Adjusting OCN Prediction Method by Invoking EOF Modes

Peitao Peng and Huug van den Dool
CPC/NCEP/NWS/NOAA, Camp Springs, Maryland

1. Introduction

The optimal climate normal (OCN) method is one of the major tools of seasonal climate predictions at the climate prediction center (CPC). With this method, the climate prediction for a given season of the next year is essentially given by the average of the most recent K years. The K is determined according to the hindcast skill. Such an optimally determined number K is usually spatially and seasonally varying (Huang et al. 1996). In the current operational OCN seasonal temperature forecasts, a constant time period K=10 years is used for all locations and all seasons. The major purpose of using the constant time period is to make the forecast spatially consistent. However, as a cost of that purpose, the prediction skill becomes significantly lower than that using the spatially varying time period.

By analyzing the US climate variability, it is found that most variance (> 85%) of the seasonal mean US surface air temperature can be explained by a few empirical orthogonal function (EOF) modes. These EOF modes are different not only in their spatial pattern, but also in their dominant time scales. This finding has guided us to construct an EOF based OCN prediction scheme. With this scheme, the OCN prediction is conducted for each EOF component separately and independently, and the predicted EOF components are then synthesized to give a prediction for the total anomaly field. Since the dominant time scales of different EOF modes are different, so are the optimal Ks corresponding to these EOF modes. Therefore, this new OCN scheme can take account of multiple time scales of climate variability and meanwhile guarantee the forecast to be spatially consistent. The hindcasts for the last 42 years have shown that the skill of this new OCN scheme is significantly higher than the current operational one.

2. Data and method

The data used in this study are the seasonally averaged daily temperature during 1931-2002 period for 102 United States' climate divisions.

With the aid of the EOF analysis, the temperature anomaly of a given season in year *n* and at climate division i can be expressed as

$$T(i,n) = \sum_{m=1}^{M} \alpha_m(n) E_m(i), \qquad (1)$$

where, E_m and α_m are the m^{th} EOF pattern and its associated time coefficient, *M* is the number of EOF modes kept in this study. It turns out that EOF modes with *m*>6 are basically noise, so *M*=6 should be an appropriate choice.

Applying OCN prediction method to the m^{th} EOF component, we have the predicted time coefficient

$$\overline{\alpha}_{m,K}(n) = \frac{1}{K}\sum_{j=1}^{K} \alpha_m(n-j), \qquad (2)$$

where, the *K(m)* is chosen such that the temporal correlation between the predicted coefficients $\overline{\alpha}_{m,K}$ and the "observed" coefficients α_m in the history reaches its maximum. The predicted temperature anomaly thus is given by

$$T^f(i,n) = \sum_{m=1}^{M} \overline{\alpha}_{m,K(m)}(n) E_m(i) \tag{3}$$

The temporal correlation and spatially averaged temporal correlation are used as the measures of the prediction skill. Conventionally, we define temperature anomalies as the departures from the WMO recommended climate (which is the averages over a specified 30-year period updated at the start of each decade), i.e., $\hat{T}^f = T^f - C_{WMO}$ for prediction and $\hat{T}^{obs} = T^{obs} - C_{WMO}$ for observation. With these notations, the skill measures can be written respectively as

$$COR(i) = \frac{\sum_n \hat{T}^f(i,n)\hat{T}^{obs}(i,n)}{[\sum_n (\hat{T}^f(i,n))^2 \sum_n (\hat{T}^{obs}(i,n))^2]^{1/2}}, \tag{4}$$

and

$$COR = \frac{\sum_i \sum_n \hat{T}^f(i,n)\hat{T}^{obs}(i,n)}{[\sum_i \sum_n (\hat{T}^f(i,n))^2 \sum_i \sum_n (\hat{T}^{obs}(i,n))^2]^{1/2}}. \tag{5}$$

Because the first 30-year data need to be reserved for defining the climate, the OCN prediction test is only for the period 1961-2002. Likewise, the upper bound of the OCN parameter K is set to be 30 years.

3. Results

The EOF dependency of the OCN parameter K is shown in the Table1, where the K numbers for the six leading EOFs are presented for the four seasons. Take the winter season (DJF) as an example, the K numbers of the six EOFs are distinctly different. They are17 and 23 corresponding to EOF1 and EOF2 respectively, while the others are below 10. Similar patterns are also found for other seasons. The results indicate that the optimal "climate normals" associated with different EOF modes are very different. It is thus suggested that dealing with these EOF modes individually in OCN prediction is necessary and potentially beneficial.

The hindcast skill of the EOF based OCN prediction is shown in Figure 1, where the temporal correlation between the 9-month lead predictions and the observations over the past 42 years are presented for the four seasons. The same skill map but for the OCN prediction with K=10 everywhere is presented in Figure 2. Comparing the two figures, one can immediately see the big improvement by the EOF based OCN method. For all the seasons the colored area with correlation higher than 0.3 in Figure 1 is much more extensive than that in Figure 2. Not only this, the magnitude of the skill in Figure 1 is also significantly higher that in Figure 2. A more quantitative comparison is given in Table 2, where the spatially averaged temporal correlation over 102 US climate divisions are presented. Obviously, the skill of the EOF based OCN method is higher than that with

K=10 for all locations by 0.1 or more for most seasons. The annual average of the former is **0.1** higher than the latter.

Though the skill scores of the two OCN techniques are quite distinct, their geographical distribution and seasonal change are very similar. For example, in both the figures the high skill region is in the southeastern states for the winter and the summer, in the western states for the spring and in the southern states for the fall. In both cases the skill score reaches the highest values in the winter and drops to the lowest in the fall (SON). It is found that the spatial variation of the skill score coincides with the variation of the fraction of the low frequency climate variability. This is understandable since the predictions by the OCN methods are indeed in the low frequency end.

4. Summary

In this study, it is found that the ability of the OCN method in predicting the United States surface temperature can be significantly improved through invoking EOF modes. The increase in the annual skill score averaged over the 102 climate divisions is about **0.1** for the 9-month lead predictions**.** This is attributed to the use of multiple time scales of climate variability embedded in different EOF modes.

Reference

Huang, Jin, H. M. van den Dool and A. Barnston, 1996: Long-Lead Seasonal temperature prediction using Optimal Climate Normals. *J. Climate*, **9**, 809-817.

Table 1. The OCN parameter K of the six leading EOFs of the US surface temperatures. The numbers are determined based on the hindcast for the period of 1961-2002.

	EOF 1	EOF 2	EOF 3	EOF 4	EOF 5	EOF 6
DJF	17	23	4	5	8	2
MAM	15	12	7	8	5	5
JJA	18	15	8	9	7	8
SON	19	15	8	9	8	9

Table 2. Spatially averaged temporal correlation between the 1-year lead OCN predictions and the observations for the 102 US climate divisions and the period of 1961-2002.

	DJF	MAM	JJA	SON	ANNUAL MEAN
EOF Based	.36	.33	.26	.15	.28
K=10	.30	.23	.15	.02	.18

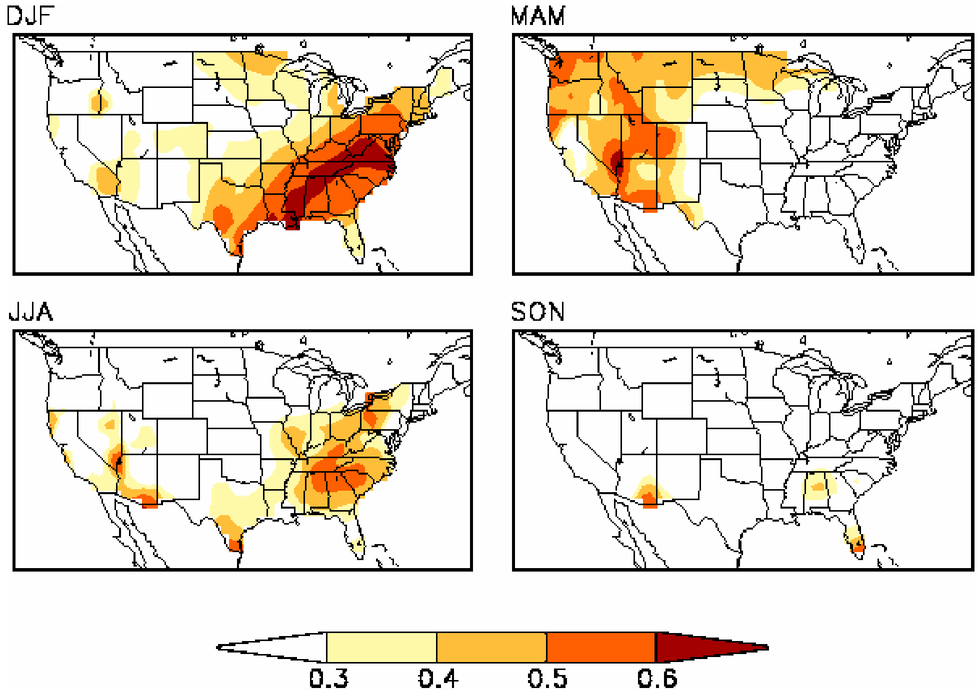

Figure 1. The temporal correlation between the predicted surface temperatures with the EOF based OCN and the observations for the period of 1961-2002.

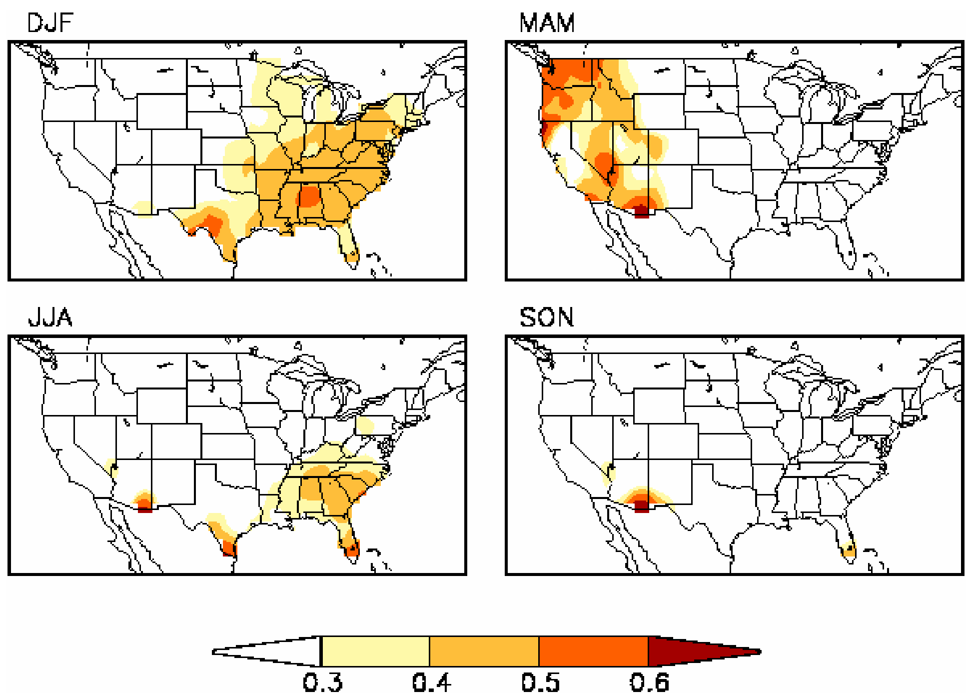

Figure 2. The temporal correlation between the predicted surface temperatures with the OCN with K=10 over all climate divisions and the observations for the period of 1961-2002.