A New Multidimensional Time Series Forecasting Method Based on the EOF Iteration Scheme

Zhang Banglin (张邦林), *Liu Jie* (刘洁)
Institute of Atmospheric Physics, Beijing 100080
and *Sun Zhaobo* (孙照渤)
Nanjing Institute of Meteorology, Nanjing 210044
Received November 14, 1991; revised August 24,1992

ABSTRACT

In this paper a new .mnultidimensional time series forecasting scheme based on the empirical orthogonal function (EOF) stepwise iteration process is introduced. The scheme is tested in a series of forecast experiments of Nino3 SST anomalies and Tahiti-Darwin SO index. The results show that the scheme is feasible and ENSO predictable.

I. INTRODUCETION

As a pattern of short-term climatic fluctuations with a characteristic period of 2-7 years, ENSO is closely associated with the global climatic anomaly. An overall relationship between ENSO and the climate has been extensively studied (Chen, 1981; Rasmusson and Carpenter, 1982; Ropelewski and Halpert, 1987), and many of the recognized effects are widely accepted as being statistically significant, physically plausible, and of sufficient magnitude to be of considerable economic concern in a number of regions. Therefore, the researches of the predictability and long-range forecasting technique have been considerable interest and are paid more attention by meteorologists. A lot of papers have been published on the predictability and predicition of ENSO. For example, Fraedrich (1988) estimated predictability of two SO time series with the conclusion that predictions are possible for at least one to two years. Inoue and O'Brien (1984) demonstrated the feasibility of a forecasting scheme for predicting the onset of a major ENSO, which is based on simple physics and statistics. Using a purely statiscal technique—the linear prediction theory, Barnett (1981) was also able to show that SST anomalies off Peru are predictable one year in advance for the period 1976-1979. Recently, Xu and Storch (1990) introduced a forecasting scheme based on principal oscillation pattern (POP) analysis. A series of hindcasting experiments show that the POP perdiction scheme is skillful for a lead time of two to three seasons. Thus, it seems that ENSO is predictable. Obviously, the prediction schemes are methodologically quite different, but they are all based on observational data and theoretical considerations of recognizing and understanding such large-scale behavior. Since many of the theoretical principles are known, scheme construction has been and continues to be a primary research field for meteorologists.

The purpose of this paper is to intro duce a new statistical forecasting scheme based on empirical orthogonal function (EOF) iteration process and illustrates its use by applying the scheme to some of the ENSO indices.

A description of the scheme based on EOF iteration process is given in Scetion 2. This is followed by an application of the scheme to the forecasts of ENSO indices in Section 3. The paper is concluded with discussion in Section 4.

II. TIME-SERIES FORECAST SCHEME BASED ON EOF ITERATION PROCESS

Given a discretely-sampled L-dimensional time series X_i of length $1 \leqslant i \leqslant N$, the scheme based on EOF iteration process for predicting X_{N+1} at time $N+1$ is as follows: at first we construct an initial augmented matrix of the form

$$\underset{(N-M+2)\times LM}{F^{(0)}} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+1} & X_{N-M+2} & \cdots & X_{N-1} & X_N \\ X_{N-M-2} & X_{N-M+3} & \cdots & X_N & X_{N+1}^{(0)} \end{bmatrix},$$

where M is dimension of embedding space, $X_{N+1}^{(0)}$ denotes the initial L-dimensional vector. EOF-expansion of the matrix $F^{(0)}$ yields its fitting field $\hat{F}^{(o)}$ in the form

$$\underset{(N-M+2)\times LM}{\hat{F}^{(0)}} = \sum_{K=1}^{K^{(0)}} T_K^{(0)} V_K^{(0)} = \begin{bmatrix} \hat{X}_1^{(0)} & \hat{X}_2^{(0)} & \cdots & \hat{X}_{M-1}^{(0)} & \hat{X}_M^{(0)} \\ \hat{X}_2^{(0)} & \hat{X}_3^{(0)} & \cdots & \hat{X}_M^{(0)} & \hat{X}_{M+1}^{(0)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{X}_{N-M+1}^{(0)} & \hat{X}_{N-M+2}^{(0)} & \cdots & \hat{X}_{N-1}^{(0)} & \hat{X}_N^{(0)} \\ \hat{X}_{N-M+2}^{(0)} & \hat{X}_{N-M+3}^{(0)} & \cdots & \hat{X}_N^{(0)} & \hat{X}_{N+1}^{(0)} \end{bmatrix},$$

in which $K^{(o)}$ represents the truncated order number of the iteration at the zero-th step, $T_k^{(o)}$ and $V_k^{(o)}$ are time expansion coefficients and empirical orthogonal functions of the matrix $F^{(o)}$, respectively. After replacing $X_{N+1}^{(o)}$ in $F^{(o)}$ with $\hat{X}_{N+1}^{(0)}$ in $\hat{F}^{(o)}$, we have the matrix of the first approximation

$$\underset{(N-M+2)\times LM}{F^{(1)}} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+1} & X_{N-M+2} & \cdots & X_{N-1} & X_N \\ X_{N-M+2} & X_{N-M+3} & \cdots & X_N & \hat{X}_{N+1}^{(0)} \end{bmatrix}.$$

The EOF-expansion of $F^{(1)}$ yields its fitting $\hat{F}^{(1)}$ of the form

$$\underset{(N-M+2)\times LM}{\hat{F}^{(1)}} = \sum_{K=1}^{K^{(1)}} T_K^{(1)} V_K^{(1)} = \begin{bmatrix} \hat{X}_1^{(1)} & \hat{X}_2^{(1)} & \cdots & \hat{X}_{M-1}^{(1)} & \hat{X}_M^{(1)} \\ \hat{X}_2^{(1)} & \hat{X}_3^{(1)} & \cdots & \hat{X}_M^{(1)} & \hat{X}_{M+1}^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{X}_{N-M+1}^{(1)} & \hat{X}_{N-M+2}^{(1)} & \cdots & \hat{X}_{N-1}^{(1)} & \hat{X}_N^{(1)} \\ \hat{X}_{N-M+2}^{(1)} & \hat{X}_{N-M+3}^{(1)} & \cdots & \hat{X}_N^{(1)} & \hat{X}_{N+1}^{(1)} \end{bmatrix}.$$

Likewise, we obtain the matrix $F^{(2)}$ of the second approxination also expanded by EOFs for the fitting field $\hat{F}^{(2)}$. The procedure is repeated until the *n–th* iteration is performed, resulting in

$$\underset{(N-M+2)\times LM}{F^{(n)}} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+1} & X_{N-M+2} & \cdots & X_{N-1} & X_N \\ X_{N-M+2} & X_{N-M+3} & \cdots & X_N & \hat{X}_{N+1}^{(n-1)} \end{bmatrix},$$

which is then expanded by EOFs, giving

$$\underset{(N-M+2)\times LM}{\hat{F}^{(n)}} = \sum_{K=1}^{K^{(n)}} T_K^{(n)} V_K^{(n)} = \begin{bmatrix} \hat{X}_1^{(n)} & \hat{X}_2^{(n)} & \cdots & \hat{X}_{M-1}^{(n)} & \hat{X}_M^{(n)} \\ \hat{X}_2^{(n)} & \hat{X}_3^{(n)} & \cdots & \hat{X}_M^{(n)} & \hat{X}_{M+1}^{(n)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{X}_{N-M+1}^{(n)} & \hat{X}_{N-M+2}^{(n)} & \cdots & \hat{X}_{N-1}^{(n)} & \hat{X}_N^{(n)} \\ \hat{X}_{N-M+2}^{(n)} & \hat{X}_{N-M+3}^{(n)} & \cdots & \hat{X}_N^{(n)} & \hat{X}_{N+1}^{(n)} \end{bmatrix}.$$

For $\| \hat{X}_{N+1}^{(n)} - \hat{X}_{N+1}^{(n-1)} \| \leqslant \varepsilon, \hat{X}_{N+1}^{(n)}$ is the forecast value at time *N+1*. A lot of numerical expercments show that the scheme developed in this paper is in nature of convergent.

After the forecast value $\hat{X}_{N+1}^{(n)}$ at time *N+1* is gotten, we construct an initial augmented matrix

$$\underset{(N-M+3)\times LM}{F^{(0)}} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+2} & X_{N-M+3} & \cdots & X_N & \hat{X}_{N+1}^{(n)} \\ X_{N-M+3} & X_{N-M+4} & \cdots & \hat{X}_{N+1}^{(n)} & X_{N+2}^{(0)} \end{bmatrix},$$

and repeat the EOF iteration process. Then the forecast value at time *N+2* can be obtained. Clearly, we may get the forecast values up to time *N+τ*. Especially, when $L = 1$, the above scheme may be used for forecasts of single variable time series.

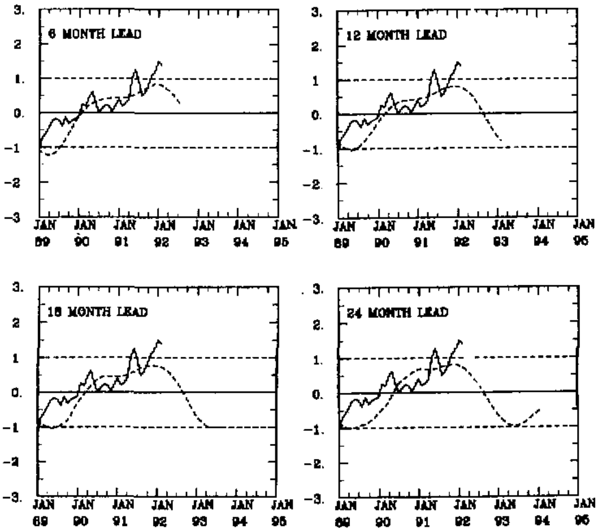

Fig.1. The forecasts for four different lead times versus observations of Nino3 SST index. Solid curve denotes observed. Dashed curve denotes predicted.

III. APPLICATIONS OF THE FORECAST SCHEME TO THE ENSO INDICES

Using the above forecasting scheme, we have done forecasts of Nino3 SST index and Tahiti–Darwin SO index based on the data set of atmospheric and SST indices provided by Climate Analysis Center of USA. Forecasts of Nino3 SST anomalies for a few different lead times are given in Fig.1. Form the figure we know that the ENSO could be predicted well up to about 24 months.

To assess the scheme's predictive capabilities, the outcome of forecast experiments can be compared with a traditional reference—persistence forecast by defining forecasting skill (Xu and Storch,1990)

$$S^p = \frac{(O,P_\tau)}{\sqrt{O^2 P_\tau^2}},$$

where $P_\tau(t)$ and $O(t)$ are predicted for lead time τ and the observed variables. If those variables are perfectly predicted, $S^p = 1$, whereas $S^p = 0$ for a useless prediction. The forecasting skills of Nino3 SST and Tahiti–Darwin SO indices at each lead time $\tau = 1,24$ by using

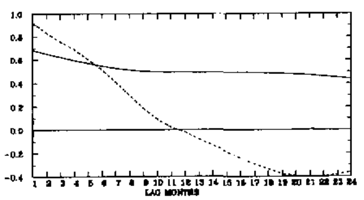

Fig.2. Correlation skill S^p of prediction of the Nino3 SST index. Solid curve: prediction based on EOF iteration process. Dashed curve: persistence.

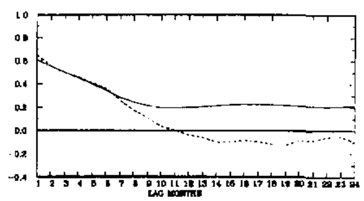

Fig.3. The same as in Fig.2, but for the Tahiti–Darwin SO index.

our forecasting scheme and persistence are respectively given in Figs. 2 and 3. It can be seen from these figures that persistence of Nino3 SST anomalies is better than our forecasts for $\tau \leqslant 5$ months. However after that, our scheme is superior. For Tahiti–Darwin SO index, skills of persistence are almost equal to our forecasts at $\tau \leqslant 6$ months. After that, our scheme is also superior. Thus the forecasting scheme based on EOF interation process is feasible.

IV. DISCUSSION

We have outlined a new time serics prediction scheme based on the EOF stepwise iteration process and demonstrated its usefulness by conducting the forecasting experiments of the ENSO. This method can be applied to the prediction of other climatic elements, e.g., temperature, precipitation etc. For operational purposes much more work, including choice of $M, K^{(i)}$ and definition of ε for the EOF iteration scheme, will have to be done. Greater skill scores could be reached if these aspects were considered in practice.

We would like to thank Prof. Zen Qingcun and Dr. S. S. Shen for their support to this study.

REFERENCES

Barnett,T.P. (1981), statistical relations between ocean / atmosphore fluctuations in the tropical Pacific, *J.Phys. Oceanogr.*, **38:** 1043–1058.

Chen, W.Y. (1981), Fluctuations in the Northern Hemisphere 700Mb height field associated with the Southern Oscillation, *Mon.Wea. Rev.*, **110:** 808–823.

Fraedrich, K. (1988), El Niño / Southern Oscillation predictability, *Mon. Wea. Rev.*, **116:** 1001–1012.

Hasselman, K. and T. P. Barrnett (1981), Techniques of linear prediction for systems with periodic statistics, *J. Atmos. Sci.*, **38:** 2275–2283.

Inoue, M., and J.J.O'Brien, (1984) A forecasting model for the onset of El– Niño, *Mon. Wea. Rev.*, **114:** 2326–2337.

Rasmusson, E. M., and T. H. Carpenter (1982), Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation / El Niño, *Mon. Wea. Rev.*, **110:** 358–384.

Ropelewski, C. F. and M. S. Halpert (1987), Global and regional scale preciptation patterns associated with the El Niño / Southern Oscillation, *Mon. Wea. Rev.*, **115:** 1606–1626.

Xu, J. and H. V. Storch (1990), Predicting the state of the Southern Oscillation using principal oscillation pattern analysis, *J Climate*, **3:** 1316–1329.