# Recent Analysis and Improvements of the Statistical Water Supply Forecasts for the Upper Klamath Basin, Oregon, and California, USA

**Jolyne K. Lea, David C. Garen, Adam M. Kennedy, John C. Risley**
Natural Resources Conservation Service, National Water and Climate Center; Portland State University; United States Geological Survey, Portland, Oregon, USA

## Abstract

Upper Klamath Basin water supplies in Oregon and California, USA, have been the focus of many competing uses and needs for the past one hundred years. Water supplies have been forecasted in the basin since the 1930s based on the relationship of streamflow with the seasonal snowpack and climate. In 2001, the 5th driest year on record, agricultural irrigation was curtailed in much of the basin, with the little available water allocated to support the survival of endangered and threatened fish. The lack of available irrigation water generated a large outcry in the local and national agricultural community, prompting collaborative research beginning in 2003 to improve the accuracy of the water supply forecasts, which would enhance water management decision-making in the watershed. The focus of the research described in the present paper was to review the current statistical forecasting techniques and investigate other statistical techniques as well as research additional data variables to use in the water supply prediction models.

## Introduction

USDA Natural Resources Conservation Service (NRCS) has been forecasting water supplies in the Klamath Basin since the 1930s. The NRCS forecasts water supplies at over 700 other stream gauge stations and reservoir inflow points throughout the western United States. The relationship between winter snowpacks and the resulting spring runoff spawned the development of statistically based water supply forecast methods beginning in the early 1900s. Water supply forecasts for the western United States have been traditionally requested and used by federal, state, and local water managers for flood control, irrigation, hydropower generation, and municipal use. Water supply forecast use has expanded significantly in the last decade to include fish and wildlife management and winter and summer recreation.

Most irrigation water in the Klamath Basin is allocated and delivered by the U.S. Bureau of Reclamation (BOR), which operates three reservoirs in the basin (Upper Klamath Lake, Clear Lake, and Gerber Reservoir). The BOR supplies water to irrigate approximately 810 $km^2$, which varies annually (Risley, et al., 2005).

In the 1990s, the U.S. Fish and Wildlife Service (USFWS) designated two Upper Klamath Basin fish species endangered: Lost River Sucker (*Deltistes luxatus*) and Short Nose Sucker (*Chasmistes brevirostris*). The Coho salmon in the Klamath River was also listed as threatened by the National Oceanic and Atmospheric Administration (NOAA) Fisheries Service. Water management plans were developed to provide the appropriate amount of water to improve these fish populations.

In 2001, the 5th driest year on record (based on data from 1901-2006), a very limited supply of water was available for irrigation, power, and endangered fish needs. Based on the April 2001 water supply forecast, the BOR determined that to comply with the Endangered Species Act, no irrigation water would be allocated to the farmers in the BOR Project, though other farmers in the area were able to irrigate. This decision caused many protests throughout the local, state, and national agricultural community and received the attention of the White House and the President of the United States. The resulting legislation provided some federal funds to enhance and conserve water supplies in the basin. The immediate actions included emergency well drilling, water conservation as well as long-term projects such as irrigation efficiency improvements and vegetation management. There was also funding to support improvements in water supply predictions, the basis for water conservation and management decisions. These improvements include additional data collection stations, hydrologic model development, and a study of the accuracy of the statistical water supply forecasts and ways to improve them. There is also a continuing effort to educate the water managers and the public on the use and limitations of water supply forecasts.

## Geography

The Upper Klamath Basin encompasses approximately 20,720 km$^2$ and is located in south-central Oregon and northeastern California. The Klamath River originates at Upper Klamath Lake in Oregon and flows in a southwesterly direction, draining the Cascade Mountain Range on the west and smaller mountains on the north and east sides of the basins, and discharges to the Pacific Ocean. The Oregon part of the basin is approximately 14,500 km$^2$ (Lea and Pasteris, 2004).

## Data Network

The Klamath Basin data collection network used to generate water supply forecasts is distributed throughout the mountainous areas of the basin as shown in Figure 1. The primary source for the climate and snowpack data used for water supply forecasting is the SNOw TELemetry (SNOTEL) network operated by the NRCS. In the Klamath Basin, the SNOTEL network consists of 19 remote stations that collect hourly precipitation, snow water equivalent (SWE), snow depth, and temperature data. Six SNOTEL sites have been augmented to provide soil moisture and soil temperature measurements at five different soil depths, and four of these sites also measure solar radiation, wind, and relative humidity.

In addition, six manually measured snow courses provide SWE and snow depth data once a month, during January through June. A snow course is a permanent site where these manual snow measurements are taken by trained observers near the first of the month during the winter and spring. Generally, the courses are about 1,000-feet long and are situated in small meadows protected from the wind. The observers take measurements along a set transect at regular intervals, averaging the measurements over the course. There are also four aerial markers in the basin consisting of poles with crossbars that indicate snow depth, which are read from a small airplane once a month, during this same period. The single SCAN (Soil Climate Analysis Network) site provides soil temperature, soil moisture, and weather data elements but does not measure snow due to its location in an agricultural field at a lower elevation where snow is ephemeral.

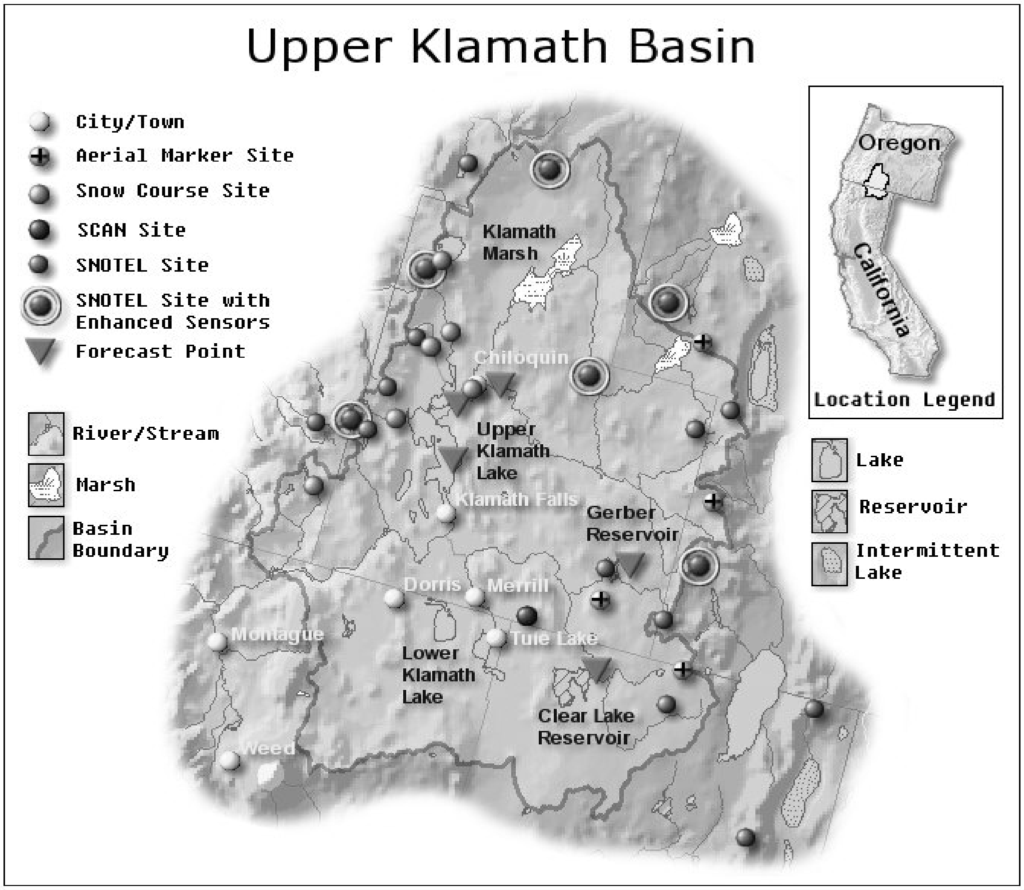

Figure 1. Upper Klamath Basin hydromet network and water supply forecast points.

Precipitation from five low elevation National Weather Service (NWS) cooperative observer sites is also used in water supply forecasts. The five water supply forecast points within the basin are located at long-term stream gauges that provide historic and current streamflow data collected by the BOR, U.S. Geological Survey (USGS) and the Oregon Department of Water Resources.

## Water Supply Forecasting

A water supply forecast is the expected volume of water available during a specific period of time at a specific location. Examples include lake inflow, reservoir inflow or flow at stream gauge over a multi-month time step or season. In the western United States, statistical forecasts are also made for annual events such as peak flow and date of the peak and for recession (low flow) dates and stage. Seasonal water supply forecasts are used for water management decision-making such as flood management, irrigation, municipal use, wildlife and fish, hydropower, and recreation. The seasonal volume forecasts are often used as an input for daily water management models.

Water supply forecasting in the Klamath Basin is based on statistical models relying on a linear regression of historic monthly hydroclimatic input variables (SWE, precipitation, streamflow) against historic observed streamflow volume. These regression equations are developed using the principal components statistical method developed by Garen (1992). The principal components method was developed to account for the intercorrelation among predictor variables (which especially affects precipitation and snow observations for a given time period among stations).

Once an initial set of candidate stations and climate elements has been selected, screening is done both manually and with the help of an automated search routine. A final set of predictor variables is selected balancing statistical optimality (i.e., minimizing the standard error) with the selection of hydrologically meaningful variables. Consistency in the variable usage from month to month during the forecast season is important to minimize forecast fluctuations and to ensure physical interpretability of the forecasts. As a robust measure of model accuracy, a jackknife test is performed. The test is an iterative procedure of removing each year's observations one at a time, recomputing the model's regression coefficients, predicting the removed year, then returning that year's observation and removing the next one. This is repeated until a series of predictions is obtained, each of which is from a model that did not include the respective year in the calibration. This test is used to evaluate each candidate model, and the standard error calculated from the jackknife predictions is used to develop confidence bounds around forecasts. The statistical models are normally developed with 20 to 40 years of data to ensure the robustness and physical representativeness of the statistical relationships. Each monthly forecast model is developed independently. Thus, a given month's forecast is not dependent on the previous month's forecast, although consistency is maintained by using similar data stations from month to month.

## Artificial Neural Network Model

As one experiment to improve water supply forecast accuracy, the USGS tested the Artificial Neural Network (ANN) model (Figure 2). This statistical method is a flexible mathematical structure capable of describing complex nonlinear relationships between input and output data sets that are typically found in natural systems (Risley, et al., 2005). The USGS also tested the autoregressive artificial neural network using past streamflow to predict future streamflow volumes for 1979 through 2003 in a weekly time step (Risley, et al., 2005). In both of these techniques, forecasts were developed for the five forecast points in the Klamath Basin for the months of January through June and were compared to the principal component method. In the comparison, the principal components model performed better at all forecast points in April, though there were mixed results in other months, suggesting there would be little to gain if the ANN method were adopted.

## New Variables for Statistical Models

Several new variables were evaluated for their potential in improving water supply forecast accuracy with principal components regression models. These variables include those representing groundwater conditions (wells and springs), the average monthly temperature during the spring season to assist in describing snowpack melt conditions, a new climate teleconnection index (Trans-Niño Index) to indicate climate conditions for the upcoming winter, and the basin mean areal precipitation. These variables were analyzed in conjunction with the current variables used in the forecast equation (SWE, precipitation, streamflow).

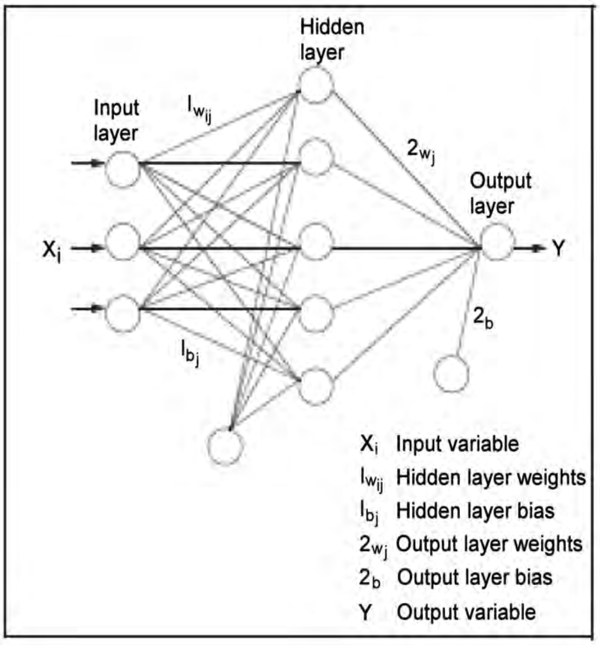

Figure 2. An example of a neural network model architecture with three input layers, five hidden later nodes, and a single output. (Risely, et al., 2006).

## Wells and Springs

From other studies in the Klamath Basin, it is known that groundwater flow and storage are significant components of the basin hydrology due to its volcanic nature. Data from wells and springs in the basin were reviewed, and it was determined that one Oregon Department of Water Resources observation well had a long-term dataset that could be edited and used, and one Oregon Department of Water Resources streamflow gauge that measures a large spring shortly after it begins to flow had a data set that was robust and quality controlled. Both of these data sets provide good correlation to the spring and summer streamflow in the Klamath basin. The single correlation between the spring streamflow at Fall River to the Williamson River streamflow ranged from 0.28 to 0.59. The correlation was better in the forecasts for later season summer flows, which is logical in that the springs would be best correlated with summer baseflow. The well level correlation to the Williamson streamflow was also good at -0.45 to -0.52, as the depth to groundwater is another good indication of baseflow conditions.

## Spring Season Temperature

While spring temperature is a critical element in physically based models, it has rarely been used in statistical models. It was surmised that the temperature during the months of March, April, and May would be of help in forecasting streamflow by indexing snowpack ripeness, melt rates, and evapotranspiration losses. Obviously, this is a negative relationship, where warmer temperatures are associated with lower streamflow. The only station with a sufficiently long temperature record was Crater Lake National Park Headquarters, located at the northwestern edge of the basin. It was found that the correlation coefficients between average temperature during March, April, and May and the subsequent seasonal streamflow volume were in the range -0.26 to -0.53. Of the three months, March provided the best correlation to subsequent streamflow. These variables, then, are useful for improving the accuracy of forecasts issued in the months of April through June.

## Climate Indices

There are several standard climate indices that are used in water supply forecasting in the western United States and elsewhere. The Southern Oscillation Index (SOI) is used in some parts of the Pacific Northwest, the northern Rocky Mountains, and the Southwestern states of Arizona and New Mexico. There is a region between these two areas that does not have a strong correlation with the SOI, and the Klamath Basin falls on the edge of this area. The Pacific Decadal Oscillation (PDO) is also not well correlated, but it is useful in identifying decadal-scale climate regimes. A new index, the Trans-Niño Index (TNI), was the focus of our work. The TNI, first published by Trenberth and Stepaniak (2001), is the standardized equatorial sea surface temperature (SST) gradient between the Niño 1+2 and Niño 4 regions (Figure 3). The evaluation was limited to 1980-2004 to align with the current climate regime as defined by the PDO. The TNI during the fall and early winter provides a good correlation (r of approximately 0.7) to streamflow in this current warm PDO phase. Recent work has shown that there is a broad regional pattern of TNI correlation to streamflow (Kennedy et al., 2005).

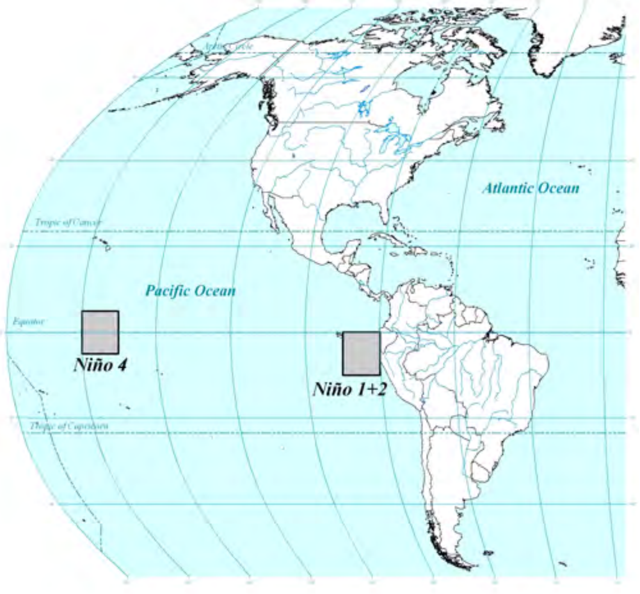

Figure 3. The TNI is the standardized equatorial sea surface temperature (SST) gradient between Niño 1+2 and Niño 4 regions (Trenberth and Stepaniak, 2001).

## Mean Areal Precipitation

Mean areal precipitation was calculated as spatial averages derived from monthly time series grids estimated with the Parameter-Regression on Independent Slopes Model (PRISM) (Daly et al., 1994; http://www.ocs.orst.edu/prism).

The monthly mean areal precipitation data series derived from PRISM grids were compared to the individual station data series. The correlations of each month's data to the subsequent April-September streamflow for the Williamson and Sprague sub-basins are shown in Figures 4 and 5. These figures indicate that the mean areal precipitation derived from the PRISM grids have better correlations with streamflow overall than any individual station. Individual stations may have a better correlation for one or more months, but for the season, the more robust and consistent correlation of the areal precipitation is preferred.

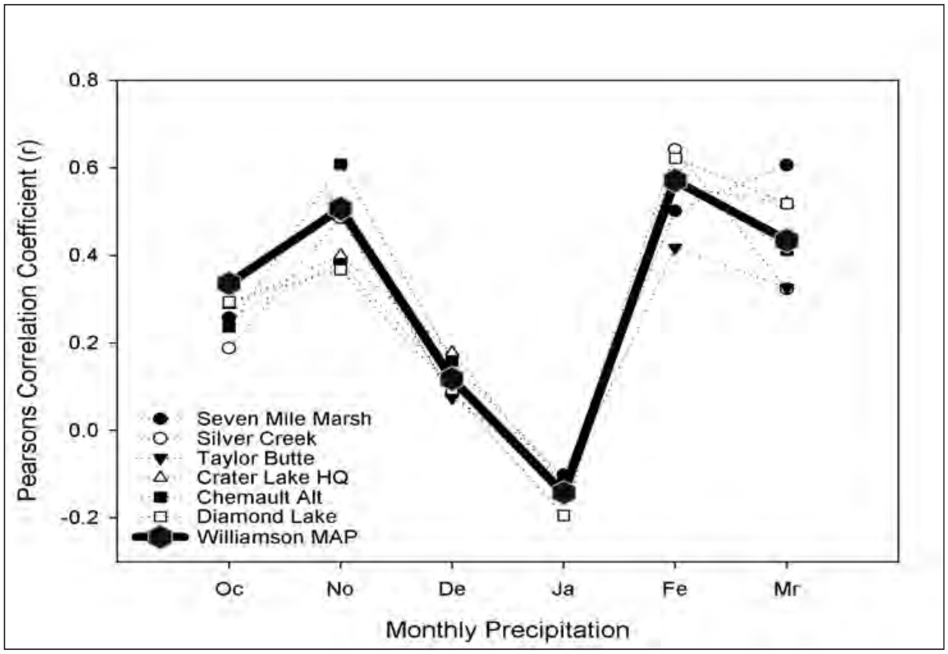

Figure 4. The correlation of the monthly mean areal precipitation and individual stations with April through September streamflow volume in the Williamson subbasin.

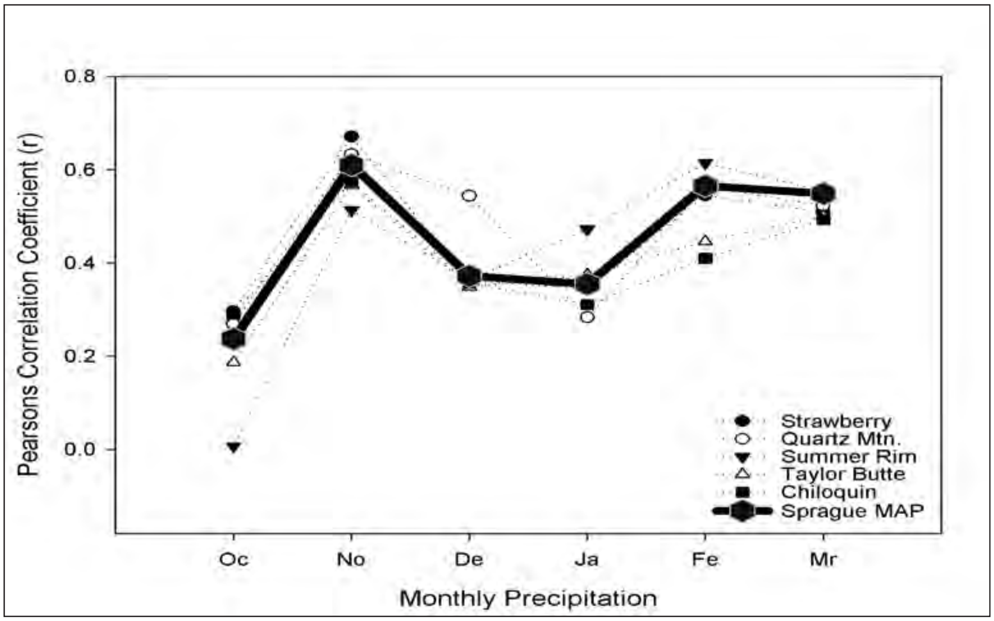

Figure 5. The correlation of monthly mean areal precipitation and individual stations with April through September streamflow volume in the Sprague subbasin. (Kennedy, et al., 2005).

## Conclusion

Forecasting future water supplies continues to be the primary planning tool for water resource management in the western United States and especially in the Klamath Basin. There will continue to be emphasis on improving the forecast accuracy for better decision making for the multiple and often conflicting water resource needs. This is complicated by unique basin characteristics, extreme weather events, and changing climate conditions.

Snowpack, precipitation, and streamflow have been long standing good predictors of future streamflow. They will continue to be the mainstay of statistical forecasting in the western United States. This study examined several new variables that improve forecast accuracy, and have potential for use in statistical streamflow forecasting models beyond the basin under study. These variables include groundwater data (wells and springs), spring season temperature, climate teleconnection indices (especially the Trans-Niño Index), and mean areal precipitation derived from spatial grids. The groundwater variables provide long- term, multi-seasonal conditions of the basin hydrology and the baseflow characteristics. The temperature and mean areal precipitation variables are related to the current weather, and they improve our knowledge of the status of the snowpack and resulting streamflow in the basin. The Trans-Niño Index provides a much needed early prediction of the future weather expected in the basin. Early season forecasts allow additional time for the implementation of conservation and mitigation measures to offset any water shortages or surplus. All of these variables have a good correlation to the streamflow period of interest and together provide an increase in forecast accuracy. The forecast techniques and variables used here may also have applicability in other basins where hydrometeorological and climate variables contain sufficient information to make useful streamflow forecasts.

**Author E-mail contact:** Jolyne.Lea@por.usda.gov

## References

Daly, C. R. P. Neilson and D L. Phillips. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*. 33:140-158 pp.

Garen, D. C. 1992. Improved techniques in regression–based streamflow volume forecasting in *Journal of Water Resources Planning and Management.* 118(6):654-670 pp.

Kennedy, A. M., D. C. Garen, and R. Koch. 2005. Effects of incorporating a climate index into the Upper Klamath operational water supply forecast: Trans Niño Index. Proceedings of American Water Resources Association 2005 Annual Meeting, Seattle, Washington.

Lea, J. and P. Pasteris. 2004. Preliminary Assessment - Water Year 2004 Upper Klamath Lake Inflow Forecasts. Internal NRCS Report, Portland, Oregon.

Risley J., M. Gannett, J. Lea, and E. Roehl. 2005. An analysis of statistical methods for seasonal flow forecasting in the Upper Klamath River Basin of Oregon and California. USGS Scientific Investigations Report. 2005-5177 pp.

Trenberth, K.E, and D.P. Stepaniak. 2001. Indices of El Niño evolution. *Journal of Climate.* 14:1697-1701 pp.