Basis for Extending Long-Term Streamflow Forecasts in the Colorado River Basin

Kenneth W. Lamb[1]; Thomas C. Piechota, M.ASCE[2]; Oubeidillah A. Aziz[3]; and Glenn A. Tootle, M.ASCE[4]

Abstract: The National Weather Service (NWS) maintains a collection of computer models used to perform various functions for managing the rivers of the United States. One function of the NWS's river forecast centers is to provide long-term resource forecasts for the main river basins in the United States. By using singular value decomposition (SVD) analysis, the research presented here identifies a new sea surface temperature (SST) index which demonstrates significant, long-lead covariance with streamflow in the Colorado River Basin. This index is compared with other existing climate indices by using the nonparametric rank sum test and by also using the index in a forecasting scenario.
DOI: 10.1061/(ASCE)HE.1943-5584.0000153.

CE Database subject headings: Streamflow; Forecasting; Water supply; Water temperature; Climates; Colorado River; River basins.

Author keywords: Streamflow; Forecasting; Water supply; Sea surface temperatures; Climate; Colorado River; Singular value decomposition; Rank sum.

Introduction

Recent research has provided many tools for the climate scientist to understand the connection among sea surface temperatures (SSTs) and sea level pressure (SLP) with surface climate over the United States (Latif and Barnett 1994; Hidalgo and Dracup 2003; McCabe and Dettinger 1999; Werner et al. 2002). Because of the recent drought conditions in the desert southwest, much of this research focuses on ways to describe the streamflow of the Colorado River Basin (CRB), by observing large-scale climate observations over the Pacific Oceans (Grantz et al. 2005; Kim et al. 2006; Kim et al. 2007). The well known relationships among SST/SLP data and streamflow of the western United States are a result of the El Niño Southern Oscillation (ENSO). However, many limitations to what information the ENSO can provide water managers of the CRB exist (McCabe and Dettinger 2002).

The first limitation is that the effects of the ENSO are most easily correlated with surface climate of the northwestern and southwestern areas of the United States. A large area covering central California, Nevada, Utah, Colorado, and Wyoming does not respond consistently to ENSO variability. The supposition is that at times, this region is influenced by northerly atmospheric conditions, and at other times, southerly conditions. Another limitation to the ENSO is that its effect on surface hydrology of the western U.S. has a relatively short lead time. For instance, McCabe and Dettinger (2002) found that fall-winter ENSO observations were correlated with snowpack measurements in the western U.S., and snowpack measurements are well correlated with spring runoff (Clark et al. 2001). Therefore, forecasts generated before a given water year, or during a non-ENSO cycle year, have a less successful forecasting ability of CRB streamflow.

Currently the National Weather Service's (NWS) Colorado River Basin River Forecast Center (CBRFC) issues three types of forecasts. These forecasts are then employed by the U.S. Bureau of Reclamation (USBR) offices in Denver and Boulder City to project the amount of water that will be stored or released from the river for the next 24 months. These forecasts include a three-month forecast of monthly water volumes, a four-month lump sum water volume for the April–July runoff period, and an "outlook" for the coming water year. The three-month forecasts are generated each month, the April–July forecast is issued monthly from January 1 to July 1, and the yearly outlook is issued only twice per year, August 1 and October 1. Although all of the forecasts are generated by using the CBRFC's Ensemble Streamflow Prediction (ESP) software, there is less confidence in the water year outlooks owing to the lack of research supporting 12-month forecasts (NWS CBRFC staff, personal communication, 2008).

The 24-month projections use the CBRFC's ESP forecasts as the primary input to their 24-month projections (referred to by USBR staff as the 24-month study) for the first 12 months of their analysis. The second 12 months or the "out" year, relies upon the 30-year average (1976-2005) streamflow values to continue their release/storage projections (USBR staff, personal communication, 2009).

One limitation of this method for long-term forecasting is the use of the 30-year average monthly data. Using the 30-year average for model calibration assumes a certain amount of climate stationarity, a concept which lately has been challenged in light of global climate trends (Milly et al. 2008). Another limitation of this method is that the NWS forecast and the 24-month study are deterministic forecasts. From a practical standpoint, this indicates the actual

[1]Civil Engineering Dept., Cal Poly Univ., Pomona, 3801 W. Temple Ave., Pomona, CA 91767; formerly, Dept. of Civil and Environmental Engineering, Univ. of Nevada, Las Vegas, 4505 Maryland Parkway, PO Box 454015, Las Vegas, NV 89154-4015. E-mail: kwlamb@csupomona.edu

[2]Associate Vice President for Interdisciplinary Research, Univ. of Nevada, Las Vegas, 4505 Maryland Parkway, Box 451087, Las Vegas, NV 89154-1087 (corresponding author). E-mail: thomas.piechota@unlv.edu

[3]Dept. of Civil and Environmental Engineering, Univ. of Tennessee, 73F Perkins Hall, Knoxville, TN 37996. E-mail: oaziz@utk.edu

[4]Dept. of Civil and Environmental Engineering, Univ. of Tennessee, 73F Perkins Hall, Knoxville, TN 37996. E-mail: gtootle@utk.edu

Note. This manuscript was submitted on June 2, 2009; approved on May 6, 2010; published online on June 7, 2010. Discussion period open until May 1, 2012; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, Vol. 16, No. 12, December 1, 2011. ©ASCE, ISSN 1084-0699/2011/12-1000–1008/$25.00.

probability of each is not reported, therefore the quality of the forecast remains unknown. A final limitation is that the forecasts created by the ESP view each of the historical inputs as equally likely in the current forecast.

This paper proposes that one improvement to generating long-term forecasts can be found by weighing the historical input data used by the ESP. Currently, after the ESP generates its forecast traces for the forecast period, certain years can be assigned more significance than others on the basis of how similar the forecast year is to the calibrated years. For example, if an El Niño event is taking place during the water year, El Niño years from the historical record can be assigned more weight (e.g., Tootle et al. 2009). Although this postanalysis weighting is sometimes performed, this information is not used to generate the official ESP forecast (CBRFC, personal communication, 2009). However, developing a forecasting approach in which the streamflow traces are weighted postanalysis provides a greater promise for implementation because the CBRFC are already acquainted with the method. The key to finding an objective weighting measure is to identify a long-lead climate indicator which provides a better indication to changes in streamflow than the 30-year mean. The principle goal of this paper is to establish a long-lead teleconnection among existing climate indices or Pacific Ocean SST's and CRB streamflow. A physical description of a long-lead teleconnection will be the subject of future research on the basis of the results from the current paper.

Much has been discovered about the connection of observed climate indices with the current water year or two-season lag time among the climate and runoff measurements (Grantz et al. 2005). The current research builds upon recent studies showing the correlation of large-scale climate index phases and streamflow or snowpack measurements in the Colorado River Basin (Tootle et al. 2005; Clark et al. 2001; Werner et al. 2002). The difference between this study and existing work (Timilsena et al. 2007; Woodhouse 2001) is that measured data will be used to ascertain the effect of large-scale climate on unimpaired/naturalized CRB streamflow data, as opposed to using tree-ring reconstructed streamflow data sets. Another departure from existing research (Werner et al. 2002) is that this study will be analyzing the correlation of lead climate indices and the following water year's volume over multiple lag times.

Description of Data

A total of six data sets were used in this analysis: naturalized streamflow, observed monthly southern oscillation index (SOI), monthly Pacific decadal oscillation (PDO) index, monthly Northern Atlantic oscillation (NAO) index, monthly Atlantic multidecadal oscillation (AMO) index, and sea surface temperatures.

Naturalized Streamflow

The USBR maintains a database of naturalized or unimpaired runoff data for the CRB and is available for download from the following URL (http://www.usbr.gov/lc/region/g4000/NaturalFlow/current.html). Streamflow gauges are considered naturalized when the effects of human interference (e.g., dams and diversions) have been removed (mathematically) from the gauge measurements. For the readers reference, there are two other kinds of gauges data, namely, unimpaired and impaired. Unimpaired gauges are those which, for all intents and purposes, measure the natural flowrate of the river. Impaired gauges measure streamflow which is affected by upstream storage or diversions. In all, 29 streamflow gauges within the upper and lower CRB, which report naturalized streamflow from 1905 (October 1905–September 1906) thru the 2004 water year, were analyzed. This particular dataset was chosen for this analysis because the USBR uses this data as input into their 24-month study, and it is referenced as the baseline for all forecasts of the CRB. For example, a given CBRFC seasonal forecast is often reported as a percentage of the 30-year average (1971-2000) streamflow volume using this data. Fig. 1 shows a map of the 29 river gauge locations.

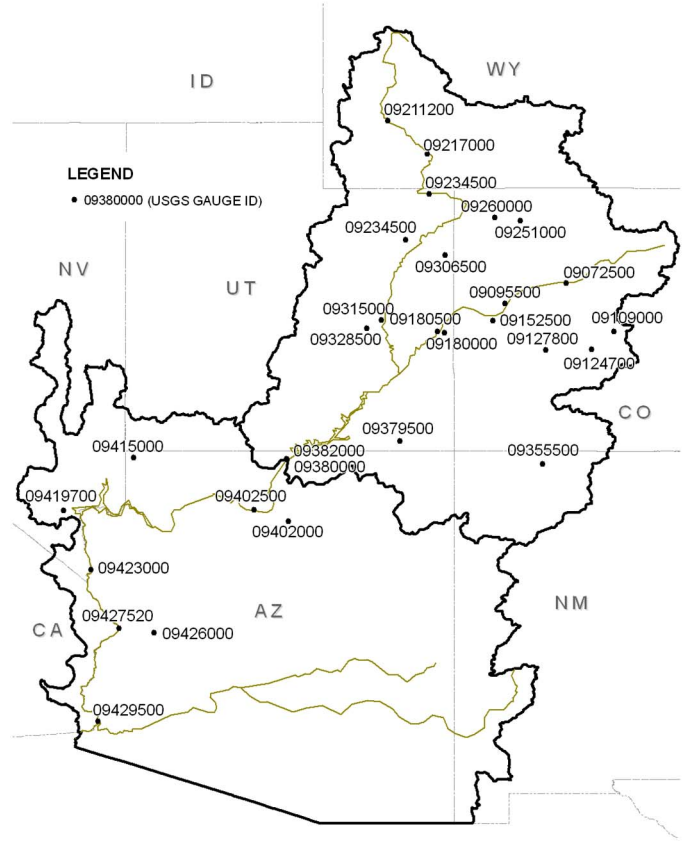

Fig. 1. Location map showing 29 river gauge locations used in this study

SST and Climate Index Data

Average monthly SST values were obtained from the National Climatic Data Center (http://www.cdc.noaa.gov/cdc/data.noaa.ersst.html). The data is presented in $2° \times 2°$ grid cell for all oceans (Smith and Reynolds 2003; Smith et al. 2008) and spans the years from 1846 to the current year. Although the dataset is a reconstruction of SST values for the period, data quality improves from 1950 on because of the increase in coverage of ocean monitoring capability. The range of ocean data used in this study covers the area from longitudes 120°E to 80°W and latitudes 20°S to 60°N, which is similar to the region studied by Tootle et al. (2006).

Climate index data can be found through the National Oceanic Atmospheric Administration's (NOAA) Climate Prediction Center (CPC) website. These indices have been used previously to demonstrate a correlation among the changes in the index and surface climate changes (Tootle et al. 2005; Shabbar et al. 1997; Barlow et al. 1993; Rassmussen & Wallace 1983; McCabe et al. 2007). The availability of continuous data for each index and for the streamflow data defines the study period. The SOI data (http://www.cpc.ncep.noaa.gov/data/indices/soi) covers monthly standardized values from 1951 to 2008 calendar years. PDO index values are available from 1900 to the current year (http://jisao.washington.edu/ pdo/PDO.latest). In addition to these data sets, the available

NAO (http://www.cdc.noaa.gov/Correlation/nao.data) and AMO (http://www.cdc.noaa.gov/Correlation/amon.us.data) index values is available from 1950 and 1948, respectively, until the current year. On the basis of the data availability, the study period starts in 1951 and ends with the 2005 water year.

Methodology

The correlation between the climate variability and SST with the gauge streamflow was completed by using two statistical methods. The first test is the rank sum test. This test is used to compare mean values from two different datasets to determine if these are significantly different. The second test is the singular value decomposition (SVD) method of finding maximum covariance between two fields. SVD was used to compare the SST data with the streamflow data.

Rank Sum Analysis

The rank sum test is an excellent way to test a "before" and "after" scenario without trying to determine if either data set has a defined distribution. Alternatively, it can be used to test for differences among two segmented data sets. Hence, this is a nonparametric test and does not make assumptions about the distributions of either dataset. In the current analysis, we are using the negative and positive phases of the indices to act as our "before" and "after conditions, respectively. The mean of the streamflow measured during a negative climate phase is compared to the mean streamflow measured during the positive climate phase. The null hypothesis of this test is that the means of the two sample sets are equal. The results of a rank sum test provide a value of the probability that the two sets are similar. For p-values less than 0.05, the null hypothesis was rejected. A theoretical discussion of this test is offered by Alder and Roessler (1977).

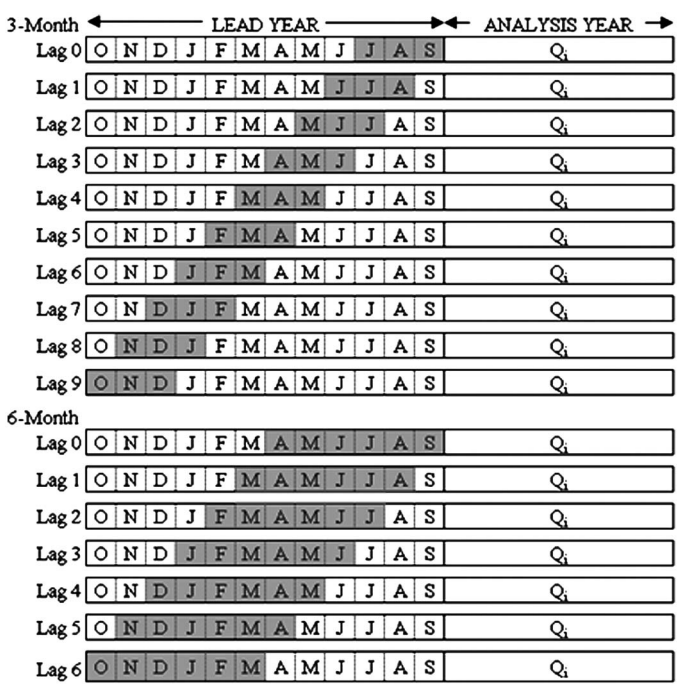

Fig. 2. Depiction of three and six-month analysis comparing lagged climate index and SST values with streamflow values from the following water year; Q_i = yearly total runoff measured at station i

To determine the effect of SOI on CRB streamflow, the three- and six-month average SOI values were used to separate the historical streamflow data into "positive" and "negative" groups. Fig. 2 shows an overview of how the rank sum test uses the historical index values to compare the positive and negative streamflow phase groups. Starting with July, August, and September (JAS) average SOI values (lag = 0), the streamflow values from years in which the monthly interval is positive were grouped and compared with those streamflow values corresponding to negative SOI values. This test was performed for the measured streamflow values from each station. After each station was compared for the JAS average of SOI values, then the three-month SOI interval was lagged back by one month (lag 1: June, July, and August) and the analysis was performed again at all measurement stations. The analysis process was repeated until lag 9 (average of October, November, and December SOI values). The three-month average rank sum test was repeated for the PDO (index number 2), NAO (number 3), and AMO (number 4) indices, which provided a total of 1,160 results (29 stations × 10 lags × 4 indices). Similarly, the same process was performed for the six-month average values, starting with the April-September interval [April, May, June, July, August, and September (AMJJAS), lag = 0] and ending with the October-March interval [October, November, December, January, February, and March (ONDJFM), lag = 6] (Fig. 1). The six-month analysis consisted of an additional 812 rank sum tests (29 stations × 7 lags × 4 indices).

Combination of Indices

The next step in the analysis was to compare the combined effect of the indices, assuming that the influence from different climatic phenomena can occur at the same time. The process used to determine the effect of an individual climate index was repeated for the combination analysis. This process is shown graphically in Fig. 3. By using the SOI as a base, the combination analysis compared the PDO-positive and PDO-negative values which also occur under an SOI-positive phase (labeled as analysis number 5). Likewise, the process was repeated for PDO-positive versus PDO-negative values under the SOI-negative phase (analysis number 8). The SOI-PDO combination of three- and six-month average values included an additional 986 rank sum tests (29 stations × 10 lags × 2 combination phases + 29 × 7 × 2). In a similar manner, the NAO (analysis numbers 6 and 9) and AMO values (analysis numbers 7 and 10) were also completed for the three- and six-month values. Therefore, the combination analyses yielded a total of 2,958 results (986 × 3 indices).

SVD Analysis of Pacific Ocean SSTs

The next step in the analysis process was to determine how Pacific Ocean sea surface temperatures vary in relation to the streamflow measured in the 29 gauges of the CRB. The goal of this analysis was to determine if any SST pattern that can be identified, which is not already represented in one of the indices used in the rank sum testing. The process is also known as maximum covariance analysis

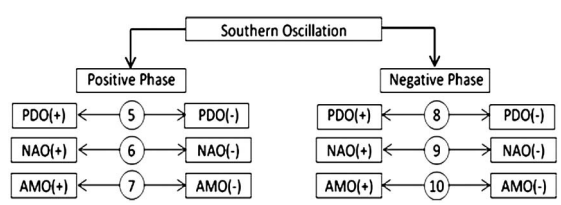

Fig. 3. Graphic showing the combination rank sum testing by using the SOI as a basis

because the goal is to identify the dimensions or modes which represent the maximum amount of covariance between two temporal fields. Because the spatial component of most climate data is quite large, using the temporal field to compare generally limits the analysis dimension to the number of days, months, or years included in the study period. In this paper the time scale employed is in years.

The use of SVD to find coupled patterns in climate data was explored by Bretherton et al. (1992), who compared five different methods for analyzing climate data. Other research (Wallace et al. 1992; Soukup et al. 2009) has utilized SVD to analyze the covariance of sea-level pressure and SST with snow-water equivalent measurements, precipitation, and, as done in this study, with streamflow. The process performed in this analysis is similar to Tootle and Piechota (2006), in which SST anomalies in the Pacific Ocean are coupled with streamflow anomalies across the United States. This study limits the analysis to the CRB streamflow only.

The following description of the SVD process is provided by several sources (Strang 1998; Bretherton et al. 1992; Wallace et al. 1992). Before any computations were performed, the raw SST and streamflow data were standardized on the basis of each grid location and the values representing land (values of -9999) were deleted from the SST matrix. Next, a covariance matrix by using the SST and streamflow anomalies was determined by using the following formula:

$$_{\mathrm{COVAR}}\mathrm{TS} = (1/n)(\mathbf{T'S})$$

In this analysis, the number of years in the temporal analysis is $n = 53$. The SST field $[53 \times 4771]$ was represented by the letter $\mathbf{T}$, and $\mathbf{S}$ represents the streamflow data field $[29 \times 53]$. The SVD was performed on the $_{\mathbf{COVAR}}\mathbf{TS}$ matrix $[4771 \times 29]$ which yielded the following three matrices:

$$\mathrm{SVD}(_{\mathrm{COVAR}}\mathrm{TS}) = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}$$

$\mathbf{U}[4771 \times k]$ and $\mathbf{V}[29 \times k]$ represent the left and right singular (heterogeneous) matrices respectively, and $\boldsymbol{\Sigma}[53 \times 29]$ represents the diagonal matrix of singular values from the SVD. The value k represents all modes (columns in the matrix) which contain a squared covariance fraction (SCF) $> 10\%$. SVD attempts to describe the variance of a given matrix (in this case, the $_{\mathrm{COVAR}}\mathrm{TS}$ matrix) by drawing new axes which are rotated in matrix space to capture as much of the variance (in essence, changes in SST or streamflow) as possible. Each column of the $\mathbf{U}$ and $\mathbf{V}$ represents a new axis and the values in those columns represent the variance captured by that axis. To measure the proportion of variance contained within a given column, Bretherton et al. (1992) used the SCF. The SCF is measured by dividing the square of each singular value in the $\boldsymbol{\Sigma}$ matrix by the sum of the squared singular values (sum of squared values from the diagonal), as is shown by the following equation:

$$\mathrm{SCF} = \frac{S_{i,i}^2}{\sum_{i=1}^{29} S_{i,i}^2}$$

In this equation, S represents the matrix of singular values $\boldsymbol{\Sigma}$ (substituted to avoid confusion with the summation sign in the denominator), and the summation is taken for values of i from 1 to 29 because this is the total number of columns of $\boldsymbol{\Sigma}$.

Finally, the left and right temporal expansion series (LTES and RTES, respectively) were obtained by projecting $\mathbf{U}$ onto $\mathbf{T}$ and $\mathbf{V}$ onto $\mathbf{S}$ in the following manner:

$$\mathbf{LTES} = \mathbf{U'T'}$$

$$\mathbf{RTES} = \mathbf{V'S'}$$

The LTES $[53 \times k]$ and RTES $[53 \times k]$ expansion series are both time series representing the maximum yearly variance of the SST data. The LTES is then standardized to reflect the unit increments of variance. In this format, the LTES can now be used in the same way that the climate indices were used to test correlation with the streamflow in the rank sum test. The LTES is run through the lagged correlation testing outlined previously, just like the testing approach to the established climate indices.

The final step in the analysis was to determine the significance of the variance observed in the left heterogeneous field (SST) and the right heterogeneous field (streamflow) by calculating the Pearson-r correlation coefficient. The testing for significance was performed by using a student t-test, which assumes that the correlation coefficients are normally distributed, with a mean value of $r = 0$. Working backward, starting with $n - 2$ degrees of freedom (i.e., 51), the computation required that $|r| \geq 0.27$ to be significant. This value was calculated on the basis of $\alpha = 0.05$.

Forecast Skill Comparison

The final analysis proposed in this paper is used to compare the forecast skill abilities up to three climate indices at a time. In contrast to the rank sum analysis, which provides an assessment of significant dissimilarities between streamflow observed under the positive and negative phases of a given index, the forecast skill score will calculate a value which represents the magnitude of contribution that a given index makes to a given forecast model. The comparison of the three indices was completed by using a model initially developed by Piechota et al. (1998), who attempted to develop long-lead-time forecasts of streamflow in eastern Australia. A detailed explanation of the method is provided in Piechota et al. (1998) and in other more recent research (Tootle and Piechota 2004; Soukup et al. 2009). A brief explanation is provided here.

A probability density function is created by using a kernel density estimator and a sorted vector of observed streamflows at a given location. For each observed streamflow, a unique probability value is obtained by using a given predictor (SOI, PDO, and so forth). The streamflow/probability data pairs are then plotted to create a probability exceedance curve for each predictor. The three exceedance curves are combined to create a composite exceedance curve which is the basis for the forecast model. This new forecast model calculates weights a, b, and c, which are applied to the appropriate predictors to optimize the forecast model's linear error in probability space (LEPS) score. The LEPS score was initially developed by Ward and Folland (1991) and subsequently revised by Potts et al. (1996). The revised LEPS score is used in this paper because the formula proposed by Potts et al. (1996) gives a higher score to forecasts of extreme values than to forecasts closer to the observe mean.

Results

One limitation influences the rank sum test and its significant results. In the current analysis, a total of 11 climate indices (four climate indices, six combinations of the same, and the LTES) which yielded a total of 5,423 rank sum test results. Of these tests, a total of 828 significant test results were generated. On the basis of the chosen α for the analysis, 272 of the significant results $(5,423 \times 0.05)$ have a probability of representing type I errors. In other words, approximately $1/3$ of the significant results can come back as false positives. To address this issue, the total number of successful results at each streamflow gauge station was ranked according to the p-value resulting from the rank sum test. P-values close to

0.05 have the highest chance of a type I error, so only the two lowest p-values at each station were reported.

Climate Indices

Fig. 4 shows the total number of successful results from each climate input. By far, the most successful index input to the rank sum test is the LTES. In fact, the January-February-March LTES rank sum test resulted in significant results at all 29 stations. The next largest number of successful results came from the NAO and the AMO index tests. As the lag time in the analysis increases, the number of significant AMO test results also increases, suggesting a lagged relationship with CRB streamflow. Table 1 shows the top two index inputs at each gauge location. The SST index in this figure is the LTES resulting from the SST-streamflow SVD analysis for the given interval length and lag number.

These simple observations are not enough to draw hard conclusions, but these do identify additional questions: Why are the NAO and AMO out performing the SOI and the PDO? Why are the Pacific Ocean SSTs performing so well when the SOI and the PDO are not?

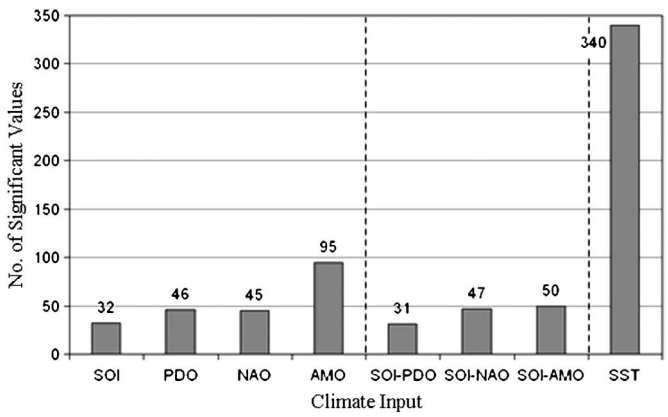

Fig. 4. Number of significant results of the rank sum test on the basis of climate input, recorded across the basin ($\alpha = 0.05$)

The answer to the first question is that this research confirms previous research in related areas. First, McCabe et al. (2007) analyzed the effects of the northern Atlantic, northern Pacific, tropical Pacific, and the Indian Ocean regions on streamflow in the upper CRB. Whereas this study's methods to determine a long-term correlation differ, McCabe et al. (2007) was able to show that although all SST input values were teleconnected to streamflow anomalies in the CRB, a PCA analysis showed that the AMO mode represented more of the variance than the PDO or Indian Ocean SST data. Secondly, researchers (McCabe and Dettinger 2002) have identified that SOI, NINO3, PDO and the Pacific/Northern America (PNA) circulation have weak correlations to streamflow in the upper CRB. Those who have observed this suppose that the typical climate of the upper CRB is not completely similar to either the Northwest or Southwest United States and therefore does not maintain the same consistent correlation signal with the tropical Pacific Ocean data.

As a result of the poor signal response Grantz et al. (2005) developed basin-specific climate indicators, as suggested by Yernall and Diaz (1986). If the established climate indices are not sufficient to determine strong correlation, then it is worth evaluating whether other regions of the Pacific may serve as predictors. It is along this line of reasoning that the second question was explored: Why does the temporal expansion series from the SVD analysis outperform the established climate indices?

Pacific Ocean SSTs

The SVD analysis provides a number of maps which provide a helpful amount of information about what might be a basin-specific climate index (Figs. 5 and 6). These figures are generated by plotting the significant correlation coefficients of the left (SST) and right (streamflow) heterogeneous matrices. Starting at lag = 0 (July-August-September and AMJJAS) a visible signal from the ENSO region of the Pacific diminishes as the lag number increases. Conversely, because the lag number increases, another area located east and south of Japan begins to grow until it reaches its maximum size at approximately lag = 6 for the three-month average interval, and lag = 4 for the six-month interval. The difference in sign between the ocean area east of Japan and the streamflow gauges

Table 1. Top Two Significant Rank Sum Test Results at Each Station

USGS					USGS				
Station number	Index/average month interval/lag				Station number	Index/average month interval/lag			
09072500	[1]	SST/3/8	[2]	NAO/3/3	09315000	[1]	SST/3/6	[2]	SP(p)/3/9
09095500	[1]	NAO/3/3	[2]	SST/3/8	09328500	[1]	SST/3/9	[2]	NAO/3/3
09109000	[1]	SST/3/8	[2]	SN(n)/3/2	09355500	[1]	SN(n)/3/2	[2]	AMO/3/8
09124700	[1]	SN(n)/3/2	[2]	NAO/3/3	09379500	[1]	SST/3/9	[2]	SN(n)/3/2
09127800	[1]	SN(n)/3/2	[2]	NAO/3/3	09380000	[1]	SST/3/6	[2]	NAO/3/3
09152500	[1]	SST/3/6	[2]	AMO/3/8	09382000	[1]	SST/3/7	[2]	SOI/6/2
09180000	[1]	SST/3/9	[2]	AMO/3/8	09402000	[1]	SST/3/0	[2]	AMO/3/9
09180500	[1]	SST/3/6	[2]	NAO/3/3	09402500	[1]	SST/3/6	[2]	NAO/3/3
09211200	[1]	SP(p)/3/9	[2]	SOI/6/6	09415000	[1]	SST/3/9	[2]	SOI/3/5
09217000	[1]	SP(p)/3/9	[2]	SST/3/6	09421500	[1]	SST/3/6	[2]	AMO/3/8
09234500	[1]	SP(p)/3/9	[2]	SST/3/6	09423000	[1]	SST/3/6	[2]	AMO/3/8
09251000	[1]	SST/3/8	[2]	NAO/3/3	09426000	[1]	SST/3/0	[2]	SOI/3/8
09260000	[1]	SST/3/6	[2]	SP(p)/3/9	09427520	[1]	SST/3/6	[2]	AMO/3/8
09302000	[1]	SST/3/6	[2]	SN(n)/3/2	09429490	[1]	SST/3/6	[2]	NAO/3/3
09306500	[1]	NAO/3/3	[2]	SST/3/6					

Note: SP(p) = PDO-positive versus PDO-negative under the SOI-negative phase; SN(n) = NAO-positive versus NAO-negative under the SOI-negative phase; SA(n) = AMO-positive versus AMO-negative under the SOI-negative phase. Shaded cells denote lower CRB gauges.

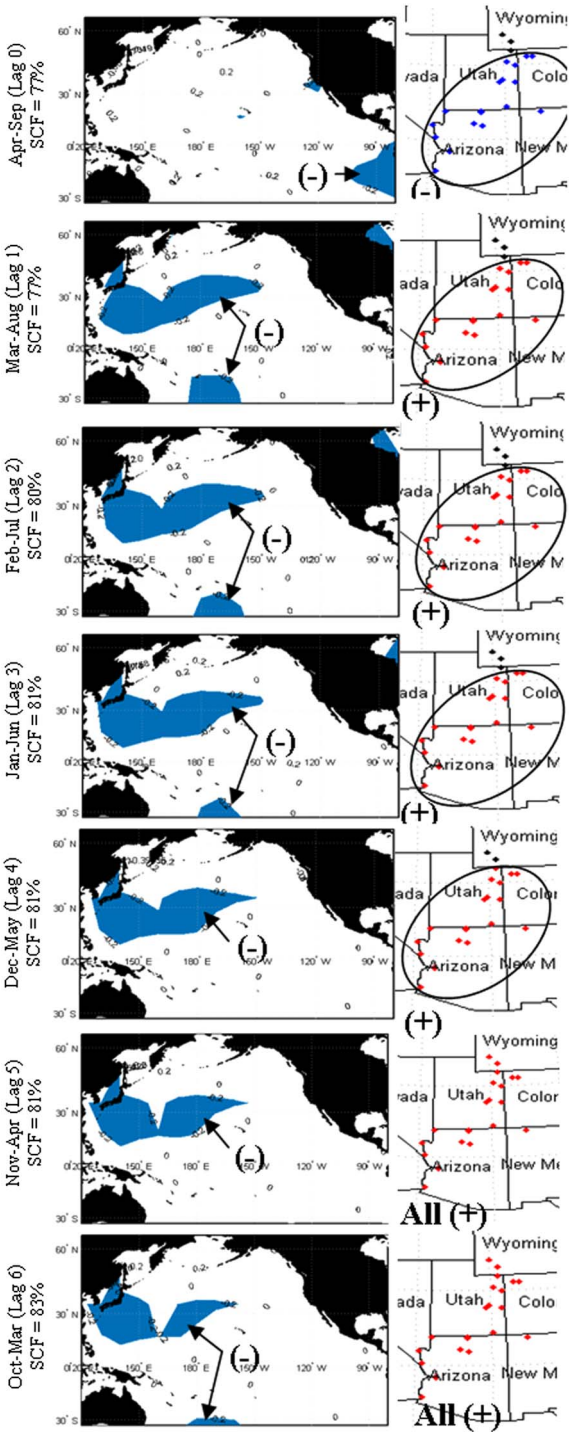

Fig. 5. Left and right heterogeneous correlation maps ($|r| \geq 0.27$) plotted from six-month average SST data, from lag 0 to lag 6; SCF values displayed for first mode

shows that they are negatively correlated, in which one is increasing because the other is decreasing.

The region in question is anchored over an area that this paper defines as 150 to 200°E (160°W) and from 24 to 34°N. This particular area is south and west of the area which is associated with the northern Pacific signals of the PDO as identified by Mantua and Hare (2002). However, this area is not original to the current research because it has been identified in other research without being specifically acknowledged or defined (Rajagopalan et al. 2000; Wang and Ting. 2000; Tootle and Piechota 2006; Soukup et al. 2009). Because the area is located adjacent to Japan, this paper refers to this area as the Hondo Region. *Hondo* is a Japanese word for "main road" or "highway". In Spanish, *hondo* is an adjective meaning "deep", and it is most commonly used when referring to water. Because the word is used in two languages which are geographically relevant to the research at hand and its meanings are relevant to the research, this region of the ocean lends itself to this designation.

The success of the LTES as an index can be explained through the way in which it is created and how it can be used. Because the LTES is the projection of the SST variance onto the normalized raw SST data, the result is a vector which represents the scaled SST variance within the given three- or six-month interval of time selected for the analysis. For all but three of the analysis periods (three-month, lags 0 and 1 and six-month lag 0), the overwhelming majority of the variance could be explained by the first mode, and is therefore captured as part of the LTES. The hypothesis of this study is that this occurs because the influence of other climate indices is stronger during these periods than the rest of the year. The LTES may seem less ideal to use because it is not simply the standardized SST values of the Hondo Region (similar to the PDO or AMO). As already shown, the LTES is the product of the SST and streamflow data, which creates another question: How is a forecast created by using a model that requires the identification of a forecast point? This study does not propose to answer this question here. However, because the Hondo Region is identified by an SVD analysis, it is best to continue discussing the SST-streamflow connection in the SVD output (in essence, the left temporal expansion series).

The teleconnection among the SSTs of the western Pacific and the streamflow may be explained from the perspective of the east Asian jet stream (EAJS). Zhang et al. (1997) demonstrated that the presence of high pressures over Siberia and cold surges from east Asia (traveling on a south-southeast direction from northern China), have a great effect on the winter monsoon season over east Asia. A strong east Asia winter monsoon is characterized by two high-pressure centers: the first is a warm high-pressure center whose midpoint is roughly east of Papua New Guinea, whereas the other center is a cold "dome adjacent to the Siberian region". Between the two high-pressure centers is the winter traveling path of the EAJS.

Athanasiadis et al. (2010) showed that the Hondo Region is actually the terminus of the EAJS (Asian/Pacific jet). By completing an EOF (i.e., PCA) of the average wintertime (December-March) wind field over the northern hemisphere, this study identified the path of the dominant Asian/Pacific jet. Fig. 4 of Athanasiadis et al. (2010) displays the dominant wintertime jet (shown in a heavy black line) on the basis of the EOF analysis superimposed upon the teleconnectivity of the 250 hPa zonal wind and the 500 hPa geopotential height fields [Wallace and Gutzler (1981) define teleconnectivity as the absolute value of the strongest negative correlation between a grid point and any other grid point]. This figure shows that the terminus of the Asian/Pacific jet is approximately 150°W, and between 30 and 35°N. This coincides fairly well with the eastern boundary of the Hondo Region. Athanasiadis et al. (2010) continue to show that whereas the first mode of the Asian/Pacific jet EOF is closely correlated with the Pacific/North American (PNA) index, the second mode is not identified as tied to a specific index. However, it is the second mode which determines the location of the Asian/Pacific jet (and therefore the EAJS) and the storm track that encroaches on the west coast of North America. It is the second mode of the teleconnectivity patterns which best approximate the location of the Hondo Region.

Yang et al. (2002) identified the consistency at which the EAJS crosses the southern Japanese island of Kyūshū and enters the

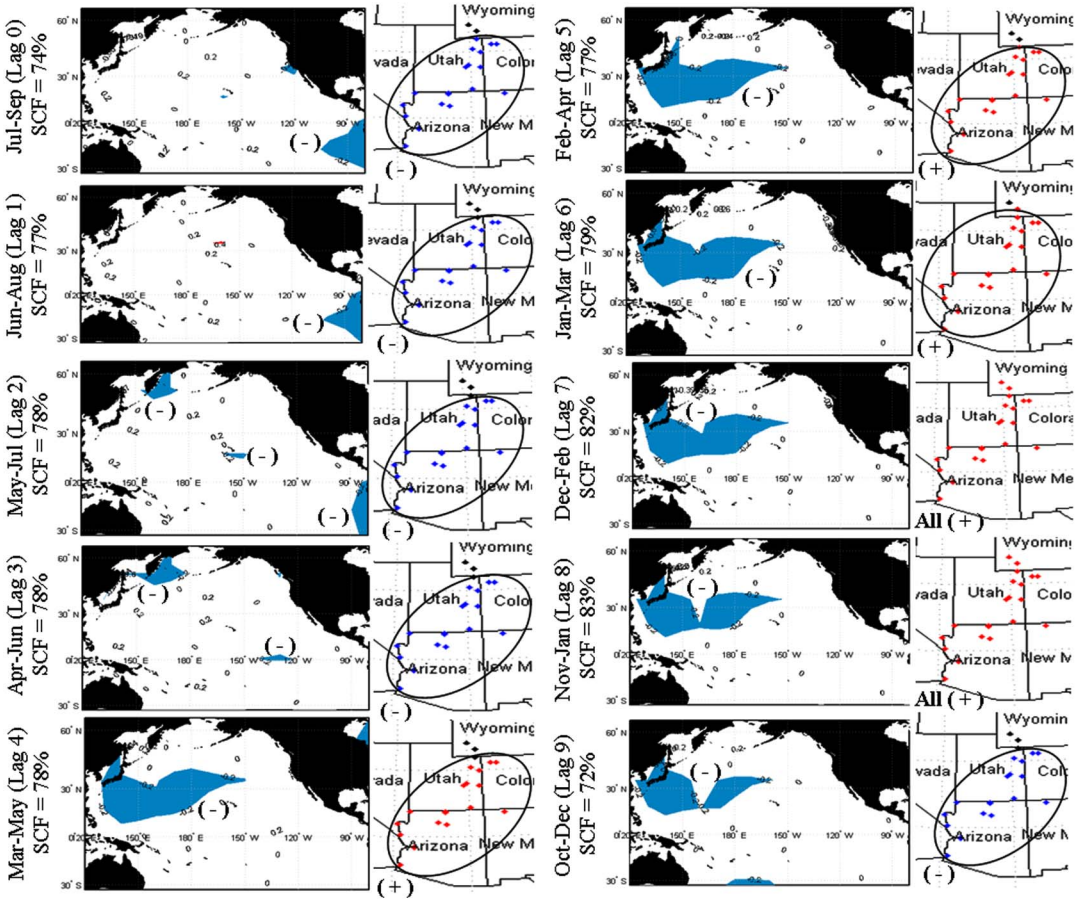

Fig. 6. Left and right heterogeneous correlation maps ($|r| \geq 0.27$) plotted from three-month average SST data, from lag 0 to lag 9; SCF values displayed for first mode

Hondo Region. Yang et al. (2002) further identified that as the position of the EAJS changes, corresponding changes in SSTs over the eastern Pacific Ocean and western North America occur. During the strong winter monsoon events, Zhang et al. (1997) showed that the surface wind patterns show that the EAJS "bifurcates" shortly after exiting the Hondo Region. A portion of the flow pattern turns southeast and enters the path of the tradewinds. The other portion is diverted north and joins the Aleutian low. Both air streams ultimately cross into the western United States thereafter.

This paper hypothesizes that the EAJS is the vehicle by which strong monsoonal moisture is carried from east Asia to western North America. Incidentally, Sun and Sun (as cited in Zhang et al. 1997) proposed that a strong winter monsoon usually precedes a drought summer season in east Asia. This study hypothesizes that operating in such conditions, moisture, which normally would remain in east Asia, is carried by the bifurcated wind flow toward North America. The northerly flow path carries some of the moisture to the western United States, where it is deposited as snowfall within the CRB (Azziz et al. 2010).

Forecast Skill Comparison Results

Once the SVD and rank sum testing has been completed, a Hondo index was defined as the LTES of the SVD analysis and then compared to the SOI and PDO index. Although the use of indexes (e.g., SOI, PDO) may result in skillful long-lead-time forecasts of streamflow, some streamflow regions may not be teleconnected to established climate signals and thus, would result in less than skillful forecasts when using such indexes. A distinct advantage of applying SVD to SST and streamflow regions is the possible identification of a new SST region that, when used in a forecast model, results in a skillful streamflow forecast.

The Forecast Skill Comparison method described earlier was employed by using the SOI, the PDO, and the SST LTES [taken from the January, February, and March (JFM) SVD analysis] as the third predictor. Fig. 7 displays the results from this analysis. At this lag time, the analysis shows that the SOI and the PDO result in negative forecast skill values (i.e., their forecasts are less accurate than the observed mean). On the other hand, the SST LTES provided an improvement to the forecast skill over the observed

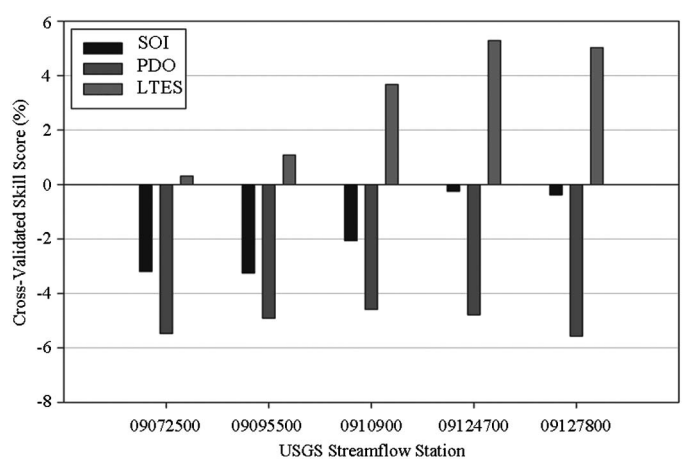

Fig. 7. Plot showing contribution of index to forecast skill. Value of '0' is a forecast as good as the mean. All stations are from the upper CRB

mean demonstrating the potential of the new region of the ocean to be used as a forecasting tool.

Conclusion

The goal of this analysis was to find which climate data or combination of climate data of the lead water year could be used to increase the skill of forecasting streamflow on the first day of the following water year. The rank sum testing of this research identified at least two different climate inputs for every river gauge in the CRB. For each of the 29 river gauges, there was a significant correlation with an established climate index. However, when comparing the number of successful tests from the different climate inputs, the SST data were more consistent in providing significant results.

Further testing by using SVD showed that a long-term correlation exists among the SST anomalies and streamflow anomalies for the Colorado River Basin. SVD testing of three-month averaged SST data showed that a region bounded by 150°E to 160°W and 24 to 34°N, named here as the Hondo Region, demonstrates a significant amount of covariance with the streamflow measured over the following water year. The significance of this region has been demonstrated in other research reports. The success of the Hondo Region lies in the strength and consistency of the Asian winter monsoon and the effectiveness of the EAJS to transport the moisture across the Pacific Ocean.

The results obtained in this research demonstrate the potential to generate long-lead forecasts of CRB streamflow from SST values observed over the Hondo Region in the Pacific Ocean. Therefore, the next steps include identifying a forecast approach which utilizes the teleconnection among the Hondo SST and CRB streamflow observations. The new work should include a variety of input alternatives to determine the simplest and most effective forecasting model for the entire water year or simply the seasonal runoff (April–July).

Acknowledgments

This research is supported through project NOAA NA070AR4310228.

References

Alder, H. L., and Roessler, E. B. (1977). *Introduction to probability and statistics*, 6th Ed., W.H. Freeman & Company Limited, San Francisco.

Athanasiadis, P. J., Wallace, J. M., and Wettstein, J. J. (2010). "Patterns of wintertime jet stream variability and their relation to the storm tracks." *J. Atmos. Sci.*, 67(5), 1361–1381.

Aziz, O. A., Tootle, G. A., Gray, S. T., and Piechota, T. C. (2010). "Identification of Pacific Ocean sea surface temperature influences of upper Colorado River Basin snowpack." *Water Resour. Res.* 46, W07536.

Barlow, L. K., White, J. W. C., Barry, R. G., Rogers, J. C., and Grootes, P. M. (1993). "The North Atlantic Oscillation signature in deuterium and deuterium excess signals in the Greenland Ice Sheet Project 2 Ice Core, 1840-1970." *Geophys. Res. Lett.*, 20(24), 2901–2904.

Bretherton, C. S., Smith, C., and Wallace, J. M. (1992). "An intercomparison of methods for finding coupled patterns in climate data." *J. Clim.*, 5(6), 541–560.

Clark, M. P., Serreze, M. C., and McCabe, G. J. (2001). "Historical effects of El Nino and La Nina events on the seasonal evolution of the Montane snowpack in the Columbia and Colorado River Basins." *Water Resour. Res.*, 37(3), 741–757.

Grantz, K., Rajagopalan, B., Clark, M., and Zagona, E. (2005). "A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts." *Water Resour. Res.*, 41(10), W10410..

Hidalgo, H. G., and Dracup, J. A. (2003). "ENSO & PDO effects on hydroclimatic variations of the upper Colorado River Basin." *J. Hydrometeorol.*, 4(1), 5–23.

Kim, T., Valdes, J. B., Nijssen, B., and Roncayolo, D. (2006). "Quantification of linkages between large-scale climatic patterns and precipitation in the Colorado River Basin." *J. Hydrol. (Amsterdam)*, 321(1-4), 173–186.

Kim, T., Yoo, C., and Ahn, J. (2007). "Influence of climate variation on seasonal precipitation in the Colorado River Basin." *Stoch. Environ. Res. Risk Assess.*, 22(3), 411–420.

Latif, M., and Barnett, T. P. (1994). "Causes of decadal climate variability over the north Pacific and North America." *Science*, 266(5185), 634–637.

Mantua, N. J., and Hare, S. R. (2002). "The Pacific decadal oscillation." *J. Oceanog.*, 58(1), 35–44.

McCabe, G. J., and Dettinger, M. D. (1999). "Decadal variations in the strength of Enso teleconnections with precipitation in the western United States." *Intl. J. Climatology*, 19(13), 1399–1410.

McCabe, G. J., and Dettinger, M. D. (2002). "Primary modes and predictability of year-to-year snowpack variation in the western United States from teleconnections with Pacific Ocean climate." *J. Hydrometeorology*, 3(1), 13–25.

McCabe, Gregory J., Betancourt, Julio L., and Hidalgo, Hugo G. (2007). "Associations of decadal to multidecadal sea-surface temperature variability with upper Colorado River flow." *J. of the Amer. Water Res. Assoc.*, 43(1), 183–192.

Milly, P. C. D., et al. (2008). "Stationarity is dead: Whither water management?" *Science*, 319(5863), 573–574.

Piechota, T. C., Chiew, F. H. S., Dracup, J. A., and McMahon, T. A. (1998). "Seasonal streamflow forecasting in eastern Australia and the El Niño-southern oscillation." *Water Resourc. Res.*, 34(11), 3035–3044.

Potts, J. M., Folland, C. K., Jolliffe, I. T., and Sexton, D. (1996). "Revised "LEPS" scores for assessing climate model simulations and long-range forecasts." *J. Clim.*, 9(1), 34–53.

Rajagopalan, B., Cook, E., Lall, U., and Ray, B. K. (2000). "Spatiotemporal variability of ENSO and SST teleconnections to summer drought over the United States during the twentieth century." *J. Clim.*, 13(24), 4244–4255.

Rasmussen, E. M., and Wallace, J. M. (1983). "Meteorological aspects of the El Nīno/southern oscillation." *Science*, 222(4629), 1195–1202.

Shabbar, A., Bonsal, B., and Khandekar, M. (1997). "Canadian precipitation patterns associated with the southern oscillation." *J. Clim.*, 10(12), 3016–3027.

Soukup, T., Aziz, O., Tootle, G., Wulff, S., and Piechota, T. (2009). "Incorporating climate into a long lead-time non-parametric streamflow forecast." *J. Hydrol.*, 368(1–4), 131–142.

Smith, T. M., and Reynolds, R. W. (2003). "Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997)." *J. Clim.*, 16(10), 1495–1510.

Smith, T. M., Reynolds, R. W., Peterson, Thomas C., and Lawrimore, Jay (2008). "Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006)." *J. Clim.*, 21(10), 2283–2296.

Strang, G. (1998). *Introduction to linear algebra*, 2nd Ed., Addison-Wesley, Reading, Mass.

Timilsena, J. A., Piechota, T. C., Hidalgo, H., and Tootle, G. A. (2007). "Five hundred years of hydrological drought in the upper Colorado River Basin." *J. Amer. Water Res. Assoc.*, 43(3), 798–812.

Tootle, G. A., and Piechota, T. C. (2004). "Suwannee River long-range streamflow forecasts based on seasonal climate predictors." *J. Amer. Water Res. Assoc.*, 40(2), 523–532.

Tootle, G. A., Piechota, T. C., and Singh, A. K. (2005). "Coupled oceanic-atmospheric variability and U.S. streamflow." *Water Resour. Res.*, 41(12), W12408.

Tootle, G. A., and Piechota, T. C. (2006). "Relationships between Pacific and Atlantic ocean sea surface temperatures and U.S. streamflow variability." *Water Resour. Res.*, 42(7), W07411.

Tootle, G. T., et al. (2009). "2009-2010 El Niño: Hydrologic relief for parts of the U.S.?" *EOS Trans. AGU*, 90(50), 481–482.

Wallace, J. M., and Gutzler, D. S. (1981). "Teleconnections in the geopotential height field during the Northern Hemisphere winter." *Mon. Weather Rev.*, 109(4), 784–812.

Wallace, J. M., Gutzler, D. S., and Bretheron, C. S. (1992). "Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies." *J. Clim.*, 5(6), 561–576.

Wang, H., and Ting, M. (2000). "Covariabilities of winter U. S. precipitation and Pacific sea surface temperatures." *J. Clim.*, 3711–3719.

Ward, M. N., and Folland, C. K. (1991). "Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea-surface temperature." *Intl. J. of Climatology*, 11(7), 711–743.

Woodhouse, C. A. (2001). "A tree-ring reconstruction of streamflow for the Colorado Front Range." *J. of Amer. Water Res. Assoc.*, 37(3), 561–569.

Yang, S., Lau, K. M., and Kim, K. M. (2002). "Variations of the east Asian jet stream and Asian-Pacific-American winter climate anomalies." *J. Clim.*, 15(3), 306–325.

Yarnal, B., and Diaz, H. F. (1986). "Relationships between extremes of the Southern Oscillation and the winter climate of the Anglo-American Pacific coast." *J. Climatology : A Journal of the Royal Meteorological Society*, 6(2), 197–219.

Zhang, Y., Sperber, K. R., and Boyle, J. S. (1997). "Climatology and interannual variation of the east Asian winter monsoon: Results from the 1979-95 NCEP/NCAR reannalysis." *Mon. Weather Rev.*, 125(10), 2605–2619.