Long Lead-Time Forecasting of U.S. Streamflow Using Partial Least Squares Regression

Glenn A. Tootle[1]; Ashok K. Singh[2]; Thomas C. Piechota[3]; and Irene Farnham[4]

Abstract: Pacific and Atlantic Ocean sea surface temperatures (SSTs) were used as predictors in a long lead-time streamflow forecast model in which the partial least squares regression (PLSR) technique was used with over 600 unimpaired streamflow stations in the continental United States. Initially, PLSR calibration (or test) models were developed for each station, using the previous spring-summer Pacific (or Atlantic) Ocean SSTs as predictors. Regions were identified in the Pacific Northwest, Upper Colorado River Basin, Midwest, and Atlantic states in which Pacific Ocean SSTs resulted in skillful forecasts. Atlantic Ocean SSTs resulted in significant regions being identified in the Pacific Northwest, Midwest, and Atlantic states. Next, streamflow stations were selected in the Columbia River Basin, Upper Colorado River Basin, and Mississippi River Basin and a PLSR cross-validation model (i.e., forecast) was developed. The results of the PLSR cross-validation model for each station varied with linear error in probability space scores of +9.5 to +51.0% where 10% is considered skillful forecasts using Pacific and Atlantic SSTs.

DOI: 10.1061/(ASCE)1084-0699(2007)12:5(442)

CE Database subject headings: Streamflow; Water temperature; Ocean water; Forecasting; Least squares method.

Introduction

Sea surface temperature (SST) variability can provide important long lead-time predictive information about hydrologic variability. In the continental United States, volumetric streamflow represents an important hydrologic parameter for water supply purposes. Water managers and planners are tasked with making critical decisions prior to the beginning of the water year (October 1st) and a skillful, long lead-time forecast of streamflow would be beneficial. Currently, long lead-time forecasts of streamflow in the continental United States are developed using both physical and statistical models.

The National Weather Service (NWS) Office of Hydrology—Advanced Hydrologic Prediction Services (http://www.nws.noaa.gov/oh/ahps/) provides a physically based prediction of streamflow for numerous rivers in the continental United States. This is termed the ensemble streamflow prediction (ESP), which utilizes data from various sources including radar, reservoir releases, river gauges, and historical/forecasted climate. These data are input into a physical hydrologic model that generates the streamflow prediction (i.e., forecast). ESP forecasts provide an exceedance probability curve of the predicted streamflow. An exceedance probability is defined as the probability that the specified value (i.e., streamflow) will be equal to or exceeded during a time period. An exceedance probability forecast can be used depending on an assumed level of risk. For example, a water manager may choose to take a 10% risk, which would correspond to a streamflow value that has a 90% probability of exceedance. A continuous exceedance probability forecast can be made by several methods including principal component analysis, regression, and linear discriminant analysis (Piechota et al. 2001).

Nonparametric methods, which do not predefine the form (i.e., linear or nonlinear) of the function, have been successfully applied to streamflow forecasting. Lall (1995) performed a detailed review of applications of nonparametric probability uses in stochastic hydrology. Piechota and Dracup (1999) applied nonparametric (kernel density estimator) methods to forecasting streamflow for long lead times. Significant improvement was found when comparing the results to the climatology (no skill) forecast (Piechota and Dracup 1999). The nonparametric 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). Several other nonparametric methods (K nearest neighbor local polynomials and local weighted polynomials) have been successfully applied to hydrologic (and streamflow) forecasting (Lall and Sharma 1996; Rajagopalan and Lall 1999; Souza and Lall 2003). In addition to nonparametric models, various statistically based regression methods (e.g., multiple linear regression, principal component regression, partial least squares regression) exist that can be applied to various disciplines, including hydrology (streamflow forecasting). Specifically, partial least squared regression methods have been used in chemometric studies that resulted in the successful identification of relationships between variables.

Chemometrics is the study of the interconnections of chemical and physical properties of compounds (Malinowski 2002). A com-

[1]Dept. of Civil and Architectural Engineering, Univ. of Wyoming, Dept. 3295, 1000 E. University Ave., Laramie, WY 82071. E-mail: tootleg@uwyo.edu

[2]Hotel Management, Univ. of Nevada, Las Vegas, 4505 Maryland Pkwy., Box 454020, Las Vegas, NV 89154-4020. E-mail: aksingh@unlv.nevada.edu

[3]Dept. of Civil and Environmental Engineering, Univ. of Nevada, Las Vegas, 4505 Maryland Pkwy., Box 454015, Las Vegas, NV 89154-4015 (corresponding author). E-mail: piechota@unlv.nevada.edu

[4]S.M. Stoller Corporation, 7710 W. Cheyenne Ave., Las Vegas, NV 89129.

Note. Discussion open until February 1, 2008. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on June 29, 2005; approved on July 24, 2006. This paper is part of the ***Journal of Hydrologic Engineering***, Vol. 12, No. 5, September 1, 2007. ©ASCE, ISSN 1084-0699/2007/5-442–451/$25.00.

mon method of determining interconnections between variables is multiple regression. However, a concern when using multiple regression is when the predictor variables are not independent and are collinear. This can result in poor model prediction due to multicollinearity in predictor data. Principal component regression (PCR) is typically utilized to account for collinearity issues and has been successfully applied to streamflow forecasting (e.g., Eldaw et al. 2003). Eldaw et al. (2003) identified seasonal values of SSTs (i.e., regions) that were highly correlated with seasonal Nile River streamflow for several long lead times. These regions were then used in a multiple-regression model to forecast streamflow. Next, principal component analysis was used to identify SST regions and PCR was used to develop streamflow forecast models. The PCR streamflow forecast models showed significant improvement over the multiple-regression models (Eldaw et al. 2003). Whereas PCR is widely used in hydrology (streamflow forecasting), partial least squares regression (PLSR) is an improved technique that has gained popularity in the field of chemometrics and is directly applicable in streamflow forecasting.

PLSR differs from PCR in that the PLSR model is based on the principal components of both the predictor (i.e., SSTs) and the predictand (i.e., streamflow). In PLSR, the principal component scores of both SSTs and streamflow are used in lieu of the original data to develop the regression model. This is an attractive feature of PLSR and could result in improved model skill. Herman Wold developed PLSR in the late 1960s for use in the field of econometrics (Wold 1966). PLSR gained importance in the field of chemistry during the 1970s (Gerlach et al. 1979). Svante Wold continued the work of his father (Herman Wold) with several PLSR applications in chemistry (Wold 1978; Wold et al. 1987). Geladi and Kowalski (1986) developed a PLSR tutorial, which outlines the nonlinear iterative partial least squares (NIPALS) and partial least squares (PLS1) algorithms used in PLSR. Frank and Friedman (1993) provided a detailed comparison of several techniques used in chemometrics, including multiple regression, PCR, and PLSR. Currently, PLSR is widely used in a variety of applications including the determination of soil properties and soil contaminants (Sorensen and Dalsgaard 2005; Wu et al. 2005).

The goal of the current research presented is to develop an improved long lead-time streamflow forecast with the main contribution being the first-time application of PLSR to streamflow forecasting. Utilizing Pacific and Atlantic Ocean SSTs as predictors, PLSR calibration models were developed for over 600 streamflow stations in the continental United States. For selected streamflow stations in the continental United States, PLSR cross-validated forecasts were developed for yearly, volumetric streamflow. For both the PLSR calibrated and cross-validated models, model skill was evaluated to determine if PLSR is a skillful method for long lead-time streamflow forecasting.

Data

The major data sets used to develop the relationships between continental U.S. streamflow and oceanic SST variability were unimpaired streamflow data for the continental U.S. and oceanic SST data for the Pacific and Atlantic Oceans.

Streamflow Data

Unimpaired streamflow stations (1,009) were identified from Wallis et al. (1991) and, utilizing the U.S. Geological Survey (USGS) NWISWeb Data retrieval (http://waterdata.usgs.gov/nwis/), the period of record was extended from 1988 to 2002. This resulted in 639 stations having monthly flow rate data for the period from 1951 to 2002 (Tootle et al. 2005) (Fig. 1). The reduction of 370 (1,009 minus 639) unimpaired streamflow stations was a result of the data not being updated on the USGS website and missing data. A review of the USGS NWISWeb resulted in 172 stations not having updated data, 184 stations missing a year (or multiple years) of data, and 14 stations missing both updated and a year (or multiple years) of data. However, extending the period of record was important because it provided both recent data and increased the number of years used when performing the PLSR. The average monthly streamflow rates (in cubic feet per second—cfs) were averaged for the water year (October of the previous year to September of the current year) and converted into streamflow volumes (km^3) with proper conversions. Water year streamflow data covering a period from 1951 to 2002 (52 years) were then used in the following analysis.

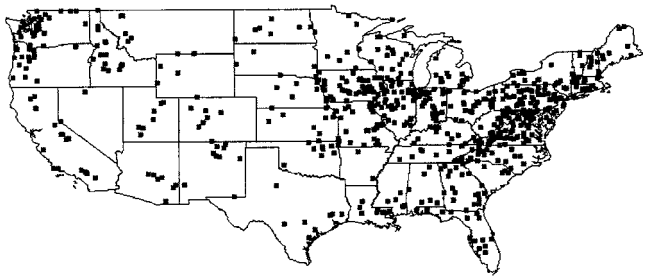

Fig. 1. Location of unimpaired U.S. Geological Survey streamflow stations in the continental United States (1951–2002)

Pacific and Atlantic Ocean SST Data

SST data for the Pacific and Atlantic Oceans were obtained from the National Climatic Data Center (http://www.cdc.noaa.gov/cdc/data.noaa.ersst.html). The oceanic SST data consist of average monthly values for a 2 by 2° grid cell (Smith and Reynolds 2002). The extended reconstructed global SSTs were based on the Comprehensive Ocean-Atmosphere Data Set (COADS) from 1854 to present (Smith and Reynolds 2002). A quality control procedure was developed by Smith and Reynolds utilizing a base period (1961–1991) to develop the reconstructed SSTs back to 1854. The uncertainty in the reconstructed data decreases through most of the period (1854 to present) with the smallest uncertainty after 1950 (Smith and Reynolds 2002) due to improved aerial coverage of the oceans.

The region of Pacific Ocean SST data used for the analysis was longitude 240°W to longitude 80°W and latitude 20°S to latitude 60°N whereas the region of Atlantic Ocean SST data used for the analysis was longitude 80°W to longitude 0° and latitude 20°S to latitude 60°N. These regions represent the majority of atmospheric/oceanic influence on U.S. climate (i.e., storm tracks such as Pacific Ocean frontal storms).

Methods

A brief discussion of PLSR and its application (in the current research) to long lead-time continental U.S. streamflow forecasting is provided.

PLSR

Eigenanalysis, which is the basis for principal factor analysis (PFA) or principal component analysis (PCA), is a procedure for decomposing matrices and calculating eigenvalues and eigenvectors. The four methods most frequently used for eigenanalysis are the power method, the Jacobi method, singular value decomposition (SVD), and NIPALS (Malinowski 2002). For example, the PCA of the matrix (X) (i.e., the matrix of predictors or independent variables) decomposes (X) into a score matrix (T) times a loading matrix (P) and a residual (i.e., error) matrix (E) (Wold et al. 1987)

$$X = T^{*}P' + E \tag{1}$$

The score and loading plots that result from the decomposition of (X) provide information about the systematic structure in (X). PCA is equivalent to SVD and is used to compute the eigenvectors of the covariance matrix $(X'X)$ or the association matrix (XX'). When concerned with only the first few principal components, NIPALS is advantageous due to calculation speed and simplicity (Wold et al. 1987).

To develop a prediction (i.e., forecast), PCA is commonly combined with multiple linear regression (MLR) when the number of predictors (X) exceeds the number of predictands (Y). When MLR is used alone with a large number of predictors, the calibration (or test) model results in a good fit for the sample data. However, for new data (verification), the MLR model results in very large standard deviations of the estimates. Inflation of the standard deviations and the estimates is a result of multicollinearity, which occurs when several of the predictors are highly correlated with each other. To reduce the effect of multicollinearity, PCA is performed on (X) to reduce the number of predictors and eliminate the collinearity between predictors. Next, MLR is performed on (Y) using the scores obtained in the PCA of (X). This method is commonly referred to as principal component regression (PCR).

PLSR differs from PCR in that the PLSR model is based on the principal components of both the predictor (i.e., independent variable) (X) and the predictand (i.e., dependent variable) (Y). PLSR generalizes and combines features from both PCA and MLR (Abdi 2003). PLSR is especially useful when there is a need to provide a prediction from a very large set of independent (predictor) variables (Abdi 2003). In PLSR, the principal component scores of both (X) and (Y) are used in lieu of the original data to develop the regression model. PLSR identifies components from (Y) that are also relevant for (X) (Abdi 2003). The generalization step results in PLSR searching for a set of components (latent vectors) that explains the maximum covariance between (Y) and (X) which is followed by a regression step where the decomposition of (X) is used to predict (Y) (Abdi 2003).

As with PCA in Eq. (1), (X) is decomposed into a score matrix (T) times a loading matrix (P) and a residual matrix (E) (Fig. 2). Similarly, (Y) is decomposed into a score matrix (U) times a loading matrix (R) and a residual matrix (F)

$$Y = U^{*}R' + F \tag{2}$$

These equations are commonly referred to as the outer relations (Geladi and Kowalski 1986). The objective of the PLSR model is to minimize (F) while maintaining the correlation between (X) and (Y), referred to as the inner relation U (Geladi and Kowalski 1986)

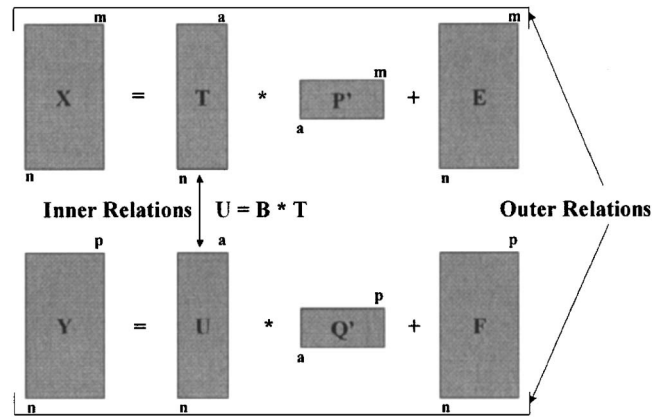

Fig. 2. PLS analysis of the relationship between X and Y matrices [adapted from Geladi and Kowalski (1986)]

$$U = B^{*}T + H \tag{3}$$

where (H) represents the error and (B)=diagonal matrix explaining the correlation between (X) and (Y). When Eq. (3) is inserted into Eq. (2), a predictive relation for (Y) is developed where (F^{*}) represents the error

$$Y = T^{*}R'B + F^{*} \tag{4}$$

Eq. (4) is sometimes referred to as the mixed relation where (F^{*}) is to be minimized (Geladi and Kowalski 1986). To perform PLSR, several methods are available including the previously mentioned SVD and NIPALS methods. The NIPALS iterative approach results in the blocks (i.e., X and Y) receiving scores from each other and thus improving the inner relation.

If the PLSR model is to be used for prediction, it is important to determine the optimal number of components needed to develop the forecast model. The prediction residual sum of squares (PRESS) statistic is a cross-validation calculation that determines the minimum number of components required (Geladi and Kowalski 1986). The cross validation consists of removing a row (or multiple rows) from the data matrix and then completing the eigenanalysis on the reduced matrix. Target testing is then performed on the removed rows using the various levels of the abstract factor space and the difference between the target points and the predicted points is calculated (Malinowski 2002). This process is repeated until every row has been deleted once and the errors in the target fit for each row are summed (Malinowski 2002). The PRESS(j) statistic is calculated for each of the j factor levels using

$$\mathrm{PRESS}(j) = \sum_{i=1}^{r}\sum_{k=1}^{c}(d_{ik} - \hat{d}_{ik}(j))^2 \tag{5}$$

where $\hat{d}_{ik}(j)$ and d_{ik}=predicted and actual values, respectively, of the deleted rows obtained with j factors and r and c=matrix dimensions (Malinowski 2002). There are several methods for using the PRESS statistic to determine the optimal number of components. The most popular method is using the PRESS statistic with the minimum value (Malinowski 2002).

A detailed discussion of the NIPALS method, including the PRESS statistic, and its use in PLSR is provided in Wold (1966); Geladi and Kowalski (1986); Wold et al. (1987), and Malinowski (2002). Of note, PLSR, applying the NIPALS method, has recently been made available in the latest release (Version 14) of the MINITAB statistical software package.

PLSR Streamflow Forecasting

Pacific (or Atlantic) Ocean SSTs were used as predictors in PLSR calibration and cross-validation models. A long lead-time approach was adopted such that spring–summer (April to September) season SSTs were used to forecast the following water year (October to September) streamflow volume for stations in the continental United States. PLSR calibration and cross-validation models were developed and are described in the following.

PLSR Calibration Model

A PLSR calibration (or test) model was developed for each streamflow station (639 stations) for Pacific (or Atlantic) Ocean SSTs. Spring–summer SSTs for the Pacific (or Atlantic) Ocean were used to develop a streamflow regression equation for each (i.e., individual) streamflow station. This procedure is referred to as PLS1 (Malinowski 2002). The number of components used in the PLSR calibration model was preselected at 25 and remained constant for all 639 stations. Whereas the previously discussed PRESS statistic can be used to optimize the number of components, the consistent use of 25 components allowed for an unbiased comparison of the PLS1 procedure for all 639 stations. Obviously, the PLSR calibration model skill can be improved if the number of components used to develop the regression equation for each station utilized the PRESS statistic. However, the PLSR calibration model was developed to identify regions that show predictability and was not used to develop a forecast. The PLSR calibration model skill (coefficient of determination—R^2 or R-squared) was determined for each station for Pacific (or Atlantic) Ocean SSTs. An R-squared value exceeding 0.80 is generally considered good skill.

PLSR Cross-Validation Model

A PLSR cross-validation model was developed for selected streamflow stations. This procedure was not performed on all 639 stations as it would have required excessive computer time and the intent of the current research was to determine if PLSR was applicable in the selected regions. The stations selected included:

- Columbia River Basin—USGS Station No. 13317000—Salmon River at White Bird, Id.;
- Upper Colorado River Basin—USGS Station No. 09304500—White River near Meeker, Colo.; and
- Mississippi River Basin—USGS Station No. 06775500—Middle Loup River at Dunning, Neb.

These stations were selected based on their location in regions of interest (i.e., Columbia River Basin, Upper Colorado River Basin, and Mississippi River Basin) for water managers and planners. Unlike the PLSR calibration model, the PLSR cross-validation model applied the PRESS statistic to determine the optimum number of components used to develop a streamflow regression equation for each (i.e., individual) streamflow station. Initially, 10 years were randomly chosen and removed, and the PRESS statistic (i.e., optimum number of components) was determined.

Next, the cross-validation model utilized a "drop one" approach in which the model removes a year, calibrates the model on the remaining 51 years, and forecasts the streamflow for the year removed. Yearly forecasted (and actual) water year streamflow volumes were reported.

Finally, as previously discussed, the loading matrix P from the decomposition of X (i.e., SSTs) provides useful information about which SST cells (i.e., regions) influence the streamflow station being forecasted. SST loading maps were developed for both Pacific and Atlantic Ocean SSTs displaying significant (loading value exceeded + or −0.5%) SST cells used to develop the PLSR model.

There are various measures available to determine forecast skill including Nash–Sutcliffe (Nash and Sutcliffe 1970) ranked probability skill score (Wilks 1995). In the research presented here, the forecast skill was determined by the linear error in probability space (LEPS) score (Ward and Folland 1991; Potts et al. 1996). First, a "no skill" or "climatology" curve was developed for the observed yearly streamflow values. The climatology curve was created by ranking observed yearly streamflow 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 a measure of skill that was developed originally to assess the position of the forecast and the position of the observed values in the cumulative probability distribution (nonexceedance probability); the LEPS score can be used for continuous and categorical variables (Ward and Folland 1991; Potts et al. 1996). Potts et al. (1996) describe the advantages of the LEPS score over traditional skill measurements such as root-mean-square error and a brief description is hereby provided. In terms of probability, the LEPS score measures the distance between the forecast and observed values. The LEPS score is defined as

$$S'' = 3^*(1 - |P_f - P_o| + P_f^2 - P_f + P_o^2 - P_o) - 1 \tag{6}$$

where P_f and P_o=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

$$\mathrm{SK} = \frac{\sum 100S''}{\sum S''_m} \tag{7}$$

where the summation S'' is for all years of record. If S'' is positive, S''_m=sum of the best possible forecast (i.e., $P_f = P_o$) for all years of record. If S'' is negative, S''_m=sum of the worst possible forecast (i.e., P_f=1 or 0) for all years of record. A LEPS SK score of greater than +10% is generally considered "good" skill. The LEPS SK score has been previously utilized as a measure of skill in streamflow forecast models (Piechota et al. 1998; Piechota and Dracup 1999; Tootle and Piechota 2004). An application of the LEPS SK score is hereby provided for the White River.

Initially, the climatology curve was developed for the observed streamflow values. Next, a similar curve was developed for the predicted streamflow values. This resulted in the year 1974 having an observed streamflow curve value (P_o) of 3.8% and a predicted streamflow curve value (P_f) of 7.5%. Thus, the predicted streamflow value was relatively close to the observed streamflow value (i.e., good forecast). Inserting P_o and P_f into Eq. (6) results in a value (S'') of 1.57. The year 1983 had an observed streamflow curve value (P_o) of 22.6% and a predicted streamflow curve value (P_f) of 81.1%. Again, inserting P_o and P_f into Eq. (6) results in a value (S'') of −0.74 (i.e., bad forecast). Each yearly S'' was then summed and, if a positive number, divided by the sum of the best possible forecast to determine the LEPS SK score.

Note, the LEPS SK score methodology rewards good forecasts of extreme events. For example, the year 1981 had an observed streamflow curve value (P_o) of 54.7% and a predicted streamflow curve value (P_f) of 60.4%. With the predicted streamflow value being relatively close to the observed streamflow value, one

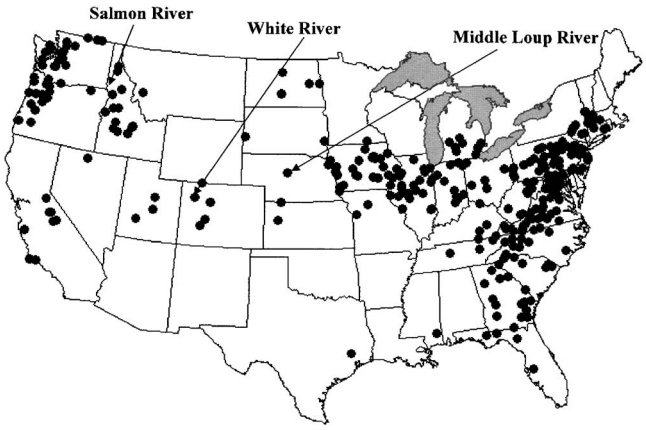

Fig. 3. PLSR calibration model results (R-squared values) for Pacific Ocean SSTs and continental U.S. streamflow. Black dots represent R-squared values greater than 0.80. Stations used in PLSR cross validation are identified.

Fig. 4. PLSR calibration model results (R-squared values) for Atlantic Ocean SSTs and continental U.S. streamflow. Black dots represent R-squared values greater than 0.80.

would acknowledge a good forecast. However, inserting P_o and P_f into Eq. (6) results in a value (S'') of 0.37, much less than the 1.57 value determined for the year 1974.

Forecast uncertainty was examined by two different approaches. Initially, for each year, the cumulative exceedance probability difference between the forecasted streamflow and actual streamflow was reported. For a given year, this value shows how good the streamflow forecast was, in terms of comparing it to the actual streamflow, based on the cumulative exceedance probability curve. Finally, confidence intervals (5 and 95%) were computed for each year forecasted. For the year being forecasted (e.g., 1951), that year (e.g., 1951) and one additional year (e.g., 1952) were removed and a forecast was developed. This procedure was repeated, again removing the year being forecasted (e.g., 1951) and one additional year (e.g., 1953) previously not removed. For the 52 years of record, 50 forecasts were developed for each year and a normal distribution based confidence interval was completed.

Results

PLSR Calibration Model of Continental U.S. Streamflow

Pacific and Atlantic Ocean SSTs were evaluated independently for the PLSR calibration model and the results (R-squared values) are hereby provided. It is noteworthy that the calibration model for each station utilizes different Pacific (or Atlantic) Ocean SST regions. These regions were identified for the selected streamflow stations for further analysis.

Pacific Ocean Sea Surface Temperatures (SSTs)

Several streamflow regions were identified in the continental United States in which Pacific Ocean SSTs, when used as predictors in the PLSR calibration model, achieved R-squared values greater than 0.80 (Fig. 3). The Pacific Northwest (Columbia River Basin) including the states of Washington, Oregon, and Idaho reported the highest R-squared values, with some streamflow stations exceeding 0.90. The Great Basin (Nevada and western Utah) and the Upper Colorado River Basin (eastern Utah and western Colorado) also reported significant R-squared values. Several significant regions (eastern Wyoming, western Nebraska, western Missouri, central Illinois, western Ohio, and eastern Kentucky) that contribute to the Mississippi River were also identified. Finally, a spatially large region from northern Florida to southern Vermont was identified in the eastern United States.

Atlantic Ocean Sea Surface Temperatures (SSTs)

Although far less robust when compared to Pacific Ocean SST streamflow regions, several streamflow regions were identified in the continental United States in which Atlantic Ocean SSTs, when used as predictors in the PLSR calibration model, achieved R-squared values greater than 0.80 (Fig. 4). The Pacific Northwest (northern Oregon and western Idaho) again reported the highest R-squared values with some stations exceeding 0.85. For the Mississippi River Basin, spatially small regions were identified in central Nebraska and southeastern Kansas/southwestern Missouri. Finally, spatially small regions were identified in the upper Atlantic states.

The results of the PLSR calibration model identified several regions (Columbia River, Upper Colorado River, and Mississippi River) in which R-squared values were significant. A streamflow station was selected in each region (Fig. 3) and a PLSR cross-validation forecast model was developed.

PLSR Cross-Validation Model of Selected Streamflow Stations

Three streamflow stations (Salmon River, White River, and Middle Loup River—Fig. 3) were selected and a PLSR cross-validation forecast model was developed for each. The results include: the optimum number of factors used based on the PRESS statistic and total variance explained; the cross-validated SK LEPS score; the yearly forecasted (solid line) versus actual (dashed line) water year streamflow volumes; the cumulative exceedance probability difference (gray bars) between the forecasted streamflow and actual streamflow; forecasted streamflow confidence levels (5 and 95% limits); and the loading factors from the PLSR model (gray shading displayed at ±5 equates to a loading factor of ±0.5%) are hereby provided for Pacific and Atlantic Ocean SSTs.

Columbia River Basin (Salmon River at White Bird, Id.)

The Salmon River USGS streamflow station (No. 13317000) is located in Idaho County, Id. and the drainage area is 35,000 km^2.

The Salmon River is a tributary of the Snake River and converges with the Snake River along the Washington–Idaho border. The Snake River is the main tributary of the Columbia River. Pacific Ocean SSTs were used to develop the PLSR cross-validation streamflow forecast model. The results are summarized in Fig. 5 and are discussed in the following.

For the Pacific Ocean SSTs, the PRESS statistic resulted in eight factors being used in the PLSR cross-validated model. The eight factors explained approximately 76% of the total variance in the Pacific Ocean SSTs. The cross-validated SK LEPS score was 9.5%, which was just below the "good skill" minimum value of 10%. The yearly forecasted versus actual water year streamflow volumes, including the cumulative exceedance probability difference between the forecasted streamflow and actual streamflow, is provided in Fig. 5(a). A review of Fig. 5(a) shows good forecast years (e.g., small gray bars in 1953, 1954, 1966, 1979, and 2001) and bad forecast years (e.g., large gray bars in 1957, 1965, and 1977). This is also displayed in Fig. 5(b) whereas for the good forecast years, the actual yearly streamflow falls within the forecasted streamflow confidence levels in most years. The SST regions identified as predictors were spatially located in several regions including the northern, western, and equatorial areas of the Pacific Ocean [Fig. 5(c)]. The northern Pacific Ocean SST region may be associated with the Pacific/North American (PNA) teleconnection pattern (http://www.cpc.noaa.gov/data/teledoc/pna.shtml) and the Pacific Decadal Oscillation (PDO) (Mantua et al. 1997), whereas the equatorial Pacific Ocean SST region may be associated with ENSO. Similar SST patterns were identified by Harshburger et al. (2002). Winter precipitation and spring streamflow data from Idaho revealed that winter precipitation in the northern Idaho mountains is negatively correlated with fall SSTs in the eastern tropical Pacific Ocean (i.e., El Niño/La Niña SST region) (Harshburger et al. 2002). Spring discharge was also negatively correlated with SSTs in the eastern tropical and northern regions of the Pacific Ocean (Harshburger et al. 2002). Tootle and Piechota (2006) found the Pacific Northwest to be positively related to central Pacific SSTs. Barlow et al. (2001) applied PCA to Pacific Ocean SSTs and then regressed the PCA factors with continental U.S. streamflow, precipitation, and drought. Barlow et al. (2001) associated the PDO and ENSO with streamflow variability in this region.

Upper Colorado River Basin (White River near Meeker, Colo.)

The White River USGS streamflow station (No. 09304500) is located in Rio Blanco County, Colo. and the drainage area is 2,000 km^2. The White River is a tributary of the Green River, which discharges into the Colorado River in eastern Utah. Pacific Ocean SSTs were used to develop the PLSR cross-validation streamflow forecast model. The results are summarized in Fig. 6 and are discussed in the following.

For the Pacific Ocean SSTs, the PRESS statistic resulted in nine factors being used in the PLSR cross-validated model. The nine factors explained approximately 78% of the total variance in the Pacific Ocean SSTs. The cross-validated SK LEPS score was 26.5%, which indicates good model skill. The yearly forecasted versus actual water year streamflow volumes, including the cumulative exceedance probability difference between the forecasted streamflow and actual streamflow, is provided in Fig. 6(a). A review of Fig. 6(a) shows several good forecast years (1953, 1961, 1962, 1976, 1987, 1990, and 2001) and bad forecast years (1965 and 1979). This is also displayed in Fig. 6(b) whereas for the good forecast years, the actual yearly streamflow falls within the forecasted streamflow confidence levels in most years. The SST regions identified included two regions near the U.S. Pacific coast and two "strong" [e.g., loading value of ±1% which equates to a ±10 in Fig. 6(c)] regions in the northern and southern Pacific Ocean [Fig. 6(c)]. Similar to the results for the Salmon River, the northern Pacific Ocean SST region may be associated with the PNA or PDO whereas the southern Pacific Ocean SST region may be associated with ENSO. Hidalgo and Dracup (2003) identified PDO and ENSO signals in the Upper Colorado River Basin. Distinct shifts in the mean values of precipitation and streamflow in the Upper Colorado River Basin coincide with shifts in the PDO whereas ENSO displays some predictability with precipitation in the Upper Colorado River Basin (Hidalgo and Dracup 2003).

Mississippi River Basin (Middle Loup River at Dunning, Neb.)

The Middle Loup River USGS streamflow station (No. 06775500) is located in Blaine County, Neb. and the drainage area is 4,700 km^2. The Middle Loup River is a tributary of the Platte River, which flows into the Missouri River, which converges with the Mississippi River. Atlantic Ocean SSTs were used to develop the PLSR cross-validation streamflow forecast model. The results are summarized in Fig. 7 and are discussed in the following.

For the Atlantic Ocean SSTs, the PRESS statistic resulted in nine factors being used in the PLSR cross-validated model. The nine factors explained approximately 85% of the total variance in the Atlantic Ocean SSTs. The cross-validated SK LEPS score was 51.0%, which indicates excellent model skill. The yearly forecasted versus actual water year streamflow volumes, including the cumulative exceedance probability difference between the forecasted streamflow and actual streamflow, is provided in Fig. 7(a). A review of Fig. 7(a) shows numerous good forecast years and few bad forecast years. Fig. 7(b) shows actual yearly streamflow falls within the forecasted streamflow confidence levels for numerous years. The four strong (e.g., loading value of + or −1% or + or −10 in the figure) SST regions were identified in the northern Atlantic Ocean and near the African coast [Fig. 7(c)]. Marshall et al. (2001) associated the SST tripole pattern with air–sea fluxes associated with the North Atlantic Oscillation (NAO) (Hurrell and Van Loon 1995) and the Atlantic Ocean SST regions identified in Fig. 7(c) were spatially similar to the tripole pattern. Thus, the NAO may be associated with streamflow variability at this station.

Conclusions

A significant contribution of the current research was an improved method (PLSR) for using spatial SSTs to perform long lead-time forecasting of streamflow. PLSR, by utilizing component scores of both SSTs and streamflow, resulted in excellent forecast skill (i.e., SK LEPS score greater than 10%) for the Upper Colorado River Basin (White River) and the Mississippi River Basin (Middle Loup River) for both Pacific and Atlantic Ocean SSTs. Utilizing Pacific Ocean SSTs, the PLSR calibration model identified significant (i.e., R-squared values greater than 0.80) regions in the Pacific Northwest, Upper Colorado River Basin, Midwest, and Atlantic states, whereas Atlantic Ocean SSTs resulted in significant regions being identified in the Pacific Northwest, Midwest, and Atlantic states.

Cross-validated forecast skill was not as robust for the Columbia River Basin (Salmon River) with Pacific Ocean SSTs achiev-

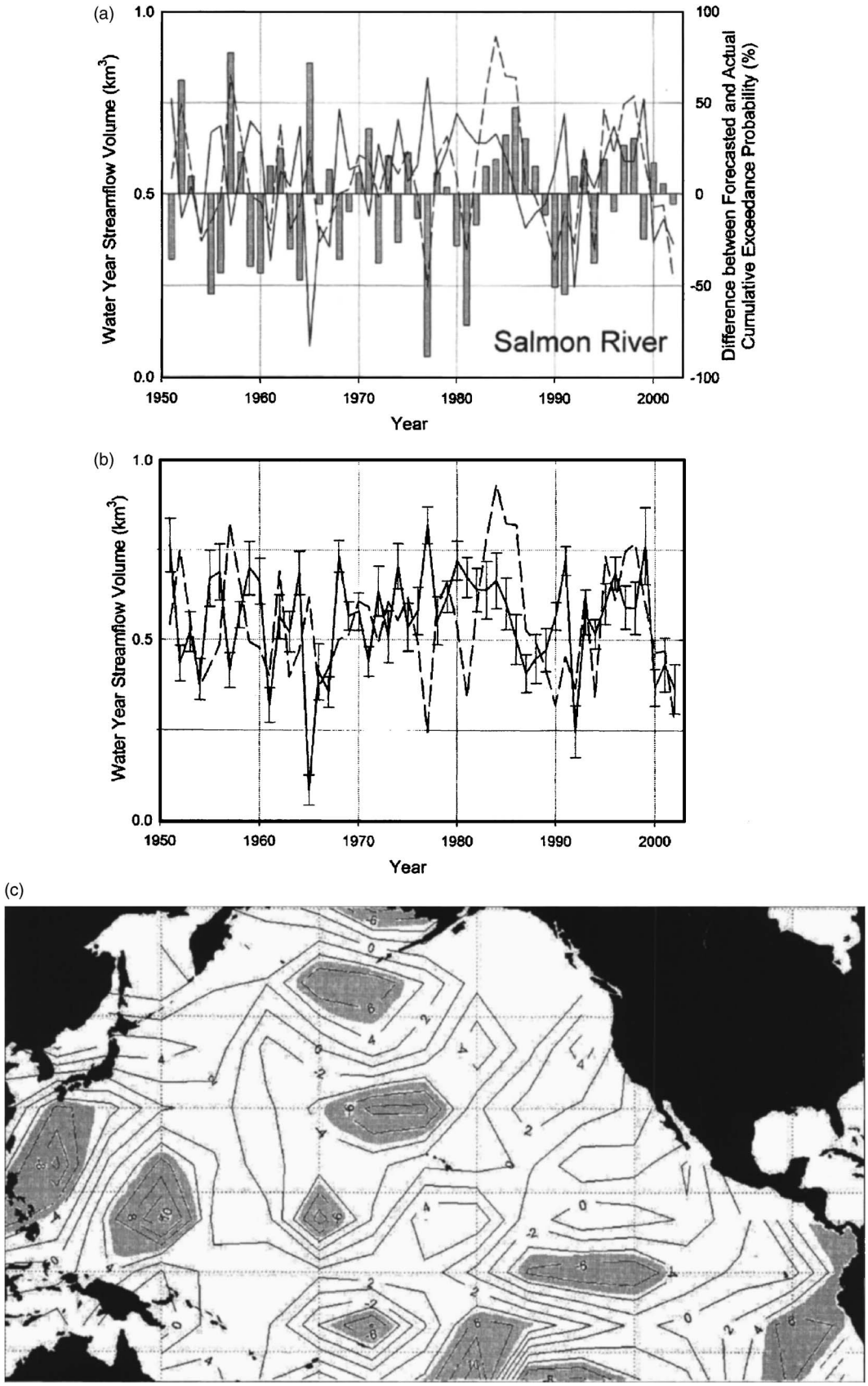

Fig. 5. Salmon River: (a) yearly forecasted (solid line) versus actual (dashed line) water year streamflow volumes with the cumulative exceedance probability difference (gray bars) between the forecasted streamflow and actual streamflow; (b) yearly forecasted (solid line) versus actual (dashed line) water year streamflow volumes with confidence levels (5 and 95% limits); and (c) loading factors from PLSR model for Pacific Ocean SSTs (gray shading displayed at ±5 equates to a loading factor of ±0.5%)

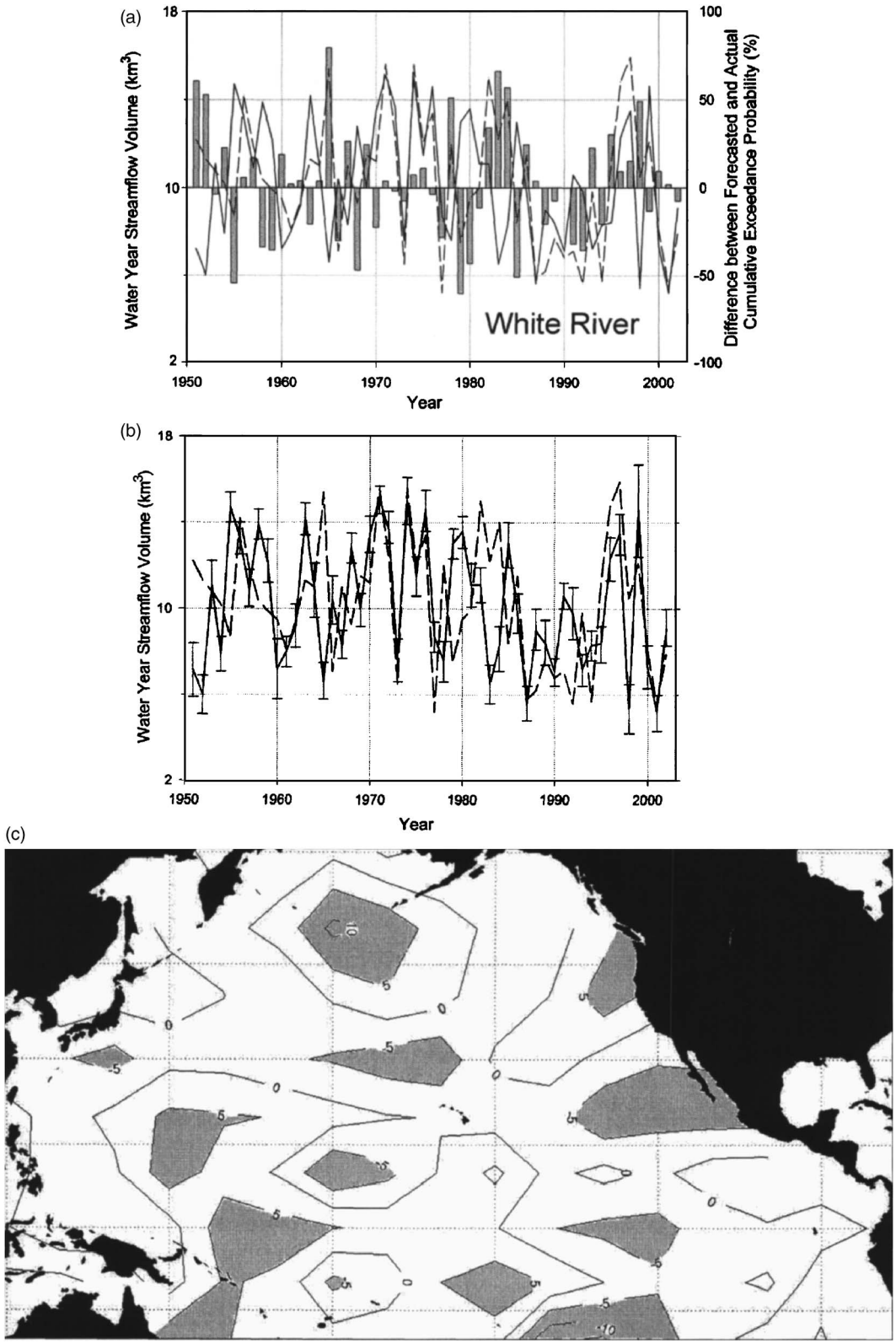

Fig. 6. White River: (a) yearly forecasted (solid line) versus actual (dashed line) water year streamflow volumes with the cumulative exceedance probability difference (gray bars) between the forecasted streamflow and actual streamflow; (b) yearly forecasted (solid line) versus actual (dashed line) water year streamflow volumes with confidence levels (5 and 95% limits); and (c) loading factors from PLSR model for Pacific Ocean SSTs (gray shading displayed at ±5 equates to a loading factor of ±0.5%)

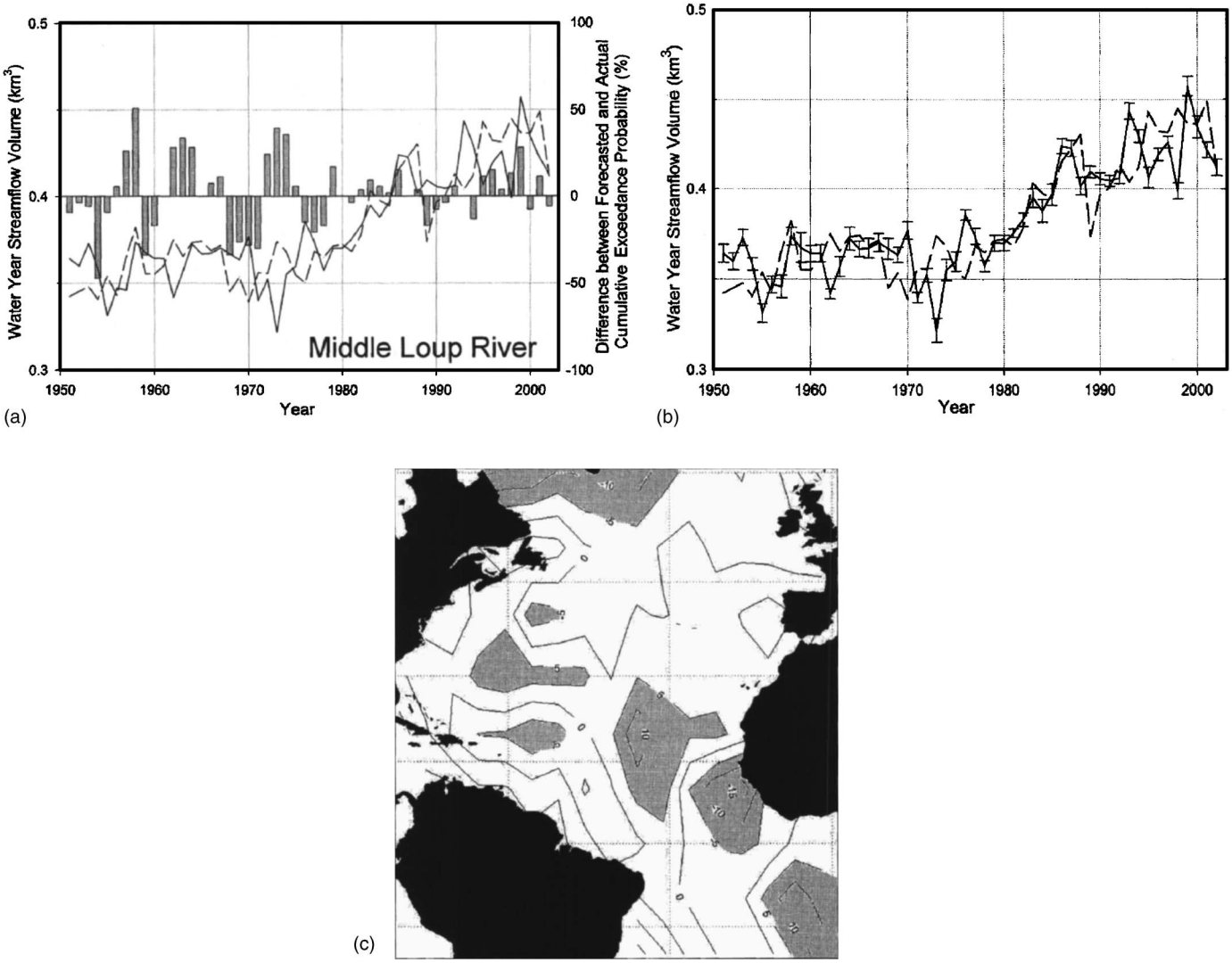

Fig. 7. Middle Loup River: (a) yearly forecasted (solid line) versus actual (dashed line) water year streamflow volumes with the cumulative exceedance probability difference (gray bars) between the forecasted streamflow and actual streamflow; (b) yearly forecasted (solid line) versus actual (dashed line) water year streamflow volumes with confidence levels (5 and 95% limits); and (c) loading factors from PLSR model for Atlantic Ocean SSTs (gray shading displayed at ±5 equates to a loading factor of ±0.5%)

ing marginal forecast skill. Interestingly, the Salmon River is in a region of known ENSO influence. The lower SK LEPS score may be explained by the nonlinearity of ENSO with streamflow response.

Varying both the predictor (i.e., SSTs) and the predictand (i.e., streamflow) seasons and lead times may improve results and should be considered. Additionally, the streamflow stations selected for the PLSR cross validation were selected at random, based on their location in regions (Columbia River, Colorado River, and Mississippi River Basins) of interest to water managers. Additional streamflow stations in these regions should be forecasted and regional skill using the PLSR cross-validation model should also be considered. Although new to hydrology, the PLSR technique provided strong forecast skill, which may result in utilization in other hydrologic applications.

Acknowledgments

This research is supported by the U.S. Geological Survey State Water Resources Research Programs of Nevada and Wyoming, National Science Foundation Award No. CMS-0239334, the National Science Foundation State of Nevada EPSCOR Fellowship, the Wyoming NASA Space Grant Consortium, and the Wyoming Water Development Commission. The writers acknowledge and thank Ms. Sang Hee Lee for her assistance in SAS programming. The writers wish to thank the Editor, Section Editor, Associate Editor, and the two anonymous reviewers for their useful comments.

References

Abdi, H. (2003). "Partial least squares regression." *Encyclopedia of social sciences research methods*, M. Lewis-Beck, A. Bryman, and T. Futing, eds.

Barlow, M., Nigam, S., and Berbery, E. H. (2001). "ENSO, pacific decadal variability, and the U.S. summertime precipitation, drought, and stream flow." *J. Clim.*, 14, 2105–2128.

Eldaw, A. K., Salas, J. D., and Garcia, L. A. (2003). "Long-range forecasting of the Nile River flows using climatic forcing." *J. Appl. Meteorol.*, 42(7), 890–904.

Frank, I. E., and Friedman, J. H. (1993). "A statistical view of some chemometrics regression tools." *Technometrics*, 35(2), 109–135.

Geladi, P., and Kowalski, B. (1986). "Partial least squares regression: A tutorial." *Anal. Chim. Acta*, 185, 1–17.

Gerlach, R. W., Kowalski, B. R., and Wold, H. O. A. (1979). "Partial least-squares path modelling with latent variables." *Anal. Chim. Acta*, 112(4), 417–421.

Harshburger, B., Ye, H., and Dzialoski, J. (2002). "Obervational evidence of the influence of Pacific sea surface temperatures on winter precipitation and spring stream discharge in Idaho." *J. Hydrol.*, 264(1–4), 157–169.

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

Hurrell, J. W., and Van Loon, H. (1995). "Decadal variations in climate associated with the North Atlantic Oscillation." *Clim. Change*, 31, 301–326.

Lall, U. (1995). "Recent advances in nonparametric function estimation." *Rev. Geophys.*, 33, 1093–1102.

Lall, U. B., and Sharma, A. (1996). "A nearest neighbor bootstrap for resampling hydrologic time series." *Water Resour. Res.*, 32(3), 679–693.

Malinowski, E. (2002). *Factor analysis in chemistry*, 3rd Ed., Wiley, New York.

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

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrell, J., McCartney, M., Saravanan, R., and Visbeck, M. (2001). "North Atlantic climate variability: Phenonmena, impacts, and mechanics." *J. Clim.*, 21, 1863–1898.

Nash, J. E., and Sutcliffe, J. V. (1970). "River forecasting using conceptual models. Part I: A discussion of principles." *J. Hydrol.*, 10, 282–290.

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

Piechota, T. C., Chiew, F. H. S., Dracup, J. A., and McMahon, T. A. (2001). "Development of exceedance probability streamflow forecast." *J. Hydrol. Eng.*, 6(1), 20–28.

Piechota, T. C., and Dracup, J. A. (1999). "Long range streamflow forecasting using El Niño–Southern Oscillation indicators." *J. Hydrol. Eng.*, 4(2), 144–151.

Potts, J. M., Folland, C. K., Jolliffe, I. T., and Sexton, D. (1996). "Revised 'LEPS' scores for assessing climate model simulations and long-range forecasts." *J. Clim.*, 9, 34–53.

Rajagopalan, B., and Lall, U. (1999). "A nearest neighbor bootstrap resampling scheme for resampling daily precipitation and other weather variables." *Water Resour. Res.*, 35(10), 3089–3101.

Smith, T. M., and Reynolds, R. W. (2002). "Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997)." *J. Clim.*, 16, 1495–1510.

Sorensen, L. K., and Dalsgaard, S. (2005). "Determination of clay and other soil properties by near infrared spectroscopy." *Soil Sci. Soc. Am. J.*, 69(1), 159–167.

Souza, F. A., and Lall, U. B. (2003). "Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: Applications of a multivariate, semi-parametric algorithm." *Water Resour. Res.*, 39, 1307–1320.

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

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

Tootle, G. A., Piechota, T. C., and Singh, A. K. (2005). "Coupled interdecadal and interannual oceanic/atmospheric variability and United States streamflow." *Water Resour. Res.*, 41(12), W12408.

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

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

Wilks, D. (1995). *Statistical methods in the atmospheric sciences*, Elsevier, New York.

Wold, H. (1966). "Estimation of principal components and related models by iterative least squares." *Multivariate analysis*, P. R. Krishnaiah, ed., Academic, New York, 391–420.

Wold, S. (1978). "Cross-validatory estimation of the number of components in factor and principal components models." *Technometrics*, 20, 397–405.

Wold, S., Geladi, P., Esbensen, K., and Ohman, J. (1987). "Muliway principal ponents and PLS analysis." *J. Chemom.*, 1, 41–56.

Wu, Y., Chen, J., Wu, X., Tian, Q., Ji, J., and Qin, Z. (2005). "Possibilities of reflectance spectroscopy for the assessment of contaminant elements in suburban soils." *Appl. Geochem.*, 20(6), 1051–1059.