A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts

Katrina Grantz[1] and Balaji Rajagopalan[2]

Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, Colorado, USA

Martyn Clark

Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA

Edith Zagona

Center for Advanced Decision Support for Water and Environmental Systems, University of Colorado, Boulder, Colorado, USA

Received 5 July 2004; revised 25 February 2005; accepted 2 June 2005; published 18 October 2005.

[1] Water managers throughout the western United States depend on seasonal forecasts to assist with operations and planning. In this study, we develop a seasonal forecasting model to aid water resources decision making in the Truckee-Carson River System. We analyze large-scale climate information that has a direct impact on our basin of interest to develop predictors to spring runoff. The predictors are snow water equivalent (SWE) and 500 mbar geopotential height and sea surface temperature (SST) ''indices'' developed in this study. We use local regression methods to provide ensemble (probabilistic) forecasts. Results show that the incorporation of climate information, particularly the 500 mbar geopotential height index, improves the skills of forecasts at longer lead times when compared with forecasts based on snowpack information alone. The technique is general and could be used to incorporate large-scale climate information into ensemble streamflow forecasts for other river basins.

Citation: Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona (2005), A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts, *Water Resour. Res.*, *41*, W10410, doi:10.1029/2004WR003467.

1. Introduction

[2] Water resource managers in the western United States are facing the growing challenge of meeting water demands for a wide variety of purposes under the stress of increased climate variability [e.g., *Hamlet et al.*, 2002]. Careful planning is necessary to meet demands on water quality, volume, timing and flowrates. This is particularly true in the western United States, where it is estimated that 44% of renewable water supplies are consumed annually, as compared with 4% in the rest of the country [*el-Ashry and Gibbons*, 1988]. The forecast for the upcoming water year is instrumental to the water management planning process. In the managed river systems of the West the skill of a streamflow forecast dramatically affects management efficiency and thus system outputs, such as crop production and the monetary value of hydropower production [e.g., *Hamlet et al.*, 2002], as well as the sustainment of aquatic species.

[3] Forecasting techniques for the western United States have long used winter snowpack as a predictor of spring runoff. Because the majority of river basins in the West are snowmelt dominated [*Serreze et al.*, 1999], winter snowpack measurements provide useful information, up to several months in advance, about the ensuing spring streamflow. More recently, information about large-scale climate phenomena such as El Niño–Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) pattern has been added to the forecaster's toolbox. The link between these large-scale phenomena and the hydroclimatology of the western United States has been well documented in the literature [*Ropelweski and Halpert*, 1986; *Cayan and Webb*, 1992; *Redmond and Koch*, 1991; *Gershunov*, 1998; *Dettinger et al.*, 1998]. *Clark et al.* [2001] showed that including large-scale climate information together with SWE improves the overall skill of the streamflow predictions in the western United States. *Souza Filho and Lall* [2003] show significant skills at long lead times in forecasting streamflows in Cearra, Brazil using climate information from the Atlantic and Pacific oceans.

[4] These teleconnection patterns, though dominant on a large scale, often fail to provide forecast skill in individual basins. For example, while streamflow in El Niño events is generally below average in the Pacific Northwest and above average in the desert southwest, ENSO information offers limited forecast skill in basins outside these core regions

[1] Also at Center for Advanced Decision Support for Water and Environmental Systems, University of Colorado, Boulder, Colorado, USA.

[2] Also at Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.

0043-1397/05/2004WR003467$09.00

[*McCabe and Dettinger*, 2002]. Moreover, relatively minor shifts in large-scale atmospheric patterns can result in large differences in surface climate [e.g., *Yarnal and Diaz*, 1986], suggesting that predictive indices may need to be basin-specific.

[5] In this paper we present a generalized framework for utilizing large-scale climate information to forecast streamflows at the basin scale. The framework first identifies the large-scale climate patterns and predictors that modulate seasonal streamflows in the given basin. It next uses the predictors to develop a forecast model of the seasonal flows and subsequently tests and validates the model. This framework is applied to forecasting spring streamflows in the Truckee and Carson river basins located in the Sierra Nevada Mountains.

[6] The paper is organized as follows. Section two presents a background on large-scale climate and its impacts on western U. S. hydroclimatology. The study region and data used are described in sections three and four, respectively. This is followed by the proposed method of climate diagnostics and identification of predictors for forecasting spring streamflows in section five. Section six presents the development of the statistical ensemble forecating model using the identified predictors. This section also discusses model testing and verification. Section seven presents the results and section eight summarzies and concludes the paper.

2. Large-Scale Climate and Western U.S. Hydroclimatology

[7] The tropical ocean-atmospheric phenomenon in the Pacific identified as El Niño–Southern Oscillation (ENSO) [e.g., *Allan et al.*, 1996] is known to impact the climate all over the world and, in particular, the western United States [e.g., *Ropelweski and Halpert*, 1986]. The warmer sea surface temperatures and stronger convection in the tropical Pacific Ocean during El Niño events deepen the Aleutian Low in the North Pacific Ocean, amplify the northward branch of the tropospheric wave train over North America and strengthen the subtropical jet over the southwestern United States [*Bjerknes*, 1969; *Horel and Wallace*, 1981; *Rasmussen*, 1985]. These circulation changes are associated with below-normal precipitation in the Pacific Northwest and above-normal precipitation in the desert Southwest [e.g., *Redmond and Koch*, 1991; *Cayan and Webb*, 1992]. Generally opposing signals are evident in La Niña events, but some nonlinearities are present [*Hoerling et al.*, 1997; *Clark et al.*, 2001].

[8] Decadal-scale fluctuations in SSTs and sea levels in the northern Pacific Ocean as described by the PDO [*Mantua et al.*, 1997] provide a separate source of variability for the western U. S. hydroclimate. Independence of PDO from ENSO is still in debate [*Newman et al.*, 2003]. Regardless, the influence of PDO and ENSO on North American hydroclimate variability has been well documented [e.g., *Ropelweski and Halpert*, 1986; *Cayan and Webb*, 1992; *Kahya and Dracup*, 1993; *Dracup and Kahya*, 1994; *Redmond and Koch*, 1991; *Cayan*, 1996; *Gershunov*, 1998; *Kerr*, 1998; *Dettinger et al.*, 1998, 1999; *Cayan et al.*, 1999; *Hidalgo and Dracup*, 2003].

[9] Incorporation of this climate information has been shown to improve forecasts of winter snowpack [*McCabe and Dettinger*, 2002] and streamflows in the western United States [*Clark et al.*, 2001; *Hamlet et al.*, 2002] while increasing the lead time of the forecasts. Use of climate information enables efficient management of water resources and provides socioeconomic benefits [e.g., *Pulwarty and Melis*, 2001; *Hamlet et al.*, 2002].

[10] Often, however, the standard indices of these phenomena (e.g., NINO3, SOI, PDO index, etc.) are not good predictors of hydroclimate in every basin in the western United States, even though these phenomena do impact the western U. S. hydroclimate. Because the canonical patterns of these climate phenomena refer to specific regions in the ocean (e.g., 5°N–5°S and 90°W–150°E for the NINO3 index) slight shifts in the patterns can result in decreased correlation values between the indices and basin hydroclimate. Furthermore, certain regions in the western United States (e.g., basins in between the Pacific Northwest and the desert Southwest) can be impacted by both the northern and southern branches of the subtropical jet, potentially diminishing apparent connections to ENSO and PDO. The Truckee and Carson basins are two such examples and previous studies have shown that the Truckee River is not significantly (or consistently) influenced by ENSO, PDO or PNA [e.g., *Tootle and Piechota*, 2004]. Hence predictors other than the standard indices have to be developed for each basin.

3. Study Region: Truckee and Carson Basins

[11] The study region of the Truckee and Carson River basins in the Sierra Nevada Mountains is shown in Figure 1. The Truckee and Carson rivers originate high in the California Sierra Nevada Mountains and flow northeastward down through the semiarid desert of western Nevada. The Truckee River originates as outflow from Lake Tahoe in California and terminates approximately 115 miles later in Pyramid Lake in Nevada. The Carson River has its headwaters approximately fifty miles south of Lake Tahoe, runs almost parallel to the length of the Truckee River and terminates in the Carson Sink area. The two rivers are connected approximately two thirds of the way down the system by the one-way Truckee Canal which brings water from the Truckee River to the Carson River. The basins' areas are comparable and are approximately 3000 square miles. The majority of the basins' precipitation falls as snow during the winter months (November–March) and the bulk of the annual streamflow arrives during spring (April–July) due to the melting of the snowpack. This is evident in the climatology of monthly precipitation and streamflows for the Truckee River (inset in Figure 1). The streamflows in Figure 1 are from the Farad gauging station on the Truckee River and the precipitation is from the national climatic data center climate division covering the headwater region of the basin (details on the data sets are provided in section 4). The Carson River exhibits similar climatology.

[12] The Truckee and Carson rivers are highly litigated and managed, and thus require strict planning for efficient management of the system. The Bureau of Reclamation (BOR) Lahontan Basin area office manages operations on the rivers and relies heavily on seasonal streamflow forecasts for planning and management. The BOR uses forecasts of the spring runoff (April to July volume) that are issued on the first of each month starting in January. Current

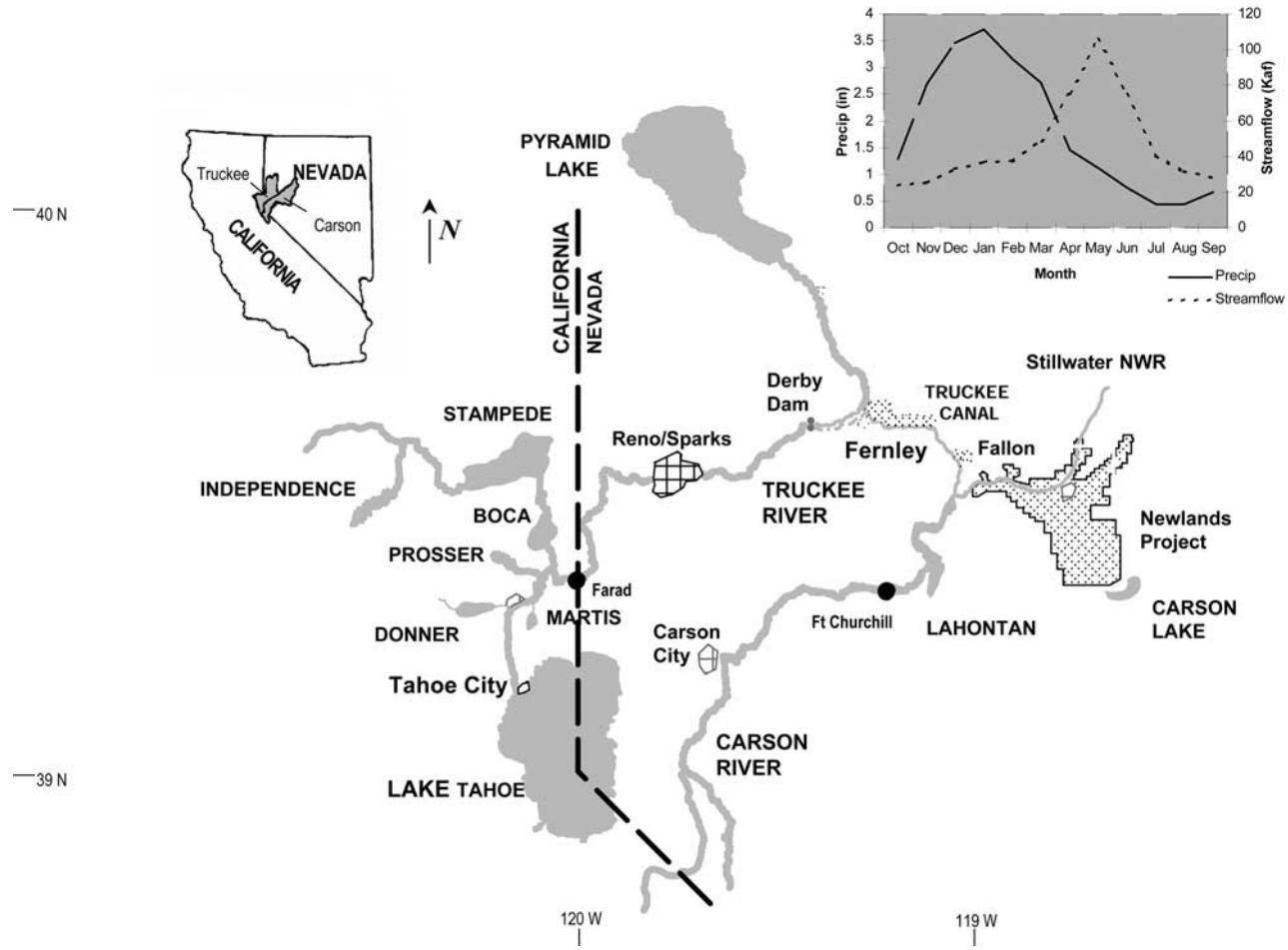

Figure 1. Truckee and Carson rivers and surrounding area. They gray shading represents rivers and lakes, and the hatched area represents irrigation areas. Annual hydrographs are shown in the inset; the dashed line represents the average annual precipitation, and the solid line represents the average annual streamflow.

forecasting techniques use multiple linear regression analysis based on factors related to the existing snowpack and, hence long-lead forecast skills are limited. Additionally, the current technique does not provide forecasts prior to January as the snowpack information is only partial. Thus improvements to the spring forecasts, both in skill and in lead time, are needed to strengthen planning and operations in the Truckee and Carson basins.

4. Data

[13] The following data sets for the period 1949–2003 were used in the analysis.

[14] 1. Monthly natural streamflow data for Farad and Ft. Churchill gauging stations on the Truckee and Carson rivers, respectively, were obtained from BOR. Natural streamflows are computed based on inflows to the seven major storage reservoirs near the top of the basin before any significant depletion have been made (J. Rieker, U. S. Bureau of Reclamation, personal communication, 2003). Spring seasonal (April–July) volume was computed for this study from the monthly streamflows.

[15] 2. Monthly SWE data were obtained from the NRCS National Water and Climate Center Web site (http://www.wcc.nrcs.usda.gov). The SWE data are gathered from snow course and snotel stations in the upper Truckee Basin (17 stations) and upper Carson Basin (7 stations). Basin averages of SWE were calculated for this study using the method employed by the NRCS for these basins: the SWE depth from every station in the basin is summed and then divided by the sum of the long-term averages for each of the stations (T. Pagano, Natural Resources Conservation Service, personal communication, 2003). All values for SWE are represented as the percent of normal.

[16] 3. Monthly winter precipitation data for the California Sierra Nevada Mountains region were obtained from the U. S. climate division data set from the NOAA-CIRES Climate Diagnostics Center (CDC) Web site (http://www.cdc.noaa.gov).

[17] 4. Monthly values of large-scale ocean atmospheric variables, SST, geopotential heights, sea level pressure (SLP), wind, etc., from NCEP/NCAR Re-analysis [*Kalnay et al.*, 1996] were obtained from the CDC Web site.

5. Climate Diagnostics

[18] The first step in the forecasting framework is to identify predictors of spring flows in the basin. To this

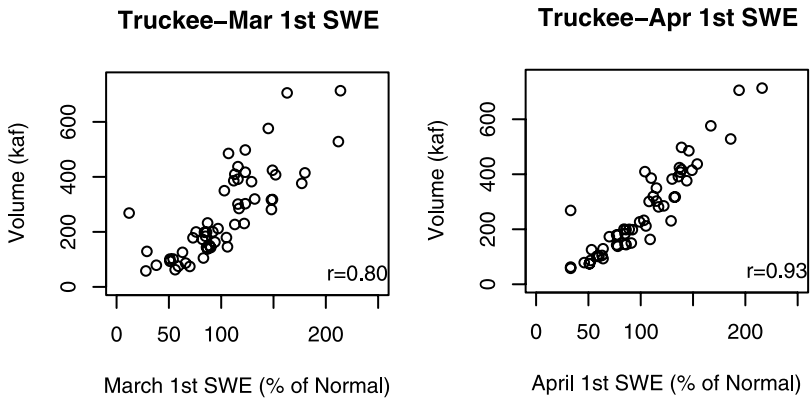

Figure 2. SWE, 1 March and 1 April, versus spring runoff volume in the Truckee River for the period 1949 to 2003. SWE is taken as a basin-wide average and represented as a percent of normal value.

end, we first examined the relationship between SWE and spring runoff in the basins. Next, we correlated spring streamflows with global climate variables from the preceeding fall and winter seasons. We chose to examine variables from fall and winter because the state of the atmosphere during this time affects the position of the jet stream, and consequently, snow deposition and the resulting spring runoff. Also, predictors from fall and winter allow for potential long-lead forecasts.

[19] Scatterplots of the end of winter SWE and spring runoff in the Truckee River are shown in Figure 2. Results are similar for the Carson River. As expected, there is a high degree of correlation between winter SWE and spring runoff, particularly with 1 April SWE as it provides a more complete representation of the end of winter snowpack in the basins. Correlation values for Truckee streamflows are 0.80 and 0.93 with 1 March SWE and 1 April SWE, respectively, and 0.81 and 0.90, respectively, with the Carson flows. High correlations of streamflows with 1 March SWE offers the opportunity for at least a 1 month lead forecast. The 1 January SWE, however, does not correlate as well with spring streamflows (0.53 for the Truckee and 0.49 for the Carson) and hence provides poorer skill as a predictor to spring runoff.

[20] Spring streamflows in the Truckee and Carson basins are likely modulated by ENSO and PDO, but the standard indices of these phenomena did not show significant correlations with spring streamflows (0.22 for the December–

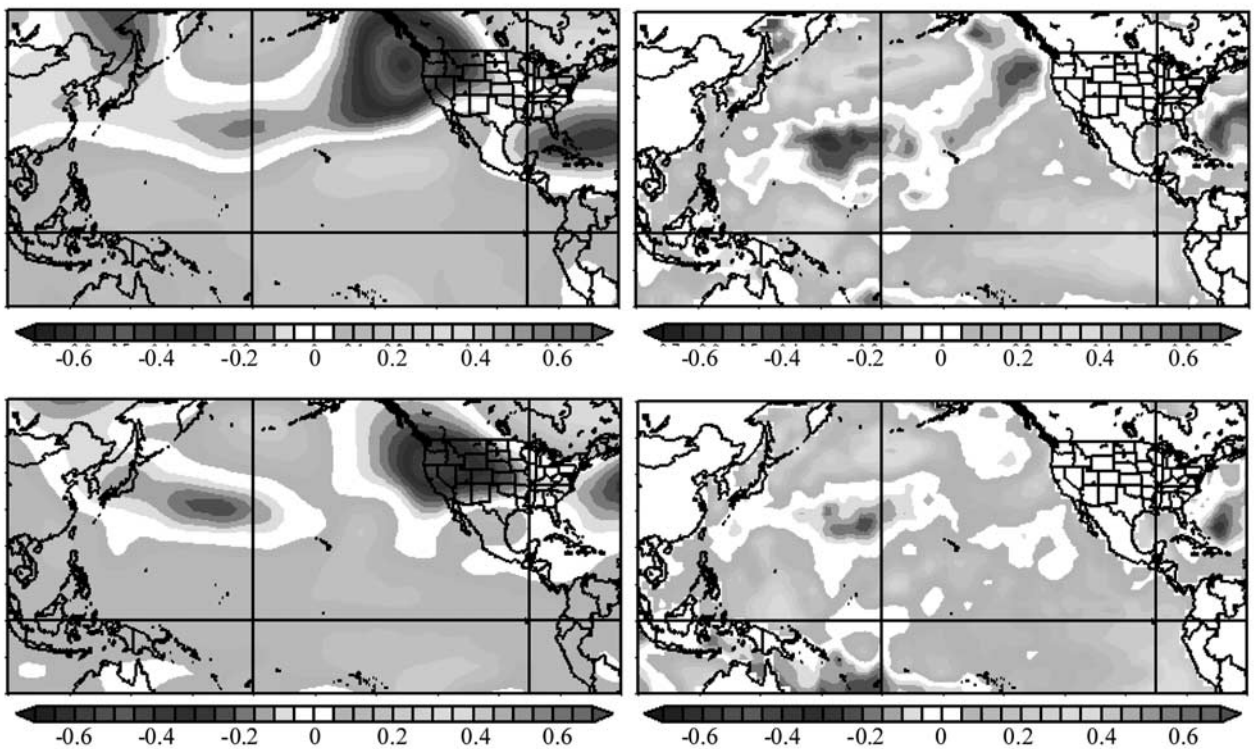

Figure 3. Correlation of Carson River spring streamflows with (top) December–February and (bottom) September–November climate variables: (left) 500 mbar geopotential height (Z500) and (right) SST. See color version of this figure at back of this issue.

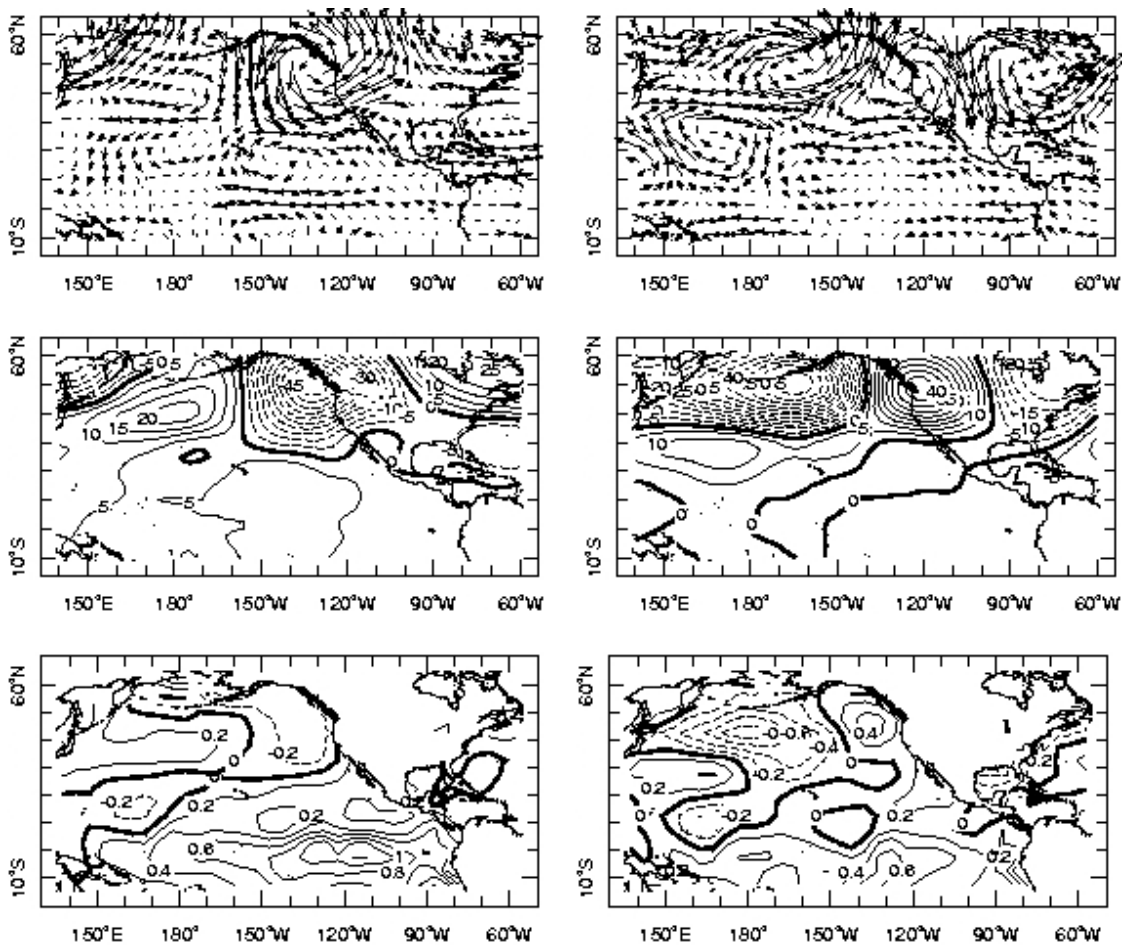

Figure 4. Composites of vector winds, SST, and Z500 during the winter of (left) high and (right) low streamflow years.

February NINO3, −0.13 for the December–February PDO, and −0.21 for the December–February SOI, for the Truckee; results are similar for the Carson). Thus we correlated the spring streamflows with the standard ocean-atmospheric circulation fields (e.g., 500 mbar geopotential height fields, SSTs, SLPs, etc.) to investigate the large-scale climate link and potential predictors.

[21] Figure 3 (top) presents the correlations between spring streamflows in the Carson River and the winter SSTs and 500mbargeopotential heights, henceforth, referred to as Z500, in the Pacific Ocean. Strong negative correlations (approximately −0.7) with Z500 in the region off the coast of Washington can be seen. The SSTs in the northern mid-Pacific Ocean exhibit a strong positive (about 0.5) correlation and to the east of this they exhibit a negative correlation. Similar, but slightly weaker correlation patterns can be seen for the preceding fall (September–November) Z500 and SSTs (Figure 3, bottom). This suggests that the physical mechanisms responsible for the correlations are persistent from fall through winter. These correlations offer hopes for a long-lead forecast of spring streamflows; at the least, they can provide significant information about the upcoming spring streamflows before SWE data is available.

[22] To understand the physical mechanisms driving the correlation patterns seen above, a composite analysis was perfomed. In this, average SST, wind and Z500 patterns for high and low streamflow years were obtained to identify coherent regions with strong magnitudes of the variables. We chose years with streamflows exceeding the 90th percentile as high years and those below the 10th percentile as low years. Figure 4 shows the composites of vector wind, Z500, and SST anomalies during the winter season preceding the high and low streamflow years. The winds in high streamflow years show a counterclockwise rotation around the low-pressure region off the coast of Washington state, the region of highest correlation seen in Figure 3. This counterclockwise rotation brings southerly winds over the Trukee and Carson basins. Southerly winds tend to be warm and moist, thus increasing the chances of enhanced winter snow and, consequently, higher streamflows the following spring. The opposite pattern is seen during low streamflow years when anomalous northeasterlies tend to bring cold, dry air and, consequently, less snow and decreased streamflows. The Z500 patterns and the vector wind anomalies in high and low streamflow years are consistent with each other. The SST patterns in high and low streamflow years (Figure 4) are a direct response to the pressure and winds. The winds are generally stronger to the east of a low-pressure region; this increases the evaporative cooling and also increases upwelling of deep cold water to the surface. Together, they result in cooler than normal SSTs to the east of the low-pressure region. The opposite is true on the west

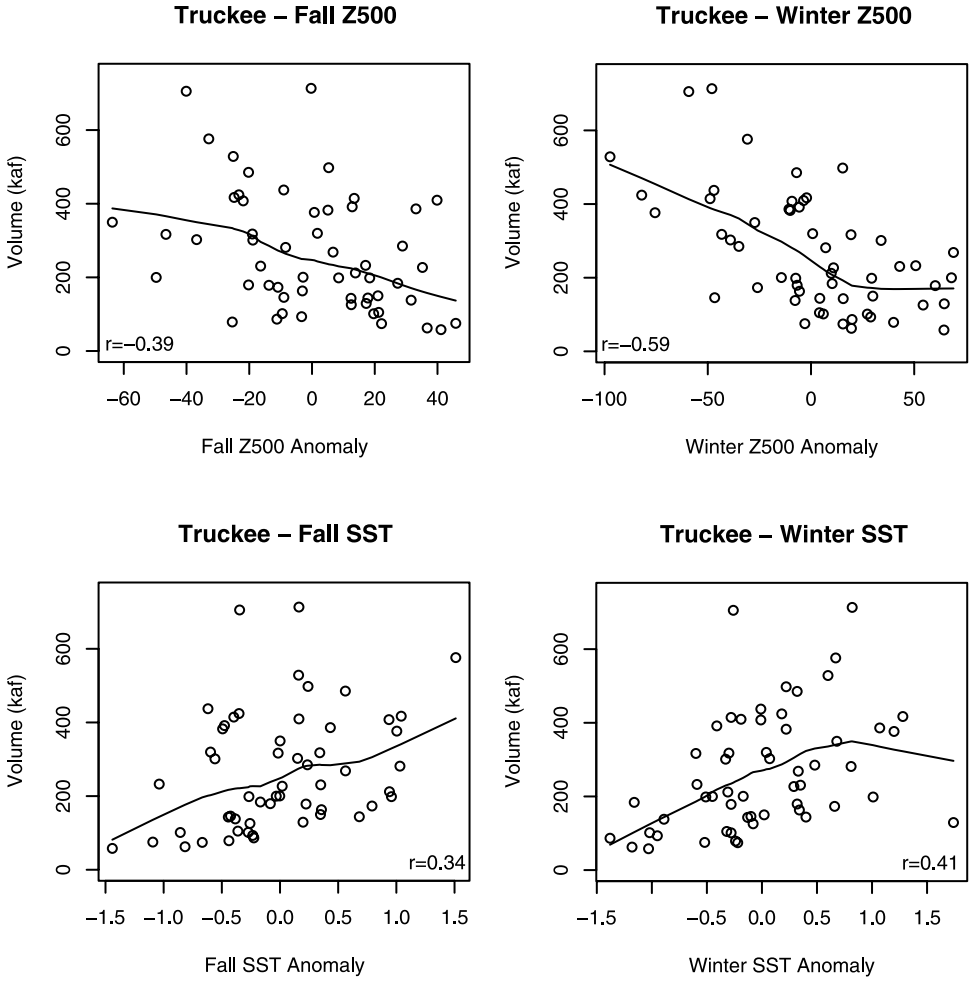

Figure 5. Scatterplots of (left) fall and (right) winter (top) geopotential height index and (bottom) SST index with the spring runoff in the Truckee River.

side of the low-pressure region. Composite maps for the fall season show similar patterns, indicating that the physical mechanisms are persistent. Results for the Truckee River streamflows are very similar [*Grantz*, 2003].

[23] It is recognized that atmospheric circulation is a response to some known (or unknown) boundary forcing (e.g., tropical sea surface temperatures, land surface processes, and other boundary forcings that have yet to be discovered). The known forcings, such as ENSO and PDO, however, do not have strong signals in all river basins. The research philosophy taken here is to use low-frequency variability in atmospheric circulation, i.e., the Z500 pattern, as a proxy for both known and unknown boundary forcings.

6. Forecast Model Development

[24] The correlation and composite analyses indicate a potential for long-lead (1 $\sim$ 2 seasons) forecasts of spring streamflows. To realize this potential, we first developed predictors to be used in the ensemble forecast model. This is the second step in the forecasting framwork and is described below.

6.1. Selection of Predictors

[25] On the basis of the correlation and composite analyses (Figures 3 and 4) we developed indices specific to the Truckee and Carson basins by averaging the ocean-atmospheric variables over the areas of highest correlation. These areas were determined by visual inspection of the correlation maps. Correlations in the selected areas had to be statistically significant and relatively high compared to other regions. Specifically, the Z500 was averaged over the region 225°–235°E and 42°–46°N and the SSTs over the region 175°–185° E and 42°–47°N. Time series of the indices were obtained to be used as predictors in the forecast model.

[26] Figure 5 shows the scatterplot of Z500 and SST indices from the preceding fall and winter seasons with the spring streamflows in the Truckee River. We used a local polynmial technique [*Loader*, 1999] to fit a smooth curve to the scatterplot. As expected, a negative relationship exists between the streamflows and Z500 index and a positive relationship with the SST index. Slight nonlinearities can also be seen from the scatterplots. Correlations between spring streamflows in the Truckee River and the two indices from preceding seasons show persistence from late summer and are statistically significant back to the August–October season [*Grantz*, 2003]. This supports the potential for longer–lead time forecasts of spring streamflows. Thus there are thee potential predictors (Z500 index, SST index, and SWE) that can be used for streamflow forecast.

[27] The selection of the best subset of predictors to be included in the model can be determined using several different approaches. Some of the standard objective methods for predictor selection are stepwise regression using the F test, Akaike Information Criteria (AIC), Mallow's Cp statistic, and Generalized Cross Validation (GCV) (discussed in section 6.2). In this study we used the GCV criteria to evaluate all possible combinations of predictors and select the best subset. The combination of predictors that produced the best GCV value was the Z500 index and SWE. This subset of predictors was used in the ensemble forecasts.

6.2. Ensemble Forecast Model

[28] Statistical forecast models can be represented as

$$Y = f(x_1, x_2, x_3, \ldots x_p) + e \quad (1)$$

Where f is a function fitted to the predictor variables (x_1,x_2,. . .,x_p), Y is the dependent variable (in this case the spring streamflows) and e is the error assumed to be normally (or Gaussian) distributed with a mean of 0 and variance σ. Traditional regression methods involve fitting a function, often linear, to the entire data set. The theory behind these methods and the procedures for parameter estimation and hypothesis testing are well developed [*Helsel and Hirsch*, 1995]. The main drawbacks, however, are the assumption of a Gaussian distribution of the data and the fitting a global relationship (e.g., a linear equation in the case of linear regression) between the variables. If the linear model is found inadequate, higher-order models (quadratic, cubic, etc.) have to be considered, which can be difficult to fit in the case of short data sets and, consequently, local nonlinear features cannot be well captured.

[29] Local functional estimation methods, in contrast, estimate the function f ''locally'' and this provides the capability to capture any arbitrary feature (linear or nonlinear) present in the data. There are several local estimation methods, such as kernel-based [*Bowman and Azzalini*, 1997], splines, K nearest neighbor (KNN) local polynomials [*Rajagopalan and Lall*, 1999; *Owosina*, 1992], locally weighted polynomials (LWP) [*Loader*, 1999], etc. *Owosina* [1992] performed an extensive comparison of a number of regression methods, both local and global, on a variety of synthetic data sets and found the local regression methods to generally perform better in capturing the underlying functions of the data. For an overview on local functional estimation methods and hydrologic applications, see *Lall* [1995].

[30] KNN and LWP methods obtain the value of the function f at any point 'x*' by fitting a polynomial to a small set of neighbors to ''x*.'' Once the neighbors are identified, there are two main options: (1) The neighbors can be resampled with a weight function that gives more weight to the nearest neighbors and less to the farthest, thus generating an ensemble [*Lall and Sharma*, 1996; *Rajagopalan and Lall*, 1999; *Yates et al.*, 2003; *Souza Filho and Lall*, 2003]. (2) A polynomial can be fit to the neighbors that can be used to estimate the mean of the dependent variable [*Rajagopalan and Lall*, 1998] and the variance of the errors. The estimate of the error variance can be used to generate random normal deviates which, when added to the mean estimate, yield ensembles.

[31] Thus the parameters to be estimated are the size of the neighborhood (K) and the order of the polynomial (p). Note that unlike the global regression alternatives, no prior assumption is made regarding the functional form of the relationship (e.g., a linear relationship).

[32] In this research, we utilize a modified version of LWP adapted by *Prairie* [2002] and *Prairie et al.* [2005] and applied to streamflow and salinity modeling. This method was later implemented by *Singhrattna et al.* [2005] for forecasting the Thailand monsoon. This technique uses the LWP to estimate the mean (expected) value and the residuals of the fit are bootstrapped (or resampled) to obtain ensembles. For details on the methodology, see *Prairie* [2002] and *Prairie et al.* [2005]. A brief description of the methodology is as follows.

[33] 1. For a given data set, the best choice of neighborhood size (K) and the order of polynomial (p) are obtained using objective criteria such as generalized cross validation (GCV) or likelihood.

[34] 2. At each observed data point, x_j, K nearest neighbors are identified and a local polynomial of order p is fitted. This fit is then used to estimate the value of the dependent variable (the conditional mean) at an observed point. (In this research, the package LOCFIT [*Loader*, 1999] was used to determine the LWP.) The residual, e_j, is then computed. This is repeated at each data point, thus obtaining the residual for all data points. This can be described as the ''fitting'' process.

[35] 3. For a new data point, x_{new}, at which a forecast is required, the conditional mean value, Y_{new}, is obtained using step 2.

[36] 4. Next, one of the neighbors of $x_{new,}$ say x_i, and the corresponding residual, e_i are selected. The residual is then added to the mean forecast ($Y_{new} + e_i$) thus obtaining one of the ensemble members. The selection of one of the neighbors is done using a weight function of the form:

$$W(j) = \frac{1}{j\sum_{i=1}^{k}\frac{1}{i}} \quad (2)$$

This weight function gives more weight to the nearest neighbors and less to the farthest neighbors. The number of neighbors to be used to resample the residuals need not be same as the number of neighbors used to perform the local polynomial in step 1. In practice, $\sqrt{N-1}$ is used to resample the residuals. Note that for small sample sizes it is not advisable to resample the residuals, but rather to use the standard error of the distribution. This alternative technique is discussed later in this section.

[37] 5. Repeat step 4 as many times as required (100 in this study) to obtain a probability density function (PDF) that does not change with more sampling. This will result in the ensemble forecast.

[38] 6. Repeat steps 3–5 for each forecast point.

[39] Figure 6 can be utilized to better visualize these steps. Figure 6 shows the scatterplot of the historical area-averaged winter Z500 index and spring runoff for the Carson River at Ft. Churchill. The solid line is the locally weighted fit through the scatter. The bootstrapping of the residuals for the ensemble forecast is depicted in the dashed box. The main advantage of this modified KNN approach is

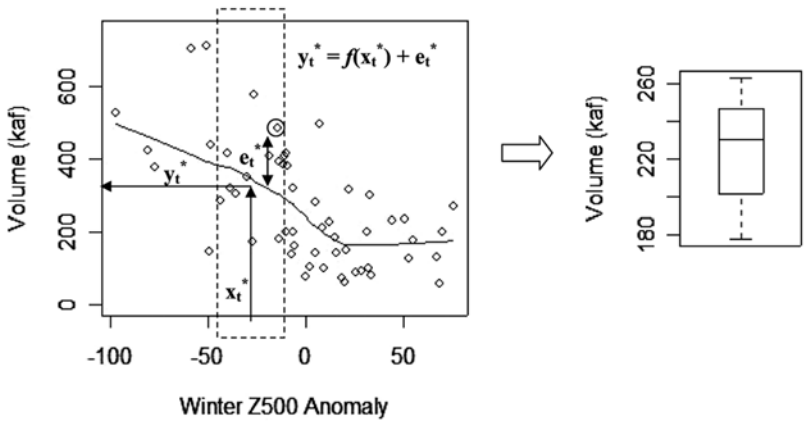

Figure 6. Residual resampling to obtain an ensemble forecast.

the ability to capture the local error structure and also to generate values not seen in the historical record unlike the straight bootstrap techniques [*Lall and Sharma*, 1996; *Souza Filho and Lall*, 2003]. However, with short data sets the residual resampling (step 4) provides limited variety in the ensembles and consequently, might not model the local error structure well. An alternative approach in such cases is to (1) assume a ''locally normal'' distribution of the errors, (2) use the standard errors from the LWP [*Loader*, 1999] to generate random normal deviates, and then (3) add these to the mean estimate from the LWP to generate the ensembles. In an application of forecasting Thailand summer rainfall [*Singhrattna et al.*, 2005] with small sample size, this approach was implemented with good results.

[40] Though the relationships in the Truckee and Carson basins are not highly nonlinear, we chose the LWP forecasting approach because it can be used for any arbitrary (linear or nonlinear) feature that might be present in the data, thus making the framework general and readily applicable to any basin.

[41] In this study, the neighborhood size and the polynomial order for the forecasting model are obtained using the objective criteria, generalized cross validation (GCV) [*Loader*, 1999]. The GCV function is a good surrogate of predictive error [*Craven and Whaba*, 1979] of the model, unlike least squares which is a measure of goodness of fit and provides no information on the predictive capability. For a range of values of the neighborhood size (K) and polynomial order (p) (usually 1 or 2), the combination that provides the minimum GCV value is selected. This is done separately for each forecasting lead time. The GCV (K, p) score function is defined as:

$$GCV(K,p) = \frac{\sum_{i=1}^{n} \frac{e_i^2}{N}}{\left(1 - \frac{m}{N}\right)^2} \quad (3)$$

where e_i is the error, N is the number of data points, m is the number of parameters.

[42] The neighborhood size for the residual resampling can be the same as obtained from the GCV criteria for fitting the local polynomial, or it can be different. In this study, the heuristic approach that calls for the neighbors for the residual resampling to be $\sqrt{N-1}$ was used [*Rajagopalan and Lall*, 1999; *Yates et al.*, 2003].

[43] The GCV function described above can also be used for selecting the best subset of predictors from all predictors. In this, one fits local polynomial for different predictor combinations along with the polynomial order and the neighborhood size and calculates the GCV value for each case. For this, equation 6.3 entails a third parameter, q, for the combination of predictors: GCV (K, p, q). The combination that produces the least GCV value is chosen as the best subset. In this study we used the GCV criteria to select the Z500 index and SWE as predictors. It is possible to have GCV models with similar GCV values close to the minimum value. In this situation, one can take all the models with the similar low GCV values, weight them based on the GCV value, then resample the models and create a multimodel superensemble to capture the model uncertainty [*Regonda et al.*, 2005].

6.3. Ensemble Forecast

[44] Using the model described above, we forecast the April to July total runoff volume in the Truckee River at Farad and in the Carson River at Ft. Chruchill. This is the third step in the forecasting framework. We develop these ensemble forecasts for the beginning of each month from November to April. Predictors from the preceeding three months are used in the model. For example, the forecast issued on 1 March uses the time-averaged Z500 and SST indices from the December–February period and the 1 March SWE. Forecasts issued in November and December use only the Z500 and SST indices as the SWE is not yet available.

6.4. Forecast and Model Verification

[45] The fourth and final step in the forecasting framework is to verify the forecasting model in a cross-validated mode. In this, the streamflow value in a given year is dropped from the data set and an ensemble of predictions is generated from the model based on the rest of the data. This is repeated for each year, producing a cross-validated ensemble forecast for each year for the 1949–2003 period. Given the relatively small sample size, we performed this leave-one-out cross validation. A more rigorous cross validation could be performed by dropping several data points from the regression, fitting the model and then forecasting the dropped points.

[46] Apart from visual inspection, the ensembles are evaluated on a suite of three performance criteria:

(1) correlation coefficient of the mean of the ensemble forecast and the observed value, which measures the skill in the mean forecast, (2) ranked probability skill score (RPSS) [*Wilks*, 1995], and (3) likelihood function skill score (LLH) [*Rajagopalan et al.*, 2002]. RPSS and LLH measure the forecast's ability to capture the PDF. The RPSS is typically used by climatologists and meteorologists to evaluate a model's skill in capturing categorical probabilities relative to climatology.

[47] For the RPSS and LLH the streamflows are divided into three categories, at the tercile boundaries, i.e., 33rd percentile and 66th percentile. Values above the 66th percentile are in the above normal category, below the 33rd percentile are in the below normal category, and the remainder fall in the normal category. The categorical probability forecast is obtained as the proportion of ensemble members falling in each category. The climatology forecast is the proportion of historical observations in each category. For the tercile categories presented here the climatological probability of each category is one third.

[48] For a categorical probabilistic forecast in a given year, $P = (P_1, P_2, \ldots P_k)$ (where k is the number of mutually exclusive and collectively exhaustive categories; here it is 3) the rank probability score (RPS) is defined as

$$\mathrm{RPS(forecast)} = \sum_{m=1}^{k} \left[\left(\sum_{j=1}^{m} P_j - \sum_{j=1}^{m} d_j \right)^2 \right] \tag{4}$$

The vector d (d_1, d_2, $\ldots$ d_k) represents the observations, such that d_k equals one if the observaion falls in the kth category and zero otherwise. The RPS of climatology is also calculated using equation (4). The RPSS is then calculated as [*Wilks*, 1995]

$$\mathrm{RPSS} = 1 - \frac{\mathrm{RPS(forecast)}}{\mathrm{RPS(climatology)}} \tag{5}$$

The RPSS ranges from positive 1 (perfect forecast) to negative infinity. Negative RPSS values indicate that the forecast has less accuracy than climatology. The RPSS essentially measures how often an ensemble member falls into the category of the observed value and compares that to a climatological forecast.

[49] The likelihood function is also used to quantify the skill of ensemble forecasts. This function compares the likelihood of the ensemble forecast falling into the observed category with respect to climtology. The likelihood skill score for the ensemble forecast in any given year is calculated as

$$L = \frac{\prod_{t=1}^{N} \hat{P}_{j,t}}{\prod_{t=1}^{N} P_{cj,t}} \tag{6}$$

Where N is the number of years to be forecasted, j is the category of the observed value in year t, $\hat{P}_{j,t}$ is the forecast probability for category j in year t, and $P_{cj,t}$ is the climatological probability for category j in year t.

[50] The LLH values range from 0 to number of categories (three in this study). A score of zero indicates lack of skill; a score of greater than 1 indicates that the forecasts have skill in excess of the climatological forecast and a score of 3 indicates a perfect forecast. The LLH is a nonlinear measure and is related to information theory [*Rajagopalan et al.*, 2002].

7. Results

[51] Using the performance measures described above we found that a model using SWE and the Z500 index performed very well and that including the SST index did not significantly improve the performance of the model. Though SST correlations are statistically significant (Figure 3), the SST pattern is, at least in part, a response to the pressure and winds and hence provides little independent information from the Z500 index. These results corroborate the results from the GCV predictor selection method (section 6.1) which also chose the Z500 index and SWE as the best subset of predictors. Thus SWE and the Z500 index were used as predictors in the forecasting model.

[52] Figure 7 provides a visual of the ensemble forecasts over the entire period of record. In Figure 7 the April to July runoff forecast is issued on 1 April using the SWE and Z500 predictors. The ensemble forecasts are shown as boxplots, one box for each forecasted year. Asymmetry in the boxes around the median indicates a skewed ensemble forecast, a feature that is captured by the residual resampling. The ensemble forecasts typically capture the observed value within the interquartile range, providing a visual qualitative measure that overall forecast skill is good. For a quantitative measure of forecast skill, we employed the skill measures described in the previous section.

[53] To validate the use of the Z500 index, we calculated and compared the skill scores from two forecasting models: a model that uses both the Z500 index and SWE information as predictors and a model that uses only the SWE as a predictor. All three skill measures were calculated for forecasts at all lead times for both the Truckee and Carson rivers and are shown in Figure 8. The results show that using the Z500 index together with SWE as predictors provides better skills at all lead times. This is a significant outcome that clearly demonstrates the importance of incorporating basin specific large-scale climate indices in streamflow forecasts.

[54] It is also apparent from Figure 8 that the forecast skills are above climtology at all lead times (the RPSS is above zero and the LLH is above 1). This indicates the presence of useful information about the spring streamflows from as early as fall. As in most forecasting models, the skills on all the measures improve with decreased lead time.

[55] To assess the performance of the model in extreme years we calculated the RPSS and LLH for wet and dry years. For this, we define years with streamflows above the 75th percentile as wet and those below the 25th percentile as dry. Roughly 12 years fall into each category. Skills for forecasts issued on 1 April and 1 December are shown in Tables 1 and 2. It is apparent that the model has a slightly higher skill in predicting the wet years relative to dry. This asymmetry in the skills is consistent with the nonlinearities seen in the relationship between the predictors and the streamflows (Figure 5). Whereas high streamflow years exhibit a strong linear relationship with the Z500 index, this relationship breaks down, i.e., flattens out, in low streamflow years. Of course, the skill is poorer for forecasts issued on 1 December.

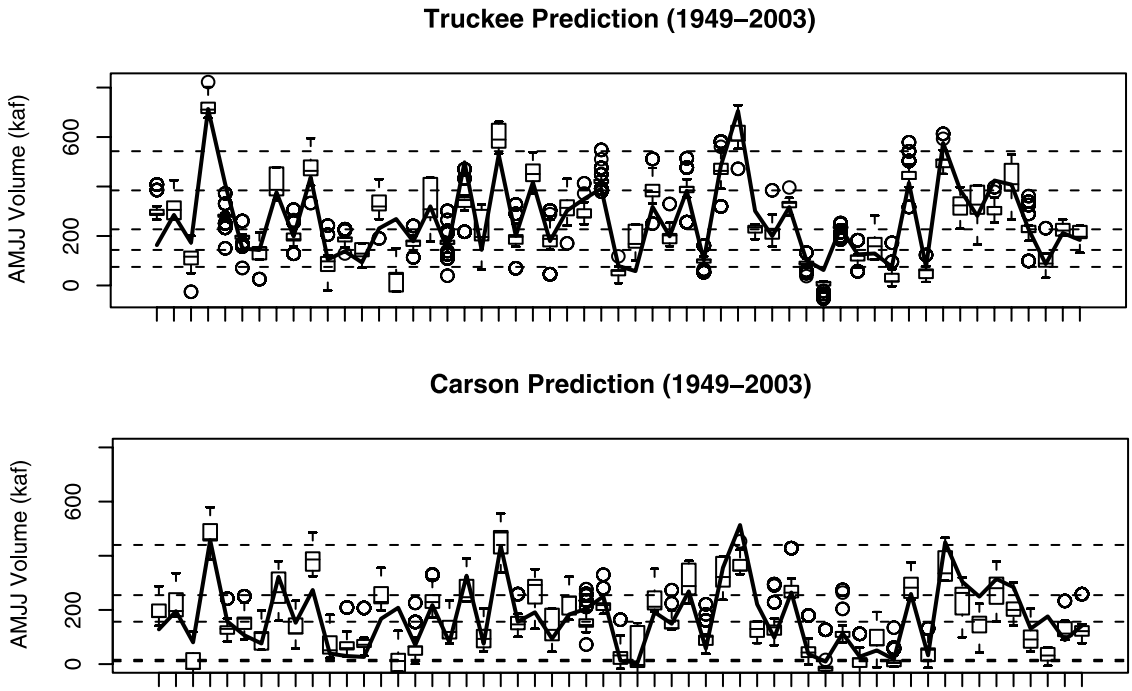

Figure 7. Time series of spring runoff with ensemble forecasts for each year (1949–2003). The thick line represents the historical time series. The box plots represent the ensemble forecast issued from 1 April in each year. The thin horizontal lines represent the quantiles of the historical data (5th, 25th, 50th, 75th, and 95th percentiles). (top) Truckee River and (bottom) Carson River.

[56] Ensemble forecasts provide a PDF, and consequently, they can be used to obtain threshold exceedence probabilities. This information is very useful for water managers. Figure 9 presents the PDF of the ensemble forecasts for 1992 and 1999, below normal and above normal streamflow years, respectively. The climatological PDF, i.e., the PDF of the historical data, is overlaid in these plots. Notice that the PDFs of the ensemble forecasts are shifted toward the observed values. In 1992, a dry year, the spring runoff in the Truckee River was 75 kaf, much below the historical average. On the basis of the climatological PDF the exceedence probability of this value is 0.92, while that from the ensemble forecasts is 0.49, much closer to the observed. Similarly, for the above average spring flow of 408 kaf in 1999, climatology suggests an exceedence probability of 0.17 while the ensemble forecasts show a much higher probability of exceedence (0.59), better capturing the probability of the observed flow value.

[57] The benefit of a forecast from fall is not that water managers know the exact volume of spring runoff, but that they have an idea of whether the coming season will be above average or below average. Because current forecasting techniques use only snowpack information, water mangers do not have the opportunity to utilize a fall forecast in their operations and decision making. BOR engineers, however, believe that a forecast in fall would be helpful in planning for the coming water season (T. Scott, U. S. Bureau of Reclamation, personal communication, 2002).

[58] As seen, incorporating identified large-scale climate predictors along with SWE information improves forecast skill. The SWE data provides important information regarding basin initial conditions, i.e., the amount of snow currently available to affect runoff. The Z500 index, however, provides information about weather yet to come in the basin, assuming that atmospheric circulation patterns persist.

8. Summary and Conclusions

[59] This paper presents a generalized framework to identify and incorporate large-scale climate information into ensemble forecasts of seasonal streamflows. The ensembles can be used to obtain streamflow threshold exceedance probabilities which are important to water resources management. The first step in the framework is to identify climate predictors that modulate the seasonal streamflows. Next, these predictors are used in stochastic local regression model to generate seasonal streamflow forecasts. The forecasting model uses a local polynomial approach for the mean forecast and residual resampling to provide ensembles. This approach is data driven with minimal assumptions unlike traditional regression alternatives. Being a local estimation scheme it also has the capability to capture any arbitrary relationship exhibited in the data. The final step in the framework presented here is to validate and test the forecasting model.

[60] The framework was applied to the Truckee and Carson River Basins located in the Sierra Nevada Mountains. Large-scale climate features in the Pacific Ocean during the preceding fall and winter seasons were found to be significantly related to the spring streamflows in the basins, thus enhancing the prospects for long-lead forecasts. In particular, the fall and winter 500 mbar geopotential

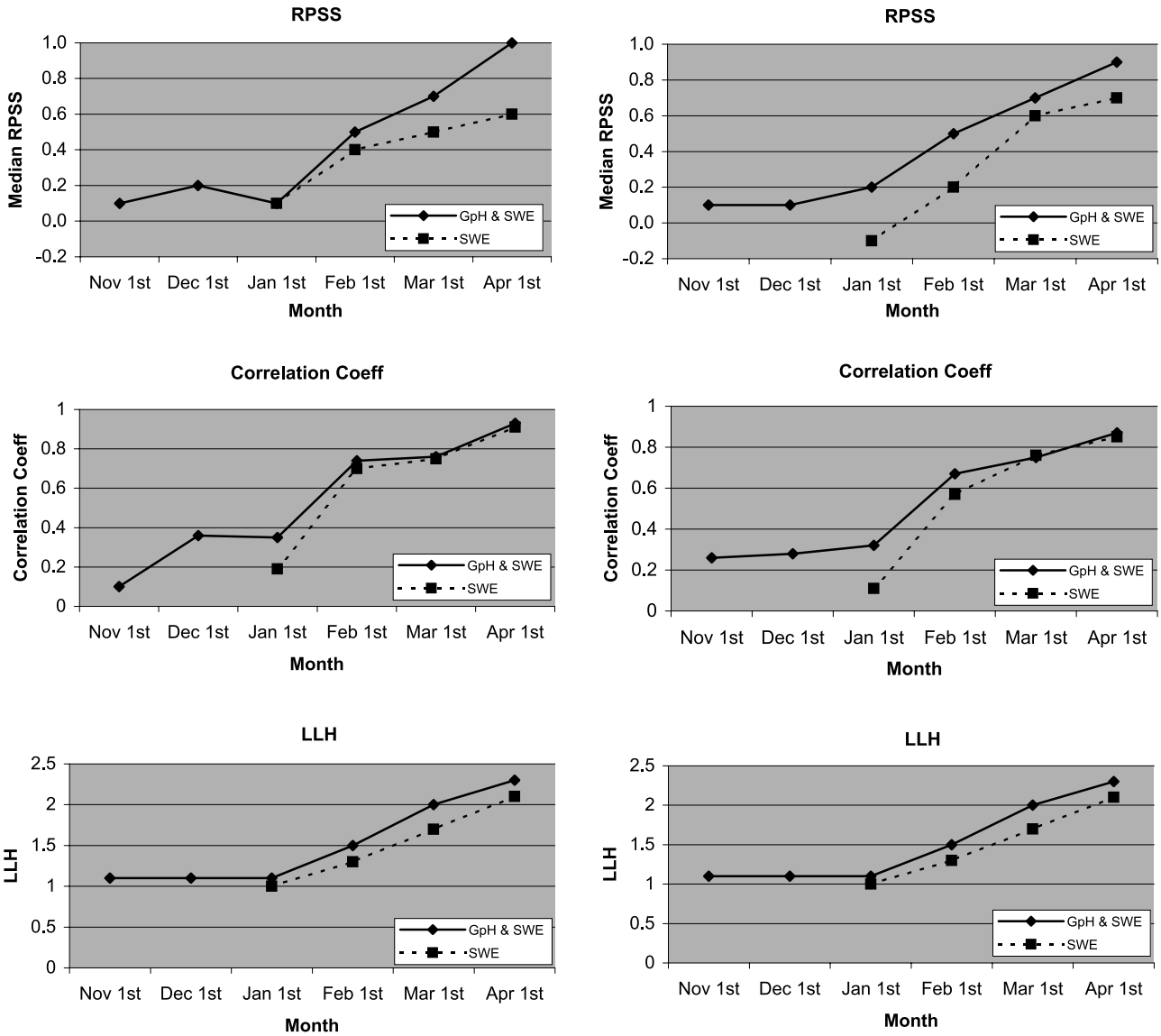

Figure 8. Skill scores of forecasts issued from the first of each month November–April for (left) Truckee and (right) Carson rivers.

height fields in the region off the coast of Washington were found to modulate the direction of the winds and the amount of moisture carried to the Truckee and Carson basins and hence the total spring runoff in the basins. Significant forecast skills at long lead times (up to 5 months) were obtained by incorporating the large-scale climate information together with the SWE as predictors. Interestingly, this was true even for the forecasts issued on 1 April. Skills for the longest–lead time forecasts were greater than climatology and the skills progressively increased with the decrease

Table 1. Skill Measure of the Ensemble Forecast Issued on 1 April in All Years, Wet Years, and Dry Years for Truckee and Carson Rivers

	Median Skill Score			
	RPSS		LLH	
	Truckee	Carson	Truckee	Carson
All years	1.0	0.9	2.3	2.3
Wet years	1.0	1.0	3.0	2.6
Dry years	0.9	0.8	2.2	2.2

Table 2. Skill Measure of the Ensemble Forecast Issued on 1 December in All Years, Wet Years, and Dry Years for Truckee and Carson Rivers

	Median Skill Score			
	RPSS		LLH	
	Truckee	Carson	Truckee	Carson
All years	0.2	0.0	1.1	1.1
Wet years	0.4	0.3	1.1	1.2
Dry years	0.0	0.0	1.1	1.1

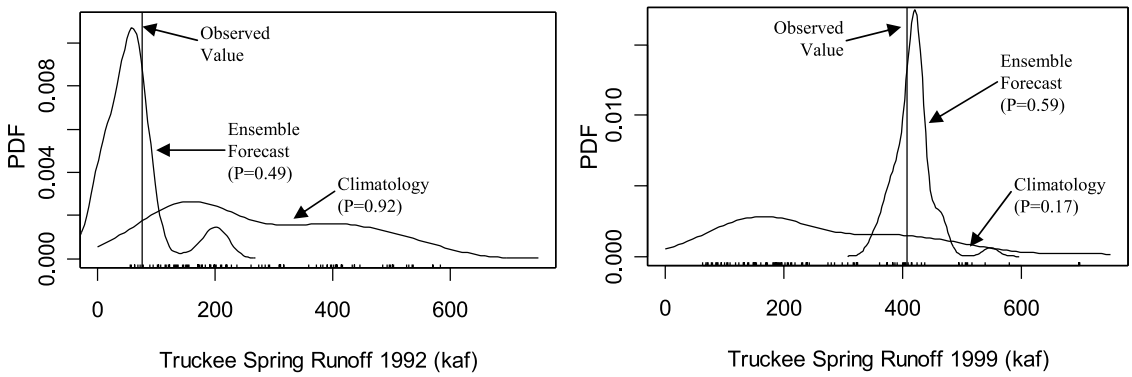

Figure 9. PDF of the ensemble forecasts in (a) a dry year (1992) and (b) a wet year (1999) for the Truckee River.

in lead time. The forecast skills were better in the extreme streamflow years. These results indicate the potential for more efficient water management in the basin. Preliminary results at demonstrating the utility of the ensemble forecasts to water management in the Truckee and Carson basins [*Grantz*, 2003] are very encouraging. Further research is underway.

[61] **Acknowledgments.** The authors thank the Bureau of Reclamation, Lahontan Basin area office, for funding this study. Funding through the CIRES Innovative Research Program at the University of Colorado at Boulder is also thankfully acknowledged. Useful discussions with Tom Scott, Tom Pagano, and Jeff Rieker are very much appreciated.

References

Allan, R., J. Lindesay, and D. Parker (1996), *El Nino, Southern Oscillation and Climatic Variability*, Commonw. Sci. and Ind. Res. Organ., Collingwood, Vic., Australia.

Bjerknes, J. (1969), Atmospheric teleconnections from the equatorial Pacific, *Mon. Weather Rev.*, *97*, 163–172.

Bowman, A., and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis,* Oxford Univ. Press, New York.

Cayan, D. R. (1996), Interannual climate variability and snowpack in the western United States, *J. Clim.*, *9*, 928–948.

Cayan, D., and R. Webb (1992), El Nino/Southern Oscillation and streamflow in the western United States, in *El Nino*, edited by H. F. Diaz and V. Markgraf, pp. 29–68, Cambridge Univ. Press, New York.

Cayan, D. R., K. T. Redmond, and L. G. Riddle (1999), ENSO and hydrologic extremes in the western United States, *J. Clim.*, *12*, 2881–2893.

Clark, M. P., M. C. Serreze, and G. J. McCabe (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*, 741–757.

Craven, P., and G. Whaba (1979), Optimal smoothing of noisy data with spline functions, *Numer. Math.*, *31*, 377–403.

Dettinger, M. D., H. F. Diaz, and D. M. Meko (1998), North-south precipitation patterns in western North America on interannual-to-decadal timescales, *J. Clim.*, *11*, 3095–4111.

Dettinger, M. D., G. J. McCabe, and J. A. Morego (1999), El Nino and the Southern Oscillation: Multiscale variability and societal impacts, in *Multiscale Hydrologic Variability Associated With El Nino–Southern Oscillation*, edited by H. F. Dica and V. Markgraf, pp. 113–147, Cambridge Univ. Press, New York.

Dracup, J. S., and E. Kahya (1994), The relationships between U. S. streamflow and La Nina events, *Water Resour. Res.*, *30*, 2133–2141.

el-Ashry, M. and D. Gibbons (1988), *Water and Arid Lands of the Western United States*, Cambridge Univ. Press, New York.

Gershunov, A. (1998), ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Implications for long-range predictability, *J. Clim.*, *11*, 3192–3203.

Grantz, K. (2003), Using large-scale climate information to forecast seasonal streamflow in the Truckee and Carson rivers, M. S. thesis, Univ. of Colorado, Boulder.

Hamlet, A. F., D. Huppert, and D. P. Lettenmaier (2002), Economic value of long-lead streamflow forecasts for Columbia River hydropower, *J. Water Resour. Plann. Manage.*, *128*(2), 91–101.

Helsel, D. R. and R. M. Hirsch (1995), *Statistical Methods in Water Resources*, Elsevier, New York.

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

Hoerling, M. P., A. Kumar, and M. Zhong (1997), El Nino, La Nina, and the nonlinearity of their teleconnections, *J. Clim.*, *10*, 1769–1786.

Horel, J. D., and J. M. Wallace (1981), Planetary scale atmospheric phenomena associated with the Southern Oscillation, *Mon. Weather Rev.*, *109*, 813–829.

Kahya, E., and J. A. Dracup (1993), U.S. streamflow patterns in relation to El Nino–Southern Oscillation, *Water Resour. Res.*, *29*, 2491–2503.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Kerr, R. A. (1998), Models win big in forecasting El Nino, *Science*, *280*, 522–523.

Lall, U. (1995), Recent advances in nonparametric function estimation: Hydrologic applications, *U.S. Natl. Rep. Int. Union Geod. Geophys. 1991–1994, Rev. Geophys.*, *33*, 1093–1102.

Lall, U., and A. Sharma (1996), A nearest neighbor bootstrap for resampling hydrologic time series, *Water Resour. Res.*, *32*, 679–693.

Loader, C. (1999), *Statistics and Computing: Local Regression and Likelihood*, Springer, New York.

Mantua, N. J., S. R. Hare, J. M. Wallace, and R. C. Francis (1997), A Pacific interdecadal climate oscillation with impacts on salmon production, *Bull. Am. Meteorol. Soc.*, *78*, 1069–1079.

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

Newman, M., G. P. Compo, and M. A. Alexander (2003), ENSO-forced variability of the Pacific Decadal Oscillation, *J. Clim.*, *16*, 3853–3857.

Owosina, A. (1992), Methods for assessing the space and time variability of groundwater data, M. S. thesis, Utah State Univ., Logan.

Prairie, J. R. (2002), Long-term salinity prediction with uncertainty analysis: Application for Colorado River above Glenwood Springs, M. S. thesis, Univ. of Colo., Boulder.

Prairie, J., B. Rajagopalan, T. Fulp, and E. Zagona (2005), Statistical nonparametric model for natural salt estimation, *J. Environ. Eng.*, *131*(1), 130–138.

Pulwarty, R. S., and T. S. Melis (2001), Climate extremes and adaptive management on the Colorado River: Lessons from the 1997–1998 ENSO event, *J. Environ. Manage.*, *63*, 307–324.

Rajagopalan, B., and U. Lall (1998), Locally weighted polynomial estimation of spatial precipitation, *J. Geogr. Inf. Decis. Anal.*, *2*(2), 48–57.

Rajagopalan, B., and U. Lall (1999), A k-nearest-neighbor simulator for daily precipitation and other weather variables, *Water Resour. Res.*, *35*, 3089–3101.

Rajagopalan, B., U. Lall, and S. Zebiak (2002), Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles, *Mon. Weather Rev.*, *130*, 1792–1811.

Rasmussen, E. M. (1985), El Niño and variations in climate, *Am. Sci.*, *73*, 168–177.

Redmond, K. T., and R. W. Koch (1991), Surface climate and streamflow variability in the western United States and their relationship to large scale circulation indices, *Water Resour. Res.*, *27*, 2381–2399.

Regonda, S., B. Rajagopalan, U. Lall, M. Clark, and Y. Moon (2005), Local polynomial method for ensemble forecast of time series, *Nonlinear Processes Geophys.*, *12*, 397–406.

Ropelweski, C. F., and M. S. Halpert (1986), North American precipitation and temperature patterns associated with El Nino–Southern Oscillation (ENSO), *Mon. Weather Rev.*, *114*, 2352–2362.

Serreze, M. C., M. P. Clark, R. L. Armstrong, D. A. McGinnis, and R. S. Pulwarty (1999), Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, *Water Resour. Res.*, *35*, 2145–2160.

Singhrattna, N., B. Rajagopalan, M. Clark, and K. Krishna Kumar (2005), Forecasting Thailand summer monsoon rainfall, *Int. J. Climatol.*, *25*, 649–664.

Souza Filho, F. A., and U. Lall (2003), Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: Applications of a mutlivariate, semiparametric algorithm, *Water Resour. Res.*, *39*(11), 1307, doi:10.1029/2002WR001373.

Tootle, G. A., and T. C. Piechota (2004), Evaluation of climate factors to forecast streamflow of the Upper Truckee River, *J. Nev. Water Resour. Assoc.*, *1*(1), 7–19.

Wilks, D. (1995), *Statistical Methods in the Atmospheric Sciences*, Elsevier, New York.

Yarnal, B., and H. F. Diaz (1986), Relationships between extremes of the Southern Oscillation and the winter climate of the Anglo-American Pacific coast, *J. Climatol.*, *6*, 197–219.

Yates, D. S., S. Gangopadhyay, B. Rajagopalan, and K. Strzepek (2003), A technique for generating regional climate scenarios using a nearest neighbor bootstrap, *Water Resour. Res.*, *39*(7), 1199, doi:10.1029/2002WR001769.

M. Clark, Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA.

K. Grantz and B. Rajagopalan, Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO 80309, USA. (grantz@cadswes.colorado.edu)

E. Zagona, Center for Advanced Decision Support for Water and Environmental Systems, University of Colorado, Boulder, CO 80309, USA.

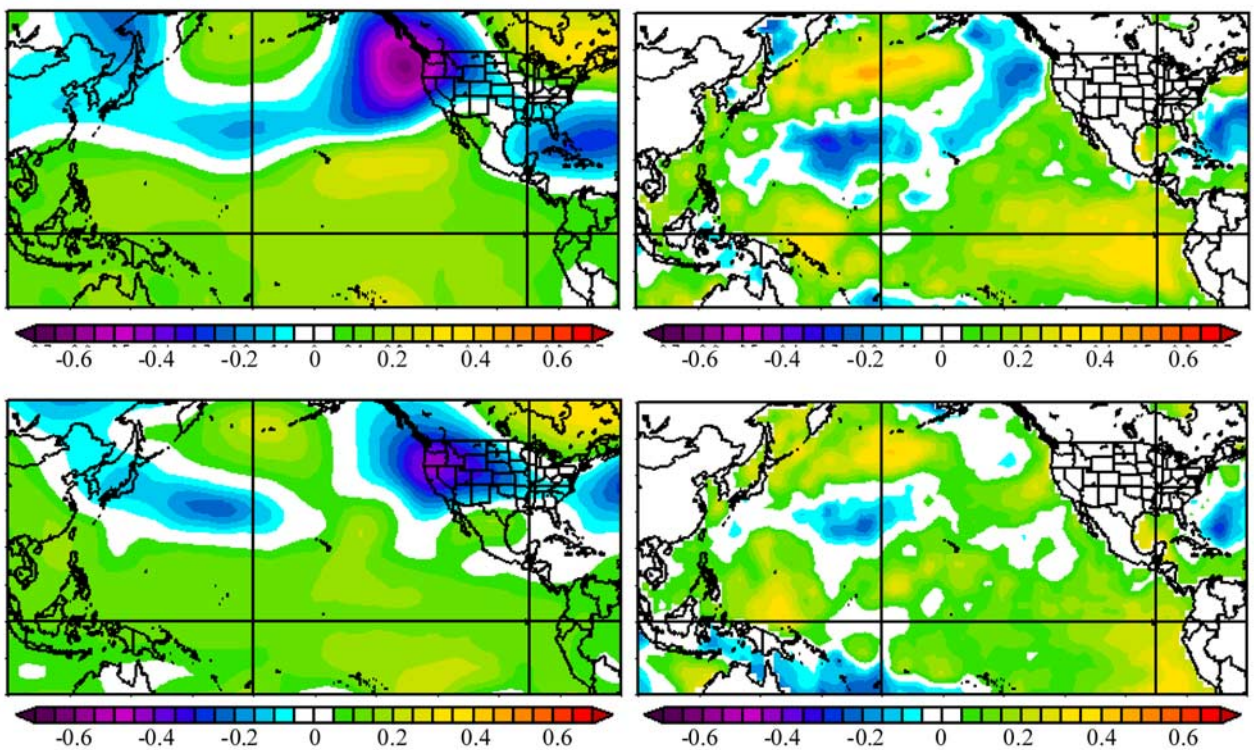

Figure 3. Correlation of Carson River spring streamflows with (top) December–February and (bottom) September–November climate variables: (left) 500 mbar geopotential height (Z500) and (right) SST.