Role of Terrestrial Hydrologic Memory in Modulating ENSO Impacts in North America

JI CHEN* AND PRAVEEN KUMAR

Environmental Hydrology and Hydraulic Engineering, Department of Civil and Environmental Engineering, University of Illinois at Urbana–Champaign, Urbana, Illinois

(Manuscript received 23 May 2001, in final form 20 May 2002)

ABSTRACT

Relationships among the terrestrial hydrologic processes over the North American continent associated with the El Niño–Southern Oscillation (ENSO) are investigated using a large-area basin-scale land surface model driven by the European Centre for Medium-Range Weather Forecasts Re-Analyses 15-yr (1979–93) dataset. The modeling approach allows for the study of the relationships of ENSO with several hydrologic variables simultaneously, such as soil water storage, basin runoff, snow-water equivalent, and precipitation. The cross-correlation coefficients between terrestrial variables and the ENSO index are computed. The runoff from the northern part of North America was found to be most often negatively correlated with ENSO, and there are four distinct coherent regions over the continent where the runoff anomalies are positively correlated. The terrestrial systems have a delayed response to the ENSO signal, as compared to the precipitation, and the delay may range from a month to a season or longer. The shorter and longer delays are typically associated with rainfall runoff, and snow accumulation and melt processes, respectively. The soil moisture storage plays a very vital role in delaying the effects of the climate variability on the terrestrial hydrologic processes and in extending the influences of the El Niño or La Niña events on the terrestrial climate.

1. Introduction

Interannual climate variabilities, such as the El Niño–Southern Oscillation (ENSO), have important impacts on various terrestrial systems (Glantz 2001) such as streamflow (e.g., Kiladis and Sinha 1991), ecology (e.g., Barber and Chavez 1983), fish habitat (e.g., White and Downton 1991), etc. The response of the hydrologic system serves to transmit the climate signal to several of these terrestrial systems. Consequently, our ability to model and predict the terrestrial hydrologic response to the climatic fluctuations would provide significant advancement to this area of research.

The phenomena and consequences of the warm phase [El Niño with its anomalously warm sea surface temperature (SST) over the eastern equatorial Pacific] and the cool phase (La Niña with its anomalously cool SST) of the ENSO signal are well studied and documented during the last few decades. The Pacific jet stream is displaced north of the normal location during a La Niña period but south of the normal location during El Niño (Trenberth and Guillemot 1996; Green et al. 1997), and the spatial distribution of the storm track over the North American continent closely relates to the jet stream location. Using observed temperature and precipitation from 1944 to 1996, Green et al. (1997) studied the North American climate patterns associated with ENSO. For instance, in the warm event winter of ENSO, positive precipitation anomalies occur in the area around the Gulf of Alaska, Oregon, and northern California, and the area around the Gulf of Mexico; meanwhile, negative anomalies occur in Washington, Indiana, and southern California. In the cold event winter, positive precipitation anomalies appear over Washington, Oregon, and the southwestern coast of Canada, and the western part of the southeast United States, and negative anomalies appear on the Gulf of Alaska, California, and the U.S. southeast (Green et al. 1997). Even though, in some areas, the ENSO–climate relationship is asymmetric between El Niño and La Niña events, and the relationships have varied in strength from one episode to the next, it is found that most of these significant areas in North America have symmetric effects for opposite phases of ENSO (Green et al. 1997).

Understanding the distinct relationships among the various systems influenced by ENSO (e.g., Horel and Wallace 1981; Wright et al. 1988; Trenberth and Guil-
lemot 1996; Liang et al. 1997) has enabled us to explain the anomalous atmospheric circulation and hydrologic fluctuations. However, the relationship between the ENSO extreme and terrestrial hydrology is not as comprehensively investigated, even though there are quite a few recordings of the consequences of El Niño and La Niña on terrestrial hydrology (Glantz 2001). Using daily precipitation and streamflow from late winter to early summer at several hundred sites in the western United States, Cayan et al. (1999) studied the relationship between ENSO and hydrologic extremes and found that streamflow responses to ENSO extremes are accentuated over precipitation responses. Applying the knowledge of the relationships among the intertropical convergence zone (ITCZ) and the Pacific jet stream associated with the sea surface temperature (SST) anomalies in the tropical Pacific Ocean, Trenberth and Guillemot (1996) concluded that the 1988 summer drought was caused by the strongest La Niña event in over 20 yr and the 1993 wettest summer was due to the mature El Niño conditions. The drought and flood episodes are the anomalous phases of the soil moisture storage and the streamflow. But little research has been done on the relationships between terrestrial hydrologic processes and ENSO at the regional to continental scales.

In this paper we present a modeling study [using a large-area basin-scale (LABs) model (Chen and Kumar 2001)] to understand the response of the terrestrial hydrologic variables to ENSO. We use the results from a 15-yr model simulation (1979–93) and compute the cross correlation of the ENSO index with the anomalies of the model output and precipitation. The areas strongly affected by ENSO extremes are found to be consistent with other studies (e.g., Green et al. 1997; Cayan et al. 1999). We show that the ENSO signals are modulated by the memory of the terrestrial hydrologic processes.

This paper is organized as follows. In section 2, we give a brief review of the LABs model, and describe the parameters that are required. The datasets used in the study are also described in this section. The methodology and results are presented in sections 3 and 4, respectively, and summary and conclusions are given in section 5.

2. Background

The LABs model implements the vertical moisture and energy transport schemes as in the Surface–Vegetation–Atmosphere Transfer (SVAT) schemes (e.g., Sellers et al. 1986; Dickinson et al. 1986; Abramopoulos et al. 1988; Pitman et al. 1991; Bonan 1996), and parameterizes the spatial heterogeneity of soil moisture using the probability distribution of topographic index. It predicts terrestrial components, such as streamflow, water table depth, soil moisture deficit, along with the moisture and energy fluxes at the land–atmosphere interface typical of the SVAT schemes. The LABs model has already been implemented for basins over the entire North American continent with the International Satellite Land Surface Climatology Project (ISLSCP) 2-yr dataset and was validated using the streamflow observations over the Mississippi River and its subbasins (Chen and Kumar 2001). The basins are delineated using the HYDRO1K dataset (Verdin and Verdin 1999) developed by the United States Geological Survey (USGS). In the present study, the model is driven using the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis 15-yr (ERA-15) dataset (Gibson et al. 1999) for the period of 1979 to 1993. The ERA-15 and observed precipitation from the National Centers for Environmental Prediction (NCEP; Higgins et al. 1996a); the observed streamflow discharge from USGS; and the anomalies of model-predicted streamflow, snow-water equivalent (SWE), and soil water deficit are correlated with the ENSO signal. Regions that show distinct correlations and the associated lags are identified.

a. Model description

The LABs model (Chen and Kumar 2001) provides the ability to model various components of terrestrial hydrology at the basin scale. The topographic influences on large-scale terrestrial hydrologic processes are explicitly taken into account in the model by integrating the TOPMODEL framework (Beven and Kirkby 1979; Beven et al. 1995) with a SVAT scheme. The SVAT component of LABs is adapted from the National Center for Atmospheric Research (NCAR) Land Surface Model (LSM) (Bonan 1996).

The essential scheme of modeling the moisture dynamics of LABs is shown in Fig. 1. Each basin is represented as a single column with six soil layers for computing the vertical soil moisture and energy transport. To simulate the dynamics of partial-variable source area of a basin, a wetness index, that is, the topographic index (Beven and Kirkby 1979), is used. The topographic index at any location in a basin is defined as $x = \ln(a/\tan \beta)$, where $a$ is the upstream contributing area, from the watershed divide, per unit contour length, and $\tan \beta$ is the local slope. The relationship between the basin average water table depth $z_T$ and the local water table depth $z$ (positive downwards from the surface) at any location within the basin is given as

$$z = z_T - \frac{\zeta}{5}(x - \lambda),$$

where $\zeta$, the soil decay factor, is a hydrogeologic constant and is a measure of the decline of the saturated hydraulic conductivity with increasing soil depth. Namely, vertical saturated hydraulic conductivity $k_v(z)$ and lateral saturated hydraulic conductivity $k_l(z)$ exponentially decline as $e^{-\zeta z}$. Here, $k_v(z)$ is used in the computation of the vertical moisture transport through the different layers, and $k_l(z)$ is used in the computation of the baseflow due to lateral moisture transport. It is
assumed that $k_s(z) = \alpha k_s(z)$, where $\alpha$ is an anisotropic ratio. The parameter $\lambda$ is the mean value of the topographic indices over the basin. Therefore, the spatial heterogeneity of soil moisture is modeled through the probability distribution of the topographic index, $f(x)$, often parameterized as a three-parameter gamma distribution (Sivapalan et al. 1987; Kumar et al. 2000). For any particular value of $z$, knowledge of the pattern of topographic index allows prediction of those areas where $z < 0$, that is, the saturated contributing area, $A_c$, where all rainfall contributes to runoff without any infiltration loss. The contributing area for direct runoff, $A_c$, in a basin of area $A$, is given as

$$A_c = \int_{x(z+\lambda)} f(x) \, dx.$$  \hspace{1cm} (2)

During rainfall, water table rise results in a larger fraction of the area contributing to direct runoff.

The streamflow consists of the surface runoff and subsurface runoff. The subsurface runoff, $Q_s$, is computed as a function of the mean water table depth, given as (Sivapalan et al. 1987; Beven et al. 1995)

$$Q_s = AT_0 e^{-\mu z_t},$$  \hspace{1cm} (3)

where $T_0 = k_s(0)/\zeta$ is the basin lateral transmissivity when the soil is just saturated, and $k_s(0)$ is the saturated lateral hydraulic conductivity of the soil at the surface.

To incorporate the heterogeneity of land use in each basin, four land use categories along with a category for water body are implemented. Each of the four land use categories can be either in the saturated or the unsaturated region. We, therefore, effectively have nine possible subregions in each basin consisting of four land use categories in both the saturated and unsaturated areas, and a water body. This results in different runoff and infiltration rates corresponding to different subregions. This enables us to model the subregion variabilities in water and energy fluxes, since evaporation and transpiration are at potential rates over the saturated subregions and at moisture-controlled rates over the unsaturated ones. Saturation and infiltration excess overland flows occur in the saturated and the unsaturated areas, respectively. The aggregated infiltration is used in the solution of moisture transport through the vertical column. Furthermore, the four land use categories in both unsaturated and saturated regions have the various properties of vegetation phenology, providing variability in the interception of precipitation and in the radiation processes in the basin.

The moisture transport in the single soil column is
described by the Richards equation. For each soil layer the discrete form of the equation is

\[
\frac{\Delta \bar{h}(i) \Delta z(i)}{\Delta t} = -\overline{q}_w(i) + \overline{q}_{sw}(i) - \overline{e}(i),
\]

where \( \bar{h}(i) \) is the \( i \)th-layer soil water content for the entire basin, and \( \Delta t \) is obtained as area-weighted averages of each subregion. The values \( \overline{q}_w(i) \) and \( \overline{q}_{sw}(i) \) are the averaged quantities of input and output water fluxes (positive in the upward direction) from layer \( i \) over the time period of \( \Delta t \); \( \overline{e}(i) \) is the averaged evapotranspiration loss; \( \Delta z(i) \) is the thickness of soil layer \( i \). This equation, with the boundary conditions of infiltration as the flux of water into the soil and the zero flux of water at the bottom of the soil column, is numerically implemented for a six-layer soil column to calculate vertical moisture transport. The model prediction for the soil moisture in each layer is the prediction for the mean state, and this mean state determines the mean water table depth (for details of the soil moisture dynamics in LABs, see Chen and Kumar 2001, and Chen 2001).

The LABs model maintains the NCAR LSM framework and requires the same forcings. The forcings from the atmosphere to drive the model are convective and large-scale precipitation, solar and longwave radiation, air temperature, air pressure, specific humidity, and zonal and meridional wind velocities at a reference height. The outputs from LABs are latent,ensible and ground heat fluxes, zonal and meridional momentum fluxes, longwave and shortwave radiations from land surface, overland and subsurface flows, soil moisture contents of the six soil layers, basin average water table depth, and saturated fraction. Required surface data for each basin is its location, surface type (which defines the vegetation type), soil color type, soil texture, and the first three statistical moments of the distribution of the topographic indices.

b. Model parameters

1) Basin delineation and topographic parameters

The extraction of the hydrographic features, that is, drainage areas and slopes, from the GTOPO30 digital elevation model (DEM) data (Gesch et al. 1999) are based on the drainage analysis algorithm of Jenson and Domingue (1988). This algorithm computes the flow accumulation for each cell using the direction of the steepest descent. A 1000 km\(^2\) accumulation threshold is then used to identify the streams (Verdin and Jenson 1996). The drainage network is then used to delineate basins and subbasins. In order to represent the basin hierarchically, a system developed by Pfafstetter (Verdin and Verdin 1999) is used. Verdin and Verdin (1999) developed descriptions at five hierarchical levels of subdivisions. For this study, level-5 delineations are chosen with the mean basin size of 3255 km\(^2\). A total of 5020 basins are required to cover the continent at this level (for details of basin delineations, please see Verdin and Verdin 1999, and Chen and Kumar 2001).

The topographic index of each basin is computed using the single-flow algorithm (Quinn et al. 1991; Wolock 1993) that utilizes the flow accumulation and the steepest slope at every grid point. The topographic index obtained using the GTOPO30 data captures the general spatial distribution over the North American continent. Using topographic index obtained from 90-m DEM data for several 1° × 1° latitude–longitude grid boxes, it is found that a simple linear relationship between the L-moments obtained at the 1-km and 90-m resolutions can be developed (Kumar et al. 2000). Using this scheme, linearly downscaled, L-moments of topographic indices obtained from the GTOPO30 data, to an equivalent 90-m resolution, for each basin are used to obtain the parameters for the three-parameter gamma distribution describing the probability distribution function of the topographic index. This parameterization is used for the computation of the water table depth for each basin in the North American continent.

2) Soil and vegetation parameters

Soil texture, soil color, and surface type of each basin in North America are obtained by interpolating NCAR LSM’s dataset, which adopts the surface types of Olson et al. (1983), the soil colors of Biosphere–Atmosphere Transfer (BATS) T42 data (Dickinson et al. 1993), and the soil textures of Webb et al. (1993). Each land surface type is comprised of bare ground and/or several vegetation types, which determine land surface moisture and energy fluxes. There are a total of 28 surface types, which include combinations of a subset of 12 vegetation types and bare ground surface. Soil color defines the saturated and dry soil albedos and is categorized into nine classes. Soil texture provides percentages of sand, silt, and clay, which define soil thermal and hydroligic properties.

c. Atmospheric forcing data

The model forcing is derived from the ERA-15 dataset (Gibson et al. 1999). We also investigated a 40-yr (1958–97) dataset from NCEP–NCAR reanalysis project and found that the precipitation is overestimated over some parts of the United States by a factor of almost 2 (Higgins et al. 1996b) and therefore is not used for this study. The ERA-15 dataset is global and available on a 2.5° × 2.5° latitude–longitude grid. The forcing for each basin in North America is obtained using an area-weighted average from all the 2.5° × 2.5° grid boxes that it straddles. The LABs model is driven using four times daily (0000, 0600, 1200, 1800 UTC) ERA-15 forcing obtained from the lowest \( \sigma \) level. The temporal interpolation schemes described in Chen and Ku-
To evaluate the quality of ERA-15 precipitation, we compare the ERA-15 precipitation with the observed precipitation. The observed values are from the 30-yr gridded hourly precipitation data developed by Higgins et al. (1996a), which is the product of a retrospective gridded analysis of the NCEP hourly precipitation gauge data, for the conterminous United States. The dataset covers the period from January 1964 to December 1993, and were gridded into 2° latitude by 2.5° longitude boxes. For our study time period (1979–93), the mean values and standard deviations (std devs) of the observed and ERA-15 precipitations, disaggregated onto the level-5 basins, are computed. Figure 2 shows the normalized mean and standard deviation differences. They are computed as the difference between observation and ERA-15 mean daily precipitation divided by the observed mean daily precipitation, and similar for normalized standard deviation difference. Further, Fig. 2 also shows the correlation between Higgins and ERA-15 daily precipitation datasets. From the figure, it is found that ERA-15 moderately underestimates the precipitation over the lower Mississippi region and the southwestern United States coast region, and overestimates the precipitation in the Rocky Mountains region. Also, this is true for the normalized standard deviation difference. The correlation between Higgins and ERA-15 daily precipitation datasets are generally larger than 0.7. The result of the above comparison indicates that the quality of ERA-15 precipitation is quite good, given that the Higgins dataset is also subjected to various errors. Hence, the ERA dataset was deemed acceptable for our modeling study. A further verification of the quality of the variability of ERA-15 precipitation is carried out in section 4c.

d. Observed ENSO data

We derive the ENSO index from the raw data of National Oceanic and Atmospheric Administration (NOAA) Niño-3 (5°N–5°S, 90°–150°W) SST (CPC 2001) for the base climatologic period of 1979–93. The ENSO index is computed by subtracting the monthly means for 1979–93 (Fig. 3).

3. Methodology

The LABs model is used to simulate the terrestrial hydrology in each of the 5020 basins. Initialization of
the model is done using a spinup method with guess initial conditions. The 1979 dataset is used to run the model twice with the conditions at the end of the first run providing the initial conditions for the second run. The conditions at the end of the second run are used to obtain the initial conditions for the 15-yr simulation starting from 1 January 1979 (see also Chen and Kumar 2001).

The ERA-15 forcing and the model output are combined into the monthly aggregates and anomalies for correlation with the ENSO index. The monthly mean of any particular variable \( y \) is computed as

\[
y_j = \frac{1}{15} \sum_{k=1}^{15} y_{kj}, \quad k = 1, \ldots, 15,
\]

where, \( k \) and \( j \) are indices for year and month, respectively. The complete 15-yr monthly anomaly series with monthly mean removed is given as

\[
y'(i) = y_{kj} - \bar{y}, \quad i = (k - 1) \times 12 + j.
\]

The sample size of this series is 180 for the 15-yr dataset. For the single monthly 15-yr dataset that groups the same months together, for each month \( j \in \{1, \ldots, 12\} \) there is a series of 15 values corresponding to each year given as

\[
y'(k) = y_{kj} - \bar{y}, \quad k = 1, \ldots, 15.
\]

For the single-anomaly series \( y'(k) \), the correlation coefficient between the anomaly of each terrestrial variable and the ENSO index is computed. The relationship between the complete anomaly series, \( y'(i) \), and the ENSO signal, \( E(i) \), is studied by computing and analyzing their cross-correlation coefficient \( \rho_{Ey}(\tau) = E[E(i), y'(i + \tau)] \) [see Eq. (A1)] and the lag \( \tau_{\text{max}} \) associated with the extremal value of the cross-correlation coefficient.

In order to locate the significant correlation areas, we use the test statistic, \( t \), to identify the confidence level \( \gamma \) of the correlations between ENSO and the terrestrial hydrologic variables. The test statistic, \( t \), is defined as (Hirsch et al. 1993)

\[
t(\rho; n^*) = \frac{\rho \sqrt{n^* - 2}}{\sqrt{1 - \rho^2}},
\]

where \( t(\rho; n^*) \) is the point on the Student’s \( t \) distribution for the effective number of degrees of freedom \( n^* \) and coefficient \( \rho \). The effective number of degree of freedom \( n^* \), when we have two autocorrelated time series, is given as follows (Emery and Thomson 2001):

\[
n^* = \frac{n}{\sum_{\tau=-\infty}^{\infty} \{ \rho_{Ey}(\tau) \rho_{y'y}(\tau) + \rho_{Ey}(\tau) \rho_{y'y}(\tau) \}}.
\]

Here, \( \rho_{Ey}(\tau) \) and \( \rho_{y'y}(\tau) \) are the cross correlations between \( E \) and \( y' \). In reality, the computation of \( n^* \) requires the substitution of sample estimates over finite lags for the correlation (Emery and Thomson 2001). Considering the characteristics of the annual cycle of terrestrial variables, we use the maximum monthly lag of 12 to estimate the effective number of degree of freedom. If each of the \( n \) values in given samples \( E \) and \( y' \) is statistically independent, the value of the denominator of the right side of Eq. (9) is equal to 1, and \( n^* \) is equal to the sample size \( n \).

The confidence level \( \gamma \) for \( t(\rho; n^*) \) is then computed as

\[
\gamma(\rho; n^*) = \int_{-r(\rho,n^*)}^{r(\rho,n^*)} f(t; n^*) \, dt,
\]

where \( f(t; n^*) \) is the probability density function of the Student’s \( t \) test. Combining Eqs. (8) and (10), the confidence interval \( \gamma \times 100\% \) are computed (see Fig. 4 for the sample size of 15). A confidence level \( \gamma \) indicates that \( (1 - \gamma) \times 100\% \) of the variability is attributable to random variability.

For the complete monthly anomaly series, the extremal values of the cross correlations and their associated time lags are obtained using the following method. The cross correlations are computed for lags of 0 to 24 months. At different lags the computations are based on \( (n - \tau) \) samples [see Eq. (A1)] where \( n \) is the maximum number of samples in a series. Consequently, the confidence level \( \gamma \) for \( \rho(\tau) \) is different for each \( \tau \). We therefore revise the \( \rho(\tau) \) to obtain \( \rho_m(\tau) \), which has the same confidence levels for all \( \tau \) with the degree of freedom of \( m \). This is obtained as (see appendix for derivation)

\[
\rho_m(\tau) = \frac{\rho(\tau) \sqrt{n^* - 2 - \tau}}{\sqrt{m - 2 - \rho^2(\tau)(\tau + m - n^*)}}.
\]

The extremal value \( \rho_m(\tau_{\text{max}}) \) (maximum absolute cross-correlation coefficient) and the corresponding lag \( \tau_{\text{max}} \) are identified from this series. The value of \( m = 15 \), that is, sample size of single monthly time series, is used as the degree of freedom to study the significance of cross correlation between ENSO and each of the terrestrial variables.
4. Results

a. Teleconnection between ENSO and terrestrial hydrology

To understand the teleconnection between ENSO and terrestrial hydrologic components, the adjusted extremal (maximum or minimum value) cross-correlation coefficient $r_m(t_{\text{max}})$ and the associated lag $t_{\text{max}}$ are computed for each of the 5020 basins using the complete 15-yr monthly anomaly series. We use a two-step process. We first compute the adjusted extremal cross-correlation coefficient $r_m(t_{\text{max}})$ and the associated time lag $t_{\text{max}}$ for each of the 5020 basins. The regions of potential impact of ENSO are located as areas where the $r_m$ is generally higher than surrounding regions and show a coherent pattern, that is, they are not randomly distributed. Each of these regions is then used for a detailed investigation to verify the impact of ENSO.

1) Runoff

The cross correlation between the ENSO index and the anomaly of the runoff over North America (Fig. 5a) shows that there is a distinct negatively correlated region in Canada, bounded by $55^\circ-65^\circ$N, $100^\circ-120^\circ$W, with absolute values in some areas greater than 0.4 and even 0.5. The confidence interval of the absolute cross correlation above 0.5 is greater than 95% (see Fig. 4). In northwestern Alaska, there is a distinct negatively correlated region as well. These two regions will henceforth be referred to as the CD and AK regions, respectively (see Fig. 6). The runoff anomalies in the north and northwest of the North American continent are most often negatively correlated with the ENSO signal. A third negative correlation region is located in south Mexico and will be referred to as the MX region.

In contrast, there are four distinct positively correlated regions in the continent. The region with the largest and the strongest correlation is located in the center of the continent, which straddles the states of Montana, North Dakota, South Dakota, Wyoming, Nebraska, and Iowa. The second region is found around the Gulf of Mexico including the Florida Peninsula. The third one is located around Hudson Bay, Canada, and the fourth around California and Nevada. These four regions will henceforth be referred to as the CN, GM, HB, and CA regions, respectively (see Fig. 6). When considering the lag $t_{\text{max}}$ corresponding to the maximum correlation (Fig. 5b), it is more than 6 months in the high land of the Rocky Mountains in the CN region. The maximum correlation in the east of the GM region has a 1- or 2-month lag with respect to the ENSO signal. The time lag in the CA region is about a season long. For the
negatively correlated AK region, the time lag is more than 6 months. The MX region has a 1- or 2-month lag.

The negatively correlated regions, such as CD, AK, and MX, demonstrate a high probability of low flow during the warm phase of an ENSO year and/or of high flow during the cool phase. The four positively correlated coherent regions have high-flow potential in an El Niño year and/or low-flow potential in a La Niña year. The spring and summer drought of 1988 over the central United States associated with La Niña and the summer flood of 1993 in the U.S. midwest associated with the mature El Niño conditions (Trenberth and Guillemot 1996) are mostly associated with the CN region. Figure 7 shows the maximum value of the cross correlation and its associated time lag between ENSO and the snow-water-equivalent anomaly. This anomaly in part of the CN region has a significant positive correlation with ENSO, and its time lag is greater than 6 months. This may explain the reason for long time lag for runoff anomaly in that region, and a more detailed explanation is given in section 4d.

2) PRECIPITATION

The forcing variable directly related to a basin’s runoff is precipitation. Figure 8 displays the extremal cross-correlation coefficient between ENSO and the precipitation anomaly and the associated time lag for each basin. The spatial correlation pattern for precipitation is consistent with that of runoff. Comparing Fig. 8 with Fig. 4, we find that the significance of some ENSO-correlated regions for precipitation is stronger than that for runoff [e.g., the area around the Middle Rio Grande (its center around 35°N and 102°W)], and vice versa for some other regions (e.g., the area around the center of the continent). This is because the terrestrial hydrologic memory modulates the effect of ENSO. The production of runoff is generally dependent on the volume of precipitation and the antecedent soil moisture. Usually, in dry areas, such as the Middle Rio Grande, the effect of anomalous pre-
TABLE 1. Seven basins, from four positively and three negatively ENSO-correlated regions over North America (see Fig. 6), are chosen for detailed study. Four streamflow stations are used for the study of the relationship between ENSO and streamflow. Basin area is obtained from the HYDRO1K data (Verdin and Verdin 1999). The drainage area is that documented by USGS for the gauge station. White, Neches, Russian, and Kobuk River basins are located in the states of SD, TX, CA, and AK, respectively.

<table>
<thead>
<tr>
<th>Region</th>
<th>Level-5 ID</th>
<th>Basin</th>
<th>Basin area (km²)</th>
<th>Stream station</th>
<th>USGS station ID</th>
<th>Drainage area (km²)</th>
<th>Lat (ddmms)</th>
<th>Lon (ddmms)</th>
<th>Phase (ENSO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN</td>
<td>89780</td>
<td>White</td>
<td>26757</td>
<td>Oacoma</td>
<td>06452000</td>
<td>26407</td>
<td>431941N</td>
<td>1014120W</td>
<td>+</td>
</tr>
<tr>
<td>GM</td>
<td>91400</td>
<td>Neches</td>
<td>25653</td>
<td>Evadale</td>
<td>04041000</td>
<td>20584</td>
<td>312023N</td>
<td>094703W</td>
<td>+</td>
</tr>
<tr>
<td>HB</td>
<td>56800</td>
<td></td>
<td>11012</td>
<td></td>
<td></td>
<td></td>
<td>542746N</td>
<td>071464W</td>
<td>+</td>
</tr>
<tr>
<td>CA</td>
<td>99720</td>
<td>Russian</td>
<td>5376</td>
<td>Guerneville</td>
<td>11467000</td>
<td>3464</td>
<td>385216N</td>
<td>122544W</td>
<td>+</td>
</tr>
<tr>
<td>CD</td>
<td>28460</td>
<td></td>
<td>3100</td>
<td></td>
<td></td>
<td></td>
<td>572215N</td>
<td>1092700W</td>
<td>–</td>
</tr>
<tr>
<td>AK</td>
<td>19230</td>
<td>Kobuk</td>
<td>24696*</td>
<td>Kiana</td>
<td>15744500</td>
<td>24646</td>
<td>617957N</td>
<td>1604031W</td>
<td>–</td>
</tr>
<tr>
<td>MX</td>
<td>95389</td>
<td></td>
<td>6657</td>
<td></td>
<td></td>
<td></td>
<td>181651N</td>
<td>0972859W</td>
<td>–</td>
</tr>
</tbody>
</table>

* The basin area includes seven level-5 basins: 19230, 19240, 19250, 19260, 19270, 19280, and 19290, and the gauge station of Kobuk is located in level-5 basin 19230.

Precipitation on runoff is attenuated. In addition, the time lag for the runoff anomaly is, in certain areas (e.g., southeast United States), longer than that for the precipitation anomaly by 1 or 2 months. In areas where runoff is primarily associated with snowmelt, such as the area around Wyoming, the lag differences between precipitation and runoff are more than 4 or 5 months.

b. Teleconnection lags

From the previous analysis it is evident that the terrestrial systems have a delayed response to the ENSO signals, as compared to the precipitation, and the delay may range from a month to a season or longer. The shorter delays are typically associated with rainfall-runoff processes where the lag is caused by soil moisture storage. The longer lags seem to be associated with snow accumulation and melt processes.

To verify that the correlation and the computed lags are not a computational or statistical artifact, several level-5 basins corresponding to the center of each ENSO-correlated region (Table 1, Fig. 6) are chosen for a detailed study. These basins correspond to the centers of the seven regions with distinct correlations between ENSO and runoff anomalies identified earlier. The product of ENSO index \( E(t) \) and the lagged time series \( y'(t + \tau_{max}) \), where \( y' \) is the runoff anomaly from the LABs simulation (Fig. 9), or the precipitation anomaly of ERA-15 data (Fig. 10), or the SWE anomaly of the model simulation (Fig. 11), illustrates that the lags of extremal correlations are indeed important in teleconnection patterns. The \( \tau_{max} \) for a given basin may be different for different variables as illustrated through Figs. 5, 7, and 8. Although there are only five ENSO events in the 15-yr study period, the correlation method is effective in identifying the influence of ENSO on the runoff response (Fig. 9), which does show markedly different characteristics during the ENSO episode periods. The 1982/83 El Niño and 1988/89 La Niña have strong signatures in basins with positive and negative correlations, respectively. In addition, basin 99720, lo-
cated in the CA region, shows a very strong correlation with the 1982/83 event and the contribution to the correlation value during the 15-yr period is primarily because of this event. This is also indicative of the fact that the regions identified as correlated with ENSO events are not necessarily affected by every episode.

The analysis of the precipitation anomaly (Fig. 10) indicates that the precipitation during ENSO episodes is generally different from other periods. The SWE anomaly (Fig. 11) is negatively correlated with ENSO, which coincides with negative correlation in the runoff anomaly in the CD and AK regions. These results also support the argument that even though the estimated correlation of hydrologic variables with the ENSO index are generally low, they do reflect the important teleconnection impacts and that the observed regional patterns are not spurious.

c. Verification

In order to verify that streamflows do indeed show an enforced response to the ENSO signal, as compared to precipitation, we analyzed the cross correlation of ENSO with observed precipitation (Higgins et al. 1996a), ERA-15 precipitation, observed streamflow, and LABs simulated streamflow for the White River, South Dakota, in the CN region; the Neches River, Texas, in the GM region; the Russian River, California, in the CA region; and the Kobuk River, Alaska, in the AK region. The area-averaged precipitation for these basins are obtained from Higgins precipitation data (Higgins et al. 1996a). Since Higgins data does not include Alaska, no data for the Kobuk River is available. The observed daily streamflow discharges for the 15-yr (1979–93) period are obtained from the USGS (EarthInfo, Inc. 1995). The details of these basins and the streamflow gauge stations of the rivers are summarized in Table 1. Figure 12 shows the cross correlation [Eq. (11)] between ENSO index and precipitation and streamflow anomalies. The cross correlation of ERA-15 precipitation closely matches that of observed precipitation and lends further credibility to our study. The LABs simulated streamflows always display stronger correlation to ENSO than observed streamflows in these basins. We surmise that regulations may be a major reason for this discrepancy. For example, in the Russian River basin, there are 62 dams along the river (California Resources Agency 1997), and the capacity of all reservoirs in California can hold 60% of the state’s total runoff every year (California Resources Agency 1997). Consequently, the observed streamflow variability usually decreases. In the study of Coe (2000), it is indicated that the man-made dams and their reservoirs can dramatically reduce the amplitude of the seasonal cycle of streamflow. However, the patterns of correlation, that is, the
peaks and associated lags are fairly consistent, thereby supporting the inference from our model study.

d. Soil moisture storage and terrestrial climate

To understand the role of soil moisture storage in determining terrestrial climate, the extremal cross-correlation coefficients and their time lags between the ENSO index and anomalies of soil water deficit are studied. Two categories of soil water deficit are considered in the study. First is total soil water deficit (TSWD) and is given as

\[
TSWD = \frac{1}{6} \sum_{i=1}^{6} \left[ \theta_{sat} - \bar{\theta}(i) \right] \times \Delta z(i), \tag{12}
\]

where \( \theta_{sat} \) is saturated soil moisture content. Here, \( \bar{\theta}(i) \) and \( \Delta z(i) \) are the soil moisture content and the thickness of layer \( i \), respectively. The second is the near-surface soil water deficit (NSWD), which refers to the soil water deficit in the first soil layer, with thickness 10 cm, obtained as

\[
NSWD = \theta_{sat} - \theta(1). \tag{13}
\]
Because the first soil layer is shallow and quickly interacts with various driving variables, NSWD reflects the rapid fluctuation of the land surface forcing from the atmosphere, and represents the short timescale memory. Meanwhile, TSWD reflects the total memory of the soil system in response to the larger timescale of the atmospheric forcing.

ENSO correlations with TSWD and NSWD anomalies (Fig. 13) provide insights in the role of slow and fast moisture dynamics in modulating the ENSO impact. The spatial correlation pattern with TSWD anomaly shows the same distribution as that of runoff. Note that the low water deficit is generally associated with higher runoff and vice versa, consequently, the sign of the correlation is opposite of each other. The patterns of $t_{\text{max}}$ are also consistent with that of lags associated with runoff. However, differences also exist. The distinct ENSO related area around California and Nevada for runoff (Fig. 5) shows correlation with NSWD (Fig. 13, right) but not with TSWD (Fig. 13, left). This indicates that the runoff anomaly can be determined by the short timescale memory of terrestrial moisture storage.

Spatial correlation pattern with the NSWD anomaly shows distribution that is similar to that of precipitation and the associated lags are similar as well. A significant difference of ENSO correlation between NSWD and precipitation is in the Middle Rio Grande. This reflects that even short timescale terrestrial memory cannot capture the signature of ENSO extremes in dry areas.

To identify the causal control on the runoff response the cross correlations between the correlation fields, that is, to compute the spatial correlation of the maps of temporal correlation, are calculated as follows:

$$
\rho(x, y) = \frac{\sum_{i=1}^{Nb} [\rho_i(x) - \bar{\rho}_x][\rho_i(y) - \bar{\rho}_y]}{\sqrt{\sum_{i=1}^{Nb} (\rho_i - \bar{\rho}_x)^2} \sqrt{\sum_{i=1}^{Nb} (\rho_i - \bar{\rho}_y)^2}},
$$

where Nb is the number of level-5 basins over North
TABLE 2. The cross-correlation coefficients between the correlation fields of ENSO and various terrestrial hydrologic variables.

<table>
<thead>
<tr>
<th></th>
<th>Precipitation</th>
<th>Runoff</th>
<th>TSWD</th>
<th>NSWD</th>
<th>Snow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>0.78</td>
<td>-0.69</td>
<td>-0.78</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Runoff</td>
<td>-0.94</td>
<td>-0.77</td>
<td></td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>TSWD</td>
<td>-0.78</td>
<td>0.78</td>
<td>-0.18</td>
<td></td>
<td>-0.23</td>
</tr>
<tr>
<td>NSWD</td>
<td>-0.30</td>
<td>0.24</td>
<td>0.18</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>

America where the ENSO correlation with runoff is greater than 0.2, and \( r_x \) and \( r_y \) are the extremal cross-correlation coefficients between ENSO and any two terrestrial variables \( x \) and \( y \). In the above equation, the difference due to the areas of the basins is not considered. The results are presented in Table 2. These results suggest that although precipitation control on runoff, TSWD and NSWD, is comparable, TSWD control on runoff is significantly larger. Based on this evidence, it seems that higher rainfall during an ENSO episode may not always correspond to extreme runoff. The moisture could get stored in the terrestrial reservoirs and less intense rainfall at later time (possibly a month to a season-long delay) may cause higher runoff due to the reduced soil water deficit.

Further corroboration of this pattern is offered by examining the ENSO correlation with runoff and precipitation, and NSWD and TSWD, using the annual series for the months of June and July [see Eq. (7)]. In June, there is a distinct positively correlated area for the precipitation anomaly over the CN and CA regions and most of the GM region (Fig. 14). A corresponding pattern is identifiable for the runoff anomaly except with stronger correlation around the GM region. As com-

![June Corr of ENSO & ERA Precipitation](image1)

![June Corr of ENSO & LABs Runoff](image2)

![July Corr of ENSO & ERA Precipitation](image3)

![July Corr of ENSO & LABs Runoff](image4)

**Fig. 14.** Monthly correlations of ENSO with (left) ERA-15 precipitation and (right) runoff anomalies for (top) Jun and (bottom) Jul. The confidence interval for the regions inside \(-0.5 \) and \( 0.5 \) contours is greater than 95%.
pared to June, precipitation anomaly shows lower correlation in July for the GM region. However, the runoff continues to show strong correlation in July. The TSWD continues to show strong correlation in July (Fig. 15) while the pattern for NSWD anomaly is more reminiscent of that of precipitation. In June, the location of the negative correlation between ENSO and the TSWD anomaly closely matches that of ENSO and the runoff anomaly. In addition, the area of the correlation between ENSO and the NSWD anomaly has not only the signature of the precipitation anomaly but also that of the TSWD anomaly. In July, the correlation area around the GM region for the TSWD anomaly shrinks and the correlation strength decreases, and the correlation area around the CN region enlarges and the correlation strength increases. Moreover, the correlation area around the GM region for the NSWD anomaly reduces dramatically and only remains a small area near the outlet of the Mississippi River. However, the area along the U.S. and Canada border with high correlation between ENSO and the precipitation anomaly also shows the distinct correlation between ENSO and the NSWD anomaly. Clearly, the NSWD anomaly is a mix representing primarily the dynamics of precipitation but it is also impacted by the total storage as characterized by TSWD.

5. Summary and conclusions

A basin-scale hydrologic model (LABs) is used to simulate the terrestrial hydrology of North America. The model is driven by ERA-15 data for the period 1979 to 1993. The continent is divided into 5020 basins using the HYDRO1K data (Verdin and Verdin 1999) with an average basin size of 3255 km$^2$. A 2-yr spinup cycle is used for initializing the model. The simulation time step
is 30 min and the model output is aggregated to monthly timescale. The anomalies are computed with respect to the average monthly climatology of the simulation period for various variables such as precipitation, runoff, snow-water equivalent, and both near-surface and total soil water deficit.

The anomaly time series of these variables for each basin are correlated with the ENSO index. Maps of extremal (maximum or minimum) correlation values and the associated lags are studied to find regional impact of ENSO. Seven distinct regions are identified over the continent. The lags corresponding to runoff anomalies in general are longer than that for rainfall. Regions where runoff is primarily produced by rainfall-runoff processes show shorter lags than those where snowmelt is the primary contributor to the runoff. The simulation results are further corroborated with precipitation and streamflow observations.

The study using both near-surface and total soil moisture deficit reveals that influence on runoff is generally communicated through the latter, while the former is more coherent with the rainfall pattern. It is hypothesized that the relatively slower dynamics of terrestrial moisture storage serves as a memory in causing the delayed peak in runoff as compared to the precipitation.

The correlation coefficients between ENSO and terrestrial variables are generally low. This is possibly due to the short time span of 15 yr used in the simulation, which included relatively few ENSO events. However, detailed examination of the time series show no spurious behavior in the regions identified. The more likely reason for low-correlation values is because the ENSO events may not explain all the variability of the hydrologic processes. This is further corroborated by the observation that a correlated region is not always influenced by ENSO episodes. It is also worth noting that in North America ENSO teleconnections are less robust and have more uncertainty than those in some tropical countries, such as those in South America.

The extreme events related to runoff have the potential to cause severe damages in the related ecological systems. The analysis presented here enables us to localize the potential areas influenced by ENSO, and it may be possible to make monthly, seasonal, or annual terrestrial climate prediction for these areas based on the knowledge and capability of predicting ENSO.

In this paper, the primary contribution is in establishing that in the regions of ENSO influence, terrestrial hydrologic variables such as runoff exhibit a response that has a delay greater than that of precipitation to ENSO signal, and this phase lag is caused by the storage characteristics of the terrestrial water reservoirs.

The correlation method used in this research is a linear analysis technique. However, it is established that ENSO signals generally have a nonlinear impact (Hohenberg et al. 1997). This implies that different ENSO extremes for the same phase will have dissimilar effects on the atmospheric system, and consequently the terrestrial system, because of differences in strength and timing and variations in the atmospheric background state. The observation that the correlated regions are not always influenced by the ENSO episode lends support to this argument. To design a method to study this nonlinearity is difficult [see Mason and Goddard (2001) for a possible alternative]. Singular spectrum analysis (Broomhead and King 1986; Vautard et al. 1992), the most proficient tool available for nonlinear analysis, is in the initial stages of being applied to study intercorrelation between variables (Varadi et al. 2000) and it does not reveal the phase-lagged information. The linear analysis study presented here is consistent with previous results by other scientists, but offers a more detailed explanation of the role of terrestrial moisture storage in modulating the climate signal.

Acknowledgments. Support for this project has been provided by NASA Grants NAG5-3361, NAG5-8555, NAG5-7170, and NSF Grant EAR 97-06121. Computational support was also provided by NCSA Grant EAR990004N. Thanks are also due to Xin-Zhong Liang for insightful discussions and to the two anonymous reviewers for their thoughtful comments, which helped to clarify and improve the presentation of this study.

APPENDIX

Adjustment of Cross-Correlation Coefficients

The cross-correlation coefficient at lag $\tau$ between any monthly anomaly of hydrologic variable $y'_t$ and the ENSO index $E_t$ is given as

$$
\rho(\tau) = \frac{\sum_{t=1}^{n-\tau} (E_t - \overline{E}) (y'_{t+\tau} - \overline{y'}_{t+\tau})}{\sqrt{\sum_{t=1}^{n} (E_t - \overline{E})^2 \times \sum_{t=1}^{n} (y'_{t+\tau} - \overline{y'}_{t+\tau})^2}},
$$

(A1)

where $(n - \tau)$ is the sample size, $\overline{E}$ is the mean of the first $n - \tau$ values of series $E_t$, and $\overline{y'}_{t+\tau}$ is the mean of the last $n - \tau$ values of series $y'_{t+\tau}$.

The sample size $(n - \tau)$ for computation of $\rho$, decreases as lag $\tau$ increases. The $\rho$, confidence level $\gamma$ is obtained as

$$
\gamma(\rho; n^*) = \int_{-\rho}^{\rho} f(t; n^*) \, dt,
$$

(A2)

where $n^*$ is the effective number of degrees of freedom and is computed using Eq. (9). In this study, the criterion for evaluating the significance of correlation is based on the confidence interval. For the convenience of obtaining a cross-correlation coefficient with the maximum confidence interval, we adjust the cross-correlation coefficient computed from Eq. (A1). For sample size $(n - \tau)$, Student’s $t$ test is given as
\[ t[\rho(\tau)] = \frac{\rho(\tau) \sqrt{n^* - 2 - \tau}}{\sqrt{1 - \rho^2(\tau)}}. \quad (A3) \]

As well, for \( m \) degrees of freedom, \( t \) is computed as follows:

\[ t[\rho_m(\tau)] = \frac{\rho_m(\tau) \sqrt{m - 2}}{\sqrt{1 - \rho_m^2(\tau)}}. \quad (A4) \]

Maintaining the equivalence of the confidence intervals computing from \( t[\rho(\tau)] \) and \( t[\rho_m(\tau)] \), we combine Eqs. (A3) and (A4) and have

\[ \rho_m(\tau) = \frac{\rho(\tau) \sqrt{n^* - 2 - \tau}}{\sqrt{m - 2 - \rho^2(\tau)(\tau + m - n^*)}}. \quad (A5) \]

This is identical to Eq. (11), which is derived for adjustment of cross-correlation coefficients. It is worthwhile to note that the effect of various numbers of degrees of freedom on \( t \) distribution is not included during the derivation of Eq. (A5). However, this effect is negligible when the degree of freedom is large enough (i.e., larger than 15) and the confidence intervals considered is not extremely high (say, less than 99%) (see \( t \)-distribution table, Box et al. 1994). In our study, the complete monthly time series meet these requirements, and using Eq. (A5) to adjust cross correlation is reasonable.

**REFERENCES**


Olson, J. S., J. A. Watts, and L. J. Allison, 1983: Carbon in live vegetation of major world ecosystems. ORNL-5862, Oak Ridge National Laboratory, Oak Ridge, TN.

1991: Description of bare essentials of surface transfer for the bureau of meteorological Research Centre AGCM. BMRC Research Rep. 32, BMRC, Australia.


Verdin, K. L., and S. Jenson, 1996: Development of continental scale DEMs and extraction of hydrographic features. *Proc. Third Int. Conf/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM, National Center for Geographic Information Analysis, University of California, Santa Barbara, CD-ROM. [Available from NCGIA, Santa Barbara, University of California, 3510 Phelps Hall, Santa Barbara, CA 93106-4060.]

