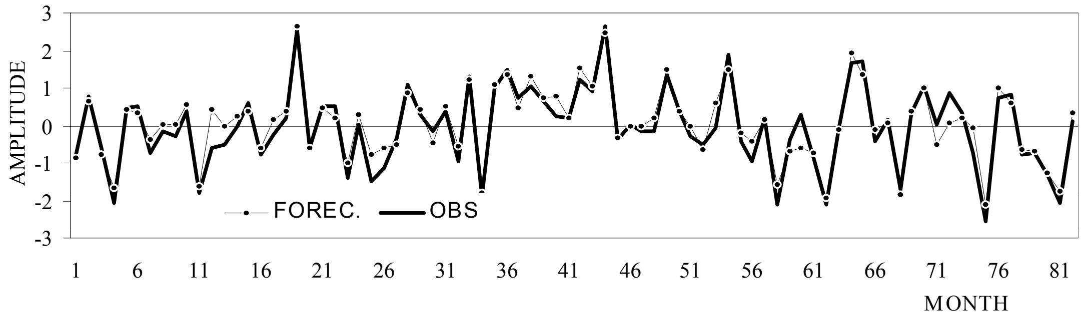

Fig. 1. Forecast and observed temperature at the Sibiu station (situated in centre of Romania) for 1991-1997.

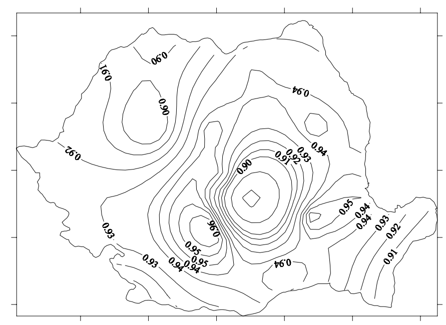

FIG. 2. Spatial distribution of correlation coefficients between the forecast temperature field filtered by the first three EEOF and observations.

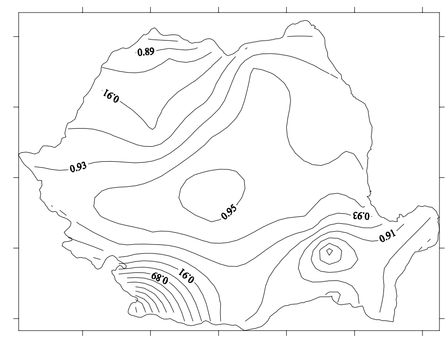

FIG. 4. As FIG. 2, but without the stations: Vf. Omu, Fagaras and Miercurea Ciuc, that have a high variability.

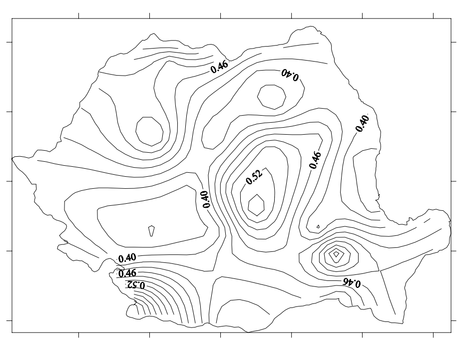

FIG. 3. RMSE associated to the correlation coefficients in FIG. 2.

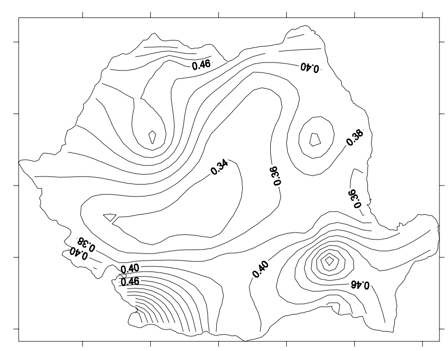

FIG. 5. As FIG. 3, but without the stations: Vf. Omu, Fagaras and Miercurea Ciuc that have a high variability.

3.2. Precipitation field

For the precipitation field, 9 EEOF modes are significant using Rule N. This means 9 components should be used for extrapolation. The same procedures used for temperature are followed for precipitation. Ideally, a detailed analysis should be performed for each retained component and this might be the aim of a separate study. Hence, only the first EEOF component will be considered here. Although this component explains only 19% of the total variance, it yields a good image of the precipitation pattern over Romania, especially if the amounts are interpreted as deviations from climatology.

Following the same procedure as for temperature, the coefficients and model orders of the AR-MEM model are selected by the 7 candidate criteria, for 1950-1990 period while verification tests have been made for 1991-1997 period.

Although the first component has 8 and 23-month periodicities with 95% significance, after extracting the periodic part of the series, extrapolating it, and then adding the periodic part back again, no model skill improvements are noted.

The results for the PC of EEOF1 for precipitation, in accordance with the 7 criteria, are presented in Table 3. The FPE, CAT, AIC and AIC_C all suggest the same order (23), while BIC and AIC_C suggest order 12, and HQ order 27. The HQ criterion gives the best result, but is worse then persistence. The Heidke score for persistence is 0.625, and for all of the criteria the Heidke score is lower. HSS for the forecast is even weaker than the reference one based on persistence. Persistence had large values because the last two decades have been unusually dry in Romania. If the extrapolated forecasts are split into two categories (above and below normal compared to climatology) results are no better because these meteorological elements do not have a Gaussian distribution. A quantiles classification may be more appropriate.

If precipitation displays a strong persistence, modeling with a high order Markov chain may be appropriate using the ergodicity coefficients of the transition matrix (Iosifescu 1980).

TABLE 3. Performance of AR-MEM models depending on the selected order for the first PC of EEOF 1 for the precipitation field.

Component	FPE-CAT-AIC-AICW	BIC-AIC_C	HQ
EEOF 1			
Model order	23	12	12
R	0.834	0.812	0.835
RMSE	2.158	2.292	2.157

4. Summary and concluding remarks

The combined statistical technique of decomposing certain meteorological fields into extended empirical orthogonal functions and, followed by temporal extrapolation, using an AR-MEM model yielded useful results for surface temperature fields in Romania. In this way, periodicities are extracted and extrapolated forecasts display some skill.

In order to examine the periodicity, both a power spectrum based on MEM, and the periodograms, with appropriate statistical significance, are used. Several spectral peaks are evident in temperature field as well as precipitation. For temperature, most significant peak has a period of 26 months. For precipitation, there are two significant peaks: 8 months and 23 months.

The 26 and 23 month-periodicities might be associated with the Quasi-Biennial Oscillations (QBO) in this part of the European continent. Similar periodicities of 2.1 and 2.4 years in the precipitation field have been observed by Tabony (1979) for the Western and Northern part of Europe. The QBO is also observed in other meteorological fields such as sea level pressure (Trenberth and Shin 1984). In addition, there is clear evidence for a coherent QBO signal in time series of surface temperatures in some zones of the Northern Hemisphere, but no persuasive statistical connection exists between the QBO in the equatorial stratosphere and the QBO in the low troposphere. Both stratospheric and tropospheric QBOs, and their possible connections, are explained in the other literature

(Lindzen and Holton 1968; Niwano and Takahashi 1998).

In this study, the 26-month period was extracted from the time series of the EEOF1 component for temperature, (as per Kim and North 1998) which improves the HSS for EEOF1. The HSS was best (0.7) for the EEOF3 component in all-7 cases of tested criteria.

Extrapolating each of the three significant principal components for the temperature field and using them to reconstruct the initial field, filtered by the significant modes, led to a forecast with good skill compared to persistence for meteorological stations situated in the central part of Romania (HSS = 0.756).

For precipitation field, although the first EEOF component describes the mean precipitation over Romania well, application of an AR-MEM did not lead to a better forecast than persistence.

Regarding the efficiency of all the 7 criteria analyzed in this study in order to select the order of an AR-MEM model, the tests led to different results, depending on whether they were applied to simulated or observed time series. For data simulated with a 2 or 4 order AR model, criteria such as BIC determined the model order with great accuracy, but worked poorly for observed data. The FPE, CAT, AIC and AICW criteria indicated for observed data, in most of the cases, the same order, while BIC, HQ and AIC_C estimated different orders. Concerning their forecast skills, BIC had the lowest one and the skills of the other criteria changed depending on the analyzed time series. The, HQ had the highest skill for the series reconstructed of the first 3 EEOF modes for lag 0 and, for some stations, also for lag 1.

The main difference between the behavior of the criteria for simulated and observed data may be due to the fact that simulations have been performed only on 2^{nd} and 4^{th} order AR-models and not on time series fitted with AR models of higher orders. It appears that the efficiency of the various criteria depends a lot on the nature of the data. Kashyap (1980) showed that the criteria for accepting an AR model are different from one application to another. For valid results, the simulations must properly represent the nature of the phenomena being considered. It appears that no single rule can be generally applied for determining the appropriate model order.

There was no skill for forecasts on lag 2 that extrapolate forward one step. This means that it is necessary to increase the length of the data window (compared to 3 months used in the present paper) to take advantage of the EEOF method for obtaining forecast estimations for some months ahead. It is important to emphasize that the analysis of EEOF components, as well as their extrapolations, led to very good results for the central part of the analyzed domain when compared to the border area, where the results are weaker.

For the temperature field in Romania, this method can be applied to the long-range weather forecasts. Unfortunately, results for precipitation are poor. The poor results for precipitation seem to be linked to the fact that precipitation is not normally distributed. In the future, we shall use a procedure of probability transformation in order to bring the precipitation distribution close to a Gaussian one.

REFERENCES

Akaike, 1970: Statistical predictor identification. *Ann. Inst. Stat. Math.*, **22,** 203-217.

Barnett, T. P., and K. Hasselmann, 1979: Techniques of linear prediction with application to oceanic and atmospheric fields in the tropical Pacific. *Rev. Geophys. Space Phys*, 17, 949-968.

Burg, J. P., 1975: Maximum entropy spectral analysis. Ph. D. dissertation, Stanford University, 123 pp.

——, 1978: A new analysis technique for time series data. *Modern Spectrum Analysis*, D.G. Childers, Ed., IEEE Press, 42-48.

Compagnucci, R. H., D. Araneo, and P. O. Canziani, 2001: Principal sequence pattern analysis: A new approach to classifying the evolution of atmospheric systems. *Int. J. Climatol.*, **21**,197-217.

Hannan, E., and B. Quinn, 1979: The determination of the order of an autoregression. *J. Roy. Statistical Soc.*, B(**41**), 190-195.

Hurvich, C. M., and C. L. Tsai, 1989: Regression and time series model selection in small samples. *Biometrika*, **76**,297-307.

Iosifescu, M., 1980: *Finite Markov Processes and Their Applications*. John Wiley & Sons, Inc., 295 pp.

Kasyap, R.L.,1980: Inconsistency of the AIC rule for estimating the order of autoregressive models. *IEEE Trans. Automat. Control*, AC**-25**, 996-998.

Kim, K. Y., 1997: Statistical interpolation using cyclostationary EOFs. *J. Climate*, **10**, 2931- 2942.

——, and G. R. North, 1998: EOF-based linear prediction algorithm: Theory. *J. Climate*, **11**, 3046-3056

——, and Q. Wu, 1999: A comparison study of EOF techniques: Analysis of nonstationary data with periodic statistics. *J. Climate*, **12**, 185-199.

Lindzen, R., and J. R. Holton, 1968: A theory of the quasi-biennial oscillation. *J. Atmos. Sci*., **25**, 1095-1107.

Mares, C., and I. Mares, 1997: Predictability studies with the ARPEGE model and a statistical local model. Preprints, *WMO International Workshop on Dynamical Extended Range Forecasting*, WMO / TD-No. 881, Toulouse, France, World Meteor. Org., 52-57.

Mares, C. and I. Mares, 2003: Improvement of long-range forecasting by EEOF extrapolation using an AR-MEM model. *Wea. Forecasting*. **18**. 411-424.

Niwano, M., and M. Takahashi, 1998: The influence of the equatorial QBO on the Northern Hemisphere winter circulation of a GCM. *J. Meteor. Soc. Japan*, **76**, 453-460.

Parzen, E., 1977: Multiple time series: Determining the order of approximating autoregressive schemes. *Mutivariate Analysis-IV*, P. R. Krishnaiah, Ed., North-Holland Publishers, Amsterdam, 283-295.

Preisendorfer, R. W., 1988 : Principal component analysis in meteorology and oceanography. *Developments in Atmospheric Science*, **17**, C. D. Mobley, Ed., Elsevier Science Publishers B.V., Amsterdam, 426 pp.

Richman, M.B.,1986 : Rotation of principal components. *Int. J. Climatol*., **6**, 293-335.

Schwarz, G., 1978 : Estimating the dimension of a model. *Ann. Stat*. , **6**, 461-464.

Tabony, R.C., 1979: A spectral and filter analysis of long-period rainfall records in England and Wales. *Meteor. Mag*., **108**, 97-118.

Tangang, F., T. Benyang, A. Monahan, and W. Hsieh, 1998: Forecasting ENSO Events: A neural network-extended approach. *J. Climate*, **11**, 29-41.

Trenberth, K. E., and W-T. K. Shin, 1984: Quasi-biennial fluctuations in sea level pressures over the Northern Hemisphere. *Mon. Wea. Rev*.,**112**,761-777.

Ulrych, T. J., and T. Bishop, 1975: Maximum entropy spectral analysis and autoregressive decomposition. *Rev. Geophys. Space Phys*., 13**,** 183-200.

Vautard, R. and M. Ghil, 1989: Singular spectrum analysis in nonlinear dynamics with application to paleoclimatic time series. *Physica D*, **35**,395-424.

Vialar, J., 1968 : *Calcul des Probabilites et Statistique. Tome IV : Statistique Etude des Series Chronologiques*. Direction de la Meteorologie Nationale,128 pp.

von Storch, H., and C. Frankignoul, 1998: Empirical modal decomposition in coastal oceanography. *The Sea***,** **10**, K. H. Brink and A. R. Robinson, Eds., John Wiley & Sons,Inc., 419-455.

Wang, R., K. Fraedrich, and S. Pawson, 1995: Phase-space characteristics of the tropical stratospheric quasi-biennial oscillation. *J. Atmos. Sci*., **52**, 4482-4500.

Weare, B. C., and J.N. Nasstrom, 1982: Examples of extended empirical orthogonal function analyses. *Mon.Wea. Rev*.,**110**, 481-485.

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

Zhang, H., and T. Casey, 2000: Verification of categorical probability forecasts. *Wea. Forecasting,* **15**, 80-89.