

Scott River Runoff Forecast Model: Investigation of Potential Formulation



A Report for the Scott River Watershed Council

Prepared by

Michael L. Deas, Ph.D.

and

Stacy K. Tanaka, M.S.

Watercourse Engineering

133 D Street, Suite F

Davis, California 95616

(530) 750 - 3072

December, 2005

Scott River Runoff Forecast Model: Investigation of Potential Formulation - Executive Summary -

Effective management of water resources requires planning and planning requires two basic concepts: (1) the status of current conditions and (2) an estimate of future conditions. Both of these pieces of information are typically benchmarked to long-term conditions to improve decision-making capabilities. The two tasks identified by the Siskiyou County Resources Conservation District (RCD) as important in water resources management – runoff forecasting and water year type designations – seek to quantify these critical pieces of information. The topic of runoff forecasting is the focus of this report. An appendix characterizing water year types based on historical Scott River data is presented in the appendix.

A unique attribute of the Scott River basin that sets it apart from other basins where large water projects have been developed is that there is no appreciable surface water storage. As such, the typical characterizations of water year type and runoff forecast models that incorporate storage as a variable do not readily apply. The methodologies vary widely with regard to water availability forecasts, ranging from physically-based models to empirical relationships. In general, physically-based models require a significant amount of studies and data to represent the local processes and to apply the model. The empirical models (linear and non-linear regression models or time series models such as autoregressive moving average models (or variations thereof)) are generally developed based on sufficiently long time series of data to create a relationship between known parameters and a desired future parameter, e.g., base future runoff timing on known snow pack.

Empirical, or statistical models, are explored herein. Relationships using precipitation, as snowfall and rainfall, as well as both impaired flow data and estimates of unimpaired flow data are used to assess potential relationships that could provide the forecasting ability to assist local landowners and resource managers in assessing annual water supply availability.

Three statistical relationships are developed within this report using linear and multiple linear regression for unimpaired flow at the Fort Jones gage versus a single variable April 1st snow water content linear regression, a multivariable April 1st snow water content linear regression, and a single variable October to March rainfall linear regression. The results from the relationships are compared with each other and with the observed total annual runoff totals to assess their performance based on average error, as well as tendency to under or over-predict runoff. In terms of the models, the under or over prediction of the annual unimpaired river flow may or may not be significant depending on what the total annual runoff is. It is necessary to determine the desired level of accuracy prior to selecting or rejecting a particular model. For this report, all relationships effectively reproduced the annual forecast trends, with the multivariable snow water content relationship illustrating the best performance. It should be noted that

significant changes in environmental, meteorological, land use or other process changes could increase the uncertainty in these models.

Overall, the forecast relationships are restricted by the limited long-term data sets for the full range of parameters that potentially correlate to annual runoff. Several of the Scott River tributaries have had gages installed in the last few years, but these records are not sufficiently long at this time to be useful in such statistical relationships. Also, extended time series of data for groundwater conditions throughout the basin and of the unimpaired runoff were identified as potentially useful parameters to include in such models, but exploration of these data was beyond the scope of this initial investigation. Nonetheless, several useful concepts have been identified in this project and as ongoing monitoring leads to longer records such relationships could be revisited.

Scott River Runoff Forecast Model: Investigation of Potential Formulation Table of Contents

1. Introduction	1
2. Modeling Background	1
2.1. Model Types	1
2.2. Modeling Approaches	2
2.3. Model Processes	2
2.4. Model Selection	3
3. Forecast Modeling	3
3.1. Introduction	3
3.2. Statistical Methods	4
3.3. Physically-Based (Theoretical) Models	7
4. Scott River Watershed Data	8
4.1. Characteristics of the Scott River Watershed	8
4.2. Data Considered for Forecasting Models	9
5. Scott River Precipitation Characteristics	13
5.1. Rainfall	13
5.2. Snow Water Content	15
6. Scott River Flow Characteristics	17
6.1. Monthly Flow Distribution	17
6.2. Annual Scott River Unimpaired Runoff Estimates	18
6.3. Annual Non-Exceedence Plots	19
7. Runoff Forecast Modeling	22
7.1. Runoff Forecast as a Function of April 1 st Snow Water Content	22
7.2. Rainfall Single Variable Linear Regression	31
7.3. Multivariable Snow Pillow and Rainfall Linear Regression	34
7.4. Comparison of Forecasting Tools	35
8. Other Factors	37
8.1. Tributary Inflows into the Scott River Valley	39
8.2. Groundwater Levels	40
8.3. Consumptive Use	41
9. Conclusions	42
10. Recommendations	43
10.1. General Recommendations	43
10.2. Detailed Recommendations	44
11. References	45
12. Appendix – Water Year Types	46
12.1. Water Year Types	46

Scott River Runoff Forecast Model: Investigation of Potential Formulation

1. Introduction

Runoff forecasting is an important component in water resources planning and management. Planners use the available information, such as precipitation as snow depth and water content or rainfall, to predict the volume and timing of runoff that will be available later in the year. These estimates help managers as they try to allocate water in an efficient and productive manner. The models that generate the estimates can range from very simple, using minimal data, to very complex that include a number of different meteorological, hydrological, and other variables.

The California Department of Water Resources (DWR) does not forecast the flows in the Scott River at present. DWR reports the April 1 snow water content yield estimates in lieu of a detailed runoff forecast. For other rivers that DWR manages, more detailed forecasts are made using established relationships that often times include current flow conditions, the previous year flow conditions, and upstream surface storage volumes. The Scott River watershed does not have appreciable surface water storage and any forecast modeling would more likely be based on snow pack or precipitation conditions.

This document contains four sections. The first section provides the reader with an overview of general forecast modeling concepts. The second section focuses on forecast modeling techniques. The third section discusses the Scott River Watershed, including a review of the available data and the presentation of a simple forecast regression model using the existing data. The fourth section provides the recommendations for future activities associated with constructing and maintaining forecast models.

2. Modeling Background

2.1. Model Types

Models are applied to a wide variety of real world problems. As such, there are many different types of models and modeling approaches that range from simple to complex. Physical models are scaled representations of actual systems (e.g., a flume experiment) and are not addressed herein, while mathematical models consist of empirical (e.g., statistical model) or theoretical (e.g., based on conservation laws) formulations (Figure 1). Empirical models use known data in conjunction with statistical relationships between system components and/or events. Theoretical law models rely on scientific principals, such as conservation of mass, momentum, and/or energy, to represent a system or event. Based on the modeling objective, critical system components are identified and an appropriate modeling approach is selected.

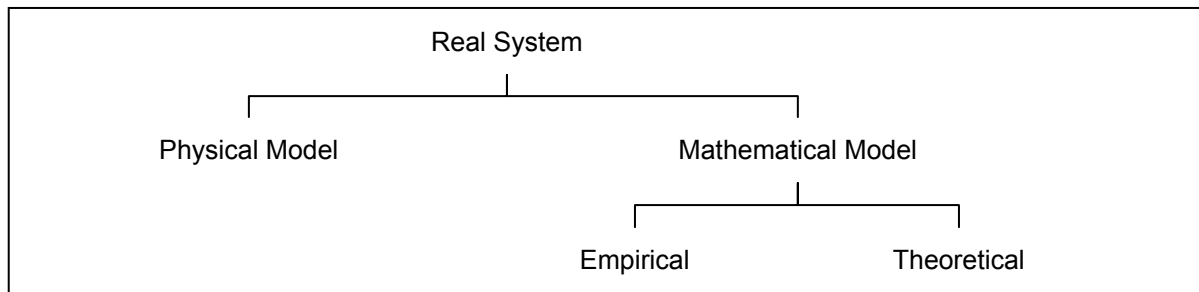


Figure 1. Model Types Flow Chart.

2.2. Modeling Approaches

Generally models can be categorized into one of three broad types. These groups are simulation, statistical, and optimization models which are described below.

Simulation models falls under the theoretical model type and are a method of replicating a system or process (e.g., rainfall, infiltration, runoff) using defined mathematical relationships. Simulation modeling is sometimes called descriptive modeling because the various system processes are described mathematically and the outcome is in response to specified conditions.

Statistical models fall under the empirical model type and are methods of replicating a system or process based on collected data using some type of statistical method, such as regression analysis. Statistical analysis can be beneficial when the processes involved are not well understood and/or when there is a large volume of data available. A drawback to statistical modeling is that the data used to construct the model is time and site specific. Thus, results of such models may not be applicable under conditions which differ dramatically from the data used to construct the model; a limitation also true of simulation models. Nonetheless, many sub-process and/or relationships within simulation and optimization models may be represented by statistical models. Further, statistical models are often a relatively quick and inexpensive way to attain a forecasting or predictive capability.

Although there are optimization-type processes that can be used in some forecast models, these are not generally used for runoff forecasting in watersheds where the uplands are largely undeveloped. Thus, optimization models will not be included in the discussion of forecasting models.

2.3. Model Processes

Within a modeling framework variables are generally considered deterministic or probabilistic. In deterministic modeling it is assumed that there is no uncertainty in any of the input parameters. A converse approach to this is stochastic or probabilistic modeling, wherein, the values of the input parameters are not known with certainty, so probabilities are used to determine the likely outcomes. Examples of processes that are modeled using probabilistic techniques include many of the hydrologic processes associated with ground and surface water interactions.

Runoff forecasting is inherently a stochastic process, because meteorological, hydrological, and watershed conditions are not known with certainty beforehand. When using a physically based model to forecast, these conditions must be estimated based, for example, on conservation laws. However, in certain cases the stochastic nature of the processes can be minimized when applying statistical models by basing the forecast on known information. For example, relating the total October through March accumulated snow pack to April through September total runoff can yield simple relationships (e.g., regression relationships) that can be useful for estimating runoff volumes and, in some cases, the timing of runoff. A key underlying assumption in such a statistical model is that there is sufficient year-to-year correlation among variables, i.e., snow falls in winter and melts in spring, to allow a statistically relevant relationship to be formulated.

2.4. Model Selection

A valuable aspect of adopting and implementing a model is the careful consideration of important system components and desired outcome. As such, modeling can be a very useful exercise in learning about a watershed or river basin. It can also be a very complex task. Prior to modeling or even selecting a modeling approach, the user(s) should consider two key concepts:

1. Identify the objective of the modeling exercise (e.g., what applications will the model be used for?);
2. Identify the type/scale of the model that is best suited to meet the modeling objective (e.g., space and time aspects of the input and desired output).

A clear, specific objective prior to model selection will reduce many of the difficulties commonly associated with application and interpretation of a modeling effort. When selecting a modeling approach, the user(s) should keep in mind construction/implementation requirements, ease of use, costs (maintenance, data collection, software), and desired accuracy (or uncertainty) of the model. For this investigation, the identified objective was to explore potential annual runoff forecast models based on existing Scott River runoff and precipitation data.

3. Forecast Modeling

3.1. Introduction

One of the most common reasons for the development of modeling capabilities is for use in future planning processes. Resource managers and other land and water use activities often require methods to predict the state of the system in the future given current and historical conditions. Most people have had exposure to forecasts; for example in the form of daily weather forecasts. In these cases, meteorologists estimate conditions such as air temperature, wind, atmospheric pressure, and precipitation in the near future (usually for a range of one to seven days) based on recent and current weather patterns. These forecasts can help people plan activities over a period of several days. Water planning makes use of forecasts to help plan activities ranging from flood routing to water allocations over short and long periods of time. Outlined below are selected statistical modeling approaches and a discussion of physically based (or theoretical) mathematical models.

3.2. Statistical Methods

Statistical relationships between historical measured data sets can be derived using a variety of different means. The simplest method is the single-variable linear regression; from there methods becoming increasingly complex. This section presents some basic statistical methods, including single and multiple-variable linear regression, polynomial regression, and autoregressive (AR) methods. For additional details the reader is referred to Yevjevich (1972), Maidment (1993), Waissi (2002), Systat (2004), and Meko (2005).

3.2.1. Linear Regression

A linear relationship between the dependent variable (e.g., runoff) and one or more independent variable(s) (e.g., snow water content, air temperatures, and/or precipitation¹) can be determined using linear regression. Regression analysis attempts to determine the relationship between the dependent variable and one or more independent variables. The general form of the equation for multivariable regression is:

$$y = m_1x_1 + m_2x_2 + \dots + m_kx_k + b$$

where ‘y’ is the dependent variable, m_i are the coefficients, x_i are the independent variables, and ‘b’ is the y-intercept. Multivariable regression can be reduced to its simplest form: single variable linear regression ($k = 1$). The general form of the single variable linear regression is:

$$y = m_1x_1 + b$$

where ‘y’ is the dependent variable, m_1 is the slope of the line (or the coefficient), x_1 is the independent variable, and b is the y-intercept.

Most statistical software packages produce a suite of statistical parameters that can be used to determine the quality of the regression relationship. Herein, a stepwise linear regression analysis was used to assess various relationships. Some of the more important values include the R-squared (R^2) value, standard error, F-statistic, and the P-value. The R^2 value (coefficient of determination) indicates how much of the variance in the dependent variable (y) is explained by the dependent variable(s) (x_i value(s)). Higher R^2 values suggest that the selected dependent variable(s) do a relatively better job of predicting the independent variable, than relationships with lower R^2 values. The standard error (Std. Error) of the estimated coefficient is the amount by which the coefficient value could vary depending on which data points from the entire sample were used to estimate the coefficient. The larger the standard error is, as a percent of the coefficient value, the less likely the variable is to be useful in predicting the value of the dependent variable. The F-Statistic is used to determine which variables should be included (e.g., $F > 4.0$) or removed (e.g., $F < 3.9$) from a stepwise linear regression. Additional information on the application and interpretation of these

¹ Dependent variables can also be “transformed” to provide a more linear distribution, an improved distribution about the regression line, or a more normal (or Gaussian) distribution. This exercise can reduce the variance in the residuals. Transformation may include taking the logarithm or square root of the data, or some other mathematical function.

statistical measures in regression analysis can be found in most elementary statistical texts, as well as under the “help” portion of most statistical software packages.

Thus, using a relatively small set of statistical parameters, a linear regression (single or multiple variable) can be formulated. Single variable regressions are useful when the outcome can be largely described by a single independent variable. Multivariable regressions are used when the results are not directly attributed to a single input. For example it might be reasonable to assume that the flow in a river is the direct result of the tributary flow if there is only one major tributary and no other sources of inflow. Now say there are two major tributaries, it might be more reasonable to suggest that the flow in the river is a function of the flows in both of the tributaries, rather than just one.

3.2.2. Polynomial Regression

Conceptually a polynomial regression is very similar to a multi-variable linear regression. The main difference is that in a polynomial regression the relationship between the independent and dependent variables include non-linear terms. The most general form of the polynomial regression equation is:

$$y = \text{Linear Terms} + 2^{\text{nd}} \text{ Degree Terms} + \dots + N^{\text{th}} \text{ Order Terms} + \text{Interaction Terms}$$

An example of a second-order, three variable poly regression equation is:

$$y = \beta_0 + \underbrace{\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3}_{\text{Linear _ Terms}} + \underbrace{\beta_4 x_1^2 + \beta_5 x_2^2 + \beta_6 x_3^2}_{\text{Second _ Degree _ Terms}} + \underbrace{\beta_7 x_1 x_2 + \beta_8 x_1 x_3 + \beta_9 x_2 x_3 + \beta_{10} x_1 x_2 x_3}_{\text{Interaction _ Terms}}$$

where ‘y’ is the dependent variable, β_i are the coefficients, and x_i are the independent variables. The polynomial above is of the order 2 ($m = 2$), with three independent variables (x_1 , x_2 , and x_3). The first and second groups of terms include the first order and second order terms, respectively. The third group of terms is the interactions between dependent variables. In this case there are two-way interactions and three-way interactions, because there are three dependent variables. If the polynomial relationship was of the third order then there would be an additional set of terms, representing third order terms and additional interaction terms.

The values for the β_i are solved for by minimizing the sum of the square errors in a similar fashion to linear regression. In essence:

$$\text{Min } \text{SSE} = \sum_{i=1}^N \varepsilon^2 = \sum_{i=1}^N (y_i - y)^2$$

where y_i is the observed value, y is the predicted value based on the regression relationship, and SSE is the sum of the square errors. It is generally recommended that m (the order of the polynomial) be much less than the number of observed data points ($m \ll N$).

3.2.3. Autoregressive Moving Average (ARMA) Models²

Autoregressive-moving-average (ARMA) models are mathematical models that are based on the persistence, or autocorrelation, of a time series. A time series is a continuous, generally fixed-interval set of observations such as daily rainfall or annual runoff. ARMA models are widely used in hydrology to create potential future time series.

In general, modeling can contribute to understanding the physical system by revealing attributes of the physical processes of interest. The reason for fitting an ARMA model to data is to take advantage of the persistence that may occur within a series. For example, a simple physical water-balance model consisting of terms for precipitation input, evaporation, infiltration, and groundwater storage can be shown to yield a stream flow series that follows a particular form of ARMA model. As such, ARMA models can also be used to predict the behavior of a time series from past values alone.

ARMA models can be described by a series of mathematical equations. One subset of ARMA models are the so-called autoregressive, or AR models. An AR model expresses a time series as a linear function of its past values plus a residual (noise term). The order of the AR model tells how many lagged past values are included. The simplest AR model is the first-order autoregressive, or $AR(1)$, model. The equation for this model is

$$y_t + a_1 y_{t-1} = \varepsilon_t$$

where y_t is the value of y at time t , y_{t-1} is the value of y at the previous time, a_1 is defined as the lag-1 autoregressive coefficient, and ε_t is the residual or error (presumed to be random in time, i.e., not correlated). The $AR(1)$ model has the form of a single value regression model where y is regressed versus its previous value, and ε_t is analogous to the regression residuals, or

$$y_t = -a_1 y_{t-1} + \varepsilon_t$$

The name *autoregressive* refers to the regression on self (auto). Higher-order autoregressive models include more lagged terms as predictors, e.g., the second-order autoregressive model, $AR(2)$, is given by

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} = \varepsilon_t$$

Another subset of the ARMA models are moving average (MA) models, where the time series is regarded as a moving average of random residuals. A first order moving average model, $MA(1)$, is given by

$$y_t = \varepsilon_t + c_1 \varepsilon_{t-1}$$

where ε_t and ε_{t-1} are residuals at time t and $t-1$, respectively, and c_1 is defined as the first order moving average coefficient. As with AR models, higher order MA models can be

² The ARMA discussion is drawn largely from Meko (2005).

implemented by using additional lagged residuals, e.g., a second order moving average model, $MA(2)$, is given by

$$y_t = e_t + c_1 \varepsilon_{t-1} + c_2 \varepsilon_{t-2}$$

The autoregressive model includes lagged terms on the time series itself, and the moving average model includes lagged terms on the residuals. By including both types of lagged terms, an autoregressive-moving-average, or ARMA, model can be developed. The order of the ARMA model is included in parentheses as $ARMA(p,q)$, where p is the autoregressive order and q the moving-average order. The simplest, and most frequently used ARMA model is $ARMA(1,1)$ model

$$y_t + a_1 y_{t-1} = \varepsilon_t + c_1 \varepsilon_{t-1}$$

Like in the AR and MA models, the residual terms (ε_t) must be an uncorrelated (independent) with a mean of zero. ARMA models also require stationary time series, i.e., those without trends, shift, or periodicity. Non-stationary signals can be removed from time series prior to ARMA modeling through a variety of statistical techniques. The reader is referred to Maidment (1993) for more details.

The process of constructing an ARMA model can generally be presented in three steps. The first step is to identify the model. Identification consists of specifying the appropriate structure (AR, MA or ARMA) and order of model. Identification is sometimes done by looking at plots of the autocorrelative function and partial autocorrelation function, and sometimes done by an automated iterative procedure such as fitting many different possible model structures and orders and using a goodness-of-fit statistic to select the best model (e.g., using a software package). The second step is to estimate the coefficients of the model. Coefficients of AR models can be estimated by least-squares regression. Estimation of parameters of MA and ARMA models usually requires a more complicated iteration procedure. In practice, estimation is accomplished automatically by a computer program with little user interaction. The third step is to check the model. This step is also called diagnostic checking or verification. Two important elements of checking are to ensure that the residuals of the model are random and to ensure that the estimated parameters are statistically significant. Usually the fitting process is guided by the principal of parsimony, by which the best model is the simplest possible model that adequately describes the data. The simplest model is the model with the fewest parameters.

3.3. Physically-Based (Theoretical) Models

In addition to the statistical models, where relations are derived based on the data provided, there are physically-based models. These models rely on theoretical principals to predict the state of a system given certain input conditions. These types of models can be very simple (e.g., spreadsheets) to very complex. A review of physically based models was preformed previously (Deas et al., 2004). Of the seven models reviewed, four are capable of estimating runoff from snowmelt and/or precipitation. However, prediction of runoff is not the primary function of those models and as such most require additional data (ranging from topography to measured water use and flow rates) to function in this role.

4. Scott River Watershed Data

4.1. *Characteristics of the Scott River Watershed*

The Scott River watershed, located in Siskiyou County of Northern California (Figure 2), encompasses approximately 425,000 acres of irrigated agriculture, wooded forest, urbanized lands, other lands, and waterways. The main economic industries in the watershed are irrigated agriculture (primarily alfalfa), timber, and cattle. Fort Jones and Etna are the main urbanized areas with populations of approximately 660 and 780, respectively, as of the 2000 Census (U.S. Census, 2003a; 2003b). Fort Jones and Etna are linked to one another and to the town of Yreka by Highway 3, which runs through the Scott Valley floor.

The Scott River watershed is dominated by a Mediterranean-type climate with summer being warm and dry and winter being cold and wet. Precipitation is mainly concentrated in the winter months and falls primarily as rainfall on the valley floor, while significant snowfall occurs on the surrounding mountain ranges, resulting in snowmelt runoff during the early spring months. The Scott Valley does not have any significant surface water storage facilities, but there is approximately 300 to 400 thousand acre-feet (taf) of groundwater storage, which is recharged annually by the Scott River, tributary streams, and by infiltration of precipitation and snowmelt (DWR, 2003).

The Scott River, a tributary to the Klamath River, is the major water resource feature in the basin. It is fed by a number of smaller tributaries, which include Kidder, Moffet, Mill, Sugar, and French Creeks. It is estimated that there are over 700 miles of streams within the basin. A river mile table depicting approximate miles from the confluence with the Klamath River for selected locations in the basin is shown below (Table 1). The two branches of the Scott River (the East Fork Scott River and the South Fork Scott River) merge at Callahan, in the southern portion of the watershed, approximately 56 miles from the confluence with the Klamath River.

Urban and agricultural water needs are met through diversions from surface water sources or from pumping of groundwater. The Scott River and tributaries also serves as habitat for coho salmon (*Oncorhynchus kisutch*) and fall run Chinook salmon (*Oncorhynchus tshawytscha*), as well as steelhead trout (*Oncorhynchus mykiss irideus*). Coho salmon are a federal and state protected species under the respective Endangered Species Acts. Runoff forecasting tools could assist local agricultural business enterprises, resource managers, other interested parties, and stakeholders to assess water availability in light of existing demands on the resource.

Table 1. River Mile (RM) Locations on the Scott River as Measured from the Confluence with the Klamath River.

Location	Approximate River Mile (RM)
Scott River at Mouth	0.0
Kelsey Creek	14.4
Canyon Creek	15.6
Scott River nr Ft. Jones (USGS Gage)	21.0
Mill Creek	23.9
Moffett Creek	31.3
Highway 3	31.9
Kidder Creek	32.3
Etna Creek	42.2
Youngs Dam	46.0
French Creek	47.6
Farmers Ditch	52.9
Sugar Creek	53.5
Wildcat Creek	55.3
East Fork/South Fork Scott River	56.1
Callahan	56.4

Source: Pacific Southwest Inter-agency Committee. 1973. *River Mile Index – Klamath River Basin, California and Oregon*. Report to the Water Management Technical Subcommittee. July.

4.2. Data Considered for Forecasting Models

For this scoping level assessment of runoff forecasting models, hydrological and meteorological data considered included stream flow, snow pillow, and rainfall data. Generally, state and federal agencies collect and disseminate such data regarding water resources throughout the state of California via published documents (such as technical reports and watershed investigations), web-based searchable databases, and from data archives. Data are also available from published and unpublished reports and databases, private entities, and other agencies and organizations.

There are 14 California Data Exchange Center (CDEC) stations located within the Scott River Watershed. Of the 14 CDEC stations, one reports stream flow, nine report snow data and five report precipitation (Figure 2). DWR has three flow stations in the Scott River Watershed. The various data types are outlined below.

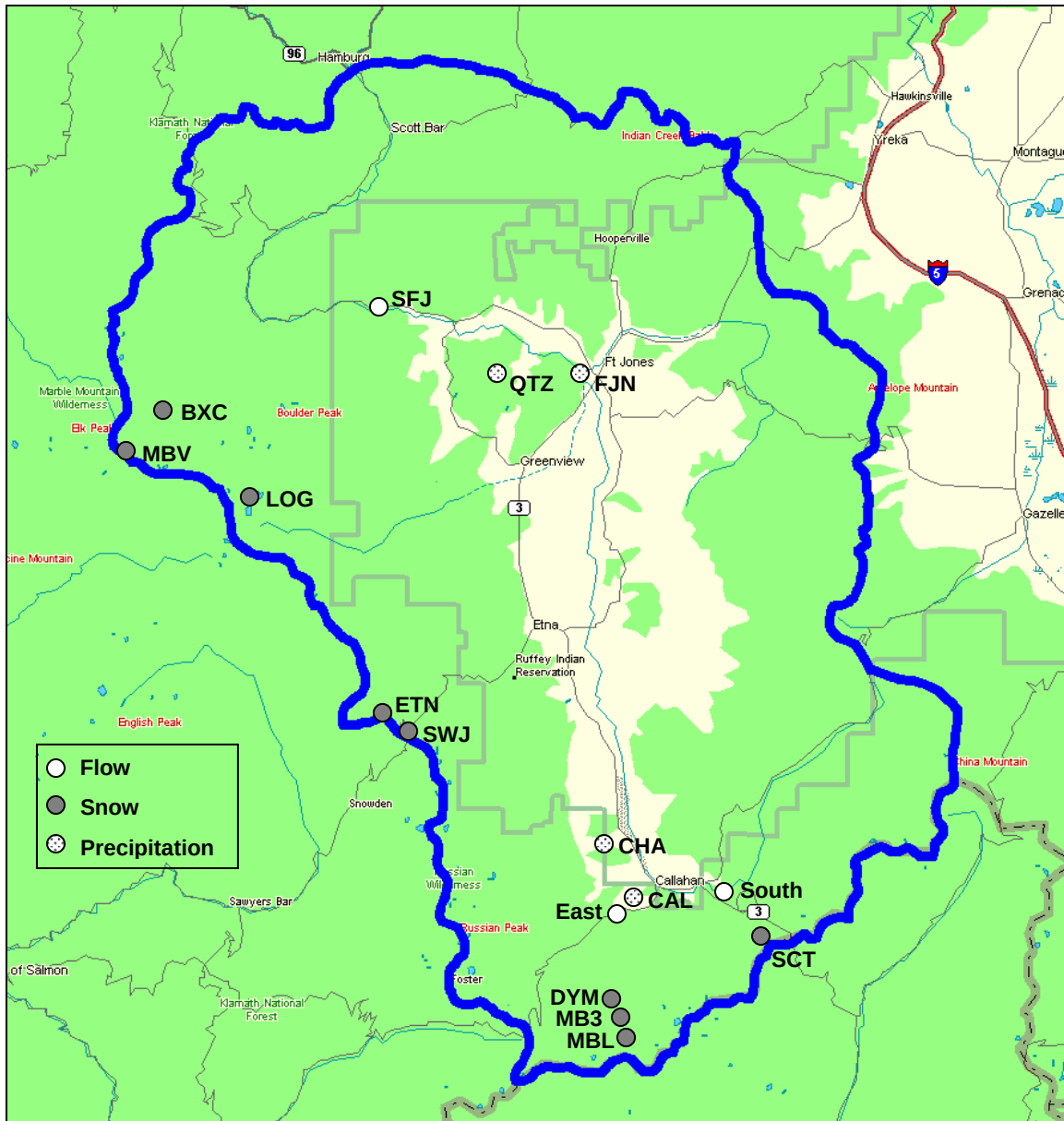


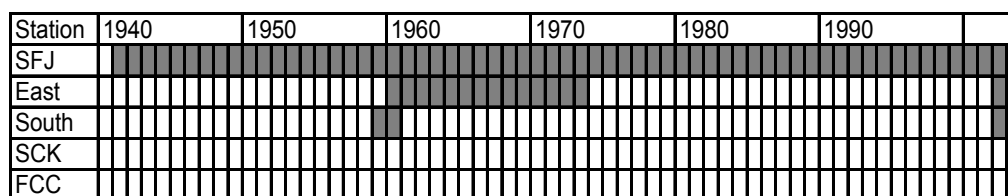
Figure 2. Data Stations in Scott River Watershed (Base Map from DeLorme® Topo Quads).

4.2.1. Stream Flow Data

The United States Geological Survey (USGS) reports the Scott River stream flow near Fort Jones (USGS 11519500) and makes that data available via CDEC and the USGS website. The most comprehensive stream flow record is for the Fort Jones gauging station. This gage is located at the downstream end of the valley, below the vast majority of water uses in the basin, and represents impaired runoff at this location. Stream flow for the East and South Forks of the Scott River are also available from USGS for a limited period of time from the late 1950's into the early 1970's. DWR maintained gages in the East and South Forks of the Scott River from June of 2002 through October of 2003 and intends to make the data available via the Water Data Library (WDL) in the future. DWR also reports a gage in the

Scott River at Scott Bar, but no additional information is available via the WDL. Figure 3 summarizes the periods of record for each station.

DWR has an unpublished estimate of unimpaired stream flow for the Scott River for the period of October 1945 to September 1994 (DWR, 1998). The available flow data for tributaries to the Scott River, Shackleford Creek, and French Creek, is limited to recent years and only stage is available from CDEC. It has also been indicated that flow data exists for Mill, Kidder, Sugar, and Etna Creeks, but this data is not available via CDEC or the WDL.



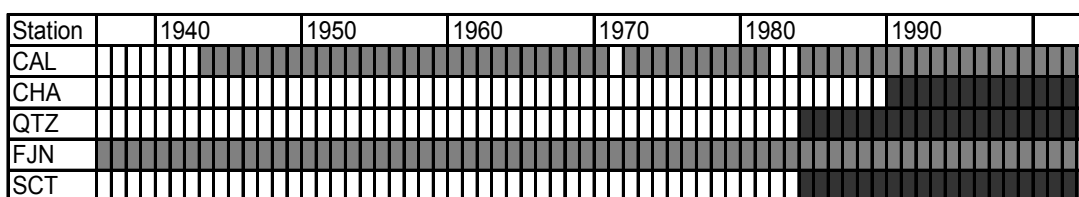
SFJ = Scott R. nr Fort Jones
East = East Fork Scott R nr Callahan
South = South Fork Scott R nr Callahan
SCK = Shackleford Ck nr Mugginsville
FCC = French Ck nr Callahan

= Continuous (or Nearly Continuous) Data

Figure 3. Stream Flow Data, Period of Record.

4.2.2. Rainfall Data

In total there are five rainfall stations in the Scott Valley, three of which are reporting at an hourly frequency (Table 2 and Figure 4). The Callahan and Fort Jones rainfall stations extend back to the early 1940's and late 1930's, respectively. The hourly records at Callahan, Quartz Valley, and Scott Mountain start in the mid- to late-1980's and continue to the present.



CAL - Callahan Station
CHA - Callahan Station
QTZ - Quartz Valley Station
FJN - Fort Jones RS Station
SCT - Scott Mountain Station

= Monthly Data
 = Hourly Data

Figure 4. Rainfall Stations in the Scott Valley, Period of Record.

Table 2. Latitude, Longitude, Elevation and Operator for CDEC Rainfall Stations.

Station ID	Station Name	Latitude	Longitude	Elevation	Operator ¹
CAL	Callahan	41.370°N	122.8000°W	3185'	NWS
CHA	Callahan	41.3330°N	122.8330°W	3136'	USFS
QTZ	Quartz Valley	41.6000°N	122.9290°W	4238'	CADF
FJN	Fort Jones RS	41.6000°N	122.8500°W	2725'	USFS
SCT	Scott Mountain	41.2720°N	122.7180°W	5900'	USBR

¹ 'NWS' = National Weather Service, 'USFS' = US Forest Service, 'CADF' = CA Department of Forestry, 'USBR' = US Bureau of Reclamation

4.2.3. Snow Fall

Snow fall as represented by water content data for the Scott River watershed are available from CDEC for several locations (Figure 5). Although snow depth information is also available, from a water supply perspective the water content is the more valuable information for this exercise. Data are generally reported monthly, but varies with time period and station. Generally, the April 1st snow water content is used as representative of the spring and summer snow pack yield (e.g., runoff). The CDEC snow stations can be grouped into three general sets based on proximity to one another. BXC, MBV, and LOG are all in the North-West portion of the Scott Valley (MBV and LOG are no longer in service), ETN and SWJ are in the West, and DYM, MBL, MB3, and SCT are in the South. The latitude, longitude, elevation, and operator are presented below for each station (Table 3).

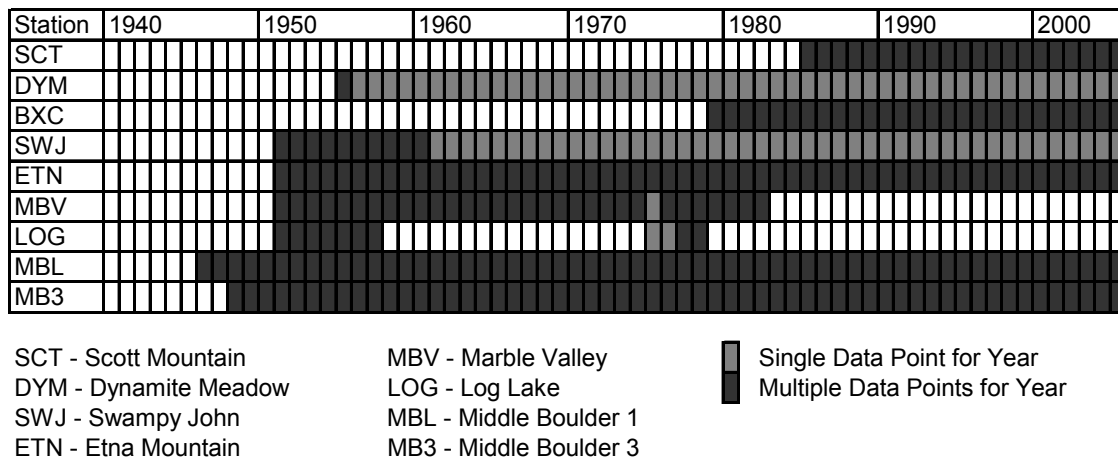


Figure 5. CDEC Adjusted Snow Water Content and Snow Depth Data, Period of Record.

Table 3. Latitude, Longitude, Elevation and Operator for CDEC Snow Stations.

Station ID	Station Name	Latitude	Longitude	Elevation	Operator ¹
BXC	Box Camp	41.5970°N	123.1650°W	6450'	SR Ranger District
MBV	Marble Valley	41.5670°N	123.1980°W	5900'	(Discontinued)
LOG	Log Lake	41.5470°N	123.1000°W	5300'	(Discontinued)
ETN	Etna Mountain	41.4000°N	123.0010°W	5900'	SR Ranger District
SWJ	Swampy Johns	41.3970°N	122.9930°W	5500'	SR Ranger District
DYM	Dynamite Meadow	41.2330°N	122.8170°W	5700'	SR Ranger District
MB3	Middle Boulder 3	41.2250°N	122.8110°W	6200'	USBR
MBL	Middle Boulder 1	41.2170°N	122.8070°W	6600'	SR Ranger District
SCT	Scott Mountain	41.2720°N	122.7180°W	5900'	USBR

¹ 'SR Range District' = Scott River Ranger District, 'USBR' = US Bureau of Reclamation

5. Scott River Precipitation Characteristics

Precipitation as rainfall and snowfall (water content and depth) data are available at several locations from CDEC (Table 2 and Table 3). The availability and quality of the data varies among stations.

5.1. Rainfall

Callahan (CLA) and Fort Jones (FJN) are the two stations where monthly accumulated rainfall records exist for a prolonged period of time. The CAL station is in the southern portion of the valley, while the FJN station is in the middle. The two stations are at relatively similar elevations (there is about a 450' feet difference) and the accumulated rainfall pattern is relatively similar (Figure 6).

A non-exceedence plot can be a useful method of depicting distributions. Although there are several methods of determining probabilities, the Weibull distribution is a common approach (Stedinger, 1993) wherein the data are ranked in ascending order and plotting positions determined using the following formula:

$$\text{Probability of Non-Exceedence} = \frac{\text{Rank}}{n + 1}$$

Where 'n' is the number of points in the data series. By definitions, non-exceedence probabilities cannot be zero or one-hundred.

The non-exceedence plots for Fort Jones and Callahan October through March³ precipitation are similar except under very low and high annual rainfall conditions (Figure 7). In this case, the lower and upper limits are about 2 percent and 98 percent, respectively. Further, for any given October through March period there is a 25 percent probability that precipitation will be less than approximately 12 inches (based on the available record). Conversely, it can be

³ The period of precipitation typically associated with the subsequent April through September runoff.

said that for any given October through March period the precipitation will there is a 75 percent probability that precipitation will be greater than approximately 12 inches. Non-exceedence plots can be used to determine the probability of any October to March precipitation within the historically observed range. Other metrics, such as annual precipitation or monthly precipitation, could also be analyzed.

Likewise, there is a fairly strong relationship between the annual accumulated rainfall at FJN and the annual accumulated precipitation at CAL (Figure 8). A linear regression equation between the two yielded an R^2 value 0.836 ($FJN_t = 1.0376CAL_t + 0.1539$). This would indicate that at the lower elevations, rainfall is fairly consistent in the middle and southern portion of the valley.

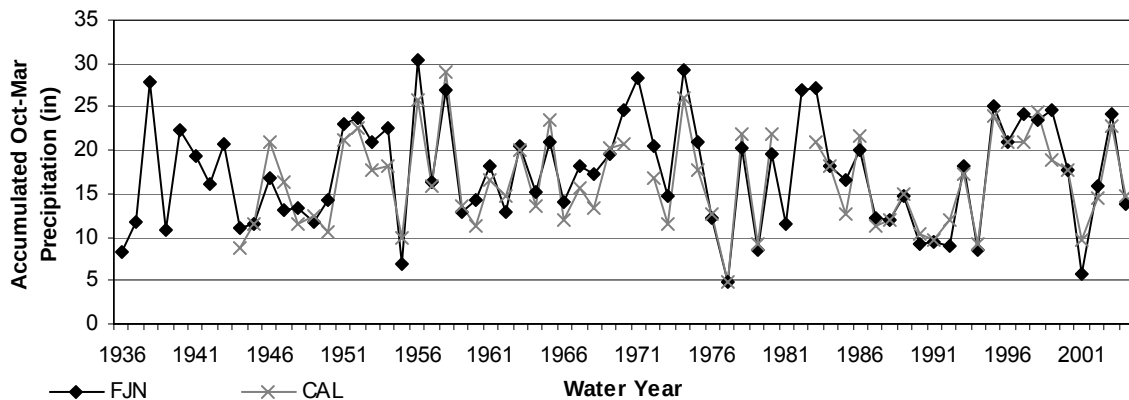


Figure 6. October Through March Accumulated Rainfall.

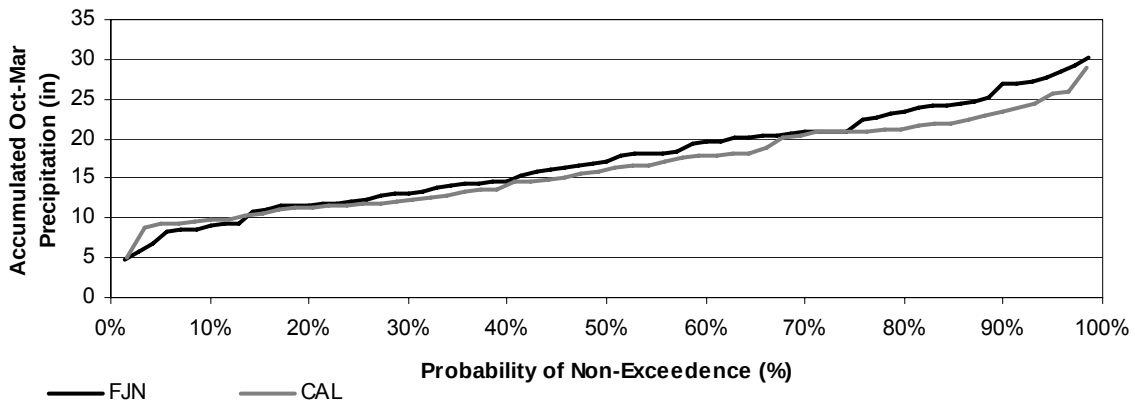


Figure 7. Probability of Non-Exceedence for CAL and FJN Rainfall.

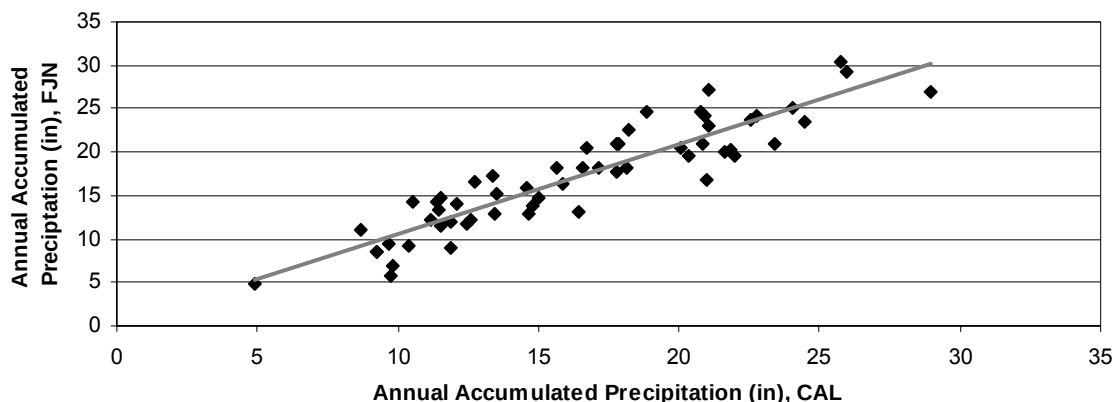


Figure 8. Annual Accumulated Rainfall at CAL versus FJN.

5.2. Snow Water Content

The nine CDEC snow stations were compared with each other to identify similarities in measured snow water content. Each station was compared against the other eight and the R^2 values are reported below (Table 4). As noted earlier the stations were roughly grouped into three sets: 1) BXC, MBV, and LOG, 2) ETN and SWJ, and 3) DYM, MB3, MBL, and SCT. The period of record for BXC did not overlap with MBV or LOG for a long enough period to determine how well the three stations relate to each other. Examining the R^2 values relating the various stations (Table 4) suggests that the five stations on the western side of the valley (BXC, MBV, LOG, ETN, and SWJ) relate relatively well ($R^2 > 0.820$). The ETN and SWJ stations related quite strongly (R^2 of 0.958). Likewise, the four stations in the southern portion of the valley also related fairly well to each other ($R^2 > 0.810$). On the other hand, the stations in the west and in the south did not indicate such strong relationships relating one to another (R^2 values ranged from 0.788 to 0.597, LOG did not have sufficient data to perform a comparably meaningful regression). In sum, this information suggests that conditions vary spatially among snow pillow data, and more detailed analyses could be carried out.

Of the nine snow stations listed in CDEC, five of them contain data for more than just April 1st. SWJ, DYM, MB3, MBL and SCT have readings taken in January, February, March, April, May, and June, or some combination thereof. From those additional data it is possible to determine the general state of the snow pack on April 1st. The status of the snow pack on April 1st can be described by four general states:

- Case 1: peaking (annual maximum),
- Case 2: increasing (continued accumulation),
- Case 3: decreasing (melting), or
- Case 4: re-accumulating (accumulating after partial melting).

Using April 1st as a benchmark these four cases are illustrated in Figure 9.

Table 4. R^2 for April 1 Snow Station Relationships.

	R^2 Value								
	BXC	MBV	LOG ¹	ETN	SWJ	DYM	MB3	MBL	SCT
BXC	1								
MBV	⁻²	1							
LOG	⁻³	0.971	1						
ETN	0.826	0.885	0.961	1					
SWJ	0.823	0.886	0.971	0.958	1				
DYM	0.631	0.641	0.799	0.689	0.703	1			
MB3	0.756	0.788	0.847	0.656	0.682	0.925	1		
MBL	0.776	0.771	0.718	0.597	0.638	0.819	0.938	1	
SCT	0.610	⁻³	⁻³	0.645	0.590	0.930	0.942	0.830	1

¹ There were only twelve years of data at the LOG station.

² There is insufficient data to do a comparison ($n < 5$).

³ There are no periods of overlapping data ($n = 0$).

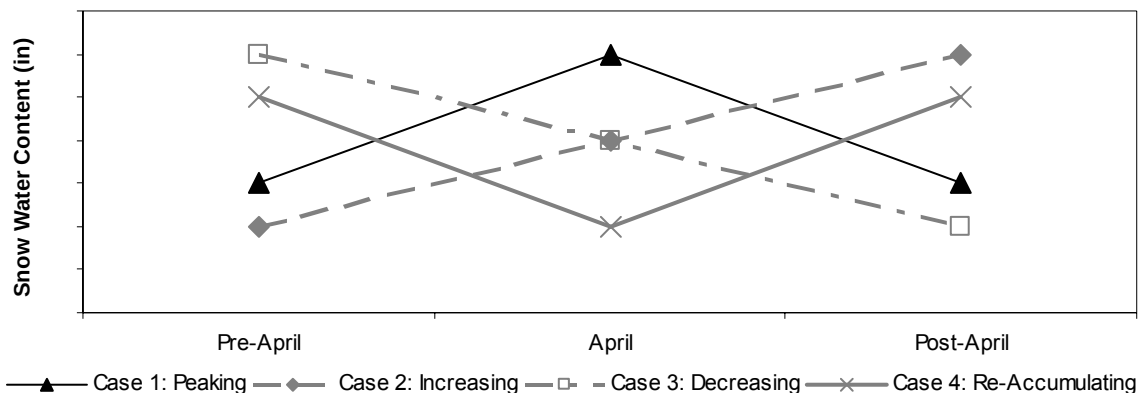


Figure 9. Example for Peaking, Increasing, Decreasing, and Re-accumulating Snow Water Content Condition.

Ideally the April 1st snow water content reading would be peaking. If the April 1st reading indicates that the snow pack is still increasing, then additional potential supply (in the form of snow) is still accumulating. Use of this value for forecasting might under-predict the potential runoff. If the value is decreasing, then runoff has already begun and using an April 1st forecast might underestimate runoff. Similarly, re-accumulation indicates that runoff had begun, but then additional snow fell after April 1st, and a forecast based on April 1st water content may underestimate runoff.

For the five stations where pre-April, post-April, and April 1st snow water content readings are available, there was considerable variability in the state of the snow pack on April 1st (Table 5). For SWJ, MB3 and MBL, the April 1st snow water content is generally peaking.

For DYM and SCT the April 1st snow water content is generally decreasing. This analysis indicates that variability in snow fall accumulations and subsequent snowmelt runoff can introduce uncertainty into forecast models based on such data.

Table 5. State of the April 1st Snow Pack.

Station	Number of Years				N/A ¹	Total
	Peaking	Decreasing	Increasing	Re-Accumulating		
SWJ	22	12	9	2	10	55
DYM	16	26	5	3	1	51
MB3	34	15	6	1	2	58
MBL	31	10	13	3	3	60
SCT	7	8	1	4	0	20

¹ Insufficient data to determine the state of the April 1st snow pack.

6. Scott River Flow Characteristics

To gain insight into the runoff characteristics of the Scott River and the annual variability in flow conditions, monthly flow distribution, unimpaired flow conditions, and annual probability of non-exceedence were assessed. The monthly flow distribution lends insight into the timing of runoff. Unimpaired flow conditions, flow without consumptive uses, indicate the total volume of runoff. The annual probability of non-exceedence provided insight into annual variability in annual runoff.

6.1. Monthly Flow Distribution

The monthly Scott River flow is highly variable through the year. The available data from three gages was used to compare monthly average flows in the Scott River basin: East Fork near Callahan, South Fork near Callahan, and Scott River near Fort Jones. The average monthly distribution of the annual total stream flow indicates that flow generally peaks in the winter and spring monthly and decreases throughout the summer. Scott River flow at the Fort Jones gage records impaired runoff due to upstream diversions. Both the East and South Fork also have diversions upstream of the gages⁴. Review of Figure 3 shows that the period of record varies widely among the three stations, with the South Fork capturing only two years of data. (The tributary data are not used due to the short flow record. In addition, unimpaired flows from DWR (1998) were only available on an annual basis and could not be compared in this monthly assessment. Flow data is also available from DWR for June of 2002 through October of 2003 for the East and South Forks of the Scott River. This data was of limited duration and thus not included herein.)

All three gages show generally similar annual runoff distributions (Figure 10). The East Fork and Scott River at Fort Jones (impaired) records indicate a dual peak hydrograph. The

⁴ The East and South Forks of the Scott River are adjudicated (Siskiyou County Superior Court Decree #30662). Diversions occur primarily in the irrigation season (April through October). The East Fork, while adjudicated, does not have a State or private watermaster service (John Clements, pers. comm.).

first peak occurs in the early winter (January) and the second in early spring (May), representing a typical combined rainfall and snowmelt hydrograph. Following spring runoff there is a steady reduction in flow through the summer. Flows begin to increase again in the fall and winter with reduced evapotranspiration and increased seasonal precipitation. The spring snowmelt peak on the impaired Fort Jones record is most likely reduced from an unimpaired condition. The South Fork shows a fairly similar pattern to the East Fork, except for the lack of a distinct peak in January, suggesting more of a snowmelt influence than rainfall; however, it is difficult to draw strong collusions from this short data set.

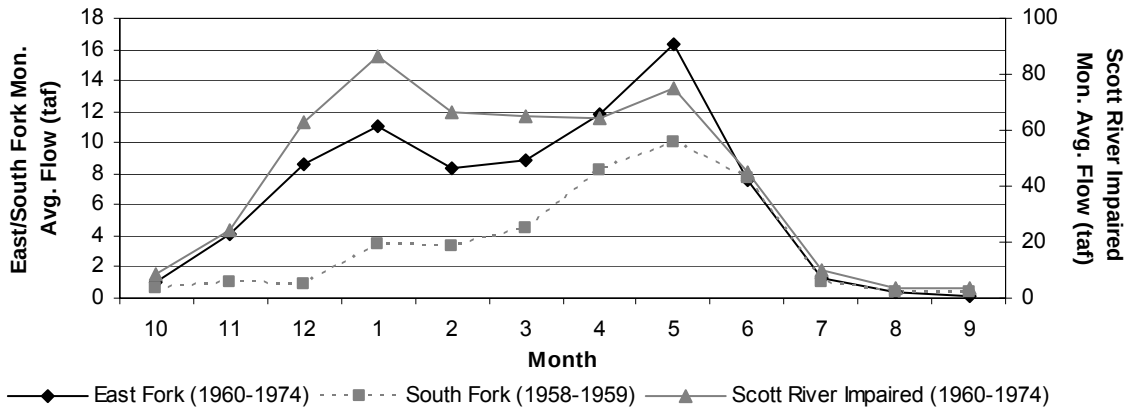


Figure 10. Monthly Average Impaired Flows (taf) from WY1960-1974.

6.2. Annual Scott River Unimpaired Runoff Estimates

In addition to the stream flow records identified previously, annual unimpaired flow estimates for the Scott River at Fort Jones were available from DWR (1998). This unimpaired flow study includes a time series of flow created using agricultural and urban water use and the impaired gauged flow at Fort Jones that spans from water year (WY: October 1 through September 30) 1945 to WY1994. For comparison, the impaired flow at the Fort Jones gage available on CDEC spans from WY1942 to WY2004. The methodology that DWR used to estimate the WY1945 through WY1994 unimpaired flow was not included in the memo (DWR, 1998) and thus no attempt was made to extend the DWR unimpaired flow study from 1994 to present.

The maximum, average, and minimum annual unimpaired and impaired flow records for the Scott River at Fort Jones are compared for a 50-year period (WY1945-94) (Table 6 and Figure 11). The highest total annual unimpaired flow during WY1945 to WY1994 was 1,150 thousand acre-feet per year (taf/yr) (in WY1974) and the lowest total annual flow was 123 taf/yr (in WY1977). In comparing the 63-year impaired hydrology with the 50-year impaired hydrology the highest and lowest flows remained unchanged, i.e., the extreme events occur prior to 1993. The annual average flow is slightly higher for the longer data series: 453 taf/yr compared to 461 taf/yr. Also included in Figure 11 is the difference between the impaired and unimpaired flows at Fort Jones. This difference represents the consumptive use and reaches a stable level at just over 70 taf/yr by the early 1990's. Consumptive use is discussed in more detail below.

Table 6. Maximum, Minimum, and Average Flows in the Scott River at Fort Jones.

Statistic	DWR Data (WY1945 – WY1994)		CDEC (WY1942-WY2004)
	Unimpaired (taf/yr)	Impaired (taf/yr)	Impaired (taf/yr)
Maximum	1149.9	1082.9	1082.9
Minimum	123.0	54.2	54.2
Average	513.6	452.7	461.1

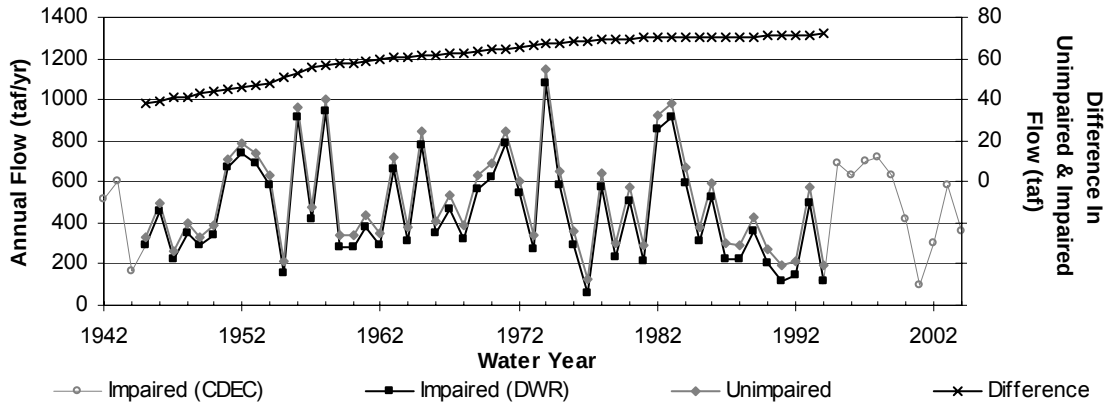


Figure 11. Impaired and Unimpaired Flow at Fort Jones.

As noted above, the difference between the unimpaired and impaired flows is due to the consumptive water use(s) in the Scott Valley (Figure 11). During the WY1945-94 period the annual consumptive water use ranged from 38.0 taf to 71.8 taf, with a 50-year average of 60.9 taf. Comparing annual consumptive use and the annual measured (impaired) flow suggest that there is no clear trend to the data (Figure 12).

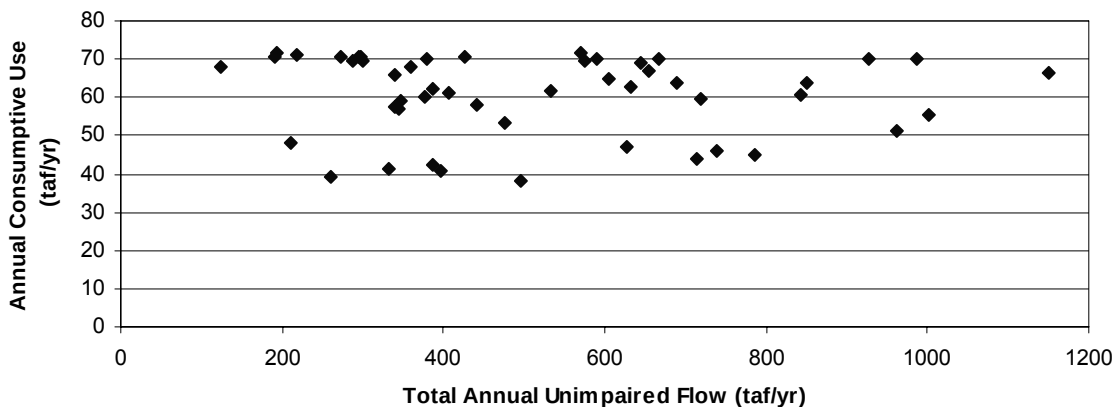


Figure 12. Total Impaired Flow versus Annual Consumptive Use.

6.3. Annual Non-Exceedence Plots

Non-exceedence plots for impaired and unimpaired Scott River flow at Fort Jones were formulated based on data from WY1945-94 (Figure 13). A non-exceedence plot for annual

runoff represents the probability that for any given year the flow will be less than a selected volume.

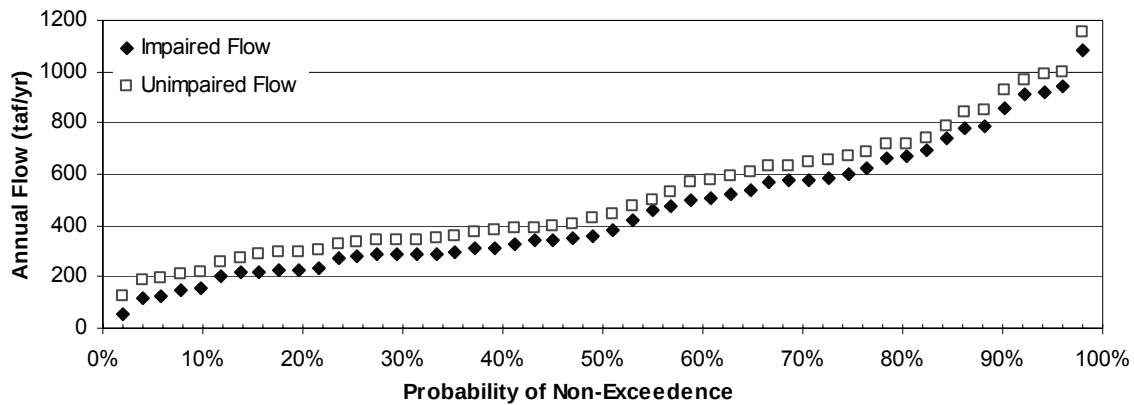


Figure 13. Probability of Non-Exceedence.

From Figure 14 the 50 percent non-exceedence annual unimpaired flow at Fort Jones is 434.3 taf/yr. The lower limit (2 percent non-exceedence) is 123 taf/yr and the upper limit (98 percent non-exceedence) is 1,150 taf/yr. The non-exceedence plot highlights an additional component of flow characteristics in the Scott River at Fort Jones. From Figure 14 it can be seen that the non-exceedence plot is relatively 'flat' for values between 200 and 400 taf/yr, compared to the range above 400 taf. It is postulated that relatively low variability below the 50 percent non-exceedence value is due to the presence of a base flow controlled by the groundwater inflows. As mentioned earlier, the Scott Valley sits atop a large alluvial groundwater basin. Base flow in the river may be sustained over a period greater than one year due to groundwater influences. Thus, the lower range may reflect the stable influences of groundwater, while the upper range represents the more variable nature of larger annual runoff events.

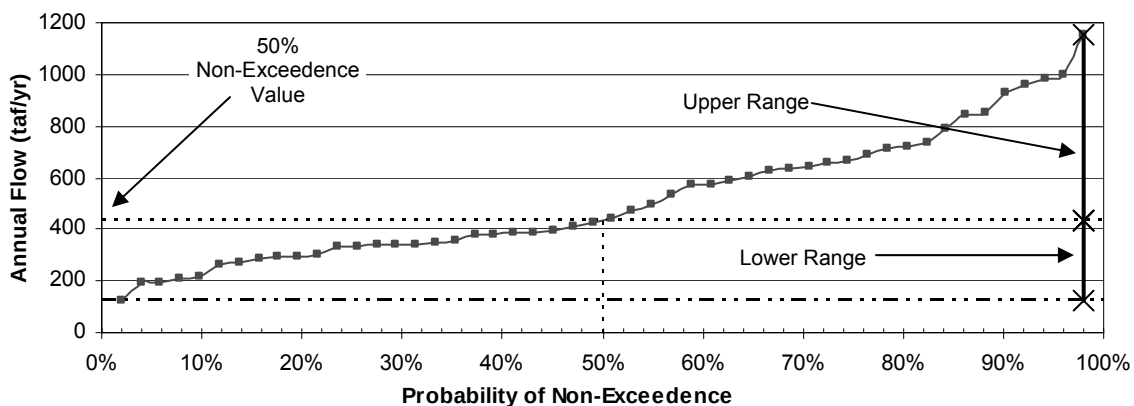


Figure 14. Probability of Non-Exceedence for Unimpaired Flows for the Scott River at Fort Jones with Upper and Lower Limits.

When reporting statistical summaries of historical flow records a common value to report is the average annual flow observed during the period of record. For the Scott River, the average unimpaired annual flow was 514 taf. The issue arises from the fact that the average

can be skewed by extreme events. The exceedence plot presented above highlights that despite the fact the average was 514 taf/yr, the median or 50% value was closer to 434 taf/yr. Exceedence plots (like the one in Figure 14) are useful, but not the most intuitive. A simpler means of highlighting the difference between the average and median flow values is presented by analyzing the frequency of certain ranges of flow (herein referred to as “binning”).

To identify the frequency distribution of flows of various magnitudes, the annual unimpaired total flows were binned into groups of 100 taf (i.e., the first group counted the number of year for which the unimpaired river flows was between 0 taf and 100 taf, the second group counted flows from over 100 taf to 200 taf, the third group counted flows from over 200 taf to 300 taf, etc.). Grouping the data in 100 taf bins indicated that there is a multi-modal distribution (Figure 15). It can be seen that the majority of the annual flows were in the 300 to 400 taf bin.

A second binning was done using increments of 200 taf. The 200 taf binning smoothed out the multiple small peaks in the 100 taf case, giving the impression of a bi-modal distribution (Figure 16). Finally, the 250 taf binning smooths out all but the main peak, highlighting the skewed distribution of annual flows (Figure 17). From Figure 17 it is clear that the most common annual unimpaired flow rate is in the 250 to 500 taf bin, which is below the long-term average of 514 taf/yr. The low variability in annual flow below the 50 percent exceedence (flatness of the exceedence curve) and the skewed distribution of the annual flows suggests that the system has a base flow component that may be influenced by groundwater contributions, extreme flow events, and/or other factors.

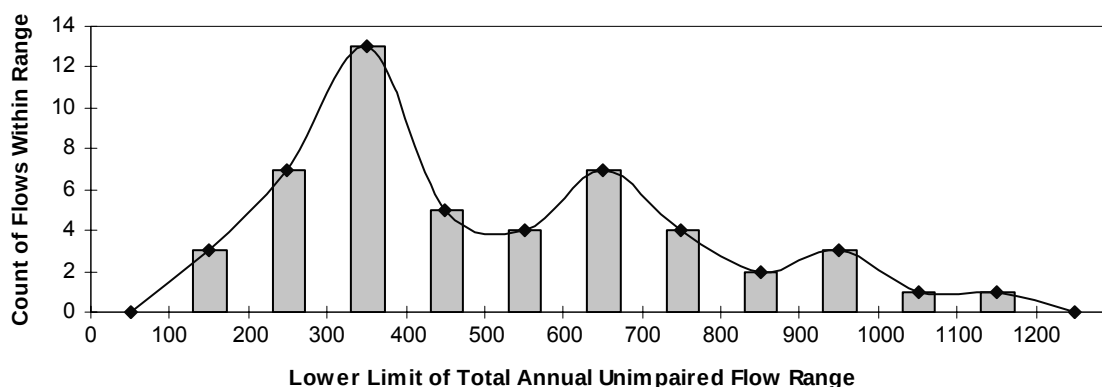


Figure 15. Distribution of Unimpaired Annual Flows by Groups of 100 taf/yr.

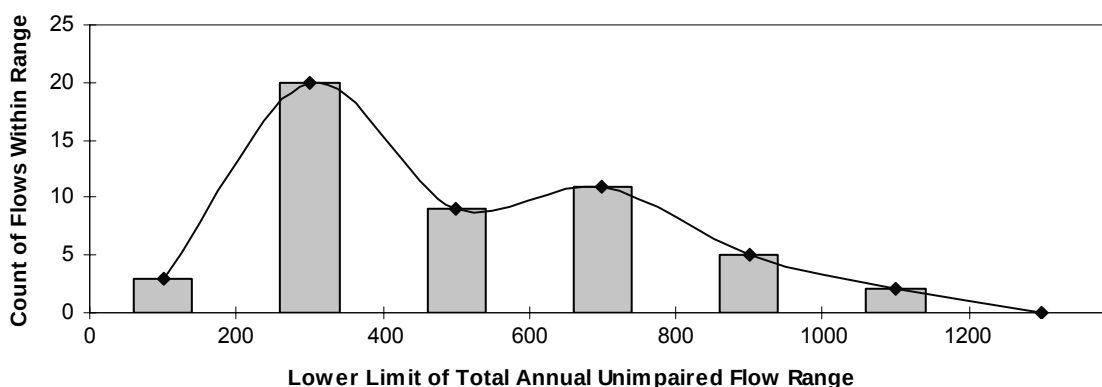


Figure 16. Distribution of Unimpaired Annual Flows by Groups of 200 taf/yr.

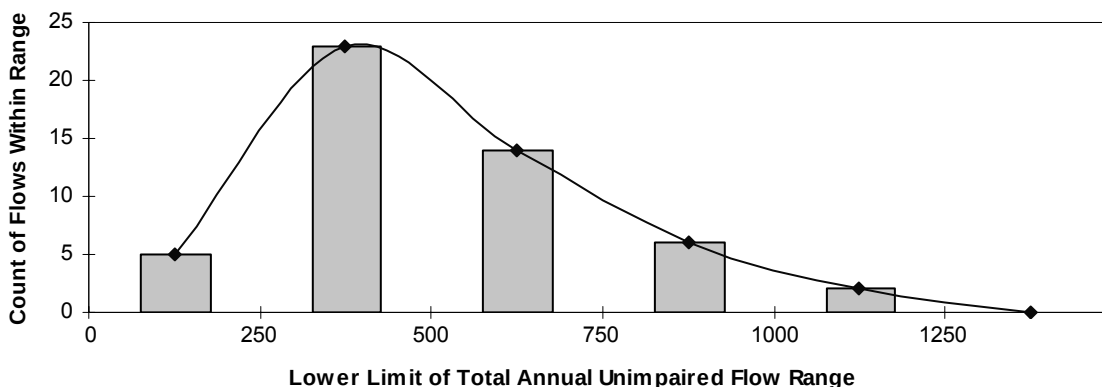


Figure 17. Distribution of Unimpaired Annual Flows by Groups of 250 taf/yr.

7. Runoff Forecast Modeling

Runoff is influenced by multiple factors including, rainfall, snow pack (water content), meteorological conditions, soil conditions, natural and man-made storage, vegetation, other water resources development, as well as other factors. For this analyses, a limited number of independent variables are examined, namely rainfall and snow water content, to forecast unimpaired flow. Although identified as a potential contributor, groundwater was not included in this effort due to limited time and resources. Outlined below are relationships that explore rainfall and snow water content as dependent variables. In the Scott River watershed there are nine snow and five rainfall records. An assessment of non-stationary time series was not completed, thus impacts of global warming, modified cropping patterns, modification of groundwater practices, or other long-term changes are not considered.

7.1. *Runoff Forecast as a Function of April 1st Snow Water Content*

7.1.1. Single Variable Snow Pillow Linear Regression (Full Record)

The April 1 snow water content values were compared with the unimpaired flow rates. The unimpaired hydrology exists from WY1945 to WY1994, while the nine CDEC stations cover

a variety of periods (the longest record being from Middle Boulder 1 (MBL)). Constructing a single variable linear regression for each of the nine stations indicates that BXC correlates to a higher degree ($R^2 = 0.7888$) than other stations by a considerable margin (Table 7).

Table 7. Results of a Single Variable Linear Regression Model for Annual Unimpaired Scott River at Fort Jones Flow as a Function of Snow Water Content at Local Snow Pillow Locations.

Station	WY	N	R^2 *
BXC	1979 – 1994	16	0.7888
MBV	1951 – 1982	32	0.2899
LOG	1951 – 1958; 1975 – 1978	12	0.4078
ETN	1951 – 1994	44	0.3091
SWJ	1951 – 1994	44	0.3565
DYM	1956 – 1994	39	0.1086
MB3	1948 – 1994	47	0.2534
MBL	1946 – 1994	49	0.3485
SCT	1986 – 1994	9	0.1517

* The analysis was completed using Microsoft Excel Data Analysis Regression Tool.

The linear regression equation relating snow water content at BXC to unimpaired runoff at the Fort Jones gage site is:

$$Y_t = 12.03BXC_t + 35.87$$

where BXC_t is April 1st snow water content (inches) and Y_t is annual unimpaired runoff (taf/yr) at Fort Jones. Although the BXC station had the highest R^2 value, the standard error of the constant was 64.320 (1.664 for the coefficient): nearly twice the magnitude of the constant value – an undesirable trait for a regression relationship.

Using the linear regression equation to calculate the predicted total annual unimpaired flow in the Scott River produced mixed results. The regression under-estimated the annual unimpaired flow 10 of the 16 periods and over-estimated it the remaining 6 (Figure 18). Of the under-estimates, the regression under predicted by as much as 188 taf/yr. For the over-estimates, it over predicted by as much as 193 taf/yr (Figure 19). While this model could be useful as a first-level screening method, the results deviate by as much as 65% when compared to historic data (Figure 20). Additionally, forecasting based on a single snow water content station can run the risk being subject to uncertainties due to spatial variation in snow fall conditions (as noted above), and in rare instances the potential for a zero snow water content value on April 1st (i.e., no snow).

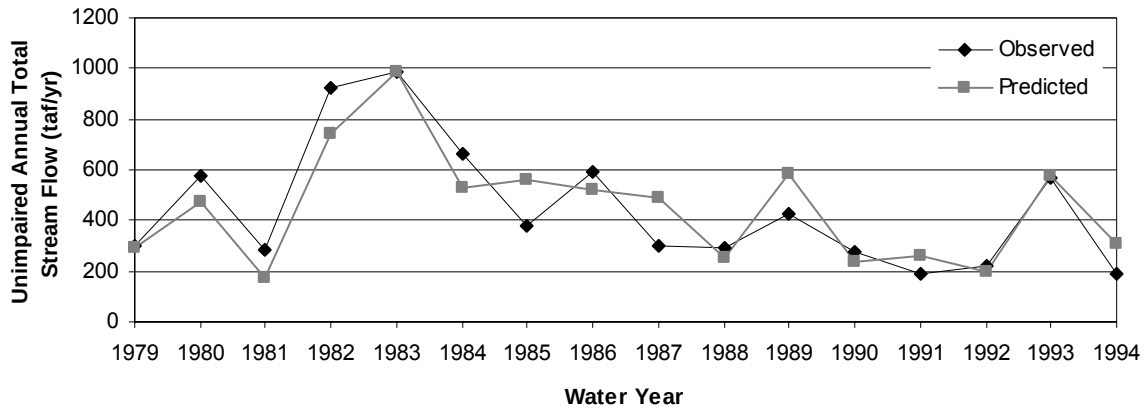


Figure 18. Observed and Predicted Unimpaired Flows Using BXC Linear Regression.

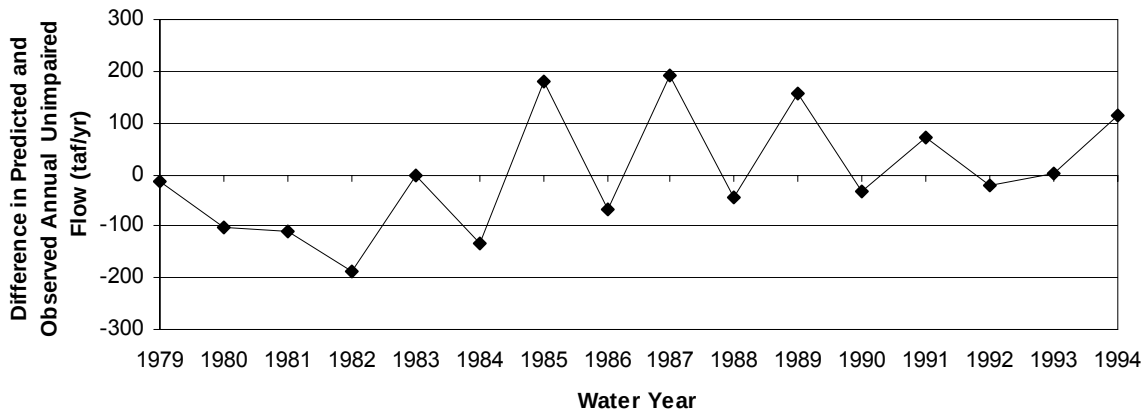


Figure 19. Differences in Predicted and Observed Unimpaired Flow Using BXC Linear Regression.

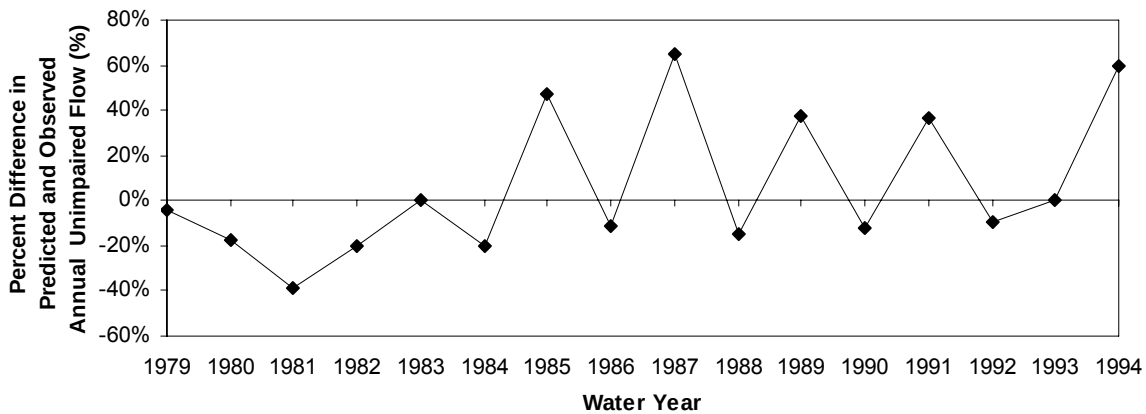


Figure 20. Percent Differences in Predicted and Observed Unimpaired Flow Using BXC Linear Regression.

The results from the BXC Linear Regression have a wide range of variability and unequal distribution. Based on the original data it can be seen that the April 1st snow water content readings fall into essential two clusters, plus two outliers (Figure 21). The reasons behind this unique distribution in the data are not understood, but may be the reason the standard

error of the constant is so large. While the BXC relationship might be a means to estimate the runoff volume, it is not necessarily recommended unless the forecaster is sufficiently aware of the limitations present.

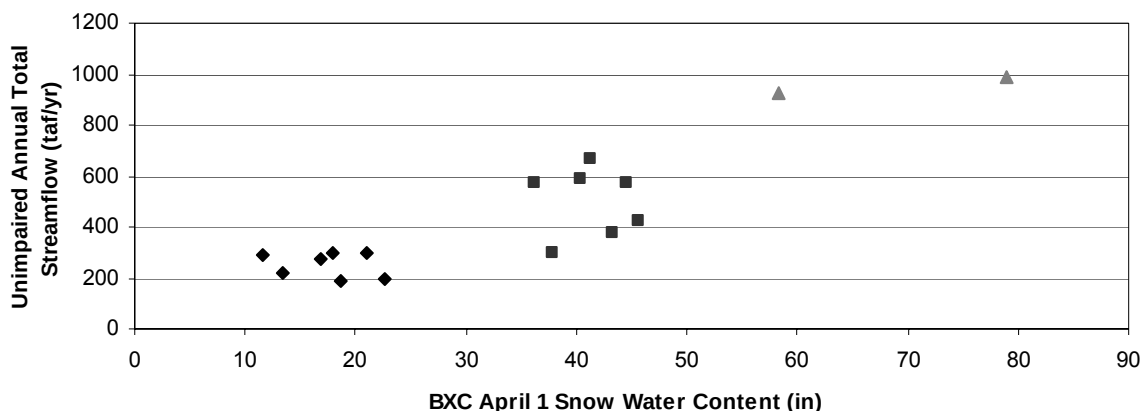


Figure 21. Clustered BXC April 1st Snow Water Content Data.

7.1.2. Single Variable Linear Regression (By State of Snow Pack)

As mentioned earlier, the April 1st snow water content should ideally be the annual peak, but this is not guaranteed due to meteorological variability. A preliminary analysis comparing the different states of the snow pack (peaking, increasing, and decreasing (re-accumulation was not assessed) was completed based on the antecedent April 1st snow water content conditions. The April 1st snow water contents within those categories were plotted against the associated unimpaired annual stream flow for the Scott River at Fort Jones to assess any potential correlation. None of the three cases appeared to indicate a consistently strong relationship from the four stations (SCT was omitted because it had only 7 data points) where sufficient data existed, i.e., there is appreciable scatter among the data as indicated in Figure 22 through Figure 25. Regression analyses of these data sets supported this observation. The lack of strong correlation may be due in part to the limited amount of data for certain cases. For example, DYM has 26 data points for decreasing snow pack, 5 for increasing snow pack, and 16 for snow pack at the peak. However, another reason is that there are other components or interactions among components that are not included in the simple linear regression that are important to runoff prediction.

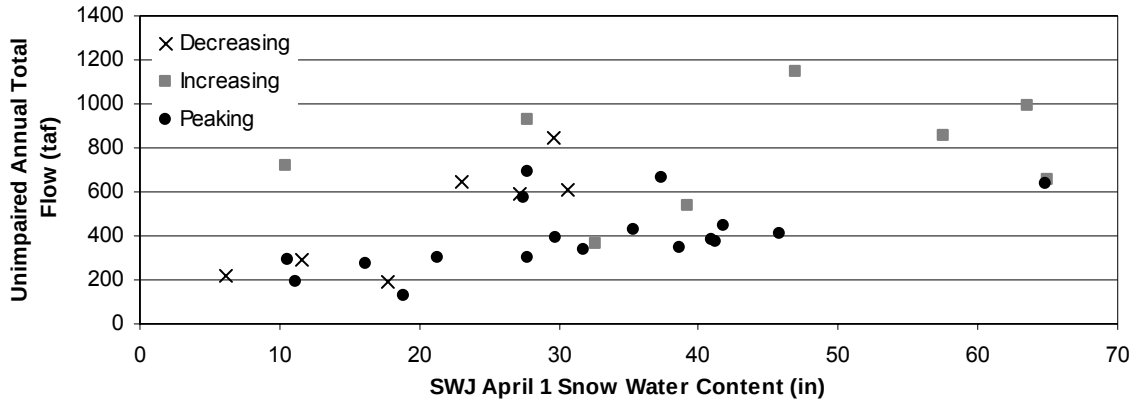


Figure 22. SWJ April 1st Snow Water Content by State of Snow Pack.

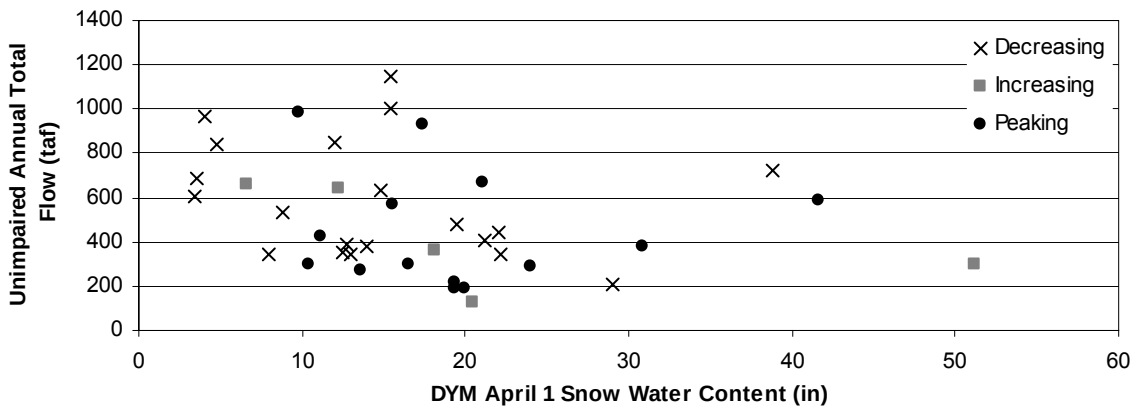


Figure 23. DYM April 1st Snow Water Content by State of Snow Pack.

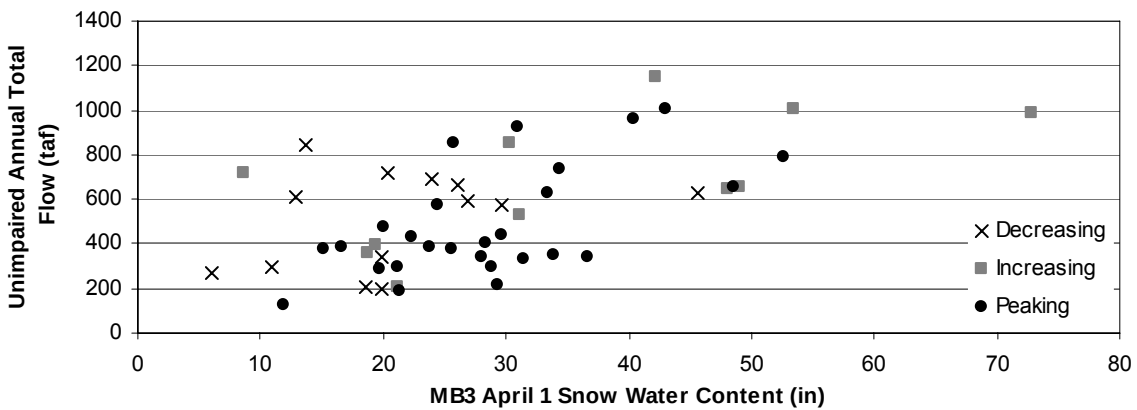


Figure 24. MB3 April 1st Snow Water Content by State of Snow Pack.

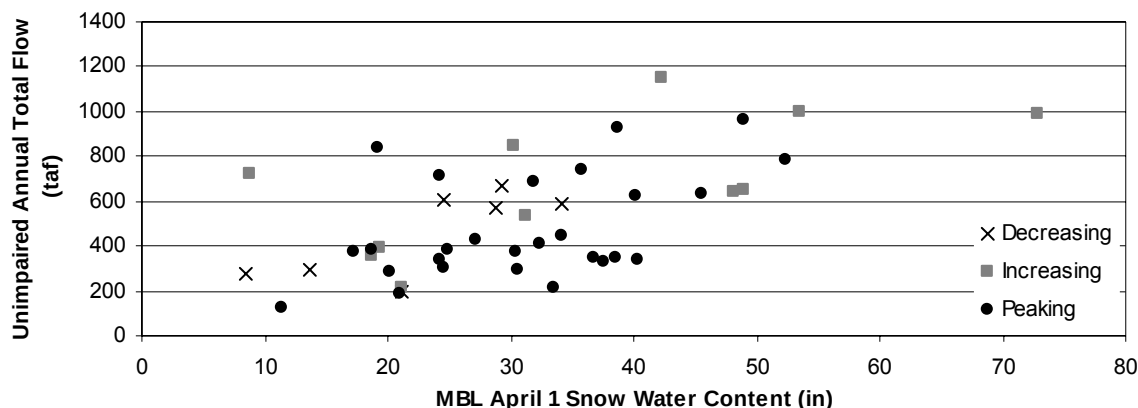


Figure 25. MBL April 1st Snow Water Content by State of Snow Pack.

7.1.3. Multiple Variable Snow Pillow Linear Regression

To determine if there may be correlations between unimpaired flow at Fort Jones and multiple snow pillow locations, a multiple variable linear regression was explored. A basic multi-variable linear regression analysis was run for five such periods. The smallest data set had 9 years of data and the largest had 39 years of data. At most seven stations were used at the same time and in the smallest model only five stations were considered. The general form of the equation is:

$$Y_t = m_1 \text{Station}_{1,t} + m_2 \text{Station}_{2,t} + \dots m_n \text{Station}_{n,t} + b$$

For example, the first model covered a period from WY1979 to WY1994 and included 16 snow water content values (one data point per year) for six of the stations (BXC, ETN, SWJ, DYM, MB3, and MBL). The snow water content readings on April 1st for each of these stations were used as the independent variables to estimate unimpaired flow at Fort Jones.

$$Y_t = m_1 \text{BXC}_t + m_2 \text{ETN}_t + m_3 \text{SWJ}_t + m_4 \text{DYM}_t + m_5 \text{MB3}_t + m_6 \text{MBL}_t + b$$

Similar equations can be formulated for any combination of the snow water content data. Determination of the coefficients (m_i) and the constant (b) requires the use of a statistical program⁵. The multiple variable linear regression analysis was done using Systat 11.

1951 – 1982 Model

The 1951 – 1982 model started with five stations (MBV, ETN, SWJ, MB3, and MBL). Ultimately MBV and SWJ were removed from the model. The adjusted R^2 value was 0.392. The coefficients, standard error, and P-values are presented below (Table 8).

⁵ Microsoft Excel has a Data Analysis Regression Tool, but it is only capable of performing basic linear regression. It is ideally suited for single variable linear regression and not multi-variable linear regression.

Table 8. Multi-Variable Linear Regression Results for the 1951-1982 Model.

	Coefficient	Std. Error	F¹	P-Value
CONSTANT	177.991	101.192	-	-
ETN	8.017	3.283	5.965	0.021
MB3	-36.095	12.549	8.273	0.008
MBL	35.028	10.354	11.444	0.002
MBV	-	-	0.243	-
SWJ	-	-	0.118	-

¹ F to enter/remove was 4.0

1956 – 1994 Model

The 1956 – 1994 model started with five stations (ETN, SWJ, DYM, MB3, and MBL). Ultimately SWJ, DYM, and MBL were included in the model. The adjusted R² value was 0.580. The coefficients, standard error, and P-values are presented below (Table 9).

Table 9. Multi-Variable Linear Regression Results for the 1956-1994 Model.

	Coefficient	Std. Error	F	P-Value
CONSTANT	27.516	75.165	-	-
SWJ	9.508	2.569	13.703	0.001
DYM	-24.296	5.613	18.736	0.000
MBL	19.942	4.130	23.316	0.000
ETN	-	-	0.058	-
MB3	-	-	0.000	-

¹ F to enter/remove was 4.0

1951-1958, 1975-1978 Model

The 1951-1958, 1975-1978 model started with six stations (MBV, LOG, ETN, SWJ, MB3, and MBL). It is two periods because the data at LOG is not continuous from 1951 through 1978. The adjusted R² value is 0.857. The coefficients, standard error, and P-values are presented below (Table 10).

Table 10. Multi-Variable Linear Regression Results for the 1951-1958, 1975-1978 Model.

	Coefficient	Std. Error	F	P-Value
CONSTANT	-293.080	130.176	-	-
MBV	-60.484	26.337	5.274	0.061
LOG	58.559	27.469	4.545	0.077
SWJ	44.792	15.472	8.382	0.028
MB3	-124.178	37.502	10.964	0.016
MBL	113.694	32.482	12.251	0.013
ETN	-	-	0.054	-

¹ F to enter/remove was 4.0

1979 – 1994 Model

The 1979 -1994 model started with six stations (BXC, ETN, SWJ, DYM, MB3, and MBL). Ultimately BXC and MB3 were removed from the model. The adjusted R² value was 0.921. The coefficients, standard error, and P-values are presented below (Table 11).

Table 11. Multi-Variable Linear Regression Results for the 1979-1994 Model.

	Coefficient	Std. Error	F	P-Value
CONSTANT	-75.111	45.427	-	-
ETN	-14.603	4.504	10.512	0.008
SWJ	16.685	3.865	18.640	0.001
DYM	-26.182	4.216	38.572	0.000
MBL	27.699	3.644	57.775	0.000
BXC	-	-	2.618	-
MB3	-	-	0.393	-
¹ F to enter/remove was 4.0				

Of the four multi-variable linear regression models, the 1979 – 1994 one had the highest adjusted R² value (0.921) and the smallest standard error values for the coefficients. From the regression analysis the following equation is generated:

$$Y_t = -14.603ETN_t + 16.685SWJ_t - 26.182DYM_t + 27.699MBL_t - 75.111$$

There are 16 known values that the above equation can be compared with. If the values for the April 1st snow water content are substituted for the variables in the equation, a predicted unimpaired Scott River total annual flow can be calculated and compared to actual historic conditions (Figure 26). For the 16 data points, the equation over-predicts 10 times and under-predicts 6 times (Figure 27 and Figure 28). The average difference is zero, but if the positive and negative values are averaged separately, then the model over-predicts by about 38 taf/yr and under-predicts by about 63 taf/yr. In general, over a long period of time the average annual flow predicted by the model will match the average annual observed flow.

In terms of using a model, this under or over prediction of the annual unimpaired river flow may or may not be a significant problem depending on the total annual runoff. For example, a 50 taf error for a total annual runoff of 1,000 taf (5%) may be acceptable, while a 50 taf error for a total annual runoff of 250 taf (20%) may be unacceptable. It is necessary to determine the desired level of accuracy prior to selecting or rejecting a particular model.

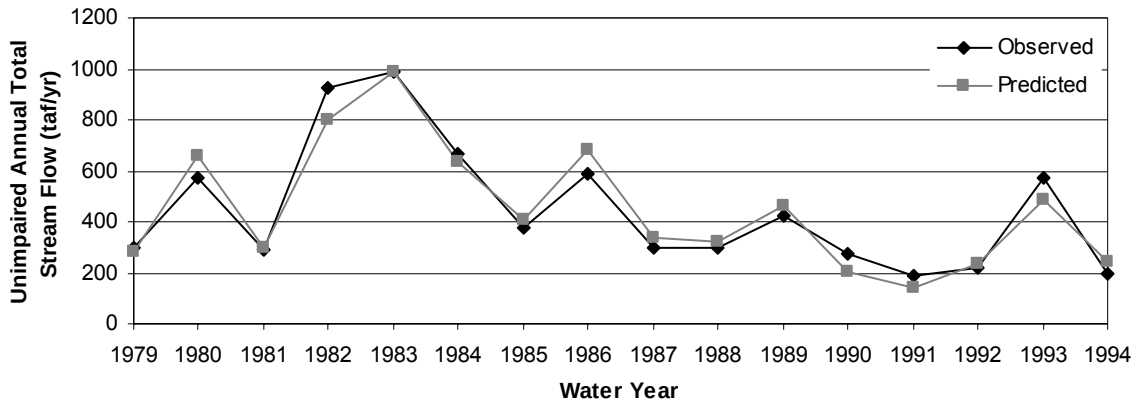


Figure 26. Observed and Predicted Unimpaired Flow Using Multi-Variable Linear Regression.

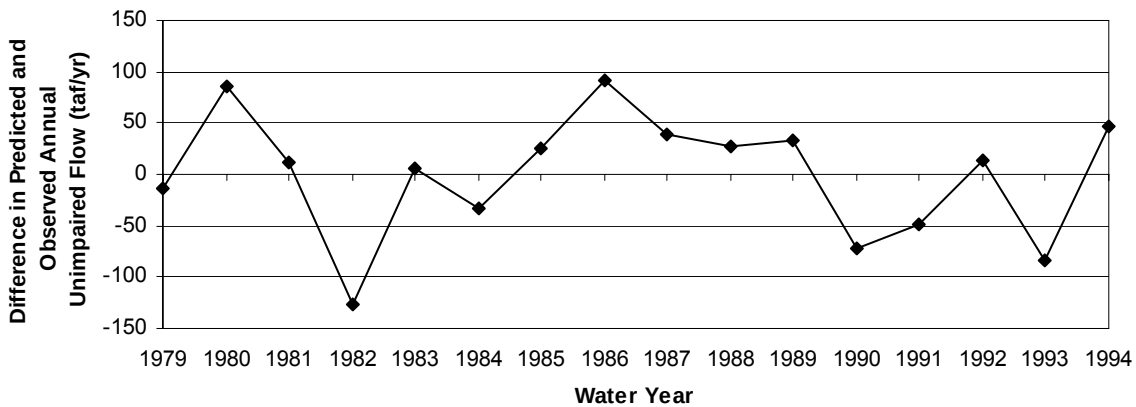


Figure 27. Differences in Observed and Predicted Unimpaired Flow Using Multi-Variable Linear Regression.

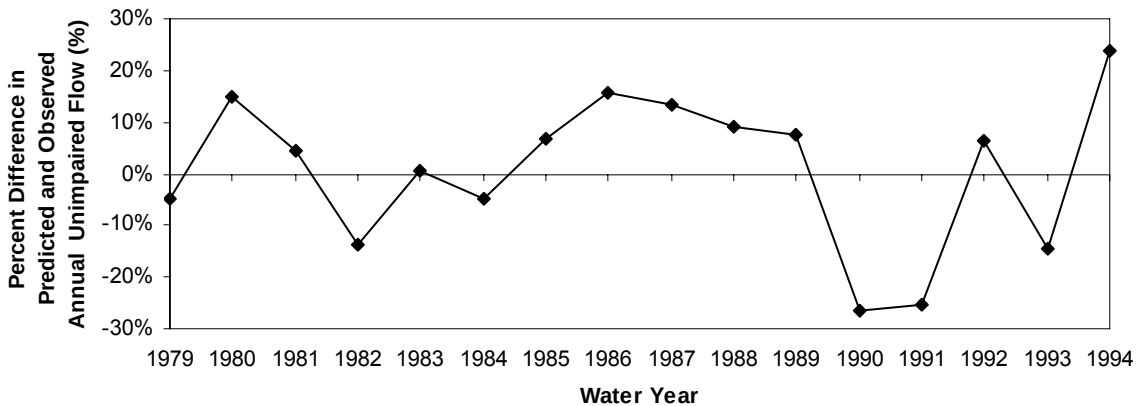


Figure 28. Percent Differences in Observed and Predicted Unimpaired Flow Using Multi-Variable Linear Regression.

After ranking the differences in ascending order and selecting the values corresponding roughly to 5% and 95% exceedence, the range of over and under prediction was identified. The Multi-Variable Linear Regression under estimates the annual unimpaired flow by up to 127 taf and over estimates by up to 92 taf. Based on the average or normal values for the

four stations used in the relationship (Table 12), the model was used to estimate unimpaired runoff at Fort Jones for conditions ranging from 10 through 250 percent of normal. The results, with the uncertainty included, are depicted graphically in Figure 29.

Table 12. Normal (Average) April 1st Snow Water Content Values for ETN, SWJ, DYM and MBL.

Station	Normal (Average) April 1 st Snow Water Content (in)
ETN	20.7
SWJ	26.2
DYM	16.2
MBL	29.4

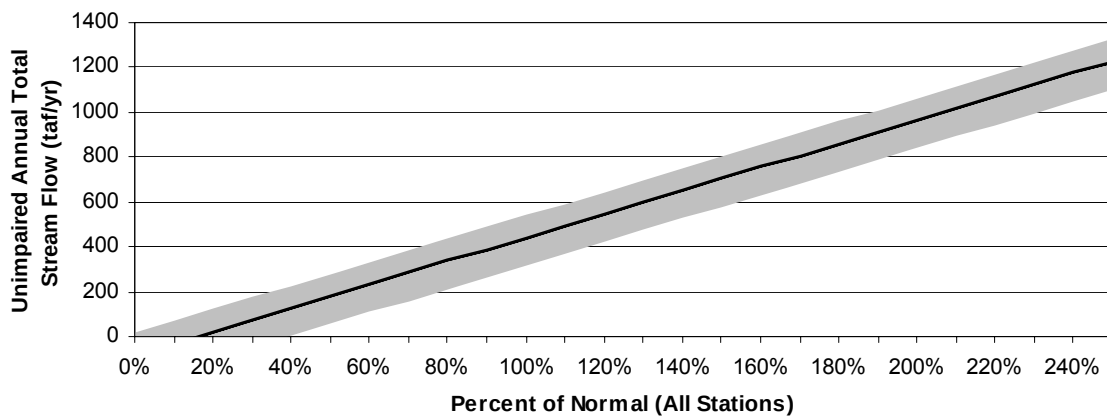


Figure 29. Multi-Variable Linear Regression Results with Error Range.

7.2. Rainfall Single Variable Linear Regression

Precipitation is another hydrological component that has direct implications on runoff. There are five precipitation stations located in and around the Scott Valley. The longest records exist for the Fort Jones station (January 1936 to the present) and the Callahan station (October 1943 through the present). Both records are presented as monthly accumulated precipitation measurements. The accumulated October through March precipitation at the two stations varies between years (Table 13). The accumulated October through March precipitation record is presented in Figure 6.

Table 13. Maximum, Minimum, and Average Accumulated October through March Rainfall.

	Callahan (CAL) ¹	Fort Jones (FJN) ²
Maximum Accumulated October – March Rainfall (in)	28.9	30.3
Minimum Accumulated October – March Rainfall (in)	4.9	4.8
Average Accumulated October – March Rainfall (in)	16.4	17.4

¹ Callahan values are based on water years 1943 – 1969, 1971 – 1979, and 1982 – 2004.

² Fort Jones values are based on water years 1936 – 2004.

Plotting the unimpaired stream flow against the October through March accumulated precipitation yielded a plot that displays a positive trend for both stations (Figure 30 and Figure 31). Graphically, it appeared that the accumulated precipitation could be used as a preliminary means of predicting the total unimpaired stream flow. The relationship appears to be better between the Fort Jones station data and unimpaired stream flow than it does for the Callahan station (Table 14).

Table 14. Regression Summary for FJN and CAL Regression Models.

	FJN	CAL
R ²	0.913	0.850
Constant	-136.267	-172.96
Standard Error on the Constant	30.75	44.51
Coefficient	38.10	39.33
Standard Error on the Coefficient	1.70	2.47

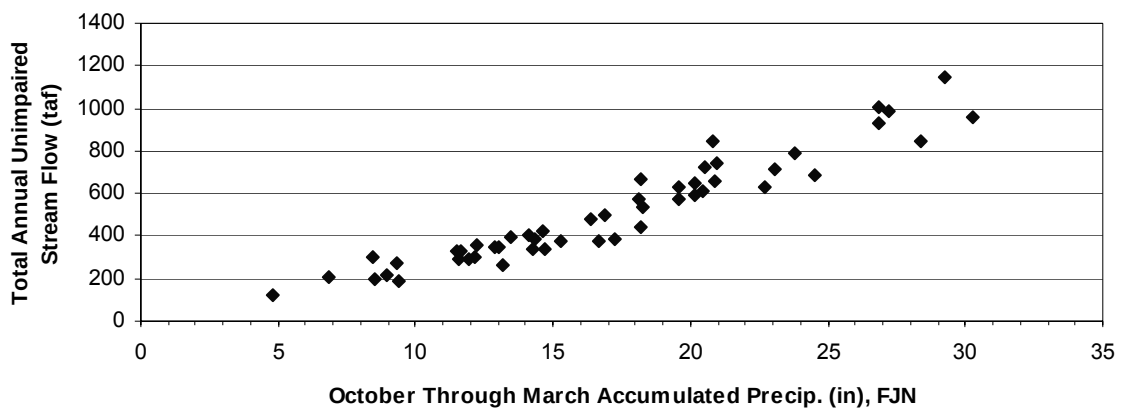


Figure 30. Unimpaired Stream Flow for the Scott River near Fort Jones and Accumulated October through March Precipitation at Fort Jones.

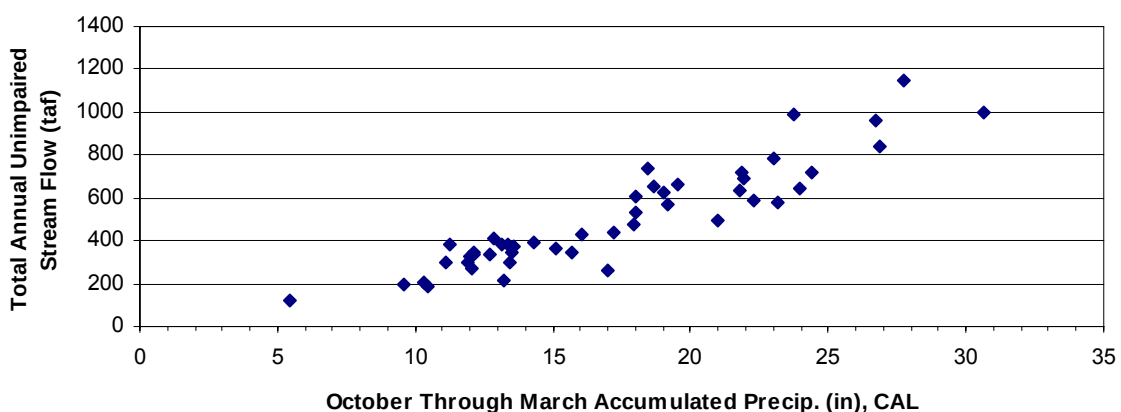


Figure 31. Unimpaired Stream Flow for the Scott River near Fort Jones and Accumulated October through March Precipitation at Callahan.

The linear regression equation relating October through March precipitation at Fort Jones to unimpaired runoff at the Fort Jones gage site is:

$$Y_t = 38.096FJN_t - 136.27$$

Where FJN_t is October through March precipitation (in) and Y_t is annual unimpaired runoff at Fort Jones. The R^2 value for the equation is 0.913 (Figure 32). It is worth noting that like the multivariable linear regression, the regression above produces a fit that matches the observed historical data over a long period of time. For the 50 years of data available, the long-term average difference is 17 taf/yr. The model over predicts the stream flow in 26 of the years and under predicts it in the remaining 24 years. The over prediction is approximately 53 taf/yr (on average) and the under prediction is approximately 58 taf/yr (on average). In any given year the regression can over or under predict the unimpaired stream flow by up to approximately 40% and 60%, respectively (Figure 33 and Figure 34).

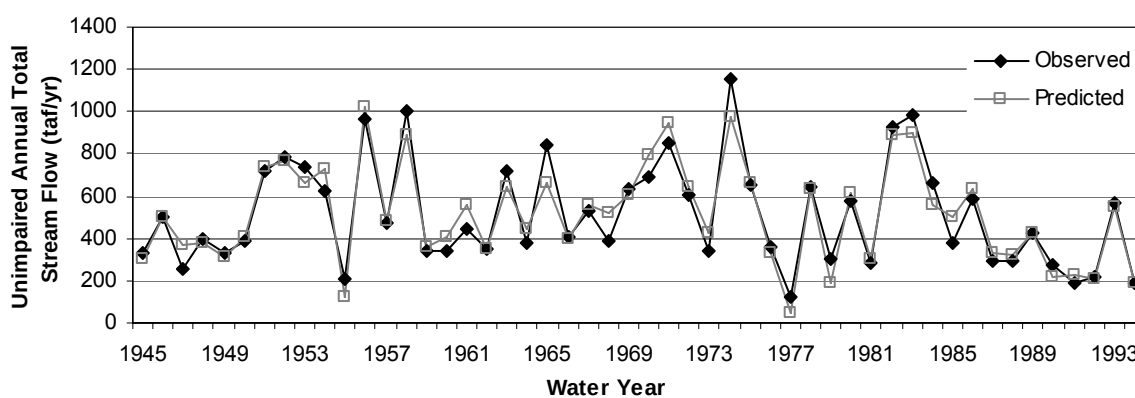


Figure 32. Observed and Predicted Unimpaired Flow Using FJN Precipitation Linear Regression.

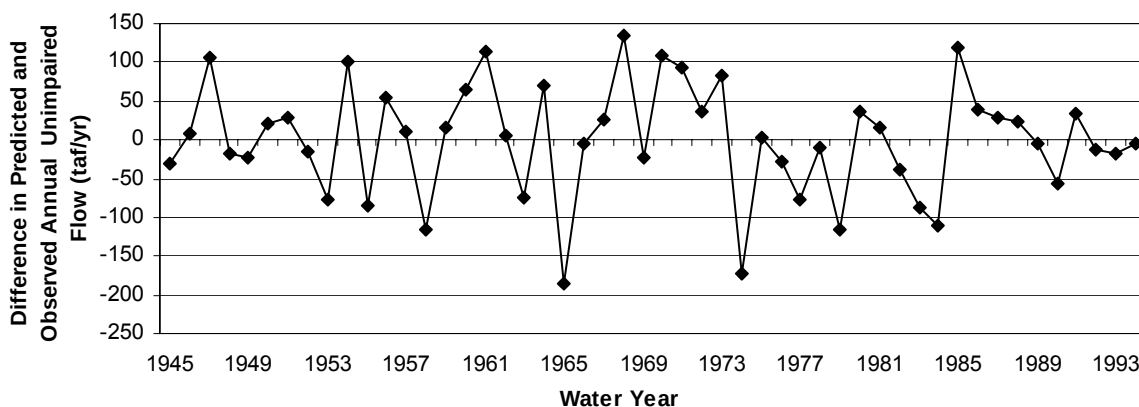


Figure 33. Difference in Observed and Predicted Unimpaired Stream Flow Using FJN Precipitation Linear Regression.

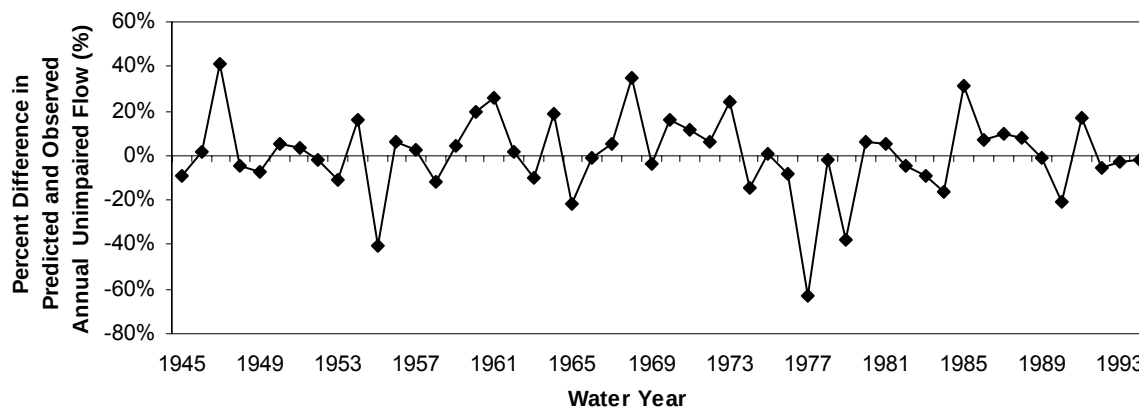


Figure 34. Percent Difference in Observed and Predicted Unimpaired Stream Flow Using FJN Precipitation Linear Regression

After ranking the differences in ascending order and selecting the values corresponding roughly to 5% and 95% exceedence, the range of over and under prediction was identified. The Multi-Variable Linear Regression under estimates the annual unimpaired flow by up to 143 taf and over estimates by up to 117 taf. Based on an average or normal values for FJN of 17.1 inches (October through March), the model was used to estimate unimpaired runoff at Fort Jones for conditions ranging from 10 through 250 percent of normal Figure 35.

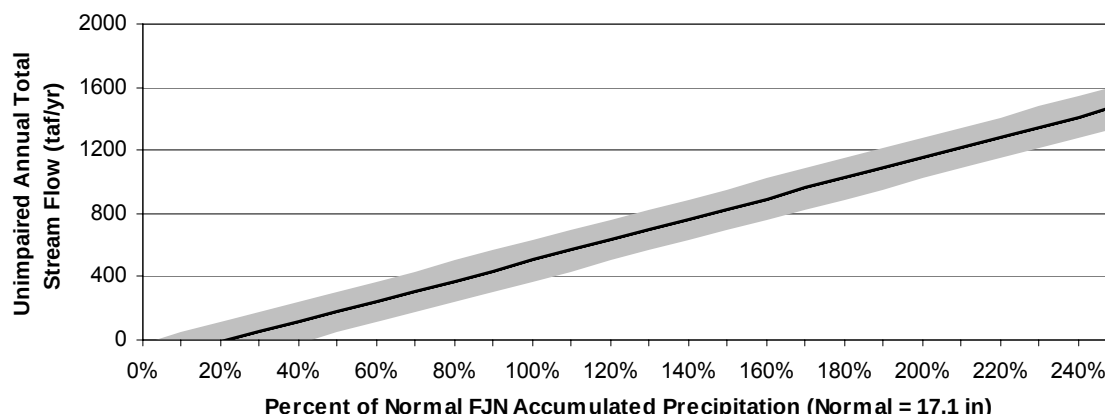


Figure 35. FJN Precipitation Linear Regression Results with Error Range.

7.3. Multivariable Snow Pillow and Rainfall Linear Regression

Models using snow pillow or rainfall data have been developed in the previous sections. A combination of snow pillow and rainfall data could be used to create a multivariable linear regression forecast. Regardless of the combination, rainfall data dominated the relationship. The snow pillow data (be it a single or multiple records) was always dropped (or not included) from the multiple regression analysis due to having an F-Statistic less than 3.9. The result is that the analysis yielded a relationship between rainfall and runoff.

The R^2 values ranged from 0.910 to 0.958. It is interesting to note that the R^2 values generally decreased as the number of data points increased (Table 15). This indicates that

some of the shorter data sets did not exhibit the variability present in the longer data sets (overall this indicates some of the limitations of data models).

While it is worth noting that rainfall and snow pillow data are both important components of runoff, statistically the relationship that can be derived from the existing data tend to exclude the snow pillow data when rainfall data is also present. As a result of this, none of the rainfall and snow pillow data regression relationships are included herein.

Table 15. Number of Data Points and R² Values for Snow Pillow and Rainfall Regressions.

Snow Pillow/Rainfall Stations Used	Water Year Range	n	R²
ETN/FJN	1951 – 1994	44	0.910
SWJ/FJN	1951 – 1994	44	0.910
MBV/FJN	1951 – 1982	32	0.910
MBL/FJN	1946 – 1994	49	0.911
DYM/FJN	1956 – 1994	39	0.911
MB3/FJN	1948 – 1994	47	0.912
BXC/FJN	1979 – 1994	16	0.936
LOG/FJN	1951 – 1958, 1975 – 1978	12	0.952
SCT/FJN	1986 – 1994	9	0.958

7.4. Comparison of Forecasting Tools

The preceding sections present three different forecasting models: a single variable snow pillow model (using BXC), a multivariable snow pillow model (using ETN, SWJ, DYM, and MBL) and a single variable rainfall model (using FJN) (Table 16). The single variable snow pillow linear regression (BXC Model) was

$$Y_t = 12.03BXC_t + 35.87$$

The multivariable snow pillow linear regression (Multivariable Snow Model) was

$$Y_t = -14.603ETN_t + 16.685SWJ_t - 26.182DYM_t + 27.699MBL_t - 75.111$$

The single variable rainfall model (FJN Model) was

$$Y_t = 38.096FJN_t - 136.27$$

The BXC model had the lowest R² value and the largest range of over and under predictions of the three models. The multivariable model had the best R² value and the smallest range of over and under predictions. The FJN model was in the middle, with a high R² value and relatively low range in over and under predictions. See Figure 36 through Figure 39 for details.

Table 16. Summary of Results from the Linear Regression Models.

	n	R ²	Average Error* (taf/yr)	Over Predict**		Under Predict**	
				Max (taf/yr)	Average (taf/yr)	Max (taf/yr)	Average (taf/yr)
Single Variable Snow Pillow Model	16	0.789	0.0 (0.0%)	193 (65%)	119 (41%)	188 (39%)	71 (15%)
Multivariable Snow Pillow Model	16	0.921	0.03 (0.8%)	92 (24%)	38 (10%)	127 (26%)	63 (15%)
Single Variable Rainfall Model	50	0.914	0.0 (0.0%)	133 (41%)	53 (13%)	184 (63%)	58 (13%)

* The average error, the over predicted maximum and average, and under predicted maximum and average values are based on the entire period of record for which the model was applicable. For the Single Variable Snow Pillow (BXC) and Multivariable model the data set was limited to WY1979 to WY1994. The data set for the Single Variable Rainfall (FJN) model was more extensive (WY1945 to WY 1994). The second line of summary statistics for the Single Variable Rainfall (FJN) model represents the values if only WY1979 to WY1994 were considered.

** The taf/yr is the volume by which the model over/under predicted the annual runoff, while the percentage indicates what percent of the observed flow the over/under-prediction represented.

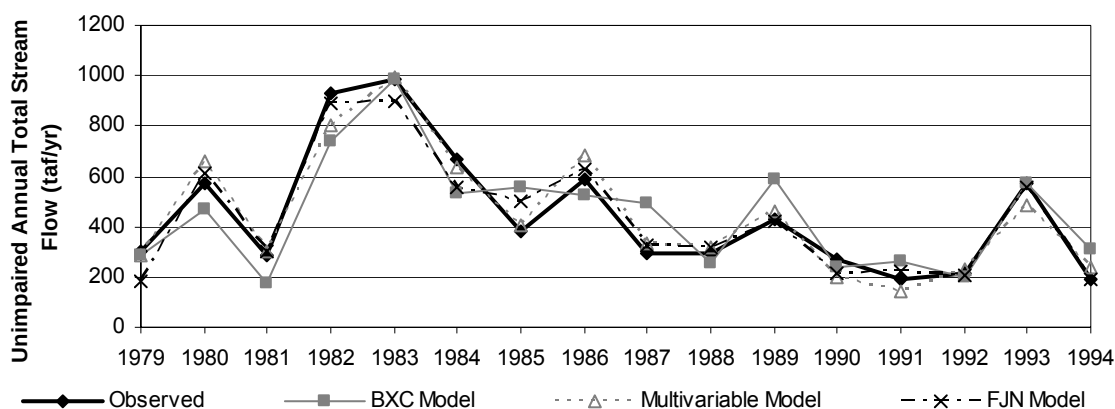


Figure 36. Comparison of the Three Linear Regression Models.

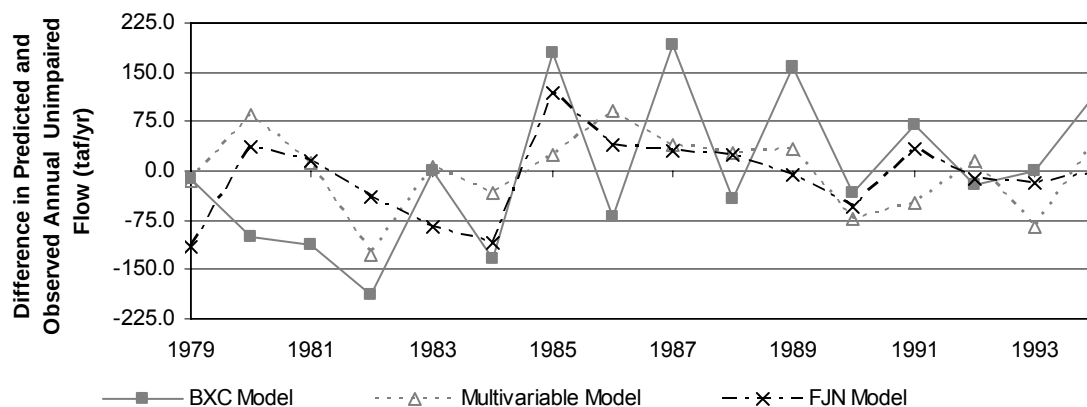


Figure 37. Difference in Observed and Predicted Unimpaired Stream Flow for the Three Models.

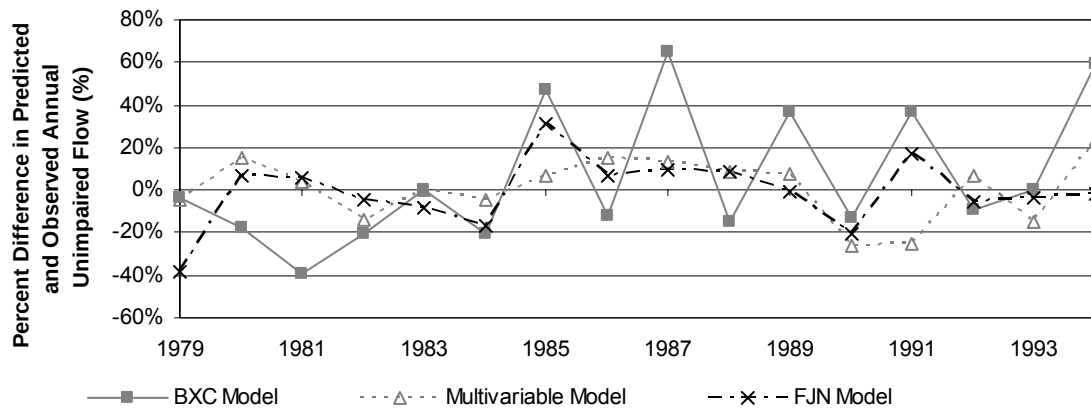


Figure 38. Percent Difference in Observed and Predicted Unimpaired Stream Flow for the Three Models.

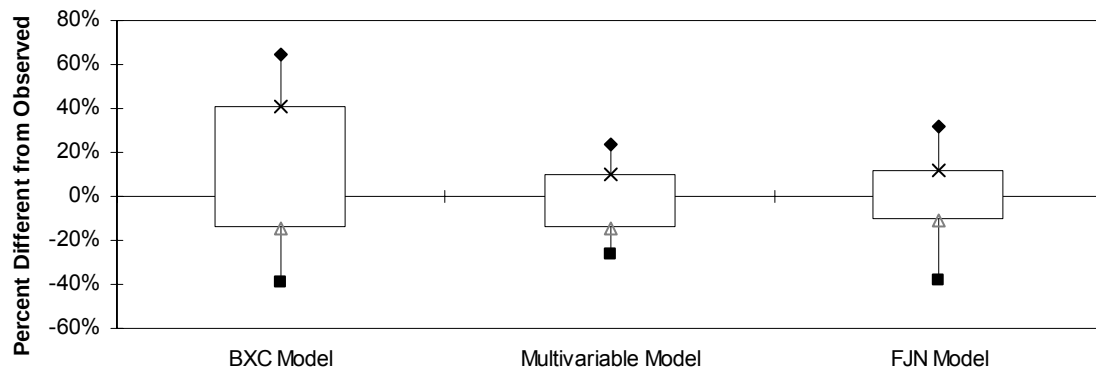


Figure 39. Maximum and Average Error for the Three Models When Comparing WY1979-1994 (solid diamond is the maximum over prediction, x is the average over prediction, grey triangle is the average under prediction, and solid square is the maximum under prediction).

While any of the three models could be used as a preliminary means of forecasting the unimpaired stream flow, it is important to be aware of the limitations and weaknesses of the models. These linear regression models are not a sufficiently robust model to use for detailed planning at a small time step (sub-annual) and would probably be best suited for long term planning. However, they are useful for providing a approximate first estimate of runoff forecast. Significant changes in environmental, meteorological, land use or other process changes could increase the uncertainty in these models.

8. Other Factors

The linear regression relationships presented above attempt to relate measured rainfall or snow water content data to the unimpaired total annual stream flow at the Fort Jones gage. However, none of the models provided a highly accurate forecast, nor do any of them address runoff timing. As evidenced from Figure 39 there is the potential for uncertainty in the forecast (either as an under prediction or an over prediction) for any of the applied models. This would indicate that precipitation as rainfall or snow are not the only factors that influence runoff in the Scott River Valley.

Factors that potentially play a role in runoff quantity and timing including the spatial and temporal variation in precipitation, antecedent conditions (e.g., dry soils, previous year rainfall and/or runoff), land use practices, water use practices, the rate and frequency of precipitation through the year, non-winter precipitation, vegetation condition, meteorological conditions, etc. Two other components in the Scott River Valley that are likely to impact the forecasting of the flow in the Scott Valley proper include groundwater levels and diversions (Figure 40).

A critical aspect of constructing a runoff forecast model is identifying the point where the forecast will apply. In the previous examples, the unimpaired runoff for Scott River near Fort Jones (USGS gage site) is used as the independent variable (the value that is forecast). However, the runoff at other locations may be desired for water resources planning and management activities (for example, total annual runoff at the edge of the valley above diversions and prior to encountering the groundwater basin or at some intermediate location in the valley). There are different data requirements and approaches depending on which forecast is desired. Statistical models would most likely employ snow water content and/or seasonal rainfall data, but there are additional data requirements for forecast models that would address areas within the valley or at the downstream end of the valley. These latter formulations may include consumptive use (or some surrogate set of parameters that may represent consumptive use, e.g., land use characterization, water application, cropping patterns, and water use) or groundwater conditions as dependent variables to represent the processes affecting runoff at the selected locations. Outlined below are additional areas of interest pertaining to forecasting unimpaired runoff in tributaries (major and minor) prior to reaching the valley floor and the role of groundwater and consumptive uses on forecasting runoff at locations within the valley.

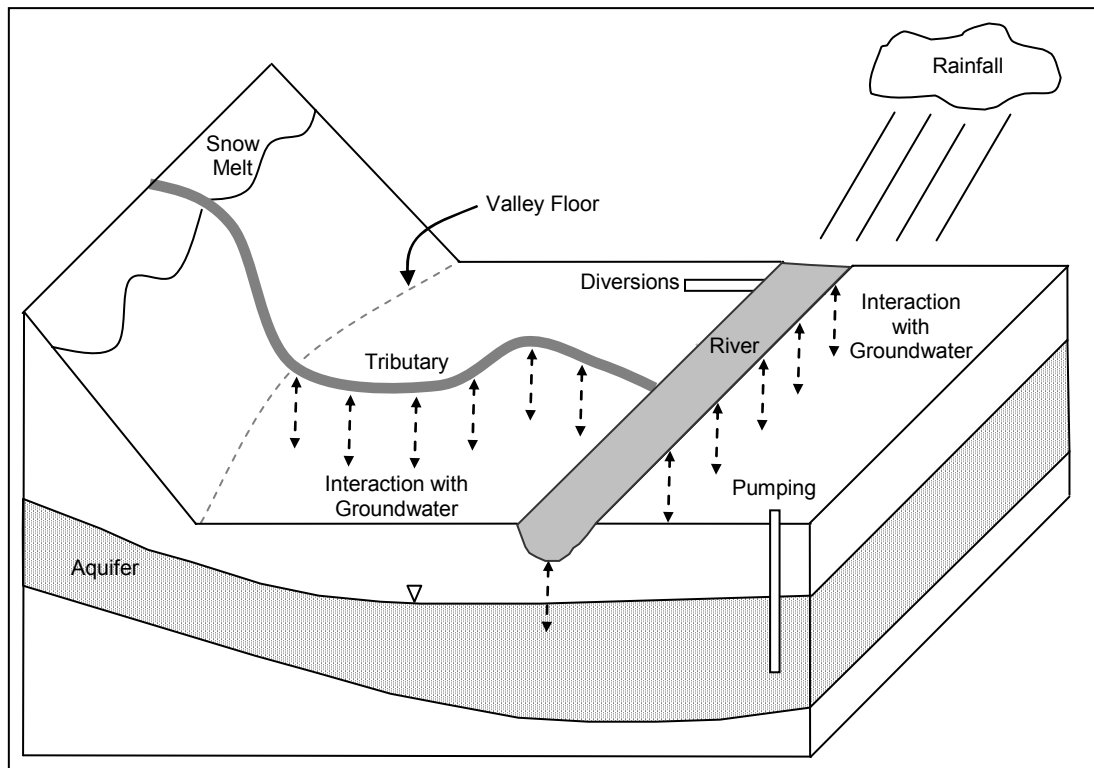


Figure 40. Major Water Interactions Schematic.

8.1. *Tributary Inflows into the Scott River Valley*

If the goal is to forecast the unimpaired inflow to the Scott Valley, then flows at the “edge” of the valley where major tributaries (and possibly minor tributaries) reach the valley floor are focal points – shown in the schematic as a grey dashed line (Figure 40). It is relatively reasonable to assume that there is little interaction between surface and groundwater at this location. Thus, tributary flows could be related to the appropriate snow water content data and an unimpaired runoff forecast could be made. Current monitoring is in place on Shackleford, Mill, Kidder, and French Creeks, on the East and South Forks of the Scott River, and in the Scott River at Fort Jones (Figure 42)⁶. Monitoring at these sites provide some estimation for the flows that reach the valley edge via these tributaries. There are other tributaries and locations in the Scott Valley that might be beneficial if monitored. Examples include, but are not limited to, Moffet, Etna, and Sugar Creeks. It may also be beneficial to monitor the creeks for which there are corresponding snow stations (such as Etna Creek and the ETN snow station). Monitoring would have to be carried out for a number of years to capture enough system variability to construct a meaningful relationship. A typical length of time is approximately 10 to 15 years. Using the Scott River unimpaired flow record for a period of less than 15 years, the running average values change considerably with the addition of a subsequent year (Figure 41). After 15-years the change in running average is relatively small as each additional year of data is added. However, it should be noted that if

⁶ The RCD, USFWS, and DWR have been monitoring flows in selected tributaries and recent additions include Shackleford Creek near the mouth, Shackleford-Mill Creek approximately 1/2 mile below Emigrant Creek (at the Quartz Valley Road Bridge), French Creek above North Fork, East and South Scott Rivers above their confluence (Deas et al., 2004).

the selected period does not provide sufficient variability then monitoring may have to be extended for additional years to capture the potential range of conditions in the basin. Likewise, while 10 to 15 years of data would provide a good basis, ongoing monitoring would be ideal.

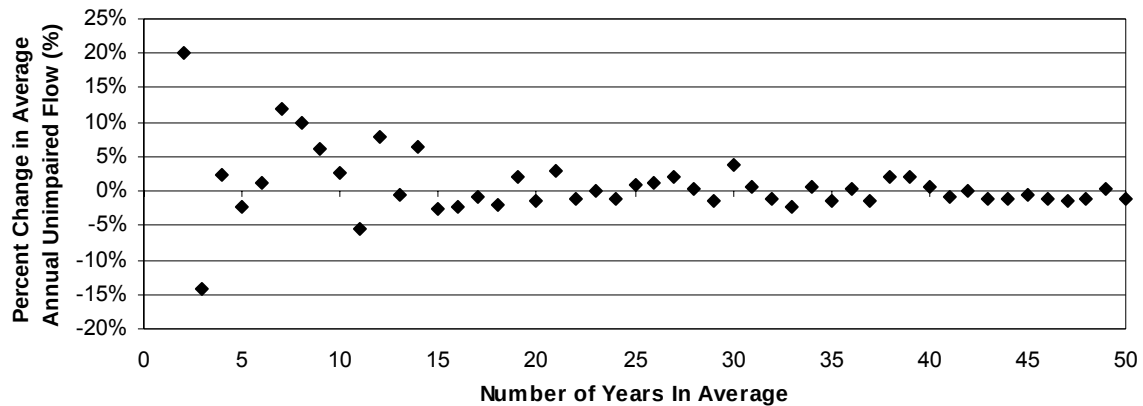


Figure 41. Percent Change in the Running Average Annual Unimpaired Flow in the Scott River.

There is probably sufficient coverage to provide enough snow water content data for the purposes of forecasting. The BXC station corresponds to the northern portion of the basin, while ETN and SWJ could be used to represent the middle of the basin, and DYM, MB3, MBL, and SCT can represent the southern portion of the basin. Specifically, with a longer flow record (e.g., 10-15 years) it might be possible to relate DYM, MB3, and/or MBL to the flows in the South Fork of the Scott River and relate SCT to the flows in the East Fork of the Scott River.

8.2. Groundwater Levels

If a runoff forecast model is desired for an intermediate location in the valley or at the downstream end of the valley, accounting for groundwater conditions would be a valuable arena to explore in any statistical model.

The Scott River Watershed is shown in Figure 42 with the valley floor shaded, which is roughly equivalent to the extent of alluvium, while the non-shaded area approximately represents the surrounding higher elevation watershed. While some of the snowmelt is lost to evaporation and/or infiltration, a large portion of tributary runoff reaches the valley floor as surface water flow. Once tributary flows reach the alluvium, a discernible portion infiltrates into the groundwater basin. This is noticeable in many of the west side tributaries which run dry in their lower reaches in mid-to-late summer. (Albeit there are complex relationships between the losing reaches, diversions, and groundwater pumping. Further mining tailings in the upper basin and local soil types, such as impermeable (e.g., clay) materials, may also affect recharge and permeability.)

Because groundwater can augment river base flow in gaining reaches or periods of the year, as well as diminish base flow in losing reaches or periods of the year, impacts on runoff forecasting are evident. In addition to intra-annual conditions, inter-annual variability can be important in groundwater systems. Although beyond the scope of this initial investigation, it

would be valuable to explore potential relationships between runoff and groundwater levels (or some surrogate) at multiple locations in the valley. Because groundwater generally reacts slower than surface waters to changes in hydrology, previous year(s) conditions can affect current conditions. Thus, to incorporate groundwater into a statistical model, multivariable linear regression model with lagged parameters or an ARMA model may be useful. Recall that the ARMA models can take into consideration the conditions in the previous periods when determining the current or future conditions.

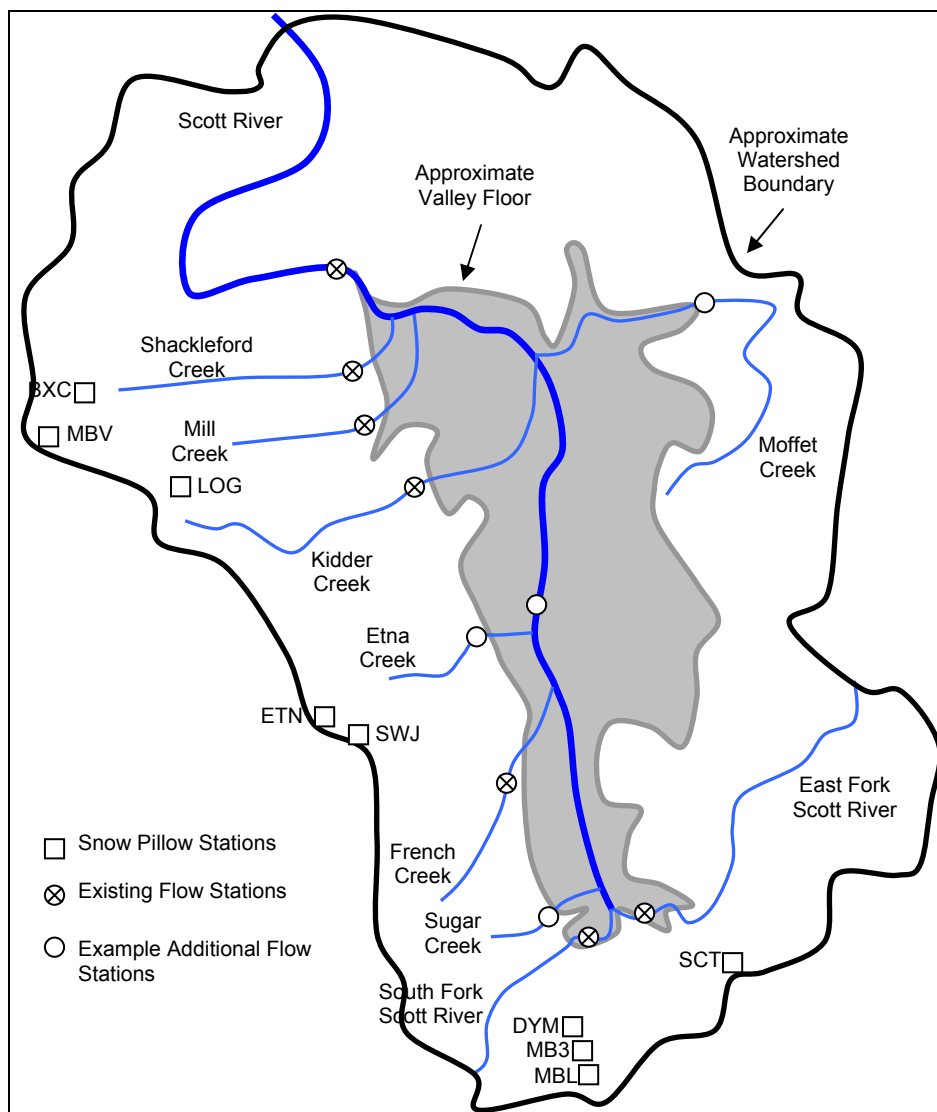


Figure 42. Schematic Representation of the Scott River Watershed with Example Monitoring Locations.

8.3. Consumptive Use

As with groundwater, if a runoff forecast model is desired for an intermediate location in the valley or at the downstream end of the valley, accounting for consumptive uses would be a valuable arena to explore in any statistical model.

As waters flow from the southern Scott Valley north, stream flow is impacted by groundwater levels, diversions, and additions from the tributaries. By the time the East and South Forks merge at Callahan, diversions are sufficient to impact the hydrograph so that the flows measured at Callahan are not representative of the 'natural' hydrograph⁷.

By the time the flow reaches the gage at Fort Jones, there has been a significant change in the volume observed compared to what the unimpaired flow would total. DWR estimated the unimpaired flow in the Scott River at Fort Jones from WY1945 to WY1994. It appears that the consumptive use has reached a relatively stable volume (Figure 11) of about 70 taf/yr on average. However, this is only an approximation and may not capture the variability in diversions due to land use changes, meteorological conditions, groundwater levels, or other factors.

For the purposes of this investigation, an unimpaired annual runoff estimate (DWR, 1998) was used to relate precipitation to runoff. This is one approach; however, more detail or an improved level of reliability may be desired. In cases where a runoff forecast model is used to estimate the flows in the Scott River downstream of substantial diversions (such as is the case at the Fort Jones gage), then some means of incorporating the effects of diversions into the model should be considered. There are a variety of ways in which the consumptive use could be explicitly incorporated into the model. Generally, these could be included as another dependent variable in the regression relationships. However, a thorough exploration of the information available from DWR on consumptive use should be considered. Such an exploration would identify the temporal and spatial characteristics of the data, any uncertainty that might be associated with the information, the length of available record, and other pertinent factors. Consumptive use is a significant activity during the year and should not be neglected when forecasting runoff on the valley floor.

9. Conclusions

Runoff is a major component of the hydrologic cycle; providing surface water supplies to recharge groundwater (via infiltration) and augment river/tributary flows. The volume and timing of runoff that reaches the river depends on the meteorological conditions (air temperature, precipitation, snowfall), watershed and groundwater basin conditions, geographic conditions (land area, slopes, vegetation cover, soil types), and a variety of other factors. These conditions vary from location-to-location and from year-to-year. Any forecast model that is needed at a fine resolution (e.g., sub-annual) would need to consider these inter-related conditions that influence runoff.

Presented herein is a suite of regression relationships (statistical models) that illustrate the potential for runoff forecasting based on seasonal rainfall and/or snow water content. From the three models developed it also appears that rainfall is an important variable in predicting stream flow. Snow water content also appears to be an important parameter because this is where the majority of the surface water flows originates. Discretion should be employed in

⁷ The allocated water rights to the South and East Fork of the Scott River upstream of the confluence at Callahan represent a significant portion of the flow during the irrigation season (John Clements, electronic mail, October 27, 2005).

the application of any such statistical model, noting that runoff is influenced by a wide variety of conditions whose influence may be more critical at different times and under different situations.

Data records exist for monthly snow water content at nine locations (seven of the stations remain active). More detailed data records exist for precipitation at various locations throughout the watershed. Flow data exists at several locations, but certain stations have short records that limit their use in the statistical models presented herein. Further, an unimpaired flow record for the Scott River is available, but it requires an update from 1994 to present to be more useful.

To generate a more complete runoff forecast model, more detailed information about the flows and precipitation (predominately rainfall) is needed. Further, quantifying surface-groundwater interaction and consumptive use in the valley proper may provide a more comprehensive and useful runoff forecast model, particularly if the runoff forecast is desired for a location within or downstream of the valley. Incorporating groundwater into such an analysis may require examining information from previous years because groundwater conditions generally respond to hydrologic conditions more slowly than surface waters.

In closing, there is sufficient data to create a preliminary forecasting model; however, there is room for improvement. The models presented in this report could be used to determine rough estimates of unimpaired annual stream flow. If the models are used, it should be noted that those values could be inaccurate by some margin of error. Despite those sometimes significant margins of error, the values generated by the first-cut forecasting models could give managers in the Scott Valley a range of annual flows to expect, which could still be useful in planning how manage the valley's water resources.

10. Recommendations

10.1. General Recommendations

A runoff forecast model would benefit from three relatively broad recommendations:

- **Forecast Location Determination** – In order to better determine what data are needed, it should be decided if the forecast is to be of runoff to the Scott River Valley (at the base of mountains), the Scott River runoff at some intermediate valley location, or runoff at the USGS Gage for Scott River at Fort Jones. Determination of what forecast is needed will greatly influence what additional data are needed.
- **Long-Term Data** – To better forecast runoff, additional long-term precipitation (as rainfall and snow water content), flow (in the tributaries and mainstem), and groundwater level data should be considered. The current record is both temporally and spatially sporadic for flow and rainfall with the exception of a few gages. Examination of local precipitation and runoff data, it is apparent that a 15-year period generally produces a sufficiently wide range of conditions to formulate statistical models. Longer data sets are always desirable.

- Groundwater Levels – Groundwater levels should be related to the stream flows because groundwater could have a relatively significant impact on the observed stream flows in the valley. To do this, monitoring of the groundwater levels at several locations throughout the valley floor are needed. Groundwater data were not explored herein, and an assessment of available information is recommended.

10.2. Detailed Recommendations

10.2.1. Data Quality and Management

There are presently nine snow water content and five precipitation data stations. The snow water content stations provide a relatively robust picture of the status of the snow pack surrounding the Scott Valley. The precipitation data is at a relatively fine resolution (hourly), but has notable quality control issues. For a forecast model of the annual stream flow in the Scott River, monthly accumulated rainfall appears to be sufficient to provide a preliminary level analysis.

Likewise, if the snow water content values are used to determine the forecasted unimpaired annual flow, April 1st forecasts should be employed. Although there are data from other months, it appears that, given the annual variability, April 1st is an acceptable estimate of the annual forecast. There is no reason that a water resources manager using such a relationships could not consider antecedent or potential future conditions. In fact, such “sensitivity” analysis would be prudent in application of any statistical model.

10.2.2. Tributary Data

The snow water content stations are located at high elevations relative to the Scott River at Fort Jones. However, these stations are often located near major tributaries that discharge to the mainstem Scott River at various locations through the valley. Additional monitoring of flows on the tributaries that have snow stations may provide reliable relationships between snow water content and runoff. A forecast model of the Scott River could be created from multiple sub-models, each representing a different major tributary. To facilitate a better understanding of the snow water content, precipitation, and runoff, the flows in the tributaries of interest should be monitored at the valley edge (when the tributary first reaches the valley floor). In addition, gaging the stream near the confluence with the Scott River would provide a preliminary means of determining the gains and losses to groundwater, irrigation, etc.

In addition to the tributary monitoring, monitoring of the flows in the South and East Fork of the Scott River before they merge at Callahan should be continued. While it is not clear that the flows in the South and East Fork of the Scott River could be used to forecast the flows at Fort Jones, monitoring should be pursued.

10.2.3. Unimpaired Stream Flow

To forecast the unimpaired stream flow in the Scott River at Fort Jones, data regarding the consumptive use of water upstream of the gage would be beneficial. Presently the gage records the impaired stream flow. Estimates of the stream flow could be achieved by monitoring the withdrawals and return flows to the Scott River, or based on an empirical relationship between land use, acreage, and water use. This information is largely available

from DWR and would lead to an extension of the DWR's work to estimate the unimpaired stream flow to bring the record up to the present. As an initial estimate, a value of 70,000 acre-feet per year (70 taf/yr) could be used. This corresponds to the approximate average level of consumptive use that was calculated by DWR from about 1979 to 1994.

10.2.4. Groundwater Data

The Scott Valley has minimal surface water storage, so the majority of the inter-annual storage occurs via the 300 to 400 taf groundwater basin. Groundwater augments the river in years when the basin is 'full' or it can withdraw water from the tributaries and river to recharge the basin. This large sub-surface water storage also provides base flow in the river. Dry years, despite having a low runoff contribution will still have flow in the river due to the presence of the groundwater inflows (although certain reaches are known to dry up). The impacts of the groundwater basin on the flows in the river and tributary is a location specific phenomena, meaning that what is happening in one location does not necessarily mean the same thing is happening at another. As an important part of the Scott River water balance, additional information on groundwater conditions throughout the basin would be beneficial.

In conjunction with increased surface water monitoring in and around the Scott Valley, groundwater monitoring should be implemented at strategic locations throughout the basin. Depth to the water table is one way to gauge response to hydrologic conditions. Groundwater levels should be monitored on a monthly basis at a minimum, to provide a reasonable time with which to correlate surface water flows.

11. References

- Bay Delta Modeling Forum (BDMF). 2000. Protocols for Water and Environmental Modeling. Prepared by the Ad hoc Modeling Protocols Committee. January. (Posted at <http://www.cwemf.org/>).
- California State Water Resources Control Board. 1974. *Report on Water Supply and Use of Water, Scott River Stream System, Scott River Adjudication, Siskiyou County*. California Division of Water Rights.
- California State Water Resources Control Board (SWRCB). 1975. *Report on Hydrogeologic Conditions, Scott River Valley, Scott River Adjudication, Siskiyou County*. California Division of Water Rights.
- California Department of Water Resources (DWR). 1998. *Klamath River Basin Natural Flow*. Unpublished Data. Red Bluff, CA.
- California Department of Water Resources (DWR). 2003. *California's Groundwater: Siskiyou County Groundwater Basins/Sub-basins, Draft*. Bulletin 118.
- California Department of Water Resources (DWR). 2005. *California Water Plan Update 2005, Volume 4, California River Indices, Draft*. Bulletin 160-05.
- Deas, Michael L., and Stacy K. Tanaka. 2004. "Scott River Water Balance Study – Data Analysis and Model Review." A Report for the Scott River Watershed Council. June.

- Drake, Daniel J., Kenneth W. Tate, and Harry Carlson. 2000. "Analysis shows climate-caused decrease in Scott River fall flows." *California Agriculture*, v. 54(6), November-December.
- Hillier, Fredrick S., and Gerald J. Lieberman. 2002. *Introduction to Operations Research*, 7th ed. McGraw-Hill Inc., New York, NY.
- Mack, S. 1958. *Geology and Ground-Water Features of Scott Valley Siskiyou County, California*, U.S. Geological Survey Water-Supply Paper 1462. Washington D.C.
- Maidment, David R. 1993. *Handbook of Hydrology*. McGraw-Hill Inc., New York, NY.
- Meko, D. 2005. Applied Time Series Analysis. Course materials for geology 585A. University of Arizona, Tucson. <http://www.ltrr.arizona.edu/~dmeko/geos595e.html>
- Stedinger, Jerry R., Robert M. Vogel, and E. Foufoula-Georgiou. 1993. *Handbook of Hydrology*. Ed. D. Maidment. McGraw-Hill Inc., New York, NY.
- Systat. 2004. *SYSTAT 11, Statistics IIII*. SYSTAT Software, Inc., Richmond, CA.
- Taha, Hamdy A. 2002. *Operations Research: An Introduction*, 7th ed. Prentice Hall, New York, NY.
- U.S. Census Bureau. 2003a. "Profile of General Demographic Characteristics: 2000 Geographic Area: Fort Jones city, California." (Posted at <http://censtats.census.gov/data/CA/1600625128.pdf>).
- U.S. Census Bureau. 2003b. "Profile of General Demographic Characteristics: 2000 Geographic Area: Etna city, California." (Posted at <http://censtats.census.gov/data/CA/1600622972.pdf>).
- Waissi, George R. 2002. *Applied Statistical Modeling, An Electronic Textbook*. University of Michigan-Dearborn, Dearborn, MI. (Posted at <http://www.som.umd.umich.edu/gwaissi/stats/regressionbook.html>).
- Yevjevich, Vujica. 1972. *Probability and Statistics in Hydrology*. Water Resources Publications, Littleton, CO.

12. Appendix – Water Year Types

Although not a formal part of this project, an example of characterizing water year types based on historical data is presented herein. Annual Scott River runoff at Fort Jones was used to create exceedance plots and sample year types are defined for illustration purposes only.

12.1. Water Year Types

Water Year types or indexes have been created for a number of river systems in California. The indices are essentially a means of describing the volume of water originating in a river or basin in response to differing hydrologic conditions (e.g., above normal, normal, and below normal conditions) (DWR, 2005).

Using the unimpaired flow estimate for the Scott River, a four year type index can be postulated based on the non-exceedence plot (Figure 13). The 50% non-exceedence criteria

is used to define above and below normal hydrologic conditions. The 20% non-exceedence criteria could be used to define the dry hydrologic conditions, and the 80% non-exceedence criteria could be used to define the wet hydrologic conditions (Table 17 and Figure 43).

Table 17. Five Category Water Year Types.

Water Year Type	Probability of Non-Exceedence	Flow Range ¹ (taf/yr)
Wet	Equal to or greater than 80%	$Q \geq 719$
Above Normal	Equal to or greater than 50%, but less than 80%	$434 \leq Q < 719$
Below Normal	Equal to or greater than 20%, but less than 50%	$297 \leq Q < 434$
Dry	Less than 20%	$Q < 297$

¹'Q' = annual unimpaired flow in thousand of acre-feet per year.

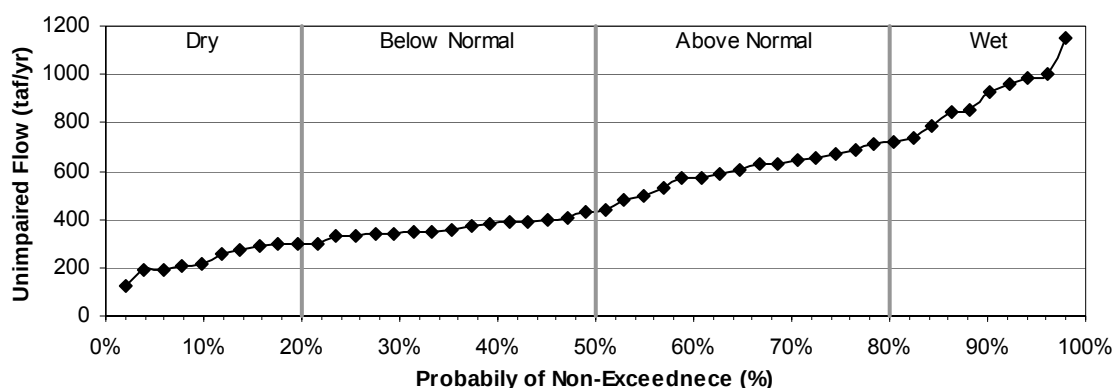


Figure 43. Four Category Water Year Types and Ranges.

Additional resolution could be added to the dry year type category. For example, both the Sacramento and San Joaquin Rivers both have a five category year type index. (Note that the Sacramento and San Joaquin River systems are very different from the Scott River system. They both have significant amounts of surface reservoir storage capacity whereas the Scott River does not.) In addition to the wet, above normal, below normal, and dry categories, there is a critical category. For the Scott River, the dry year type could be divided in half to represent a critical and a dry year type (Table 18 and Figure 44).

Table 18. Five Category Water Year Types.

Water Year Type	Probability of Non-Exceedence	Flow Range ¹ (taf/yr)
Wet	Equal to or greater than 80%	$Q \geq 719$
Above Normal	Equal to or greater than 50%, but less than 80%	$434 \leq Q < 719$
Below Normal	Equal to or greater than 20%, but less than 50%	$297 \leq Q < 434$
Dry	Equal to or greater than 10%, but less than 20%	$218 \leq Q < 297$
Critical	Less than 10%	$Q < 218$

¹'Q' = annual unimpaired flow in thousand of acre-feet per year.

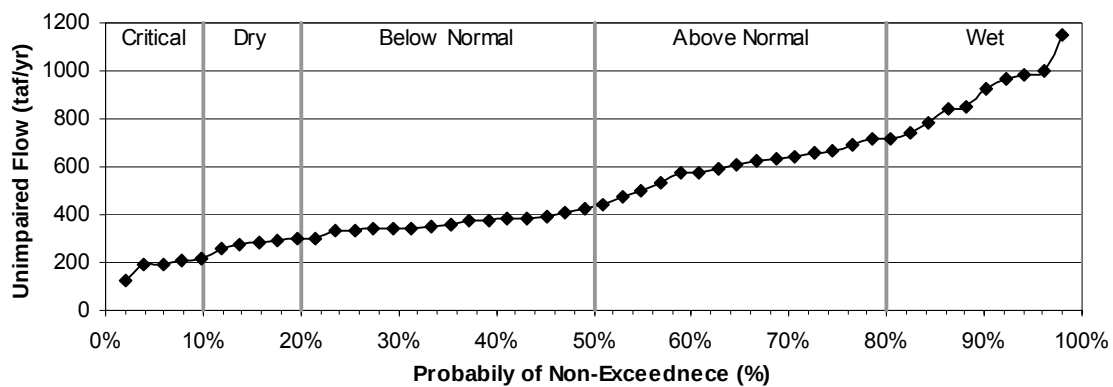


Figure 44. Five Category Water Year Types and Ranges Based On Non-Exceedence Criteria.

Similar water year type designations can be used to produce any number of desired year-types. These can be based on non-exceedence criteria or other statistics such as interquartile range, standard deviation, other measures of variance, and skewness of the distribution. Water year types can also be based on a variety of parameters, including current forecasted runoff, the runoff in previous year(s), and groundwater conditions.