Civil Engineering   School of Engineering

2-18-2009

Using synoptic-scale land surface moisture indices for hydrologic prediction

Andrew Wood

Edwin P. Maurer
Santa Clara University, emaurer@scu.edu

Follow this and additional works at: http://scholarcommons.scu.edu/ceng

Part of the Civil and Environmental Engineering Commons

Recommended Citation

Using synoptic-scale land surface moisture indices for hydrologic prediction

ANDREW W. WOOD[1] and EDWIN P. MAURER[2]

[1]Department of Civil and Environmental Engineering, University of Washington, Seattle 98195 USA

[2]Department of Civil Engineering, Santa Clara University, Santa Clara, CA 95053 USA

Abstract

Forecasting seasonal runoff is an important challenge in the western U.S. because the timing and volume of summer (dry season) streamflow plays a critical role in managing water supply and delivery systems. Beginning in December of each year, observations of the snow pack begin to provide invaluable information on the amount of runoff expected in late spring and summer. The skill of such water supply forecasts also benefits, in some locations, from the consideration of synoptic climate indices -- e.g., using sea surface temperature anomalies to characterize ENSO or PDO state, which have been shown to have predictable teleconnections to land surface hydrology months in advance. These two sources of forecast information complement each other, with climate/SST state information providing predictability at long lead times (3-9 months), and observed snow and other moisture states of local catchments providing skill at shorter lead times (1-4 months). However, because similar SST anomalies can produce widely different future hydrology outcomes, information at intermediate lead times describing how a particular teleconnection or climate pattern is evolving on a macro-scale level (especially regarding precipitation anomalies) may provide additional insight into future spring and summer runoff. In this study, we explore the use of a long-term land surface hydrology data set to define continental-scale hydrologic indices with potential predictive value, and to assess where in the western U.S. these might enhance predictability already achievable using local observations and climate indices. We find that the west-wide hydrologic indices have potential value in the Pacific Northwest and the Southwest, but yield little improvement in other areas of the western U.S., relative to existing climate and observation-based indices.

1. INTRODUCTION

Forecasting seasonal runoff is an important challenge in the western U.S. because the volume of summer (dry season) streamflow is critical in the management of water resources, the objectives of which include water supply for municipal, industrial and agricultural uses, hydropower generation, navigation, recreation, fisheries and provision for natural ecosystems (Pagano and Garen, 2005b). Beginning in December of each year, the accumulating snow pack throughout this domain begins to provide invaluable information about the amount of runoff expected in late spring and summer, and in fact, observations of snow water equivalent (SWE) are the primary predictors used by operational forecasting entities at the federal, state and local levels to forecast summer streamflow. The most common forecasting method in current operational use is to regress cumulative streamflow for the summer periods (e.g., April-July or April-September) on observed SWE, cumulative water year precipitation, and current or past streamflow, which can be a proxy for the state of the groundwater or soil moisture in a river basin. These predictors are almost always 'local' observations -- that is, measured within or just outside the boundaries of the drainage basin for the streamflow forecast point. Due to the role of snowmelt in generating streamflow, the accuracy of the locally-based predictions increases steadily throughout the winter and early spring, achieving correlations of 90 percent by April 1 in some locations in the western U.S. (Pagano and Garen, 2005b). Forecast accuracy is low enough before mid-December, however, that the lead forecasting agencies – the USDA National Water and Climate Center (NWCC) and NOAA River Forecast Centers (RFCs) – have not traditionally issued forecasts earlier than this in the water year.

The skill of summer water supply forecasts also benefits, in some locations, from the consideration of teleconnection indices that reflect the connection of variations in continental climate to variations in sea surface pressures and temperatures (SSTs) in the oceans. Foremost

among these teleconnections, from the standpoint of seasonal climate prediction, is the El Nino-Southern Oscillation (ENSO), an ocean-atmosphere phenomena stemming from an oscillation in the location of a warm pool in the equatorial Pacific that is linked to climate patterns in many extra-tropical locations (Goddard et al., 2001). Many indices describing ENSO have been developed, including the Southern Oscillation Index, which reflects zonal pressure gradients in the equatorial Pacific (Ropelewski and Jones, 1987) and the NINO3.4, which measures SSTs in the equatorial Pacific (Trenberth, 1997), and more recently, for example, the Trans-Nino Index (TNI), which has been found to improve predictions in the transition zone between the Pacific Northwest and California (Trenberth and Stepaniak, 2001). ENSO has a dramatic influence over western US climate and hydrology (e.g., McCabe and Dettinger, 1999; Cayan et al., 1999; Quan et al., 2006], and in the last decade, advances in understanding the physics and predictability of ENSO have led to improvements in hydrologic and streamflow forecasting. A number of studies (e.g., Maurer et al., 2004; Piechota and Dracup, 1999; Piechota et al., 1997) have catalogued the streamflow or hydrologic predictive skill by season and location for at least a half-dozen such ocean-based indices. Typical correlations with future summer runoff range from 0.10 to 0.45, depending on location, lead time and time of year. The ENSO indices tend to have predictive value for future streamflow in the western US beginning in late summer and peaking in late fall. In contrast to the local predictors described above, the climate teleconnection indices can be thought of as remote, measured at a large distance from the streamflow forecast locations. In some western U.S. locations, the two sources of forecast information (local and remote) complement each other, with ENSO states providing relatively higher predictability at long lead times (6-9 months), shorter than which local observations become more useful.

The skill of predicting the SSTs used to determine ENSO states is higher than that for predictions of climate (hence of future continental hydrologic variables) at the same lead times due to the variability of atmospheric conditions that develop during various ENSO phases (e.g., Gershunov, 1998). For this reason, information describing the actual evolution of a particular teleconnection or climate pattern (for example, an ENSO event) in terms of the resulting land surface moisture states may provide additional insight into future summer runoff before local predictors become central. Such information represents a third potential category of streamflow predictor that has not been widely, if at all, explored: synoptic to continental-scale hydrologic indices reflecting the strength and location of moisture transport onto land, as well as the persistence of previously accumulated land surface moisture in various states. Unlike the oceanic indices, the large scale hydrologic index is not primarily intended to describe conditions that drive climate, work to understand land atmosphere feedback mechanisms (e.g., Koster et al., 2000; Hong and Kalnay, 2000) notwithstanding. Rather, it reflects the integration of observed climate patterns affecting a region over a period of preceding months. A hydrologic index is 'downstream' of oceanic effects on climate, hence offers less lead time for predicting hydrologic response than an ocean-based index. By integrating the climate effects of large scale circulation patterns, however, such an index may identify significant seasonal moisture accumulation patterns that augment the information provided by local predictors.

The investigation of synoptic-scale hydrologic indices is motivated by recent hydroclimatic experience in the western US. For example, in water year 2005 (categorized as a weak La Nina event) a striking divergence of dry hydrologic conditions in the north and wet conditions in the south (illustrated in Wood and Lettenmaier, 2006) developed, due in large part to persistence of the dominant storm track from the Pacific Ocean over southern California and a high pressure system

just off the coast of Washington and Oregon. This north-south dipole of relative moisture conditions, which forms a prominent mode of hydroclimatic variability in the western US (Dettinger et al., 1998; Cayan et al., 2003), was evident as early as October 2004, lasted through at least May 2005, and produced both record-setting snowpacks and streamflows throughout the domain – high in the south, particularly the Colorado Basin, and low in the north.

Recent work focusing on the western U.S. provides a foundation for understanding the region's spatial and temporal hydro-climatic variability, despite not addressing summer streamflow prediction. Working with datasets of point observations (e.g., of SWE, precipitation, temperature), researchers have analyzed spatial and temporal variability of hydroclimatic variables in the western U.S. and elsewhere (e.g., Mote et al., 2005; Cayan et al., 2001; Cayan, 1996; Hamlet et al., 2005). Cayan et al. (2003) investigated regional connections of high and low discharge, and found statistically significant co-incidence of high flow years in the Sierras and the Columbia Basin, as well as years of contrasting extremes in Columbia and Colorado basin flows. Using streamflow gage data, others have identified long-term trends in streamflow and streamflow variability (Jain et al., 2004; Pagano and Garen, 2005a), in some cases relating them to trends in climate patterns. Spatially distributed model-based (gridded) datasets have also been the basis for investigations of land surface moisture variability. Maurer et al. (2004), for instance, assessed potential land surface and climate index sources of future runoff predictability.

The forecast-centered analysis presented in this paper follows a line of inquiry pursued by Keyantash and Dracup (2004), who found value in the use of first principal component of moisture variables to characterize droughts in three California basins. Here, we use a retrospective land surface hydrology data set to define synoptic-scale hydrologic indices with potential predictive value for summer streamflow, and to assess where and when these might augment the predictability

achieved using local observations and remote climate teleconnection indices. One significant difference from the Keyantash and Dracup (2004) work is the derivation of these indices using the entire western U.S. domain, hence allowing the incorporation of information external to the streamflow basins of interest. The central research question is whether the assessment of the entire west-wide domain can lead to earlier identification of the types of streamflow years highlighted by Cayan et al. (2003), hence actual streamflow during the peak runoff period of the year, than traditional basin-focused prediction approaches.

2.   APPROACH

Two types of indices, both based on hydrologic information from the western U.S. domain, are investigated as predictors for summer streamflow in the western US: (1) the principal components of moisture over the entire domain, defined as the sum of SWE and soil moisture; and (2) the average moisture within defined sub-regions of the domain. These are used in a multiple linear regression framework to forecast April-July streamflow at 4 major river locations, and compared with predictions based on NINO3.4 and average moisture within only the drainage areas of each respective flow location. This section describes the dataset from which the indices are derived, the index derivation procedure, and the streamflow forecast locations and regressions.

2.1   Land surface moisture dataset

Gridded long-term climatologies of land surface variables such as SWE and soil moisture are less thoroughly established as a data resource than are atmospheric (re)analyses (a prime example is the NCEP reanalysis, which has supported a vast array of research and operational activities). Maurer et al. (2002) provides one of the earliest and most comprehensive continental-scale hydrologic analyses archived for public use: a retrospective model-based simulation of over 40 land surface energy and moisture variables at a 3 hourly timestep and 1/8 degree latitude by longitude spatial resolution, for a region encompassing the continental US and portions of Mexico

and Canada. The Maurer et al. (2002) dataset was generated using the distributed water and energy balance Variable Infiltration Capacity (VIC) macroscale hydrologic model (Liang et al., 1994; 1996, among others), which has been featured in many analyses of regional and continental scale hydrology. Since the Maurer et al. (2002) implementation, the model has also been applied to the domain of this study at 1/8 degree longitude by latitude spatial resolution by Christensen et al. (2004), Payne et al. (2004), Wood et al. (2005), Van Rheenen et al. (2004), Wood and Lettenmaier (2006). Andreadis et al. (2005) also aggregated the 1/8th degree parameter sets to create a ½ degree resolution VIC simulation for the continental US.

In 2005, the VIC parameter sets from the ½ degree implementation were used in combination with an index station approach for generating real-time gridded meteorological forcings (Tang et al., 2007) to create a daily-updating land surface simulation for the continental US, called the Experimental Surface Water (SW) Monitor (described in Wood et al., 2007; see http://www.hydro.washington.edu/forecast/monitor/). The period of this simulation is from January 1, 1915 to the present. The current products of the system are surface hydrologic and climate fields (grid cell average precipitation, air temperature, evapotranspiration, runoff, soil moisture for three model layers, and snow water equivalent), data and maps of soil moisture, runoff and SWE, as well as of their anomalies and percentiles relative to a historical period. The work of this paper is almost entirely based on the SW Monitor simulation of soil moisture and SWE (the sole deviation from this approach is discussed in Section 2.3). One aspect of the SW Monitor system that makes its simulation results appropriate for this work is that the temporal and large-scale variability of the meteorological forcings for the simulation are primarily derived from a set of NWS Cooperative Observer stations that were selected for their long historical records (45 years at a minimum, and in many cases much longer) and reliable real-time reporting. This

consideration, to some extent, protects the resulting simulations from potentially spurious effects of inhomogeneities in the station-based observing network. For the purpose of real-time estimation of hydrologic conditions, which is necessary for forecasting applications, the real-time and retrospective consistency also helps to legitimize the assessment of current simulated conditions in the context of the multi-decade simulated historical climatology.

2.2 Derivation of Moisture Indices for Use as Predictors

A common approach to characterizing land surface moisture is to use accumulated precipitation over a period of several months to a year – e.g., the Standardized Precipitation Index (SPI; McKee et al, 1993) is often used as a hydrologic indicator of drought. Although the SPI also describes the strength and location of moisture anomalies in the western US, however, it does not account for the partitioning of rain into snowpack, surface runoff and infiltration, or for evaporation, all processes that regulate land surface moisture storage and ultimately streamflow response to precipitation. For that reason, our study focuses on soil moisture and snow water equivalent as streamflow predictors.

Principal components analysis (PCA) has been used previously to reduce dimensionality in large data sets and to extract spatially coherent modes of variability of many hydrological variables (e.g., Bartlein, 1982); Cayan, 1996; Cook et al., 1999; Derksen et al, 1997; Guetter and Georgakakos, 1993; Lins, 1997; Maurer et al., 2004; Piechota et al., 1997; Wittrock and Ripley, 1999). In this study, we apply PCA to characterize the variability of monthly surface moisture over the Western U.S. Principal components (PCs) were computed on the gridded monthly sum of the land surface's stored water (SWE added to soil moisture). The sum of the two fields was used primarily to avoid the complicating issue of the seasonally varying partitioning of moisture between the two states (e.g., the transfer from snow to soil during the melt season). The moisture

data were arranged into a set of 12 matrices, each with one column for each grid cell in the domain (comprising 1220 1/2 degree grid cells) and one row for each year. The correlation rather than covariance matrix was used in the computations so that high PC loadings identified high moisture variability, rather than regions with relatively high surface moisture – i.e., the mountainous and coastal regions. The original moisture data matrices were then projected onto the PC loadings matrix to obtain a time series for each PC and each month.

Figures 1, 2 and **3** show the PC loading patterns for each month for the first three PCs, which together account for 49-53% of the total variance for October-July, and 41-44% for August-September. Most of this variance is captured in the first two PCs, which consequently are chosen as predictors in the flow prediction regression described in the following section. The highest loadings for the first PC appear in the center of the domain, and exhibit relatively small changes in amplitude and location throughout the year. The first PC loading's timeseries contains variability at a number of higher frequencies (monthly to interannual), and a remarkable change beginning in the late 1970s toward a more pronounced lower frequency mode of variation. This finding is consistent with the increasing synchroneity of western US climate reported by Jain et al. (2005). Also notable within the first PC is the correspondence of low periods with the century's major droughts, such as the 1930s, 1977 (which defined many water supply shortage records in the western US), and the recent 2000s drought in the Colorado R. basin.

The second PC loadings define a dipole with the axis oriented from the US northwest (NW) to southwest (SW), which is the primary mode of western UW moisture variation associated with ENSO. The second PC timeseries tends to exhibit an interannual frequency variability throughout the record. The third PC loadings, mostly by construct of the PCA technique, show a dipole with the axis oriented from California to Montana, and the associated timeseries is also dominated by

high frequency variation. Both the second and third PC loading patterns also show relatively modest variation throughout the year. The first PC may reflect, to some extent, the overall strength of moisture transport into the western US, whereas the second and third PCs characterize variations in the location of the moisture transport, which is affected by the ENSO cycle via the jet stream positioning.

The spatial loadings are also used to guide the selection of sub-regions within the domain containing the highest loadings for each of the first two PCs, within which the average moisture is specified as the basis for an predictive index. The rationale for this alterative approach is that fixed regions that reflect the behavior of the larger spatial field are easier to monitor than the entire field, hence a researcher or practitioner without the wherewithal to conduct or access land surface simulations or perform PCA might approximate that approach by averaging the more widely available point observations of snow, precipitation and streamflow (a potential proxy for soil moisture) within the sub-regions. The sub-regions are analogous to the implementation of latitude-longitude boxes for SST measurement to generate indices such as Nino1, Nino 3 and so forth, and are also similar to the use of river-basin averages in the PCA frameworks of studies cited earlier.

Figure 4 delineates one sub-region labeled CNTR that captures the high loadings from the first PC, and the difference of the moisture in the two subregions labeled NW and SW (i.e., NW minus SW) captures the behavior of the second PC. These predictors are termed "indexes" in the subsequent discussion. The delineation of the sub-regions was the result of visual inspection of the nodes and antinodes resulting from PCA (as in Wallace and Gutzler, 1981), but this choice could also be accomplished using more objective techniques that would, perhaps, optimize the correlation of the subareas with the PCs that they approximate.

2.3   Development of regression models for streamflow prediction

The performance of two experimental sets of predictors are contrasted with that of more traditional predictors in a multiple linear regression framework for forecasting summer flow in several river basins of the western US. The traditional predictors for streamflow from a drainage basin are moisture observations (typically estimated from accumulated precipitation, streamflow and/or SWE at measuring stations) within the drainage area. As a surrogate for this approach, we use the moisture averaged over the grid cells within the drainage basin, termed "local moisture" (LM). Because the Columbia River basin contains significant drainage area in Canada, outside the SW monitor boundaries, LM for the Columbia River was calculated from a VIC simulation for that basin taken from the west-wide streamflow forecasting system described in Wood and Lettenmaier (2006). The underlying parameters and forcings for the two simulations are nearly identical, differing mainly in resolution and extent. The inclusion of the Canadian portion of the domain improves the LM estimate for the Columbia River. The last, by-now traditional type of predictor for streamflow is an ENSO teleconnection index, for which we use the Nino3.4 SST index, here termed "teleconnection" (TC). Where their predictive skill warrants, teleconnection indices are now combined with LM indices by operational streamflow forecasting agencies. The five forecasting models evaluated in this paper are summarized in **Table 1**.

Summer streamflows (April-July volumes) are denoted by Q, and are from the four basins shown in **Figure 5**. The TC, PC and index variables are the same in each equation for each basin, whereas the LM variable is calculated for the drainage basin for the flows (Q) being predicted. Forecast equations are developed using monthly predictor values for each month from July of the year previous to the summer flow predictand to June of the same year. The July prediction equations use moisture from the entire month of July, hence the effective date of the earliest forecast evaluated in each year is August 1.

The single and multiple linear regression prediction equations were solved using the open

source R statistical software package. A leave-one-out cross-validation approach was used to test

the performance of the models derived in this manner, and the results given for predicted summer

streamflow are to the extent possible based on the cross-validated performance of the forecast

models. The verifying observations for the forecasts were the naturalized monthly streamflow

records for each river location, aggregated to the summer period of April-July. The length of

record used in the calculations spanned from 1950 to the end of the available naturalized record for

each basin (2005 for the Columbia and Feather Rivers, 2004 for the Colorado River, and 1997 for

the upper Rio Grande River), hence the correlations are generally applicable for the second half of

the 20th century.

3.   RESULTS

Before presenting the results of the forecast model comparison, we describe the behavior of the

PCs and indices relative to the interannual variation in a flow predictand, using as an example the

Columbia River location, which yielded the best predictability results of the locations evaluated.

Figure 6 illustrates the temporal variation of the predictors used in Table 1 relative to the summer

flow predictand for the Columbia River at the Dalles, Oregon. All timeseries are normalized to the

range 0-1 for comparison purposes. The PCs and their corresponding indices (PC1 and CNTR,

PC2 and NW-SW) have similar values, with correlations of 0.87 and 0.84 for the first and second

PCs with the corresponding index, respectively. A qualitative appraisal suggests that the first PC

and index tend to mirror the lower frequency variation in the flow timeseries, particularly in the

latter part of the record, while the second PC and index tend to mirror the interannual changes in

the summer flow variable. The LM predictor varies in time with the flow predictand, as might be

expected, while the teleconnection index has an inverse correlation with the flow predictand (in this location), and the poorest relationship with it of all the predictors.

For interannual prediction purposes, it is notable in Figure 6 that the changes in the second PC and index often anticipate the changes in the summer flow (between the previous summer and the forecasted summer) by a several months. A strong correlation between the change of the second PC or index leading into the fall months in particular and the change from the prior year's summer flow to the future summer flow would be useful for water supply forecasting applications. One way of assessing this relationship quantitatively is to estimate a linear trend through the PC or index values for several seasons leading up to the fall and winter, and then relate that trend to the year-over-year change in summer flow. **Figure 7** shows the relationship between trend in the second index (after normalizing the index to a fraction between 0 and 1) during 6-, 7- and 8-month periods leading up to (a) December 1 and (b) January 1. The strongest relationship is obtained for longer trend period (8 months) leading up to January 1, the date at which traditional water supply forecasting regression techniques are beginning to show skill. In this case, the trend is for the period from May through December, and is compared to the change in summer flow from the current to the following year.

The strong relationship between the behavior of the second index and PC with the interannual variation in the Columbia River streamflow illustrates the source of good correlations between these indices early in the water year (October 1 – September 30) with the summer flow toward the end of it. For the Columbia River location, **Figure 8** compares the forecast performance of the five prediction approaches listed in Table 1. The first statistic (top) used for comparison is the cross-validated R^2, which represents the forecast variance explained by the prediction approach. Note, cross-validated skill measures such as R^2 can (and this case, do) suffer large biases when the

regression relationship is weak (i.e., R^2 less than 0.05), hence we show the regression calibration R^2 in those cases (e.g., for LM before December 1), with negligible effect on the results. The August 1 through November 1 forecasts, the teleconnection (TC) affords the only source of predictability, and at a modest level that declines as the following summer approaches. Beginning December 1, the traditional local moisture (LM) based forecast begins to register snow accumulation and steadily increases its correlation with summer flow as it approaches and enters the forecast period (the May 1 and June 1 forecasts are within the summer forecast period, and fall toward the end or after the snowmelt period). The improvement of LM forecast with the teleconnection (LM+TC) declines as the water year develops and the variability explained by the teleconnection begins to be incorporated into LM in the form of snowpack and soil moisture. The west-wide predictors coupled with the teleconnection (labeled PC+TC and NDX+TC for the PCs and indices, respectively) yield the best overall prediction skill. They do not improve upon the LM+TC or TC only approach until after November 1. From December 1 through February 1, their R^2s are 10-20 percent higher than the best of the other approaches (LM+TC), after which their skill falls nearer to that of the LM and LM+TC approaches. The use of indices appears to give slightly higher R^2s than the use of PCs, but the difference is minimal.

Figure 8 also shows the results for mean relative absolute error (bottom) – where the absolute error each year is expressed relative to the observed value in that year – that mirror those for R^2. The approximately 11 percent relative error for the January 1 Columbia River forecast is about equal to that which was achieved in 2002-2006 (years for which data were available) by the Northwest RFC for the official forecasts of April-August flow. The official forecasts are produced from objective statistical forecasts that are subjectively adjusted by teams of experts.

Like the Columbia River flow location, the Rio Grande River at Lobatos, CO, derives flow

from a drainage area that is near a region of higher loadings in the second PC (and near one NW-

SW index sub-region). **Figure 9** shows that the upper Rio Grande River forecasts gains little or no

predictability from the teleconnection at any time during the year, and the local moisture (LM)

approaches show only minimal skill through April 1, the start of the summer flow period. This is

largely because the summer flow in this basin is driven only in part by winter snow accumulation,

and the late spring and summer rainfall that also drives the summer runoff response has little

inherent predictability. The west-wide moisture based prediction schemes substantially improve

the skill of summer flow predictions from December 1 until May 1 (after which they are similar to

the LM-based approach), but the maximum skill does not approach that achieved by the Columbia

River forecasts. The degree of variance in summer flow for the upper Rio Grande is explained by

both the first and second PCs (and similarly by the two indices), with the first predictor

contributing more prediction skill.

A comparison of Figures 1, 2, 4 and 5 shows that the Colorado River basin upstream of Lees

Ferry, AZ lies between the regions of high loading in the first and second PCs. This is the likely

reason that the west-wide moisture predictors, which again produce very similar forecast

performance to each other, do not lead to a great improvement for this flow location (**Figure 10**) in

comparison to the traditional LM approaches. Most of the predictability for the Colorado River

flow comes from the first PC and index, rather than the second pair of predictors. The ENSO

teleconnection (in this case, Nino3.4) is not known to provide significant predictability for

Colorado River streamflow, and as a result its inclusion with the LM predictor adds no appreciable

forecast skill (and even slightly reduces the cross-validated skill of the prediction combining LM

with TC). The additional variance explained by the west-wide predictors is about 10 percent on

December 1, however, which could help to extend the lead time at which usable information for water supply purposes is available.

The fourth basin selected for analysis, the Feather River at Oroville Reservoir, CA, lies outside the boundaries of the high loading areas for the first west-wide index predictors (PC1 and CNTR), and lies between the two high loading areas (the dipole) of the second indices (PC2 and NW-SW). As shown in **Figure 11**, the ENSO teleconnection provides essentially no information about this summer flow throughout the water year, and no prediction approach has skill before December 1. This is a well-known phenomenon in this region, where local impacts of ENSO events can vary substantially (e.g., Maurer et al., 2006). From December 1 onward, the traditional LM (and LM+TC) approach is the most skillful prediction, leading to high levels (nearly 80 percent) of explained variance as the summer period approaches. This is because the winter (December-February) moisture, trapped in the form of SWE, is the major component of summer flow. In contrast, the west-wide moisture predictors never attain R^2s of even 50 percent. Because the third PC of west-wide moisture contains high loadings in the Feather River drainage, a prediction approach that included the third PC was also evaluated for this basin only. It did not improve significantly on the performance of the west-wide prediction schemes, possibly because the relative percentage of variability explained by the third PC is low.

The ENSO influence on western US hydrology is arguably captured by the second PC and index, both of which characterize a dipole of moisture between the US Pacific Northwest and desert southwest (centered on Arizona and New Mexico) regions. As shown in **Table 2**, the variation in this dipole is the source of all of the skill in the west-wide predictors for the Columbia River in November and December, but the first PC and index contribute moderately to Columbia River prediction skill beginning in January. For the upper Rio Grande flow, the dipole in the end of the

year is moderately correlated with summer flow, after which the first and second predictors play a more balanced role. The Colorado River basin, in contrast, gains predictability almost entirely from variation in the first west-wide predictors. The small difference in performance between the local moisture and west-wide predictors, coupled with the geographic proximity of the area of high loadings in PC1 with the upper Colorado River basin drainage, suggest that variations in the remainder of the west-wide domain add little or no useful information for prediction in this location. The results from Table 2 for the Feather River basin indicate a primary influence of PC2 during the fall, but this is inconsequential because neither PC1 nor PC2 explain more than a few percent of summer flow variance until February 1, after which the Feather River is more correlated with PC1.

4.   DISCUSSION AND CONCLUSIONS

Land surface moisture stored in the soil and in the form of snow integrates climate effects over periods of months to years and reflects the non-linear dynamics of the land surface's control on moisture fluxes. As a result, the evolution of the moisture state of the land surface exhibits inertia, hence auto-correlation that gives rise to predictability. The synoptic weather systems that affect the western US often have a sub-continental scale that lead to climate patterns, hence land surface moisture patterns, that are manifested over the entire domain considered in this study. This observation is the rationale for investigating the use of indices of moisture variability for the entire domain (as defined using PCA) as predictors for the streamflow variables that are widely used for water supply management. Stated another way, our premise is that because the climate that determines water supply conditions in one region is part of synoptic-scale systems that affect the entire western US, a consideration of moisture conditions over a domain extending beyond any particular region may yield greater insight into the evolution of those conditions.

From the foregoing analysis, we find that predictors based on synoptic scale hydrologic conditions have the potential during the fall months to increase prediction skill for future summer streamflow in parts of the western U.S., relative to using local estimates of moisture either alone or in combination with a SST-based teleconnection index. The four basins selected for use in this analysis varied in their geographic relationship to the west-wide moisture indices and the high loading regions of the two west-wide moisture PCs, and their usefulness as predictors for summer flow in the basins varied as a consequence. Where the basin's drainage area was adjacent with or overlapped the index area or the high loading regions, such as for the Columbia and upper Rio Grande rivers, the west-wide moisture indices yielded, at times, higher predictive skill than the teleconnection index and local moisture. Elsewhere, such as in the Feather River basin, the local moisture predictors were equal or superior at all lead times. The largest benefit of the west-wide predictors comes early in the water year (between October and January 1), and at least in the Pacific Northwest, augments the longer lead, but moderate level, prediction information of teleconnection indices such as the Nino3.4 region SSTs. Where the west-wide predictors produced forecasts with higher skill, significant advantages over local moisture were largely eliminated by mid-winter.

Several factors that bear on the results of the analysis are worth mentioning. First, the west-wide predictors were developed based on the combination of soil moisture and SWE, two variables that have distinct roles in the hydrologic cycle and contributed differently to the predictability of runoff. Particularly during the spring in snowmelt regions, the relative warmth of the year may determine whether moisture resides in the snowpack or in the soil, with consequences for the timing of the annual rise in streamflow. It is possible, therefore, that strategies making use of local or west-wide moisture in both of its primary forms would achieve higher prediction skill,

particularly for schemes based on local moisture based. Also, local moisture schemes that weight

internal areas of a drainage basin, analogous to the traditional use of many separate measurement

stations within the basin, may reap added skill from the greater distribution of input information.

As a result, the relative advantages found in some areas from using west-wide predictors may be

smaller in comparison to strategies not considered here. Second, the index-based scheme uses

predictors that are slightly inter-correlated with each other (in contrast to the PCs). As a result, the

actual skill of the scheme may be lower when used for forecasting in practice than the correlations

found here suggest, and the PC-based scheme is likely to afford higher prediction skill even where

the results here appear to be comparable. The PC-based regression scheme requires more

sophistication to implement than would a regional index approach for forecasting, hence there may

be a tradeoff between ease of use and forecast benefit in preferring one scheme to the other. Lastly,

the apparent nonstationarity in predictors such as PC1 (Figure 1) suggest that, as in forecasting

from traditional predictors, a careful consideration of the historical period that is used in training

the prediction schemes is warranted.

5.   REFERENCES

Andreadis, K.M., E.A. Clark, A.W. Wood, A.F. Hamlet, and D.P. Lettenmaier, 2005, 20th Century
    Drought in the Conterminous United States , J. Hydrometeor., 6(6): 985-1001.

Bartlein, P.J., 1982, Streamflow anomaly patterns in the U.S.A. and southern Canada-1951-1970},
    J. Hydrol. 57, 49-63.

Cayan, D. R., M. D. Dettinger, K. T. Redmond, G. J. McCabe, N. Knowles, and D. H. Peterson,
    2003: The transboundary setting of California's hydropower systems: Linkages between Sierra
    Nevada, Columbia, and Colorado hydroclimates. Climate and Water: Transboundary
    Challenges in the Americas, H. F. Diaz and B. J. Morehouse, Eds., Kluwer Academic, 237–262.

Cayan, D.R., 1996, Interannual climate variability and snowpack in the western United States, J. Climate 9, 928-948.

Cayan, D.R., K.T. Redmond, and L.G. Riddle, 1999, ENSO and hydrologic extremes in the western United States, J. Climate 12, 2881-2893

Cayan, D.R., Kammerdiener, S.A., Dettinger, M.D., Caprio, J.M. and D.H. Peterson, 2001. Changes in the onset of spring in the western United States, Bull. Amer. Met. Soc. 82, 399-415.

Christensen, N.S., Wood, A.W., Voisin, N., Lettenmaier, D.P. and R.N. Palmer, 2004, Effects of climate change on the hydrology and water resources of the Colorado River Basin, Clim. Change 62, 1-3, 337-363.

Cook, E.R., D.M. Meko, D.W. Stahle and M.K. Cleaveland, 1999, Drought reconstructions for the continental United States, J. Climate 12, 1145-1162.

Derksen, C., K. Misurak, E. LeDrew, J. Piwowar, and B. Goodison, 1997, Relationship between snow cover and atmospheric circulation, central North America, winter 1988, Annals of Glaciology 25, 347-352.

Dettinger, M. D., D. R. Cayan, H. F. Diaz, and D. M. Meko, 1998: North–south precipitation patterns in western North America on interannual-to-decadal timescales. J. Climate, 11, 3095–3111.

Gershunov, A., 1998, ENSO influence on intraseasonal extreme rainfall and temperature frequencies., Journal of Climate, 11:3192-3203.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane. 2001. Current approaches to seasonal-to-interannual climate predictions. Int. J. Climatology, 21, 1111-1152.

Guetter, A.K. and K.P. Georgakakos, 1993, River outflow of the conterminous United States, 1939-1988, Bull American. Meteorol. Soc. 74, 1873-1891.

Hamlet A.F., P.W. Mote, M.P. Clark, and D.P. Lettenmaier, 2005. Effects of temperature and precipitation variability on snow pack trends in the western U.S. J. Clim., 18, 4545-4561.

Hong, S-Y and E. Kalnay, 2000, Role of sea surface temperature and soil-moisture feedback in the 1998 Oklahoma–Texas drought, *Nature* 408, 842-844, doi:10.1038/35048548.

Jain, S., M. Hoerling And J. Eischeid, 2005. Decreasing reliability and increasing synchroneity of western North American streamflow, J. Climate 18, 613-618.

Kevin E. Trenberth and David P. Stepaniak: J. Climate, 14, 1697-1701.

Keyantash, J. A. and J. A. Dracup, 2004. An Aggregate Drought Index: Assessing Drought Severity Based on Fluctuations in the Hydrologic Cycle and Surface Water Storage, Water Resources Research, 40, W09304, doi:10.1029/2003WR002610.

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

Liang, X., Lettenmaier, D.P., Wood, E. and Burges, S.J. (1994) A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research 99(D7): 14415-14428.

Liang, X., Lettenmaier, D.P. and Wood, E.F. (1996) One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model. Journal of Geophysical Research 101(D16): 21403-21422.

Lins, H.F., 1997, Regional streamflow regimes and hydroclimatology of the United States, Water Resour. Res. 33, 1655-1667.

Maurer E.P., D.P. Lettenmaier and N.J. Mantua, 2004, Variability and potential sources of predictability of North American runoff Water Resour. Res., Vol. 40, No. 9, Art. No W09306, September.

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier and B. Nijssen, 2002. A long-term hydrologically-based data set of land surface fluxes and states for conterminous United States, J. Clim. 15(22), 3237-51.

Maurer, E.P., S. Gibbard, and P.B. Duffy, 2006, Amplification of streamflow impacts of El Niño by increased atmospheric greenhouse gases, Geophysical Research Letters, Vol. 33, No. 2, L02707 10.1029/2005GL025100

McCabe G.J. and M.D. Dettinger, 1999, Decadal variations in the strength of ENSO teleconnections with precipitation in the western U.S., Intl. J. of Climatology, 19, 1399-1410

McKee, T.B., N. J. Doesken, and J. Kliest, 1993, The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference of Applied Climatology, 17-22 January, Anaheim, CA. American Meterological Society, Boston, MA. 179-184.

Mote, P.W., A.F. Hamlet, M.P. Clark and D.P. Lettenmaier, 2005. Declining mountain snow pack in western North America. BAMS, 86, 39-49.

Pagano T. and D. Garen, 2005a, A recent increase in western U.S. streamflow variability and persistence. J. Hydrometeorology 6, 173-179.

Pagano, T and D. Garen, 2005b, Integration of climate information and forecasts into Western US water supply forecasts, *Climate Variations, Climate Change, and Water Resources Engineering*, J.D. Garbrecht and T.C. Piechota, Eds., ASCE, 86-103.

Payne, J.T., A.W. Wood, A.F. Hamlet, R.N. Palmer and D.P. Lettenmaier, 2004, Mitigating the effects of climate change on the water resources of the Columbia River basin, Clim. Change Vol. 62, 1-3, 233-256.

Piechota, T.C., and J.A. Dracup, 1999. Long Range Streamflow Forecasting Using El Niño-Southern Oscillation Indicators. Journal of Hydrologic Engineering, 4(2), 144-151.

Piechota, T.C., J.A. Dracup, and R.G. Fovell, 1997, Western US streamflow and atmospheric circulation patterns during El Niño-Southern Oscillation, J. Hydrol. 201, 249-271.

Quan, X., M. Hoerling, J. Whitaker, G. Bates, and T. Xu, 2006: Diagnosing sources of U.S. seasonal forecast skill. J. Climate, 19, 3279-3293.

Ropelewski, C.F. and P.D. Jones, 1987, An extension of the Tahiti-Darwin Souther Oscillation Index, Mon. Wea. Rev. 115, 2161-2165.

Tang, Q., Wood, A.W. and D.P. Lettenmaier, 2006, An index-station based approach for real-time hydrologic simulation in the continental U.S., J. of Hydrometeorology (in preparation).

Trenberth, K. E., and D. P. Stepaniak, 2001, Indices of El Niño evolution, J. Climate, 14, 1697-1701.

Trenberth, K. E., 1997, The Definition of El Niño, Bull. Amer. Met. Soc., 78, 2771-2777.

Van Rheenen,N.T., A.W. Wood, R.N. Palmer and D.P. Lettenmaier, 2004, Potential Implications of PCM Climate Change Scenarios for Sacramento - San Joaquin River Basin Hydrology and Water Resources, Clim. Change Vol. 62, Issue 1-3, 257-281.

Wallace, J.M. and D.S. Gutzler, 1981, Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter, Mon. Weather Review 109, 784-812.

Wittrock, V. and E.A. Ripley, 1999, The predictability of autumn soil moisture levels on the

    Canadian prairies, International J. of Climatology, 19, 271-289.

Wood, A.W. and D.P. Lettenmaier, 2006, A testbed for new seasonal hydrologic forecasting

    approaches in the Western U.S., *Bull. Amer. Meteorol. Soc.*, (2007), 87:12,

    DOI:10.1175/BAMS-87-12-1699, 1699-1712.

Wood, A.W., A. Kumar and D.P. Lettenmaier, 2005, A retrospective assessment of NCEP climate

    model-based ensemble hydrologic forecasting in the western U.S. J. Geophys. Res. 110 (D4).

Wood, A.W., Akanda, A.S., Andreadis, K., and D.P. Lettenmaier, 2007, The Surface Water

    Monitor: a real-time land surface hydrologic simulation for monitoring and prediction of

    droughts in the continental US., *Bull. of the Amer. Met. Soc.* (in preparation).

LIST OF FIGURES

9   Value of five linear regression based schemes for predicting summer streamflow (April-July) at the Rio Grande River at Lobatos, CO location, based on the 1950-1997 period. Statistics shown are cross-validated R2 (top) and mean relative absolute error (bottom).

10   Value of five linear regression based schemes for predicting summer streamflow (April-July) at the Colorado River at Lees Ferry, CO location, based on the 1950-2004 period. Statistics shown are cross-validated R2 (top) and mean relative absolute error (bottom).

11   Value of five linear regression based schemes for predicting summer streamflow (April-July) at the Feather River at Oroville Reservoir, CA location, based on the 1950-2005 period. Statistics shown are cross-validated R^2 (top) and mean relative absolute error (bottom).

(a) SM+SWE principal component 1 loadings

(b) SM+SWE principal component 1 time series
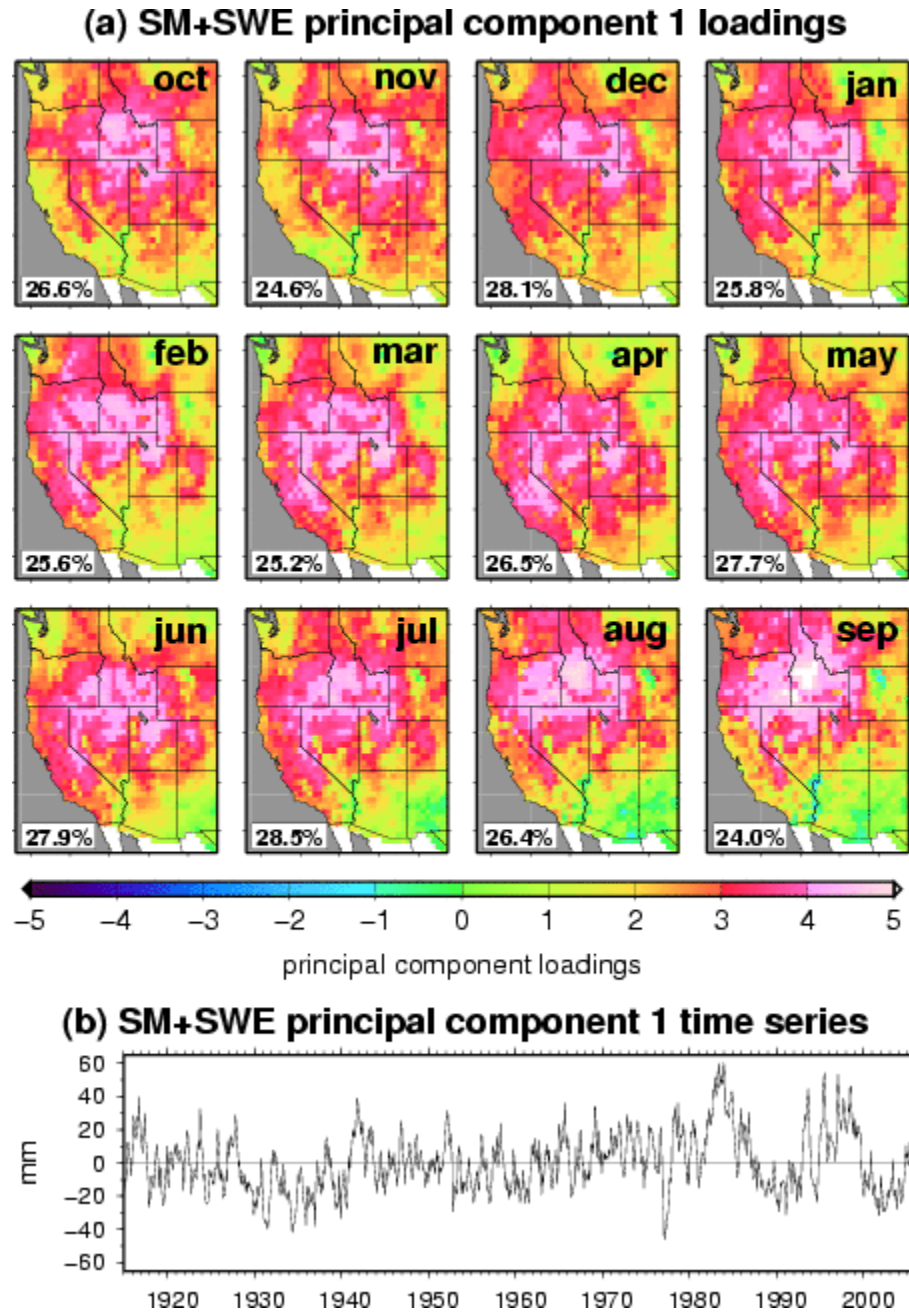
Figure 1 The first PC of western US moisture: (a) spatial loadings, with percent variability explained inset; and (b) timeseries.
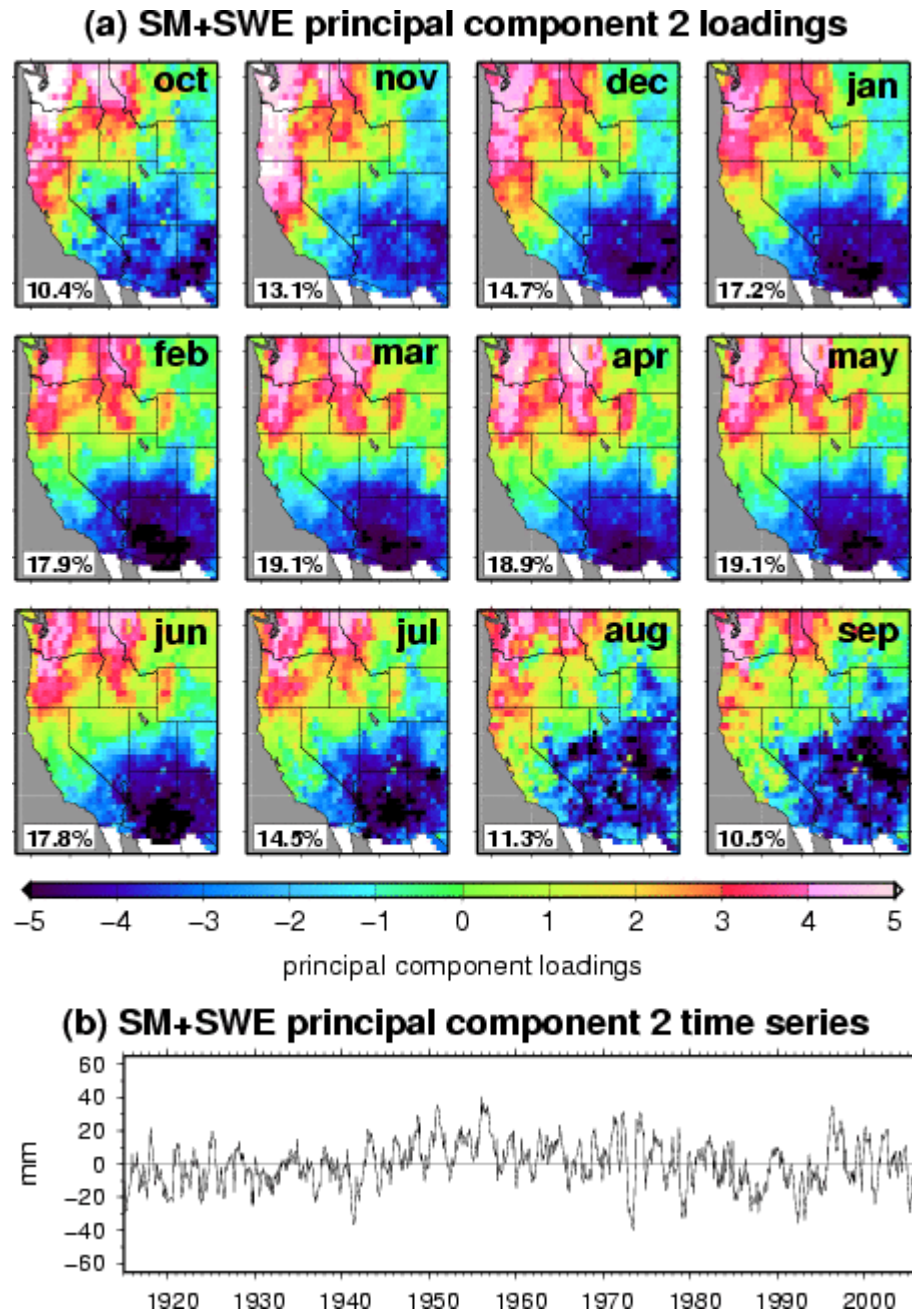
Figure 2 The second PC of western US moisture: (a) spatial loadings, with percent variability explained inset; and (b) timeseries.
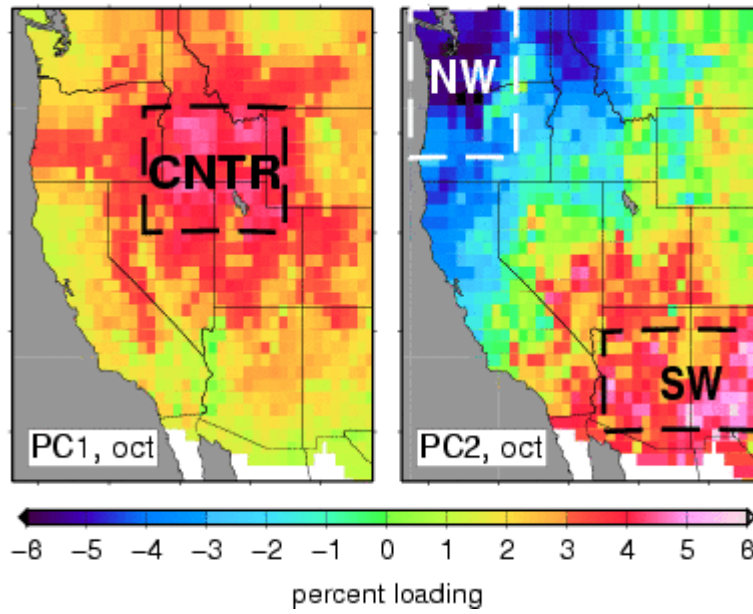
(a) SM+SWE principal component 3 loadings

(b) SM+SWE principal component 3 time series
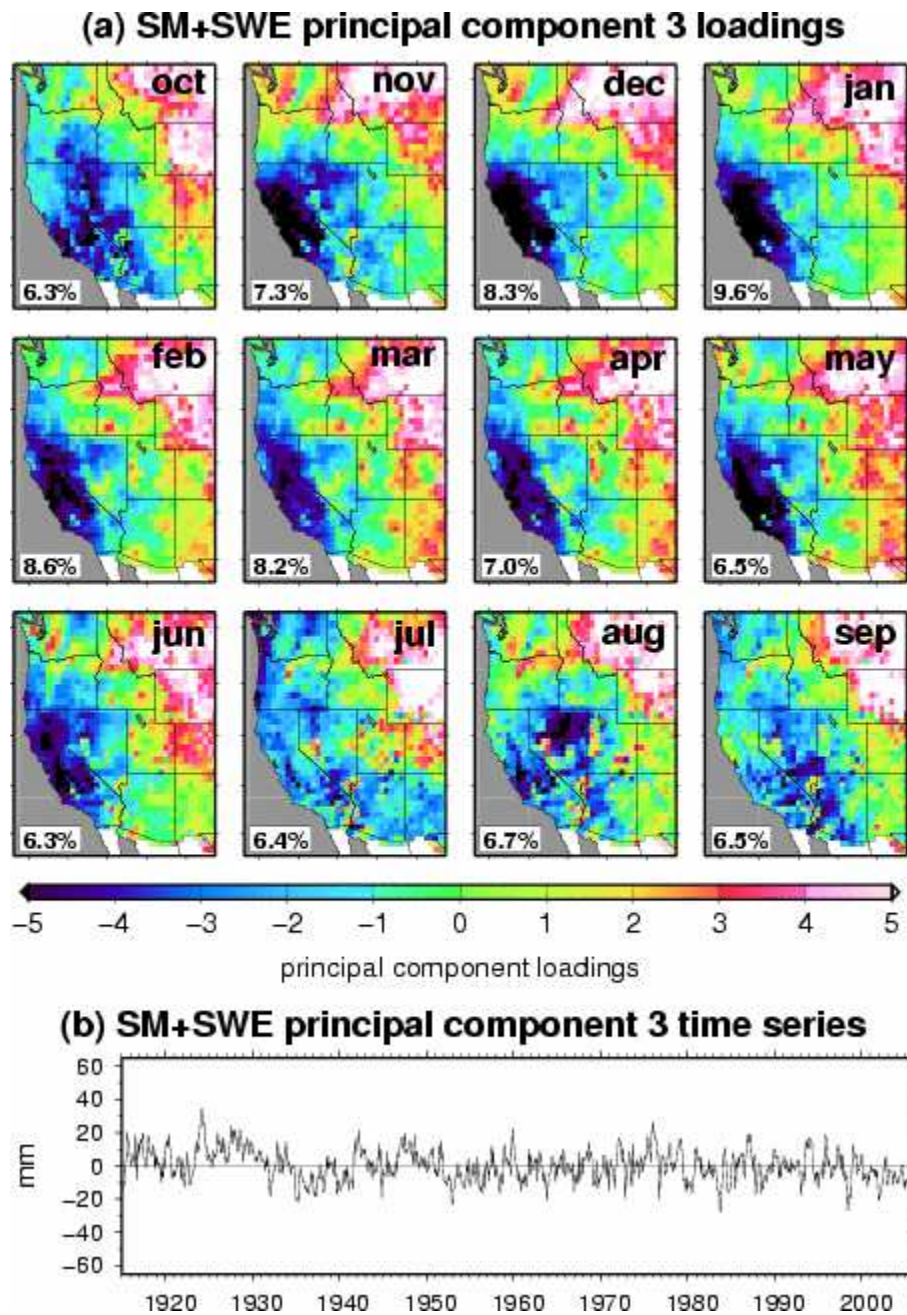
Figure 3   The third PC of western US moisture: (a) spatial loadings, with percent variability explained inset; and (b) timeseries.

Figure 4 Definition of index sub-regions based on the first and second PC loading patterns (for reference, the October PC loading patterns are shown).
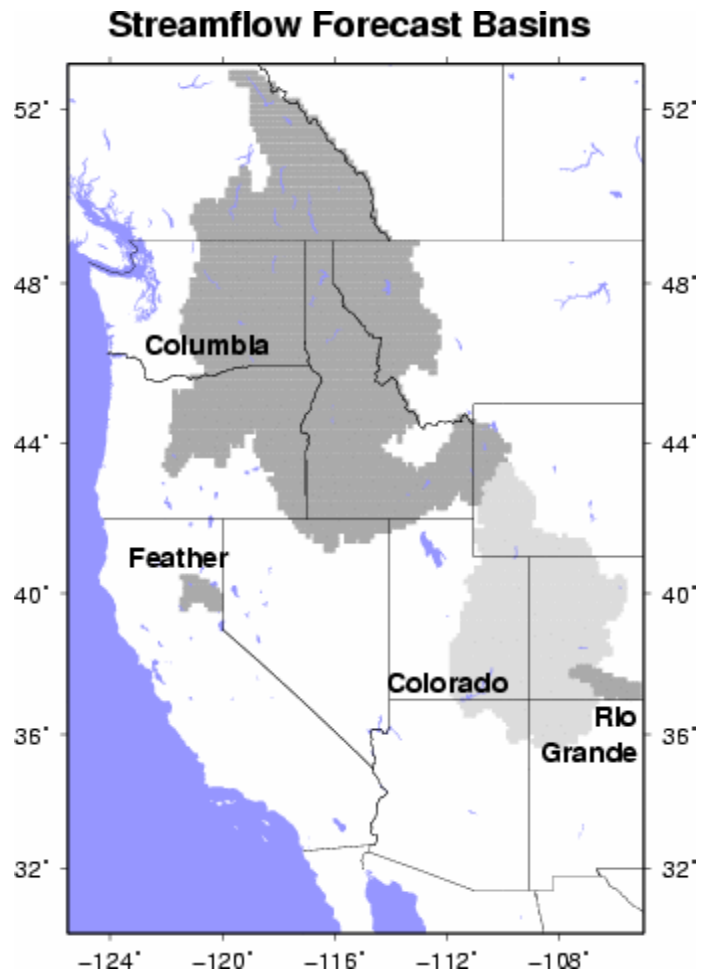
Figure 5 Four streamflow forecasting basins

Figure 6   Time series of predictors plotted against the summer flow (April-July) for the Columbia River at the Dalles, Oregon. The predictors are first two principal components (PC1 and PC2), their associated indices (CNTR and NW-SW), local moisture within the Columbia River drainage basin, and the teleconnection index, Nino3.4 (TC). All time series are all normalized to the range 0-1 for comparison. The summer flow in each year is plotted as a bar during the months April through July, with a thin dotted line connecting the bars.
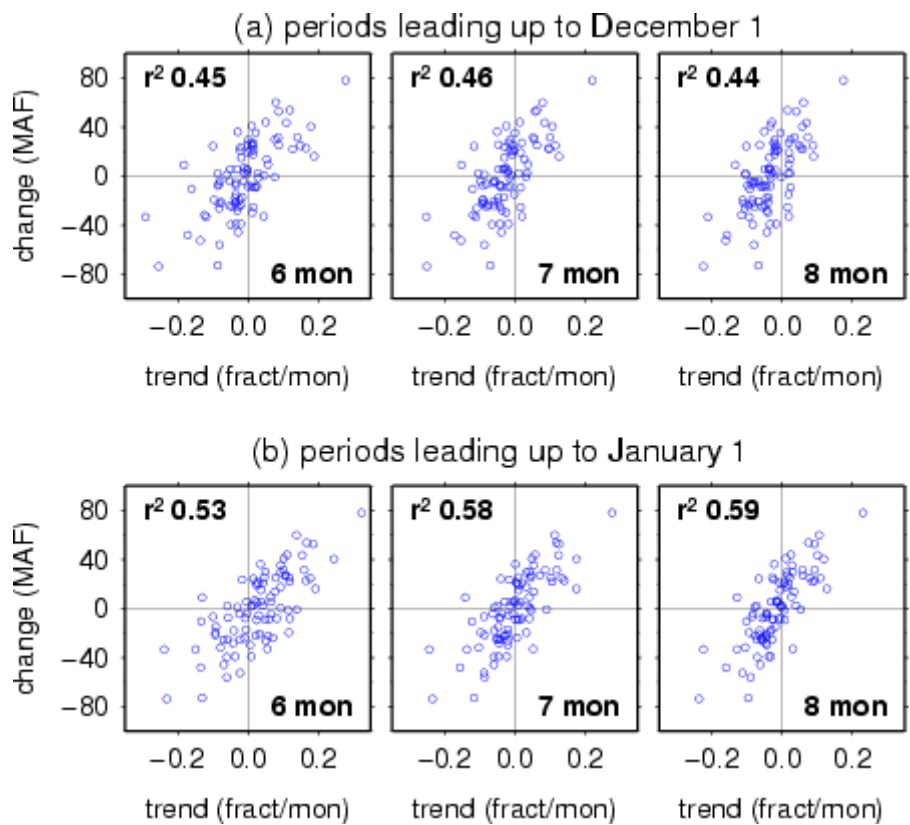
Figure 7   Comparison of trend in the second (i.e., NW-SW) index near the start of the water

near with the change in streamflow from the previous summer to the coming summer, for the

Columbia River at the Dalles, Oregon.

Value of Alternate Predictors for Summer Period Streamflow
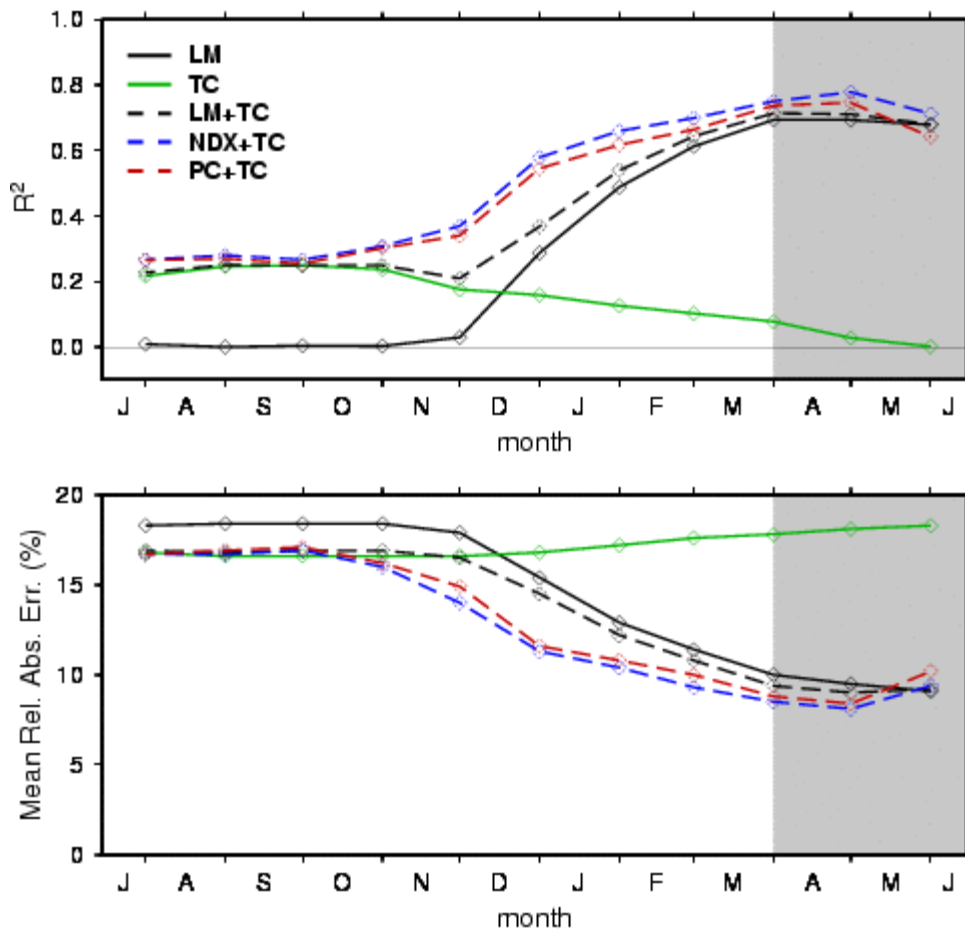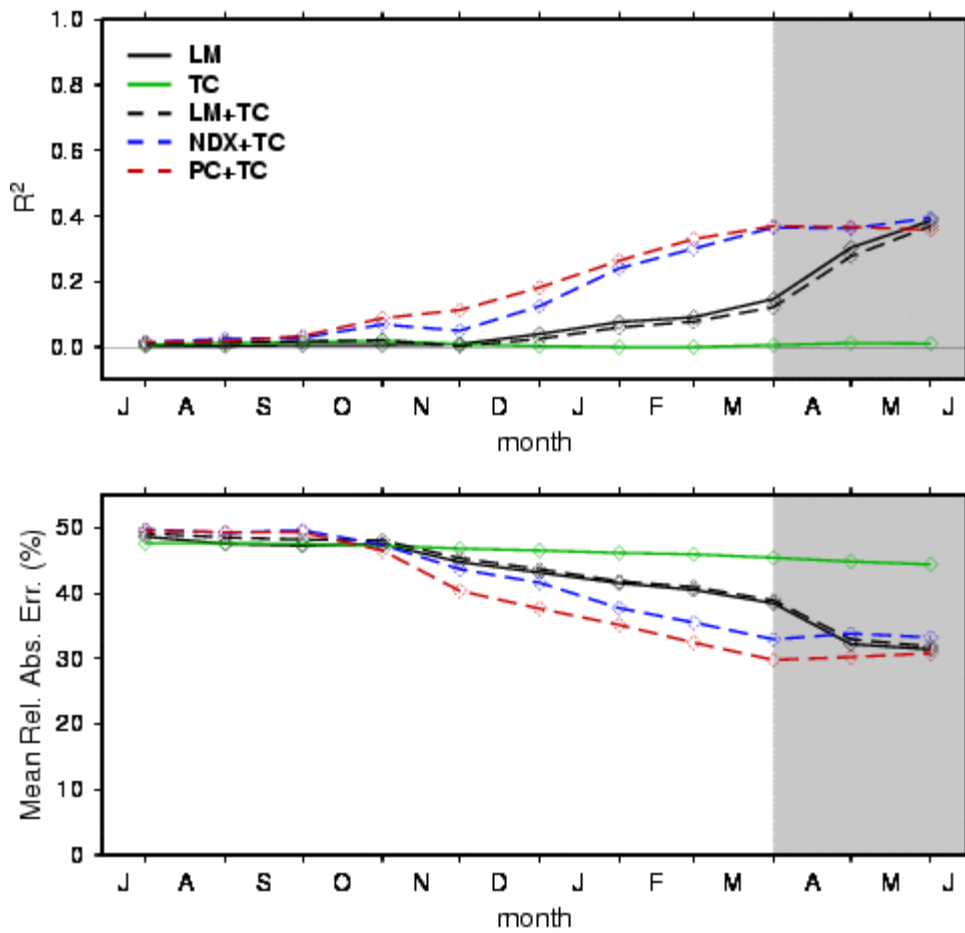*Columbia River at The Dalles, OR*

Figure 8 Value of five linear regression based schemes for predicting summer streamflow (April-July) at the Columbia River at the Dalles, OR location, based on the 1950-2005 period. Statistics shown are cross-validated R2 (top) and mean relative absolute error (bottom).

Figure 9 Value of five linear regression based schemes for predicting summer streamflow

(April-July) at the Rio Grande River at Lobatos, CO location, based on the 1950-1997 period.

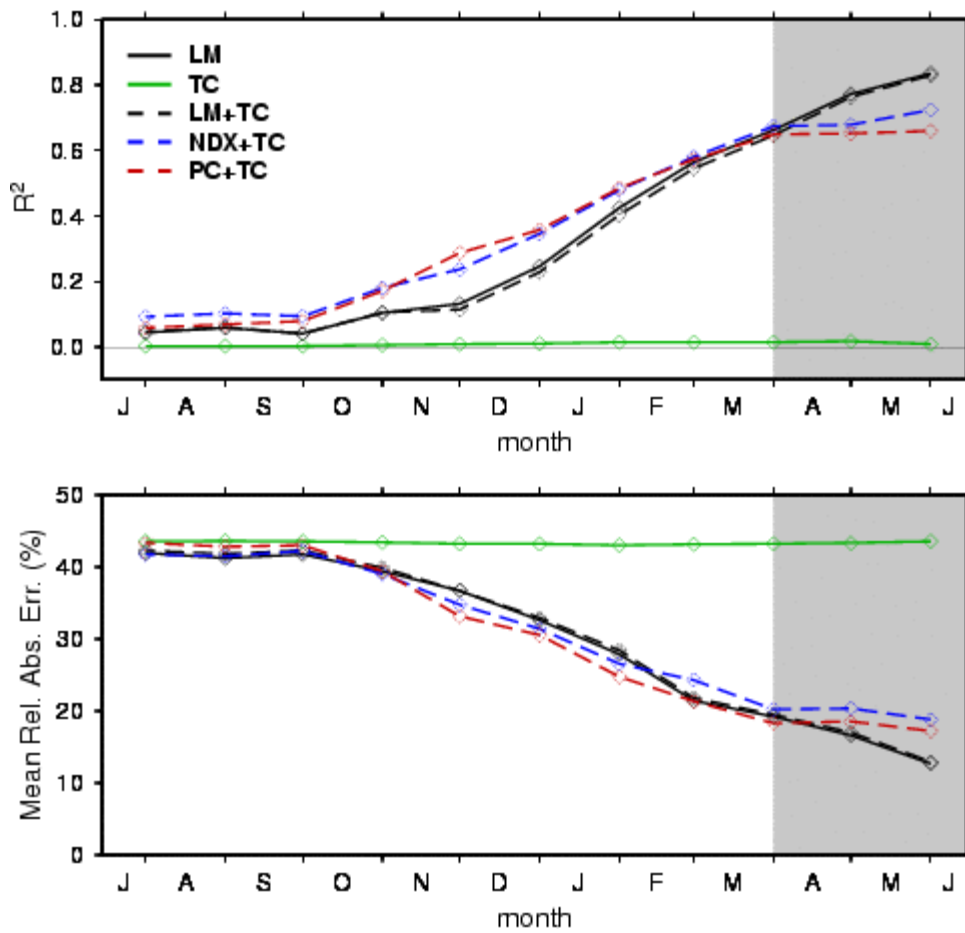Statistics shown are cross-validated R^2 (top) and mean relative absolute error (bottom).

Figure 10 Value of five linear regression based schemes for predicting summer streamflow (April-July) at the Colorado River at Lees Ferry, CO location, based on the 1950-2004 period. Statistics shown are cross-validated R^2 (top) and mean relative absolute error (bottom).
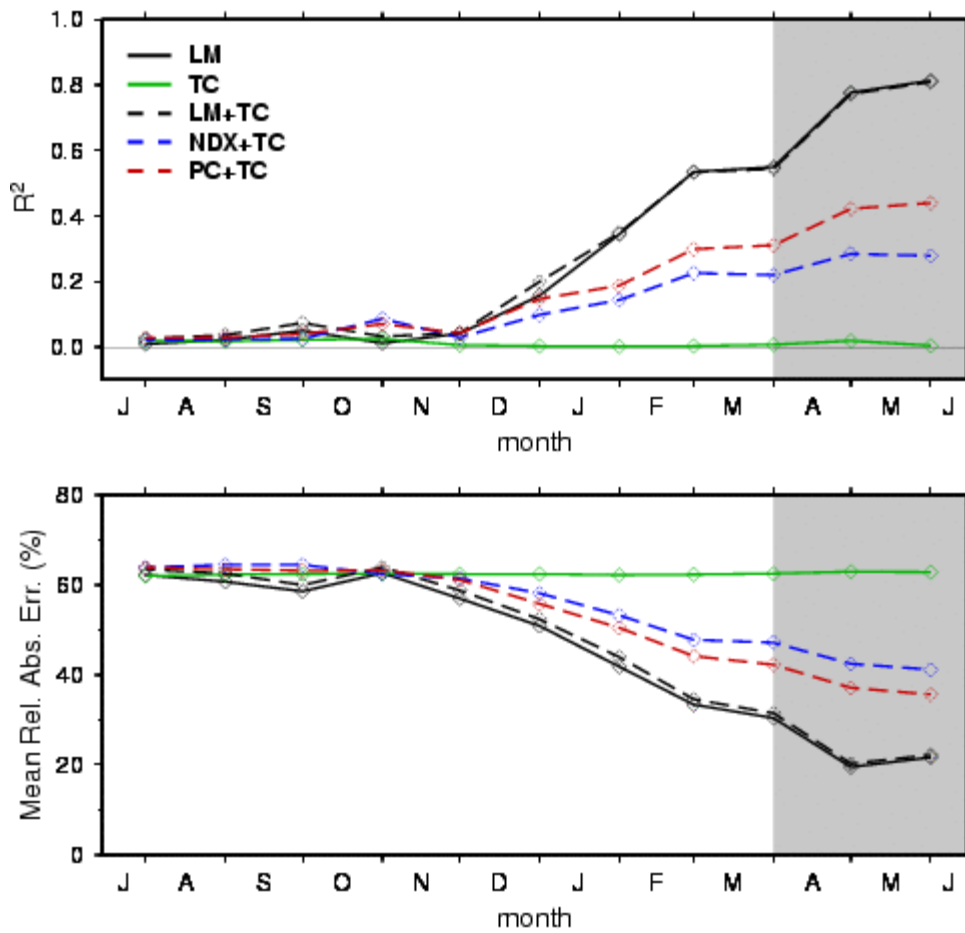
Figure 11   Value of five linear regression based schemes for predicting summer streamflow (April-July) at the Feather River at Oroville Reservoir, CA location, based on the 1950-2005 period. Statistics shown are cross-validated R^2 (top) and mean relative absolute error (bottom).

TABLES

Table 1   Linear regression equations for predicting summer streamflow volumes

description (label)	equation
teleconnection only (TC)	Q = a + b(TC)
local moisture only (LM)	Q = a + b(LM)
local moisture and teleconnection (LM+TC)	Q = a + b(LM) + c(TC)
principal components and teleconnection (PC+TC)	Q = a + b(PC1) + c(PC2) + d(TC)
indices and teleconnection (NDX+TC)	Q = a + b(CNTR) + c(NW-SW) + d(TC)

Table 2   Percentage contribution of PC2 to total variance explained by both PC1 and PC2.

RIVER FORECAST LOCATIONS

Forecast Date	Columbia	Rio Grande	Colorado	Feather
Nov 1	100	28	1	82
Dec 1	100	20	1	62
Jan 1	83	50	3	31
Feb 1	76	48	5	14
Mar 1	71	55	10	2
Apr 1	83	42	4	2
average	*85*	*41*	*4*	*32*