# Analysis of Model-Calculated Soil Moisture over the United States (1931–1993) and Applications to Long-Range Temperature Forecasts

JIN HUANG, HUUG M. VAN DEN DOOL, AND KONSTANTINE P. GEORGAKAKOS*

*Climate Prediction Center, National Centers for Environmental Prediction, Washington, D.C.*

(Manuscript received 5 January 1995, in final form 8 November 1995)

ABSTRACT

A long time series of monthly soil moisture data during the period of 1931–1993 over the entire U.S. continent has been created with a one-layer soil moisture model. The model is based on the water budget in the soil and uses monthly temperature and monthly precipitation as input. The data are for 344 U.S. climate divisions during the period of 1931–1993. The main goals of this paper are 1) to improve our understanding of soil moisture and its effects on the atmosphere and 2) to apply the calculated soil moisture toward long-range temperature forecasts.

In this study, the model parameters are estimated using observed precipitation, temperature, and runoff in Oklahoma (1960–1989) and applied to the entire United States. The comparison with the 8-yr (1984–1991) observed soil moisture in Illinois indicates that the model gives a reasonable simulation of soil moisture with both climatology and interannual variability.

The analyses of the calculated soil moisture show that the climatological soil moisture is high in the east and low in the west (except the West Coast), which is determined by the climatological precipitation amounts. The annual cycle of soil moisture, however, is determined largely by evaporation. Anomalies in soil moisture are driven by precipitation anomalies, but their timescales are to first order determined by both climatological temperature (through evaporation) and climatological precipitation. The soil moisture anomaly persistence is higher where normal temperature and precipitation are low, which is the case in the west in summer. The spatial scale of soil moisture anomalies has been analyzed and found to be larger than that of precipitation but smaller than that of temperature.

Authors found that generally in the U.S. evaporation anomalies are much smaller in magnitude than precipitation anomalies. Furthermore, observed and calculated soil moisture anomalies have a broad frequency distribution but not the strongly bimodal distribution indicative of water recycling.

Compared to antecedent precipitation, soil moisture is a better predictor for future monthly temperature. Soil moisture can provide extra skill in predicting temperature in large areas of interior continent in summer, particularly at longer leads. The predictive skill of soil moisture is even higher when the predictand is daily maximum temperature instead of daily mean temperature.

## 1. Introduction

### a. *Influences of soil moisture on temperature*

The importance of soil moisture in *long-range* temperature forecasts has been emphasized by several authors including Namias (1952, 1962). The soil moisture controls the partitioning between the sensible and latent heat fluxes, and consequently influences the temperature of the surface and the lower atmosphere. Further feedback may occur through changed cloudiness, relative humidity, surface albedo, roughness, and upper-level atmospheric circulation.

A series of numerical sensitivity experiments on the atmospheric response to soil moisture anomalies have been carried out [see Mintz (1984) and Yeh (1989) for a review]. Model-simulated climates are sensitive to soil moisture presumably because it affects evapotranspiration. Recently, two multiyear integrations have been made at the National Centers for Environmental Prediction (NCEP: formerly National Meteorological Center) with NCEP's Medium Range Forecast (MRF) model. It was found (Huang and Van den Dool 1993, hereafter referred to as HD) that precipitation has a large impact on the next month's near-surface temperature in summer in the model with interactive soil moisture (see their Figs. 3b and 3c). Moreover, month to month temperature persistence is much higher when the soil moisture feedback is active.

Observational analyses of the effects of soil moisture on temperature have been made by Walsh et al. (1985),

* Also affiliated with Hydrologic Research Center, San Diego, California, and Scripps Institution of Oceanography, University of California at San Diego, San Diego, California.

*Corresponding author address:* Dr. Jin Huang, Climate Prediction Center, NCEP/NWS/NOAA, 5200 Auth Road, Camp Springs, MD 20233.

Karl (1986), Van den Dool et al. (1986a), Georgakakos et al. (1995), and Cayan and Georgakakos (1995), to mention only a few. Most of these observational studies lack explicit soil moisture measurements. Therefore, the observational evidence is usually indirect. Walsh et al. (1985) made a series of objective specification experiments (Klein 1983) with monthly 700-mb height and surface temperature. They found that the errors in the specifications have a significant relation to soil moisture. During summer, the errors that can be attributed to soil moisture are 0.3°–0.7°C; that is, the surface is warmer than that anticipated from the large-scale flow if the soil is dry. Karl (1986) calculated the relationship between temperature and soil moisture indices. The results indicate that the soil moisture may provide some skill in predicting monthly and seasonal temperature during the spring and summer in the interior U.S. continent.

The influences of soil moisture on temperature are also suggested by several precipitation–temperature data analyses (Van den Dool 1988, 1989; Lyons 1990; HD 1993). In these studies, antecedent precipitation acts presumably as a proxy for soil moisture. Huang and Van den Dool summarized the evidence found in their previous studies as follows. 1) The observed, usually negative, $P$–$T$ correlation, with precipitation ($P$) leading temperature ($T$) by a month, is much larger than the lagged $T$–$P$ correlation. This rules out the possibility that $P$ and $T$ are correlated only because each of them is correlated to some common slowly varying cause. 2) The negative $P$–$T$ correlation is present mostly in the warmer months when the potentially large latent heat flux is subject to large change as a result of soil moisture variation. 3) The positive $T$–$T$ correlation (i.e., enhanced temperature persistence) in summer appears to be physically caused, to a large extent, by the interactive soil moisture.

### *b. Existing soil moisture datasets*

Systematically compiled observed quantitative soil moisture data over large areas and multiyear periods are lacking. In the United States, measured soil moisture data with substantial record length are only available in a few locations (Hollinger and Isard 1994).

There are several calculated soil moisture indices for monitoring drought–wetness over the United States and elsewhere. Considering that drought is caused by absence of rain, a simple minded approach is to use easily available precipitation anomalies to compute an index called precipitation anomaly classification (PAC) to detect drought. One version of the index originally developed by the Australian Bureau of Meteorology (Lee 1980) has been applied to the United States by Janowiak et al. (1986). However, as we will see later, the typical duration of a drought period is determined strongly by evaporation, and so more precise soil moisture indices should be based on a full water budget. Among the most popular indices based on the water budget is the Palmer Drought Severity index (PDSI) (Palmer 1965). However, the calculation of PDSI uses future temperature and precipitation, a so-called look-ahead procedure. This procedure limits their usefulness in prediction of future temperature. The PDSI was developed using records from only central Iowa and Kansas, so that their applicability to other regions is not quite clear. Furthermore, the index (and a host of other) gives only qualitative soil moisture data and are discontinuous in time. The prediction of temperature requires continuous quantitative soil moisture data.

The next step is to use a physical model. We could use comprehensive models (e.g., Sellers 1986), but it will be impossible to create a long-term soil moisture dataset over large areas because of a lack of input data (primarily radiation). As an alternative, Mintz and Serafini (1981) and Mintz and Walker (1993) computed global soil moisture data with a bucket model driven by easily available monthly mean surface temperature and precipitation.

The principal goals of this study are as follows.

1) To create a long-term quantitative soil moisture dataset for the United States using as input only the observed historic temperature and precipitation at 344 *climate divisions* for 1931–1993. Our model is equally simple and practical as the Mintz–Serafini model (1981). But our model has a more physically reasonable runoff formulation, and the model parameters can be calibrated using observed runoff data.

2) To improve physical understanding by studying the spatial distribution of soil moisture and its variation; studying the timescale of soil moisture anomalies; examining if precipitation could possibly be locally recycled to make soil moisture anomalies persist; and studying the difference between using daily and monthly data (precipitation and temperature) as model input.

3) To apply the calculated soil moisture historical data in predicting temperature statistically. We have reasons to expect improved long-range temperature forecasts primarily in spring and summer over continental areas on monthly to seasonal scales. Furthermore, because soil moisture affects atmospheric temperature through evaporation, which is strongest in the daytime, we expect soil moisture to have a stronger impact on daily maximum temperature than on minimum temperature (Georgakakos et al. 1995). This hypothesis will be tested in this study.

## 2. The soil moisture model and data

### *a. Model*

Soil moisture is calculated based on the water balance in the soil. The components of the water balance in the model are precipitation, evaporation, runoff (or streamflow divergence), and groundwater loss. We model the soil moisture to the extent it participates in land-surface processes, that is, usually in the upper 1–2 m of soil. The soil moisture budget over an area A can be expressed as

$$\frac{dW(t)}{dt} = P(t) - E(t) - R(t) - G(t), \quad (1)$$

where

$W(t)$ the soil water content at time $t$
$P(t)$ the mean areal precipitation over area $A$
$E(t)$ the mean areal evapotranspiration over area $A$
$R(t)$ the net streamflow divergence from area $A$
$G(t)$ the net groundwater loss (through deep percolation) from area $A$.

The streamflow divergence $R(t)$ consists of a surface runoff component $S(t)$ and a subsurface (base flow) runoff component $B(t)$:

$$R(t) = S(t) + B(t). \quad (2)$$

Following practice in operational hydrologic forecasting (e.g., Georgakakos 1986), the surface runoff and the baseflow are parameterized as follows:

$$S(t) = P(t)\left[\frac{W(t)}{W_{\max}}\right]^m \quad (2a)$$

$$B(t) = \frac{\alpha}{1+\mu} W(t), \quad (2b)$$

where $W_{\max}$ is a measure of the capacity of soils to hold water in millimeters, $m$ is a parameter with values greater than 1, $\alpha$ is the inverse of the response time of the baseflow, and $\mu$ is a dimensionless parameter that determines the portion of the subsurface flow that becomes baseflow in the channels draining out from the area of interest. The remaining portion is lost as unobserved groundwater flow, which is then given as

$$G(t) = \frac{\mu\alpha}{1+\mu} W(t). \quad (2c)$$

The evapotranspiration $E(t)$ is estimated in this model as follows:

$$E(t) = E_p \frac{W}{W_{\max}}, \quad (3)$$

where $E_p$ is the potential evapotranspiration rate in millimeters per month. The potential evapotranspiration depends mainly on the net radiative heating on the surface. However, measurements or sufficiently accurate calculation of the net radiation on the surface (needed for aerodynamic formulas) are inadequate or absent over large areas for long times. Term $E_p$ can also be estimated from pan evaporation, but the observations are lacking and the estimation has a number of problems (Sellers 1965). In this study, we calculate the potential evapotranspiration from the observed air temperature and duration of sunlight using Thornthwaite's method (1948). The rationale is that air temperature does, to a considerable extent, serve as a parameter of the net radiation. This is a shortcut of replacing a comprehensive atmospheric model as well as some interactions by prescribing observed temperature and precipitation.

Following Thornthwaite (1948), the formulas of $E_p$ are expressed as

$$E_p = \begin{cases} 0 & \text{when} \quad T < 0°\text{C} \\ 16L\left(\dfrac{10T}{I}\right)^a & \text{when} \quad 0 \leq T < 26.5°\text{C} \\ -415.85 + 32.25T - 0.43T^2 & \\ & \text{when} \quad T \geq 26.5°\text{C}, \end{cases} \quad (3a)$$

with

$$L = \frac{d}{30}\frac{h}{12}$$

$$I = \sum_{M=1}^{12} i, \quad \text{where} \quad i = \left(\frac{T_M}{5}\right)^{1.514} \quad \text{and} \quad i_{\min} = 0$$

$$a = (6.75 \times 10^{-7} I^3) - (7.71 \times 10^{-5} I^2) + (1.79 \times 10^{-2} I) + 0.49, \quad (3b)$$

where $d$ is the number of days in the month, $h$ the number of hours of daylight in the middle day of the month, and $T$ the monthly mean surface air temperature in each month of the year.

## b. Parameter estimation from historical data

Parameter estimation (model calibration) was performed using both manual and automated search–optimization procedures. The hydrologic literature is replete with methods and applications of parameter-estimation procedures to conceptual hydrologic rainfall–runoff models (e.g., reviews in Sorooshian 1991; Rajaram and Georgakakos 1989). Parameter estimation of conceptual models can be thought to be analogous to experimental calibration of assumed functional relationships in the laboratory, only with large-scale data from long historical periods. The data is used to obtain values for the parameters of the functional forms of the model that represent hydrologic processes, such as surface runoff, baseflow, etc.

Initial runs establish an initial set of parameters. The automatic downhill simplex search method implemented by Press et al. (1989) was used with modifications pertaining to the establishment of infeasible regions for the parameters. The automated search procedure estimates parameters based on the changes in the value of a quantifiable error criterion function for runoff.

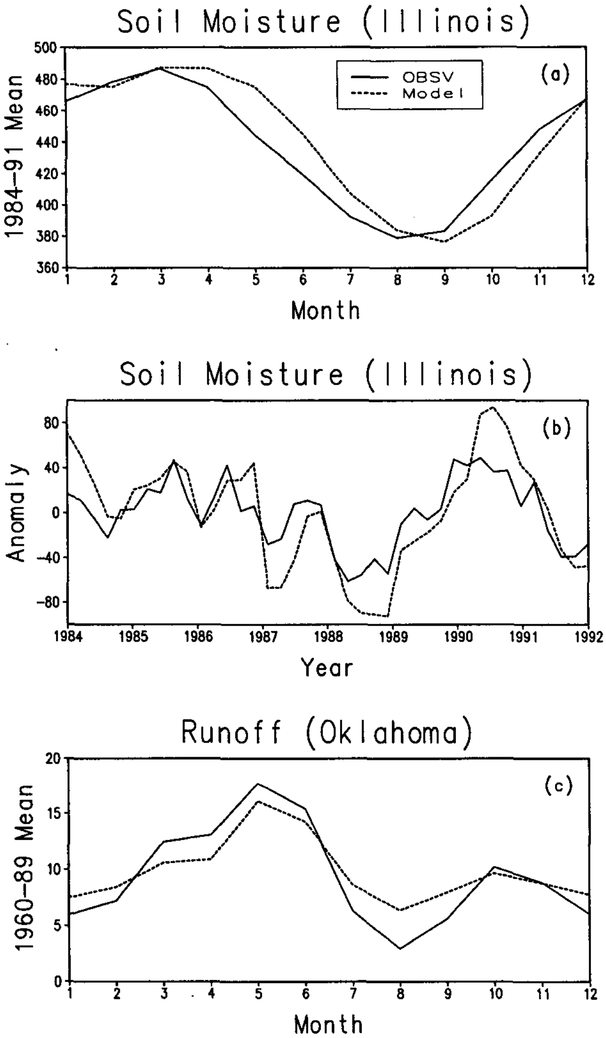


FIG. 1. (a) Observed (solid line) and calculated (dashed line) soil moisture climatology (unit: mm) averaged over 1984–1991 in Illinois as a function of month from January to December. (b) The time series of observed and calculated soil moisture anomalies (departure from 1984–1991 average) in Illinois for May to September only. (c) Observed (solid line) and calculated (dashed line) runoff climatology (unit: mm) averaged over 1960–1989 in Oklahoma as a function of month from January to December.

The automated search procedure may converge to local optima or may lock near the boundary of the feasible region. It is then necessary to start the search from several initial parameter estimates and select the convergence region with the lowest criterion value. Once feasible parameter estimates have been obtained that possess the lowest error, runs of the model are made to examine model skill in simulating hydrologically important features of the record (i.e., magnitude and timing of flood peaks, extended low flow periods, etc.). Changes in parameter estimates may be necessary to accommodate both a low error and a good reproduction of the hydrologically important features of the runoff record. There is considerable judgement involved in the parameter estimation procedure, which requires hydrologic expertise and familiarity with model applications.

*c. Data and numerical procedure*

The atmospheric datasets used to drive the soil model are monthly surface air temperature ($T$) and monthly total precipitation ($P$) at 344 U.S. climate divisions during the period 1931–1993. Daily station $T$ and $P$ data at 138 cooperative stations over 1931–1991 are also used for some specialized questions that require the use of daily data.

The observed soil moisture data in Illinois, which is the only available long-term verification dataset, are used to show the model performance. The observed soil moisture data are from 16 stations in Illinois for an 8-yr period (1984–1991) at 11 layers of soil from the top down to 2 m in the soil (Hollinger and Isard 1994). The observed runoff data (1960–1989) from a 3° by 3° area in Oklahoma (from 34° to 37°N and 96° to 99°W) are used to calibrate the model parameters.

The ideal way to use the model for the entire United States is to estimate the model parameters for various U.S. regions of similar soil and land cover and then use these parameters for each region. *In this paper, we apply the parameters estimated in Oklahoma to the entire United States* and leave the model with spatially varying parameters for future work. The model parameters estimated using the Oklahoma data are $W_{max} = 760$ mm, $\mu = 5.8$, $\alpha = 0.093$ mo$^{-1}$, $m = 4.886$, with the variables $E$, $P$, $R$, and $G$ in units of millimeters per month. The model integration starts from uniform 200 mm in January 1931. The spinup time is about 3–4 months.

Because of nonlinear surface runoff generation, the integration step is varied so as not to allow the input precipitation to exceed 2 mm per time step. This technique is also used in the operational NWS Sacramento model (Peck 1974) to prevent unreasonable runoff. We use the instantaneous[1] soil moisture estimated by the model at the end of a month for most of the analyses and to predict future temperature. However, when comparing with observed soil moisture (in Fig. 1), we use monthly mean soil moisture.

## 3. Results

*a. Analyses of the calculated soil moisture*

1) MODEL PERFORMANCE

Since the only observed soil moisture is available in Illinois, the model performance in simulating soil mois-

[1] The end of month *instantaneous* values are also very smooth in time because the input driving data are monthly mean and because the accumulation of precipitation and evaporation anomalies by (1) works like a low-pass filter.

ture can only be shown in Illinois. The comparison between simulated and observed runoff was made where the calibration was done, which is for Oklahoma in this study.

We run the model starting from 1931, but the comparison to the observed soil moisture is made for the period of 1984–1991 (verification period) only. We also produce a spatial average for the entire state of Illinois based on 16 stations with observations and nine climate divisions with the model output.

After several experiments, we found that, as far as climatology is concerned, the calculated soil moisture (with $W_{max} = 760$ mm) agrees best with the observed soil moisture in the top 1.3 meters of soil. The soil moisture climatology (8-yr average) (unit: mm) as a function of month is shown in Fig. 1a for the model results (dashed line) and for observations (solid line). The seasonal cycles are broadly similar for the model and the observation, with the maximum in winter–spring and the minimum in late summer. This suggests that the soil moisture seasonal cycle basically follows the temperature seasonal cycle, which is mostly through evapotranspiration (i.e., solar radiation). The slight difference in phase between the calculated and the observed soil moisture is probably caused by using temperature instead of net radiation to estimate the potential evapotranspiration.

Figure 1b shows the time series of soil moisture anomalies (departure from the 8-year climatology for each month; units: mm) from 1984 through 1991 for warm months (May to September) only (some jumps in the graph are due to connecting September to May by a straight line). The correlation between the simulated and observed soil moisture anomaly during May to September during 1984–1991 is 0.84. Figures 1a and 1b suggest that the model can simulate reasonably the soil moisture climatology and soil moisture anomalies for Illinois, although the model parameters are not locally calibrated. We consider this an independent verification of the validity of the model.

Figure 1c shows the modeled and observed runoff climatology (1960–1989) in Oklahoma. The model can catch the seasonal cycle and the magnitude of the monthly runoff. We have to notice, however, that the verification for runoff is not independent since the calibration and verification periods are the same.

The above results are from the model with monthly precipitation as the model input. We will also examine whether or not using daily precipitation makes for a better simulation of observed soil moisture. Because of nonlinearity, daily precipitation should be better in principle. The daily data at cooperative stations (from NCDC) are used, because, unfortunately, daily climate division data for the entire United States are lacking. It should be noted that using station data, precipitation in particular, is probably much less representative than climate division data. The comparison between daily precipitation (daily temperature is also used to calculate the potential evapotranspiration) and monthly precipitation is conducted at 12 cooperative stations in Illinois. The two model results (not shown here) are very close and both differ from the observations. The results do not show significant improvement by using daily instead of monthly precipitation and temperature. This is partly due to the fact that the length of time steps for the integration of the model-governing equation is input dependent (precipitation processed during each computational time step does not exceed 2 mm). It may also be that the present model is too insensitive or that the *monthly* soil moisture is not overly sensitive to the details of when precipitation was falling.

### 2) SPATIAL DISTRIBUTION OF THE MODEL-CALCULATED SOIL MOISTURE

We have examined the spatial distributions of the soil moisture climatology and interannual variation over 1931–1993 in United States for the annual mean and for each season. The spatial pattern of soil moisture climatology is similar for each season. Figure 2a shows the map for summer (May to August), which is the 4-month "season" we are most interested in. Generally, the soil moisture content is higher in the eastern half of the continent than that in the western half (except close to the west coast). This spatial pattern is virtually the same as the distribution of annual mean precipitation (not shown). The spatial distributions of the interannual variation of the summer soil moisture, precipitation, and evaporation (plotted in Figs. 2b–d) show similar patterns. The three variables have a large variation in the east and a small variation in the west in summer. Note that larger standard deviations of soil moisture and evaporation in California in summer may be due to the high variations in precipitation in winter. The magnitude of interannual variation of evaporation is generally smaller than that of precipitation by a factor of 2 to 3 except in California.

Figure 3a shows the soil moisture autocorrelation in summer with a 3-month lag; that is, the last day of May is correlated with the last day of August based on 63-yr data. Compared to monthly precipitation persistence at the same lead (not shown, all zero essentially), the month to month soil moisture persistence is much higher. For example, the May to August autocorrelation is up to 0.8 in Utah for soil moisture, while it is less than $\pm 0.2$ for precipitation everywhere. In summer, the largest soil moisture anomaly persistence ($>0.9$) is in California, the only U.S. climate zone where a preexisting soil moisture anomaly can be wiped out only during a few winter months.

An important feature in Fig. 3a is that the soil moisture anomaly persistence in summer is much higher in the west than in the east. In other words, the same soil moisture anomaly (dry or wet) generally persists longer in the west than in the east. In order to understand this spatial pattern, we go back to the model equations. After inserting (2) and (3) into (1), we get

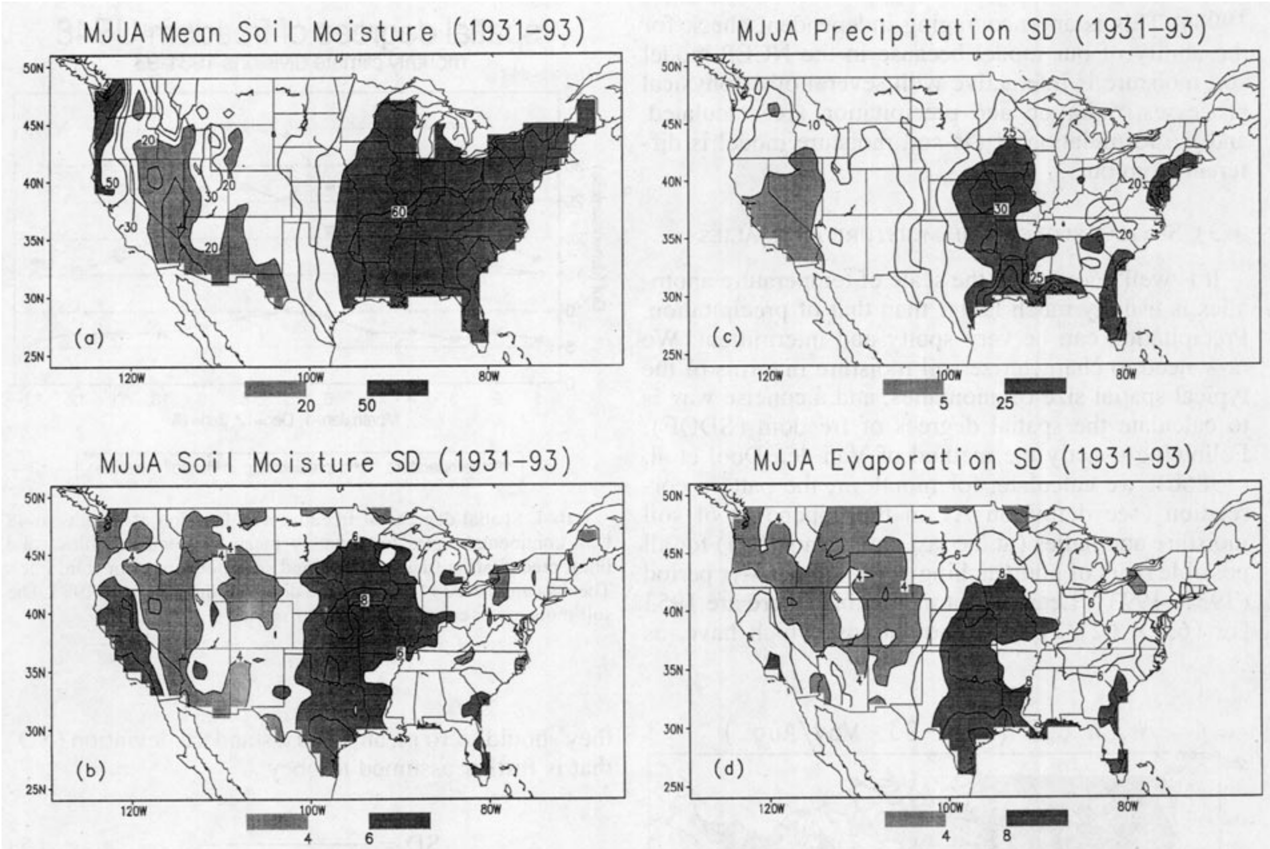


FIG. 2. U.S. maps of (a) calculated soil moisture climatology ($W/W_{max}$ 100) for months MJJA over 1931–1993 at 344 U.S. climate divisions. Areas with values lower than 20 (i.e., dry areas) are lightly shaded, and values higher than 50 (i.e., wet areas) are heavily shaded. The contour interval is 10; (b) standard deviation of 4-month mean (MJJA) calculated soil moisture ($W/W_{max}$ 100). Areas with values lower than 4 are lightly shaded, and values higher than 6 are heavily shaded. The contour interval is 2; (c) standard deviation of 4-month mean precipitation (mm mo$^{-1}$). Areas with values lower than 5 are lightly shaded, and values higher than 25 are heavily shaded. The contour interval is 5; (d) standard deviation of 4-month mean calculated evaporation (mm mo$^{-1}$). Areas with values lower than 4 are shaded, and values higher than 8 are heavily shaded. The contour interval is 2.

$$\frac{dW}{dt} = P - E_p \frac{W}{W_{max}} - \dot{P}\left(\frac{W}{W_{max}}\right)^m - \alpha W. \quad (4)$$

Then the linearized anomaly (denoted as ^, a departure from the mean) equation is

$$\frac{d\hat{W}}{dt} = \hat{P} - \bar{E}_p \frac{\hat{W}}{W_{max}} - m\bar{P}\left(\frac{\bar{W}}{W_{max}}\right)^{m-1} \frac{\hat{W}}{W_{max}} - \left(\frac{\bar{W}}{W_{max}}\right)^m \hat{P} - \alpha\hat{W} \quad (4a)$$

or

$$\frac{d\hat{W}}{dt} = a\hat{P} - c\hat{W} \quad \text{approximately;}$$

hence

$$\hat{W} = \hat{W}(0)e^{-ct}. \quad (4b)$$

If precipitation ($\hat{P}$) is assumed to be white noise, which is a reasonable assumption, (4a) becomes a first-order Markov process with evaporation and other losses ($-c\hat{W}$) being the restoring force. The decay timescale ($1/c$) of the soil moisture anomaly is then proportional to $W_{max}$ and inversely proportional to $E_p$ and $\bar{P}$. Because the field capacity $W_{max}$ is taken to be spatially constant and $E_p$ is a function of temperature only, *the timescale of the soil moisture anomaly in the model is determined to good approximation by the climatological mean temperature and precipitation.* In the western United States, the mean precipitation in summer is low. The mean temperature is relatively low in Colorado, Wyoming, and Montana along the Rocky Mountains. Areas with low temperature (see Fig. 3b) and low precipitation indeed generally agree with areas with higher soil moisture persistence. In winter, the lower persistence of soil moisture anomalies in California (not shown) is consistent with the high normal precipitation.

The spatial pattern of the soil moisture persistence from this model is similar to that obtained from the output of an NCEP model multiyear integration (HD

1993). This is an encouraging independent check for the ability of our model because in the NCEP model soil moisture is interactive with several other physical processes, radiation and precipitation are calculated, and the formulation of the soil moisture model is different from ours.

### 3) Spatial size of soil moisture anomalies

It is well known that the scale of temperature anomalies is usually much larger than that of precipitation. Precipitation can be very spotty and intermittent. We now need to characterize soil moisture in terms of the typical spatial size of anomalies, and a concise way is to calculate the spatial degrees of freedom (SDOF). Following exactly the method of Van den Dool et al. (1986b), we calculate, for month $m$, the pattern correlation (see definition A1 in the appendix) of soil moisture anomalies (at the last day of month $m$) for all possible pairs of nonmatching years in the 63-yr period (1931–1993). Leaving out the doubles there are 1953 [or $(63 \times 62)/2$] such correlations, which have, as they should, zero mean,[2] and a standard deviation (SD) that is further assumed to obey

$$\mathrm{SD} = \frac{1}{(\mathrm{SDOF} - 2)^{1/2}}. \qquad (5)$$

From our calculated SD, we can thus solve for SDOF. The results are shown in Fig. 4. For comparison the SDOF for monthly climate division temperature and precipitation are shown as well. It is clear that the soil moisture anomalies are of a spatial scale somewhat larger than precipitation, particularly in summer, but always considerably smaller than temperature. All three variables have the largest scale (low SDOF) in winter and have the smallest scales (high SDOF) in summer. For further interpretation, it should be noted that SDOF is roughly equal to the number of EOFs needed to explain 90% of the combined spatial–time variance.

Compared to Van den Dool et al. (1986b), the estimate for SDOF for precipitation and temperature is similar but numerically lower because here we use smoother climate division data, while station data was used in the previous study. With station data, the SDOF for precipitation in summer is approaching the number of stations used. This very undesirable property disappears when using climate division data.

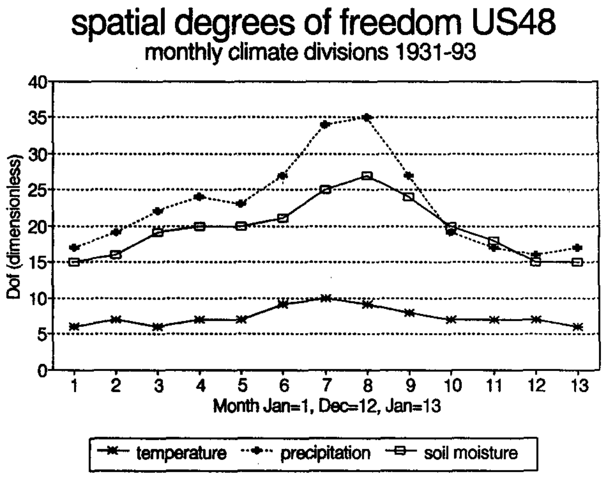


Fig. 4. Spatial degrees of freedom as a function of months in 48 U.S. continental states for monthly mean temperature (thick solid line), precipitation (dashed line), and soil moisture (thin solid line). The data are from 344 U.S. climate divisions during 1931–1993. The soil moisture is calculated from our model.

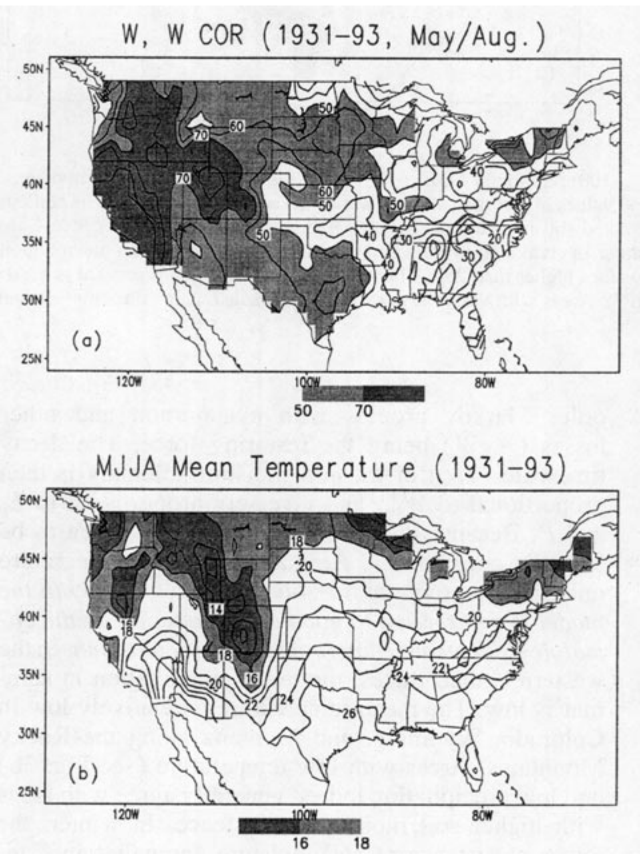


Fig. 3. U.S. maps of (a) soil moisture lagged correlation (%) between May and August with heavier shading for values greater 70% and lighter shading for values greater than 50%; (b) observed monthly mean temperature (°C). Areas with heavier shading for values lower than 16°C and lighter shading for values lower than 18°C are shaded and the interval is 2°C.

[2] The mean correlation ro ($ja, jb$) for month $m$ (averaged over all possible nonmatching years $ja, jb$) is only zero if the anomalies are expressed relative to a "climatological" mean not including year $ja$; that is, the mean needs to be recalculated.

4) LOCAL WATER RECYCLING?

The results in section 3a.2 indicated that the persistence of the soil moisture anomalies is enhanced by the low normal precipitation and low normal evaporation. The lifetime of the soil moisture anomalies has been speculated to be influenced by the so-called local water recycling. Several authors (Entakhahi et al. 1992) have argued that if precipitation is partly or largely due to local evaporation, drought or flood can perpetuate itself.

We here take an ''anomaly point of view'' to discuss this issue. From Figs. 2c and 2d we have found that the standard deviation of evaporation is smaller than that of precipitation by a factor of 2 to 3. Figure 5 shows two randomly picked time series of soil moisture ($\hat{W}$), evaporation ($\hat{E}$), and precipitation anomalies ($\hat{P}$) during 1930–1940 at one climate division in Colorado (Fig. 5a) and during 1960–1970 at one climate division in Illinois (Fig. 5b). It is clear that $\hat{W}$ basically responds to $\hat{P}$ but has much longer persistence because $\hat{W}$ is the integral of $\hat{P}$. We believe that $\hat{W}$ is mainly driven by $\hat{P}$ because (i) the magnitude of $\hat{P}$ is much larger than $\hat{E}$, and (ii) $\hat{E}$ is a restoring force, that is, it is very strongly positively correlated to $\hat{W}$ (the U.S. pattern correlation averaged in summer is about 0.88), thus driving $W$ back to climatology (in agreement with 4b). If $\hat{E}$ was returned to the soil in the form of $\hat{P}$, as assumed in recycling, $\hat{E}$ would not be the restoring force it is observed to be. Although our calculations do not address recycling directly, there is little evidence to support the suggestion that $\hat{P}$ is determined, in a substantial way and locally, by $\hat{E}$.

Entekhabi et al. (1992) showed a bimodal probability distribution of soil moisture obtained from a stochastic water balance model to support the recycling issue. Taking a bimodal distribution of $\hat{W}$ as a sure sign of recycling, we searched our soil moisture datasets for this evidence at randomly selected states in the interior continent. We did not find obvious bimodality in the frequency distribution of soil moisture anomalies during summer 1984–1991 from either the observed soil moisture in Illinois (shown in Fig. 6a) or the calculated soil moisture in Illinois and Oklahoma (shown in Figs. 6b and 6c). This is not to say that recycling does not exist or that if it exists it is not important. After all, $\hat{E}$ is small and yet plays a large role in determining the $\hat{W}$ timescale. Likewise, recycling, even if small, could enhance the $\hat{W}$ timescale. The broad shoulders of the frequency distributions such as in Fig. 6 could be argued to be the result of superposing a unimodal and a bimodal distributions.

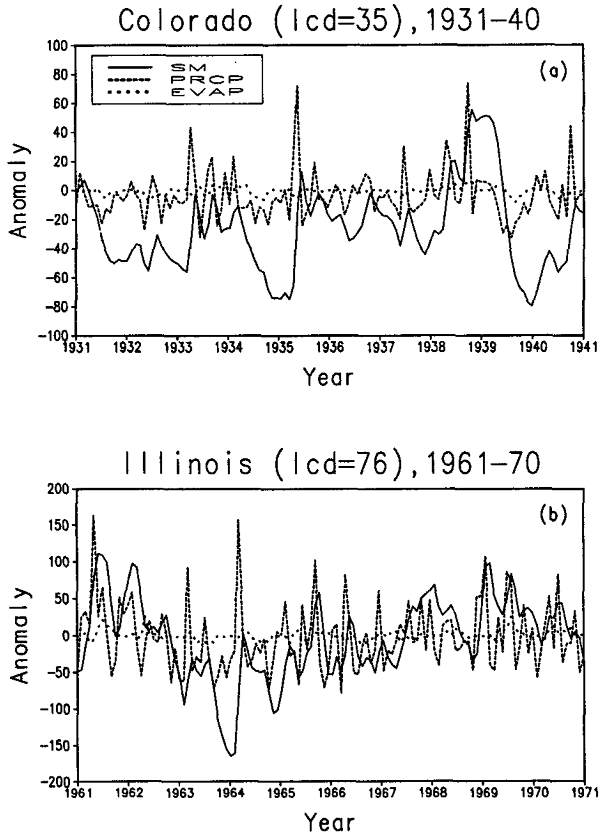


FIG. 5. Time series of anomalies of precipitation (dashed line), evaporation (dotted line), and soil moisture (solid line) at one climate division in (a) Colorado for all months during 1931–1940; (b) Illinois for all months during 1961–1970.

*b. Applications to long-range temperature forecasts*

1) TEMPERATURE PREDICTION

Our previous results (HD 1993) indicate that antecedent precipitation helps in forecasting temperature in large parts of the interior United States. In this study, we will compare the predictive skill of soil moisture to that of precipitation for the temperature forecasts. To this end, we calculate the lagged temporal correlations for soil moisture–temperature ($W$–$T$) and for precipitation–temperature ($P$–$T$) at the 344 U.S. climate divisions during the period of 1931–1993 with lags from 1 to 6 months.

Figure 7 shows the 63-yr-averaged pattern correlations (see definition A2 in the appendix) between precipitation and temperature ($P$–$T$: upper panel) and soil moisture and temperature ($W$–$T$: lower panel) as a function of the predictand month and lags in the interior United States. It is found that the $W$–$T$ correlation is always higher (i.e., more negative) than the $P$–$T$ correlation for any month and lag, particularly in spring–summer. Therefore, soil moisture is a better (local) predictor than precipitation for temperature forecasts. Similar to the $P$–$T$ correlation, the $W$–$T$ correlation is the highest when a warm month's temperature is the target. The best month in the year to predict $T$ from antecedent soil moisture is July, and rather importantly this can be done with a considerable lead. The temperature in all months from May through September can be estimated by soil moisture conditions several months earlier. The $W-T$ correlations do have some strange variations with August being less predictable than July and September for no obvious reason. We have also studied the $W-T$ correlations with lags up to 2 years (not shown) and found significant correlation out to 2 years. This perhaps explains the summer temperature year to year persistence noted a long time ago by Namias (1952), Madden and Shea (1978), and Van den Dool et al. (1986b).

The lagged temporal $W-T$ and $P-T$ correlations for May–July (shown in Fig. 8) are taken as a good example to show the spatial distribution. The $W-T$ correlation is significantly negative in most of the U.S. inland and is more negative than $P-T$ correlation almost everywhere. The appearance of the $P-T$ and $W-T$ maps is very similar, the $W-T$ values just being larger. It appears that the soil model has filtered the $P$ time series into a better predictor ($W$). To some degree the $W-T$ correlation is highly negative in places where $W$ anomalies have high persistence (see Fig. 3a), such as in Utah, Colorado, and Nebraska. The $W-T$ correlation from our model has a spatial pattern very similar to that from the output of the MRF multiyear integration (see Fig. 3c in HD 1993).

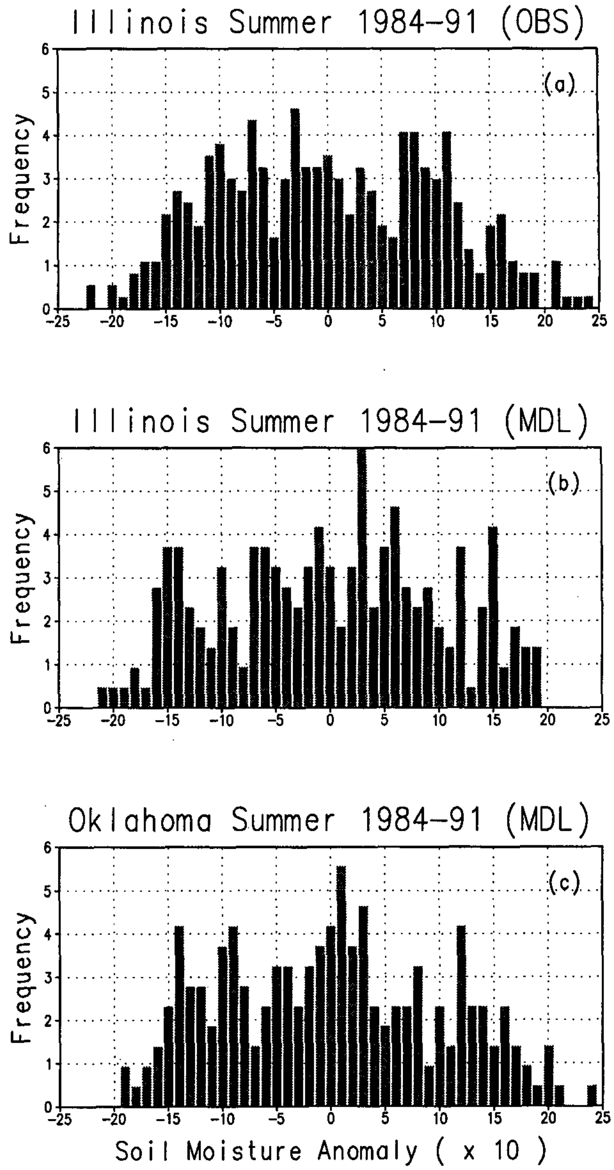


FIG. 6. Frequency distribution (%) of soil moisture anomalies during June to August 1984–1991: (a) observed in Illinois; (b) calculated in Illinois; (c) calculated in Oklahoma.

#### 2) CHOOSING THE PREDICTOR(S)

In the operational long-range temperature forecasts, temperature persistence is a common tool. We have investigated whether use of soil moisture as another predictor increases the forecast skill in the temperature forecast when previous temperature is used as a predictor as well. The degree of the success depends on 1) how much of the temperature persistence is caused by soil moisture and 2) whether antecedent temperature is perhaps a better proxy for the true state of soil moisture than is our calculated soil moisture. Our investigation has been made through a multiple regression approach. The choice of predictors is soil moisture, temperature, or both, depending on their predictive skill in the past (based on cross-validation, see HD 1993). Precipitation was not admitted as a predictor. We found that soil moisture can provide extra skill in predicting future temperature in a large area of the interior continent in warm months, especially at longer lags (such as May → July shown in Fig. 9; "no skill" is defined by the cross-validation correlation being lower than 0.2). The spatial pattern, the choice of predictors, and the level of skill depend on month and lag.

Ideally and in a real world with perfect measurements, we expect evaporation (say, averaged over the last 10 days of a month) to be a better predictor (for future $T$) than soil moisture for short lags, because evaporation represents both the atmosphere and soil moisture and affects the future temperature in a more direct way. Applied to (monthly) climate divisions, however, the potential evaporation is set to be constant in a month and the actual evaporation varies with soil moisture only. With the present datasets at hand, $E$ is thus not better than $W$ for predicting future $T$.

In the numerical experiments (e.g., Mintz 1984) and probably also in the real world, soil moisture has some impact on future precipitation by interacting with the large-scale circulation and also by supplying some of the moisture for recycling. However, the correlation between the calculated soil moisture and future precipitation has been found to be very small at any lead (slightly higher than precipitation persistence). At this

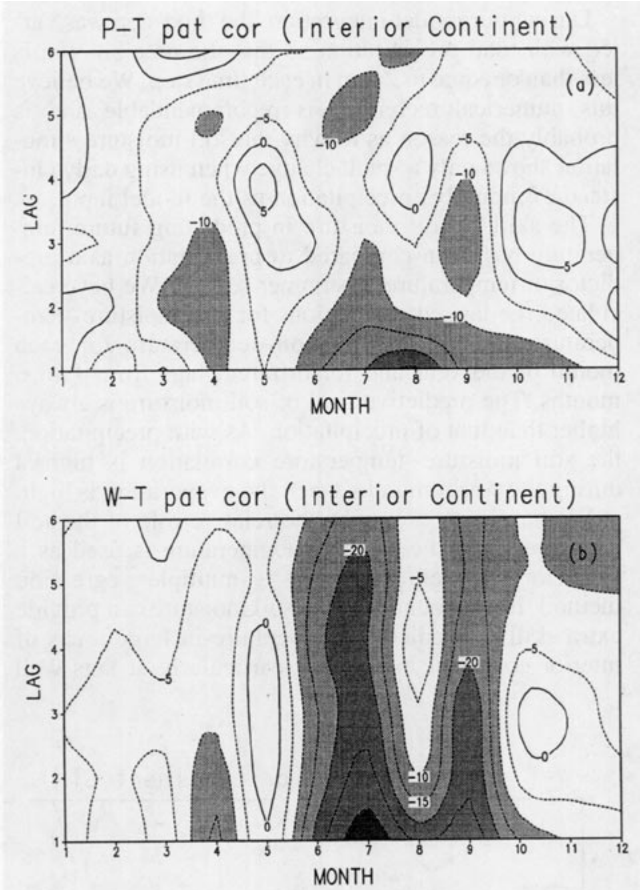


FIG. 7. Pattern correlation (%) over the interior United States (all climate divisions with 30° ≤ lat ≤ 45° and 80° ≤ long ≤ 115°) at lags from 1 to 6 months as a function of season (a) $P-T$ correlation and (b) $W-T$ correlation. The $x$ axis is predictand's month and $y$ axis is lead; for example, lead = 1 at month = 7 is for June–July case. The contour interval is 5. The light shading is for −10% to −20%, the medium shading is for −20% to −30%, and the darkest shading is for −30% or less.

point, there is no hope for improving a regression-type precipitation forecast by using antecedent soil moisture.

### 3) COMPARISON BETWEEN $T_{max}$ AND $T$ AS THE PREDICTAND

In most long-range temperature forecasts, monthly or seasonal-averaged *daily mean* temperature ($T$) is the predictand. When the soil moisture is the predictor, however, the daily mean $T$ may not be the best predictand. The daily maximum temperature ($T_{max}$) may be a better target because the soil moisture affects the surface temperature through evapotranspiration, which mainly occurs during the daytime when the temperature reaches its maximum.

To test this idea, we have compared the lagged correlations between soil moisture and monthly mean temperature ($W-T$) to the correlation between soil moisture and monthly mean daily maximum temperature ($W-T_{max}$). The monthly data was obtained from daily cooperative station data because we, unfortunately, do not have $T_{max}$ at our disposal for climate divisions. Figure 10 shows the pattern correlation in the interior continent for both $W-T$ and $W-T_{max}$ correlations. Note first that the correlations in Fig. 10a are quite a bit less than in Fig. 7b because of using station data. It can also be seen that the $W-T_{max}$ correlation is indeed always more negative than $W-T$ correlation, hence increasing the physical realism of our calculated soil moisture data.

## 4. Summary

A historical (1931–1993) soil moisture dataset at the 344 U.S. climate divisions has been calculated by a one-layer soil moisture model. The model is based on the water budget in the soil. The model uses the monthly mean temperature and monthly total precipitation as input to calculate soil moisture, surface runoff, evapotranspiration, baseflow, and groundwater loss to deep layers. In this model, potential evapotranspiration is calculated using observed surface air temperature

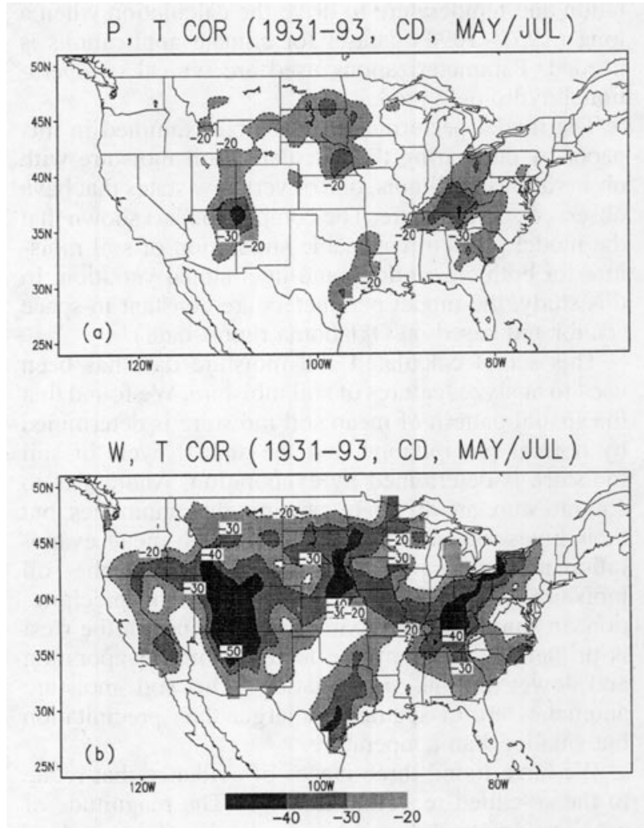


FIG. 8. Temporal correlation (%) over the 344 U.S. climate divisions (a) between May precipitation and July temperature; (b) between soil moisture of 31 May and July temperature. The data are 1931–1993. Solid lines are for positive correlation and dashed lines for negative correlation. The contour interval is 10%.

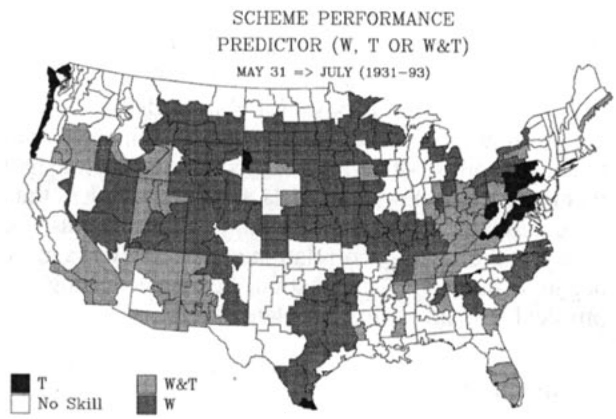


FIG. 9. Geographical map of the winner of the three regression schemes based on cross validation using 1931–1993 data for May–July case. In medium-shaded area, soil moisture is the best predictor for future temperature. In lighter-shaded area, combination of soil moisture and temperature is the best predictor. In the dark-shaded area, the temperature persistence is the best.

and duration of sunlight (astronomical). The model is a compromise between physical complexity and the practicality of having to rely on just observed precipitation and temperature to drive the calculation when a long (1931–1993) dataset for climate applications is needed. Parameterizations used are typical of operational hydrologic models.

The model performance has been examined in this paper by comparing the calculated soil moisture with observations in Illinois, one of very few states that have observed soil moisture. The comparison has shown that the model gives a reasonable simulation of soil moisture for both climatology and interannual variation. In this study, the model parameters are constant in space (calibrated based on Oklahoma runoff data).

This set of calculated soil moisture data has been used to analyze features of soil moisture. We found that the spatial pattern of mean soil moisture is determined by normal precipitation, but the annual cycle of soil moisture is determined by evaporation. Anomalies in soil moisture are driven by precipitation anomalies, but their timescales are determined by both mean evaporation and mean precipitation. As expected, the soil moisture has much longer persistence than precipitation. In summer, the maximum persistence in the west is primarily caused by the lower normal evaporation and lower normal precipitation. The soil moisture anomalies are of spatial size larger than precipitation but smaller than temperature.

We have found three pieces of evidence that relate to the so-called recycling issue. 1) The magnitude of evaporation anomalies is too small to explain much of the precipitation anomalies. 2) Term $\hat{E}$ acts as a restoring force on $\hat{W}$. 3) Neither observed nor calculated soil moisture anomalies have an obviously bimodal frequency distribution.

During our model integration, the time step was varied with total precipitation so that the precipitation is less than or equal to 2 mm in each time step. We believe this numerical technique is recommendable and is probably the reason as to why the soil moisture simulation shows only a small change when using daily (instead of monthly) precipitation as the model input.

The skill of soil moisture in predicting future temperature has been compared to precipitation as a predictor of temperature in summer season. We have calculated the lagged correlations for soil moisture–temperature and for precipitation–temperature for each month of the year and for different lags from 1 to 6 months. The predictive skill of soil moisture is always higher than that of precipitation. As with precipitation, the soil moisture–temperature correlation is highest during the warm months when the evaporation is high.

We have also examined the relative role of the soil moisture when the previous temperature is used as a predictor as well by using a multiple regression method. It was found that the soil moisture can provide extra skill in predicting temperature in large areas of interior continent in summer particularly at lags well

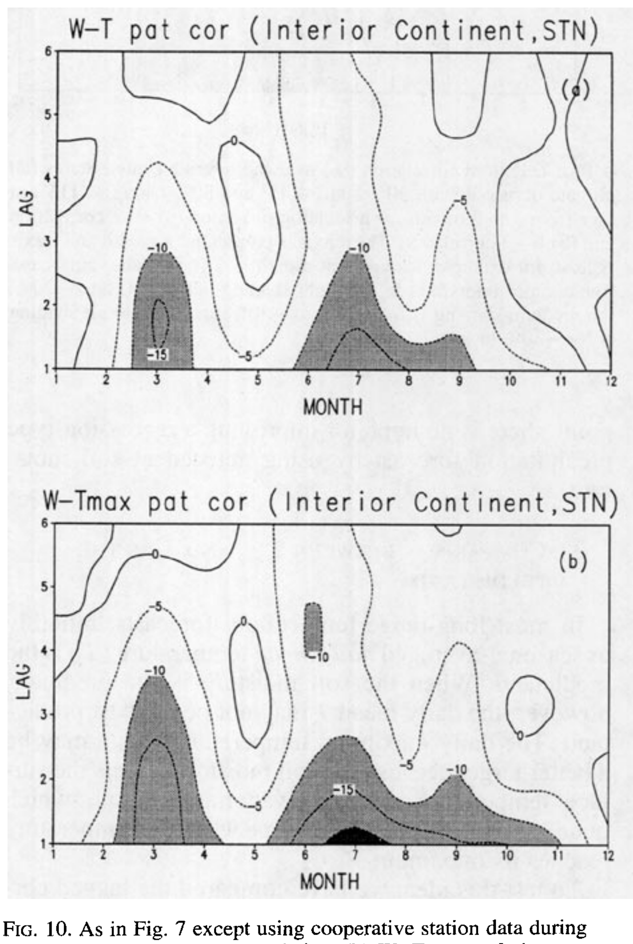


FIG. 10. As in Fig. 7 except using cooperative station data during 1931–1991: (a) $W$–$T$ correlation; (b) $W$–$T_{max}$ correlation.

beyond one month. Furthermore, the predictive skill of soil moisture is higher when the predictand is daily maximum temperature instead of daily mean temperature.

In the future, we want to apply this dataset in the operational long-range temperature forecasts at CPC/NCEP, which would require the monthly climate division data to be updated in "real time" for such use. Ideally we would like to use *daily climate division* data since 1) $W-T$ correlation is higher at climate divisions (which is a spatial averaging of stations) than that at stations, 2) the monthly mean daily maximum temperature is a better predictand, and 3) for the nonlinearity of precipitation it is more reasonable to use daily precipitation instead of monthly mean precipitation to drive the soil-moisture model.

In the future, we will determine model parameters for various U.S. regions of similar soil and land cover based on the regional runoff data and then use these parameters for each region. Moreover, we hope to get more observed soil moisture data to verify our model results. The observed evaporation, if available, could be used in verification of the model as well.

We suggest in no way that any of the issues raised are definitively settled. Results of using soil moisture in temperature forecasts may keep improving as the soil hydrology model improves. A severe limitation is that we have used only observed temperature and precipitation to drive the model. A possibly big step forward will be the so-called reanalysis of soil moisture in a soil hydrology model coupled interactively to an atmospheric model. The observations driving such a system could then include global pressure, temperature, humidity, and wind in the free atmosphere. NCEP has an ambitious plan for reanalysis for the period of 1958–present (Kalnay et al. 1991). While radiation and clouds were not observed directly for 1931–1993, they would be calculated from the atmosphere model equations through physical parameterizations. The present soil moisture dataset would be a benchmark upon which to improve by more ambitious reanalysis.

*Acknowledgments.* The first author wishes to express her thanks to Dr. D. Cayan in Scripps Institute of Oceanography for inviting her for the visit and helpful discussions. The work of the third author was supported by the NOAA Experimental Climate Forecast Center of the Scripps Institution of Oceanography. Thanks are also due to Dr. J. Schemm for helpful comments, Dr. S. E. Hollinger for supplying observed soil moisture in Illinois, and Dr. S. Saha for her help in preparing some of the maps. This work was supported in part by the National Oceanic and Atmospheric Administration Office for Climate and Global Change under the Atmospheric and Land Surface Process (ALSP) element.

## APPENDIX

### Definition of a Pattern Correlation Coefficient

Let $T(s, m, j)$ stand for monthly mean temperature at station or climate division $s$ ($s = 1, N$), month $m$, and year $j$ ($j = 1, M$). The definition of a *pattern correlation* between year $ja$ and year $jb$ for month $m$ is expressed as

$$PC(ja, jb) = \frac{\sum_{s=1}^{N} \hat{T}(s, m, ja)\hat{T}(s, m, jb)}{[\sum_{s=1}^{N} \hat{T}^2(s, m, ja)\hat{T}^2(s, m, jb)]^{1/2}}. \tag{A1}$$

The $PC(ja, jb)$ measures the degree of similarity of standardized temperature anomaly patterns over the coterminous United States between year $ja$ and year $jb$ for month $m$,
where

$$\hat{T}(s, m, j) = \frac{T(s, m, j) - \bar{T}(s, m)}{sd(s, m)}$$

$$\bar{T}(s, m) = \frac{1}{M}\sum_{j=1}^{M} T(s, m, j)$$

$$sd(s, m) = \sqrt{\frac{1}{M}\sum_{j=1}^{M} (T(s, m, j) - \bar{T}(s, m))^2}.$$

Similarly, a *pattern correlation* between *two months separated* by $\tau$ and summed over all $M$ years is defined as

$$PC(m, \tau) = \frac{1}{NM}\sum_{s=1}^{N}\sum_{j=1}^{M} \hat{T}(s, m, j)\hat{T}(s, m + \tau, j) \tag{A2}$$

(or summing over part of the country $s = n1, n2$; $n1 \geq 1, n2 \leq N$).

## REFERENCES

Cayan, D. R., and K. P. Georgakakos, 1995: Hydroclimatology of continental watersheds: spatial analyses. *Water Resour. Res.*, **31**(3), 677–697.

Entekhabi, D., I. Rodriguez-Iturbe, and R. L. Bras, 1992: Variability in large-scale water balance with land surface-atmosphere interaction. *J. Climate*, **5**, 798–813.

Georgakakos, K. P., 1986: A generalized stochastic hydrometeorological model for flood and flash-flood forecasting, 1: Formulation. *Water Resour. Res.*, **22**(13), 2083–2095.

——, D.-H. Bae, and D. R. Cayan, 1995: Hydroclimatology of continental watersheds: Temporal analyses. *Water Resour. Res.*, **31**(3), 655–675.

Hollinger, S. E., and S. A. Isard, 1994: A soil moisture climatology of Illinois. *J. Climate*, **7**, 822–833.

Huang, J., and H. M. van den Dool, 1993: Monthly precipitation–temperature relations and temperature prediction over the United States. *J. Climate,* **6,** 1111–1132.

Janowiak, J. E., C. F. Ropelewski, and M. S. Halpert, 1986: The precipitation anomaly classification: A method for monitoring regional precipitation deficiency and excess on a global scale. *J. Climate Appl. Meteor.,* **25,** 565–574.

Kalnay, E., and R. Jenne, 1991: Summary of the NMC/NCAR reanalysis workshop of April 1991. *Bull. Amer. Meteror. Soc.,* **72,** 1897–1904.

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

Klein, W. H., 1983: Objective specification of monthly mean surface temperature from mean 700-mb heights in winter. *Mon. Wea. Rev.,* **111,** 674–691.

Lee, D. M., 1980: On monitoring rainfall deficiencies in semidesert regions. *The Threatened Drylands—Regional and Systematic Studies of Desertification,* J. A. Mabbutt and S. M. Berkowics, Eds., Fujinomiya, Japan, 953 pp. [Available from Australian Bureau of Meteorology, Melbourne, Australia.]

Lyons, S. W., 1990: Spatial and temporal variability of monthly precipitation in Texas. *Mon. Wea. Rev.,* **118,** 2634–2648.

Madden, R. A., and D. J. Shea, 1978: Estimates of natural variability of time-averaged temperatures over the United States. *Mon. Wea. Rev.,* **106,** 1695–1703.

Mintz, Y., 1984: The sensitivity of numerically simulated climates to land-surface boundary conditions. *The Global Climate,* J. T. Houghton, Ed., Cambridge University Press, 79–105.

——, and Y. V. Serafini, 1981: Global fields of soil moisture and surface evapotranspiration. NASA/Goddard Flight Center Tech. Memo. 83 907, Research Review—1980/81, 178–180.

——, and G. K. Walker, 1993: Global fields of soil moisture and land surface evapotranspiration derived from observed precipitation and surface air temperature. *J. Appl. Meteor.,* **32,** 1305–1334.

Namias, J., 1952: The annual course of month-to-month persistence in climatic anomalies. *Bull. Amer. Meteror. Soc.,* **33,** 279–285.

——, 1962: Influences of abnormal heat sources and sinks on atmospheric behavior. *Proc. Int. Symp. on Numerical Weather Prediction,* Tokyo, Meteor. Soc. Japan, 615–627.

Palmer, W. C., 1965: *Meteorological Drought.* Res. Paper 45, U.S. Weather Bureau, 58 pp. [NOAA Library and Information Services Division, Washington, D.C. 20852].

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1989: *Numerical Recipes, The Art of Scientific Computing (Fortran Version),* Cambridge University Press, 702 pp.

Rajaram, H., and K. P. Georgakakos, 1989: Recursive parameter estimation of hydrologic models. *Water Resour. Res.,* **25**(2), 281–294.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher, 1986: A simple biosphere model (SIB) for use within general circulation models. *J. Atmos. Sci.,* **43,** 505–531.

Sellers, W. D., 1965: *Physical Climatology.* University of Chicago Press, 272 pp.

Sorooshian, S., 1991: Parameter estimation, model identification, and model validation, Conceptual-type models. *Recent Advances in the Modeling of Hydrologic Systems,* D. S. Bowles and P. E. O'Connell, Eds., Kluwer, 443–467.

Thornthwaite, C. W., 1948: An approach toward a rational classification of climate. *Geogr. Rev.,* **38,** 55–94.

Van den Dool, H. M., 1988: A utilitarian atlas of monthly and seasonal precipitation–temperature relations, the United States, 1931–1987. University of Maryland SR-88-17, 40 pp.

——, 1989: Monthly precipitation-temperature relations, the United States, 1931–1987. *Proc. 13th Annual Climate Diagnostics Workshop,* Cambridge, MA, Climate Analysis Center, 210–213.

——, J. E. Walsh, R. Yang, and W. H. Klein, 1986a: Feedbacks on soil moisture anomalies: A mechanisms to perpetuate air temperature anomalies in hot continental areas? *Proc. 11th Annual Climate Diagnostics Workshop,* Champaign, Illinois State Water Survey, 233–242.

——, W. H. Klein, and J. E. Walsh, 1986b: The geographical distribution and seasonality of persistence in monthly mean air temperatures over the United States. *Mon. Wea. Rev.,* **114,** 546–560.

Walsh, J. E., W. H. Jasperson, and B. Ross, 1985: Influences of snow core and soil moisture on monthly air temperature. *Mon. Wea. Rev.,* **113,** 756–768.

Yeh, T.-C., 1989: Sensitivity of climate model to hydrology. *Understanding Climate Change,* A. Berger, R. E. Dickinson, and J. W. Kidson, Eds., Amer. Geophys. Union, 101–108.