Prediction of monthly-seasonal precipitation using coupled SVD patterns between soil moisture and subsequent precipitation

Yongqiang Liu
Forestry Sciences Laboratory, USDA Forest Service, Athens, Georgia, USA

Received 8 May 2003; revised 24 June 2003; accepted 7 July 2003; published 15 August 2003.

[1] It was suggested in a recent statistical correlation analysis that predictability of monthly-seasonal precipitation could be improved by using coupled singular value decomposition (SVD) patterns between soil moisture and precipitation instead of their values at individual locations. This study provides predictive evidence for this suggestion by comparing skills of two statistical prediction models based on the coupled SVD patterns and local relationships. The data used for model development and validation are obtained from a simulation over East Asia with a regional climate model. The results show a much improved skill with the prediction model using the coupled SVD patterns. The seasonal prediction skill is higher than the monthly one. The most remarkable contribution of soil moisture to the prediction skill is found in warm seasons, opposite to that of sea surface temperature. *INDEX TERMS:* 1854 Hydrology: Precipitation (3354); 1866 Hydrology: Soil moisture; 1869 Hydrology: Stochastic processes; 3322 Meteorology and Atmospheric Dynamics: Land/atmosphere interactions. **Citation:** Liu, Y., Prediction of monthly-seasonal precipitation using coupled SVD patterns between soil moisture and subsequent precipitation, *Geophys. Res. Lett.*, *30*(15), 1827, doi:10.1029/2003GL017709, 2003.

1. Introduction

[2] With the capacity to retain anomalous signals over a long period [*Delworth and Manabe*, 1988; *Vinnikov et al.*, 1996], soil moisture can contribute to long-term variability of the surface temperature and precipitation by passing its relatively slow anomalous signals to the atmosphere. Observational and modeling studies have indicated close relationships of initial soil moisture conditions with anomalies in subsequent monthly and seasonal surface temperature and precipitation [e.g., *Huang et al.*, 1996; *Eltahir*, 1998]. Thus, it is possible to improve predictability of long-term variability of the two atmospheric variables by using soil moisture.

[3] *Karl* [1986] illustrated the great value of soil moisture to monthly and seasonal objective forecasts of the surface temperature. However, it is difficult to determine soil moisture's role in improving predictability of long-term precipitation. Precipitation is controlled by large-scale atmospheric circulations, whose long-term variability is in turn related to other factors such as sea surface temperature (SST), and by local land-atmospheric interactions. The atmospheric circulations play a predominant role in most cases, while soil moisture becomes important only under certain circumstances. Therefore, identification of such circumstances using both dynamic [e.g., *Koster et al.*, 2000] and statistical techniques [e.g., *Wang and Kumar*, 1998] is of essential importance for demonstrating and developing a capacity in making long-term precipitation prediction using soil moisture.

[4] In a recent study concerning the above issue [*Liu*, 2002], the coupled patterns between soil moisture and precipitation were identified using singular value decomposition (SVD) [*Bretherton et al.*, 1992; *Ting and Wang*, 1997]. It was indicated that the correlation with soil moisture preceding precipitation is much more significant for the SVD expansion series than original data series, suggesting that predictability of monthly-seasonal precipitation could be improved by using soil moisture in the form of its coupled SVD patterns with precipitation. The present study provides predictive evidence for this suggestion by comparing skills of statistical forecasts using SVD patterns and local relationships.

2. Method

[5] Prediction models are developed for both the SVD pattern (SVD model) and local relationship (local model). The SVD model is derived from a linear regression relationship between soil moisture [$u(x, t)$] and precipitation [$v(x, t + n)$] with application of a lag SVD analysis, where x and t are space and time indices, respectively, and n is a period of time (see the Appendix A for the derivation). The precipitation series for the SVD analysis lags the soil moisture series by one month/season (i.e., $n = 1$). The resulting SVD patterns therefore represent spatial relationships of soil moisture with subsequent precipitation. In the prediction model, precipitation of a coming month/season is determined mainly by the SVD expansion coefficients of soil moisture of the current month/season, and SVD spatial patterns of precipitation.

[6] The local model is composed of a set of linear regression between soil moisture and precipitation at individual locations, $v(x, t + n) = D(x)u(x, t)$, where u and v are normalized and D is the regression coefficient. The local model has two major differences from the SVD model: it has a separate regression relationship at each location x, and its regression coefficient at a location is obtained independent of soil moisture and precipitation at other locations.

[7] The method to build the SVD model is similar to that using canonical correlation analysis (CCA) [e.g., *Barnett and Preisendorfer*, 1987]. *Mo* [2002] combined CCA with an assemble technique to predict U.S. rainfall with soil moisture and other predictors. Major features are common between SVD and CCA. SVD is adopted mainly considering that SVD is favored when a predictor and a predictand

0094-8276/03/2003GL017709$05.00

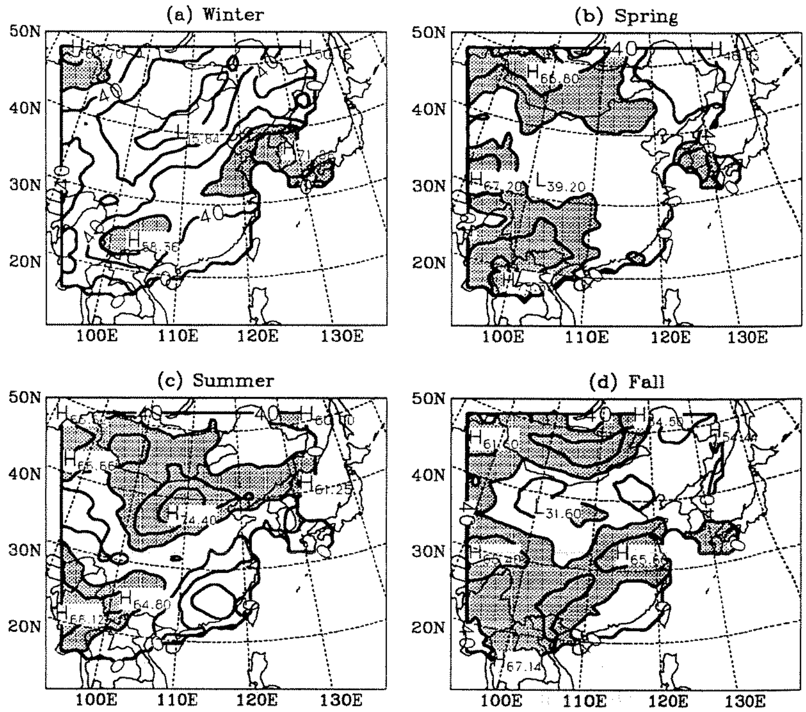

Figure 1. Skill of seasonal precipitation prediction with SVD model. Panels (a–d) represent winter through fall. The contour interval is 10 and the unit is %. The areas with skill above 50% are shaded.

participating in each pattern linkage are similar to those found in the individual data set empirical orthogonal function (EOF) patterns [*Barnston*, 1994].

[8] A prediction with the SVD model is made in the following steps: (1) to construct a validation data set $u_{valid}(x, t)$ and $v_{valid}(x, t + n)$ by extracting data of the kth year (k = 1, 2, ..., K; K = 10 for this study) from the original data set $u(x, t)$ and $v(x, t)$, and a modeling data set $u_{model}(x, t)$ and $v_{model}(x, t + n)$ from the remaining data in the original data set; (2) to build the SVD model using the modeling data set; (3) to calculate the SVD time coefficient using $a_i(t) = \sum_{j=1}^{M} u_{valid}(j, t)p_i(j)$ and to predict precipitation of the validation period using equation (A7) in the Appendix A; and (4) to repeat the above steps for k = 1 *through K*, which generates K separate data sets for each of modeling, validation, and prediction. For a specific month/season at each location, $v_{valid}(x, t + n)$ and the corresponding predicted values each are equally divided into normal, above and below normal categories. A prediction is regarded as correct if it has the same category as $v_{valid}(x, t + n)$. The prediction skill is measured by the ratio of correct to total number of prediction made.

[9] Because of the unavailability of systematic measurements of soil moisture, this study uses model output data. As used by *Liu* [2002], soil moisture and precipitation over East Asia simulated with the National Center for Atmospheric Research regional climate model (RegCM) [*Giorgi et al.*,1993] are used in this study. The simulation has a domain of 90 × 79 grid points with a horizontal resolution of 60 km, and is integrated for the period from January 1987 to December 1997 with the first year regarded as a spin-up time.

[10] A regional climate model (RCM) like RegCM is usually integrated over a short period up to a few years [*Liu*, 2002], which makes it difficult to use its output data assessing statistical significance of prediction skill. Longer simulations have been performed with some global climate models coupled with land-surface processes [e.g., *Bonan et al.*, 2002]. These models could be an alternate tool to produce soil moisture and precipitation needed for the SVD analysis. A RCM has been used mainly in the consideration that, with the boundary conditions updated every 12 hours during the integration period primarily using meteorological observations, it is expected to produce relatively realistic regional circulation patterns and hydrological processes.

[11] A major difference in the validation method stated above from the one used in, e.g., *Barnett and Preisendorfer* [1987] is that, because the original data set for this study has a short period of 10 years, continuous monthly/seasonal data are used to build models or to validate results. Thus, the serially uncorrelated condition required by a strict cross validation is not met here. A significance test does not make much sense for the same reason. Thus, the results obtained here are used only as a criteria to judge which model (the SVD or local model) has better skill. In addition, the prediction models are built using all months/seasons of a year in a modeling data set, while validation is made for each month/season of a year in a validation data set to briefly look at seasonal dependence of the prediction skill.

3. Results and Discussion

[12] Monthly and seasonal precipitation forecasts are made separately, using monthly soil moisture and seasonal soil moisture as predictors, respectively. Four leading SVD patterns are used. An experiment with the pattern numbers from 2 to 10 indicates that the results are not sensitive to the number of patterns. Figure 1 shows geographic distribution

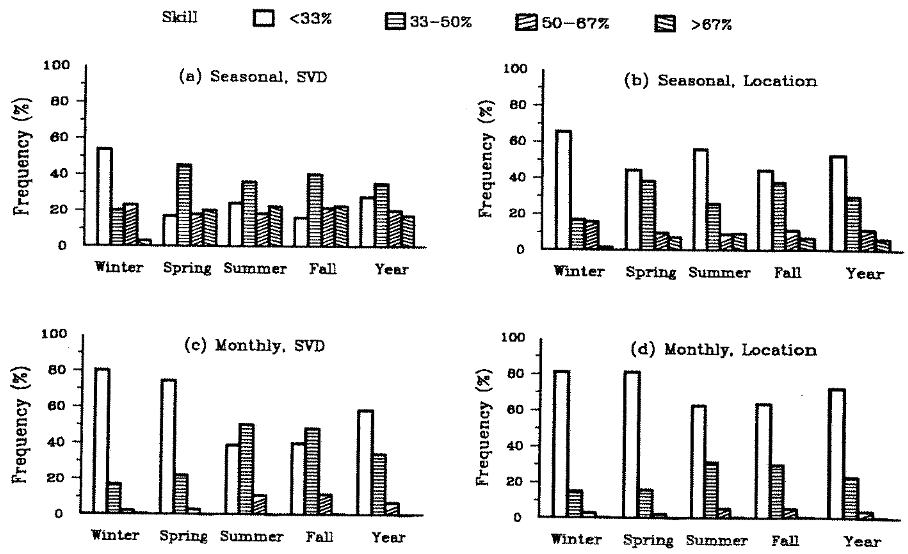

Figure 2. Frequency percentage of various skill levels of precipitation prediction. Panels (a–b) represent seasonal and monthly forecasts with the SVD model, and (c–d) the corresponding forecasts with the local model. The seasonal values in panels (c–d) are averages of Dec–Feb (winter), Mar–May (spring), June–Aug (summer), and Sept–Nov (fall).

and seasonal dependence of the SVD model prediction skill of seasonal precipitation. The areas with a skill greater than 50% are shaded. This skill level, selected arbitrarily, is considered remarkably different from the skill level of 33% for a random prediction. The skill exceeds 50% over two areas. One is in northern China, which covers the western corner of the region in winter, extends eastward in spring and summer, and retreats in fall. The other is in southern China, which is limited in the western corner of the region in winter, and extends northeastward in three other seasons.

[13] The skill of the local model (not shown) is much lower than that of the SVD model. The spatially averaged skill for the seasonal prediction is 46% with the SVD model, but only 36% with the local model. To have a more detailed comparison, the entire range of skill is divided into Levels I–IV (<33, 33–50, 50–67, and >67%, respectively). The higher the frequency percentage for Level I, the worse the prediction skill. It is opposite for Level IV. For seasonal prediction (Figure 2), the frequency percentages of Level I with the SVD model are about 50 (winter) and 10s–20s (other seasons and annual average), compared with the corresponding values with the local model of nearly 70 and 40s–50s; Those of Level III or IV for all seasons except winter are about 20 with the SVD model, compared with only about 10 with the local model.

[14] For monthly prediction, the overall skill of the SVD model is also higher than that of the local model. The annual frequency percentages of Level I are 59 and 72 for the SVD and local models, respectively, while those of Level II are 36 and 25, respectively. The difference between the two models is the most significant in summer and fall. The skill is lower for monthly than seasonal prediction. Its annual average frequency percentage with the SVD model is twice as much as that of seasonal prediction for Level I, while only about half of that of seasonal prediction for Level III or IV.

[15] Based on the results for the case of this study, the predictability of monthly and seasonal precipitation is indeed improved by using the coupled SVD patterns of soil moisture and subsequent precipitation, which therefore supports the suggestion made in *Liu* [2002]. As pointed out by *Barnston* [1994], precipitation at a given location is determined by the combined effects of systematic relationships, which mostly are of large spatial scale, and identification of its patterns as wholes can enhance predictive skill at individual locations.

[16] The seasonal dependence of the skill of precipitation prediction using soil moisture is opposite to that using SST, whose contribution to prediction skill of long-term precipitation was found profound in winter [*Barnston*, 1994]. Because of little water exchange on the land surface and weaker land-atmospheric interactions, the impact of soil moisture on precipitation variability is small in winter. The East Asian monsoon may also have an adverse impact on the prediction skill. The skill is low in northern China during the cool seasons and southeastern China during the warm seasons, where the winter and summer monsoon circulations prevail, respectively.

[17] The typical length of timescale of soil moisture variability is about 2–3 months [*Vinnikov et al.*, 1996] and is longer in an interactive land-atmospheric system [*Liu and Avissar*, 1999], suggesting that the role of soil moisture in precipitation variability may be more important at seasonal than monthly scale. In fact, the correlation coefficients of the four leading SVD expansion series with soil moisture preceding precipitation are 0.765, 0.835, 0.735, and 0.779 for the seasonal data series, compared with 0.658, 0.777, 0.574, and 0.653 for the monthly data series. This may explain the higher prediction skill for seasonal than monthly precipitation.

Appendix A: SVD Prediction Model

[18] A regression equation is built to predict $\boldsymbol{v}(t+n) = [v(x, t+n)]$ with $\boldsymbol{u}(t) = [u(x, t)]$ (both normalized),

$$v(t+n) = \boldsymbol{D}\,\boldsymbol{u}(t) \qquad \text{(A1)}$$

where $x = 1, 2, \ldots, M$ with M being the number of space locations; $t = 1, 2, \ldots, N(= N_0 - n)$ with N_0 and n being the length of the original data set and a period of time (one month/season for this study), respectively. Determining the coefficient matrix $\boldsymbol{D}$ using the least squaresd approximation (LSA), we have,

$$\boldsymbol{D}\sum_{t=1}^{N} \boldsymbol{u}(t)\boldsymbol{u}(t)^{\mathrm{T}} = \sum_{t=1}^{N} \boldsymbol{v}(t+n)\boldsymbol{u}(t)^{\mathrm{T}} \tag{A2}$$

Applying SVD [*Bretherton et al.*, 1992] to $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t+n)$,

$$\boldsymbol{u}(t) = \sum_{k=1}^{M} a_k(t)\boldsymbol{p}_k \tag{A3}$$

$$\boldsymbol{v}(t+n) = \sum_{k=1}^{M} b_k(t+n)\boldsymbol{q}_k \tag{A4}$$

where $\boldsymbol{p}_k = [p_k(x)]$ and $\boldsymbol{q}_k = [q_k(x)]$ are spatial patterns, and $a_k(t)$ and $b_k(t+n)$ temporal coefficients. Applying the properties $\sum_{t=1}^{N} a_i(t)b_j(t+n) = \sigma_i\delta_{ij}$ and $\boldsymbol{p}_i\boldsymbol{p}_j^T = \delta_{ij}$, where σ is singular value, and devoting $\overline{a_i^2} = \sum_{t=1}^{N} a_i^2(t)$, we have,

$$\boldsymbol{v}(t+n) = \sum_{i=1}^{M} \sigma_i a_i(t)\boldsymbol{q}_i / \sum_{i=1}^{M} \overline{a_i^2} \tag{A5}$$

An empirical factor f_i is adopted to approximate the summation using M_l ($<M$) leading patterns,

$$\boldsymbol{v}(t+n) = \sum_{i=1}^{M_l} f_i\sigma_i a_i(t)\boldsymbol{q}_i / \sum_{i=1}^{M} \overline{a_i^2} \tag{A6}$$

and applying LSA again, we have the final form of the prediction model,

$$\boldsymbol{v}(t+t) = \sum_{i=1}^{M_l} \left[\sigma_i a_i(t)\boldsymbol{q}_i / \overline{a_i^2}\right] \tag{A7}$$

[19] **Acknowledgments.** The author would like to thank Dr. Hui Wang for providing the SVD computer code, two anonymous reviewers for their valuable comments, and Dr. Gary Achtemeier and Ken Forbus for conducting internal review and editorial check. This research was supported by the USDA Forest Service, and by NASA under grant NAG8-1513.

References

Barnett, T. P., and R. Preisendorfer, Origins and levels of monthly and seasonal forest skill for United States surface air temperatures determined by conanical correlation analysis, *Mon. Wea. Rev*, *115*, 1825–1850, 1987.

Barnston, A. G., Linear statistical short-term climate predictive skill in the Northern Hemisphere, *J. Clim.*, *7*, 1513–1564, 1994.

Bonan, G. B., K. W. Oleson, M. Vertenstein, S. Levis, X. Zeng, Y. Dai, R. E. Dickinson, and Z.-L. Yang, The land surface climatology of the community land model coupled to the NCAR community climate model, *J. Clim*, (accepted) 2002.

Bretherton, C. S., C. Smith, and J. M. Wallace, An intercomparison of methods for finding coupled patterns in climate data, *J. Clim.*, *5*, 541–560, 1992.

Delworth, T., and S. Manabe, The influence of potential evaporation on the variabilities of simulated soil wetness and climate, *J. Clim.*, *1*, 523–547, 1988.

Eltahir, A. B., A soil moisture-rainfall feedback mechanism. 1. Theory and observations, *Water Resour. Res.*, *34*, 765–785, 1998.

Giorgi, F., M. R. Marinucci, G. De Canio, and G. T. Bates, Development of a second generation regional climate model (RegCM2), I, Boundary layer and radiative transfer processes, *Mon. Wea. Rev.*, *121*, 2794–2813, 1993.

Huang, J., H. M. van den Dool, and K. P. Georgakakos, Analysis of model calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts, *J. Clim.*, *9*, 1350–1362, 1996.

Karl, T. R., The relationship of soil moisture parameterizations to subsequent seasonal and monthly mean temperature in the United States, *Mon. Wea. Rev.*, *114*, 675–686, 1986.

Koster, R. D., M. J. Suarez, and M. Heiser, Variance and predictability of precipitation at seasonal-to-interannual timescales, *J. Hydrometeor.*, *1*, 26–46, 2000.

Liu, Y.-Q., Spatial patterns of soil moisture connected to monthly-seasonal precipitation variability in a monsoon region, *J. Geophys. Res.*, (accepted) 2002.

Liu, Y.-Q., and R. Avissar, A study of persistence in the land-atmosphere system with a fourth-order analytical model, *J. Clim.*, *12*, 2154–2168, 1999.

Mo, K. C., Ensemble canonical correlation prediction of summer season precipitation over the United States, in "Mississippi River Climate & Hydrology Conference", May 13–17 , New Orleans, LA, pg. 36, 2002.

Ting, M. F., and H. Wang, Summertime United States precipitation variability and its relation to Pacific sea surface temperature, *J. Clim.*, *10*, 1853–1873, 1997.

Vinnikov, K., A. Robock, N. A. Speranskaya, and C. A. Schlosser, Scales of temporal and spatial variability of midlatitude soil moisture, *J. Geophys. Res.*, *101*, 7163–7174, 1996.

Wang, W. Q., and A. Kumar, A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America, *J. Geophys. Res.*, *103*, 28,637–28,646, 1998.

Y. Liu, Forestry Sciences Laboratory, USDA Forest Service, 320 Green St., Athens, GA 30602, USA. (yliu@fs.fed.us)