Atmósfera (2000), 13, pp. 223-244

Seasonal-range forecasting of the ozark climate by a principal component regression scheme with antecedent sea surface temperatures and upper air conditions

JAE -WON LEE* and ERNEST C. KUNG

School of Natural Resources University of Missouri-Columbia, Columbia, MO 65211

(Manuscript received June 9, 1999; accepted in final form January 13, 2000)

RESUMEN

Con base en un análisis de componentes principales de registros climatológicos históricos, se formulan y se efectúan modelos de regresión pronósticos experimentales de temperatura y precipitación para el área del altiplano llamado Ozark, un macizo montañoso del medio oeste sur central de los E. E. U. U. Los predictores incluyen temperaturas globales de las temperaturas superficiales del mar, así como campos hemisféricos del aire superior y las observaciones del clima local. Los experimentos para todos los meses del año se llevan a cabo a partir de datos de los segmentos de 15 años de 1961-75 y 1980-94 para los años más allá de los segmentos de datos respectivos.

Mediante análisis de correlación cruzada, se investigan las relaciones entre variables climáticas de escala regional y de gran escala a fin de identificar teleconexiones útiles para el pronóstico a largo plazo. La predictabilidad del clima de las montañas Ozark se examina con los esquemas de regresión lineal múltiple y el de componentes principales. Se muestra que el comportamiento del segundo esquema es superior al del primero. Los resultados de los experimentos de pronóstico extensivo revelan la utilidad y estabilidad del pronóstico de los elementos climáticos del terreno elevado de dichas montañas. La validez de los modelos de pronóstico se verifica hasta para 10 años después del periodo de datos usado en la formulación de la regresión.

ABSTRACT

On the basis of principal component analysis of long-term climatological records, regression models are formulated and forecast experiments are conducted for monthly temperature and precipitation of the Ozark Highlands area, a large area of low mountains and plateau in the south central midwestern United States. Predictors include global sea surface temperatures, hemispheric upper air fields and the local climate observations. The experiments for all months of the year are performed with the data from continuous 15-year segments of 1961-75 to 1980-94 for those years beyond the respective data segments.

Relationships between regional-scale and large-scale climate variables are investigated by cross-correlation analysis to identify useful teleconnections for seasonal-range forecasting. The predictability of the Ozark Highlands climate is examined with the multiple linear regression scheme and the principal component regression scheme. It is shown that the forecast performance by the latter is superior to that of the former. The results of the extensive forecast experiments reveal the useful and stable predictability of the Ozark Highlands climate elements. The validity of the forecasting models is verified for up to 10 years after the data period of regression formulation.

* Current affiliation: Korean Meteorological Administration, Seoul, South Korea

1. Introduction

The regression approach in seasonal-range forecasting is based on the established teleconnection patterns of the predictands with the forcing function that is the sea surface temperatures (SSTs) and their immediate manifestations in the general circulation. The feasibility of such forecasting schemes was reported by Kung and Tanaka (1985), Kung *et al.* (1995) and Unger (1996 a, b). A systematic study on the large-scale mode of SST variations and Northern Hemisphere tropospheric responses was documented by Kung and Chern (1994, 1995) as a basis for further improvement in the regression approach.

Despite the apparent effectiveness of multiple linear regression forecasting scheme at the seasonal range, especially for the year immediately following the data period used for model formulation, the forecast skill deteriorates noticeably afterward. As the prevailing patterns of the general circulation undergo continuous change with various scales of quasi-cyclic variations, the regression formulae require a constant updating of regression coefficients with the immediate past data. Given the data-processing power of modern computers, it is not difficult to annually update the regression model. However, there is a scientific, as well as technical, interest to delve into the nature of the commonly observed deterioration of regressions, and further to attempt a construction of the stable regression models. The purpose of this paper is to formulate and examine such a scheme that will be useful regionally for the south central midwestern area of the United States.

Park and Kung (1988), in their principal component analysis, discovered that the midwestern US summer temperature is essentially determined by its first principal component. Ting and Wang (1997) showed the correlation features between a US precipitation index and sea surface temperatures using the first two principal components. If a regional climatological variable is dominated by certain dominant components, we expect that these components may be effectively used to relate to the predictand in providing a stable forecast performance with a high degree of skill. Through the use of principal components we may reduce a large number of variables into a small manageable set of variables while retaining the essential information of the original variables, yet rejecting the nonessential information that causes the fluctuation in forecast performance. In this paper we examine the principal components of the temperature and precipitation of the Ozark Highlands area. They are used to obtain the expected values of predictands in a regression scheme that relate the principal components of regional climate variables with antecedent SSTs and upper air parameters. The forecast experiments are conducted with 15-year data segments from 1961-75 to 1980-1994 with a total of 20 data periods. The results of the principal component regression forecasting are compared with those of multiple linear regression forecasting, and are also examined in the light of teleconnections that have led to the model formulation.

2. Data

The Ozark Highlands (OZ) is located in the south central midwestern area of the United States, composed of low mountains and plateau in southern Missouri and northern Arkansas. The OZ area named in this study is bounded by 33-41°N and 89-96°W, including a large area of the states of Missouri and Arkansas. The monthly mean temperature (TMP) and monthly total of precipitation (PPT) in OZ are climate elements defined over this region. The daily U.S. cooperative station network data from 1940 to 1995 were obtained from the National Climatic Data Center (NCDC 1993). Datasets over a uniform 0.5° latitude-longitude grid mesh containing 255 (17×15) points without missing data were constructed by interpolating grid values from surrounding observations with the distance as a weighting factor for interpolation. More than 600 cooperative stations in the OZ area were involved in computing the monthly TMP and PPT.

Monthly sea surface temperatures (SSTs) of the Comprehensive Ocean-Atmosphere Data Set (COADS) (Woodruff *et al.*, 1987) and the reconstructed SSTs from the National Oceanic and Atmospheric Administration (Smith *et al.*, 1996) were utilized in this study for the period of 1960 to 1996. The SST data were reanalyzed on a 2° latitude-longitude grid mesh without missing data on all grids in the oceanic regions between 40°S and 60°N. At the University of Missouri-Columbia, daily 1200 GMT Northern Hemisphere octagonal grids data of upper air from 1960 to 1996 were transformed to 2° latitude-longitude grids. This study utilizes the 700 and 500 hPa temperatures (T7 and T5) and geopotential heights (Z7 and Z5) data from 18°N through 88°N latitude. The Z7 and Z5 are also utilized to obtain u-components (U7 and U5) and v-components (V7 and V5) of geostrophic wind from latitude 18°N to 70°N. Monthly values of observations constitute the basic datasets in this study, including the TMP, PPT, SSTs and hemispheric upper air parameters of T7, T5, Z7, Z5, U7, U5, V7, and V5. At the commencement of this study, the climate reanalysis data by the National Center for Environmental Prediction were not yet available for the entire data period of this study. However, the comparison of the available reanalysis data and our grid upper-air data indicates the general compatibility of these two data sets.

3. Methods of analysis and forecast

a. General scheme

The structure of the analysis and model construction is illustrated in Figure 1. Principal component analysis (PCA) was performed on all calendar months of monthly temperature (TMP) and precipitation (PPT) in the United States. The general procedures and discussions on PCA are available in Kutzback (1967), Walse and Richman (1981), Park and Kung

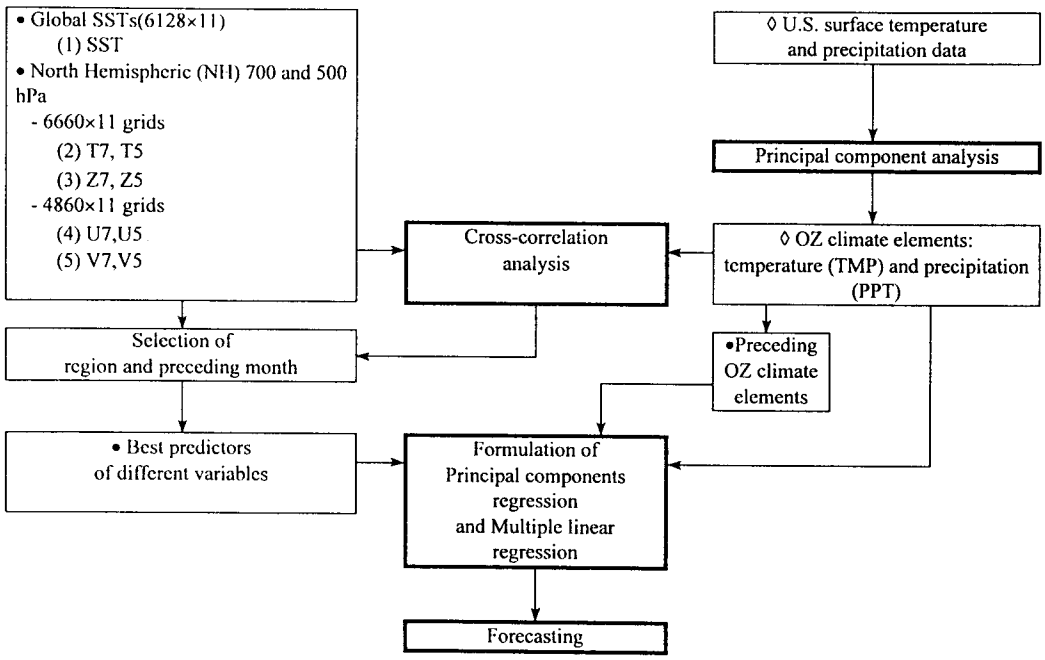


Fig. 1. Schematic illustration of procedures in regression model construction. Heavy-lined boxes indicate computational steps. The • and ◇ signs are for predictors and predictands.

(1988), Preisendorfer and Mobley (1984), Kung and Chern (1994), Montroy (1997) and Ting and Wang (1997). The data sets were analyzed with PCA being derived from the correlation matrix. January variations of the first components of TMP and PPT over the United States in Figure 2 exemplify a homogenous spatial pattern of the first component in the OZ region. The pattern independently obtained for OZ with the subset of data yields an identical spatial pattern of the first component, although the second and higher components show very different patterns between the US set and OZ subset of data. The associated time series are also consistent between the continental United States and OZ. Though not specifically illustrated, the other months of the year also indicate a similar nature of the first components. The TMP and PPT of OZ and their principal components were employed to form the predictands of regression models.

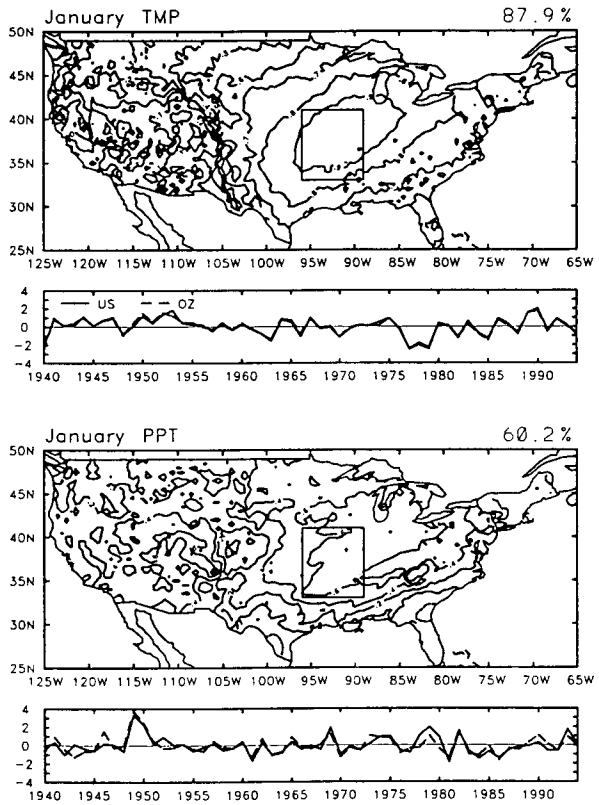
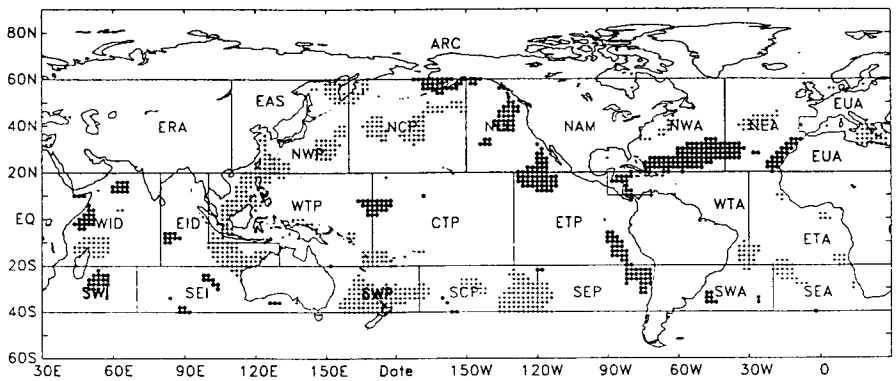


Fig. 2. Pattern vectors and time coefficients of the first principal components of January temperature and precipitation during the period of 1940-94. Solid and dashed lines respectively indicate positive and negative values. Percentages indicated are for the first components in the total variance.

Cross-correlation analysis was performed between monthly OZ climate elements and global/hemispherical climate parameters of preceding months from one to eleven months. The latter include SSTs and 700 and 500 hPa upper air parameters. The analysis of all calendar-year months was performed with the data of 15-year segments from 1961-75 to 1980-94. Grid points of SSTs and upper air parameters that show high correlation ($| r | \geq 0.8$), where r is the correlation coefficient, were identified; and the extent of regions of high correlation was examined. For example, Figure 3 illustrates correlation patterns of January TMP and preceding December SSTs.



ARC	Arctics	WTP	Western tropical Pacific
ERA	Eurasia	CTP	Central tropical Pacific
EAS	East Asia	ETP	Eastern tropical Pacific
NWP	Northwestern Pacific	WTA	Western tropical Atlantic
NCP	North central Pacific	ETA	Eastern tropical Atlantic
NEP	Northeastern Pacific	SWI	Southwestern Indian Ocean
NAM	North America	SEI	Southeastern Indian Ocean
NWA	Northwestern Atlantic	SWP	Southwestern Pacific
NEA	Northeastern Atlantic	SCP	South central Pacific
EUA	West Europe and North Africa	SEP	Southeastern Pacific
WID	Western Indian Ocean	SWA	Southwestern Atlantic
EID	Eastern Indian Ocean	SEA	Southeastern Atlantic

Fig. 3. Regions of the global/hemispherical data coverage. The plus-sign and dot-sign indicate positive and negative correlation points where preceding December SST records are correlated with OZ January temperature during the 15-year data segment 1971-85 at the 95% significance level.

In order to minimize the multicollinearity among predictors (see Kung and Tanaka, 1985; Myers, 1990), the predictors with high correlation with other predictors in temporal and spatial domains were rejected. The obtained five best correlation grids and OZ climate elements of the preceding year were employed as possible predictors in the multiple linear regression (MLR) and principal components regression (PCR) model.

b. Multiple linear regression (MLR)

A MLR scheme for long-range forecasting of middle latitude climate variables was developed by Kung and Tanaka (1985). They selected the best possible combination of predictors to give the largest value of coefficient of determination and smallest mean square error. They demonstrated how multicollinearity could be reduced among highly correlated large-scale climate data through a stepwise formulation of regressions. The MLR scheme in this study follows this procedure.

In the general multiple linear regression, the predictand y is expressed as a function of n predictors:

$$y = a_o + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \tag{1}$$

where a is the regression coefficient and x the predictor. The first predictor x_1 is selected by identifying the variable that yields the highest correlation with the predictand. After the selection of x_1, a simple regression is formed by the least squares fitting,

$$y = a_1 x_1 + \varepsilon_1 \tag{2}$$

where ε_1 is the first residual component and the correlation $r(x_1, \varepsilon_1) = 0$. The second predictor

x_2 is selected by examining the correlation between ε_1 and remaining possible predictors. The procedure is repeated until the x_n is selected when the residual $\varepsilon_n \approx o$. After the selection of predictors x_1 through x_n, the regression coefficients a_n are recomputed with the least-squares fitting.

MLR models of this study use three predictors ($n = 3$) as determined to be sufficient for the OZ climate variables. In this stepwise regression analysis, the collinearity among predictors is effectively eliminated through regression on the residuals.

c. Principal component regression (PCR)

In this scheme the first principal components of OZ climate elements are regressed on possible predictors. This PCR scheme follows the procedures as developed by Draper and Smith (1981), Preisendorfer and Mobley (1988), Basilevsky (1994) and Lee (1997). The first step of the procedure is to obtain a standardization of variables to estimate the correlation matrix and find the eigenvalues for the correlation matrix. The next step is to estimate the principal component scores (C, time coefficient in this study) from the eigenvector and the standardized matrix. Then the multiple regression coefficients (B) are obtained from principal component scores by,

$$\mathbf{B}_{(M\times 1)} = [\mathbf{C}^T_{(M\times N)}\mathbf{C}_{(N\times M)}]^{-1}\mathbf{C}^T_{(M\times N)}\mathbf{Y}_{(N\times 1)}, \tag{3}$$

where $\mathbf{C}^T_{(M\times N)}$ is the transpose of $\mathbf{C}_{(N\times M)}$, $\mathbf{Y}_{(N\times 1)}$ the matrix form of the dependent variable, M the reduced number of variables from the eigenvalue problem, and N the number of the recording period. In this study ten predictors are used with $M = 10$. Finally the estimate of the predictand $\hat{\mathbf{Y}}_{(N\times 1)}$ is obtained by

$$\hat{\mathbf{Y}}_{(N\times 1)} = \mathbf{C}_{(N\times M)}\mathbf{B}_{(M\times 1)}. \tag{4}$$

The PCR avoids the computational instability that occurs during multiplying or inverting the matrix of the predictor variables. It also reduces the number of calculations necessary to perform all possible linear regression of $2^k - 1$, where k is the number of predictors, especially when k is large (Basilevsky, 1994). The PCR model may be viewed as a special case of the canonical correlation model, which can be projected as a generalization of PCA since what it investigates accounts for the multidimensional correlation structure between two sets of multivariate variables observed for the same sample. In this study, OZ climate elements are composed of three linearly independent pieces of principal component scores inferred from a large pool of preceding global SSTs, hemispheric upper air variables, and records of the predictand in the preceding year.

d. Evaluation of forecast performance

After models are formulated and forecast experiments are performed, validation steps follow. The root mean square error (RMSE), absolute error (AE) and correlation (R) between forecasted recorded values, and skill score (SS) are utilized to evaluate forecast performance (Nicholls 1984; Davis, 1976).

RMSE is defined by

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (X_{i,\ forecast} - X_{i,\ record})^2}{N}}, \tag{5}$$

where $X_{i,\ forecast}$ and $X_{i,\ record}$ are respectively forecasted and recorded value and N is the number of observations. AE is a summation of the size of the error between the forecasted and the actual recorded values.

$$AE = \sum_{i=1}^{N} \frac{| X_{i\ forecast} - X_{i,\ record} |}{N}. \tag{6}$$

R is a test of the linear relationship between forecasted and recorded value.

$$R = \frac{\sum_{i=1}^{N} X_{i,\ forecast} X_{i,\ record} - \sum_{i=1}^{N} X_{i,\ forecast} \sum_{i=1}^{N} X_{i,\ record}/N}{\sqrt{\{\sum_{i=1}^{N} X_{i,\ forecast}{}^2 - (\sum_{i=1}^{N} X_{i,\ forecast})^2/N\}\{\sum_{i=1}^{N} X_{i,\ record}{}^2 - (\sum_{i=1}^{N} X_{i,\ record})^2/N\}}} \tag{7}$$

SS is an evaluation of the efficiency of a forecasting model, as defined by

$$SS = (1 - \frac{\sum_{i=1}^{N} (X_{i,\ forecast} - X_{i,\ record})^2}{\sum_{i=1}^{N} X_{i,\ record}{}^2}) \times 100. \tag{8}$$

A large value of SS indicates that the model provides reliable forecast values.

4. Teleconnections of OZ climate elements with global SSTs and hemispheric upper air conditions

During the past two decades major El Niño and La Niña events have dominated the global climate pattern, and a substantial effort by the scientific community has been devoted to investigate El Niño/Southern Oscillation (ENSO). Although this study treats SSTs without specifically focusing on the problem of ENSO, the mechanisms involved in ENSO naturally are involved in the manifestation of the teleconnection phenomena that are utilized in regression forecasting. The ways the local SST anomalies in the equatorial Pacific are extended vertically to the tropopause and horizontally to extra tropical latitudes have not been definitely understood. As reviewed by Tribbia (1991), they may involve the east-west Walker circulation, north-south Hadley circulation, wavetrains to disperse stationary Rossly waves, anomalous low level convergence/divergence led by anomalously cold/warm low level convergence, anomalous upper level divergent/convergent flow by anomalous mid level heat release, vorticity transport due to anomalous upper level outflow, etc. As pointed out by Trenberth (1991), however, the tropical Pacific coupled system is fundamentally unstable and multiple possible triggers appear to exist. Thus, the exact trigger of teleconnection may not be pinpointed readily.

The assortment of possible triggers of teleconnections requires that the strength and regions of teleconnections must be examined to a global extent. Studies by Blackmon *et al.* (1983), Geiser *et al.* (1985) and Tribbia (1991) collectively indicated that the Ozark Highlands area of this study only shows a rather weak response to the ENSO in comparison to other areas, which is consistent with the observed fact that the area climate shows a near-normal winter in El Niño or La Niña years. For these reasons, this study adopts a general cross-correlation analysis of the continuous time series as the basis of multiple regression schemes without a specific assumption of particular physical mechanisms. The spatial and temporal relationship between OZ climate elements and the global SSTs and hemispherical upper air conditions were examined to establish teleconnections useful for regression forecasting. The data were utilized in each of 15-year segments which were piecewise continuous (1961-1975, 1962-1976, ..., 1980-1994). Cross-correlations of OZ climate elements were calculated with the total 6128 grid points of global SSTs, 6660 grid points of upper air T and Z, and 4860 grid points of upper air U and V for each 15-year period.

Table 1 exemplifies cross-correlation between SSTs and OZ temperature. Very high correlations ($| r | \geq 0.8$) between OZ January temperature and preceding SSTs are recognized with a significance level greater than 99.9% for the period 1971-85 to 1980-94. Data regions can be identified where preceding SSTs show consistently high correlations (Fig. 3 for data regions). The southwestern Pacific region (SWP) in preceding July is listed eight times. In preceding October and June it is listed seven times and in the preceding September and May it is listed six times. Other preceding SSTs of high occurrence in Table 1 are the Eastern Tropical Atlantic (ETA) in the preceding June and April, the North Central Pacific (NCP) in preceding December and May, the South Central Pacific (SCP) in preceding August and March, SWP in preceding August, and the Southeastern Indian Ocean (SEI) in the preceding September. It is noted that the interannual variations of OZ January temperature are most closely associated with preceding SSTs in the Southern Hemisphere.

Table 1. Significant cross-correlation ($|r| \geq 0.8$) between OZ January temperature and preceding SSTs during each 15-year period. The table lists the preceding months, areas of chosen grid points, and signs of correlation coefficients.

1971-85	1972-86	1973-87	1974-88	1975-89	1976-90	1977-91	1978-92	1979-93	1980-94
Oct, SWP +	Oct, SWP +	Oct, SWP +	Dec, SCP +	Dec, NWP +	Dec, NCP +	Dec, NCP +	Dec, NCP +	Oct, SWI +	Dec, NWA +
Sep, SWP +	Sep, SWP +	Sep, SWP +	Oct, SWP +	Oct, SWP +	Dec, ETI +	Nov, SWP +	Aug, SWP +	May, NCP +	Oct, NEP+
Sep, NEA +	Sep, NEA +	Jul, SWP +	Oct, NWP +	Oct, SCP +	Dec, NEP +	Oct, SWP +	Aug, ERA +	Mar, SCP +	Aug, NWA+
Aug, SWP +	Aug, SWP +	Jun, SWP +	Sep, SWP +	Sep, SWP +	Dec, WTI +	Sep, SEI +	Aug, NEA +		Jun, NCP -
Jul, SWP +	Jul, SWP +	May, SWP +	Aug, SWP +	Sep, SEI +	Dec, CTP +	Aug, SCP +	Jul, SWP +		May, NCP +
Jun, SWP +	Jun, SWP +	Mar, NCP +	Jul, SWP +	Aug, SWP +	Oct, SWP +	Jul, SWP +	May, NCP +		Mar, SCP +
May, SWP +	May, SWP+	Feb, SWP +	Jun, SWP +	Aug, SCP +	Oct, WTP +	Jun, SWP +	Mar, SCP +		
			Jun, ETA +	Jul, SWP +	Sep, SWP +	Feb, SWP +	Feb, SWP +		
			May, SWP +	Jun, SWP +	Sep, SEI +				
			Apr, ETA +	Jun, ETA +	Aug, SWP +				
			Apr, SWP +	May, SWP +	Aug, SCP +				
			Feb, SWP +	Apr, ETA +	Jul, SWP +				
				Feb, ETP +	Jul, WTA +				
				Feb, NCP +	Jun, WTA +				
				Feb, NWP +	Jun, SWP +				
					Jun, ETA +				
					May, SWP +				
					Apr, ETA +				
					Feb, ETP +				

High cross-correlations between OZ January temperature and preceding upper air temperatures (T7 and T5), geopotential heights (Z7 and Z5) and geostrophic winds (U7, U5, V7 and V5) are shown in Table 2 for preceding (predictor) months, area of chosen grids and signs of correlation coefficients. The areas whose T7 and/or T5 repeatedly show high cross-correlations with OZ January TMP include Northeastern Atlantic (MEA), Eurasia (ERA), Europe and North Africa (EUA), NCP, Arctic (ARC), Northwestern Atlantic (NWA). Likewise, the areas whose

Table 2. Same as Table 1, but for cross-correlation between OZ January temperature and upper air data (temperature, geopotential height, and geostrophic wind) of preceding months.

1971-85	1972-86	1973-87	1974-88	1975-89	1976-90	1977-91	1978-92	1979-93	1980-94
T7: Sep, ERA - Aug, NEA + Aug, NWA+ Aug, NWP+ Jul, NCP + May, ERA + May, NCP + Apr, NWP + Feb, NEP + Feb, EAS +	T7: Sep, NEP + Sep, ERA - Aug, NWA + Aug, NEA + Aug, NWP + Aug, EUA + Jul, NCP + May, NCP + Feb, NEP + Feb, EAS +	T7: Sep, NEP + Sep, ERA - Aug, NEA+ Aug, NWP+ Aug, EUA+ May, NCP+ May, ERA - Feb, NWP + Feb, EAS +	T7: Dec, ARC + Sep, NEP + Sep, ERA - Sep, NWA + Aug, NEA + Aug, EUA + Aug, NWA + May, NCP + May, ERA - Feb, NWP + Feb, NEP +	T7: Dec, ARC - Sep, NAM + Sep, ERA - Sep, NWA + Sep, NEA + Aug, NEA + Aug, EUA + Jul, NWP + Jul, NWA + Jul, ERA - Jul, NCP + Jun, NWP + May, NWP + May, ERA - Apr, NWP + Apr, ERA + Mar, NEP + Feb, NWP + Feb, NAM +	T7: Dec, ARC - Sep, EUA + Sep, ERA - Sep, NWA + Sep, NEA + Sep, NAM + Aug, NEA + Aug, EUA + Aug, NWA + Jul, NCP + Jun, NWP + Jun, NWA + Jun, NAM + May, NWP + May, ERA - Apr, NWP + Apr, ERA + Apr, NAM + Mar, NEP + Feb, NWP + Feb, NAM + Feb, NEP +	T7: Dec, NAM + Dec, ARC + Sep, EUA + Sep, NWP + Sep, NEP + Sep, NCP + Sep, EUA + Aug, ARC + Aug, NEA + Jul, NCP + Jun, NWP + May, ERA - May, NWP + May, NCP + Apr, NWP + Mar, NEP + Feb, NAM + Feb, ARC +	T7: Dec, NAM+ Dec, ARC + Sep, NWP + Sep, NEP + Aug, ARC+ Aug, NEA+ Jul, NCP + May, ERA - Feb, NEP + Feb, ARC +	T7: Sep, NWP + Sep, ERA + Sep, ARC + Aug, ARC+ Apr, ERA + Mar, NWP+	T7: Aug, ARC + Aug, ERA + Apr, ERA +
T5: Dec, ARC - Aug, ARC - Jul, ARC - Jun, ERA - May, ARC -	T5: Dec, ARC - Aug, ARC - Aug, EUA + Jul, ARC -	T5: Dec, ARC - Aug, EUA+ Jul, ARC - Jul, ERA -	T5: Aug, EUA + Aug, NEA + Jul, ERA -	T5: Aug, EUA + Aug, NEA + Jul, NWA + Jul, ERA -	T5: Aug, EUA + Aug, NEA + Jul, NWA + Jul, ERA - Jul, NEA +	T5: Dec, NAM + Aug, EUA + Aug, NEA + Jul, NWA + Jul, NEA +	T5: Dec, NAM+ Aug, ERA + Jul, NWA +	T5: Dec, NAM+ Sep, ERA + Aug, ERA +	T5: Sep, ERA + Aug, ERA + Feb, NAM -
Z7: Dec, ARC - Aug, ARC - May, ARC -	Z7: Dec, ARC - Sep, NWA + May, ARC -	Z7: Dec, ARC - Sep, NWA+ May, ARC -	Z7: Dec, NEA + Sep, NWA + Aug, NWP + Aug, ARC - May, ARC +	Z7: Sep, NWA + Aug, NWP + Aug, EAS + Aug, EUA - May, ARC + Mar, NEP +	Z7: Aug, NWP + Aug, ARC - Jul, NEP + Jul, NAM +	Z7: Dec, NAM + Oct, ERA - Jul, NWA + Feb, NEP + Feb, NWA +	Z7: Dec, NAM+ Oct, ERA - Jul,, NEP + Feb, NEP + Feb, NWA+ Feb, NAM+	Z7: Aug, NWA+ Mar, NCP + Mar, NWP+ Feb, NWA+ Feb, NAM+	Z7: Dec, ARC + Dec, ERA - Aug,NWA+ Mar, NCP + Mar, NWP+
Z5: Dec, ARC - Sep, NWA+ Aug, ARC -	Z5: Dec, ARC - Sep, NWA + Jul, ARC - May, ARC -	Z5: Dec, ARC - Sep, NWA+ Aug, NEA+ Jul, ARC - May, ARC -	Z5: Sep, NWA + Aug, NEA + Aug, ERA -	Z5: Sep, NWA + Aug, NEA + Aug, NWP + Aug, EUA -	Z5: Dec, ARC - Sep, NWA + Aug, NEA + Jul, NWA + May, NAM + May, NEP +	Z5: Dec, NAM + Sep, NWA + Aug, NEA + Aug, NCP + Jul, NWA + Feb, NAM +	Z5: Dec, NAM+ Sep, NWA+ Aug, NEA+ Jul, NWA +	Z5: Dec, ARC + Aug, NEA+ Mar, NCP + Mar, NWP+ Feb, NAM -	Z5: Dec, ARC + Apr, ERA + Mar, NCP + Mar, NWP+ Feb, NAM -
U7: Dec, NEA + Dec, EUA + Jun, ERA +	U7: Dec, NEA + Dec, EUA + Sep, NWA -	U7: Dec, NEA + Dec, EUA + Sep, NWA -	U7: Dec, NEA + Dec, EUA + Aug, NEA + Jun, ERA +	U7: Dec, NEA + Dec, EUA + Sep, NAM - Aug, ERA + Jun, ERA +	U7: Dec, NEA + Dec, EUA + Aug, NCP - Aug, NEP - Aug, EUA + Jul, NWP - Jun, ERA +	U7: Dec, NEA + Dec, EUA + Oct, ERA + Oct, NEA + Sep, EAS - Aug, EUA + Jul, EAS - Jul, NWP - Jul, ERA +	U7: Dec, NEA + Oct, ERA + Oct, NEP + Oct, NAM+ Oct, NEA +	U7: Nov,NAM+ Oct, EUA + Sep, ARC + Aug, NEA - Apr, ERA -	U7: Nov,NAM+ Oct, NEP + Sep, ARC + Apr, ERA - Mar, ARC + Feb, NAM+
U5: Dec, NEA + Dec EUA + Dec, NEA + May, ERA +	U5: Dec, NEA + Dec EUA + Sep, NWA - Aug, EUA + May, ERA +	U5: Dec, NEA + Dec EUA + Sep, NWA - Aug, EUA+ Aug, NEA - May, ERA+	U5: Dec, NEA + Dec EUA + Sep, NWA - Aug, NEA -	U5: Dec, NEA + Dec EUA + Sep, NWA - Aug, EUA - Aug, NEA - Aug, ERA + Aug, NWP +	U5: Dec, NEA + Aug, ERA + Aug, NWP + Aug, NEA-	U5: Dec, NEA + Aug, EUA - Aug, NEA- Jul, EAS - Jul, NWP - Feb, NAM +	U5: Aug, NEA - Jul, EAS - Jul, NWP -	U5: Sep, ARC + Aug, NEA - Mar, NWP - Feb, NAM+	U5: Sep, ARC + Aug, NEA - Feb, NAM+
V7: Aug, ERA + Aug, ARC - Jul, NWP - Apr, NCP -	V7: Sep, NWP - Aug, ERA + Aug, ARC - Jul, NCP - Feb, NCP +	V7: Nov, EUA - Aug, ERA + Aug, ARC - Jun, EUA - May, NEP + Apr, NCP - Feb, NCP +	V7: Dec, ARC - Aug, ERA +	V7: Dec, NAM - Dec, NWA - Sep, ARC - Aug, ERA + May, NEP + May, ERA - Apr, ERA- Mar, EAS +	V7: Dec, NAM - Dec, NWA - Dec, ERA + Oct, ERA + Sep, ERA + Sep, ARC - Sep NCP +	V7: Dec, NAM - Dec, NWA - Dec, ARC - Oct, ERA + Sep, ARC - Sep, NCP + Sep, ERA + Aug, ERA + Jul, NWP - Jul, EAS - Jun, NAM - Mar, ARC + Feb, ARC -	V7: Dec, NAM - Dec, NWA - Dec, ARC - Sep, NAM+ Jul, NWP - May, ARC+ Feb, ARC - Feb, NAM -	V7: Dec, NAM - Dec, ERA + Dec, NWA - Dec, ARC + Aug, ERA + Jul, NWP - May, ARC+ Mar, NWP+	V7: Dec, ARC + Aug, ERA + Jul, EUA + Jul, NEA + Jul, ARC + May, ARC+ Mar, NWP+
V5: Dec, NWA - Dec, NAM - Sep, NCP + May, ERA + Apr, NCP -	V5: Apr, NCP -	V5: Dec, NWA - Sep, NCP +	V5: Oct, NCP + Aug, ERA - Aug, ARC -	V5: Sep, ARC - Sep, NCP + Aug, ERA - Mar, ARC +	V5: Sep, ARC - Sep, NCP + Aug, ERA + Aug, NEA - Mar, ARC +	V5: Dec, NWP - Dec, EAS - Dec, ARC - Sep, EUA + Sep, NCP + Jul, NWP - Jul, EAS - Jul, EUA +	V5: Dec, NWP - Dec, EAS - Dec, ARC - Sep, EUA + Sep, NCP + Jul, NWP - Jul, EAS - Feb, NAM - Feb, ARC -	V5: Dec, ERA + Dec, EAS - May, ARC+ Mar, NWP+ Feb, NAM -	V5: Dec, ARC + Nov, NWP+ Aug, ERA + Jul, NEA + May, ARC+ Mar, NWP+ Feb, NAM -

Z7 and/or Z5 show high cross-correlations repeatedly are ARC, NWA and NEA. For V7 and V5, North America (NAM), NWA, ERA and NCP stand out. To exemplify the cross-correlation patterns of the OZ precipitation, significant correlations of January PPT with preceding SST and upper air parameters are presented in Tables 3 and 4. Cross-correlation tables for other months are not specifically presented, but the general patterns are similar as with January TEP and PPT.

Table 3. Same as Table 1, but for cross-correlation between OZ January precipitation and SSTs of preceding months.

1971-85	1972-86	1973-87	1974-88	1975-89	1976-90	1977-91	1978-92	1979-93	1980-94
Dec, WTP - Oct, NWA - Sep, NCP + Sep, SEI - Jul, SEI - Jul, SWI - Jul, SWP + Jul, SCP + May, SEA -	Dec, WTP - Nov, NEA - Oct, NWA - Sep, NCP + Sep, NWP + Sep, SEI - Jul, SEI - Jul, SWP + Jul, SCP + May, SEA -	Dec, WTP - Nov, NEA - Oct, NWA - Sep, NCP + Jul, SEI - Jul, SWP + Jul, SCP + Jun, SWP + May, SEA -	Dec, WTP - Oct, NWA - Sep, NCP + Sep, NWP + Jul, SWP + Jul, SCP + Jun, SWP + May, NWA - May, SEA - Apr, NCP + Mar, WTP -	Sep, NCP + Sep, NWP + Jul, SWP + Jul, SCP + Jun, SWP + May, SEA - Apr, NCP +	Dec, WTP - Oct, NCP + Sep, NCP + Sep, NWP + Aug, WTP + Jul, WTP + May, SEA - Apr, NCP + Feb, NCP +	Dec, WTP - Oct, NCP + Oct, NWP + Oct, NEA + Sep, NCP + Sep, NWP + Sep, WTP + Aug, WTP + Aug, SWP - Jul, SWP + Jul, SCP + Jul, WTP + Apr, NCP + Apr, SWI - Mar, ERA +	Dec, WTP - Oct, NCP + Oct, NWP + Oct, NEA + Oct, WTP + Oct, NCP + Sep, NCP + Sep, NWP + Sep, WTP + Aug, WTP + Jul, SWP + Apr, SWI -	Dec, SWA + Dec, WTP - Oct, NCP + Oct, NWP + Sep, SEA - Aug, WTP + Aug, SWP + Apr, SEP + Apr, SWI - Mar, NCP -	Oct, NCP + Sep, SEA - Aug, WTP + Apr, SEP + Apr, SWI - Mar, NCP -

Table 4. Same as Table 1, but for cross-correlation between OZ January precipitation and upper air data (temperature, geopotential height, and geostrophic wind) of preceding months.

1971-85	1972-86	1973-87	1974-88	1975-89	1976-90	1977-91	1978-92	1979-93	1980-94
T7: Nov, EUA - T5: Jul, EAS + Jul, NWP + Z7: Oct, ERA + Oct, NWA - Sep, ERA + Jul, EAS + Jul, NWP + Z5: Oct, NWA - Oct, ERA + Jul, NWP + Jul, EAS + Jun, EUA + U7: Nov, EUA + Nov, NEA + Oct, ERA - Jul, EAS - Mar, NAM - U5: Oct, ERA - Aug, ARC - Jul, NWP - Jul, EAS - V7: Nov, NEA - Nov, EUA - Oct, ARC - Oct, ERA - Oct, NWA - May, ERA + Feb, NAM - Feb, NEA + V5: Oct, ARC - May, ERA +	T7: Nov, EUA - Jul, EAS + Jul, NWP + T5: Aug, NCP - Jul, EAS + Jul, NWP + Jun, ERA + Z7: Nov, NEA + Nov, EUA + Oct, ERA + Sep, ERA + Jul, EAS + Jul, NWP + Jun, ARC - Z5: Oct, NWA - Oct, ERA + Jul, NWP + Jul, EAS + U7: Dec, NCP - Oct, ERA - U5: Oct, ERA - Aug, ARC - Jul, NWP - Jul, EAS - V7: Dec, ARC + Nov, NEA - Nov, EUA - Oct, ERA - Oct, NWA - May, ERA + Apr, EAS - Feb, NAM - V5: May, ERA +	T7: Nov, EUA - Feb, ARC - T5: Sep, NCP - Sep, NEP - Jul, EAS + Jul, ARC + Mar, ERA + Feb, ARC - Z7: Nov, NEA + Nov, EUA + Sep NCP - Sep, NEP - Jul, EAS + Jul, NWP + Jun, ARC - Jun, NAM - Z5: Oct, NWA - Sep, NCP - Jul, NWP + Jul, EAS + U7: Sep, NEP + Sep, NCP + Apr, NWP + U5: Oct, ERA - Oct, NWA + Sep, NEP + Sep NCP + Aug, ARC - V7: Dec, ARC + Dec, NAM + Nov, NEA - Nov, EUA - Oct, ERA - Sep, NCP - Sep, NWA + May, ERA + Feb, ERA + V5: Nov, NEA - Sep, NCP - May, ERA +	T7: Nov, EUA - Feb, ARC - Feb, ERA - T5: Sep, NCP - Jul, ARC + Jun, ERA + Feb, ARC - Z7: Apr, NWP - Z5: Oct, NWA - U5: Oct, ERA - Oct, NWA + Sep, NEP + Aug, ARC - V7: Nov, NEA - Nov, EUA - Oct, ERA - Sep, NCP - May, ERA + Feb, ERA- V5: Oct, ARC - Sep, NCP -	T7: Mar, ERA + Feb, ARC - Feb, ERA - T5: Dec, EUA + Sep, NCP - Sep, NEP - Aug, ARC + Jul, ARC + Mar, ERA + Feb, ARC - Z7: Oct, NWA - Sep, NWA - Jul, EAS + Jul, NWP + Apr, NWP - Z5: Oct, NWA - Sep, NEP - Jul, NWP + Jul, EAS + Apr, NWP - U7: Sep, NEP + Sep, NWA+ Sep, NCP + May, ERA - U5: Oct, NWA + Sep, NEP + Sep NCP + Aug, ARC - May, ERA - V7: Nov, NEA - Nov, EUA - Sep, NCP - Sep, NAM + Aug, ARC + Feb, NWA- V5: Oct, ARC + Sep, NCP - Sep, NAM -	T7: Jun, ERA + Mar, ERA + Feb, ARC - Feb, ERA - T5: Aug, ARC + Jul, ARC + Jul, ERA + Jun, ERA + Mar, ERA + Z7: Oct, NWA - Sep, NWA - Jul, EAS + Jul, NWP + Apr, NWP - Z5: Oct, NWA - Jul, NWP + Jul, EAS + Apr, NWP - U7: Sep, NWA+ May, ERA - U5: Oct, NWA + Aug, ARC - Jul, EAS + V7: Sep, NAM + May, ARC - Feb, NWA- V5: Oct, ARC + Sep, NWA + Sep, NAM -	T7: Jul, EAS + Mar, ERA + Feb, ARC - Feb, ERA - T5: Dec, EUA + Sep, NCP - Sep, NEP - Aug, ARC + Jul, ARC + Mar, ERA + Z7: Nov, ARC - Sep, NWA - Sep, NAM - Jul, EAS + Jul, NWP + Jul, ERA + Apr, NWP - Z5: Oct, NWA - Jul, NWP + Jul, EAS + Apr, NWP - U7: Sep, NWA+ May, ERA - U5: Oct, NWA + Jul, NWP - Jul, EAS + V7: Nov, NEA - Nov, EUA - Sep, NAM + May, ARC - Feb, NWA- Feb, NAM - V5: Sep, NAM - Aug, ARC + Aug, ERA +	T7: Jul, EAS + Jul, NWP + May, ARC + Mar, ERA + Feb, ARC - T5: Dec, EUA + Jul, NWP + Jul, EAS + Mar, ERA + Z7: Jul, EAS + Jul, NWP + Jul, ERA + Apr, NWP - Z5: Oct, NWA - Jul, NWP + Jul, EAS + Apr, NWP - U7: May, ERA - U5: Oct, NWA + Jul, NWP - Jul, EAS - V7: Nov, NEA - Nov, EUA - Aug, ARC + Jul, NCP - Jul, ERA + Apr, EAS - V5: Dec, ERA - Jul, NCP - Apr, NWP +	T7: Sep, NWA + May, ARC + May, NCP + Feb, ARC - T5: Aug, ARC + Jul, ERA + Jul, EAS + May, ERA - Mar, ARC - Z7: Nov, ERA - Aug, ARC + Jul, EAS + Jul, NWP + Jul, ERA + Z5: Aug, ARC + Jul, NWP + Jul, EAS + May, ERA - U7: May, ERA - U5: Jul, NWP - Jul, EAS - Jul, ARC + May, ERA - May, NWP - V7: Nov, NEA - Nov, EUA - Aug, ARC + Jul, NCP - Jul, EUA + V5: Aug, ARC + Jul, ARC + Jul, EUA +	T7: Sep, NWA + T5: Aug, ARC + Jul, ERA + Jul, EAS + Mar, ARC - Z7: Nov, NCP + Aug, ARC + Jul, NWP + Z5: Nov, NCP + Aug, ARC + Jul, NWP + Jul, EAS + May, EAS - May, NWP - U7: Nov, NCP - Jul, ARC + May, ERA - U5: Nov, NCP - Oct, NWA + Oct, ERA + Jul, ARC + May, ERA - May, NWP + V7: Nov, ERA + Oct, NWA - Oct, ERA - Aug, ARC + Jul, NCP + V5: Jul, NCP +

The significance and consistency of these spatial and temporal teleconnection patterns justify the construction of regression forecasting models on the basis of such correlations. However, by the very nature of teleconnection study, in no way can such a correlation analysis alone definitively reveal or confirm teleconnection mechanisms involved. In the present analysis it is particularly noteworthy that the Southern Hemisphere SSTs show strong teleconnection with TMT and PPT of OZ, although no Southern Hemisphere upper air parameters are involved in such teleconnections. This seems to indicate that the Southern Hemisphere forcing of SSTs is dispersed to the Northern Hemisphere. Yet, no ready explanation is available for such a mechanism. An anonymous reviewer of this paper comments that stationary waves may be the answer, although a wave triggered in the Southern Hemisphere will meet easterlies in the tropics acting as a barrier, except in the westerly duct region. It is an intriguing question that arises from the present global teleconnection analysis, which demands thorough synoptic-dynamic analyses in the future.

It should be pointed out that all predictors utilized in this study, both SSTs and upper air parameters, are known to undergo interdecadal variations. This is a well-recognized characteristic of the general circulation, but the generally accepted comprehensive theory on this observed fact still does not exist. It is conceivable that both the internal dynamic-thermodynamic processes and the external forcing are involved and interact at various time scales. Among many possible linkages of such non- linear processes, it is most likely the surface and deep ocean currents act through SSTs. The problem is further complicated by the possible forcing of human-induced changes in the atmospheric composition. In the recent review of this problem, Barnett *et al.* (1999) conclude that recent changes in global climate inferred from surface air temperature are not likely due solely to natural causes, and that at present it is not possible to make a very confident statement about the relative contribution of specific natural and anthropogenic forcing to observed climate change. One essential reason for the use of continuous 15-year data periods for the construction of regression models is indeed to detour this problem at the onset to capture the prevalent modes of variations during the 15 years.

5. MLR and PCR schemes

Using the 15-year data periods 1961-75 through 1980-94, OZ climate elements were predicted for the years beyond the 15-year data periods with multiple linear and principal component regression models. For each month the maximum of 20 MLR and 20 PCR models were formulated for temperature and precipitation for the different 15-year time periods. Table 5 lists the number of models that have selected preceding SSTs and predictands of the previous year as predictors.

Table 5. Comparison of MLR and PCR models for selection of predictors by the number of models which selected each category of predictor among 20 models of each forecast month.

Predictand	Temperature (TMP)				Precipitation (PPT)			
Predictor	SST		TMP 12 months ago		SST		PPT 12 months ago	
Model	MLR	PCR	MLR	PCR	MLR	PCR	MLR	PCR
January	15	20	5	12	18	20	4	10
February	12	20	1	8	10	20	6	6
March	15	20	3	7	16	20	5	7
April	13	20	3	11	15	20	1	14
May	10	20	5	11	10	20	1	9
June	15	20	5	9	12	20	3	10
July	11	20	2	9	7	20	7	8
August	13	20	10	11	11	20	3	10
September	14	20	3	9	9	20	7	9
October	14	20	4	12	9	20	4	14
November	10	20	4	14	11	20	0	11
December	14	20	5	11	17	20	2	14

SST is selected in all cases of PCR and is the forcing function in the chain of the physical processes to form climate patterns through heat releases to the atmosphere. The PCR scheme is expected to focus in the most fundamental mode of climate variations, and thus it is reasonable to see all PCR cases choose SST as a predictor. Three leading predictors selected for the OZ January PCR models are shown in Table 6. Various upper air patterns are basically a reflection of the SST patterns, and their selection by PCR models is not as overwhelming as SSTs. The recorded value of the predictand in the previous year is also employed for approximately half of the PCR models, representing the persistence of the climatic pattern in their gradual change.

Table 6. List of predictors used for OZ January models: name, preceding month, and region of predictors. Predictors are listed in the order of their contribution to percentage variance.

Model	Fitting period	Three leading predictors		
Temperature	1961-75	SST(Jul, ETA)	TMP(Jan, OZ)	T5(Feb, EUA)
(TMP)	1962-76	SST(Aug, NEA)	T7(Aug, NWP)	U7(Aug, NWA)
	1963-77	U7(Oct, ERA)	SST(Apr, SWP)	TMP(Jan, OZ)
	1964-78	V5(Oct, ERA)	SST(May, SWP)	TMP(Jan, OZ)
	1965-79	V7(Jun, ERA)	SST(Apr, SWP)	Z5(Jul, ERA)
	1966-80	TMP(Jan, OZ)	SST(May, SWP)	Z7(Dec, NAM)
	1967-81	SST(Jul, SWP)	TMP(Jan, OZ)	V7(Aug, NAM)
	1968-82	SST(Jul, SWP)	TMP(Jan, OZ)	V7(Aug, ERA)
	1969-83	SST(Jul, SWP)	V5(Aug, NWA)	U7(Dec, NEA)
	1970-84	SST(Jul, SWP)	TMP(Jan, OZ)	Z7(Dec, NAM)
	1971-85	SST(Jul, SWP)	V7(Aug, ERA)	Z5(Dec, ARC)
	1972-86	SST(Jul, SWP)	Z5(Dec, ARC)	TMP(Jan, OZ)
	1973-87	SST(Jul, SWP)	TMP(Jan, OZ)	Z7(May, ARC)
	1974-88	TMP(Jan, OZ)	SST(Jul, SWP)	U5(Aug, NEA)
	1975-89	SST(May, SWP)	U5(Aug, NEA)	V7(Aug, ERA)
	1976-90	SST(Aug, SWP)	U5(Aug, NEA)	Z5(Aug, NEA)
	1977-91	V7(Jul, NWP)	SST(Jul, SWP)	Z7(Dec, NAM)
	1978-92	Z5(Aug, NEA)	SST(Mar, SCP)	TMP(Jan, OZ)
	1979-93	SST(Oct, SWI)	Z5(Aug, NEA)	U5(Aug, NEA)
	1980-94	TMP(Jan, OZ)	SST(Aug, NWA)	Z7(Mar, NCP)
Precipitation	1961-75	SST(Feb, NEA)	PPT(Jan, OZ)	Z7(Nov, NWA)
(PPT)	1962-76	V5(Jul, ERA)	SST(Mar, NWP)	PPT(Jan, OZ)
	1963-77	SST(Mar, WTP)	U5(Aug, NWP)	Z5(Jun, ARC)
	1964-78	SST(Oct, SEA)	V7(Jul, ERA)	U7(Dec, ARC)
	1965-79	SST(Jul, NWP)	V7(Oct, ERA)	U5(Sep, ARC)
	1966-80	SST(Dec, WTP)	PPT(Jan, OZ)	Z7(Oct, ERA)
	1967-81	U5(Jul, NWP)	SST(Dec, WTP)	PPT(Jan, OZ)
	1968-82	V5(Aug, ARC)	SST(Dec, WTP)	Z7(Oct, ERA)
	1969-83	PPT(Jan, OZ)	SST(Jul, SEI)	Z7(Oct, ERA)
	1970-84	SST(Oct, NWA)	V7(Aug, ARC)	PPT(Jan, OZ)
	1971-85	SST(Sep, NCP)	U5(Oct, ERA)	V7(May, ERA)
	1972-86	SST(Sep, NCP)	Z5(Oct, NWA)	U5(Aug, ARC)
	1973-87	SST(Sep, NCP)	U5(Aug, ARC)	Z5(Oct, NWA)
	1974-88	U5(Aug, ARC)	SST(Sep, NCP)	V7(Sep, NCP)
	1975-89	V7(Feb, NWA)	SST(Sep, NCP)	PPT(Jan, OZ)
	1976-90	V5(Oct, ARC)	SST(Sep, NCP)	PPT(Jan, OZ)
	1977-91	PPT(Jan, OZ)	SST(Sep, NCP)	V7(Feb, NWA)
	1978-92	PPT(Jan, OZ)	SST(Aug, WTP)	Z7(Apr, NWP)
	1979-93	V7(Jul, ERA)	SST(Aug, WTP)	U7(May, ARC)
	1980-94	SST(Sep, SEA)	U7(Nov, ERA)	V7(Nov, NCP)

Comparisons of MLR and PCR with OZ January model forecasts for two years beyond the data periods are contained in Table 7. The MLR and PCR are compared in terms of

Table 7. Comparison of MLR and PCR for the OZ January models with RMSE, absolute error and correlation between the forecasts and actual records. The listing is for the first and second year forecasts beyond the 15-year data periods.

	RMSE		Absolute error		Correlation	
Model	MLR	PCR	MLR	PCR	MLR	PCR
Temperature						
1st year	0.33	0.20	1.02	0.60	0.89	0.96
2nd year	0.69	0.66	2.35	2.11	0.24	0.33
Precipitation						
1st year	2.78	2.26	9.48	7.47	0.92	0.95
2nd year	8.09	7.62	27.91	25.56	0.50	0.37

RMSE, absolute error, and correlation between recorded and forecast values. PCR clearly produced better forecast skill than MLR. PCR objectively selects several new orthogonal reference frames among predictors after reducing the problem of multicollinearity, providing more efficient regression coefficients. It is also apparent that the first year forecasts after the data periods are superior to those of the second year. As generally expected, the model deteriorates with time, and deterioration is more serious for the MLR than for the PCR. As will be presented in the next section, the PCR generally retain considerable forecast skill beyond the data periods. For this reason, the forecast experiments presented in the remainder of this paper are only those of the PCR scheme.

6. Forecast experiments

Regression forecast of OZ TMP and PPT were performed with each of the 15-year period datasets from 1961 to 1995. Separate regression models were formulated for TMP and PPT of each month of the year for each data period. The data beyond the 15-year data period were used only as the real time data with the models without updating model coefficients. Forecasts were conducted for all years beyond the 15-year data period whenever the real time data were available.

As listed in Table 6, some predictors appear repeatedly in many 15-year periods throughout the entire data period. For instance, predictor July SST in SWP for OZ temperature is a common predictor from 1967-81 to 1974-88 models. January OZ TMP of the previous year for OZ temperature prediction is another common predictor appearing from 1966-80 to 1968-82 and from 1972-86 to 1974-88. The September SST in NCP is common for OZ precipitation from 1971-85 to 1977-91, as is the December SST in WTP from 1966-80 to 1968-82. The January OZ PPT of the previous year is also common for the OZ precipitation forecast from 1975-89 to 1978-92. The repeated appearance of the same predictors in models may indicate the stability of the scheme and their utility in synoptic analysis.

Table 8 (a). RMSE and SS for up to ten years beyond 15-year data periods for PCR forecast of OZ temperature. NS in the SS section indicates no skill score. For the data periods after 1970-84 have less than 10 years beyond the periods. Only the available years are used in compilation.

	Temperature model											
Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
RMSE												
1st yr	0.20	0.22	0.24	0.23	0.21	0.15	0.19	0.31	0.32	0.23	0.49	0.63
2nd yr	0.66	0.68	0.61	0.49	0.38	0.29	0.31	0.37	0.36	0.37	0.48	0.59
3rd yr	0.81	0.76	0.78	0.51	0.46	0.28	0.27	0.33	0.41	0.42	0.54	0.89
4th yr	0.82	0.93	0.46	0.56	0.46	0.31	0.33	0.42	0.43	0.48	0.51	0.79
5th yr	0.67	0.80	0.42	0.64	0.57	0.35	0.39	0.48	0.44	0.53	0.46	0.78
6th yr	0.70	0.86	0.54	0.71	0.66	0.31	0.31	0.48	0.45	0.42	0.52	0.99
7th yr	0.80	1.00	0.46	0.58	0.60	0.37	0.31	0.46	0.45	0.50	1.12	0.97
8th yr	0.71	1.00	0.58	0.58	0.59	0.33	0.30	0.53	0.47	0.48	0.52	1.01
9th yr	0.78	0.98	0.61	0.66	0.42	0.41	0.47	0.47	0.54	0.65	0.48	0.80-
10th yr	0.73	1.22	0.56	0.73	0.65	0.47	0.38	0.42	0.59	0.76	0.65	0.43
SS												
1st yr	91.7	94.6	98.5	99.4	99.8	99.9	99.9	99.7	99.5	99.5	93.3	44.4
2nd yr	12.5	36.9	90.9	97.6	99.3	99.7	99.7	99.6	99.5	98.8	94.1	56.2
3rd yr	NS	33.3	87.0	98.0	99.0	99.8	99.8	99.7	99.4	98.7	93.3	18.5
4th yr	NS	8.7	96.0	97.7	99.1	99.8	99.8	99.6	99.4	98.3	94.6	42.3
5th yr	NS	40.2	96.9	97.2	98.8	99.7	99.7	99.5	99.4	98.2	95.8	47.7
6th yr	NS	39.7	95.4	96.8	98.4	99.8	99.8	99.5	99.4	98.9	95.2	25.3
7th yr	NS	24.4	96.9	97.8	98.8	99.7	99.8	99.6	99.4	98.6	79.3	36.1
8th yr	NS	37.1	95.5	98.0	99.0	99.8	99.8	99.5	99.4	98.8	95.9	33.4
9th yr	NS	44.3	95.5	97.9	99.5	99.7	99.6	99.6	99.3	98.0	96.8	59.7
10th yr	24.0	12.8	96.7	97.5	99.0	99.7	99.8	99.7	99.3	97.5	94.7	87.2

The performance of OZ temperature and precipitation forecast for up to 10 years beyond the 15-year data period in terms of RMSE and SS is contained in Table 8. The values of RMSE and SS as listed in the table are averages of all models from 1961 to 1994. After the 1970-84

Table 8 (b). Same as Table 8 (a), but for precipitation.

	Precipitation model											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
RMSE												
1st yr	2.39	1.93	2.65	4.36	7.04	5.17	5.58	5.81	9.28	10.24	9.11	14.63
2nd yr	8.07	6.46	6.40	10.83	11.57	15.67	9.11	5.49	11.96	13.45	11.72	13.84
3rd yr	7.00	6.39	6.94	8.15	12.64	7.82	11.23	7.35	10.79	13.46	13.87	14.00
4th yr	9.72	6.66	8.53	10.64	13.39	6.52	9.97	13.68	12.99	16.21	11.38	13.04
5th yr	11.21	7.63	8.22	11.31	13.36	10.14	10.12	12.15	15.24	15.95	28.32	17.66
6th yr	9.87	9.11	11.51	9.89	15.01	10.23	11.28	12.88	16.89	19.16	15.33	18.15
7th yr	10.76	7.44	10.99	11.97	14.32	9.04	10.53	17.77	13.64	15.64	14.97	17.80
8th yr	10.33	10.49	13.32	13.24	14.95	10.70	11.95	14.91	11.62	18.79	18.65	16.66
9th yr	7.18	7.26	11.96	14.54	12.91	10.92	12.52	9.57	12.50	22.26	15.41	11.21
10th yr	9.30	9.87	15.43	19.26	15.52	14.16	9.81	12.25	15.35	15.42	13.66	15.47
SS												
1st yr	97.6	98.6	98.8	97.1	94.3	95.8	94.8	92.8	84.2	80.6	89.3	61.5
2nd yr	72.6	84.8	93.4	83.3	86.1	63.3	86.3	93.7	76.1	71.4	84.1	65.5
3rd yr	81.0	87.5	92.3	91.4	84.9	91.4	80.7	88.6	80.0	70.7	80.9	71.5
4th yr	65.7	87.9	88.4	86.5	83.0	94.5	85.9	59.3	73.6	62.1	86.2	74.0
5th yr	54.7	83.8	89.2	84.8	83.9	87.8	85.8	69.8	67.5	66.6	26.0	62.6
6th yr	66.4	79.6	79.1	89.6	81.6	88.8	84.4	68.9	62.4	56.4	80.0	65.4
7th yr	65.8	87.8	83.7	86.5	83.2	91.3	86.5	44.7	76.8	72.4	84.2	71.5
8th yr	63.5	78.9	78.0	85.3	82.5	88.2	84.3	63.4	86.9	69.7	77.9	70.1
9th yr	85.1	91.1	84.2	82.3	87.0	88.9	87.0	86.2	86.8	61.6	84.8	87.6
10th yr	79.2	85.2	75.6	71.2	82.2	83.9	93.2	82.2	82.5	77.3	89.4	78.6

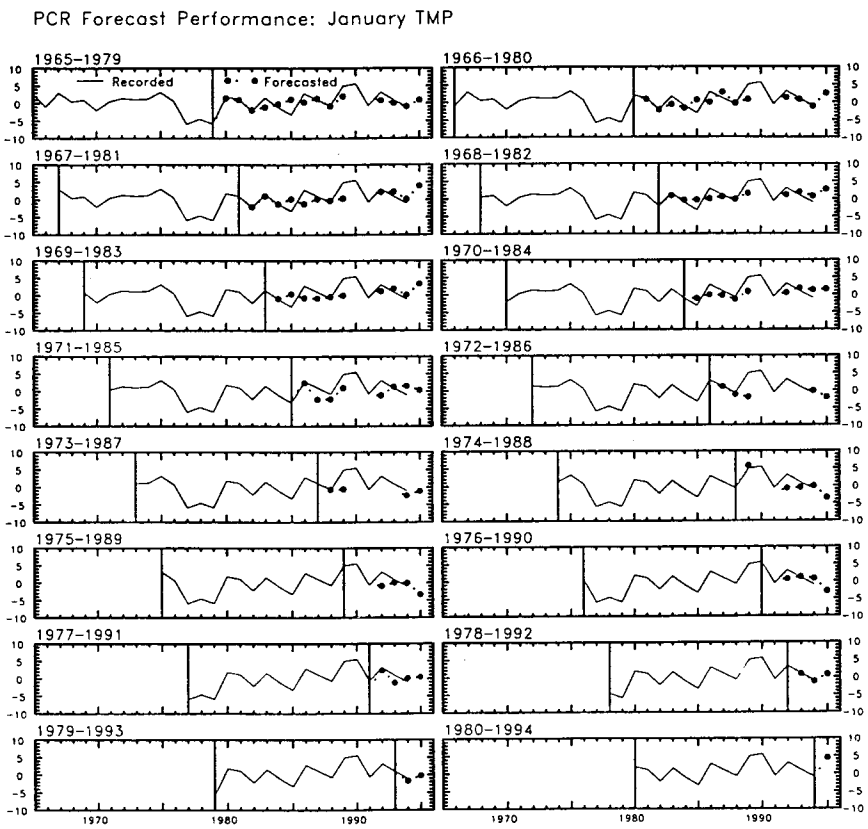


Fig. 4. Forecast performances for OZ January temperature for the period 1965-79 through 1980-94. Vertical lines identify the data periods. Dots are the forecast values and the light lines are for the recorded values. The unit of temperature is °C.

period less than 10 years were available beyond the 15-year period and only the available years were used in averaging. The small values of RMSE and high values of SS indicate high forecast skill. For OZ temperature forecasts, the first forecast years generally show much higher skill scores than those of the second forecast years and beyond. Exceptions are the November and December models, in which the first year forecast skills are slightly lower than those of the second year forecasts for these months, implying a close second year forecast skill with that of the first year. In general, the PCR forecast of the first year after the data period yields a very high forecast skill. This is consistent with the results, as shown in Figures 4 through 11, in which the forecasts for the first years are almost identical as the recorded values. Also as seen in these figures, although somewhat inferior to the forecasts of the first year, the forecasts for the second year and thereafter still generally show a useful level of forecast skill. In precipitation forecast models, the first year forecasts show their RMSE from 1.93 to 14.63 and SS from 61.5 to 98.8, and the second year forecasts show RMSE from 5.49 to 15.67 and SS from 63.3 to 93.7.

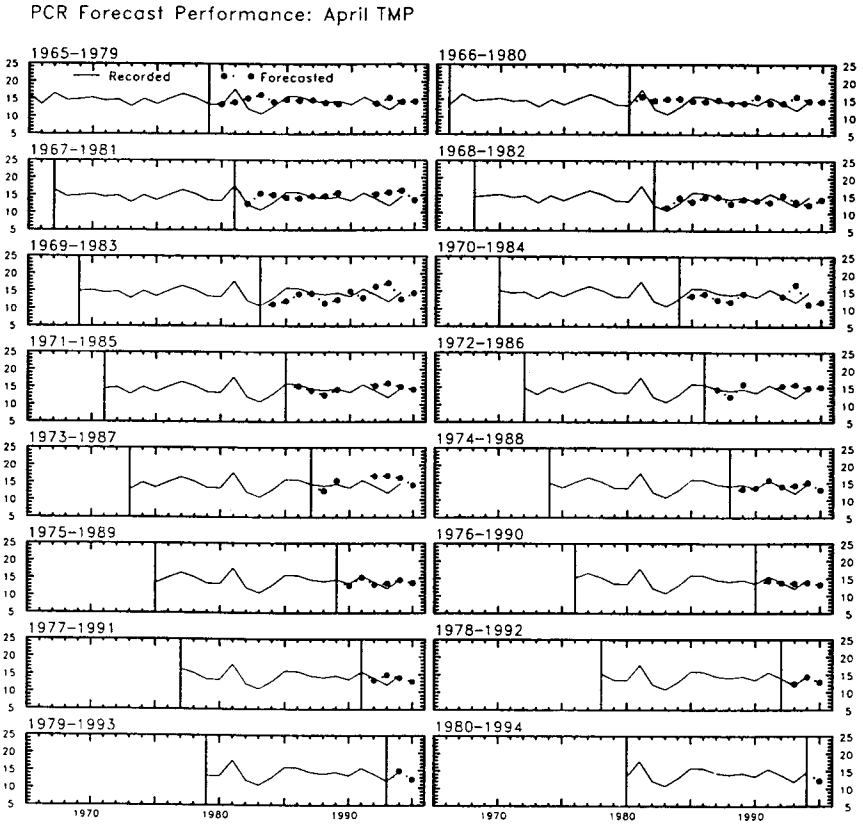


Fig. 5. Same as Figure 4, but for OZ April temperature.

Performance of the OZ temperature and precipitation forecasts by PCR, as illustrated for January, April, July, and October in Figures 4 through 11, show the forecast results beyond the respective 15-year data periods and the corresponding recorded values: the former by dots and the latter by lines. The forecasts were performed only for those years beyond the 15-year data periods whose data was the basis of regression formulation; thus the validity of forecast experiment is ensured. The forecasts were not conducted for years where the real-time values for regression models were not available.

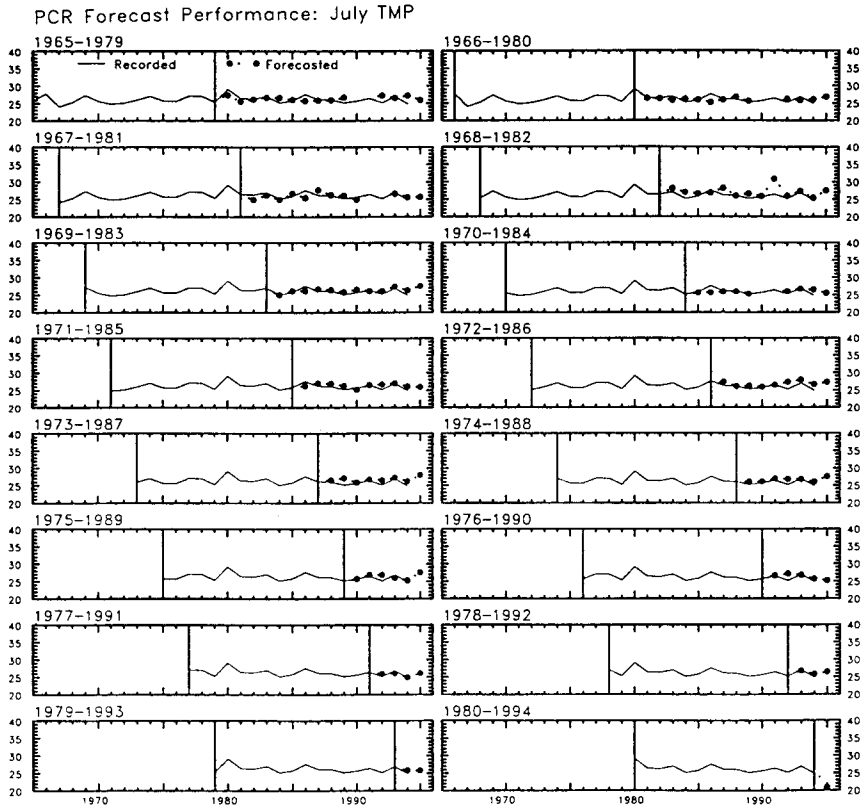


Fig. 6. Same as Figure 4, but for OZ July temperature.

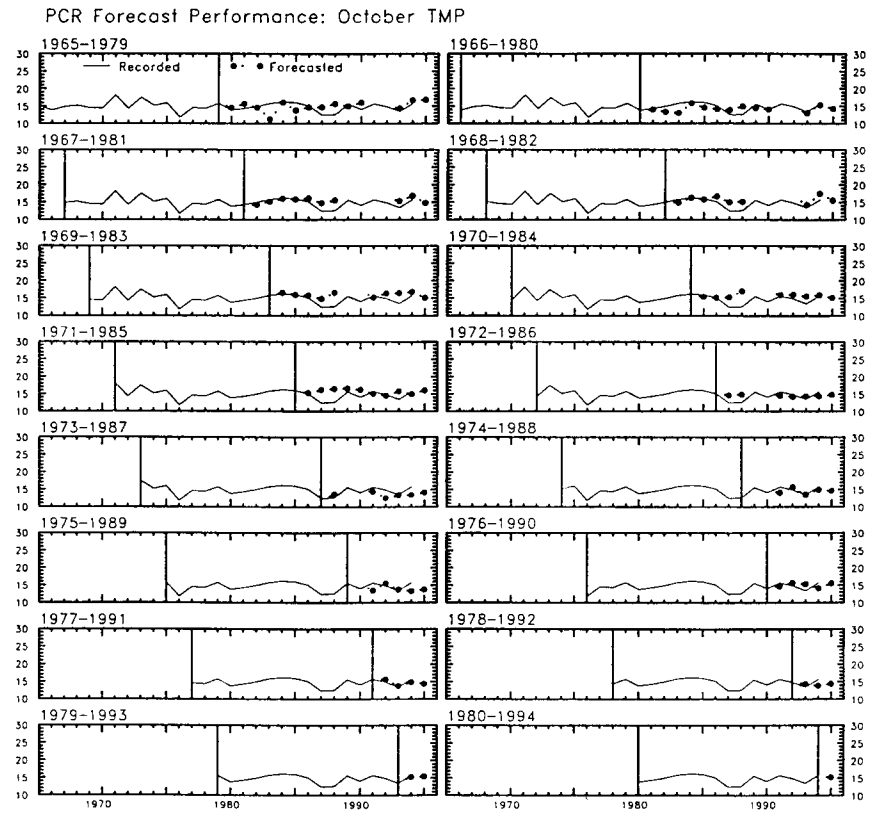


Fig. 7. Same as Figure 4, but for OZ October temperature.

The information contained the these figures and Table 8 collectively indicate a considerable level of forecast skill of respective regression models beyond the 15-year data period. This indicates the stability of the PCR regression scheme of this study. It is desirable that the forecast regression models be updated annually by incorporating the data of the current years for the next years. Yet without updating, a reasonable degree of forecast skill still may be expected.

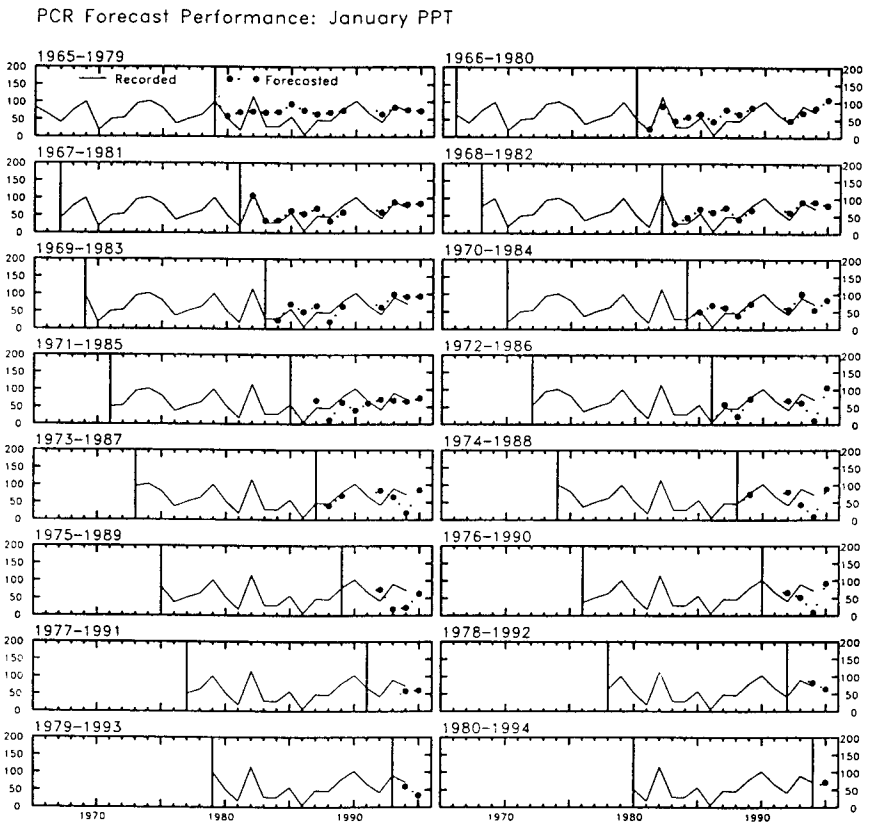


Fig. 8. Same as Figure 4, but for OZ January precipitation. The unit of precipitation is mm.

Temperature forecasts, as shown in Figures 4 through 7, are more accurate than precipitation forecasts in Figures 8 through 11. This is expected because of complexity in the precipitation environment and processes. Forecast skill shows some seasonality for some models. For instance, April and July temperature forecasts seem more stable than those of January; and April precipitation forecast seems more accurate than January, July and October. These phenomena are apparently related to the prevailing circulation patterns of the seasons.

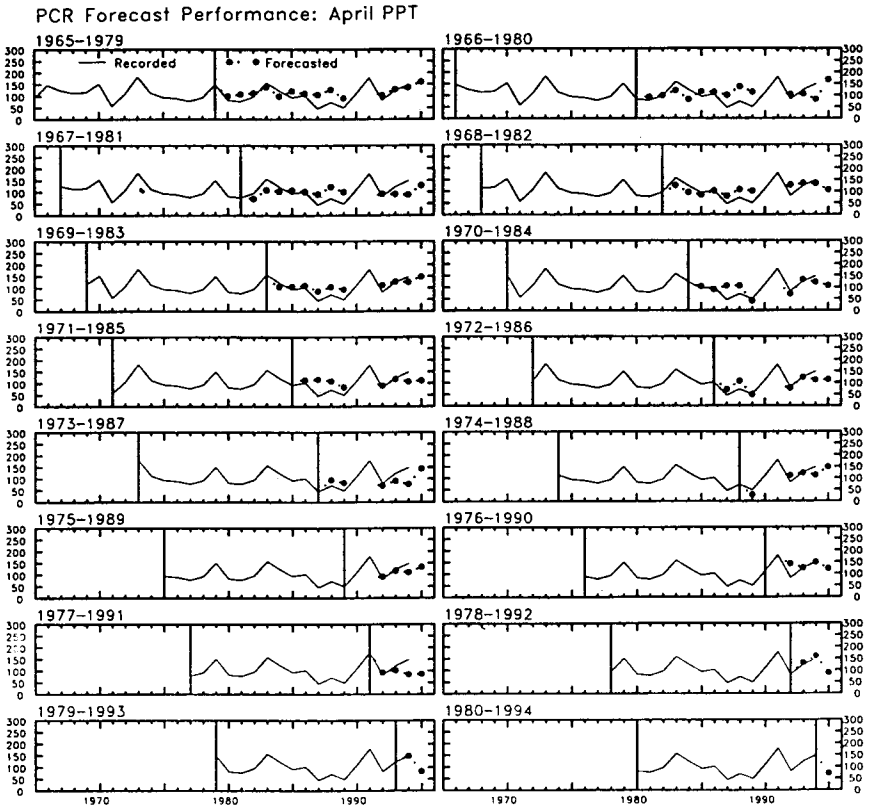


Fig. 9. Same as Figure 4, but for OZ April precipitation.

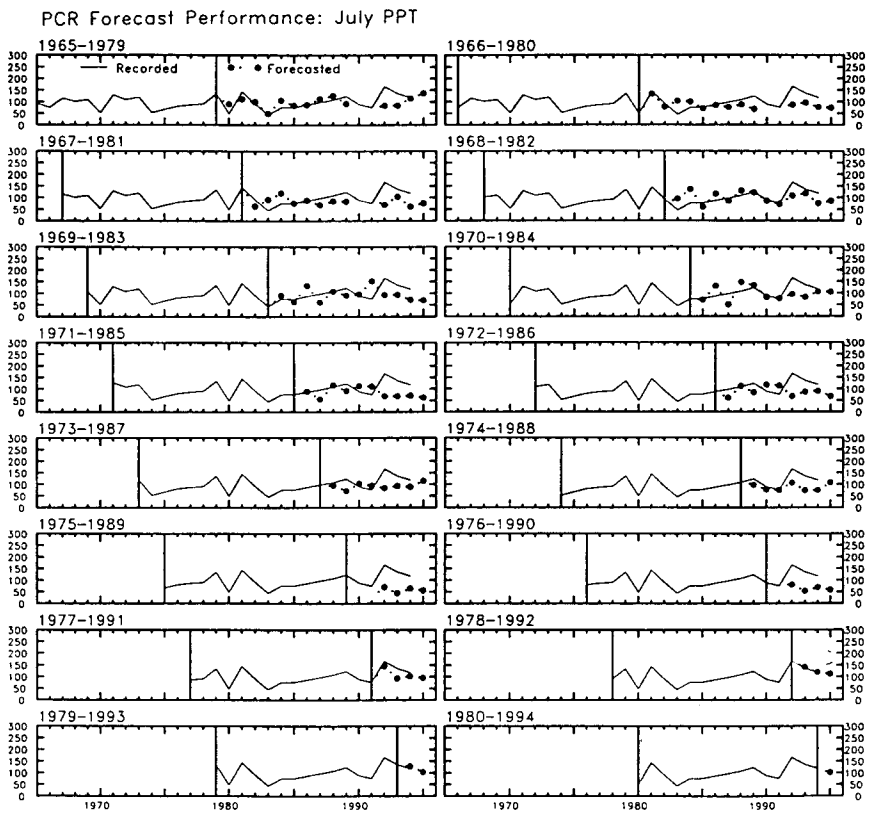


Fig. 10. Same as Figure 4, but for OZ July precipitation.

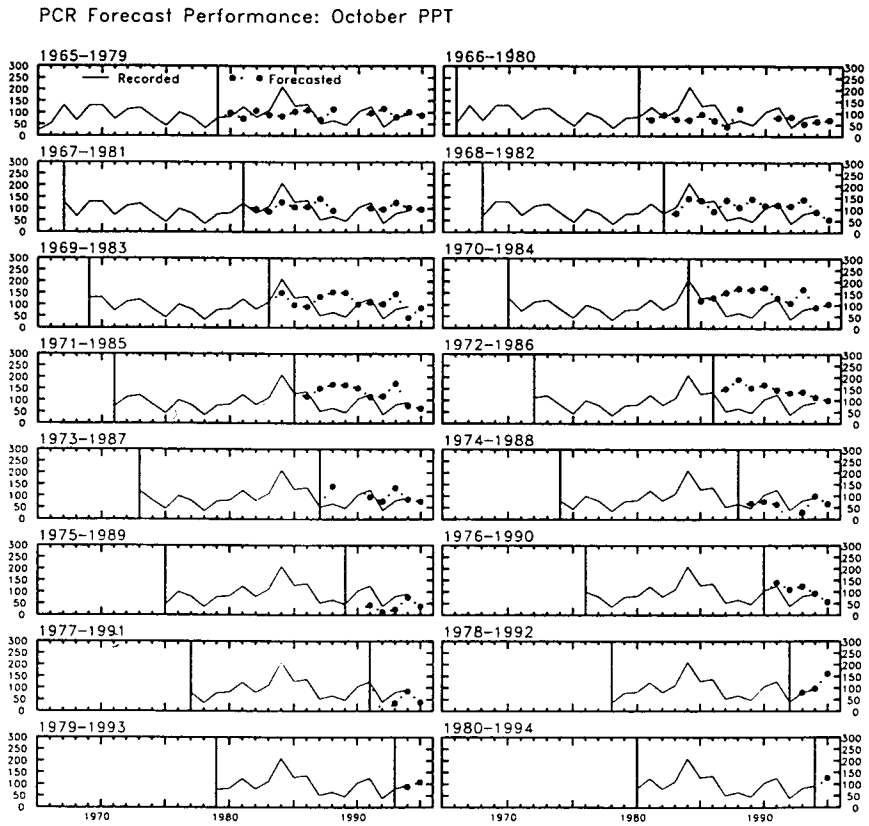


Fig. 11. Same as Figure 4, but for OZ October precipitation.

The least successful cases in the forecast as identified by differences between recorded and forecasted value are listed in Table 9. Those cases for OZ temperature are listed for the difference of absolute value of $\pm 4^{\circ}$C or larger. May, June, July, September, and October do not appear in this list for OZ temperature. Eight cases of the total 14 OZ temperature cases are associated with El Niño and two cases with La Niña. The less successful cases in OZ precipitation forecast are those cases of difference with absolute value of ± 80 mm or larger. January, February, March, April and August do not appear in the table. Eleven of total 21 OZ precipitation cases for the least successful forecast are also associated with El Niño, and two cases with La Niña. The identification of prevailing conditions of El Niño and La Niña follows the definition by Trenberth (1997). Examination of correlation patterns (for example, Table 1) reveals that most of the strong teleconnections with OZ climate of the Southern Hemisphere occur before the prevalent El Niño period of the late 1980s and 1990s. During the period of El Niño dominance during the last decade, teleconnection with the Southern Hemisphere is very limited. The case study of these less successful cases may be an interesting endeavor in the future.

Table 9. The least successful cases of PCR forecasts by the difference of recorded (REC) and forecasted (FCT) value. Prevailing SST condition is by El Niño (EN), La Niña (LN), or normal.

Temperature |Diff| >4°C	Data period for fitting	FCT Year	Difference (°C): REC-FCT	SST Condition
January	1961-75	1977	+4.88	EN
	1962-76	1978	+4.29	EN
	1963-77	1979	+4.63	normal
February	1961-75	1976	-4.02	LN
	1978-92	1994	+4.08	normal
March	1962-75	1978	+4.90	EN
	1978-92	1993	+4.42	EN
April	1967-81	1983	+4.63	EN
August	1964-78	1980	-4.06	EN
November	1961-75	1976	+5.36	LN
	1975-89	1991	+4.01	EN
December	1963-77	1979	-4.22	EN
	1969-83	1984	-4.71	LN
	1975-89	1990	-4.59	normal
Precipitation |Diff| >80 mm	Data period for fitting	FCT Year	Difference (mm): REC-FCT	SST Condition
May	1972-86	1988	+81.31	EN
	1973-87	1988	+94.79	EN
	1974-88	1990	-90.82	normal
	1976-90	1992	+99.86	EN
June	1975-89	1991	+238.50	EN
July	1976-90	1992	-85.99	EN
September	1966.80	1981	+94.94	normal
	1970-84	1986	-110.94	EN
	1977-91	1993	-115.62	EN
October	1971-85	1987	+98.01	normal
	1972-86	1987	+97.84	normal
	1972-86	1988	+125.00	LN
	1975-89	1991	-83.22	EN
November	1962-76	1978	-110.07	normal
	1969-83	1985	-84.09	normal
	1970-84	1985	-96.58	normal
December	1966-80	1982	-161.38	EN
	1967-81	1982	-155.16	EN
	1971-85	1986	+100.91	EN
	1973-87	1989	+102.70	LN
	1975-89	1990	-91.60	normal

7. Concluding remarks

The principal component analysis of long-term monthly temperature and precipitation records with a dense observation network over the United States indicates the utility of the principal components to provide the seasonal-range forecast for the region.

The cross-correlation analysis of the OZ temperature and precipitation suggests the useful teleconnection patterns between OZ climate variables and the global antecedent SSTs and upper air circulation. A principal component regression forecasting scheme was constructed in which the predictands were related to their first components that were regressed on antecedent SSTs, upper air parameters and the preceding OZ record. It was shown that the principal component regression scheme is superior to the multiple linear regression scheme.

The forecast experiments were performed for each month of the year with 15-year segments of data for the periods from 1961-75 to 1980-94 for years beyond the 15-year data periods. The examination of the results reveals useful and stable forecasting of the principal component regression models up to 10 years beyond the 15-year data period without significant deterioration of models with time.

The forecast models provide area mean forecast values in the Ozark Highlands area. It should be possible to derive the local forecast by considering the regional differences within the area through additional climate analysis.

Acknowledgments

This research has been supported under the cooperative agreement USDI GS #1445-CA09-95-0069 #2 between Biological Resources Division, U. S. Geological Survey and the University of Missouri-Columbia. The authors are indebted to Dr. G. D. Willson for his comments and consultation, to Dr. W. B. Kurtz for his thorough reading and comments on the manuscript and to the anonymous reviewer whose constructive comments have led to a great improvement of the paper. They also want to thank Mr. H. Kim and Ms. Tanya Akyüz for their technical assistance.

REFERENCES

Barnett, T. P., K. Hasselmann, M. Chelliah, T. Delworth, G. Hegerl, P. Jones, E. Rassmusson, E. Roeckner, C. Ropelewski, B. Santer and S. Tett, 1999. Deletion and attribution of recent climate change: a status report. *Bull. Amer. Meteor. Soc.*, **80**, 2631-2659.

Basilevsky, A., 1994. *Statistical Factor Analysis and Related Methods.* John Wiley & Sons, Inc., New York, 737pp.

Blackmon, M. J., J. E. Geisler and E. J. Pitcher, 1983. A general circulation model study of January climate anomaly associated with interannual variation of equatorial Pacific sea surface temperatures. *J. Atmos. Sci.*, **40**, 1410-1425.

Davis, R. E., 1976. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, **6**, 249-266.

Draper, N. R. and H. Smith, 1981. *Applied Regression Analysis.* John Wiley & Sons, Inc., New York, 709 pp.

Geisler, J., M. L. Blackmon, G. T. Bates and S. Muños, 1985. Sensitivity of January climate response to the magnitude and position of equatorial Pacific sea surface temperature anomalies. *J. Atmos. Sci.*, **42**, 1037-1049.

Kung, E. C., and J. -G. Chern, 1994. Large-scale modes of variations in the global sea surface temperatures and Northern Hemisphere tropospheric circulation, *Atmósfera*, **7**, 143-158.

Kung, E. C., and J. -G. Chern, 1995. Prevailing anomaly patterns of the Global Sea Surface temperatures and tropospheric responses, *Atmósfera*, **8**, 99-114.

Kung, E. C., and H. Tanaka, 1985. Long-range forecasting of temperature and precipitation with upper air parameters and sea surface temperature in a multiple regression approach. *J. Meteor. Soc. Japan*, **63**, 619-631.

Kung, E. C., J. -G. Chern, and D. E. Smith, 1995. Global sea surface temperatures and associated long-range predictability of the Northern Hemisphere circulation and local climatological variables. *Terrestrial, Atmospheric and Oceanic Sciences*, **6**, 553-577.

Kutzbach, J. E., 1967. Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. *J. Appl. Meteor.*, **6**, 791-802.

Lee, J. -W., 1997. *Long-Range Variability and Predictability of the Ozark Highlands Climate Elements.* Ph. D. Dissertation, University of Missouri-Columbia, 116 pp.

Montroy, D. L., 1997. Linear relation of central and eastern North American precipitation to tropical sea surface temperature anomalies. *J. Climate*, **10**, 541-557.

Myers, R. H., 1990. *Classical and Modern Regression with Applications.* Duxbury Press, Belmont, 488pp.

NCDC, 1993. *Summary of The Month Co-operative T-D 3220*, NCDC, Asheville, 29pp.

Nicholls, N., 1984. The stability of empirical long-range forecast techniques: a case study. *J. Climate Appl. Meteor.*, **23**, 143-147.

Park, C. -K., and E. C. Kung, 1988. Principal components of the North American summer temperature field and the antecedent oceanic and atmospheric condition. *J. Meteor. Soc. Japan,* **66**, 677-690.

Preisendorfer, R. W., and C. D. Mobley, 1988. *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, Amsterdam, 425pp.

Smith T. M., R. W. Reynolds, R. E. Livezey, and D. Stoker, 1996. Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate,* **9**, 1403-1420.

Ting, M., and H. Wang. Summertime U. S. precipitation variability and its relation to Pacific sea surface temperature. *J. Climate,* **10**, 1853-1873.

Trenberth, K. E., 1997. The definition of El Niño. *Bull. Amer. Meteor. Soc.,* **78**, 2771-2777.

Tribbia, J. J., 1991. The rudimentary theory of atmospheric teleconnections associated with ENSO. *Teleconnections Linking Worldwide Climate Anomalies.* (M.H. Glantz, R.W. Katz and N. Nicholls, Eds.) Cambridge University Press, 185- 305.

Unger, D. A., 1996a. Skill assessment strategies for screening regression predictions based on a small sample size. Preprints, Thirteenth Conference on Probability and Statistics in the Atmospheric Sciences. San Francisco, CA, February 21-23, 1996. 260-267

Unger, D. A., 1996b. Forecasts of surface temperature and precipitation anomalies over the U.S. using screening multiple linear regression. *Experimental Long-Lead Forecast Bulletin,* **5**, 44-45.

Walsh, J. E. and M. B. Richman, 1981. Seasonality associations between surface temperatures over the United States and the North Pacific Ocean. *Mon. Wea. Rev.,* **109**, 767-783.

Woodruff, S. D., R. J. Slutz, R. L. Jenne and P. M. Steurer, 1987. A comprehensive ocean-atmosphere data set . *Bull. Amer. Meteor. Soc.,* **68**, 1239-1250.