CHAPTER 7

Ensemble Streamflow Forecasting: Methods & Applications

Balaji Rajagopalan[+1,2], Katrina Grantz[1,3],
Satish Regonda[+1,2], Martyn Clark[2] and Edith Zagona[3]

[1]Dept of Civil, Environmental & Architectural Engineering (CEAE), University of Colorado, Boulder, CO, USA
[2]CIRES, University of Colorado, Boulder, CO, USA
[3]Center for Advanced Decision Support for Water and Environmental Systems (CADSWES)/CEAE, University of Colorado, Boulder, CO

Key words: Streamflow, Climate Variability, Climate Diagnostics, Ensemble Forecast, Local Polynomials, Bootstrap

7.1. Introduction

The chapter is organized as follows. The theme of the chapter is introduced in Section 7.1 . Section 7.2 presents a background on large-scale climate and its impacts on the western US hydroclimatology. The basins studied and data used are described in sections 7.3 and 7.4 , respectively. This is followed by the climate diagnostics and identification of predictors for forecasting spring streamflows in section 7.5. Section 7.6 presents the development of the statistical ensemble forecating model using the identified predictors. Model application and validation are

[+] *Corresponding author address*: Balaji Rajagopalan, Dept. of Civil, Env. and Arch. Engineering, University of Colorado, ECOT-541, Campux Box 428, Boulder, CO, 80309-0428
E-mail: balajir@colorado.edu

described in section 7.7 . The last section (7.8) concludes the presntation with a summary and discussion of the results.

Water resources worldwide are faced with increasing stresses due to climate variability, population growth and competing growth – more so in the Western US (e.g., Hamlet et al., 2002; Piechota et al., 2001). Careful planning is necessary to meet demands on water quality, volume, timing, and flow rates. This is particularly true in the western US, where it is estimated that 44% of renewable water supplies are consumed annually, as compared with 4% in the rest of the country (el-Ashry and Gibbons, 1988). Consequently, the forecast for the upcoming water year is crucial to the water management planning process involving system outputs such as crop production and the monetary value of hydropower production (e.g., Hamlet et al., 2002), as well as the sustenance of aquatic species.

Majority of river basins in the western USA are snowmelt driven in that, snow accumulates in the winter and melts in the spring thus producing a peak in the streamflow. Therefore, it is intuitive to use winter snowpack as a predictor of the runoff in the following spring (Serreze et al., 1999). More recently, information about large-scale climate phenomena such as El Niño Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) pattern has been added to the forecaster's toolbox. The link between these large-scale phenomena and the hydroclimatology of the western US has been well documented in the literature (e.g. Gershunov, 1998). Clark et al. (2001) showed that including large-scale climate information together with Snow Water Equivalent (SWE) improves the overall skill of the streamflow predictions in the western United States. Souza and Lall (2003) showed significant

skills at longer lead times in forecasting streamflows in Cearra, Brazil using climate information from the Atlantic and Pacific oceans.

Typically, streamflow forecasts are issued by fitting a linear regression with SWE and sometimes with standard indices that describe the ENSO and PDO phenomena. The disadvantages with this approach are (i) the relationship is not always linear, (ii) the teleconnection patterns from ENSO and PDO though dominant on a large scale, often fail to provide forecast skill on the individual basin scale. This is so because the surface climate is sensitive to minor shifts in large-scale atmospheric patterns (e.g., Yarnal and Diaz, 1986), and (iii) inability to provide realistic ensemble forecasts and thus, the probability of exceedences of various thresholds useful for water resources management.

Evidently there is a need for a generalized framework for ensemble streamflow forecast that utilizes large-scale climate information. We propose such a framework in Fig. 7.1. In this, large-scale climate predictors are first identified via climate diagnostics. The identified predictors are then used in a nonparametric framework to generate ensemble of streamflow forecast. The ensembles can then be incorporated in a decision support system for water resources management. In this chapter we focus primarily on the climate diagnostics and ensemble forecast methods, and then demonstrate their utility on the Truckee/Carson River basin and Gunnison River basin, both located in the western USA.

7. 2. Large Scale Climate and Western US Hydroclimatology

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

Decadal-scale fluctuations in SSTs and sea levels in the northern Pacific Ocean as manifested by the PDO (Mantua et al., 1997) provide a separate source of variability for the western US hydroclimate. Independence of PDO from ENSO is still in debate (Newman et al., 2003). Regardless, the influence of PDO and ENSO on North American hydroclimate variability has been well documented (e.g., Regonda et al., 2004a).

Incorporation of this climate information has been shown to improve forecasts of winter snowpack (McCabe and Dettinger, 2002) and streamflows in the western US (Clark et al., 2001, Hamlet et al., 2002) while increasing the lead-time of the forecasts. Use of climate information enables efficient management of water

resources and provides socio-economic benefits (e.g., Pulwarty and Melis, 2001; Hamlet et al., 2002).

Often, however, the standard indices of these phenomena (e.g., NINO3, SOI, PDO index, etc.) are not good predictors of hydroclimate in every basin in the western US- even though these phenomena do impact the western US hydroclimate (as described earlier). Furthermore, certain regions in the western US (e.g., basins in between the Pacific Northwest and the desert Southwest) can be impacted by both the northern and southern branches of the subtropical jet, potentially diminishing apparent connections to ENSO and PDO. The Truckee and Carson basins are two such examples, hence, predictors other than the standard indices have to be developed for each basin.

7.3. Water Management Issues in the Basins Studied

Our motivation for the development of the ensemble streamflow approaches stems from the need to develop tools for efficient water management on two basins (i)Truckee/Carson River basins in Nevada (shown in Fig. 7.2), western USA and (ii) Gunnison River basin, a tributary of Colorado River, also in the western USA that can be seen in Fig. 7.3. On the Truckee/Carson basin flows at two gaging stations are to be forecast, while in the Gunnison streamflow forecasts are required at six sites simultaneously. In both the basins, for that matter over much of the western USA, the bulk of the annual streamflow arrives during spring (April – July) from the melting of snowpack accumulated over winter. This is evident in the climatology of precipitation

and streamflows for the Truckee River (Fig. 7.4) – similar feature is observed on the Gunnison as well.

7.3.1 Truckee/Carson

The Truckee and Carson Rivers originate high in the California Sierra Nevada Mountains and flow northeastward down through the semiarid desert of western Nevada. The Truckee River originates as outflow from Lake Tahoe in California and terminates approximately 115 miles (185 km.) later in Pyramid Lake in Nevada. The Carson River has its headwaters approximately fifty miles (80 km) south of Lake Tahoe, runs almost parallel to the length of the Truckee River and terminates in the Carson Sink area. The areas of the basins are comparable and are approximately 3000 sq. miles (7770 km^2). The Bureau of Reclamation (BOR) Lahontan Basin area office manages operations on the Truckee and Carson Rivers and relies heavily on seasonal (i.e. spring) streamflow forecasts for planning and management. One of the key management issues is the interbasin transfer of water from the Truckee Basin to Lahontan Reservoir in the Carson Basin through the one-way Truckee Canal (Horton, 1995). This transfer augments storage in Lahontan Reservoir for later use by the Newlands Project irrigation district and other water users. If managers divert too much water into the Truckee Canal, they leave insufficient flows in the Truckee River to support other water users, including endangered fish populations, along the last reach of the river. Yet, if managers divert too little water, farmers in the Newlands Project district will have insufficient water in storage to sustain their crops throughout the season. The multiple users with competing objectives coupled with limited canal

capacity and the short water season require that managers use seasonal forecasts for planning and management. Recently implemented policies limit diversions through the Truckee Canal and require specific reservoir releases to aid in the protection of the endangered fish populations – adding further constraints to the reservoir operations and management. The accuracy of forecasts has become evermore important to the efficient management of the water-stressed Truckee and Carson River Basins.

The BOR currently implements forecasts of the spring runoff (April to July volume) into seasonal planning and basin management. These forecasts are issued on the first of each month starting from January. The January forecast affects flood control operations and is used to estimate the irrigation demand for the coming season and, thus, affects reservoir releases and diversions into the Truckee Canal. Updated forecasts in the ensuing months up to April 1st and throughout the runoff season continue to guide operations throughout the basin. Current forecasting techniques use multiple linear regression analysis based on factors related to the existing snowpack and, hence, long-lead forecast skills are limited. Additionally, the current technique does not provide forecasts prior to January as the snowpack information is only partial. Thus, improvements to the spring forecasts, both in skill and in lead-time, are needed to strengthen planning and operations in the Truckee and Carson basins.

7.3.2 Gunnison

The Gunnison River Basin (Fig. 7.3) resides largely in the South Western part of the state of Colorado and, is a major tributary of the Colorado River. It consists of

six sub-basins, i.e., East-Taylor (760 sq.mi; 1968 km^2), Upper Gunnison (2380 sq.mi; 6164 km^2), Tomichi (1090 sq.mi; 2823 km^2), North Fork (959 sq.mi; 2484 km^2), Lower Gunnison (1630 sq.mi; 4222 km^2), and Uncompahange (1110 sq.mi; 2875 km^2). The basin has a drainage area of approximately 20,534 km^2 and basin elevations are extremely variable, ranging from 1387 to 4359 m (McCabe, 1994). It contributes approximately 42% of the streamflow of the Colorado River at the Colorado-Utah Stateline (Ugland et al, 1990). Like Truckee/Carson, almost all of the annual flow in the basin occurs during spring (April-July) due to snowmelt from the higher elevations. The streamfows on the Gunnison impact municipal water supply, power generation and flow release for endangered species. Therefore, like on the Truckee/Carson skilful forecast of spring seasonal streamflows in the basin are key to improvement water management.

7. 4. Data

The following data sets for the period 1949 – 2003 are used in the analysis:

(i) Monthly natural streamflow data for Farad and Ft. Churchill gaging stations on the Truckee and Carson Rivers, respectively, obtained from USBR. Natural streamflows are computed based on inflows to the seven major storage reservoirs near the top of the basin before any significant depletion have been made (*pers. comm.*, Jeff Rieker, 2003). Spring seasonal (April – July) volume is computed from the monthly streamflows that are used in this study.

(ii) Gunnison basin streamflows, at six locations (Fig. 7.3) are selected from the Hydro Climate Data Network (HCDN). This network, HCDN, was developed by

USGS (http://water.usgs.gov) to analyze the climate impacts on the rivers and it has more than 1000 streamflow stations across the conterminous USA that is not affected by human activities (Slack and Landwehr, 1992).

(iii) Monthly SWE (Snow Water Equivalent) data obtained from the NRCS National Water and Climate Center website (http://www.wcc.nrcs.usda.gov). The SWE data is gathered from snow course and snotel stations in the upper Truckee Basin (17 stations) and upper Carson Basin (7 stations). For Gunnison too we had thirteen SWE stations. Basin averages of SWE are calculated using the method employed by the NRCS: the SWE depth from every station in the basin is summed and then divided by the sum of the long-term averages for each of the stations (*pers. comm.*,Tom Pagano, 2003).

(iii) Monthly winter precipitation data for the California Sierra Nevada Mountains region. This is obtained from the U.S. climate division data set from the NOAA-CIRES Climate Diagnostics Center (CDC) website (http:// www.cdc.noaa.gov).

(iv) Monthly values of large-scale ocean atmospheric variables – Sea Surface Temperatures (SST), Geopotential heights (Z500, Z700), Sea Level Pressure (SLP), wind, etc., from NCEP/NCAR Re-analysis project (Kalnay et al., 1996) also obtained from the CDC website.

7. 5. Climate Diagnostics and Predictor Selection

The first step in the forecasting framework is to identify large-scale climate predictors of spring streamflows in the basin. To this end, we first examined the relationship between SWE and spring runoff in the basins. Next, we correlated spring

streamflows with global climate variables from preceeding Fall and Winter seasons. We chose to examine variables from Fall and Winter because the state of the atmosphere during this time affects the position of the jet stream, and consequently, snow deposition and the resulting spring runoff. Also, predictors from Fall and Winter allow for potential long lead forecasts.

7.5.1 Truckee/Carson Basin

As expected, there is a high degree of correlation between winter SWE and spring runoff, particularly with April 1st SWE as it provides a more complete representation of the end of winter snowpack in the basins. Correlation values for Truckee spring streamflows are 0.80 and 0.9 with March 1st SWE and April 1st SWE, respectively, and 0.81 and 0.9, respectively, with the Carson flows. High correlations of streamflows with March 1st SWE offers the opportunity for at least a one month-lead forecast. January 1st SWE, however, does not correlate as well with spring streamflows (0.53 for the Truckee and 0.49 for the Carson) and, hence, provides less skill as a predictor of spring runoff. The snow information by January 1st is only partial and hence, the weak correlation with spring flows.

Spring streamflows in the Truckee and Carson basins are likely modulated by ENSO and PDO, but their standard indices of these phenomena did not show significant correlations with spring streamflows (0.22 for the NINO3, -0.13 for the PDO, and –0.21 for the SOI, for the Truckee; results are similar for the Carson). Thus, we correlated the spring streamflows with the standard ocean-atmospheric circulation fields (e.g., 500mb geopotential height fields, SSTs, SLPs, etc.) to

investigate the large-scale climate link and potential predictors.

Correlation map of Carson River spring streamflows and the preceding winter SSTs and 500mb geopotential heights, henceforth, referred to as Z500, are shown in Fig. 7.5. Strong negative correlations (approximately -0.7) with Z500 in the region off the coast of Washington can be seen. The SSTs in the northern mid-Pacific Ocean exhibited a strong positive (about 0.5) correlation and to the east of this, a negative correlation. Similar, but slightly weaker correlation patterns were seen with preceding Fall (Sep – Nov) Z500 and SSTs (Grantz, 2003), suggesting that the physical mechanisms responsible for the correlations are persistent from Fall through Winter. These correlations offer hopes for a long-lead forecast of spring streamflows – at the least, they can provide significant information about the upcoming spring streamflows in Fall, before SWE data is available.

To understand the physical mechanisms driving the correlation patterns seen above, a composite analysis was performed. In this, average SST, wind and Z500 patterns for high and low streamflow years were obtained to identify coherent regions with strong magnitudes of the variables. We chose years with streamflows exceeding the 90^{th} percentile as "high" years and those below the 10^{th} percentile as "low" years. Fig. 7.6 shows the composites of vector wind, Z500 and SST anomalies during the winter season preceding the high and low streamflow years. The winds in high streamflow years show a counterclockwise rotation around the low pressure region off the coast of Washington - the region of highest correlation seen in Fig. 7.5. This counterclockwise rotation brings southerly winds over the Trukee and Carson Basins. Southerly winds tend to be warm and moist, thus increasing the chances of enhanced

winter snow and, consequently, higher streamflows in the following spring. The opposite pattern is seen during low streamflow years when anomalous northeasterlies tend to bring cold dry air and, consequently, less snow and decreased streamflows. The Z500 patterns and the vector wind anomalies in high and low streamflow years are consistent with each other. The SST patterns in high and low streamflow years (Fig. 7.5) are a direct response to the pressure and winds. The winds are generally stronger to the east of a low pressure region-- this increases the evaporative cooling and also increases upwelling of deep cold water to the surface. Together, they result in cooler than normal SSTs to the east of the low pressure region. The opposite is true on the west side of the low pressure region. Composite maps for the Fall season show similar patterns – indicating that the physical mechanisms are persistent. Results for the Truckee River streamflows are very similar (Grantz, 2003).

Thus, based on the correlation and composite analyses we developed predictors of the Truckee and Carson basins by averaging the ocean-atmospheric variables over the areas of highest correlation (e.g. as in Fig. 7.4). These areas were determined by visual inspection of the correlation maps. Specifically, the Z500 index was obtained as the average over the region 225°-235° E and 42°-46° N and the SST predictor index as the average over 175°-185° E and 42°-47° N. Time series of these indices were obtained to be used as predictors in the forecast model.

7.5.2 Gunnison

In the Gunnison river basin we have streamflows from six locations (Fig. 7. 3) that are highly correlated with each other (correlation coefficients of 0.75 and higher).

One option is to create a basin streamflow series by averaging the flows across the six locations. The other is to perform a Principal Component Analysis (PCA) on the streamflow data and retain the first principal component (PC), which is in essence the average of the six streamflows. In fact, the first principal component (PC1) is correlated 0.99 with the basin average streamflows. The PCA is briefly described in the context of ensemble forecast in the following section.

The first PC of spring flows is correlated with large-scale climate variables from preceding seasons to identify predictors. Fig. 7.7 shows the correlation map of PC1 with winter Geopotential height at 700mb (Z700) and SST. A strong negatively correlated region of Z700 can be seen over the South Western US also, highly correlated regions of SSTs observed in the Central and Northern Pacific. We also correlated PC1 with other circulation variables such as, zonal and meridional winds (figures not shown). These correlation patterns were persistent in the preceding Fall season as well. To understand the physical mechanisms, composite maps of vector winds in the wet and dry years are shown in Fig. 7.8. The wet and dry years are defined as the years with PC values greater than the 90^{th} and below the 10^{th} percentile, respectively. Notice that a strong counter clock wise flow around the region with strong negative correlation with Z700 (Fig. 7.7a) for the wet years and vice-versa in the dry years. This implies, advection of warm moist air from the south to the basin, thus tending to produce more snow and consequently, higher streamflows in the following spring. These are very similar to the findings from the Truckee/Carson streamflows (Figs. 7. 5 and 7.6).

Based on the correlation maps, we selected four potential predictors of PC1 to be the time series of average values of difference between the positive and negative correlation regions in (i) Z700 (Fig. 7.7a) (ii) SST (Fig. 7.7b), (iii) Zonal winds, and (iv) Meridional winds. Often, the difference between the regions of strong positive and negative correlation constitutes a much better predictor. For example, in the case of the Z700 predictor we took the difference between the average value over the region with negative correlation (32.5°N-42.5°N latitude and 230°-250°E longitude) and positive correlation (42.5°N – 55°N latitude: 275°-297.5°E longitude) as can be seen from Fig. 7.7a.

7.6. Forecast Models

Statistical forecast models can be represented as:

$$Y_t = f(\boldsymbol{x_t}) + e_t \qquad [7.1]$$

where, $\boldsymbol{x_t} = (x_{1t}, x_{2t}, x_{3t}, \ldots x_{pt)}$, t = 1,2,…$N$, f is a function fitted to the predictor variables ($x_1, x_2, \ldots, x_p$), Y is the dependent variable (e.g., Spring seasonal streamflow), and e_t is the errors assumed to be Normally (or Gaussian) distributed with a mean of 0 and variance s. Traditional parametric methods involve fitting a linear function, also known as linear regression. The theory behind the parametric methods, procedures for parameter estimation and hypothesis testing are well developed (e.g., Helsel and Hirsch, 1995; Rao and Toutenburg, 1999). The main drawbacks, however, are: (i) the assumption of a Gaussian distribution of data and errors, (ii) the assumption of a linear relationship between the predictors and the dependent variable, (iii) higher

order fits (e.g., quadratic or cubic) require large amounts of data for fitting and, (iv) the models are not portable across data sets, i.e., sites.

Nonparametric methods, in contrast, estimate the function f "locally". There are several nonparametric approaches, such as kernel-based (Bowman and Azzalini, 1997), splines, K-nearest neighbor (K-NN) local polynomials (Rajagopalan and Lall, 1999; Owosina, 1992); locally weighted polynomials (Loader, 1999), etc. The K-NN local polynomials and the local weighted polynomial (LOCFIT)[1] approaches are very similar. Owosina (1992) performed an extensive comparison of a number of regression methods both parametric and nonparametric on a variety of synthetic data sets and found that the nonparametric methods out-perform parametric alternatives. Below we describe a few nonparametric methods for ensemble forecast.

7.6.1 Local methods of Ensemble Forecast

The local polynomial methods obtain the value of the function f at any point '$\boldsymbol{x}_t^*$' by fitting a polynomial to a small set $K(= a * N, a = (0,1])$ of neighbors to '$\boldsymbol{x}_t^*$'. The neighbors can be identified based on the Euclidean distance (Rajagopalan and Lall, 1999; Lall and Sharma, 1996) or Mahalanobis distance (Yates et al., 2003). Other approaches include weighting the predictors differently in the distance calculation, such as weighted obtained via coefficients from a linear regression between the dependent variable and predictors (Souza and Lall, 2003). Once the neighbors are identified, there are two main options for generating ensembles:

[1]LOCFIT is the package to perform local polynomial fits developed by Loader (1999) and available freely at http://cm.bell-labs.com/cm/ms/departments/sia/project/locfit/index.html

(i) The neighbors can be re-sampled with a weight function that gives more weight to the 'nearest' neighbors and less to the farthest, thus generating ensemble (e.g.. Souza and Lall, 2003). This has been widely applied for stochastic weather and streamflow simulation in the above mentioned references.

(ii) A polynomial or order *p* can be fit to the neighbors that can be used to estimate the mean of the dependent variable (Rajagopalan and Lall 1998; Loader, 1999) using LOCFIT and the local variance σ_{le}^2 of the errors around the mean. The local error variance can be used to generate random normal deviates which, when added to the mean estimate, yield ensembles. Thus, the parameters to be estimated are the size of the neighborhood (*K*) and the order of the polynomial (*p*), which is obtained using objective criteria such as Generalized Cross Validation (GCV).

$$GCV(K,p) = \frac{\sum_{i=1}^{N} \frac{e_i^2}{N}}{\left(1 - \frac{m}{N}\right)^2} \qquad [7.2]$$

where e_i is the error, N is the number of data points, m is the number of parameters. For stability purposes, the minimum neighborhood size should be twice the number parameters to be estimated in the model. The *K, p* combination to be selected is the one that has the minimum GCV score. This was used in the Thailand summer rainfall forecast (Singhrattna et al., 2004)

(iii) In (ii) above the local errors are assumed to be normally distributed. Often times this may not be the case. To get over this problem, Prairie (2002) and Prairie et al.(*in press*) proposed an interesting variation and applied it for streamflow and salinity modeling on the Colorado Rive Basin. This was later applied to the Thailand summer

rainfall forecasting (Singhrattna et al., *in press*). In this variation, the mean value, Y_i of the predictor vector '$\boldsymbol{x_t^*}$' is obtained from the steps described in (ii) above. Then, one of the neighbors of '$\boldsymbol{x_t^*}$' is selected and the corresponding residual, $\boldsymbol{e_t^*}$ is picked up. This residual is then added to the mean forecast $Y_t^*+e_t^*$, thus obtaining one of the ensemble members, repeating this several times results in an ensemble. This is pictorially shown in Fig. 7. 9 for the Truckee river spring streamflows and the Z500 index as the predictor. The solid line is the mean fit using LOCFIT, the points are the observations and the dashed rectangle is the neighborhood size from which the residuals are resampled. The neighbors are obtained using any of the distance metric described in (ii). Furthermore, the selection of one of the neighbors is done using a weight function of the form:

$$W(j) = \frac{1}{j\sum_{i=1}^{k}\frac{1}{i}} \qquad [7.3]$$

This weight function gives more weight to the nearest neighbor and less to the farthest neighbors. The number of neighbors to be used to resample the residuals can be the same as (*K)* that used in fitting the local polynomial or could be different. In practice, the $\sqrt{N}$ neighbors for resampling the residuals seem to work fine. This heuristic rule is justified by theoretical arguments of Fukunaga (1990).

Being a local estimation scheme, these methods have the ability to capture any arbitrary local features. Furthermore, unlike the parametric alternatives, no prior assumption are need be made regarding the functional form of the relationship (e.g., a linear relationship, Gaussian distribution, etc.). Other variations using nonlinear dynamics based time series analysis (Regonda et al., 2004b) can also be explored.

7.6.2 Multi-site Ensemble forecast

Often times forecasts are required at several sites simultaneously that captures the spatial correlation structure. For example, ensemble forecasts will be required at all the streamflow locations in a basin for use in decision support system. The local methods described above can be used in this case but there needs to be a pre-processing step prior to using them. The steps are as follows:

(i) A principal component analysis (PCA) (von Storch and Zwiers, 1999; and Preisendorfer, 1988) is performed on the seasonal steamflows at all the sites. PCA provides an orthogonal space-time decomposition, with spatial part represented as Eigen Vectors (EV) and the temporal part as Principal Components (PC). The theory and implementation of this is widespread in climate analysis and the above references offer a detailed exposition of this and other related approaches. Typically, the leading PC captures almost all of the variance, especially in homogeneous basins.

(ii) Predictors are identified for the leading PC

(iii) For a given predictor vector (i.e. a given year) ensembles of the leading PC are generated from the local methods described in the previous section.

(iv) The PC ensembles are back transformed to the original flow space by multiplying with the appropriate Eigen Vector. Thus, resulting in ensembles of streamflows at all the locations simultaneously, preserving the spatial correlation structure.

This approach was recently developed and encouraging results from preliminary application to the Gunnison River Basin in the western US are presented later in this chapter.

7.6.3 Subset Selection

As we saw from the previous section several potential predictors are identified for forecasting the streamflows in two basins. The task then is to select the best predictor subset. In the linear regression framework this is done using stepwise regression (e.g, Rao and Toutenburg 1999; Walpole et al., 2002), where in the smallest subset that explains most variance in the dependent variable is selected. Other methods use score functions such as Mallow's Cp, Akaike Information Criteria (AIC) (Rao and Toutenburg, 1998) etc., which favor parsimony.

We propose the use of GCV (equation 7.2) as a tool for subset selection. In this, one fits local polynomial for different predictor combinations along with the polynomial order and the neighborhood size and the GCV value computed in each case. The combination that produces the least GCV value is chosen as the best subset. The GCV function is a good surrogate of predictive error (Craven and Whaba, 1979) of the model, unlike least squares which is a measure of goodness of fit and provides no information on the predictive capability. Hence, we feel that the GCV will be a better alternative for subset selection.

Applying the GCV criteria we selected Z500 index and SWE as the best set of predictors for the Truckee/Carson streamflows and SST predictor and SWE for the Gunnison. These respective subsets will be used in the ensemble forecast.

7.7. Model Validation

The large-scale climate predictors identified for the spring seasonal streamflows are used in the ensemble forecast models. Each year is dropped from the

record and ensemble forecasts are made for the dropped year based on the rest of the data. This is repeated for all the years, thus obtaining cross-validated ensemble forecasts. The forecasts are issued at several lead times.

7.7.1 Skill Measures for Validation

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

(i) Correlation coefficient of the mean of the ensemble forecast and the observed value. This measures the skill in the mean forecast.

(ii) Ranked Probability Skill Score (RPSS) (Wilks, 1995).

(iii) Likelihood Function Skill Score (LLH) (Rajagopalan et al., 2002).

RPSS and LLH measure the forecast's ability to capture the probability distribution function (PDF). The RPSS is typically used by climatologists and meteorologists to evaluate a model's skill in capturing categorical probabilities relative to climatology.

We divided the streamflows into three categories, at the tercile boundaries, i.e., 33^{rd} percentile and 66^{th} percentile. Values above the 66^{th} percentile are in the "above normal" category, below the 33^{rd} percentile are in the "below normal" category, and the remainder fall in the "normal" category. Of course, one can divide into unequal categories as well. The categorical probability forecast is obtained as the proportion of ensemble members falling in each category. The "climatology" forecast is the proportion of historical observations in each category. For the tercile categories presented here the climatological probability of each category is 1/3.

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

$$RPS(p,d) = \frac{1}{k-1}\left[\left(\sum_{i=1}^{k} P_i - \sum_{i=1}^{k} d_i\right)^2\right] \qquad [7.4]$$

The vector d ($d_1, d_2, \ldots d_k$) represents the observations, such that d_k equals one if the observation falls in the k^{th} category and zero otherwise. The RPSS is then calculated as (Toth, 2002):

$$RPSS = 1 - \frac{RPS(\text{forecast})}{RPS(\text{climatology})} \qquad [7.5]$$

The RPSS ranges from positive 1 (perfect forecast) to negative infinity. Negative RPSS values indicate that the forecast has less accuracy than climatology. The RPSS essentially measures how often an ensemble member falls into the category of the observed value and compares that to a climatological forecast. The likelihood function is also used to quantify the skill of ensemble forecasts. This function compares the likelihood of the ensemble forecast falling into the observed category with respect to climtology. The likelihood skill score for the ensemble forecast in any given year is calculated as:

$$L = \frac{\prod_{t=1}^{N} \hat{P}_{j,t}}{\prod_{t=1}^{N} P_{cj,t}} \qquad [7.6]$$

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

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

7.7.2 Results

Truckee/Carson

For the Truckee/Carson basins we had identified in the previous section, two best predictors – Z500 index and SWE. The SWE, however, is only available from Jan 1st onwards, as prior to that there is little snow in the basin and hence, of limited use for forecasting spring streamflows.

Given that SWE is highly correlated spring streamflows, the utility of Z500 index in the model could be doubted. To investigate this, ensemble forecasts are issued for the Truckee and Carson Rivers on the 1st of each month from Nov. through April from two models – one using Z500 and SWE as predictors and another with only the SWE. The skill scores were computed for the forecasts and are shown in Fig. 7.10. The results show that using the Z500 index together with SWE as predictors provides better skills at all lead times. This is a significant outcome in that it clearly demonstrates the importance of incorporating basin specific large-scale climate indices in streamflow forecasts. It is also apparent from Fig. 7.10 that the forecast skills are above climtology at all lead times (the RPSS is above zero and the LLH is above 1), indicating the presence of useful information about the spring streamflows

from as early as Fall. As in most forecasting models, the skills on all the measures improve with decreased lead-time.

To assess the performance of the model in extreme years we calculated the RPSS and LLH for wet and dry years. We define years with streamflows above the 75^{th} percentile as wet and those below the 25^{th} percentile as dry. Roughly 12 years fall into each category. Median skills for forecasts issued on April 1^{st} and December 1^{st} are shown in Table 7. 1. It is apparent that the model has a slightly higher skill in predicting the wet years relative to dry. This asymmetry in the skills is consistent with the nonlinearities seen in the relationship between the predictors and the streamflows (Fig. 7.9). Whereas high streamflow years exhibit a strong linear relationship with the Z500 index, this relationship breaks down, i.e., flattens out in low streamflow years. The skills for forecast issued on December 1^{st} are relatively poor but there are substantial, especially in the extreme years – providing useful long lead forecast for water resources planning.

Ensemble forecasts provide the probability density function (PDF) and consequently, they can be used to obtain threshold exceedence probabilities. This information can be very useful for water managers in preparing for extreme events. Fig. 7.11 presents the PDF of the ensemble forecasts for 1992 and 1999, below normal and above normal streamflow years, respectively. The climatological PDF, i.e., the PDF of the historical data, is overlaid in these plots. Notice that the PDFs of the ensemble forecasts are shifted toward the observed values. In 1992, a dry year, the observed streamflow in the Truckee River was 75 kaf (~ 93 M m^3), much below the historical average. Based on the climatological PDF the exceedence probability of

this value is 0.92, while that from the ensemble forecasts is 0.49, indicative of drier conditions. Similarly, for the above average flow of 408 kaf (~ 504 M m^3) in 1999, climatology suggested an exceedence probability of 0.17 while the ensemble forecasts show a much higher probability of exceedence (0.59), thereby better capturing the probability of the observed flow value.

Gunnison

In this basin as described earlier, we generate ensemble forecast of the first PC and then multiply it with the Eigen vector to provide the ensemble streamflow forecast at all the six locations. Fig. 7.12 shows the skill scores (RPSS) for forecast issued on January 1^{st} and April 1^{st}. It can be seen that the skills increase with lead time and they are quite good overall. To evaluate the forecasts in extreme (wet and dry) and average years, the ensemble forecasts for the East River site is shown in Fig. 7.13. We loosely defined, for this purpose, the wet years as those with streamflows above the upper tercile and dry as those below the lower tercile and the rest as average. The ensembles are shown as boxplots with the boxes being the interquartile range, the whiskers at the 5^{th} and 95^{th} percentile and the points are outside this range, the observed values are shown as thick solid points. If the observed values fall within the box it implies that the ensembles are better able to capture the PDF of the flows in that year. The horizontal lines are the 25^{th}, 50^{th}, 75^{th}, 90^{th} and 95^{th} percentile values of the historical data. It can be seen that the boxplots are shifted in the right direction in the extreme years. Notice that the skills in the extreme years are particularly good that

can be of great importance in water management in the basin. The median skill scores at all the locations for all years, wet and dry years are shown in Table 7.2.

7.8. Summary and Discussion

We presented a framework for ensemble forecast that uses large-scale climate information and demonstrated its utility on providing seasonal streamflow forecasts in two river basins in the western USA. Climate diagnostics is first performed to obtain a suite of potential large-scale climate predictors. Local polynomial based nonparametric methods can then be used to identify the best subset of predictors and use them to generate ensemble forecasts. We developed methods for ensemble forecasts at a single site and also for multi-site preserving the spatial correlation. The proposed nonparametric methods are data driven and provide a flexible and powerful alternative to traditional parametric (i.e. linear regression) methods in capturing any arbitrary relationship between the predictors and the dependent variable and error structure. Application to Truckee/Carson and Gunnison river basins show that significant long-lead skill in forecasting seasonal streamflows can be achieved, especially in extreme years. This has tremendous impact on improving the water resources management and planning in these basins. Our preliminary application of these forecasts on the Truckee/Carson basin to improve the operations of Truckee canal (descried earlier in the chapter) gives encouraging results (Grantz, 2003). The proposed framework can be easily applied to any other basin and other variables – e.g., we applied this to Thailand summer rainfall forecast with good success (Singhrattna, et al., 2004).

Acknowledgments

We thank the Bureau of Reclamation Lahontan Basin area office for funding the Truckee/Carson study. Funding through the CIRES Innovative Research Program at the University of Colorado at Boulder is also thankfully acknowledged. Useful discussions with Tom Scott, Tom Pagano, and Jeff Rieker are very much appreciated.

References

Allan, R., J. Lindesay, and D. Parker (1996) El Nino Southern Oscillation & Climatic Variability, National Library of Cataloguing-in-Publication, Collingwood.

Bowman, A. and A. Azzalini (1997) *Applied smoothing techniques for data analysis*, Oxford, UK.

Cayan, D. and R. Webb (1992) El Nino/Southern Oscillation and Streamflow in the Western Unites States, In: *El Nino*, Henry F. Diaz and Vera Markgraf (Editors), Cambridge University Press, Cambridge, Great Britain, 29-68, 1992.

Clark, M. P. and M. C. Serreze, (2001) Historical effects of El Nino and La Nino events on the seasonal evolution of the montains snowpack in the Columbia and Colorado River Basins, *Water Resources Research*, **37**, 741-757.

Craven, P. and G. Whaba (1979) Optimal smoothing of noisy data with spline functions, *Numerische Mathematik*, **31**, pp. 377 - 403.

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

Fukunaga, K. (1990) *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, Calif..

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

Grantz, K. (2003) *Using Large-Scale Climate Information to Forecast Seasonal Streamflow in the Truckee and Carson Rivers*, M.S. Thesis, Colorado, University of Colorado at Boulder.

Hamlet, A. F., D. Huppert, and D. P. Lettenmaier (2002) Economic Value of Long-Lead Streamflow Forecasts for Columbia River Hydropower, *Journal of Water Resources Planning and Management*, March/April, 91-101.

Helsel, D. R., and R. M. Hirsch (1995) *Statistical Methods in Water Resources*, Elsevier Science Publishers B.V., Amsterdam.

Hidalgo, H. G., and J. A. Dracup (2003) ENSO and PDO Effects on Hydroclimatic Variation of the Upper Colorado River Basin, *Journal of Hydrometeorology*, **4**, 5-23.

Hoerling, M. P., A. Kumar, and M. Zhong (1997) El Nino, La Nina, and the Nonlinearity of their Teleconnections, *Journal of Climate*, **10**, 1769-1786.

Horton, G. A. (1995) *Truckee River Chronology*, Division of Water Planning, Department of Conservation and Natural Resources, Carson City, Nevada

Kalnay, E., et al (1996) The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*, **77**, 437-471.

Lall, U. and A. Sharma (1996) A Nearest Neighbor Bootstrap for Resampling Hydrologic Time Series, *Water Resources Research*, **32** No. 3, pp 679-693.

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

Mantua, N. J., S. R. hare, J. M. Wallace, and R. C. Francis (1997) A Pacific interdecadal climate oscillation with impacts on salmon production, *Bulletin of the American Meteorological Society*, **78**, 1069-1079.

McCabe, G. J., 1994: Relationships between atmospheric circulation and snowpack in the Gunnison River Basin, Colorado. *J. Hydrol.*, **157**, 157-175.

McCabe, G. J. and M. D. Dettinger (2002) Primary Modes and Predictability of Year-to-Year Snowpack Variation in the Western United States from Teleconnections with Pacific Ocean Climate, *Journal of Hydrometeorology*, **3**, 13-25.

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

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

Pagano, T. (2003) Water Supply Forecaster, Natural Resources Conservation Service. Personal communication, October 2003.

Piechota, T. C., H. Hidalgo, and J. Dracup (2001) Streamflow Variability and Reconstruction for the Colorado River Basin. Proceedings of the EWRI World Water & Environmental Resources Congress, May 20-24, 2001, Orlando, Florida, American Society of Civil Engineers, Washington D.C.

Prairie, J. R. (2002) *Long-term Salinity Prediction with Uncertainty Analysis: Application for Colorado River above Glenwood Springs*, M.S. Thesis, Colorado, University of Colorado at Boulder.

Prairie, J.R., B. Rajagopalan, T. Fulp, and E. Zagona (in press) Statistical nonparametric model for natural salt estimation, *ASCE Journal of Environmental Engineering*, 2005.

Preisendorfer R. W. (1988) *Principal component analysis in meteorology and oceanography*, Elsevier, New York, 425p.

Pulwarty, R. S. and T. S. Melis (2001) Climate extremes and adaptive management on the Colorado River: Lessons from the 1997-1998 ENSO event, *Journal of Environmental Management*, **63**, 307-324.

Rajagopalan, B. and U. Lall (1999) A Nearest Neighbor Bootstrap Resampling Scheme for Resampling Daily Precipitation and other Weather Variables, *Water Resources Research,* **35 (10),** 3089-3101.

Rajagopalan, B., U. Lall, and S. Zebiak (2002) Optimal Categorical Climate Forecasts through Multiple GCM Ensemble Combination and Regularization, *Monthly Weather Review*, **130**, 1792-1811.

Rao, C. R., and H. Toutenburg (1999) *Linear models: least squares and alternatives*, Springer, New York

Rasmussen, E.M. (1985) El Nino and variations in climate, *American Scientist*, **73**, 168-177.

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

Regonda, S., B. Rajagopalan, M. Clark, and J. Pitlick (2004 a) Seasonal cycle shifts in hydroclimatology over the Western US, *Journal of Climate (in press)*

Regonda, S., B. Rajagopalan, U. Lall, M. Clark and Y. Moon (2004 b) Local polynomial method for ensemble forecast of time series, *Nonlinear Processes in Geophysics,* Special issue on "Nonlinear Deterministic Dynamics in Hydrologic Systems: Present Activities and Future Challenges" (in press).

Rieker, J. (2003) Water Resources Engineer, US Bureau of Reclamation. Personal communication.

Ropelewski, C. F. and M. S. Halpert (1986) North American precipitation and temperature patterns associated with El Nino-Southern Oscillation (ENSO), *Monthly Weather Review*, **114**, 2352-2362.

Scott, T., Water Resources Engineer, US Bureau of Reclamation. Personal Communication, 2002.

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

Singhrattna, N., B. Rajagopalan, M. Clark and K. Krishna Kumar (2004) Forecasting Thailand Summer Monsoon Rainfall, *International Journal of Climatology (in press)*

Slack, J. R., and J. M. Landwehr (1992) Hydro-Climatic Data Network (HCDN): A U.S. Geological Survey streamflow data set for the United States for the study of climate variations, 1874-1988. *U.S. Geol. Surv. Open File Rep.,* 92-129, 200 pp.

Souza F. A., and U. Lall (2003) Seasonal to Interannual Ensemble Streamflow Forecasts for Ceara, Brazil: Applications of a Multivariate, Semi-Parametric Algorithm, *Water Resources Research,* **39**, 1307-1320.

Toth, Z. (2002) Assessing the Value of Probabilistic Forecasts from a Scientific Perspective, Validation of Probabilistic Forecasts, Predictability Seminar, ECMWF, Sept 9-13 2002.

Ugland, R.C., Cochran, B.J., Hiner, M.M., Kretschman, R.G., Wilson, E.A. and Bennett, J.D. (1990) *Water Resources Data for Colorado*, Water Year 1990, Vol. 2, Colorado River Basin. U.S Geol. Surv. Water-Data Rep. CO-90-2, 344 pp.

Von Storch, H., and F. W. Zwiers (1999) *Statistical Analysis in climate research*, Cambridge University Press, New York, 484p.

Wilks, D. (1995) *Statistical Methods in the Atmospheric Sciences*, Academic Press, San Diego.

Walpole, R.E., R.H. Myers, S.L. Myers, K. Ye, and K. Yee (2002) Probability and Statistics for Engineers and Scientist, Prentice Hall, Upper Saddle River, N.J.

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

Yates, D.S., Gangopadhyay, S., Rajagopalan, B., and Strzepek, K (2003) A technique for generating regional climate scenarios using a nearest neighbor bootstrap, *Water Resources Research*, 39(7), SWC 7(1-15).

	Median Skill Score			
	RPSS		LLH	
	Truckee	Carson	Truckee	Carson
All Years	1.0(0.2)	0.9(0.0)	2.3(1.1)	2.3(1.1)
Wet Years	1.0(0.4)	1.0(0.3)	3.0(1.1)	2.6(1.2)
Dry Years	0.9(0.0)	0.6(0.0)	2.2(1.1)	2.2(1.1)

Table 1: Median Skill Scores for ensemble forecast issued on April 1st, for all years, wet years and dry years for the Truckee/Carson basin. The values in parenthesis are for forecast issued on December 1st.

USGS Station	RPSS (LLH)		
	Total years	Wet years	Dry years
09110000	0.72 (1.91)	1.00 (3.00)	0.92 (2.35)
09112500	0.92 (2.58)	1.00 (3.00)	0.98 (2.68)
09119000	0.73 (1.91)	0.98 (2.72)	0.97 (2.64)
09124500	0.73 (1.84)	0.93 (2.39)	0.91 (2.32)
09132500	0.91 (2.35)	1.00 (3.00)	0.92 (2.35)
09147500	0.87 (2.27)	1.00 (3.00)	0.97 (2.68)

Table 2: Median Skill Scores for ensemble forecast issued on April 1st, for all years, wet years and dry years for the Gunnison basin. The values in parenthesis are the LLH.

Figure Captions:

Figure 7. 1: Flow chart of the forecast framework.

Figure 7.2: Map of the Truckee/Carson basin

Figure 7.3: Gunnison River Basin and stream flow locations. USGS gage locations and river names are mentioned below.

1. 09112500 East River at Almont, CO
2. 09110000 Taylor River at Almont, CO
3. 09119000 Tomichi Creek at Gunnison, CO
4. 09124500 Lake fork at Gate view, CO
5. 09147500 Uncompahgre River at Colona, CO
6. 09132500 North fork Gunnison river near Somerset, CO

Figure 7.4: Climatology of streamflows and precipitation in the Truckee River, at the gauging station Farad (based on data for the 1949 – 2003 period).

Figure 7.5: Correlation of Carson River spring streamflows with winter (Dec – Feb) climate variables (a) 500mb geopotential height (Z500) and (b) SST.

Figure 7.6: Composites of vector winds, SST and Z500 during the winter of high and low streamflow years.

Figure 7.7: Correlation between PC1 of Spring flows and Nov-Mar Climate indices (a) Geopotential Height – 700mb and (b) Sea Surface Temperature.

Figure 7.8: Composite of Vector wind at 700 mb for (a) wet years and (b) dry years.

Figure 7.9: Residual resampling to obtain an ensemble forecast.

Figure 7.10: Skill scores of forecasts issued from the 1st of each month Nov – April for Truckee and Carson Rivers.

Figure 7.11: PDF of the ensemble forecasts in a (a) dry year (1992) and (b) wet year (1999) for the Truckee River.

Figure 7.12: Median RPSS score for forecasts issued on January 1st and April 1st for the six streamflow sites.

Fig. 7.13: Boxplots of ensemble streamflow forecasts at the station East River, Almont, for the dry, wet and average years. The wet, dry and average years are divided based on the terciles – i.e. years with streamflows values below the lower tercile are dry, those above the upper tercile are wet and the rest as average. The horizontal lines are the 25^{th}, 50^{th}, 75^{th}, 90^{th} , and 95^{th} percentiles of the historical data.

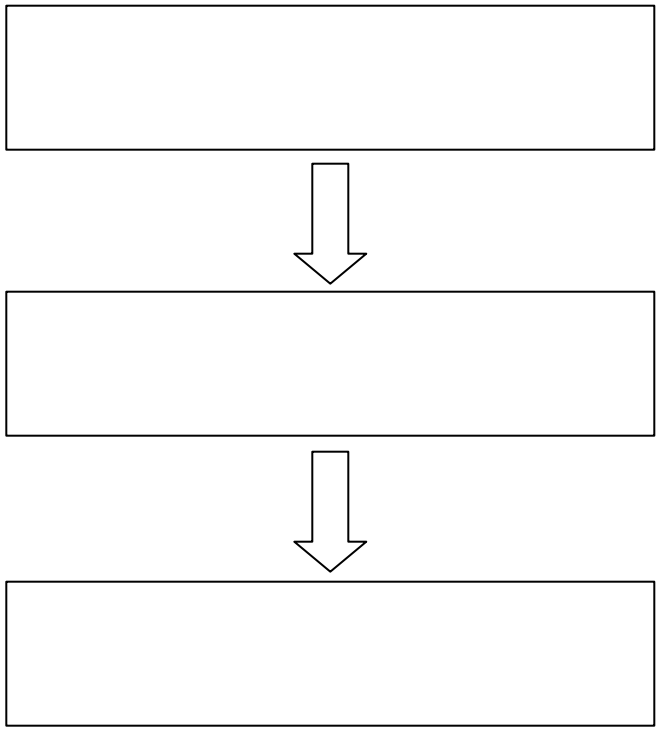


Fig. 7.1: Flow chart of the forecast framework.

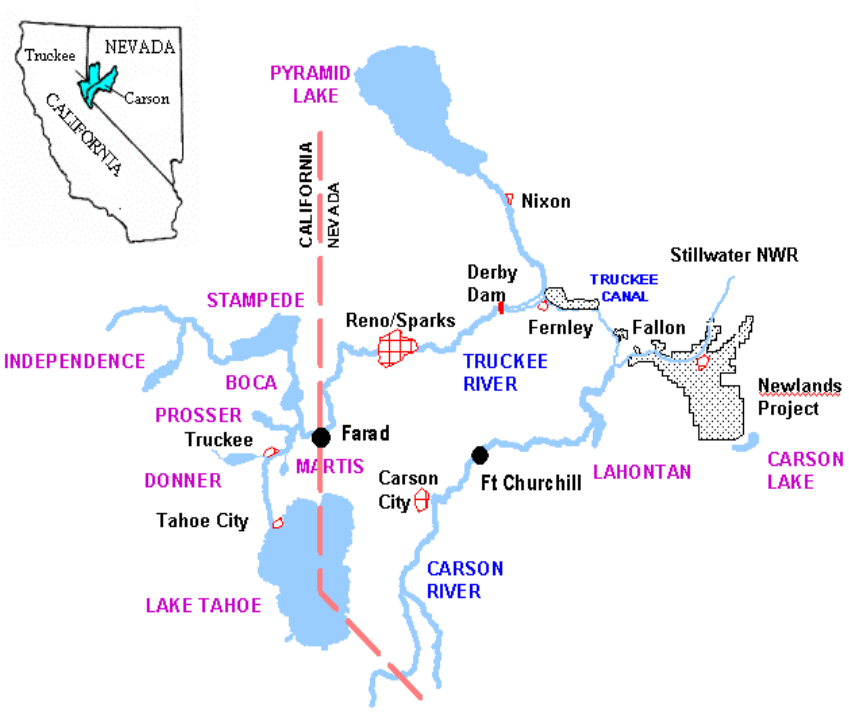


Fig. 7.2: Map of the Truckee/Carson basin

Fig. 7.3: Gunnison River Basin and stream flow locations. USGS gage locations and river names are mentioned below.

1. 09112500 East River at Almont, CO
2. 09110000 Taylor River at Almont, CO
3. 09119000 Tomichi Creek at Gunnison, CO
4. 09124500 Lake fork at Gate view, CO
5. 09147500 Uncompahgre River at Colona, CO
6. 09132500 North fork Gunnison river near Somerset, CO

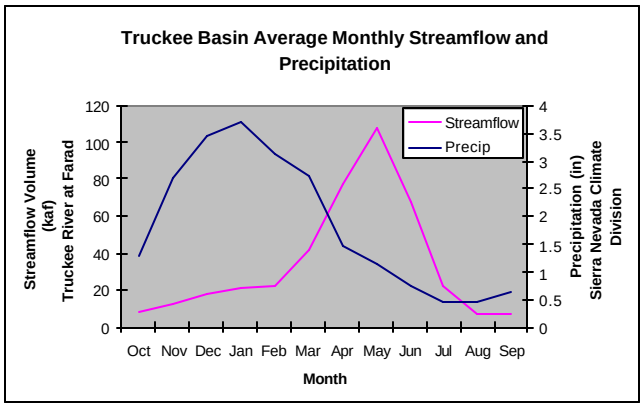


Fig. 7.4: Climatology of streamflows and precipitation in the Truckee River, at the gauging station Farad (based on data for the 1949 – 2003 period).

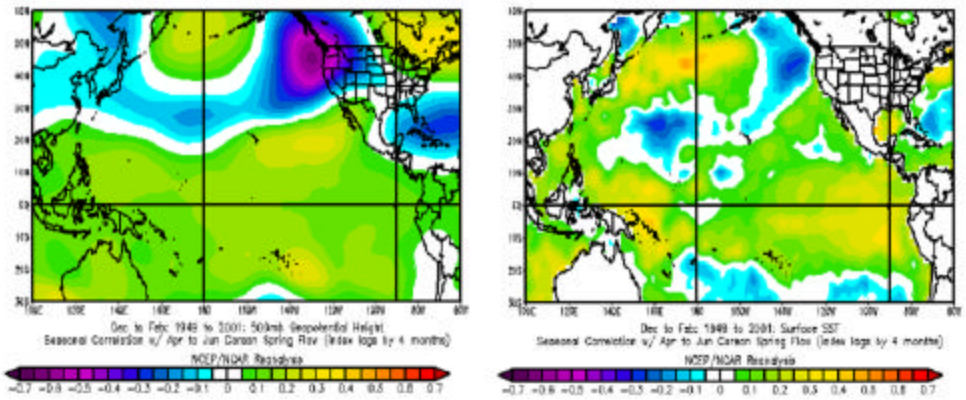


Fig. 7.5: Correlation of Carson River spring streamflows with winter (Dec – Feb) climate variables (a) 500mb geopotential height (Z500) and (b) SST.

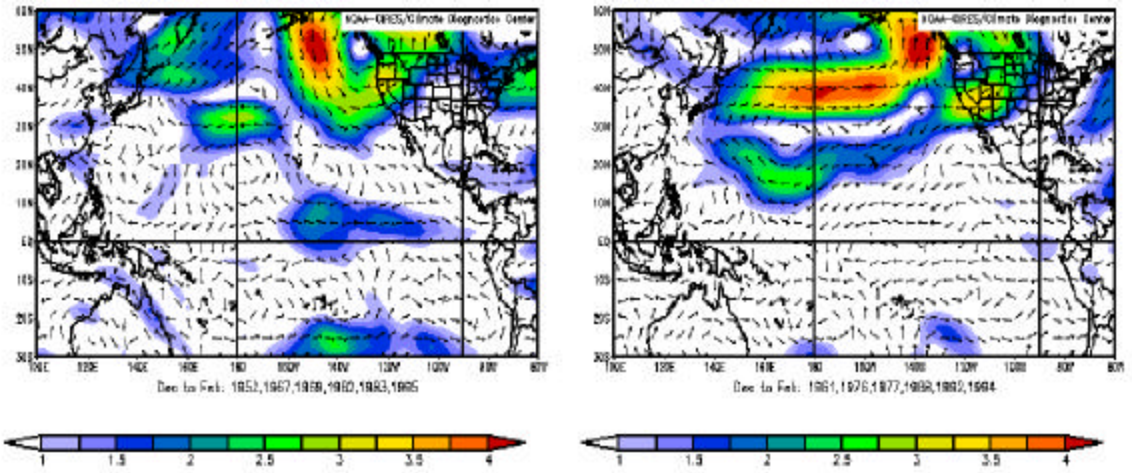


Fig.7.6: Composites of vector winds, SST and Z500 during the winter of high and low streamflow years.

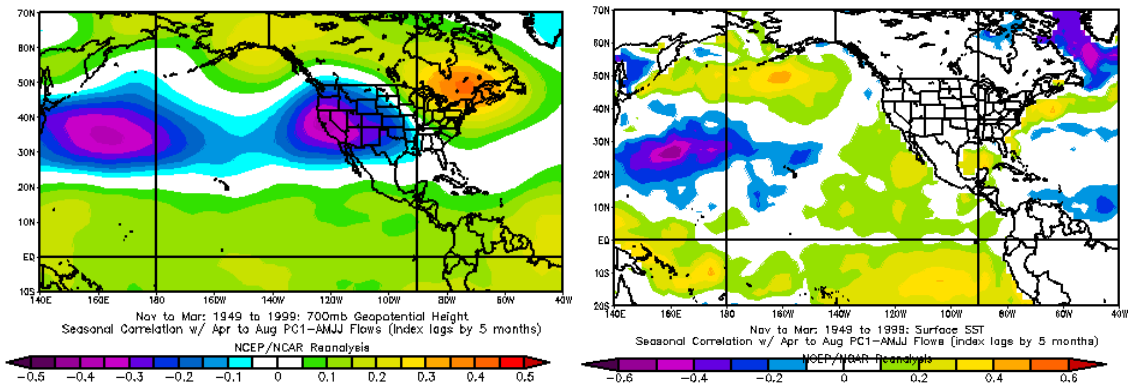


Fig. 7.7: Correlation between PC1 of Spring flows and Nov-Mar Climate indices (a) Geopotential Height – 700mb and (b) Sea Surface Temperature.

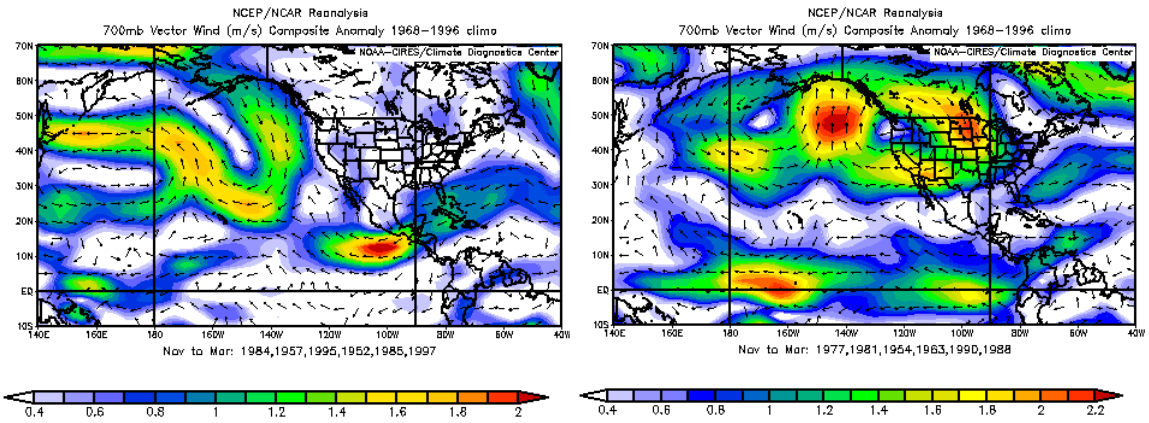


Fig. 7.8: Composite of Vector wind at 700 mb for (a) wet years and (b) dry years.

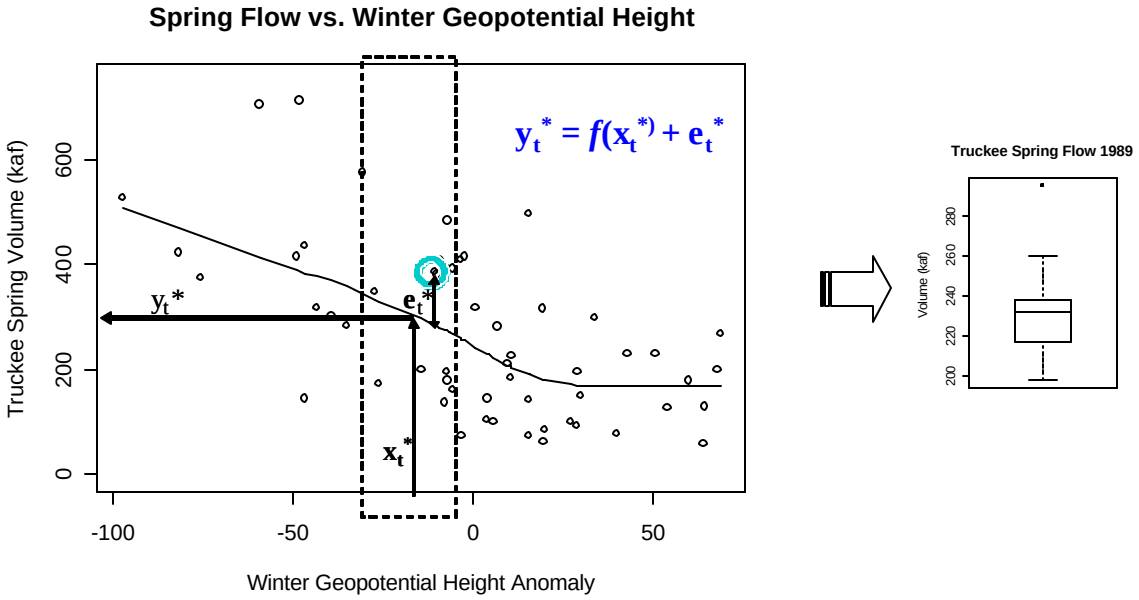
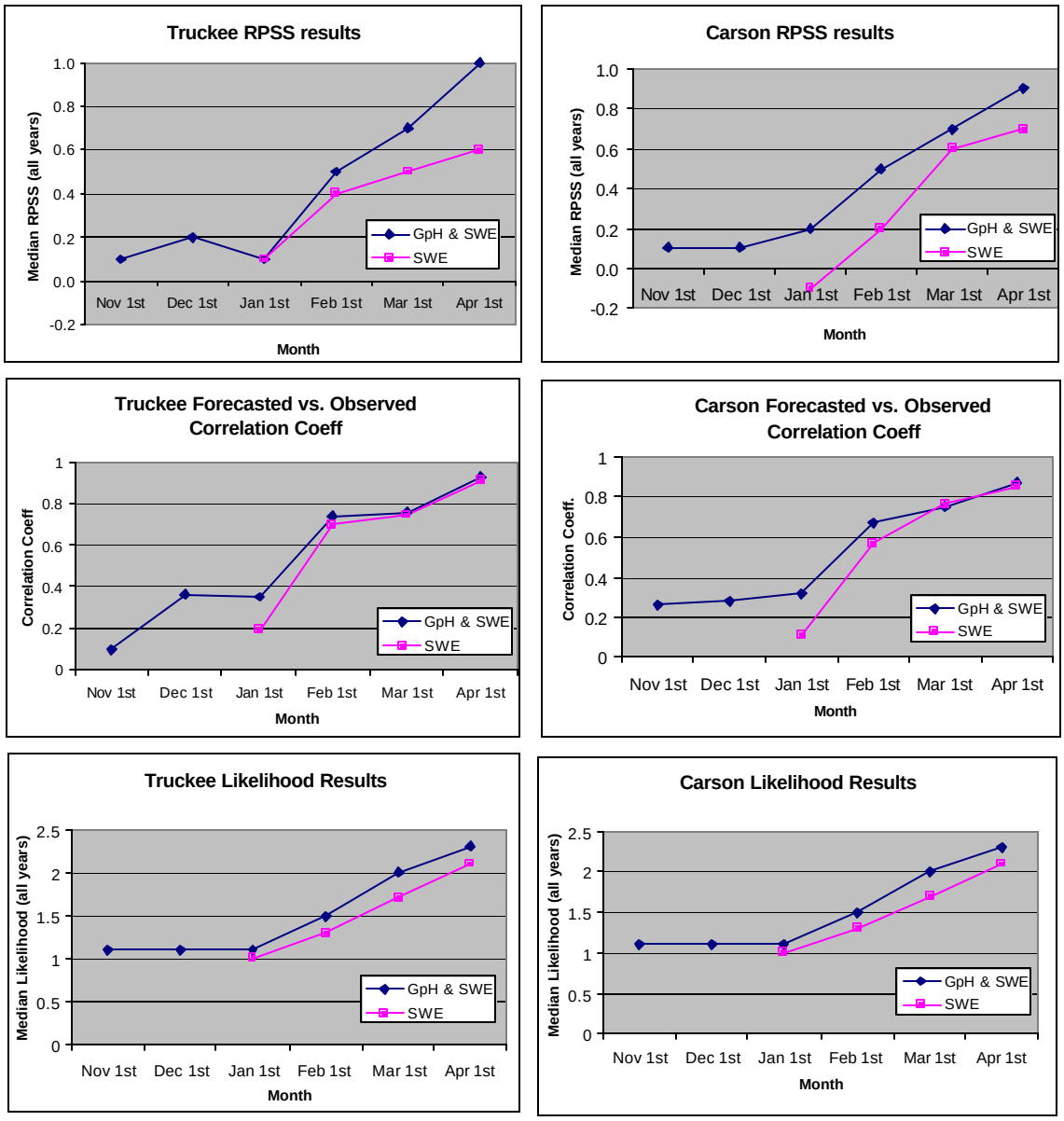


Fig. 7.9: Residual resampling to obtain an ensemble forecast.



.Fig. 7.10: Skill scores of forecasts issued from the 1st of each month Nov – April for Truckee and Carson Rivers.

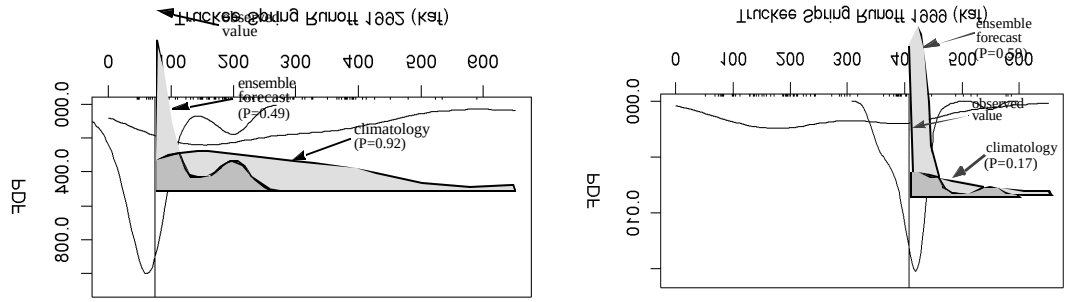


Figure 7.11: PDF of the ensemble forecasts in a (a) dry year (1992) and (b) wet year (1999) for the Truckee River.

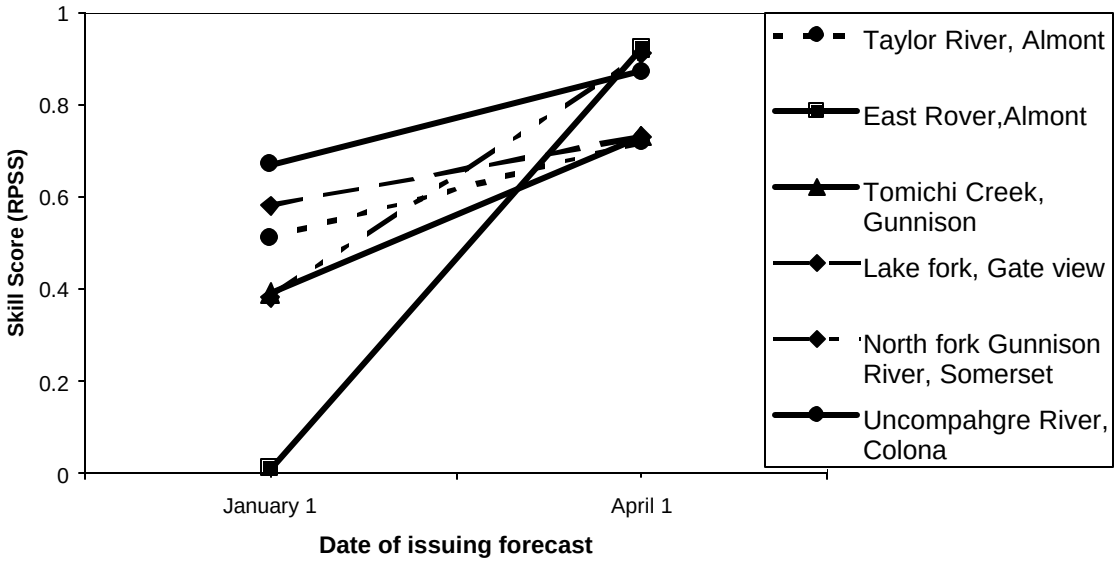


Fig. 7.12: Median RPSS score for forecasts issued on January 1st and April 1st for the six streamflow sites.

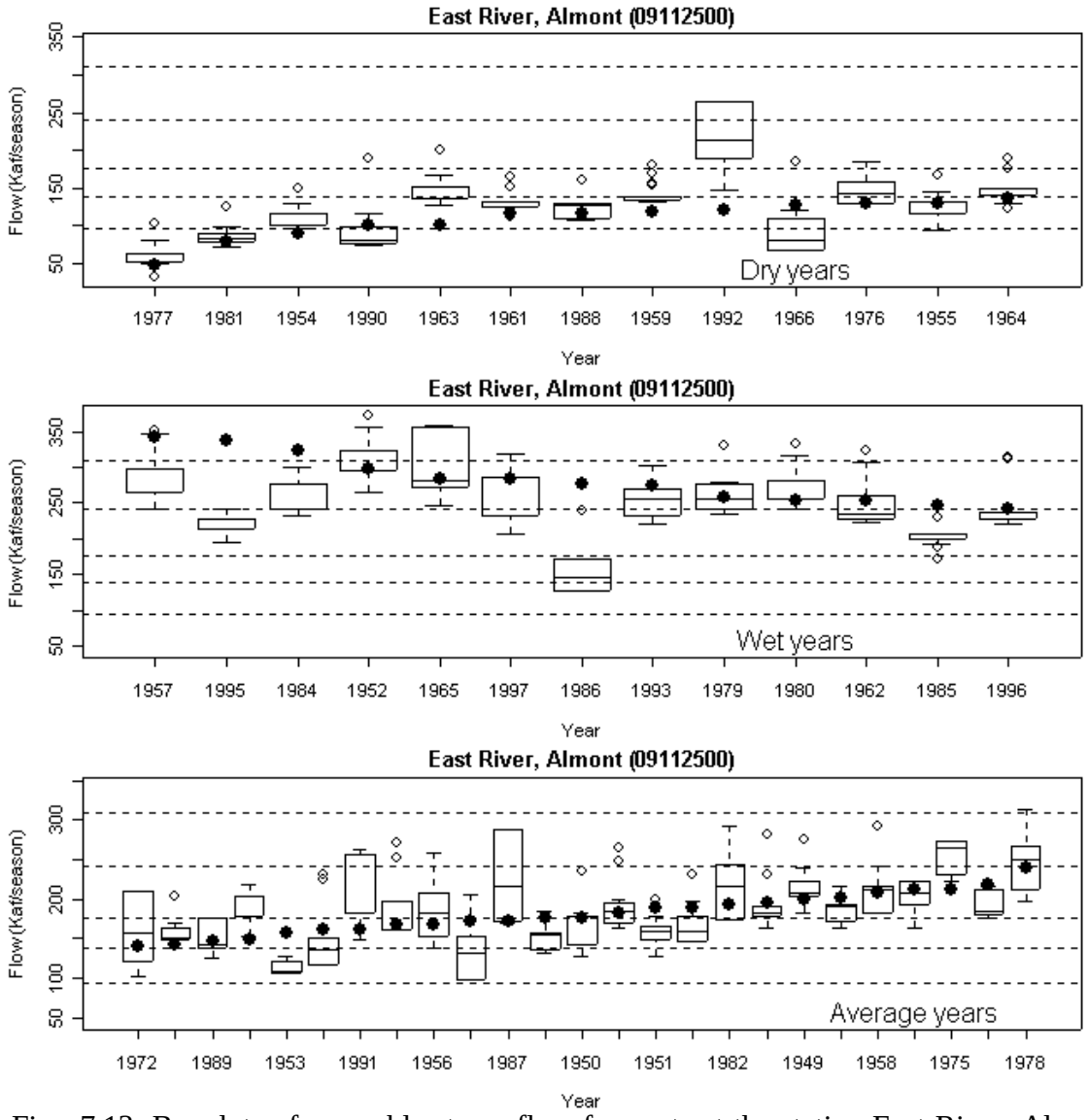


Fig. 7.13: Boxplots of ensemble streamflow forecasts at the station East River, Almont, for the dry, wet and average years. The wet, dry and average years are divided based on the terciles – i.e. years with streamflows values below the lower tercile are dry, those above the upper tercile are wet and the rest as average. The horizontal lines are the 25^{th}, 50^{th}, 75^{th}, 90^{th} , and 95^{th} percentiles of the historical data.