Journal of Hydrology 410 (2011) 169–177

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Upper Colorado River and Great Basin streamflow and snowpack forecasting using Pacific oceanic–atmospheric variability

Abdoul A. Oubeidillah [a,*], Glenn A. Tootle [a], Cody Moser [a], Thomas Piechota [b], Kenneth Lamb [c]

[a] Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, 223 Perkins Hall, Knoxville, TN 37996, United States
[b] Interdisciplinary Research, University of Nevada, Las Vegas, 4505 Maryland Parkway, PO Box 4541087, Las Vegas, NV 89154, United States
[c] Department of Civil and Environmental Engineering, University of Nevada, Las Vegas, 4505 Maryland Parkway, PO Box 454015, Las Vegas, NV 89154, United States

ARTICLE INFO

Article history:
Received 21 June 2011
Received in revised form 3 September 2011
Accepted 7 September 2011
Available online 28 September 2011
This manuscript was handled by Geoff Syme, Editor-in-Chief

Keywords:
SVD
Streamflow
Snowpack
ENSO
PDO
Forecasting

SUMMARY

Water managers in western U.S., including areas such as the State of Utah, are challenged with managing scarce resources and thus, rely heavily on forecasts to allocate and meet various water demands. The need for improved streamflow and snowpack forecast models in the Upper Colorado River and Great Basin is of the utmost importance. In this research, the use of oceanic and climatic variables as predictors to improve the long lead-time (three to nine months) forecast of streamflow and snowpack was investigated. Singular Value Decomposition (SVD) analysis was used to identify a region of Pacific Ocean SSTs and a region of 500 mbar geopotential height (Z_{500}) that were teleconnected with streamflow (and snowpack) in Upper Colorado River and Great Basin headwaters. The resulting Pacific Ocean SSTs and Z_{500} regions were used to create indices that were then used as predictors in a non-parametric forecasting model. The majority of forecasts resulted in positive statistical skill, which indicated an improvement of the forecast over the climatology or no-skill forecast. The results indicated that derived indices from Pacific Ocean SSTs were better suited for long lead-time (six to nine month) forecasts of streamflow (and snowpack) while the derived indices from Z_{500} improved short-lead time (3 month) forecasts. In all, the results of the forecast model indicated that incorporating Pacific oceanic-atmospheric climatic variability in forecast models can lead to improved forecasts for both streamflow and snowpack.

1. Introduction

The State of Utah provides a critical source for streamflow and water supply with the accumulation and spring–summer melt of snowpack in several mountain ranges. One such range, the Wasatch Range, stretches approximately 260 km from the Idaho–Utah border south into central Utah. Snowpack driven streamflow from the west slope provides the primary water supply for nearly 2 million residents including Salt Lake City. Approximately 85% of Utah's population lives within 30 km of the Wasatch Range with the majority being near the west slope which contributes to the Great Basin. Equally important, the east slope of the range contributes directly to the Upper Colorado River Basin which provides water to over 25 million residents. The Uinta Mountains are located in the northeast part of the state and are noted for their unique east–west direction. Similar to the Wasatch Range, the snowpack driven streamflow from the Uinta Mountains contributes both to the Upper Colorado River Basin (south and east slopes) and the Great Basin (north and west slopes). Several smaller mountainous regions (Henry Mountains, Tushar Mountains, Pahvant Range) also provide snowpack driven streamflow to these basins.

Over the past several decades, hydrologists and climatologists have developed relationships between large scale oceanic–atmospheric variability and climate (hydroclimatology). Atmospheric–oceanic climatic and sea surface temperature (SST) variability can provide important predictive information about hydrologic variability in regions around the world. Significant research has focused on identifying Pacific atmospheric–oceanic climatic phenomena such as the El Niño–Southern Oscillation (ENSO) (Philander, 1990) and the Pacific Decadal Oscillation (PDO) (Mantua et al., 1997). Further research has identified what influence these phenomena have on US hydrology, including streamflow (Tootle et al., 2005) and snowpack (McCabe and Dettinger, 2002). These established atmospheric–oceanic climate variability relationships may be used as long lead-time (e.g., 3–9 months) predictors (forecasters) of various hydrologic responses, including streamflow and snowpack.

However, previous research efforts have failed to identify ENSO or PDO teleconnections in hydrologic (streamflow and snowpack headwaters gages) response in the majority of Utah. Tootle et al. (2005) evaluated ENSO and PDO signals in 639 unimpaired streamflow gages in the continental US which included several gages located in Utah. When applying the non-parametric rank-sum test, a statistically significant ENSO (or PDO) signal was not detected in water-year streamflow for the majority of these gages (Tootle

* Corresponding author. Tel.: +1 865 243 0342.
E-mail address: oaziz@utk.edu (A.A. Oubeidillah).

0022-1694/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2011.09.030

et al., 2005). McCabe and Dettinger (2002) evaluated 323 snowpack (April 1st snow water equivalent – SWE) stations in the western continental US and southwest Canada. Similar to the streamflow results, McCabe and Dettinger (2002) did not identify an ENSO or PDO signal in the majority of snowpack stations in Utah. Thus, the use of ENSO or PDO as a long lead-time predictor of streamflow (or snowpack) in headwaters gages in Utah will likely result in poor forecast skill.

The identification of a Pacific oceanic–atmospheric driver of headwaters gages (streamflow or snowpack) in Utah may provide a useful predictor in long lead-time forecast models. Aziz et al. (2010) evaluated western US snowpack and Pacific Ocean SSTs using the Singular Value Decomposition (SVD) statistical method and identified a region in the Pacific Ocean that was teleconnected with snowpack in many of Utah's mountain ranges. Lamb et al. (in press, submitted for publication), again using SVD, identified a similar Pacific Ocean region when evaluating streamflow in these same ranges. A similar pattern was identified in previous research efforts (e.g., Rajagopalan et al., 2000; Wang and Ting, 2000; Grantz et al., 2005; Tootle and Piechota, 2006; Soukup et al., 2009) but, the importance of this Pacific Ocean SST region as a possible driver of hydrology in Utah headwaters was not acknowledged.

In addition to Pacific Ocean SSTs, Grantz et al. (2005) and Soukup et al. (2009) identified 500 mbar geopotential height (Z_{500}) values as a useful long lead-time predictor of streamflow. Geopotential height is the height to the pressure zone of interest, as measured above the mean sea surface elevation. Blackmon et al. (1977) explored the behavior of the 500 mbar wind statistics upon northern hemisphere wintertime circulation. The results of these studies suggested that Z_{500} index values can be attributed to substantial impacts on climate. Xoplaki et al. (2000) determined that the link between precipitation over Greece and changes in large scale atmospheric circulation were strong in relation to 500 mbar geopotential heights. Serreze et al. (1998) evaluated the relationship between snowfall and low frequency atmospheric variability and found that the troughs and ridges associated with the 500 mbar zone do play a role in the characteristics of snowfall over the eastern United States.

Streamflow (snowpack) forecasting is the process of predicting the volume of water (snow water equivalent) at a specific location for a specific time period. An alternative to typical parametric regression techniques is a non-parametric approach. Non-parametric routines avoid the usual assumption that the data comes from a normal distribution (or any specific distribution). Essentially, a non-parametric model is derived from the data and does not pre-define the form (i.e. linear or non-linear) of the function. Non-parametric methods have been successfully applied to streamflow forecasting. Piechota and Dracup (1999) applied non-parametric (kernel density estimator) methods to forecasting streamflow for long lead-times and showed significant improvement when comparing the results to the climatology (no skill) forecast (Piechota and Dracup, 1999). The non-parametric kernel density estimator was also successfully applied to El Niño–Southern Oscillation (ENSO) affected streams in eastern Australia and Florida (Piechota et al., 1998; Tootle and Piechota, 2004). The exceedance probability forecast developed provides an example of applying non-parametric techniques to forecasting. An exceedance probability forecast explains the likelihood that a certain streamflow volume will be equaled or exceeded during a certain period of time. Exceedance probability forecasts are used for the design and operation of water resource systems that require a high degree of system reliability (Piechota et al., 2001). The idea of developing exceedance probability forecasts of snowpack (April 1st snow water equivalent – SWE) is novel and, given the importance of snowpack in water supply, worthy of investigation.

The motivation of the current research was the Wasatch, Uinta and adjacent ranges represent an extremely challenging area in identifying traditional Pacific Ocean climate teleconnections (El Niño–Southern Oscillation and Pacific Decadal Oscillation) in hydrologic (streamflow and snowpack) response. Given this challenge and the likely inability to use ENSO and PDO as a skillful forecaster of streamflow and snowpack, SVD was utilized to verify the Pacific Ocean SST region identified in previous research efforts and this Pacific Ocean SST region was used as a predictor in a long lead-time non-parametric forecast model. SVD was also applied to identify 500 mbar geopotential height (Z_{500}) values which were also utilized as predictors. Using the Pacific Ocean SST and Z_{500} values, the forecast skill was determined for various lead-times (3–9 months) to evaluate if a skillful forecast of streamflow (or snowpack) could be obtained for a vital water supply region of the western United States.

2. Data

2.1. Streamflow

Data from seven unimpaired streamflow stations (Fig. 1, Table 1) were obtained from the US Geological Survey (USGS) National Water Information System (NWIS). They were identified from Wallis et al. (1991). The average monthly were retrieved from the NWIS website (http://waterdata.usgs.gov/nwis). The average monthly streamflow rate in cfs for April, May, June and July (AMJJ) were summed and converted into streamflow volumes using appropriate conversions. The period of streamflow volume used in the analysis was 1961–2002 (42 years).

2.2. Snowpack

April 1st snow water equivalent (SWE) data in inches for six stations (Fig. 1, Table 1) were obtained from the Natural Resources Conservation Service (NRCS) website (http://www.wcc.nrcs.usda.gov/snotel/). Each of the snowpack stations selected had complete records and were used in previous research efforts (Aziz et al., 2010; Hunter et al., 2006). Similar to streamflow, the period of record was 1961–2002 (42 years).

2.3. Pacific Ocean climate indices

Two pre-defined datasets representing Pacific oceanic–atmospheric climatic phenomena are the Niño 3.4 index and the Pacific Decadal Oscillation (PDO) index. Similar to Soukup et al. (2009), the average monthly values for the climatic indices (Niño 3.4 and PDO) were averaged for the various 6-month predictor periods January–February–March–April–May–June (JFMAMJ), April–May–June–July–August–September (AMJJAS), and July–August–September–October–November–December (JASOND). The period of record was 1960–2001, the previous year to AMJJ streamflow or April 1st snowpack. This resulted in lead-times of 3, 6 and 9 months.

The Niño 3.4 (Trenberth, 1997) SST region is located along the equatorial Pacific Ocean ($5P^{oP}S$–$5P^{oP}N$, $170P^{oP}$–$120P^{oP}W$) and monthly index data were obtained from the NOAA ESRL Physical Sciences Division (http://www.cdc.noaa.gov/Pressure/Timeseries/Nino34/). The Niño 3.4 index was used since it is an overall representation of the El Niño–Southern Oscillation (ENSO). The PDO is a oceanic/atmospheric phenomena associated with persistent, bimodal climate patterns in the northern Pacific Ocean (poleward of $20P^{oP}$ north) that oscillate with a characteristic period on the order of 50 years (a particular phase of the PDO will typically persist for about 25 years) (Mantua et al., 1997; Mantua and Hare, 2002). PDO

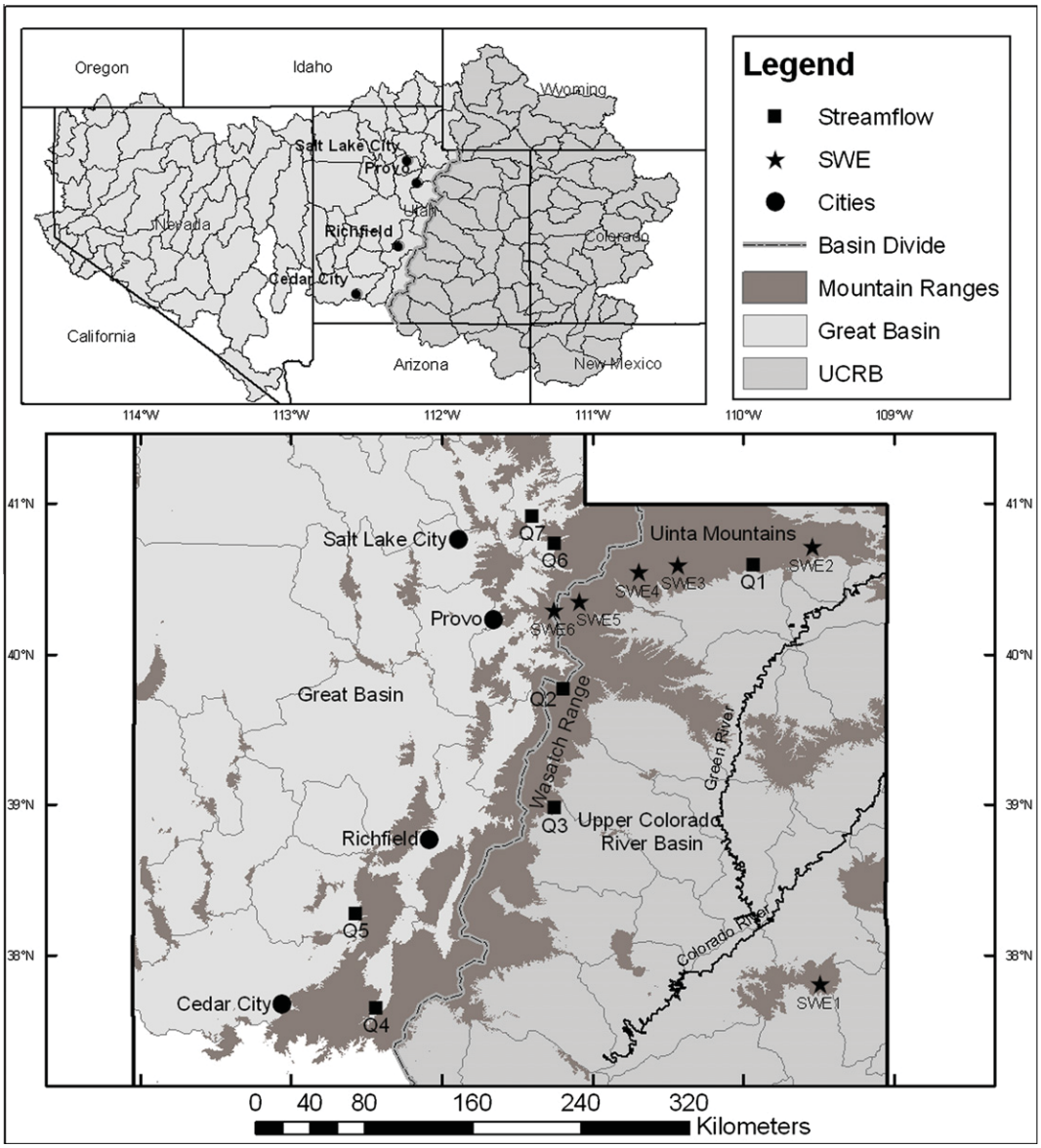

Fig. 1. Map of the study area showing the location of the streamflow gage stations and snowpack measuring stations. The identifiers correspond to Table 1.

Table 1
Streamflow and snowpack stations. The station identifier corresponds to those in Fig. 1.

Station	ID	Name	Lat.	Lon.
Q1	09299500	Whiterocks River Near Whiterocks, UT	40.59	−109.93
Q2	09310500	Fish Creek Above Reservoir, Near Scofield, UT	39.77	−111.19
Q3	09330500	Muddy Creek Near Emery, UT	38.98	−111.25
Q4	10174500	Sevier River At Hatch, UT	37.65	−112.43
Q5	10234500	Beaver River Near Beaver, UT	38.28	−112.57
Q6	10128500	Weber River Near Oakley, UT	40.74	−111.25
Q7	10131000	Chalk Creek At Coalville, UT	40.92	−111.40
SWE1	09m02s	Camp Jackson Pillow, UT	37.81	−109.49
SWE2	09j01s	King's Cabin Pillow, UT	40.72	−109.54
SWE3	10j10s	Lakefork #1 Pillow, UT	40.60	−110.43
SWE4	10j18s	Rock Creek Pillow, UT	40.55	−110.69
SWE5	11j32s	Currant Creek Pillow, UT	40.36	−111.09
SWE6	11j23s	Daniels-Strawberry Pillow, UT	40.30	−111.26

index (Mantua et al., 1997; Hare and Mantua, 2000) values were obtained from the Joint Institute for the Study of the Atmosphere and Ocean, University of Washington (http://tao.atmos.washington.edu/pdo/).

2.4. Pacific Ocean sea surface temperatures

Sea surface temperature (SST) data for the Pacific Ocean was obtained from the Kaplan SST V2 SST anomalies provided by the

NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, and obtained from their website at http://www.esrl.noaa.gov/psd/ (Kaplan et al., 1998). This dataset consists of monthly average values at a resolution of 5° by 5° and ranges from latitude 30° south to 70° north and 120° east to 80° west. This resulted in 538 active Pacific Ocean SST cells. Average Pacific Ocean SSTs were calculated for three 6-month windows (JFMAMJ, AMJJAS, and JASOND). Similar to the Nino 3.4 and PDO, this resulted in lead-times varying from 3 to 9 months.

2.5. 500 mbar geopotential height

500 mbar geopotential height (Z_{500}) data from the National Centers for Environmental Protection (NCEP) Reanalysis Derived data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, and were obtained from their web site at http://www.esrl.noaa.gov/psd/ (Kalnay et al., 1996). This dataset consisted of monthly average values at a resolution of 2.5° by 2.5° and ranges from latitude 30° south to 70° north and 120° east to 80° west. This results in 2665 active 500 mbar cells. Average 500 mbar geopotential height (Z_{500}) data were calculated for three 6-month windows (JFMAMJ, AMJJAS, and JASOND).

2.6. Data preparation

Pacific Ocean SSTs, 500 mbar geopotential height (Z_{500}) values, AMJJ streamflow and April 1st SWE anomalies, defined as the deviation of the seasonal mean from the long time average, were standardized. These standardized anomalies were used in the current research. Detrending was based on the least squares fit of a straight line to the datasets and subtracting the resulting function from the data. Detrending the Pacific Ocean SSTs, 500 mbar geopotential height (Z_{500}) values, AMJJ streamflow and April 1st snowpack data removed any trend in the data sets that may bias the analysis and mask the underlying variability (Aziz et al., 2010; Soukup et al., 2009; Tootle et al., 2008; Tootle and Piechota, 2006).

3. Methods

Pacific Ocean climate variability [ENSO, PDO, SSTs and 500 mbar geopotential height (Z_{500})] teleconnections with unimpaired streamflow were examined for various lead-times (Soukup et al., 2009; Tootle et al., 2008; Tootle and Piechota, 2006; Grantz et al., 2005). For the current research, in addition to streamflow, Pacific Ocean teleconnections with snowpack (April 1st SWE) were also examined. Singular Value Decomposition (SVD) was utilized to examine the relationships between Pacific Ocean SSTs and streamflow (or snowpack). An index was developed for the Pacific Ocean SST regions identified by SVD for the various predictor periods and predictands (streamflow and snowpack). SVD was repeated with 500 mbar geopotential height (Z_{500}) values being used in lieu of Pacific Ocean SSTs and an index was developed for the various predictor periods and predictands (streamflow and snowpack). Similar to Tootle and Piechota (2004), initially, the ENSO index and PDO index were used as predictors in a non-parametric model (Piechota et al., 2001) to determine forecast skill for streamflow (and snowpack) for the various lead-times (3–9 months). Next, the Pacific Ocean SST index was used in the non-parametric model to determine forecast skill. Finally, the 500 mbar geopotential height (Z_{500}) index was used to determine forecast skill. A comparison was made between ENSO/PDO, Pacific Ocean SSTs and 500 mbar geopotential height (Z_{500}) forecasts to determine which climatic driver provided the best skill in long lead-time forecasts of streamflow and snowpack for the various predictor periods.

3.1. Singular Value Decomposition (SVD)

While Bretherton et al. (1992) provides a detailed discussion of the Singular Value Decomposition (SVD) statistical method, a brief explanation is of this method as applied in this research is provided. A cross-covariance matrix is created from the two standardized spatial–temporal matrices used for analysis. The matrices must have the same temporal size while the spatial elements may vary. SVD is then applied to the cross-covariance matrix. This results in two orthogonal matrices, arbitrarily named Left and Right, and a singular values matrix. The left orthogonal matrix is projected into the SST (or 500 mbar) to create the SST (or 500 mb) temporal expansion series also called left temporal expansion series (LTES) and the right orthogonal matrix is projected to the snowpack (or streamflow) matrix to create the right temporal expansion series (RTES). The LTES and RTES are analogous to the principal component scores of the SST (or 500 mbar) and snowpack (or streamflow) respectively. The singular value matrix is used to create the square covariance fraction (SCF) by squaring each element of the matrix and then dividing it by the sum of all squared elements. The SCF is arranged in descending order and is analogous to the eigenvector in principal component analysis.

Heterogeneous correlation maps showing the SST (or 500 mbar) region that are significantly correlated with the snowpack (or streamflow) stations are obtained by correlating the RTES with the SST (or 500 mbar) matrix. The maps showing the snowpack (or streamflow) that are significantly correlated with the SST (or 500 mbar) are obtained by correlating the LTES with the snowpack (or 500 mbar) matrix.

SVD has the advantage of allowing two spatial–temporal matrices to interact simultaneously to identify regions that behave similarly in both fields.

3.2. Pacific Ocean SST and 500 mbar geopotential height (Z_{500}) index

The SVD analysis identified Pacific Ocean SST regions and 500 mbar regions that are significantly ($p < 0.05$) correlated with snowpack or streamflow. These cells were then averaged to create Pacific Ocean SST and 500 mbar indices for the various predictor periods and predictands. These time series were used in the non-parametric forecast model to evaluate their skills as predictors to streamflow and snowpack

3.3. Forecast methodology

The streamflow (snowpack) forecast developed is a continuous exceedance probability curve that can be used for any assumed risk level and was developed by Piechota et al. (2001). The "no skill/climatology" forecast curve is generated by dividing the rank of each historical value by the total number of years in the record. Two advantages are found using the model developed by Piechota et al. (2001): it considers the continuous relationship between the predictand and the predictor, and it does not assume a particular model structure. It suffers, however, from its semi-empiricism; fitting the model to the data points assumes that the historical data represents the entire population. A detailed description of the methodology and model can be found in Piechota et al. (2001) and Piechota et al. (1998). A brief description of the model (for one predictor) is provided below (Soukup et al., 2009):

1. The climate predictor values (P_i) for each year and the corresponding streamflow (or snowpack) predictand values (Q_i) for each year are compiled, where (P_i) represents the ENSO index, PDO index, Pacific Ocean SST index and 500 mbar geopotential height (Z_{500}) index.

2. The streamflow (or snowpack) values (Q_i) are ranked in ascending order and the corresponding climate predictor (P_i) for the corresponding year of the streamflow (or snowpack) are noted.
3. The first data point for analysis occurs immediately after the five lowest streamflow (or snowpack) values (Q_i) and the last point for analysis occurs immediately prior to the five highest streamflow (or snowpack) values (Q_i). This is required since a minimum of five values are needed to generate a probability density function.
4. The first data point for analysis is the sixth ranked streamflow (or snowpack) value (lowest to highest) based on #3 above. Using the kernel density estimator (Silverman, 1986; Piechota et al., 1998), a probability density function is developed for all climate predictor values below the first data point and a probability function is developed for all climate predictor values above the first data point. Whereas $f(x)$ is the probability density function expressed as,

$$f(x) = \frac{1}{hn}\sum_{i=1}^{n} k\left(\frac{x - x_i}{h}\right)$$

$$h_i = .9A_i n_i^{-\left(\frac{1}{5}\right)}$$

$$A_i = \min\left(\sigma_i,\ \frac{\text{interquartile range}}{1.34}\right)$$

where x_1 to x_i is a set of n observations, $k()$ is the kernel function, h is the bandwith, optimal $h = h_i$, σ_i is the stdev of predictor data in each subset i, n_i is the # of observations in each subset and the Bayes probability theorem is expressed as:

$$Prob\left(\frac{Q_i}{x}\right) = \frac{P_i f_i(x)}{\sum_{i=1}^{k} P_i f_i(x)}$$

where x = predictor value, Q_i = streamflow, P_i = prior probability streamflow, $f_i(x)$ = probability density function of prior x value
5. A unique probability value is determined for each predictor value, given the sixth ranked streamflow (or snowpack) value. These values are single points on the exceedance probability curve (probability versus streamflow or snowpack). The procedure is then repeated for the seventh ranked streamflow (or snowpack) value and so on.
6. An exceedance probability is then determined for each predictor value. The forecast curve will represent the probability of exceeding a value of streamflow (or snowpack), based on the value of the predictor.
7. The final exceedance probability forecast is found by combining the three individual forecasts into one combination forecast that has better overall skill. The combination forecast is found by applying weights a, b, and, c to the three models so that the weights add up to one. The optimal forecast is found by applying more weight to individual forecasts that better predicts streamflow and less weight to poor individual forecasts. These optimal weights are determined by an optimization procedure that evaluates the Linear Error in Probability Space (LEPS) score for all possible combinations, using weighting increments of 0.02 in which the weights vary between 0 and 1 for each model. The final combination forecast is the model with the highest LEPS score.

The skill of the forecast, as produced by the model, was measured using the Linear Error in Probability Space (LEPS) score. The LEPS score is a measure of skill that was originally developed to assess the position of the forecast and the position of the observed values in the cumulative probability distribution (non-exceedance probability); the LEPS score can be used for continuous and categorical variables (Ward and Folland, 1991; Potts et al., 1996). A modified LEPS score is required due to the absence of a convenient measure of skill for an exceedance probability forecast. A better measure of skill is one in which more weight is given to a forecast that effectively predicts low or high values and less weight to a forecast that successfully predicts average values. The application of the LEPS score is desirable here because it is less sensitive to changes near the center of the cumulative probability distribution and more sensitive to forecasts of high or low values. Essentially, it rewards a successful forecast of extreme values (Piechota et al., 2001). The developmental steps and the equations used to generate a LEPS score for an exceedance probability forecast can be reviewed in Piechota et al. (2001) and a brief description is hereby provided. In terms of probability, the LEPS score measures the distance between the forecast and observed values. First, a "no skill" or "climatology" curve was developed for the observed yearly streamflow (or snowpack) values. The "climatology" curve was created by ranking observed yearly streamflow (or snowpack) values in decreasing order (i.e., exceedance probability) of magnitude and dividing the rank of each observed value by the total number of years in the record. The LEPS score is defined as

$$S'' = 3 * (1 - |Pf - Po| + Pf2 - Pf + Po2 - Po) - 1$$

where Pf and Po are the forecasted and observed cumulative probabilities, respectively. The LEPS score was calculated for each year and "good" or "bad" forecast years were identified. The average skill (SK) is defined as

$$SK = \frac{\sum 100S''}{\sum S''_m}$$

where the summation S'' is for all years of record. If S'' is positive, S''_m is the sum of the best possible forecast (i.e. $Pf = Po$) for all years of record. If S'' is negative, S''_m is the sum of the worst possible forecast (i.e. $Pf = 1$ or 0) for all years of record. A LEPS SK score of greater than +10% is generally considered good skill.

The skill associated with each individual forecast is calculated for calibration and cross-validation analyses. The LEPS score for the calibration analysis does not provide an independent skill score because it is based on the same data in which the model was calibrated. To evaluate the skill scores, each individual yearly calibrated and cross validated LEPS skill score (CV Skill) was averaged over the entire 42 year period of record to develop an overall average forecast skill. The more rigorous CV Skill for the various predictor periods and predictands was determined and reported in the following section.

4. Results

4.1. Singular Value Decomposition (SVD)

4.1.1. Pacific Ocean SSTs and streamflow/snowpack

The SVD analysis of the relationship between Pacific Ocean SSTs and streamflow identified a similar region of significant ($p < 0.05$) cells for each of the three predictor periods (Fig. 2a). The Pacific Ocean SST region identified was similar to that reported in Aziz et al. (2010) and Lamb et al. (in press, submitted for publication). Significant SST cells in the Pacific Ocean region identified were selected for the various predictor periods (JFMAMJ, AMJJAS, and JASOND) and were averaged for each year (1960–2001). Thus, an index was developed and this index was used as a predictor in the non-parametric forecast model. The SVD analysis was repeated for Pacific Ocean SSTs and snowpack and a similar region to that of streamflow was identified (Fig. 2b). In Fig. 2b a second significant region is visible which begins over Hawaii and extends in a north easterly direction toward the west coast of the United States. This

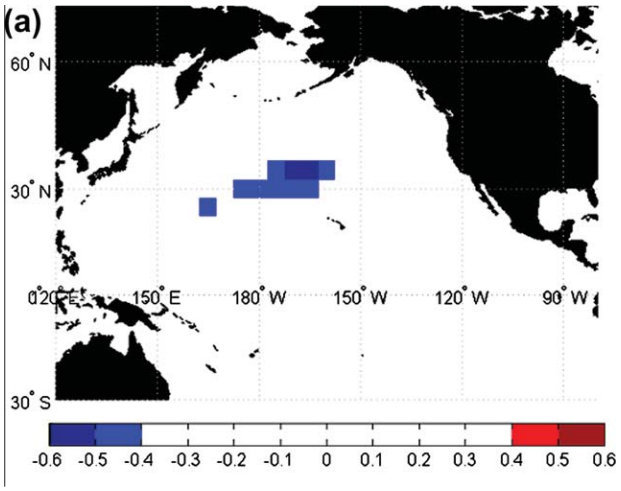

Fig. 2. Heterogeneous correlation ($p < 0.05$) figure (1st mode) when applying SVD to the previous year AMJJAS Pacific Ocean SSTs with (a) AMJJ streamflow and (b) April 1st snowpack. The color bar indicates positive or negative significant ($p < 0.5$) correlations for the SST or Z_{500} cells with snowpack or streamflow gages. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

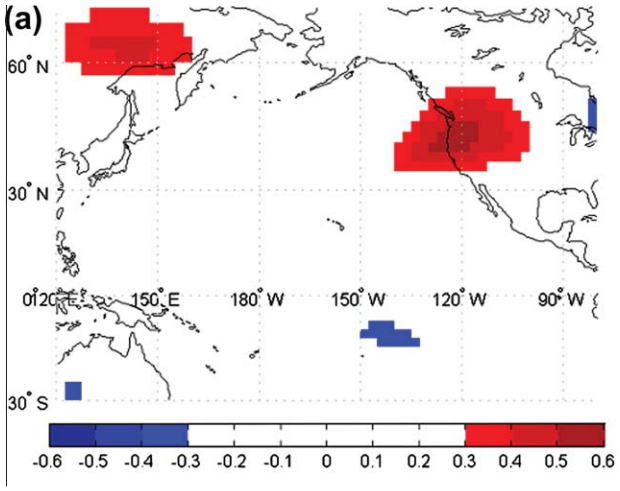

Fig. 3. Heterogeneous correlation ($p < 0.05$) figure (1st mode) when applying SVD to the previous year JASOND geopotential height at 500 mb pressure (Z_{500}) on the left side with (a) AMJJ streamflow and (b) April 1st snowpack. The color bar indicates positive or negative significant ($p < 0.5$) correlations for the SST or Z_{500} cells with snowpack or streamflow gages.

region is associated with the North American Jet Stream which is stronger during the non-winter months. This figure shows that when the Asian Jet Stream is strong, and the North American Jet is weak, we see an increase in snowpack This circulation pattern is consistent with the work shown by Lamb et al. (in press, submitted for publication). The Pacific Ocean SST region was again similar for all three predictor periods and an index was developed for use as a predictor in the non-parametric forecast model. It should be noted that the Pacific Ocean SST region displays an opposite relationship to streamflow/snowpack in that increased (decreased) SSTs result in decreased (increased) streamflow/snowpack.

4.1.2. SVD – 500 mbar geopotential height (Z_{500}) and streamflow/snowpack

The SVD analysis of the relationship between 500 mbar geopotential height (Z_{500}) and streamflow/snowpack identified a similar region of significant ($p < 0.05$) cells for each of the three predictor periods (Fig. 3a and b). Two regions were identified, one near the northeast corner of the grid over Russia and the second over the northwest US and southwest Canada. The region of northeast Asia is explained in the work by Zhang et al. (1997) who identified this region as a source for storm tracks entering the East Asian Jet Stream. The second pressure cyclone, also identified by Soukup et al. (2009) is associated with precipitation in the inter-mountain west and the CRB in work by Wang et al. (2011) and Lamb et al. (in press, submitted for publication). The two patterns both display an opposite relationship to streamflow/snowpack in that high (low) 500 mbar geopotential height (Z_{500}) results in decreased (increased) streamflow/snowpack. An index was developed and used as predictors in the non-parametric forecast model.

The results showing the correlation (r) values ($p < 0.05$) between SST and Z_{500} indices with streamflow and snowpack gages for the various predictor periods are shown in Table 3.

4.2. Forecast model

4.2.1. Climate indices (ENSO and PDO)

As displayed in Table 2, the use of ENSO and PDO as predictors consistently resulted in negative cross-validated skill for the three predictor periods, for both streamflow and snowpack. These results were expected based on previous research (Tootle et al., 2005; McCabe and Dettinger, 2002) that failed to identify a teleconnection between ENSO or PDO and hydrologic response (streamflow and snowpack) in this region. These results were also observed by Grantz et al. (2005) and Soukup et al. (2009) in the Truckee/Carson Basins and North Platte River Basin, respectively. Thus, if limited to only ENSO and PDO, a skillful forecast could not be achieved

Table 2
Cross-validated skills (CV Skill) from the results of the forecast models for different predictor periods of ENSO/PDO, Pacific Ocean SSTs and Z_{500} (JFMAMJ, AMJJAS, and JASOND) with AMJJ streamflow and April 1st snowpack.

Period station	JFMAMJ CV skill			AMJJAS CV skill			JASOND CV skill		
	ENSO/PDO	SST	Z_{500}	ENSO/PDO	SST	Z_{500}	ENSO/PDO	SST	Z_{500}
Q1	−11.0	−2.2	−1.4	−11.0	−0.2	2.4	−6.7	−8.2	−2.8
Q2	1.4	8.9	7.0	−2.5	9.8	6.3	−2.4	5.8	12.4
Q3	−9.6	4.5	5.8	−3.6	1.5	3.9	2.8	−5.8	10.6
Q4	−4.7	6.2	−0.4	−5.5	12.2	4.7	−3.1	8.8	7.8
Q5	−9.2	11.1	−1.2	−2.2	9.7	−0.4	−3.9	9.1	8.4
Q6	−6.5	3.9	12.2	−9.6	−2.7	1.9	−6.1	0.8	8.9
Q7	−6.8	2.5	8.2	0.7	−4.3	−3.5	−8.5	−2.1	8.5
SWE1	−3.5	3.5	−6.9	−2.8	14.9	5.7	−3.0	8.1	6.6
SWE2	−1.8	8.8	0.0	−4.9	4.6	3.8	−8.6	2.6	16.7
SWE3	−2.5	−5.5	−10.6	0.0	8.1	−12.0	−2.2	−3.2	6.0
SWE4	−2.3	2.9	−7.2	0.0	7.4	−0.4	0.1	−4.1	5.4
SWE5	−5.5	7.3	−1.8	−9.3	6.6	−2.3	−10.1	1.7	13.7
SWE6	−6.7	7.2	−0.7	−5.4	3.8	8.8	−4.9	−2.0	15.5
	Legend		0–5			>10			
			5–10			<0			

Table 3
Correlation (r) values ($p < 0.05$) between SST and Z_{500} indices with streamflow and snowpack gages for the various predictor periods. P1, P2 and P3 represent the three predictor periods JFMAMJ, AMJJAS, and JASOND.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	S1	S2	S3	S4	S5	S6
SST													
P1	−0.30	−0.44	−0.36	−0.52	−0.48	−0.36	−0.3	−0.45	−0.48	−0.34	−0.34	−0.47	−0.45
P2	−0.38	−0.48	−0.40	−0.60	−0.53	−0.35	−0.30	−0.56	−0.56	−0.42	−0.45	−0.50	−0.48
P3		−0.39	−0.31	−0.57	−0.48			−0.59	−0.50	−0.31	−0.35	−0.40	−0.37
Z_{500}													
P1		−0.34		−0.34	−0.32	−0.38	−0.31		−0.36	−0.28	−0.36	−0.37	−0.38
P2	−0.36	−0.37	−0.31	−0.41	−0.29	−0.34	−0.25	−0.48	−0.46	−0.40	−0.42	−0.44	−0.48
P3	−0.31	−0.49	−0.48	−0.47	−0.41	−0.48	−0.42	−0.47	−0.56	−0.37	−0.46	−0.63	−0.62

in this region. Therefore, there was a need to investigate Pacific Ocean SSTs and 500 mbar geopotential height (Z_{500}) as predictors, as identified in the SVD analysis.

4.2.2. Pacific Ocean SSTs and streamflow/snowpack

Using the Pacific Ocean SST indices determined with SVD as predictors in the non-parametric forecasting model yielded mostly positive values for streamflow and snowpack stations (Table 2). For the seven streamflow stations, the JFMAMJ forecast model CV LEPS score (skill) was positive for six stations with one station (Q5) exceeding +10. The AMJJAS and JASOND resulted in four stations displaying positive CV LEPS scores with one station (Q4) exceeding +10 for the AMJJAS predictor period. In reviewing Table 2, it appears Pacific Ocean SSTs consistently produce moderate to good skill for all three predictor periods. For the three predictor periods and seven streamflow stations (e.g., 21 forecast models), ENSO/PDO exceeded Pacific Ocean SST indices in CV LEPS scores only twice. Thus, identifying a specific Pacific Ocean SST region (teleconnection) with streamflow increased forecast skill.

For the six snowpack stations, the JFMAMJ forecast model CV LEPS score (skill) was positive for five stations. For the AMJJAS predictor period, all six stations were positive with one station (SWE1) exceeding +10. For the JASOND predictor period, only three stations were positive and it appears that the use of Pacific Ocean SSTs as a predictor for snowpack was stronger for longer lead-times (JFMAMJ and AMJJAS).

4.2.3. 500 mbar geopotential height (Z_{500}) and streamflow/snowpack

Using the Z_{500} indices determined with SVD as predictors in the non-parametric forecasting model yielded mostly positive values for streamflow and snowpack stations (Table 2). For the seven streamflow stations, the JFMAMJ and AMJJAS forecast model CV LEPS score (skill) was positive for four stations with one station (Q6) exceeding +10 in the first period. The JASOND resulted in six stations displaying positive CV LEPS scores with two stations (Q2, Q3) exceeding +10. In reviewing Table 2, it appears Z_{500} consistently produce moderate to good skill for all three predictor periods. For the three predictor periods and seven streamflow stations (e.g., 21 forecast models), ENSO/PDO exceeded Z_{500} indices in CV LEPS scores only once. Thus, identifying a specific Z_{500} region (teleconnection) with streamflow increased forecast skill.

For the six snowpack stations, no positive forecast model CV LEPS score (skill) was observed for the JFMAMJ predictor period. For the AMJJAS predictor period, only three stations showed positive skills. For the JASOND predictor period, all six stations were positive with three stations forecast skill exceeding +10. It appears that the use of Z_{500} as a predictor for snowpack was stronger for short lead-times (JASOND).

To verify the results obtained using the indices created by averaging the significant cells by year, new indices using the same significant cells were created by using principal component analysis (PCA) on the group of cells and then the principal component scores were used as indices representing the significant cells. The results obtained from the forecast model using the PC scores as indices were similar to those obtained using the cell averages. In addition, new indices were created by first weighting the significant cell values (SST or Z_{500}) by the cosine of the latitude of the cell and then taking the average of all the cells. In this case, the forecast model results were also similar.

5. Discussion and conclusions

The results of the SVD analysis identified a similar Pacific Ocean region as in Aziz et al. (2010), Soukup et al. (2009) and Grantz et al. (2005). As in Aziz et al. (2010), ENSO and PDO indices were not found to influence snowpack and streamflow in this region.

A physical explanation was provided in Grantz et al. (2005) who concluded that the SST patterns in high and low streamflow years were a direct response to pressure and winds, resulting in evaporative cooling that results in cooler than normal SSTs in this region (Aziz et al., 2010). Wang et al. (2011) went further by analyzing the moisture flux streamfunction data, which is a vertically integrated and gridded dataset, measuring the amount of moisture being transported at each grid cell. They connected the Pacific Ocean to precipitation in the Inter-mountain west, and Lamb et al. (in press, submitted for publication) identified that the same circulation pattern is teleconnected to streamflow in the CRB.

In Aziz et al. (2010), the authors argued that the region of Pacific Ocean SST identified as the driver of snowpack and streamflow could be used as predictor in a long lead-time forecast or streamflow and snowpack. The results of the non-parametric forecast model support that assessment. The current research showed that the majority of forecasts resulted in positive LEPS score (skill) and that the model skills, when the identified Pacific Ocean region SST indices were used as predictors, were the highest in the first two periods (JFMAMJ, AMJJAS). That resulted in a 6–9 month lead time forecasts for April 1st snowpack and AMJJ streamflow. It is important to note that a positive skill indicates an improvement of the forecast as compared to the climatology or "zero skill" forecast, which is none other than a distribution of the observed data.

The results of the SVD analysis of geopotential height at 500 mb (Z_{500}) with streamflow and snowpack yielded results similar to those found in Soukup et al. (2009) and Grantz et al. (2005). LEPS score (skill) of the model, when the identified Z_{500} region indices were used as predictors, were highest in the more recent period (JASOND or 3 month lead-time). These results were consistent with the ones observed in Soukup et al. (2009) and explained by the observation that winter precipitations (Wang et al., 2011; Lamb et al., in press, submitted for publication) are strongly influenced by Z_{500} and the response to the Z_{500} effects are not delayed but that

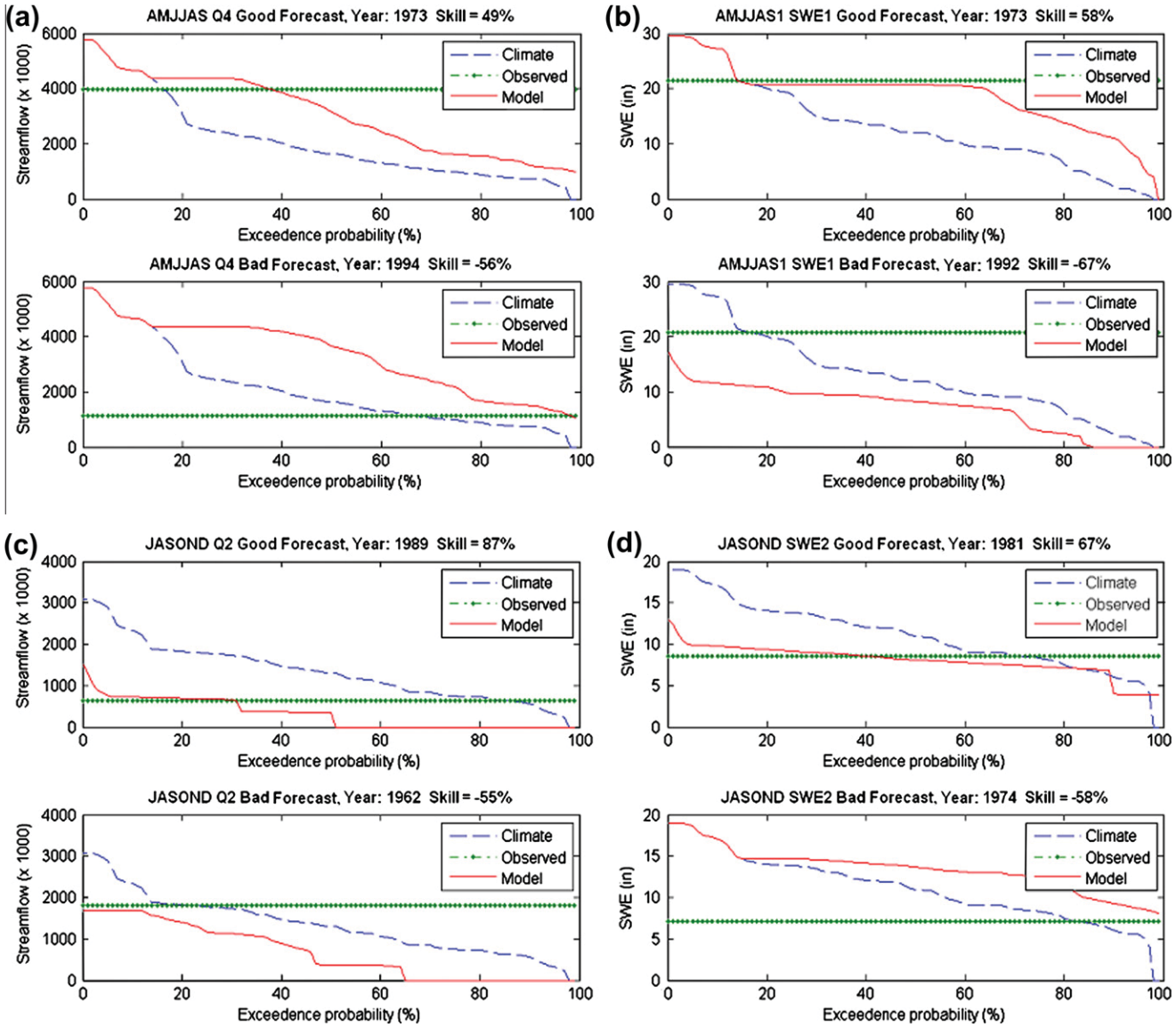

Fig. 4. Exceedence probability plots for AMJJAS Pacific Ocean SSTs with (a) Q4 and (b) SWE1 and JASOND Z_{500} with (c) Q2 and (d) SWE2. Streamflow is in cubic feet and snowpack is in inches.

instead they are more immediate. To explain how this short lag time between Z_{500} and precipitation can be related to streamflow, it was argued that the good forecast skill displayed by the Z_{500} and streamflow in this period were an indirect result of the Z_{500} induced winter precipitations, which were mainly in the form of snow, melting for a period of time before feeding streams and rivers the following Spring–Summer (Soukup et al., 2009) and thus causing a delayed effect of Z_{500} on streamflow.

One of the advantages of using the non-parametric forecast model was the development of exceedance probability forecasts with "good" and "bad" forecasts provided for both streamflow and snowpack (Fig. 4). The exceedance probability curves give a clear indication to water managers of the volume of streamflow (snow water equivalent) to expect when compared to the climatology forecast. The simplicity and clarity of these exceedence probability curves are able to give a clear picture of what to expect while at the same time showing the climatology distribution for comparison and for making informed decisions. Since the SVD analysis results did not identify an ENSO or PDO region of influence, the use of these indices as predictors in the forecast model did not provide better skill than the climatology forecast. Almost all of the forecast model skills were negative when using ENSO and PDO as predictors.

References

Aziz, O.A., Tootle, G.A., Grey, S.T., Piechota, T.C., 2010. Identification of Pacific Ocean sea surface temperature influences of the Upper Colorado River Basin snowpack. Water Resources Research 46, W07436.

Blackmon, M.L., Wallace, J.M., Lau, N.C., Mullen, S.L., 1977. An observational study of the northern hemisphere wintertime circulation. Journal of the Atmospheric Sciences 34, 1040–1053.

Bretherton, C.S., Smith, C., Wallace, J.M., 1992. An Intercomparison of methods for finding coupled patterns in climate data. Journal of Climate 5, 541–560.

Grantz, K., Rajagopalan, B., Clark, M., Zagona, E., 2005. A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts. Water Resources Research 41, W10410.

Hare, S., Mantua, N., 2000. Empirical evidence for North Pacific regime shifts in 1977 and 1989. Progress in Oceanography 47, 103–145.

Hunter, T., Tootle, G.A., Piechota, T.C., 2006. Oceanic–atmospheric variability and western US snowfall. Geophysical Research Letters 33, L13706.

Kalnay, E., Kanamitsu, M., et al., 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society 77 (3), 437–471.

Kaplan, A., Cane, M., Kushnir, Y., Clement, A., Blumenthal, M., Rajagopalan, B., 1998. Analyses of global sea surface temperature 1856–1991. Journal of Geophysical Research 103, 18567–18589.

Lamb, K.W., Piechota, T.P., Wang, S.-Y., submitted for publication. Connecting the Pacific Ocean to the Colorado River Basin. Journal of Hydrology.

Lamb, K., Piechota, T., Aziz, O., Tootle, G., submitted for publication. Establishing a basis for extending long-term streamflow forecasts in the Colorado River Basin.

Mantua, N.J., Hare, S.R., 2002. The Pacific decadal oscillation. Journal of Oceanography 59 (1), 35–44.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., Francis, R.C., 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society 78, 1069–1079.

McCabe, G.J., Dettinger, M.D., 2002. Primary modes and predictability of year-to-year snowpack variations in the Western United States from teleconnections with Pacific Ocean climate. Journal of Hydrometeorology 3, 13–25.

Philander, S.G., 1990. El Niño, La Niña and the Southern Oscillation. Academic Press Inc., San Diego, CA.

Piechota, T.C., Dracup, J.A., 1999. Long range streamflow forecasting using El Niño southern oscillation indicators. Journal of Hydrologic Engineering 4 (2), 144–151.

Piechota, T.C., Dracup, J.A., Chiew, F.H.S., McMahon, T.A., 1998. Seasonal streamflow forecasting in Eastern Australia and the El Niño–southern oscillation. Journal of Water Resources Research 34 (11), 3035–3044.

Piechota, T.C., Chiew, F.H.S., Dracup, J.A., McMahon, T.A., 2001. Development of exceedance probability streamflow forecast. Journal of Hydrologic Engineering 6 (1), 20–28.

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

Rajagopalan, B., Cook, E., Lall, U., Ray, B.K., 2000. Spatiotemporal variability of ENSO and SST teleconnections to summer drought over the United States during the twentieth century. Journal of Climate 13, 4244–4255.

Serreze, M.C., Clark, M.P., McGinnis, D.L., Robinson, D.A., 1998. Characteristics of snowfall over the eastern half of the United States and relationships with principal modes of low-frequency atmospheric variability. Journal of Climate 11, 234–250.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall, New York.

Soukup, T., Aziz, O., Tootle, G., Wulff, S., Piechota, T., 2009. Incorporating climate into a long lead-time non-parametric streamflow forecast. Journal of Hydrology 368, 131–142.

Tootle, G.A., Piechota, T.C., 2004. Suwannee river long-range streamflow forecasts based on seasonal climate predictors. Journal of the American Water Resources Association 40 (2), 523–532.

Tootle, G.A., Piechota, T.C., 2006. Relationships between Pacific and Atlantic Ocean sea surface temperatures and US streamflow variability. Water Resources Research 42, W07411.

Tootle, G.A., Piechota, T.C., Singh, A.K., 2005. Coupled oceanic/atmospheric variability and United States streamflow. Water Resources Research 41, W12408.

Tootle, G.A., Piechota, T.C., Gutierrez, F., 2008. The relationships between Pacific and Atlantic Ocean sea surface temperatures and Colombian streamflow variability. Journal of Hydrology 349 (3–4), 268–276.

Trenberth, K.E., 1997. The definition of El Niño. Bulletin of the American Meteorological Society 78, 2271–2777.

Wallis, J.R., Lettenmaier, D.P., Wood, E.F., 1991. A daily hydroclimatological data set for the continental United States. Water Resour. Res. 27 (7), 1657–1663.

Wang, H., Ting, M., 2000. Covariabilities of winter US precipitation and Pacific sea surface temperatures. Journal of Climate 13, 3711–3719.

Wang, S.-Y., Gillies, R.R., Hipps, L.E., Jin, J., 2011. A transition phase teleconnection of the Pacific quasi-decadal oscillation. Climate Dynamics 36, 681–693.

Ward, N.M., Folland, K.K., 1991. Prediction of seasonal rainfall in the North Nordeste of Brazil using Eigenvectors of sea-surface temperature. International Journal of Climatology 11, 711–743.

Xoplaki, E., Luterbacher, J., Burkard, R., Patrikas, I., Maheras, P., 2000. Connection between the large-scale 500 hPa geopotential height fields and precipitation over Greece during winter time. Climate Research 14, 129–146.

Zhang, Y., Sperber, K.R., Boyle, J.S., 1997. Climatology and interannual variation of the East Asian winter monsoon: results for the 1979–95 NCEP/NCAR reanalysis. Monthly Weather Review 125, 2605–2619.