

2000

Improved Regression-Based Streamflow Forecasting Considering Large-Scale Climate Variability

Austin R. Fisher
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Civil Engineering Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

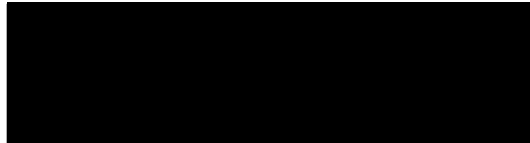
Fisher, Austin R., "Improved Regression-Based Streamflow Forecasting Considering Large-Scale Climate Variability" (2000). *Dissertations and Theses*. Paper 6493.
<https://doi.org/10.15760/etd.3629>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

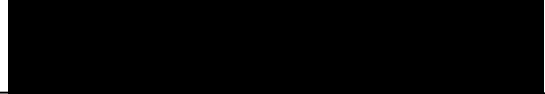
THESIS APPROVAL

The abstract and thesis of Austin R. Fisher for the Master of Science in Civil Engineering were presented July 6, 2000, and accepted by the thesis committee and the department.

COMMITTEE APPROVALS:



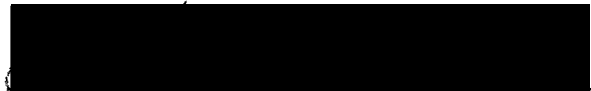
Roy W. Koch, Chair



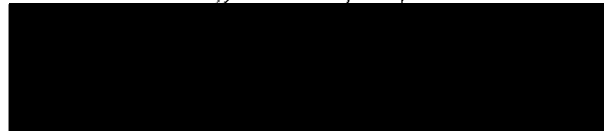
William Fish



David C. Garen



Andrew M. Fraser
Representative of the Office of Graduate Studies



DEPARTMENT APPROVAL:

Franz Rad, Chair
Department of Civil Engineering

ABSTRACT

An abstract of the thesis of Austin R. Fisher for the Master of Science in Civil Engineering presented July 6, 2000.

Title: Improved Regression-Based Streamflow Forecasting Considering Large-Scale Climate Variability

The Natural Resources Conservation Service (NRCS) produces water supply forecasts for most of the streams in the western United States. The NRCS produces forecasts starting in January, when snow course measurements of snow water equivalent were available. Although the seasonal streamflow volume forecasts made by the NRCS are useful, many water supply managers need information at the beginning of the water year in October and would like to see forecasts in the form of a monthly hydrograph.

An investigation into the effect of decadal scale variability, as reflected by the Pacific Decadal Oscillation (PDO), show important relationships that may be useful in forecasting. Data from three basins, the Sandy, Skykomish, and Rogue Rivers were split based on the warm and cool phases of the PDO and correlated to the Southern Oscillation Index (SOI) as a measure of the inter-annual climate phenomenon El Nino-Southern Oscillation (ENSO). All three basins have similar annual hydrographs with a peak in the winter around November due to direct winter runoff and a peak in the spring or summer due to snow runoff. The results show that in cool phases of the

PDO, seasonal streamflow is above average and is significantly correlated with the SOI. However, in warm phases of the PDO, streamflow is lower than normal and not as influenced by the SOI as measured by the correlation coefficient. Further, the PDO influences the distribution of flow within the year.

As a result, a new seasonal streamflow volume forecasting method is proposed. The new method fits regression equations for both phases of the PDO and mixes the two forecasts by the probability of the current state of the PDO. The model was verified by comparison to a control model that was fit to all of the data, and by a water year 2000 forecast. The results show that the mixed seasonal streamflow volume forecasts better estimate the historical mean. Further, the disaggregated mixed volume forecasts resulted in better estimates of the historical monthly mean and reduced the overall variability of the forecasts.

IMPROVED REGRESSION-BASED
STREAMFLOW FORECASTING CONSIDERING
LARGE-SCALE CLIMATE VARIABILITY

by

AUSTIN R. FISHER

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
CIVIL ENGINEERING

Portland State University
2000

ACKNOWLEDGMENTS

I would like to thank Dr. William Fish and Dr. Andrew Fraser for participating in my committee and providing the input that I needed to complete this document. I would also like to thank Dr. David Garen of the Natural Resources Conservation Service for his patience and time in helping me to gather most of the data used in this study. In addition, Dr. Garen helped me to understand how the NRCS model was intended to be used, and provided invaluable advice.

I would like to thank the United State Geological Survey for funding my research for the past 2 years. None of this would have been completed without there help and support.

I would also like to acknowledge the time, patience, and instruction given to me over the past 2 years by my advisor and committee chair, Dr. Roy Koch. I am thankful to him for giving me the opportunity to work on the project that has resulted in this thesis. I am also grateful for the time he devoted to this research and to my education.

Finally, I would like to thank my wife, Brandie, for her patience over the last 2 years. With her love and support I am able to dedicate the success of this thesis to her.

Table of Contents

Section	Page
List of Tables.....	v
List of Figures.....	vi
Chapter	
1. Introduction.....	1
1.1 Motivation	1
1.2 Current Practices	1
1.3 Outline.....	3
2. Background.....	4
2.1 Physically Based Hydrologic Forecasting.....	5
2.2 Statistical Hydrologic Forecasting.....	6
2.3 Large Scale Climate Variability.....	9
2.3.1 El Nino Southern Oscillation and the SOI.....	10
2.3.2 The Pacific Decadal Oscillation and Index.....	11
2.4 Use of Large-Scale Climate Features in Streamflow Forecasting.....	13
2.5 Observations in the Pacific Northwest.....	15
2.5.1 The Influence of the ENSO and the PDO on Pacific Northwest Streamflow.....	15
2.5.2 Intra-annual Variation in Reponse to the ENSO and the PDO.....	20

Table of Contents

(Continued)

2.5.3	Discussion of Pacific Northwest Observations.....	25
3.	Seasonal Streamflow Volume Forecast Model.....	26
3.1	Outline of Forecast Methodology.....	27
3.2	Forecasting the Sandy River nr. Marmot.....	28
3.2.1	Model Selection.....	29
3.3	Forecast Volume Results.....	32
3.3.1	Estimating the Future Phase of the PDO.....	36
3.3.2	Estimating α and β	37
3.3.3	Cross Validation Results.....	39
3.3.4	Forecast Mean and Evaluation of Bias.....	40
4.	Disaggregation of Volume Forecasts.....	43
4.1	Disaggregation Modeling.....	43
4.2	Disaggregating Volume Forecasts.....	48
4.2.1	Methods.....	48
4.2.2	Results from the Sandy River nr Marmot.....	49
4.2.3	Variability of Monthly Forecasts.....	56
5.	Forecasting Process and a Y2K Forecast.....	63
5.1	Methods.....	63
5.2	Results.....	64
6.	Summary and Conclusions.....	66

References.....	69
Appendix: Forecast Equations.....	73

List of Tables

Table	Page
1. Summary of data used in the study.....	16
2. p values for rank-sum tests on the fraction of flow occurring in each month for low and high PDO conditions. Values in bold are significantly different at the 5% level and values in italic at 10%.....	24
3. Data used as dependent and predictor variables to fit the 3 sets of equations with Reg.exe.....	31
4. Results of the 3 models fit to the Sandy River nr. Marmot data over the entire period of record (all data), high, and low PDO periods.....	33
5. Results of the F-tests performed on the standard errors.....	34
6. Contingency values predicting the phase of the PDO given the previous 1 to 4 year average of the PDO index. Oct-1 (+) indicates 1 previous year average from October to September, etc.....	37
7. Contingency tables showing the average results of the cross validation of the all data, above, and below equations respectively.....	39

List of Figures

Figures	Page
1. Relationship between predictors and streamflow volumes used in statistical regression equations.....	8
2. Map of the relationship of the SOI and October to March precipitation over the western U.S.....	9
3. Southern Oscillation Index (SOI) from 1901 to 1999.....	11
4. The PDO index over the last century. Persistent positive (negative) deviations are related to the warm (cool) phase of the PDO.....	13
5. Hydrographs for the three study basins in the Pacific Northwest. (a) the Skykomish River, (b) the Sandy River, and (c) the Rogue River.....	17
6. Correlation of 3-month average SOI starting with the April – June average that precedes the November to June Skykomish streamflow volume.....	19
7. Correlation of 3-month average SOI starting with the April – June average that precedes the November to June Sandy streamflow volume.....	19
8. Correlation of 3-month average SOI starting with the April – June average that precedes the November to June Rogue streamflow volume.....	20
9. Composite monthly hydrographs for the Sandy River nr Marmot. The data were composited by PDO and again by the SOI. La Nina is represented by values of the SOI greater than 0.5 and El Nino is represented by values of the SOI less than -0.5.....	22
10. Sandy River nr Marmot fraction of annual sum composites by high and low PDO periods.....	23
11. Sandy River nr Marmot fraction of annual sum composites for the high PDO period by SOI.....	23
12. Sandy River nr Marmot fraction of annual sum composites for the low PDO period by SOI.....	24

List of Figures

(continued)

Figures	Page
13. Jackknife standard error for the high and low PDO equations and the re-computed JSE in the high and low PDO periods of the all data equations.....	35
14. Standard error for the high and low PDO equations and the re-computed JSE in the high and low PDO periods of the all data equations.....	35
15. Alpha and Beta probabilities as estimated by the average of the PDO index over 1 to 4 previous years from October to September.....	38
16. Comparison of the mean forecast during cool PDO phase. Split data are the warm and cool phase equations that have been mixed by the PDO.....	41
17. Comparison of the mean forecast during warm PDO phase. Split data are the warm and cool phase equations that have been mixed by the PDO.....	41
18. Schematic illustration of the disaggregation terminology and the relationship of the volume forecast to the disaggregated volumes.....	45
19. Disaggregated mean monthly flow from October to September in acre-feet during the Cool Phase of the PDO (low PDO index).....	50
20. Disaggregated mean monthly flow from October to September in acre-feet during the Warm Phase of the PDO (high PDO index).....	51
21. Average Disaggregated mean monthly flow from October to September in acre-feet.....	51
22. Disaggregated mean monthly flow from November to September in acre-feet during the Cool Phase of the PDO (low PDO index).....	52
23. Disaggregated mean monthly flow from November to September in acre-feet during the Warm Phase of the PDO (high PDO index).....	52
24. Average Disaggregated mean monthly flow from November to September in acre-feet.....	53
25. Disaggregated mean monthly flow from December to September in acre-feet during the Cool Phase of the PDO (low PDO index).....	53

List of Figures

(continued)

Figures	Page
26. Disaggregated mean monthly flow from December to September in acre-feet during the Warm Phase of the PDO (high PDO index).....	54
27. Average Desegregated mean monthly flow from December to September in acre-feet.....	54
28. Desegregated mean monthly flow from January to September in acre-feet during the Cool Phase of the PDO (low PDO index).....	55
29. Desegregated mean monthly flow from January to September in acre-feet during the Warm Phase of the PDO (high PDO index).....	55
30. Average Desegregated mean monthly flow from January to September in acre-feet.....	56
31. Variability by month for the October to September volume during the cool phase of the PDO (low PDO index).....	58
32. Variability by month for the October to September volume during the warm phase of the PDO (high PDO index).....	58
33. Variability by month for the November to September volume during the cool phase of the PDO (low PDO index).....	59
34. Variability by month for the November to September volume during the warm phase of the PDO (high PDO index).....	59
35. Variability by month for the December to September volume during the cool phase of the PDO (low PDO index).....	60
36. Variability by month for the December to September volume during the warm phase of the PDO (high PDO index).....	60
37. Variability by month for the January to September volume during the cool phase of the PDO (low PDO index).....	61

List of Figures (continued)

Figures	Page
38. Variability by month for the January to September volume during the warm phase of the PDO (high PDO index).....	61
39. Year 2000 Forecast made by mixing warm and cool PDO disaggregated seasonal streamflow volume forecasts.....	65
A1 Variables used in the 10 control forecast equations. Dependent variables are in the far left column and the location of the variables is located in the column headings.....	73
A2 Variables used in the 10 warm PDO forecast equations. Dependent variables are in the far left column and the location of the variables is located in the column headings.....	73
A3 Variables used in the 10 cool PDO forecast equations. Dependent variables are in the far left column and the location of the variables is located in the column headings.....	74

Chapter 1

Introduction

1.1 Motivation

Public agencies and the private power industry use seasonal streamflow volume forecasts to manage much of our nation's surface water supply. For instance, many private corporations in the power industry manage and utilize river basins in the Pacific Northwest and elsewhere in the western U.S. to generate hydropower. Hydropower producers operate with many socio-political constraints as well as under climatological uncertainty that results in a variable water supply. A common risk-based decision faced by these companies is how much water to release to prevent floods during the wet season while maximizing the storage in the reservoirs for power production during low flow times of the year. Forecasts based mostly on statistical regression techniques using variables such as precipitation, temperature and snow water equivalent are employed to predict seasonal volumes of flow. The use of statistical techniques allows for the quantification of the risk associated with the forecasts and thus the risk of forecast based decisions.

1.2 Current Practices

Currently the National Weather Service Climate Prediction Center (NWS-CPC) produces long-lead climate outlooks, and the Natural Resources Conservation Service National Water and Climate Center (NRCS-NWCC) in cooperation with the National Weather Service River Forecast Centers produce

streamflow forecasts over much of the western United States. The NWS-CPC long-lead climate outlooks and NRCS seasonal streamflow volume forecasts make use of a relationship between the large-scale climate feature, the El Niño Southern Oscillation (ENSO), and climate in the western U.S. The Southern Oscillation Index (SOI) is an index of ENSO and has already been integrated into climate and streamflow forecasting where possible.

Recently, there have been efforts by Garen (1998) and Modini (2000) to integrate the NWS-CPC long-lead climate outlooks and statistical seasonal streamflow forecasting. Further, researchers at the University of Washington have successfully integrated two large-scale climate features, the ENSO and the Pacific Decadal Oscillation (PDO), into physical streamflow forecasting procedures (Hamlett and Lettenmaier 1999).

While Hamlet and Lettenmaier (1999) were able to successfully integrate current information from large-scale climate features into streamflow forecasting, the physical scale of the model is coarse, typically at a resolution of 1° latitude and longitude. The spatial limitations of forcing physical hydrologic forecasting models with coarse climate models limits their use for now and suggests that statistical forecast models will continue to provide the most reliable information for the time being. In this study, current forecasting methods will be expanded to make use of the ENSO as measured by the SOI and the PDO as measured by the PDO index.

The NRCS-NWCC in Portland, Oregon uses regression on principal components to fit forecast equations for streams in the western U.S. The current methods used by the NRCS-NWCC will be extended to fit forecast equations that account for the influence of the ENSO and the PDO on streamflow. Further, the volume forecasts produced by the NRCS-NWCC forecast equations will be disaggregated into sub-seasonal volumes of one month.

Water supply managers operate reservoirs and rivers in real-time, therefore the usefulness of seasonal volume forecasts limits planning. The use of a disaggregation model to distribute the seasonal volume throughout its monthly constituents will enable water supply managers to make better risk-based decisions.

1.3 Outline

The following chapter briefly describes the history of water supply forecasting, current practices, and concludes with some recent observations of climate and streamflow in the Pacific Northwest. After that, Chapter 3 describes a new seasonal streamflow forecasting method and compares the performance of the new approach to that of current statistical modeling practices used by the NRCS. Chapter 4 details the disaggregation models and again compares the disaggregation of the new forecast model volumes of flow with the volumes produced using current methods. Finally, the process for applying the new forecast methodology is discussed, and a number of conclusions are drawn based on those results.

Chapter 2

Background

The Soil Conservation Service (SCS, now known as the NRCS) was the first agency chartered to collect data on snow and provide forecasts of future water supplies (CRWMG 1993). In the 1950's the U.S. Army Corps of Engineers developed some of the first computer system models for streamflow forecasting by improving the field of snow hydrology as we know it today (CRWMG 1993). Streamflow forecasting can now be separated into two approaches based on the type of models: physical and statistical.

Physical-based models such as SSARR (Streamflow Synthesis and Reservoir Regulation) developed by the Corps of Engineers mathematically represent the hydrologic processes and attempt to “simulate” the movement of water through a system. The Extended Streamflow Prediction (ESP) method developed by the National Weather Service uses the combination of a conceptual watershed model and statistical properties of historical climate time series to produce a forecast. Sequences of historic climate data are then used to develop an ensemble of outflow hydrographs.

Statistical models, such as regression, have also been widely used to provide volume forecasts of streamflow by expressing future streamflow as a conditioned response to prior hydrologic events such as snow accumulation.

Thereby correlation between hydrologic variables such as antecedent flow, snow pack, and precipitation are used to predict future volumes of streamflow.

2.1 Physically Based Hydrologic Forecasting

Physically based hydrologic models make use of mathematical representations of the hydrologic system that compute streamflow as the output given precipitation as the input. The Sacramento Soil Moisture Accounting Model (SAC-SMA) for instance uses a two-layer representation of the soil-water system to distribute water between the surface and subsurface and a unit hydrograph to convert surface runoff into streamflow. The upper layer of the SAC-SMA model represents surface and interception storage, while the second layer represents soil storage or groundwater. The inputs to the model are mean areal precipitation (MAP) and potential evapotranspiration (PET). Models such as the SAC-SMA can be used for hydrologic forecasting if calibrated successfully to a particular watershed and reasonable estimates of MAP are made. Therein lies one limitation of physical hydrologic modeling. The estimates of MAP used to calibrate physical models are often estimated from point measurements made by the many rain gages across the U.S. The interpolation methods used to estimate MAP from point measurements can often generalize the natural variability of MAP over a watershed due to land cover and orographic variation. However, there have been improvements in estimating MAP (Garen et al. 1994).

Another limitation of physical hydrologic forecasting models is downscaling. General circulation models (GCM's) have been used to predict

climate and to force regional scale hydrologic models (Liang et al. 1994). The typical spatial resolution of GCM's is on the order of 1⁰ latitude and longitude, which generalizes the natural spatial variability of climate and therefore the outflow hydrograph.

2.2 Statistical Hydrologic Forecasting

Statistical hydrologic forecasting makes use of significant statistical relationships between seasonal streamflow volumes and measured hydrologic variables such as snow water equivalent (SWE), winter precipitation, and antecedent streamflow to fit regression equations that can then be used to forecast future volumes of flow. Statistical modeling includes a representation of uncertainty, and therefore it is possible to quantify the risk of the forecasts.

The NRCS-NWCC uses winter precipitation, SWE, and antecedent flow to predict various seasonal streamflow volumes starting in January. The NRCS produces these forecasts using regression equations fit to different seasonal volumes for the many streams in the U.S. SWE is the primary predictor used in regression equations employed by the NRCS. Prior to the use of the SNOTEL system, SWE measurements used to be available the 1st January when snow courses were measured and is subsequently the reason for forecasts being made at that time.

The NRCS currently uses regression on principal components described by Garen (1992) and others (Haan and Allen 1972, Haan 1977, McCuen 1985). In addition, Garen (1992) showed that a near optimal set of predictors should be used when fitting the regression equations such that the equations represent the

maximum accuracy that can be obtained by statistical methods. Principal component analysis is a statistical method that inherently assures that the variables used in the regression equations are independent. Inter-correlation of predictor variables is common in statistical hydrologic models because most of the predictors are spatially correlated at any given time.

Redmond and Koch (1991) showed that large-scale climate indices such as the Southern Oscillation Index (SOI) could be used as predictors in some areas of the Western U.S. Further, Garen (1992) and Koch and Buller (1993) have shown that the use of SOI can significantly improve forecasting skill and extend the lead-time when forecasts are produced by as much as 3 months or as early as October.

However, the NRCS still produces forecasts starting 1 January for volumes of seasonal streamflow. For instance, a common forecast made by the NRCS on 1 March would be for the April-September volume. The April-September forecast equation would typically include the average July to September or July to October SOI, antecedent streamflow, SWE, and in some instances, winter precipitation. In general, the predictors used in a seasonal streamflow volume forecast are observed prior to the forecast date. Therefore, only the late summer to fall SOI, winter precipitation through December, antecedent streamflow, and the January 1st SWE are available for a forecast made on January 1st (see Figure 1).

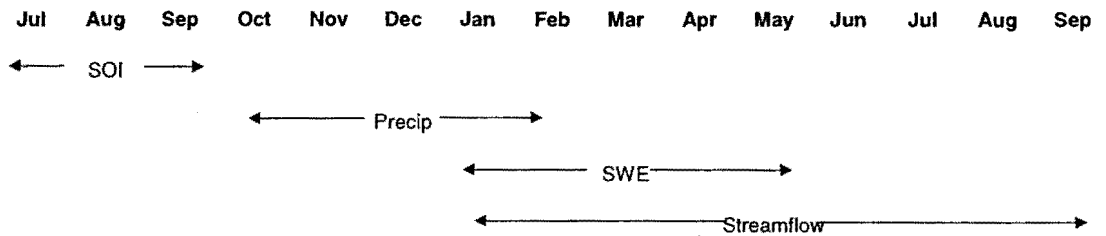


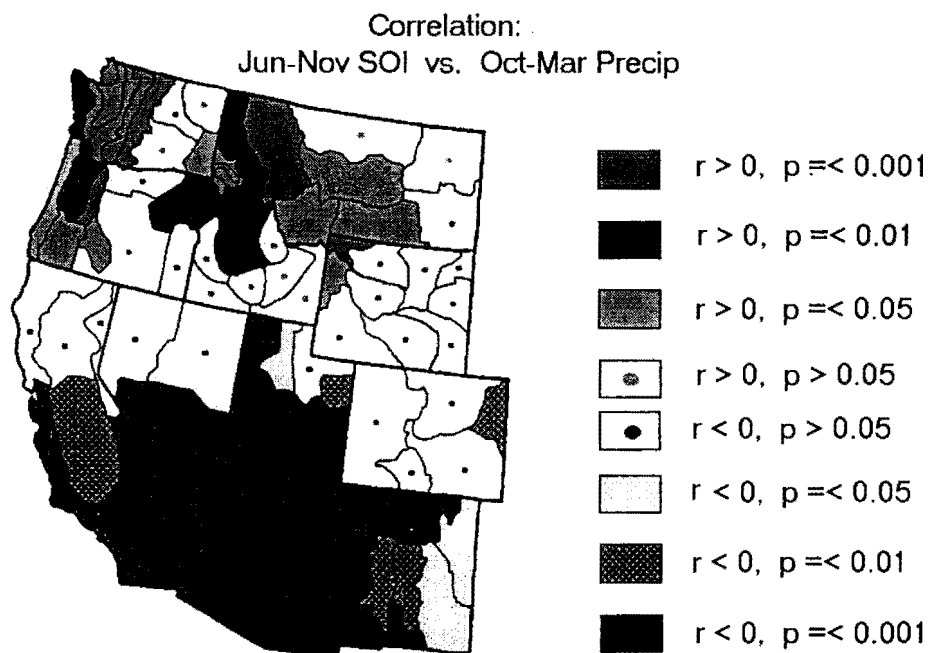
Figure 1: Relationship between predictors and streamflow volumes used in statistical regression equations.

The inclusion of large-scale climate features such as the El Nino Southern Oscillation into streamflow forecasting has increased the lead-time of forecasts because prior to its use there was little information available before the January 1st SWE measurement.

Much of the success in forecasting with statistical regression equations has been limited to seasonal volume forecasting over several months. Time resolutions as fine as one month, however, can be obtained from the combination of a seasonal streamflow volume forecast model and a disaggregation model. The seasonal streamflow volume forecasting model can be used to generate several volume forecasts that can be disaggregated into monthly values. The disaggregation model preserves the historical statistical properties between the volume or key series and the monthly or sub-key series (Valencia and Schaake 1972, Mejia and Rousselle 1976, Lane 1982, and Pei et al. 1987). To date, disaggregation has been primarily used in synthetic streamflow data generation, but it can also be used for forecasting purposes (Salas et al. 1980).

2.3 Large Scale Climate Variability

The use of the SOI in streamflow forecasting is the first direct use of large-scale climate features in producing volume forecasts. Redmond and Koch (1991) demonstrated that precipitation in the western U.S. is significantly related to large-scale climate features in the tropics such as the El Nino Southern Oscillation (ENSO) as shown by Figure 2. Since then several studies have been published that relate ENSO to western U.S. climate and streamflow (Cheiw et al. 1998, Kahya et al. 1993, Koch et al. 1991, and Piechota 1997).



Updated from Redmond and Koch (1991). Winters of 1933/34 - 1994/95.
 Reddish: El Nino associated with wet winters, La Nina with dry winters.
 Bluish/greenish: El Nino associated with dry winters, La Nina with wet winters.

Redmond, K.T., and R.W. Koch, 1991. Surface climate and streamflow variability in the western United States and their relationship to large-scale circulation indices. *Water Resources Research*, 27(9), 2381-2399.

Figure 2: Map of the relationship of the SOI and October to March precipitation over the western U.S.

2.3.1 El Nino Southern Oscillation and the SOI

The El Nino Southern Oscillation (ENSO) is characterized by a change in the Sea Surface Temperatures (SST) and Sea Level Pressure (SLP) over time in the equatorial waters of the Pacific Ocean. First discovered by Sir Gilbert Walker, the ENSO is recognized as a periodic weakening (El Nino) and intensification (La Nina) of the “normal” SLP difference between Tahiti and Darwin, Australia. During “normal” conditions the SST’s in the eastern end of the Pacific off the coast of South America are cool and gradually warm toward the western end of the Pacific off the coast of Australia. Further, during “normal” and La Nina conditions the climate pattern termed the “Walker Cell” is fully developed with convective rising of air in the western end of the Pacific near Australia and is the source of the dominant wind pattern called the Westerlies (Rasmusson 1985, Trenberth 1997). The exact opposite occurs during El Nino conditions: the SST’s warm to the east towards the coast of South America and weaken the Walker Cell and the Westerlies, leading to major Equatorial weather pattern changes at both ends of the Pacific.

A measure of the ENSO is the Southern Oscillation Index (SOI), which is a standardized index of the SLP difference between Tahiti and Darwin, Australia. The SOI is negative during El Nino and positive during La Nina. The July to September average SOI time series since 1901 is shown in Figure 3. The SOI is one of the common indexes used as a measurement of the state of the ENSO and is also a common variable in long lead climate and streamflow forecasting.

July - September Average SOI

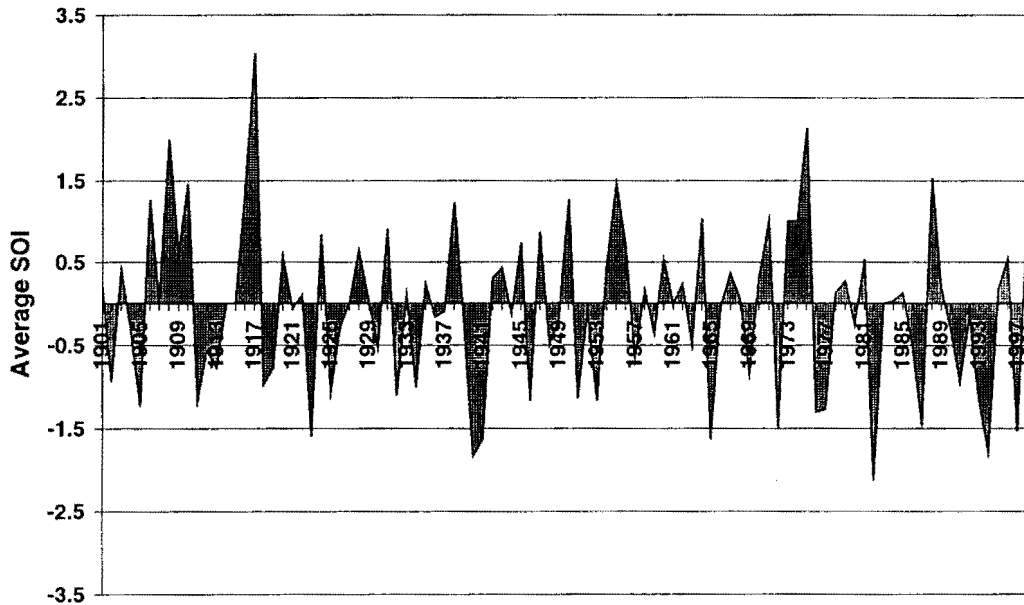


Figure 3: Southern Oscillation Index (SOI) from 1901 to 1999.

2.3.2 The Pacific Decadal Oscillation and Index

Recently another large-scale climate pattern has been observed and labeled the Pacific Decadal Oscillation (PDO). It is quantitatively represented by the PDO index (Mantua et al. 1997). The PDO pattern has been described as a low frequency oscillation in the SST's measured over the past century in the Pacific Ocean north of 20°N latitude. Unlike the inter-annual oscillation of the SOI, the PDO oscillates over two or more decades between warm and cool Alaskan SST's. Decades of warmer than normal waters off the coast of Alaska are "warm" PDO conditions and are represented by positive values of the PDO index, whereas decades of cooler than normal SST's off the coast of Alaska are termed the "cool"

PDO conditions and are represented by negative values of the PDO index (Figure 4). For the remainder of this document, high PDO refers to the relatively high values of the PDO index and will be synonymous with the warm phase of the PDO and vice versa.

The PDO has been a persistent decadal pattern of North Pacific SST's over the past century marked by shifts from below average conditions to above average conditions in 1925 and 1977 and vice versa in 1947 (Mantua et al. 1997). Mantua et al. (1997) specified the shifts in the PDO by an intervention analysis performed on the PDO index. Subsequently, the PDO index has been shown to be significantly related to Pacific Northwest climate and streamflow (Mantua et al. 1997, Koch and Fisher 2000).

Other research has looked at the relationship of the PDO and ENSO and how the two patterns may simultaneously influence climate. Particularly, it has been found that divisional precipitation and flood frequency are influenced more by ENSO during cool phases of the PDO or below average periods of the PDO index (McCabe and Dettinger 1998).

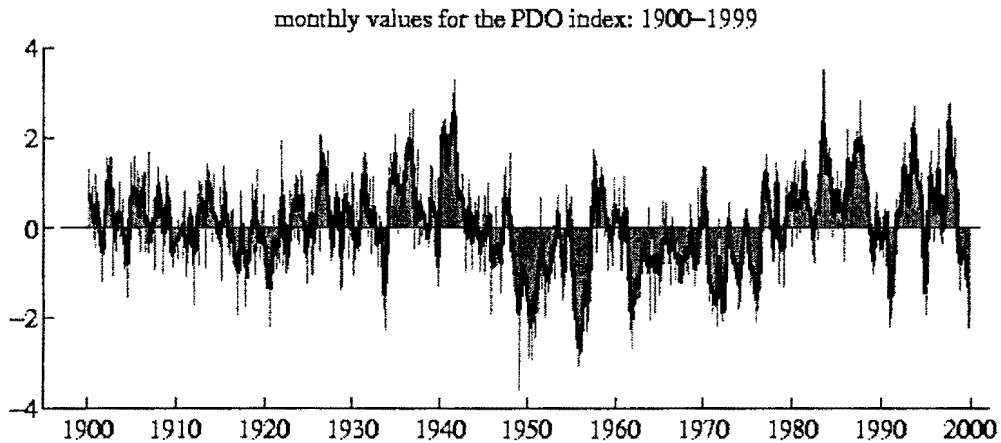


Figure 4: The PDO index over the last century. Persistent positive (negative) deviations are related to the warm (cool) phase of the PDO.

2.4 Use of Large-Scale Climate features in Streamflow Forecasting

Customarily, the NRCS issues seasonal streamflow volume forecasts for many western U.S. streams beginning in January. Lead-time and temporal resolution are the two factors that limit the use of statistical streamflow forecasts in water supply management. Although average SOI can provide sufficient lead-time, in many basins such as the Willamette in Oregon, the relationship to streamflow is not significant.

Modini (2000) followed the work of Garen (1998) in integrating the National Weather Service Climate Prediction Center's (NWS-CPC) long lead climate outlooks into statistical streamflow forecasting in an effort to extend the lead time of his forecasts of the Blue River in the Willamette basin, Oregon. Modini (2000) concluded that the use of the CPC outlooks significantly reduced the standard error of forecasts of the Blue River, particularly early on in the forecast

season. Modini (2000) and Garen (1998) have both shown that, at least indirectly, information from large-scale climate features such as the SOI can be integrated into streamflow forecasting on a regional basis.

There has been further progress towards including predictive information from large-scale climate features like the ENSO and the PDO in physical hydrologic modeling (Croley 1996, Hamlet and Lettenmaier 1999). Hamlet and Lettenmaier (1999) used the NWS-ESP method in conjunction with the 2-layer VIC model. Their study made use of the PDO and ENSO in assembling forecast ensembles produced from the historic record as the resulting streamflow from various groups of climate states. In this instance, there were 6 ensembles that corresponded to the warm, neutral, and cold states of ENSO (high, moderate, and low SOI respectively), and the warm and cool phases of the PDO respectively. The results of the Hamlet and Lettenmaier (1999) study show that the PDO could be successfully integrated into existing physical streamflow modeling schemes and used in streamflow forecasting.

The fine temporal resolution of physical models will continue to support its development. Hamlet and Lettenmaier (1999) have shown that large-scale climate features can be integrated into the ESP modeling approach by conditioning the forecast ensembles on states of the PDO and ENSO. The discrepancies in the spatial resolution of the climate models used to force physical hydrologic models such as VIC or SAC-SMA, however, still remain.

2.5 Observations in the Pacific Northwest

Koch and Fisher (2000) showed that the relationship between ENSO as measured by the SOI and streamflow is not consistent over time. Rather, the relationship between the SOI and seasonal streamflow in three basins in the Pacific Northwest as measured by the correlation coefficient is much higher during the cool phase of the PDO (prior to 1925, and 1947 – 1976) than during warm PDO phases (1925 – 1946 and 1977 – Present). The results support those found by McCabe and Dettinger (1998) and are summarized below.

2.5.1 The influence of the ENSO and the PDO on Pacific Northwest streamflow

Data from three rivers in the western slopes of the Cascade Mountains in Washington and Oregon were gathered to determine if the observations found by McCabe and Dettinger (1998) for precipitation are consistent with streamflow in the Pacific Northwest. The Skykomish River (Washington Cascade Mountains) and the Sandy and Rogue Rivers (Oregon Cascade Mountains) have relatively high flow in both the winter and in the spring. The records date back to 1912, 1924, and 1925 for the Sandy, Rogue, and Skykomish respectively allowing for decadal trends to be analyzed over much of the past century. These three basins were selected because of the length of record and lack of upstream diversions. Figure 5 shows the average monthly hydrographs for the study basins, and Table 1 presents information about the period of record and the watershed characteristics of the basins.

Table 1: Summary of data used in the study

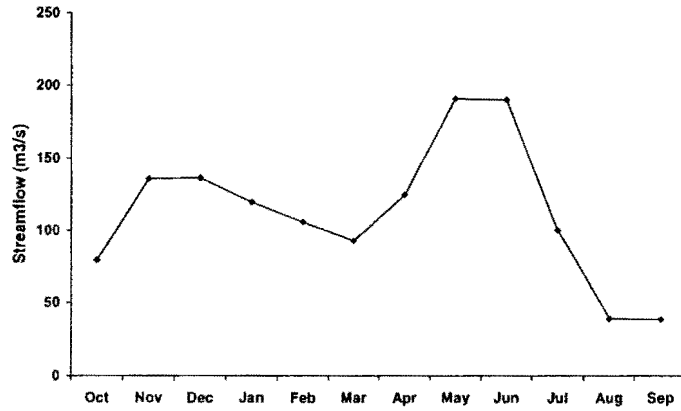
Station/Data	USGS Station Number	Period of Record	Drainage Area (km ²)	Missing data	Gage Elevation (m)	Average Annual Flow (m ³ /s)
Skykomish	12134500	1929-98	1386	none	232	112
Sandy	14137000	1912-98	681	1916, 1919	223	38
Rogue	14328000	1924-98	808	none	799	23
SOI ¹		1882-98				
PDO ²		1900-98				

¹Available at <http://www.cpc.ncep.noaa.gov>

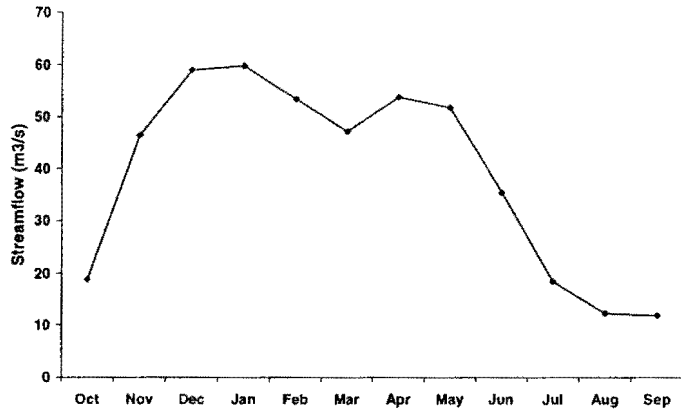
²Available at ftp://ftp.atmos.washington.edu/mantua/pnw_impacts/INDICES/PDO.latest

The November to June streamflow volume for the three basins was selected for the analysis because that period encompasses the bulk of the annual flow and includes both the precipitation driven winter runoff and the snow driven spring runoff (Figure 6). In addition, forecasting streamflow in November and December has previously been difficult in basins in the Pacific Northwest because precipitation is fairly unpredictable, and appreciable amounts of snow data are not usually available until January, at which point the November and December flows have already been observed.

(a) Skykomish River, WA: Average Monthly Streamflow



(b) Sandy River, OR: Average Monthly Streamflow



(c) Rogue River, OR: Average Monthly Streamflow

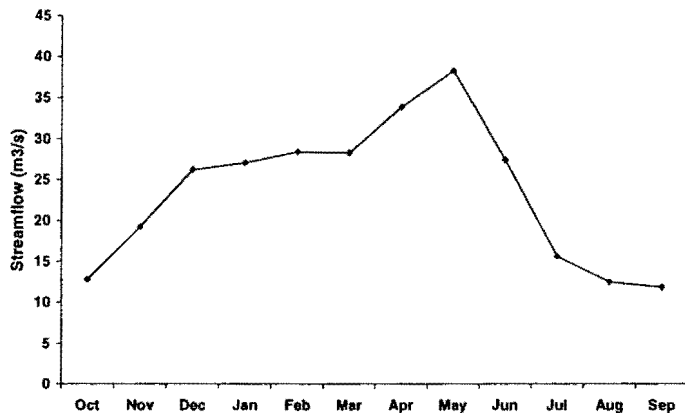


Figure 5: Hydrographs for the three study basins in the Pacific Northwest. (a) the Skykomish River, (b) the Sandy River, and (c) the Rogue River.

Recent research by McCabe and Dettinger (1998) showed that the correlation of SOI and divisional precipitation in the Pacific Northwest is not consistent over time. During predominantly cool phases of the PDO the correlation to precipitation is much higher than during the warm phases of the PDO. Therefore, the data were split into two groups: one for the warm phase of the PDO and one for the cool phase of the PDO. The data were then correlated with moving 3-month averages of the SOI starting with the April – June average preceding the streamflow volume, and the results are shown in Figures 6 – 8. Strikingly, the correlation between the SOI and the various seasonal volumes is much stronger during the cool phase of the PDO than it is during the warm phase of the PDO. These observations are consistent with those of McCabe and Dettinger (1998) for precipitation.

The teleconnection between the ENSO process as described by the SOI appears to be much stronger during the the cool phase of the PDO than during the warm phase of the PDO. The effect of conditioning by the PDO is the most dramatic in the Rogue basin, where the moving correlation coefficients between the SOI and the November to June volume is statistically insignificant when computed for all the data and for the warm phase of the PDO. The relationship is quite strong, however, in the cool phase of the PDO with correlation coefficients as high as 0.6. Furthermore, in all of the basins shown here (Figures 6 – 8), the correlation coefficients for the cool phase of the PDO is much higher than that for the warm phase of the PDO and for all years taken together.

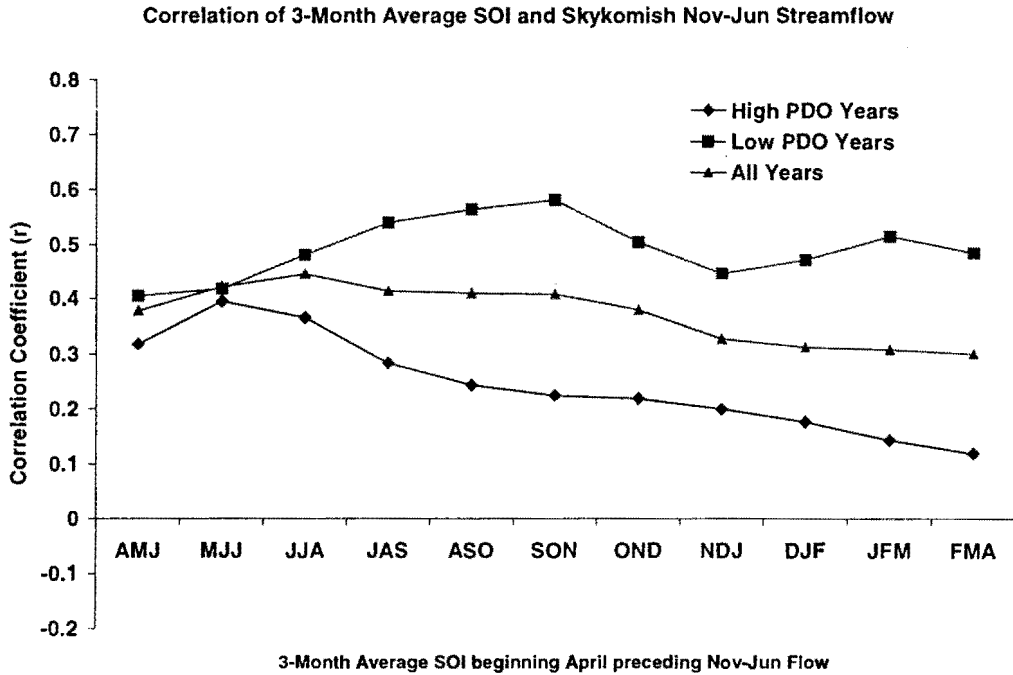


Figure 6: Correlation of 3-month average SOI starting with the April – June average that precedes the November to June Skykomish streamflow volume.

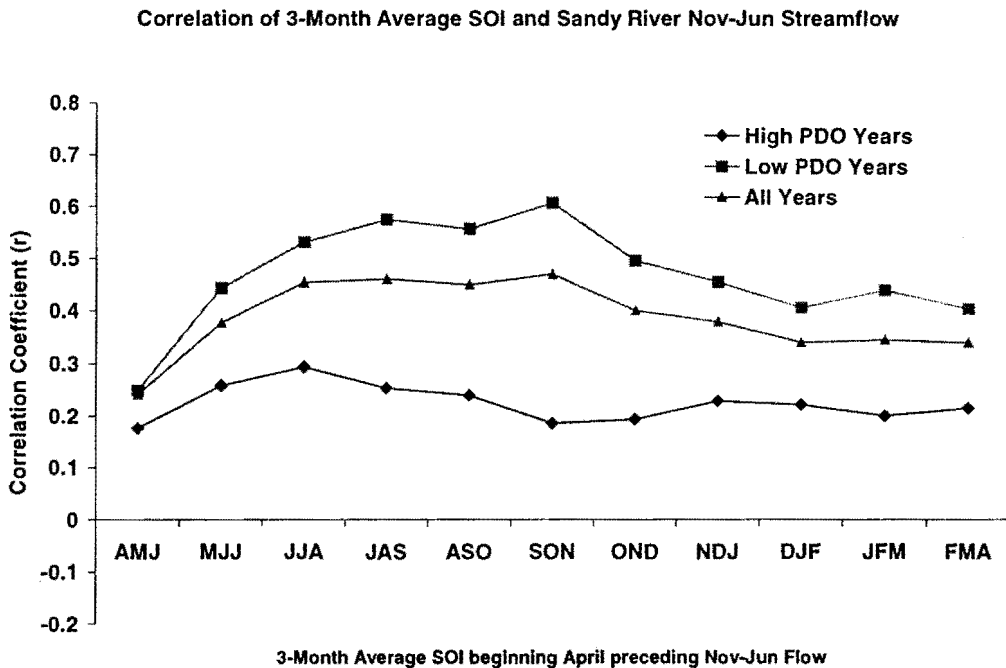


Figure 7: Correlation of 3-month average SOI starting with the April – June average that precedes the November to June Sandy streamflow volume.

Correlation of 3-Month Average SOI and Rogue River Nov-Jun Streamflow

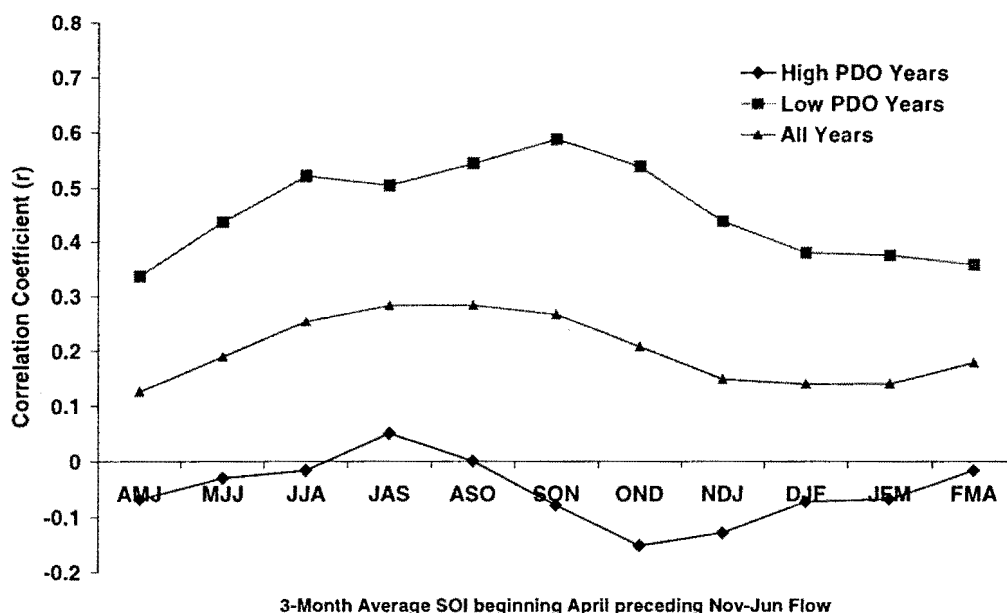


Figure 8: Correlation of 3-month average SOI starting with the April – June average that precedes the November to June Rogue streamflow volume.

Figures 6 – 8 show that the SOI is not a significant predictor of streamflow in the Sandy, Rogue, and Skykomish basins when the PDO is in the warm phase. However, they also show that the SOI is a significant predictor during the cool phase of the PDO. Previously the SOI would not have been considered a significant predictor of streamflow in either the Skykomish or Rogue basins. The influence of ENSO on western U.S. climate may indeed be more comprehensive than once thought.

2.5.2 Intra-annual variation in response to the ENSO and the PDO

Another aspect of streamflow variability analyzed by Koch and Fisher (2000) was the intra-annual (within year) variation on a monthly time scale. Most would agree that seasonal volumes of flow are only useful for long-term planning,

while monthly forecasts, if available, would significantly reduce the planning horizon. Therefore, disaggregation models will be used to distribute the predicted volumes of streamflow over their respective series of months. Koch and Fisher (2000) found that the distribution of flow within the year varies with the PDO.

As an example, Figure 9 is a composite hydrograph for the Sandy River. What is clear from Figure 9 is that the flow is highest during La Nina's that occur in the cool PDO phase, however what is not clear is how different the intra-annual variation in these periods are. Therefore, the data from each basin were manipulated such that each monthly observation was converted to the fraction it represented of the annual sum. The data were then split by the PDO to ascertain how, on average the distribution of flow varied throughout the year in the two PDO phases. The results are plotted in Figure 10.

The distribution was found to be different, particularly in October, January, March, and April. In addition, the peaks occur in December and April during high PDO periods versus peaks in January and May during low PDO periods. Although the plots are not shown, the results of the Mann-Whitney test on the monthly fractions from the Skykomish and Rogue rivers also indicate significant differences in the inter-annual distribution of flow given the state of the PDO. The results of these tables are presented in Table 2.

Figures 11 and 12 are similar to Figure 10, however, the conditioning is by high and low SOI within warm and cool PDO periods. Based on these analyses, ENSO does not provide any additional information regarding the disaggregation of

the volume forecasts. No statistically significant differences were found in the fractions in Figures 11 and 12.

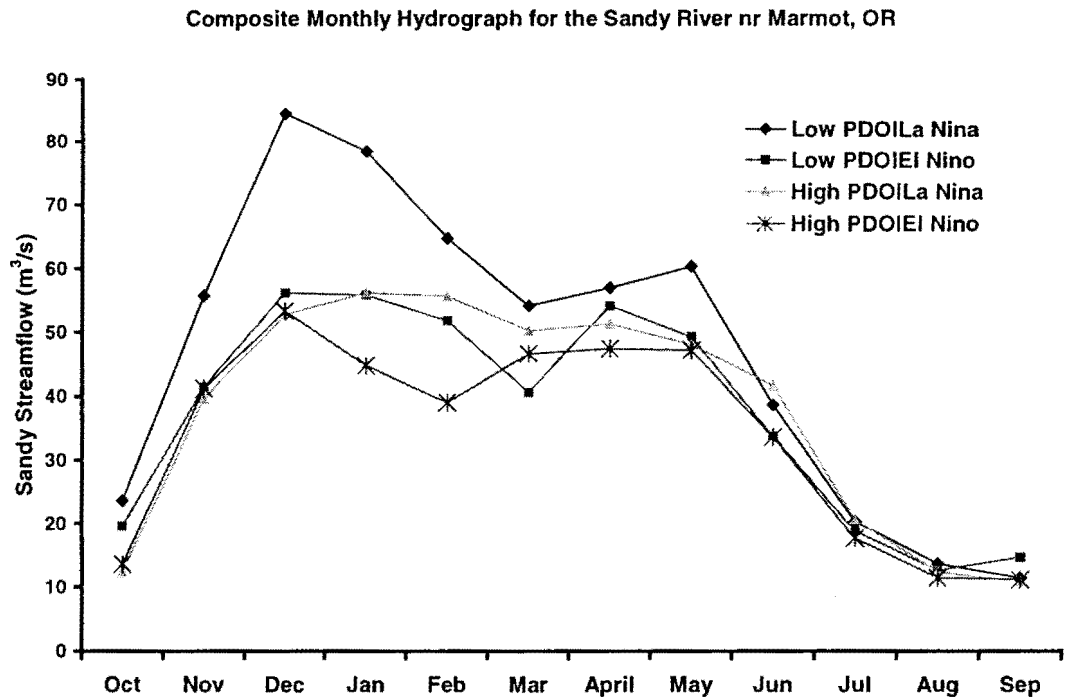


Figure 9: Composite monthly hydrographs for the Sandy River nr Marmot. The data were composited by PDO and again by the SOI. La Nina is represented by values of the SOI greater than 0.5 and El Nino is represented by values of the SOI less than -0.5.

Monthly fraction of annual flow for the Sandy River

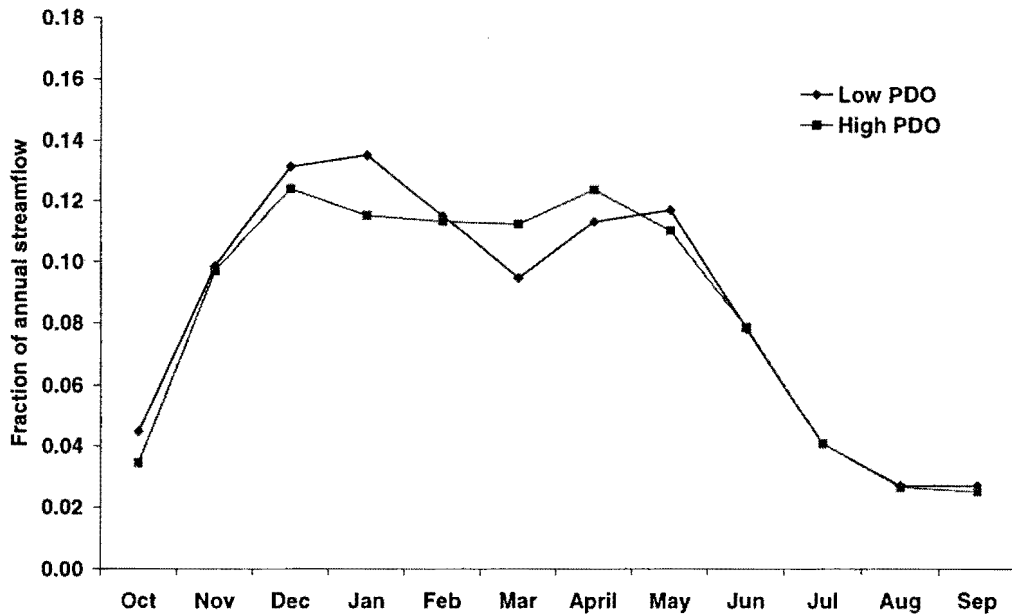


Figure 10: Sandy River nr Marmot fraction of annual sum composites by high and low PDO periods.

Sandy River - High PDO

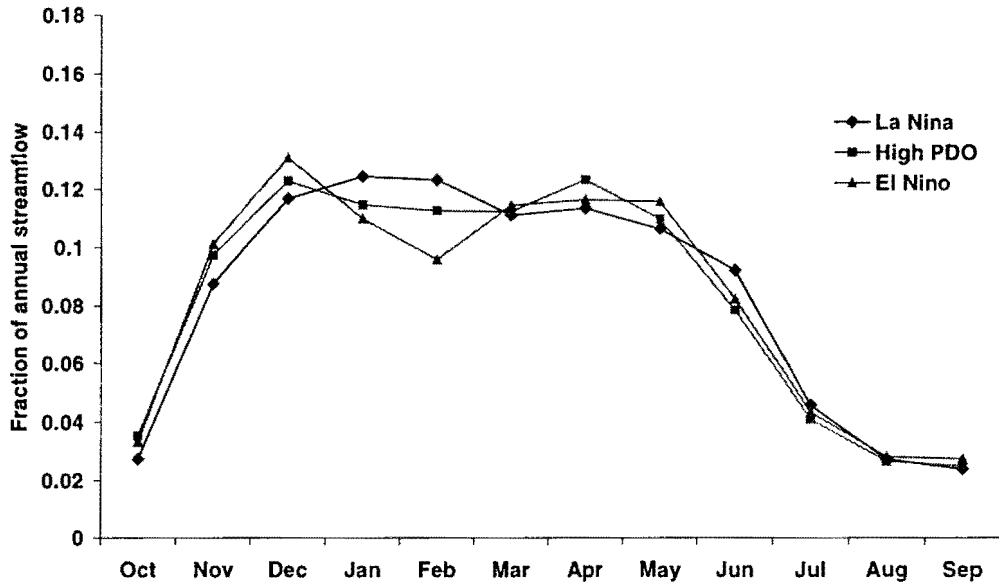


Figure 11: Sandy River nr Marmot fraction of annual sum composites for the high PDO period by SOI.

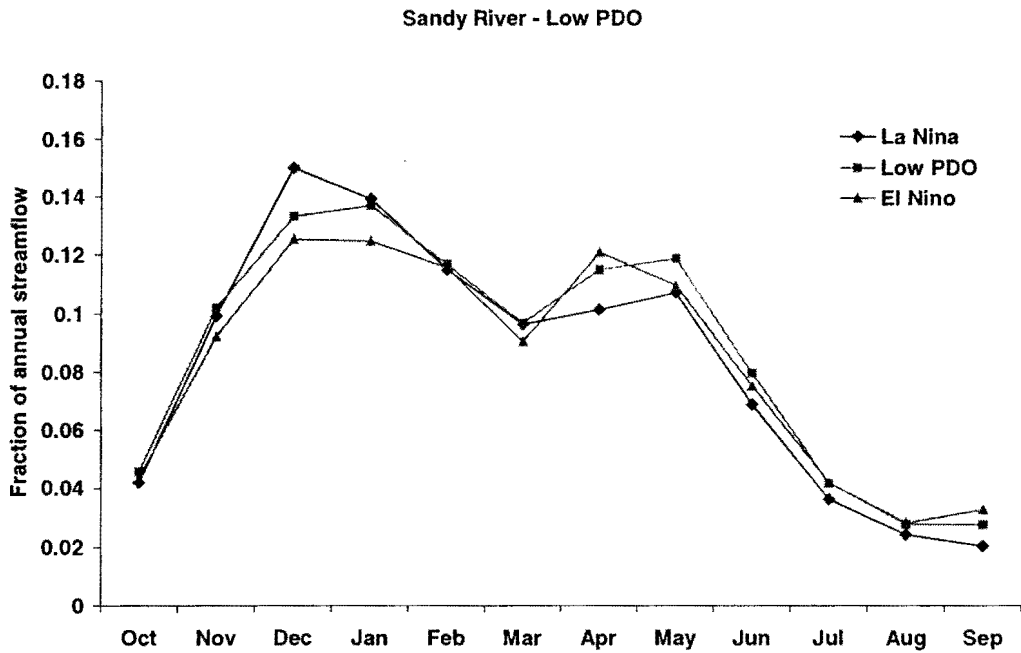


Figure 12: Sandy River nr Marmot fraction of annual sum composites for the low PDO period by SOI.

Month	Skykomish	Sandy	Rogue
Oct	0.08	0.03	0.26
Nov	0.80	0.75	0.20
Dec	0.83	0.19	0.70
Jan	0.99	0.05	0.97
Feb	0.74	0.61	0.97
Mar	0.00	0.01	0.01
Apr	0.00	0.04	0.04
May	0.53	0.59	0.43
Jun	0.01	0.72	0.28
Jul	0.00	0.28	0.87
Aug	0.00	0.93	0.07
Sep	0.21	1.00	0.50

Table 2. p values for rank-sum tests on the fraction of flow occurring in each month for low and high PDO conditions. Values in bold are significantly different at the 5% level.

2.5.3 Discussion of Pacific Northwest Observations

The Hamlet and Lettenmaier (1999) study suggested that the PDO and the ENSO are important indicators of streamflow in the Pacific Northwest: in this case the Columbia River Basin. The results from McCabe and Dettinger (1998) and Koch and Fisher (2000) establish a statistical conditioning of the influence of the ENSO on seasonal streamflow in the Pacific Northwest by the PDO. Therefore, a new seasonal streamflow volume forecasting method is needed to include this new climate information.

A new method is proposed in the following chapter that suggests conditioning forecasts on both the warm and cool phases of the PDO, and then mixing those forecasts based on an estimate of the probability of the current phase of the PDO the forecast is made in. In addition, Koch and Fisher (2000) suggest that the forecasted volumes should be disaggregated consistently with the phase of the PDO. In other words, a disaggregation model should be fit for both the warm and cool phases of the PDO and be used to disaggregate its respective volume of flow.

Chapter 3

Seasonal Streamflow Volume Forecast Model

As a result of the observations discussed in Chapter 2 a new seasonal streamflow volume forecast model methodology is developed and presented below. The NRCS forecast model, REG, will be used to fit the forecast equations and will simply be referred to as the NRCS model for the remainder of the document. The new methodology will incorporate both the PDO and the ENSO into existing regression-based statistical streamflow forecasting model development techniques.

The forecast methodology will be to use existing statistical methods to fit two near optimal forecast equations, one for each phase of the PDO. Prior to fitting the models the data are split based on the phase of the PDO (1900 – 1924 & 1947 – 1976 (Cool Phase) and 1925 – 1946 & 1977 – 1998 (Warm Phase)). Models are then developed to predict volumes of flow for each PDO phase. Validation of the observations outlined above is possible if a model is developed from all the data and compared to the models developed for the two phases of the PDO.

There are three advantages in developing a new forecasting method consistent with the observations outlined above.

1. The new method will better represent the data because of the discrepancy found in using SOI as a predictor of streamflow at all times. This should result in a better estimate of the uncertainty of all seasonal streamflow volume forecasts.

2. The new method will produce forecasts that include the winter (early) runoff, previously not available due to the unpredictability of winter precipitation. Further, the new method will make use of the SOI as a long lead indicator in basins where it was previously not used.
3. Seasonal volumes can be disaggregated to produce forecasts of monthly flows. These will be used in a decision support system to help make optimal water resource decisions.

The results produce three equations at each forecast date: one for all the data, and two for each phase of the PDO. The variability of the equations is represented by the standard error (SE). The SE of the all data equation will be re-computed for those years corresponding to the warm and cool phases of the PDO for comparison to the SE of the equations fit for the warm and cool phases respectively. In addition, the forecast results from the warm and cool PDO equations will be mixed by an estimate of the current phase of the PDO to form a new volume forecast at each forecast date, which will be compared to the all data results.

3.1 Outline of Forecast Methodology

Generally a forecast model is $V_t = f(\text{SOI, Precipitation, SWE, etc.})$ Where V is the forecast volume at time t (month forecast is made). For example, if the operational year is from October to September, then a forecast would be made on October 1st for the volume of flow between October and September. At $t = 2$ the forecast volume would be reduced by one month and therefore be for the

November to September volume. The new methodology produces equations for both phases of the PDO.

The forecast process is as follows:

1. Specify the operational year (e.g. October to September).
2. Estimate the regression parameters for both the warm and cool phases of the PDO for the seasonal volume at time t .
3. Produce a volume forecast.
4. Disaggregate the forecasted seasonal volume into its respective monthly values with a disaggregation model(s).
5. Estimate the probability that the PDO is in the warm or cool phase given its current state.
6. Mix the disaggregated monthly sequences based on the probability calculated in (5).

3.2 Forecasting the Sandy River nr Marmot

The NRCS uses regression on principal components described in Garen (1992). The methodology is first to perform a correlation analysis on the predictor variables for each seasonal volume and keep only those predictors that are significant as measured by the correlation coefficient. The NRCS model uses a search algorithm to find an optimal set of independent variables. The algorithm starts with all the one variable equations and adds variables in a systematic process that ends when additional variables do not improve the standard error. For each equation tested, principal component(s) of the candidate variables are computed

and regression is then performed on the components. The regression coefficients are later transformed back into terms of the original variables. The NRCS model lists the top 20 equations ranked by the jackknife standard error.

The jackknife standard error is the standard error of the jackknife residuals. Jackknife results are obtained by an iterative process of removing the first year of the data prior to fitting the regression equation and making a prediction. Each subsequent year is removed so that the jackknife predictions are made without the benefit of the historical observation. The result is a more robust estimate of the standard error of the forecast model and thus a more robust estimate of the model's forecasting skill.

3.2.1 Model Selection

Selecting an equation from the top 20 equations produced by the NRCS model is the most subjective piece of the forecasting process. Historically, the hydrologists at the NRCS-NWCC pay particular attention to the variables included in the each of the forecast equations. There are two simple guidelines that are used when making the selection.

1. Similar variables should persist from one forecast equation to the next to maintain month to month stability in the otherwise independent forecasts.
2. Some similarity of variables should exist between the all data and two PDO models, however variation is to be expected since the underlying relationships differ somewhat, given the observations noted in Chapter 2.

The first guideline is particularly important because each forecast is produced independent of the others. Therefore, persistent variable selection from one forecast to the next will increase the stability of the volume forecasts from one month to the next. The second guideline accounts for the fact that if stability is to be expected within a set of forecasts, should that same stability be expected between the warm and cool PDO equations? The observations made in Chapter 2 indicate that during cool PDO phases the flow is higher than during warm PDO phases and that the relationship of the flow to the SOI differs. Therefore, variables were selected for the warm and cool PDO equations based more on consistency within a set of equations rather than between.

Ten dependent variables were used to fit the NRCS model to Sandy River streamflow. The variables used as predictors are shown in Table 3. The top 20 equations for each of the 10 dependent variables were analyzed for consistency within a common set of forecast equations. In other words, equations were selected from the top 20 for each dependent variable based on those variables that were persistent or hydrologically important such as 1st of month SWE. The variables selected in each equation are presented in Tables A1 – A3. The differences between the 3 sets of equations are a slightly higher occurrence of Government Camp precipitation in the low PDO equations versus the high PDO or all data equations. Another difference, one that would be expected, is the inclusion of SOI in the October to February and October to March forecasts for the cool PDO and the all data equations, respectively, and not in the warm PDO equations. SOI was

selected at times by the warm PDO equations, however, SOI is not significantly related to any of the dependent variables in the warm PDO data set and is therefore not an optimal predictor.

Variables	Period	Station	Record
Dependent Streamflow Volumes	Oct - Sep	14137000, Sandy nr Marmot	1912-1999
	Nov - Sep	14137000, Sandy nr Marmot	1912-1999
	Dec - Sep	14137000, Sandy nr Marmot	1912-1999
	Jan - Sep	14137000, Sandy nr Marmot	1912-1999
	Feb - Sep	14137000, Sandy nr Marmot	1912-1999
	Mar - Sep	14137000, Sandy nr Marmot	1912-1999
	Apr - Sep	14137000, Sandy nr Marmot	1912-1999
	May - Sep	14137000, Sandy nr Marmot	1912-1999
	Jun - Sep	14137000, Sandy nr Marmot	1912-1999
Snow Water Equivalent	jan	21D12S, Clear Lake Snotel	1958-1999
	feb	21D12S, Clear Lake Snotel	1941,1943,1953-1999
	feb	21D04S, Red Hill Snotel	1948-1999
	mar	21D13S, Clackamas Lake Snotel	1938-1941,1943-1944,1946,1948-1999
	mar	21D12S, Clear Lake Snotel	1931,1941,1944-1999
	mar	21D04S, Red Hill Snotel	1948-1999
	apr	21D13S, Clackamas Lake Snotel	1938-1943,1946,1948-1999
	apr	21D12S, Clear Lake Snotel	1932-1999
	apr	21D04S, Red Hill Snotel	1948-1999
may	21D12S, Clear Lake Snotel	1950-1953,1955-1999	
Monthly Volumes of Precipitation	Oct	OR3770, Headworks, Portland OR	1932-1999
	Nov	OR3770, Headworks, Portland OR	1932-1999
	Dec	OR3770, Headworks, Portland OR	1932-1999
	Jan	OR3770, Headworks, Portland OR	1932-1999
	Feb	OR3770, Headworks, Portland OR	1932-1999
	Mar	OR3770, Headworks, Portland OR	1932-1999
	Apr	OR3770, Headworks, Portland OR	1932-1999
	May	OR3770, Headworks, Portland OR	1932-1999
	Jun	OR3770, Headworks, Portland OR	1932-1999
	Oct	OR3402, Government Camp, OR	1954-1999
	Nov	OR3402, Government Camp, OR	1954-1999
	Dec	OR3402, Government Camp, OR	1954-1999
	Jan	OR3402, Government Camp, OR	1954-1999
	Feb	OR3402, Government Camp, OR	1954-1999
	Mar	OR3402, Government Camp, OR	1954-1999
	Apr	OR3402, Government Camp, OR	1954-1999
	May	OR3402, Government Camp, OR	1954-1999
	Jun	OR3402, Government Camp, OR	1954-1999
SOI	jun-sep	Southern Oscillation Index	1911-1998

Table 3: Data used as dependent and predictor variables to fit the 3 sets of equations with Reg.exe.

3.3 Forecast Volume Results

The two methods, fitting models to the split data based on the PDO vs. the entire data set, are evaluated based on standard error, bias, and cross validation. Throughout the remainder of this document the model fit to all of the data will be referred to as the control or all data equations, and the two models fit to the data based on the phase of the PDO will be referred to as the split data model.

The jackknife and regression standard errors of the all data equations were compared to those of the warm and cool PDO equations. The comparison is not straightforward. The jackknife standard errors (JSE) and regression standard errors (SE) had to be recalculated from the results of the all data equations specific to the warm and cool PDO years. The SE and JSE from the warm and cool PDO equations can then be compared to the re-calculated SE and JSE of the all data equations in the warm and cool PDO years respectively. Some improvement is expected by the warm and cool PDO equations because the models are specifically fit to the warm and cool PDO years and are optimal. There are no significant predictors of the October – September volume during the warm phase. Therefore, the historical standard deviation is substituted in Table 4 and the JR values are not applicable (NA).

Also, the correlation of the jackknife predictions and the historical data were recalculated for the warm and cool PDO years of the all data results. The Jackknife R (JR) and JSE and SE values are shown in Table 4.

Volume	Warm PDO Results			All Data Results						Cool PDO Results		
	JSE*	SE*	JR	Warm	Warm	Warm	Cool	Cool	Cool	JSE*	SE*	JR
Oct - Sep	200	200	NA	212	207	NA	141	137	0.59	141	133	0.56
Nov - Sep	196	188	0.28	188	183	0.39	164	158	0.43	137	130	0.57
Dec - Sep	157	150	0.35	158	153	0.37	125	122	0.57	126	116	0.66
Jan - Sep	145	132	0.1	145	138	0.23	120	112	0.61	116	103	0.65
Feb - Sep	109	100	0.22	113	108	0.22	106	102	0.46	112	102	0.44
Mar - Sep	78	74	0.52	78	75	0.53	92	89	0.51	89	85	0.55
Apr - Sep	59	55	0.59	62	59	0.47	59	54	0.77	57	53	0.78
May - Sep	45	42	0.47	50	48	0.36	43	40	0.86	38	35	0.89
Jun - Sep	32	30	0.29	35	33	0.22	35	32	0.75	31	28	0.81
Jul - Sep	12	11	0.59	15	15	0.22	17	16	0.57	16	15	0.65

*Thousand Acre-Feet (KAF)

Table 4: Results of the 3 models fit to the Sandy River nr. Marmot data over the entire period of record (all data), warm, and cool PDO periods.

The JSE and SE of the high and low PDO equations are only modestly lower and sometimes higher when compared with the recalculated SE and JSE in the cool and warm PDO periods of the all data equations. The only exception is that of the cool PDO phase for the November to September volume.

The SE of both the warm and cool PDO equations were tested against the re-computed SE's in the warm and cool PDO years of the all data equations respectively. Only the difference of 28 KAF for the November to September cool PDO phase equation is statistically significant at a P value of 0.03 (Table 5).

F-Test on SE Results		
Volume	Warm P Values	Cool P Values
Oct - Sep	NA	0.42
Nov - Sep	0.03	0.43
Dec - Sep	0.36	0.45
Jan - Sep	0.36	0.42
Feb - Sep	0.5	0.36
Mar - Sep	0.43	0.47
Apr - Sep	0.45	0.38
May - Sep	0.28	0.27
Jun - Sep	0.24	0.31
Jul - Sep	0.32	0.1

Table 5: Results of the F-tests performed on the standard errors.

The comparison of the SE and JSE is presented in Figures 13 and 14 for each forecast date. There are two particular observations shown by Figure 13 and 14 that are important. Clearly the largest differences between the JSE and SE of the high and low PDO equations and the re-computed JSE and SE of the all data equations is during the early forecasts, particularly in the cool PDO equation in November. The other important observation is that the JSE and SE are much lower overall in the cool PDO phase versus the warm PDO phase for the October-September through the January – September forecast. The differences between the overall warm and cool PDO equations are significant at levels much less than an alpha of 0.05 through the January - September forecast and are insignificant from the February – September forecast on.

Sandy River nr Marmot forecast equations

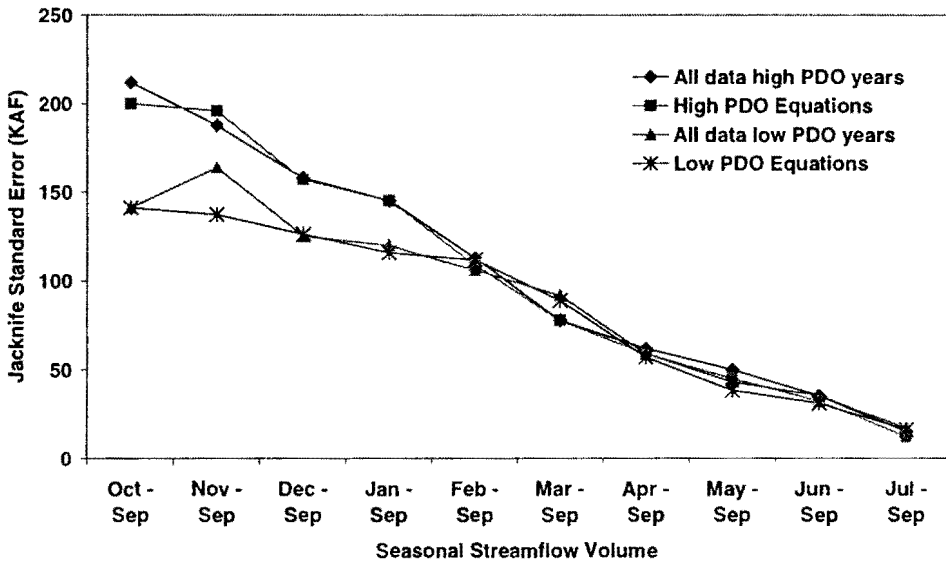


Figure 13: Jackknife standard error for the high and low PDO equations and the re-computed JSE in the high and low PDO periods of the all data equations.

Sandy River nr Marmot forecast equations

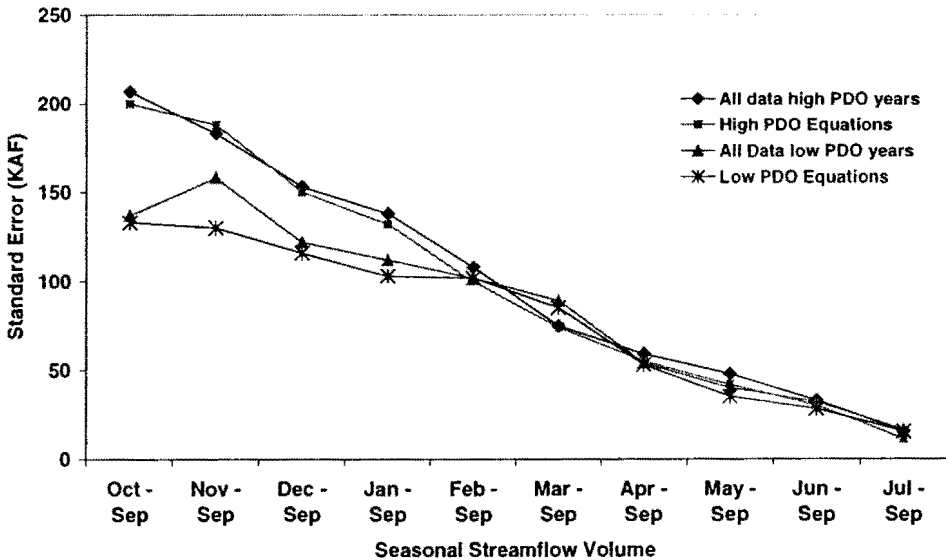


Figure 14: Standard error for the high and low PDO equations and the re-computed JSE in the high and low PDO periods of the all data equations.

The forecast model reproduces the observations discussed in Chapter 2. Early in the year, the uncertainty of streamflow predictions is much lower during the cool PDO phase than during the warm PDO phase. Furthermore, the overall reduction in the JSE and SE for all three equations during the cool PDO phase suggests that prior to the February to September volume forecast, the variability of volume forecasting during warm phases of the PDO is relatively high.

3.3.1 Estimating the Future Phase of the PDO

Given a perfect forecast of the phase of the PDO, forecasts could be made directly with the warm and cool PDO equations. However, the reality is that the PDO index is not perfectly related to the phase of the PDO. In other words, the PDO index has, at times, been positive during cool phases and vice versa. Therefore, the forecasts will have to be made with both the warm and cool PDO equations and the results will have to be mixed consistent with the probability that the PDO is going to be in one phase versus the other. The mixing is accomplished by mixing generated forecast values in proportion to an estimate of the phase of the PDO.

3.3.2 Estimating α and β .

The PDO index is used as the indicator of the current state of the PDO. Consequently, values of the PDO index averaged over periods prior to the forecast date need to be related to the phases of the PDO. Therefore, an analysis was done, where the PDO index was averaged between October and September for increasing number of years prior to a forecast date. The three year average was selected

because the results shown in Table 6 and Figure 15 show an 89% probability that the PDO will be in the warm phase given that the average PDO index over the past three years is positive, and a 93% probability that the PDO will be in the cool phase given that the previous three years of the PDO index have on average been negative.

	Warm Phase	Cool Phase
Oct-1 (+)	0.80	0.20
Oct-1 (-)	0.13	0.87
Oct-2 (+)	0.80	0.20
Oct-2 (-)	0.10	0.90
Oct-3 (+)	0.89	0.11
Oct-3 (-)	0.07	0.93
Oct-4 (+)	0.89	0.11
Oct-4 (-)	0.07	0.93

Table 6: Contingency values predicting the phase of the PDO given the previous 1 to 4 year average of the PDO index. Oct-1 (+) indicates that 1 previous year from October to September was on average positive, etc.

Therefore, α and β are constants that merely depend on the average over the past three water years be computed. If the PDO is expected to be in the warm phase during the forecast then equation 1 will be used, and if the PDO is expected to be in the cool phase equation 2 will be used:

$$V_t = \alpha(V_{t,PDO>0}) + (1-\alpha)(V_{t,PDO<0}) \quad (1)$$

$$V_t = \beta(V_{t,PDO<0}) + (1-\beta)(V_{t,PDO>0}) \quad (2)$$

Where V_t is the volume of flow at time t .

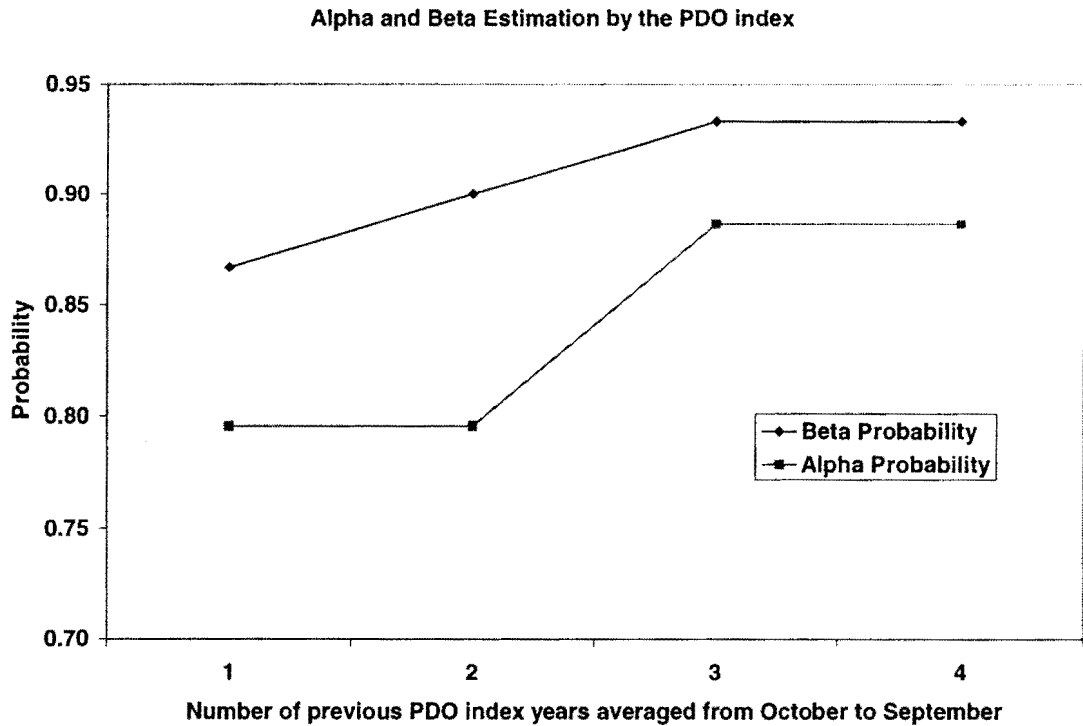


Figure 15: Alpha and Beta probabilities as estimated by the average of the PDO index over 1 to 4 previous years from October to September.

Mixed forecasts were produced over the same period of record as the all data equation results for each forecast date. Both the warm and cool forecast equations were used to generate 100 forecasts for each year. The previous three years of the PDO index were then averaged and if the result was positive then 89 warm forecasts were mixed with 11 cool. On the other hand, if the previous three years of the PDO index were on average negative, 93 cool forecasts were mixed with 7 warm.

3.3.3 Cross Validation Results

Cross validation is also useful in evaluating forecast models. Cross validation as applied here is the ability of the model to predict in the appropriate tercile of the data. The historical observations are split into equal thirds labeled as below, normal, and above for the lower, middle, and upper terciles of the data respectively. The mixed forecasts and the all data forecasts are compared to the historic data. The results are presented in Table 7.

Oct - Sep					Nov - Sep				
Case		Q Estimated w\ All Data			Case		Q Estimated w\ All Data		
		Below	Normal	Above			Below	Normal	Above
Observed Q	Below	30%	63%	7%	Observed Q	Below	48%	52%	0%
	Normal	7%	86%	7%		Normal	14%	79%	7%
	Above	4%	63%	33%		Above	0%	89%	11%
Case		Q Est. w\ Split Data			Case		Q Est. w\ Split Data		
		Below	Normal	Above			Below	Normal	Above
Observed Q	Below	33%	52%	15%	Observed Q	Below	41%	59%	0%
	Normal	18%	64%	18%		Normal	11%	86%	4%
	Above	26%	26%	48%		Above	0%	81%	19%
Dec - Sep					Jan - Sep				
Case		Q Estimated w\ All Data			Case		Q Estimated w\ All Data		
		Below	Normal	Above			Below	Normal	Above
Observed Q	Below	33%	53%	13%	Observed Q	Below	50%	36%	14%
	Normal	31%	69%	0%		Normal	36%	43%	21%
	Above	0%	53%	47%		Above	7%	64%	29%
Case		Q Est. w\ Split Data			Case		Q Est. w\ Split Data		
		Below	Normal	Above			Below	Normal	Above
Observed Q	Below	40%	60%	0%	Observed Q	Below	43%	50%	7%
	Normal	50%	44%	6%		Normal	21%	64%	14%
	Above	13%	33%	53%		Above	21%	36%	43%

Table 7: Contingency tables showing the average results of the cross validation of the all data, above, and below equations respectively.

The results are quite variable over the 4 forecast periods shown in Table 7. Overall both methods produce forecasts that tend toward normal. The mixed split forecasts better predict upper tercile events overall, but only show better results of lower tercile predictions during the October and December forecasts. The only

clear pattern of improvement shown by the cross validation results is the improvement in forecasting the upper terciles of the historical data.

3.3.4 Forecast Mean and Evaluation of Bias

Finally, it is important to evaluate not only the variability of the forecasts and their ability to forecast in relatively high and low PDO years, but also that the mean forecast is similar to the observed data. The split data forecast, which is the integrated warm and cool phase volume forecast after consideration of the PDO, predicts the mean better than the all-data in both the warm and cool phases. Figure 16 shows that the all-data forecasts are biased low in October, November, and December during the cool PDO phase, whereas the split data are not as biased. Figure 17 shows that during the warm phase, the split data are unbiased through the January forecast, while the all data equations are biased high.

Seasonal streamflow volume forecasting without consideration of the PDO has been shown to produce forecasts that are biased high during warm phases of the PDO and biased low during cool phases of the PDO resulting in a tendency to under-predict potential drought and flood years respectively. Although the bias is not completely removed by splitting the data as was shown here, there is a significant improvement particularly during the early forecasts produced from October to December.

Forecast Model Comparison (Cool PDO Phase)

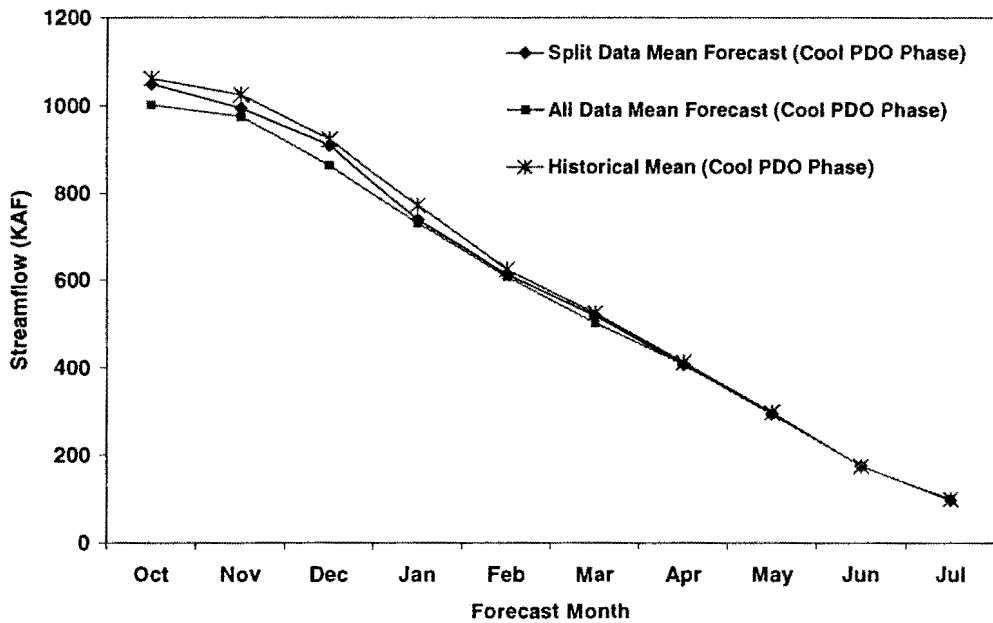


Figure 16: Comparison of the mean forecast during cool PDO phase. Split data are the warm and cool phase equations that have been mixed by the PDO.

Forecast Model Comparison (Warm PDO Phase)

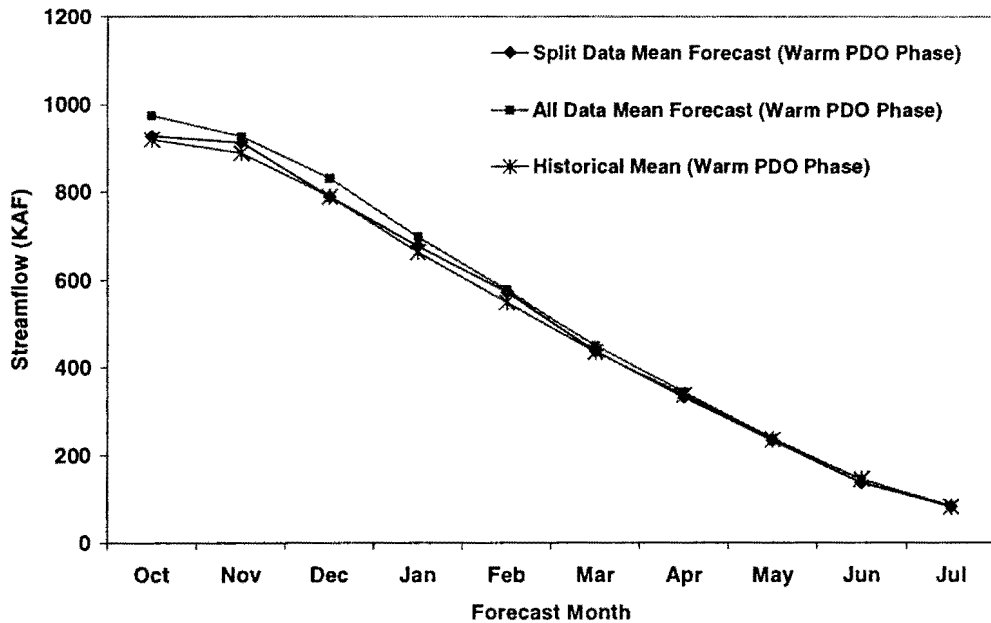


Figure 17: Comparison of the mean forecast during warm PDO phase. Split data are the warm and cool phase equations that have been mixed by the PDO.

Even though the split forecast equations for the October through January volumes was shown still to tend to normal conditions in the cross validation results, the discrepancy in bias shown in Figures 16 and 17 suggest that the split forecast equations better estimate seasonal streamflow in the Sandy River. It is unfortunate that the split-forecast results did not show significant improvements in cross-validation during both phases of the PDO. However, the new method produces unbiased forecasts that take into consideration the change in overall variability given the state of the PDO.

The next step is to develop disaggregation models consistent with the findings of Koch and Fisher (2000) that will be used to distribute generated volumes of flow throughout the forecasted operational year.

Chapter 4

Disaggregation of Volume Forecasts

In addition to a better estimate of seasonal streamflow volume, water resource managers also need a forecast product that is more suitable for the decisions they are required to make. While a seasonal volume prediction and estimate of uncertainty is helpful to operators of reservoir systems, seasonal volumes are limited to long term seasonal planning, whereas much of the operation of reservoirs is on finer time scales. Even though statistical forecasting is limited in its refinement in time, disaggregation models can be used to resolve a seasonal volume into time increments as small as one month. Based on conversations with several reservoir operators, monthly estimates of flow are more useful than seasonal volumes. Therefore, a disaggregation model will be used here to disaggregate the seasonal volumes into sequences of monthly flows.

4.1 Disaggregation Modeling

Disaggregation models are designed to produce a sub – aggregate series from an aggregate series by preserving the historical statistical relationships that exist between the two series. For instance a forecast model may produce a seasonal forecast volume for the months April to September. The disaggregation model would be used to estimate the flows for each month of the season, i.e. April, May, June, July, August, and September.

Some common terms and symbols will be used to represent the statistical moments as well as the temporal scale of a time series. The term “key” series refers to the temporal scale of the aggregate series (independent variable) and the term “sub”-series refers to the temporal scale of the disaggregated series. An example would be an annual series disaggregated into its monthly constituents, where the annual series is the “key” series and the monthly series is the sub – series. The forecast model will forecast the key series (seasonal volumes), which will subsequently be disaggregated into the sub – series (monthly volumes) by the disaggregation model.

The disaggregation model preserves certain statistical relationships that exist between the key and sub – series. Those statistical relationships are used to estimate the parameters of the models by the method of moments (MOM). A lag zero covariance matrix between the independent key series X and dependent sub – series Y is represented by S_{xy} . The lag – one covariance matrix between the same two series is denoted $S_{xy}(1)$, where the Y series is lagged by one time unit to the X series.

Having selected a time interval for disaggregation, the disaggregated streamflow is represented as:

$$Q_t(j) \text{ for } j = 1, 2, \dots, K \quad (3)$$

where $Q_t(j)$ is the disaggregated volume of streamflow,

t is the time the forecast is produced, j is the time period

corresponding to the disaggregated flow, and K is the total number of time periods into which the volume forecast is disaggregated.

The relationship between the volume forecast and the disaggregated volumes are shown schematically in Figure 18.

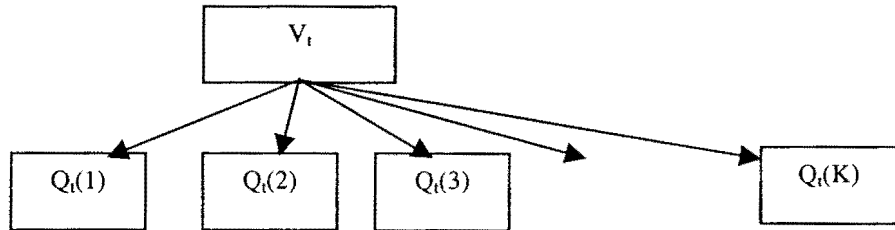


Figure 18: Schematic illustration of the disaggregation terminology and the relationship of the volume forecast to the disaggregated volumes

“Disaggregation modeling is a process by which time series are generated dependent on a time series already available” (Salas et al. 1980). The utility of disaggregation modeling is that most synthetic forecasting models are more accurate at an annual or seasonal level, but the operation of water resource systems require a more refined volume such as monthly forecasts. Valencia and Schaake (1973) introduced a disaggregation model of the form shown below in equation (4).

$$\mathbf{Y} = \mathbf{AX} + \mathbf{B}\boldsymbol{\varepsilon} \quad (4)$$

Where \mathbf{Y} is a vector of correlated random variables, \mathbf{X} is a vector of correlated random variables, \mathbf{A} and \mathbf{B} are coefficient matrices, and $\boldsymbol{\varepsilon}$ is a vector of random normal variables independent of \mathbf{X} (Valencia and Schaake 1973). Equation 4 assumes that all random variables are normal and have mean 0. Using the Method of Moments (MoM), the historical statistical relationships between the key (\mathbf{X}) and sub-series (\mathbf{Y}) are preserved.

Parameter Estimation: Valencia & Schaake (1973)

Given (4):

$$S_{XY} = E[XY^T]$$

$$S_{YX} = E[YX^T]$$

$$S_{YY} = E[YY^T]$$

Post multiplying by X^T gives, $E[YX^T] = AE[XX^T]$, from which the parameter matrix A is solved for in (5).

$$A = S_{YX} S_{XX}^{-1} \quad (5)$$

Next, post multiplying by Y^T gives, $E[YY^T] = AE[XX^T]A^T + BB^T$, and the BB^T matrix is given by (6).

$$BB^T = S_{YY} - S_{YX} S_{XX}^{-1} S_{XY} \quad (6)$$

The BB^T matrix is proven to be positive semi-definite in Valencia and Schaake (1973), pp.582 – 583. The solution of BB^T for the parameter matrix B is described by equations 3.43-3.45 on page 87 of Salas et al. (1980) or Lane (1982) page 510.

Mejia and Rousselle (1976), and Lane (1982) have since offered improvements on the form of the Valencia and Schaake (1973) model. The Valencia and Schaake (1973) model is limiting in that it does not preserve the over-year serial correlation of the dependent (sub – key) series. In (4) there are no provisions for what occurred prior to the first month of the Y series. Mejia and Rousselle (1976)

proposed an additional component of the model that would better preserve the historical serial correlation in the dependent series.

$$\mathbf{Y}_t = \mathbf{A}\mathbf{X}_t + \mathbf{B}\varepsilon + \mathbf{C}\mathbf{Y}_{t-1} \quad (7)$$

The only addition to (4) is the $\mathbf{C}\mathbf{Y}_{t-1}$ component, which preserves the lag $t = 1 \dots k$ over-year serial correlation in the dependent series \mathbf{Y}_t . Lane (1982) further improved the Mejia and Rousselle (1976) model, which was shown to have an inconsistency in its parameter estimation scheme. In following the method of moments to solve for A, B, and C in (7) the lag one covariance between the X and \mathbf{Y}_{t-1} series ($\mathbf{S}_{xy}(1)$) appears. Lane (1982) showed that an adjustment can be found for the $\mathbf{S}_{xy}(1)$ and $\mathbf{S}_{yy}(1)$ moments to correct for the inconsistency if the key series is modeled as first order auto-regressive. Equation 8 and 9 show the Lane correction to the $\mathbf{S}_{xy}(1)$ and $\mathbf{S}_{yy}(1)$ moments, which will preserve the moments and preserve additivity in the dependent series. The five equations to estimate the parameters as proposed by Lane (1982) are:

$$\mathbf{S}_{xy}^*(1) = \mathbf{S}_{xx}(1)\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy} \quad (8)$$

$$\mathbf{S}_{yy}^*(1) = \mathbf{S}_{yy}(1) + \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}[\mathbf{S}_{xy}^*(1) - \mathbf{S}_{xy}(1)] \quad (9)$$

$$\mathbf{A} = [\mathbf{S}_{yx} - \mathbf{S}_{yy}^*(1)\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}^T(1)][\mathbf{S}_{xx} - \mathbf{S}_{xy}^*(1)\mathbf{S}_{yy}^{-1}\mathbf{S}_{xy}^T(1)]^{-1} \quad (10)$$

$$\mathbf{C} = [\mathbf{S}_{yy}^*(1) - \mathbf{A}\mathbf{S}_{xy}^*(1)]\mathbf{S}_{yy}^{-1} \quad (11)$$

$$\mathbf{B}\mathbf{B}^T = \mathbf{S}_{yy} - \mathbf{A}\mathbf{S}_{xy} - \mathbf{C}\mathbf{S}_{yy}^T(1) \quad (12)$$

4.2 Disaggregating Volume Forecasts

Volume forecasting was discussed in Chapter 3, and the result was the demonstration that seasonal streamflow volume forecasting can be improved by mixing forecasts from two models that predict flow based on the state of the PDO. An additional improvement in seasonal streamflow volume forecasting is the disaggregation of those forecasts into finer time resolutions, which are much more practical to use in water supply management. The following section will discuss and evaluate the disaggregation of seasonal streamflow volume forecasts produced by the equations fit in Chapter 3 for the Sandy River. Again the results of disaggregating forecasts fit to all the data will be compared to splitting the data and disaggregating the forecasts produced by the warm and cool PDO forecast equations.

4.2.1 Methods

Data from each of the 10 dependent variables (seasonal volumes) in Table 3 are disaggregated by the control method, which uses all the data to fit both the forecast and disaggregation equations, and the split method, which splits the data based on the PDO and uses forecasts and disaggregation equations fit specifically to streamflow data during the warm and cool PDO states. The split data disaggregation sequences will then be mixed by the probability of the state of the PDO given by Table 6.

The process is to generate 100 forecasts for each year over a common period of record for all 3 equations. The control scenario is simply to disaggregate the 100 generated forecasts in each year and average them to produce the final results. The

split scenario is more complicated. In the split scenario 100 forecasts are produced by both the warm and cool PDO forecast equations. A disaggregation model specific to the state of the PDO (i.e. warm and cool phase disaggregation models) then disaggregates each of the 100 forecasts in every year. Finally, the disaggregated forecasts are mixed based on the state of the PDO with equation 13 if the previous 3 years of the PDO index are on average positive and with equation 14 if the PDO index is on average negative:

$$Q_{t,PDO>0}(j) = \alpha(Q_{t,PDO>0}(j)) + (1 - \alpha)(Q_{t,PDO<0}(j)) \quad (13)$$

$$Q_{t,PDO<0}(j) = \beta(Q_{t,PDO<0}(j)) + (1 - \beta)(Q_{t,PDO>0}(j)) \quad (14)$$

Experience in disaggregating the volume forecasts produced by all three sets of forecast equations has proven the Lane model to be difficult to use. Particularly, the BB^T matrix was found at times to be negative semi-definite, which is in clear violation of the principles within which the model was developed. Therefore, the Valencia and Schaake model, (4), was used for disaggregating the forecasts produced by the Sandy River forecast equations discussed in Chapter 3. While the Lane (1982) model is theoretically superior to the Valencia and Schaake (1973) model, the intent here is not to produce the best possible forecasts but rather to objectively compare the two methods.

4.2.2 Results from the Sandy River nr Marmot

Figures 19 and 20 show the mean monthly results of disaggregating the October to September forecasts for water years 1913 to 1999 in cool and warm phases of the PDO, respectively. Figure 21 is the average results for the period of

record. The split disaggregation scheme better estimates the mean monthly flow during cool and warm PDO phases (low and high PDO index respectively). Further, the control (all data) results are biased low on average, whereas the split results are not. The results are the same for the November-September (Figures 22 – 24) and December-September (Figures 25 – 27) volumes, however the control begins to estimate the mean monthly flow better during the warm phase beginning in the January-September (Figures 28 – 30) volume. Generally, the split results are noticeably better during the cool PDO phases when the control (all data) is typically biased low.

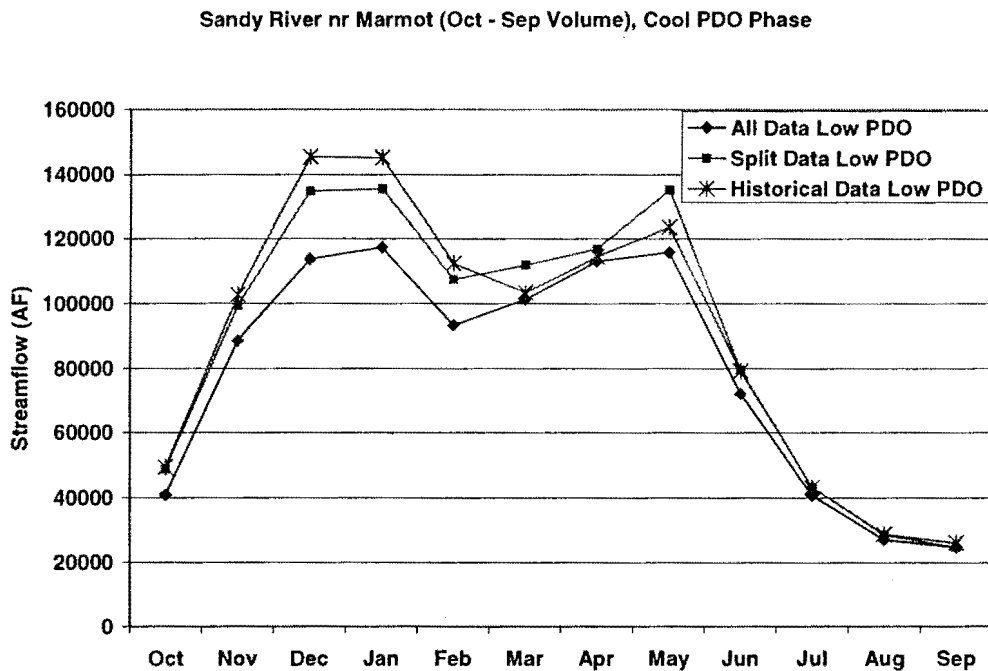


Figure 19: Disaggregated mean monthly flow from October to September in acre-feet during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Oct - Sep Volume), Warm PDO Phase

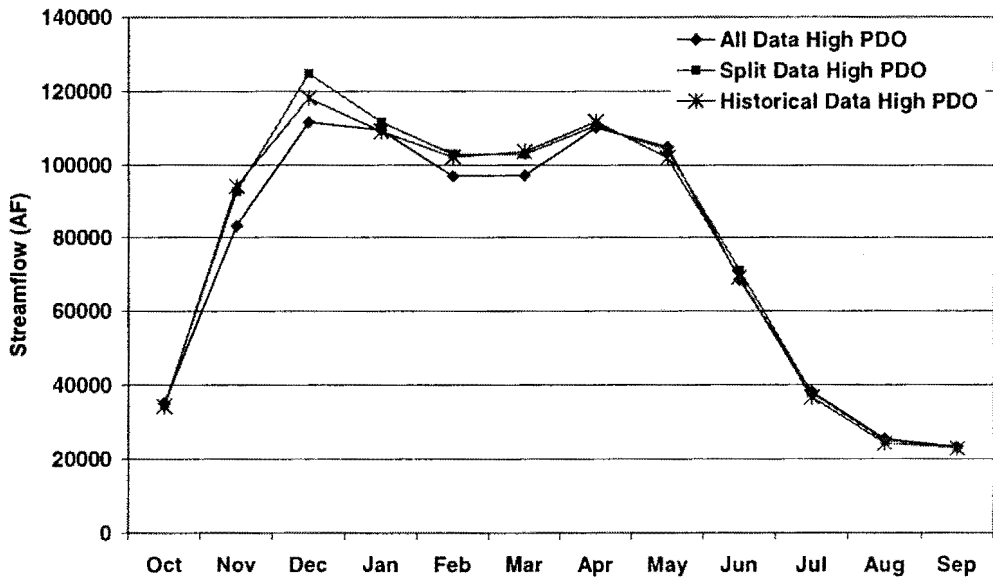


Figure 20: Disaggregated mean monthly flow from October to September in acre-feet during the warm phase of the PDO (high PDO index).

Sandy River nr Marmot (Oct - Sep Volume), Average Results

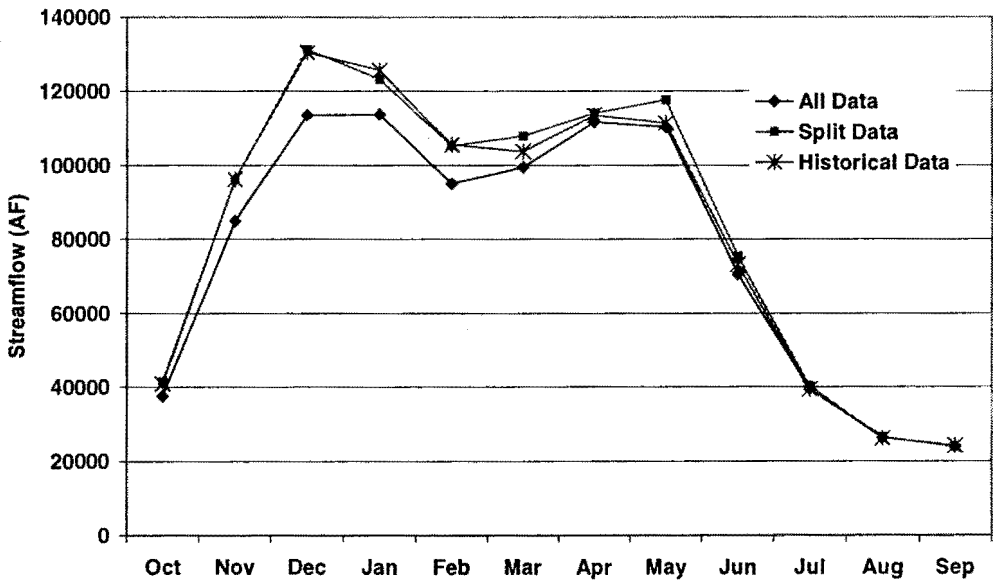


Figure 21: Average disaggregated mean monthly flow from October to September in acre-feet.

Sandy River nr Marmot (Nov - Sep Volume), Cool PDO Phase

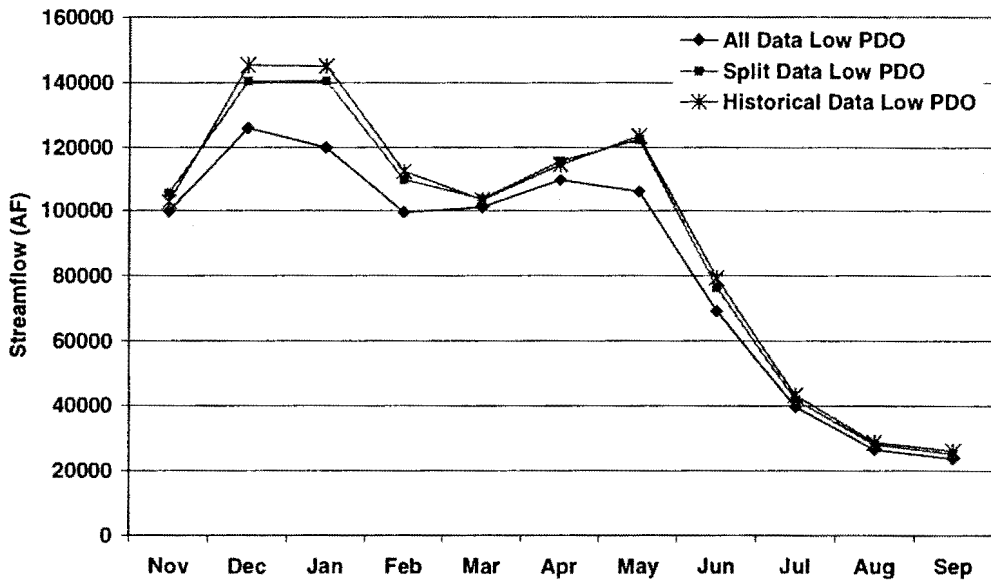


Figure 22: Disaggregated mean monthly flow from November to September in acre-feet during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Nov - Sep Volume), Warm PDO Phase

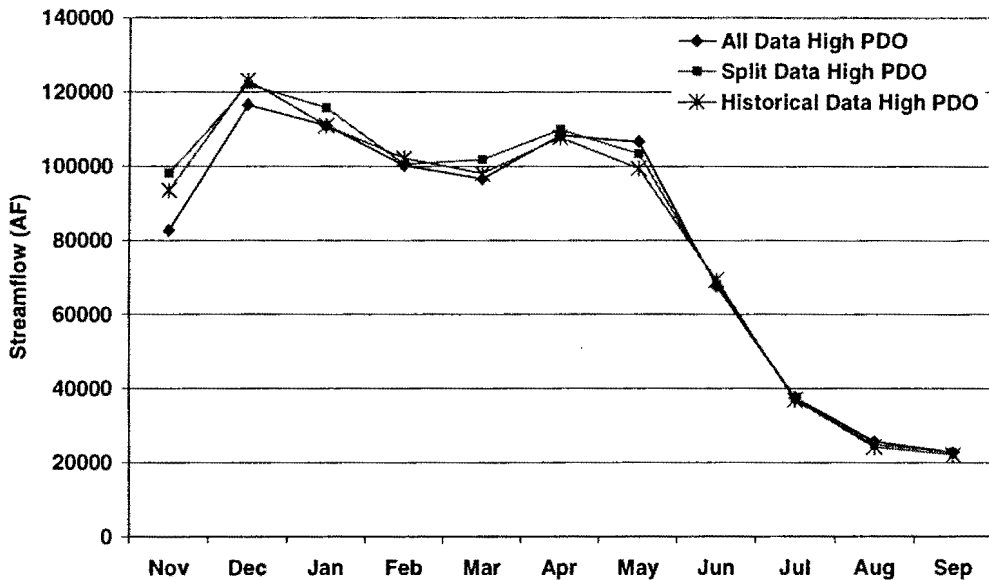


Figure 23: Disaggregated mean monthly flow from November to September in acre-feet during the warm phase of the PDO (high PDO index).

Sandy River nr Marmot (Nov - Sep Volume), Average Results

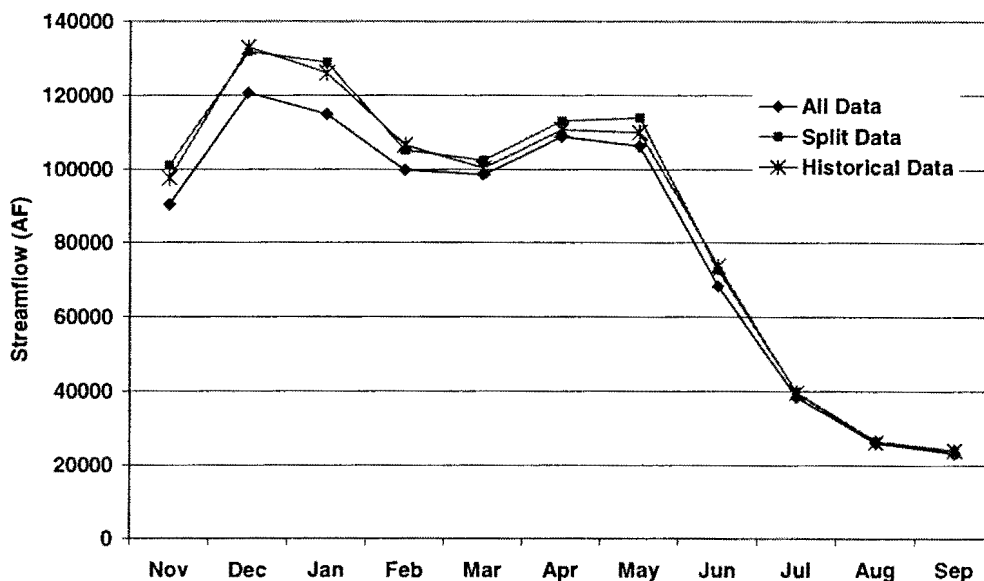


Figure 24: Average disaggregated mean monthly flow from November to September in acre-feet.

Sandy River nr Marmot (Dec - Sep Volume), Cool PDO Phase

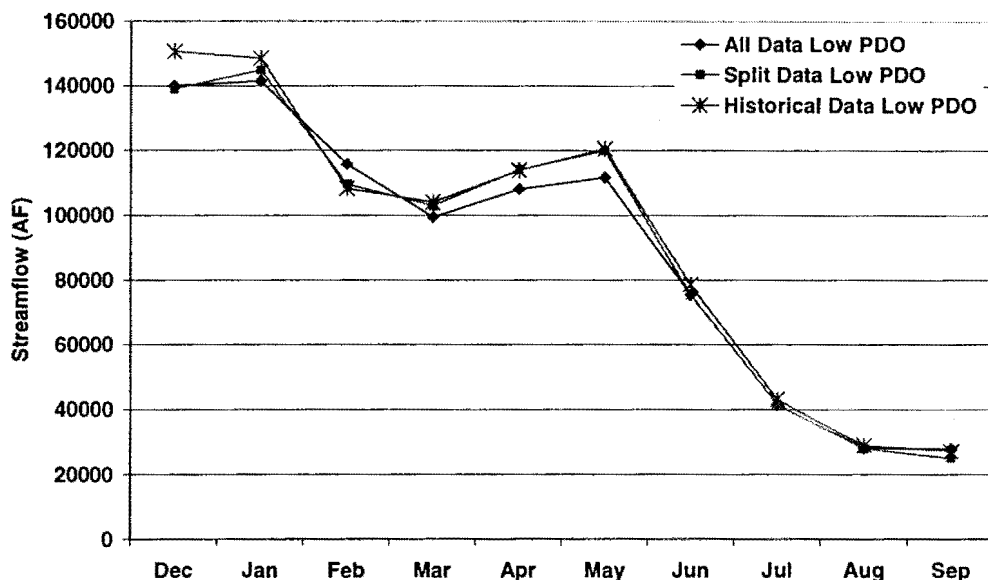


Figure 25: Disaggregated mean monthly flow from December to September in acre-feet during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Dec - Sep Volume), Warm PDO Phase

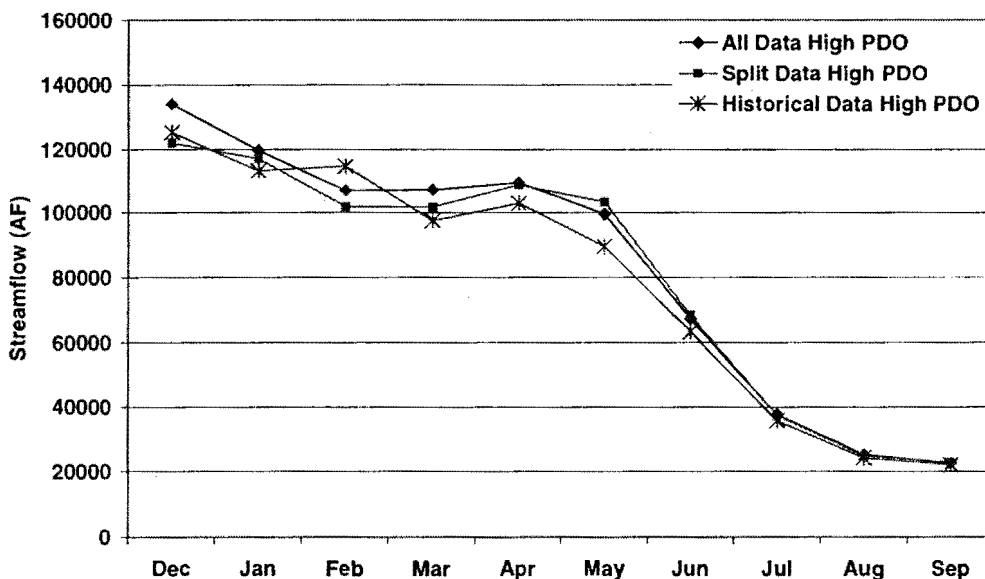


Figure 26: Disaggregated mean monthly flow from December to September in acre-feet during the warm phase of the PDO (high PDO index).

Sandy River nr Marmot (Dec - Sep Volume), Average Results

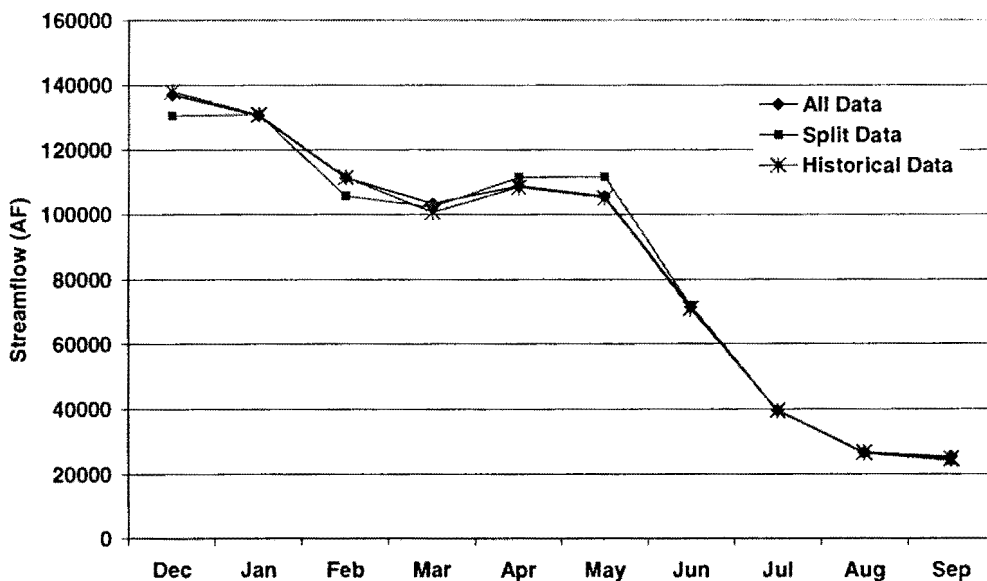


Figure 27: Average disaggregated mean monthly flow from December to September in acre-feet.

Sandy River nr Marmot (Jan - Sep Volume), Cool PDO Phase

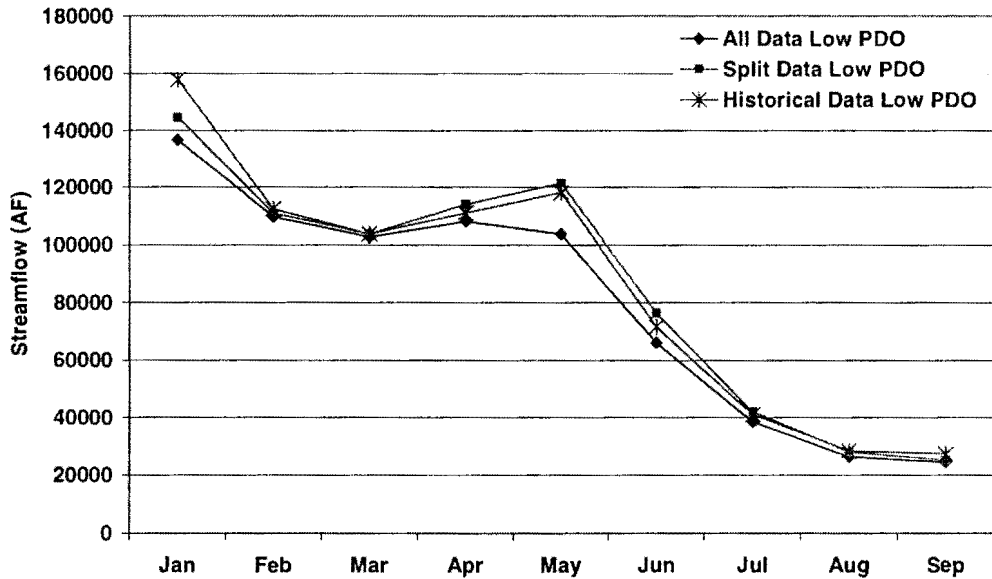


Figure 28: Disaggregated mean monthly flow from January to September in acre-feet during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Jan - Sep Volume), Warm PDO Phase

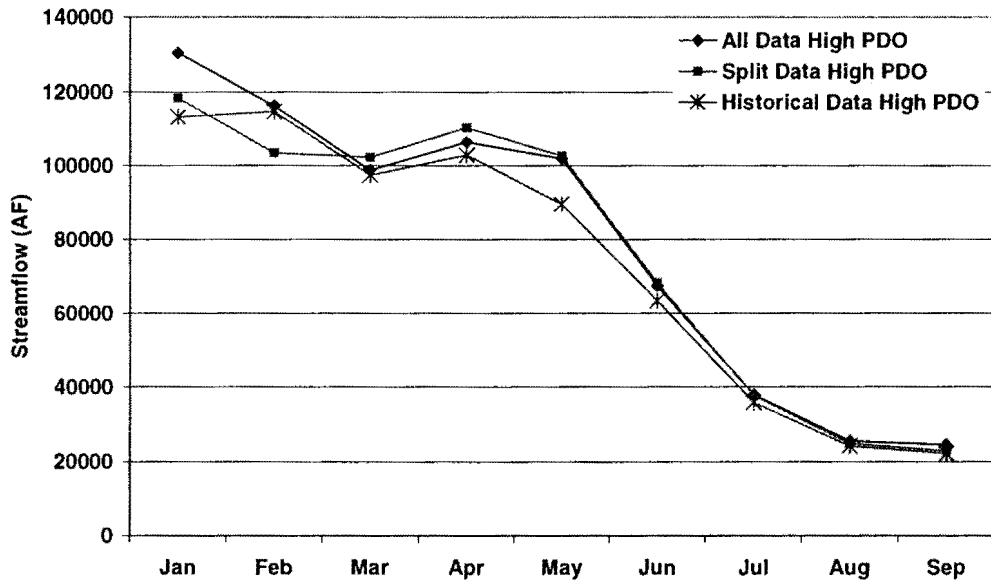


Figure 29: Disaggregated mean monthly flow from January to September in acre-feet during the warm phase of the PDO (high PDO index).

Sandy River nr Marmot (Jan - Sep Volume), Average Results

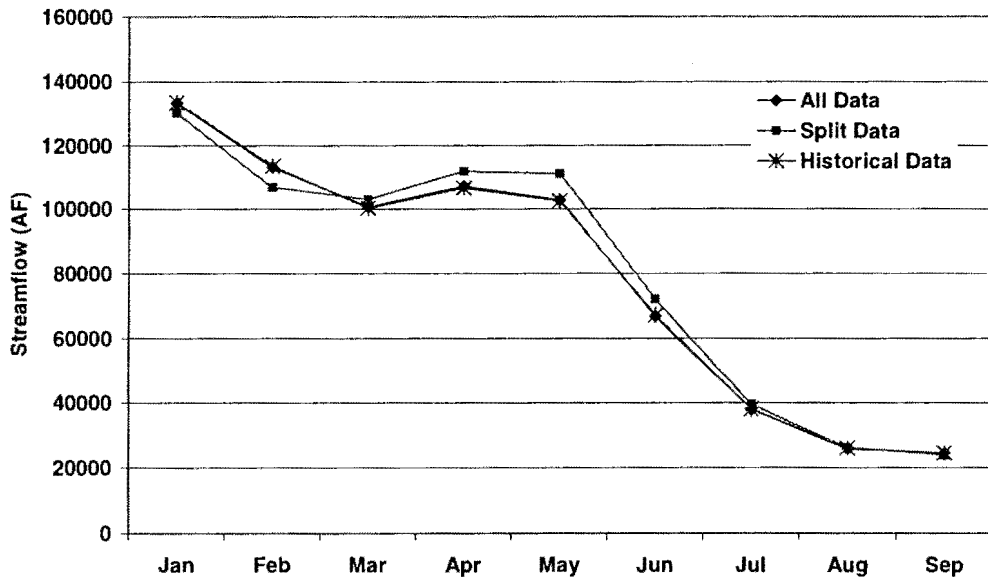


Figure 30: Average disaggregated mean monthly flow from January to September in acre-feet.

4.2.3 Variability of Monthly Forecasts

The variability of the monthly forecasts is also important. The mixed forecasts should follow the normal distribution prior to mixing by the PDO. However the mixed forecasts would not be expected to follow a normal distribution. Therefore, a non-parametric measure of the variability of the forecasts was used to ascertain the relative variability of the control and split results. The results shown below are all measurements of the 90% range, which is a measurement of the variability that does not assume a normal distribution.

The upper and lower ranks are computed for any probability ϕ and for any number of years of data, n , by 15 and 16 respectively (Helsel and Hirsch 1992).

$$R_u = (1 - \phi)(n + 1) \tag{15}$$

$$R_l = \phi(n + 1) \tag{16}$$

The 90% range is the difference between the flow that corresponds to the upper 95% rank and the lower 5% rank.

It is apparent from Figures 31 – 38 that the split results are far less variable than the historical data and the control. The results from the October-September volume are the only exception and are likely the result of the historical mean and standard deviation being used to generate the forecasts for the warm PDO phase. The historical mean and standard deviation had to be used in this case because there were not any statistically significant predictors of the October-September volume during warm PDO phases.

Sandy River nr Marmot (Oct - Sep Volume), Cool PDO Phase

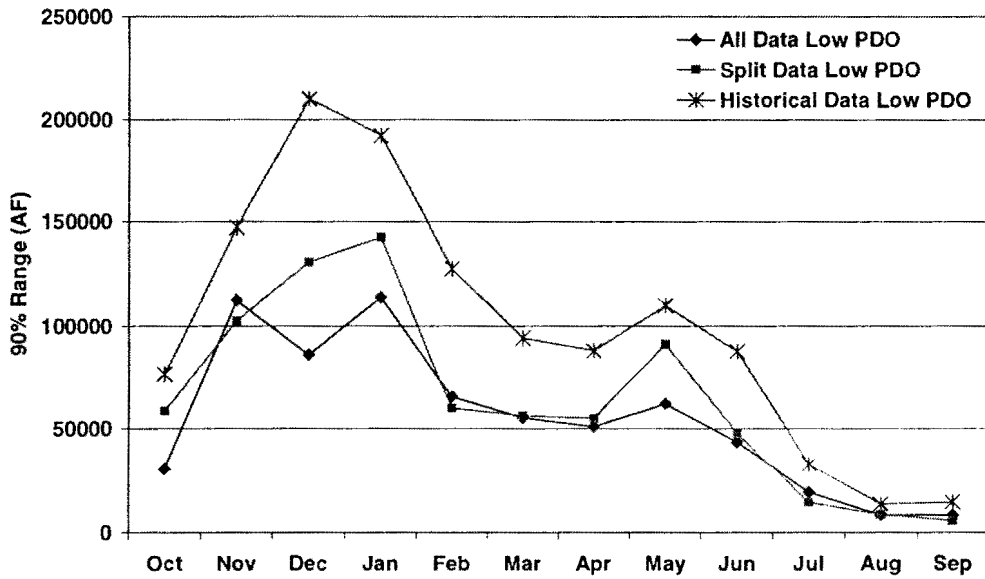


Figure 31: Variability by month for the October-September volume during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Oct - Sep Volume), Warm PDO Phase

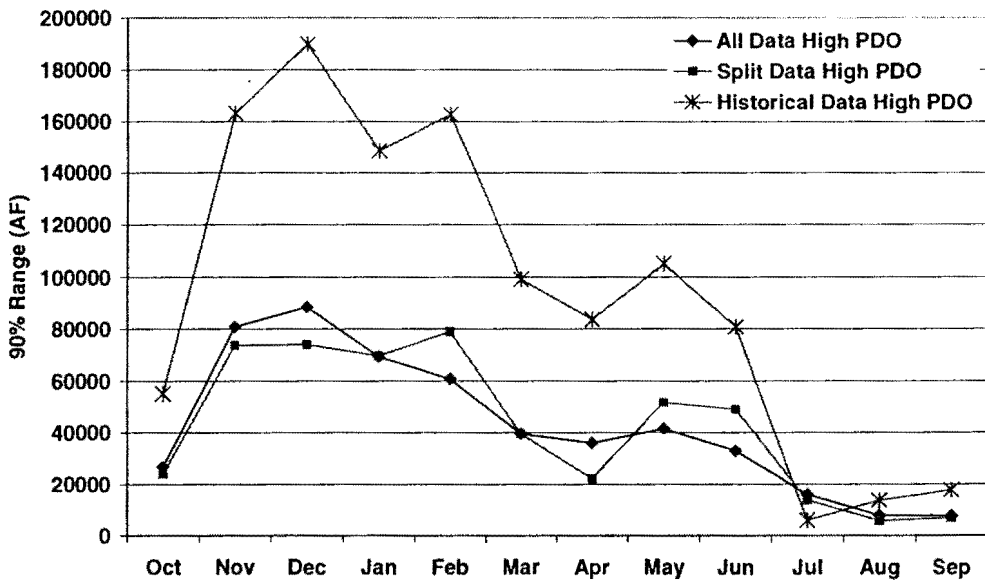


Figure 32: Variability by month for the October-September volume during the warm phase of the PDO (high PDO index).

Sandy River nr Marmot (Nov - Sep Volume), Cool PDO Phase

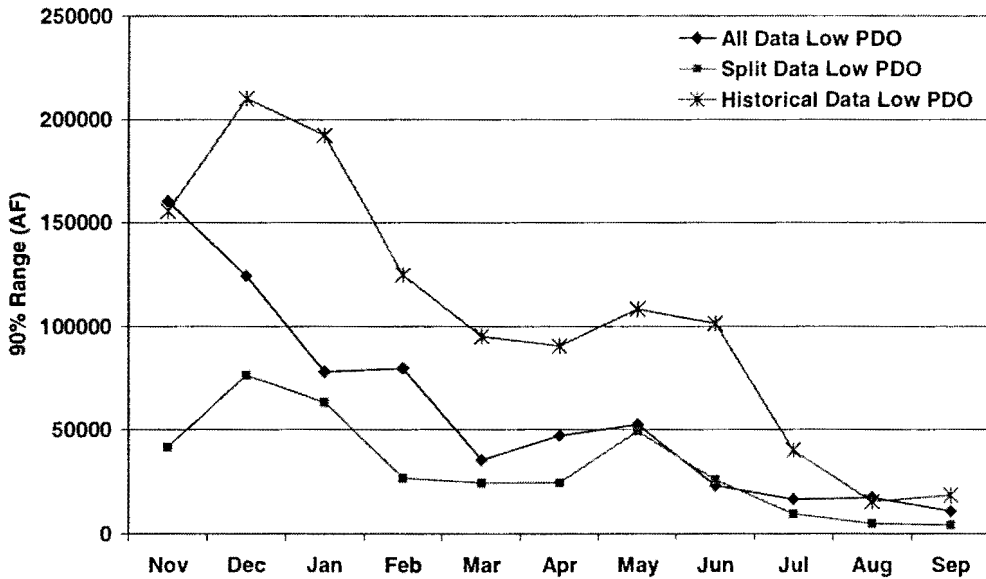


Figure 33: Variability by month for the November-September volume during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Nov - Sep Volume), Warm PDO Phase

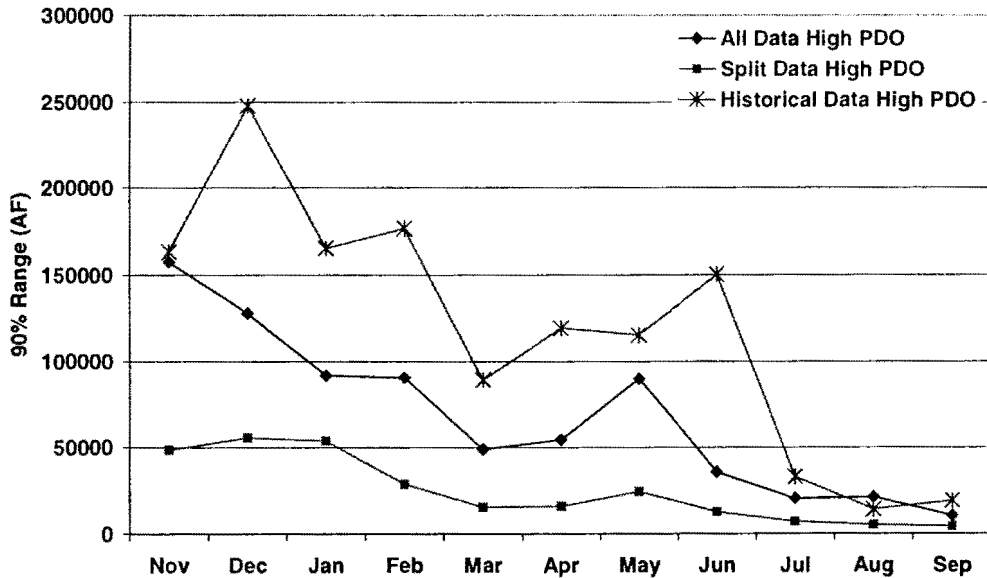


Figure 34: Variability by month for the November-September volume during the warm phase of the PDO (high PDO index).

Sandy River nr Marmot (Dec - Sep Volume), Cool PDO Phase

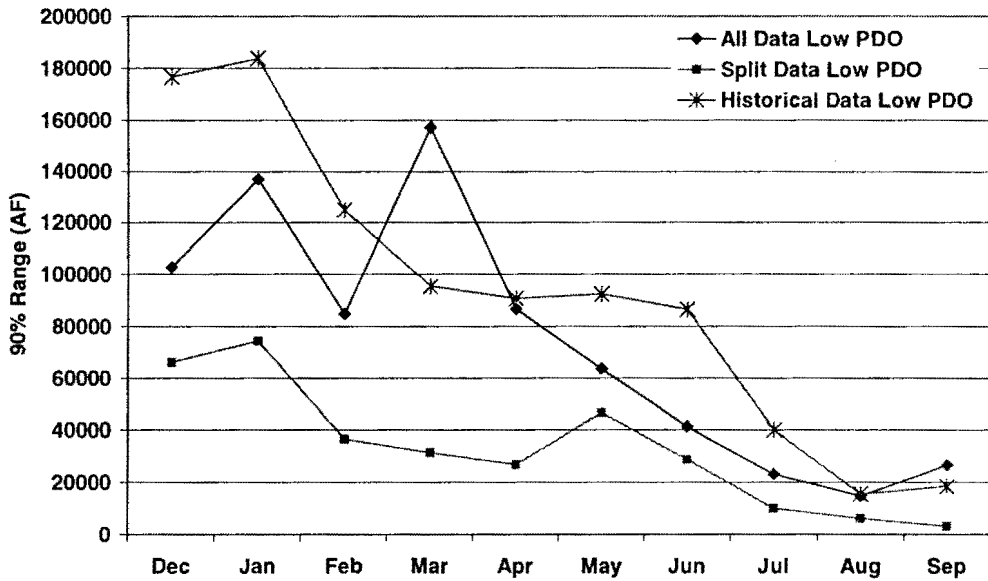


Figure 35: Variability by month for the December-September volume during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Dec - Sep Volume), Warm PDO Phase

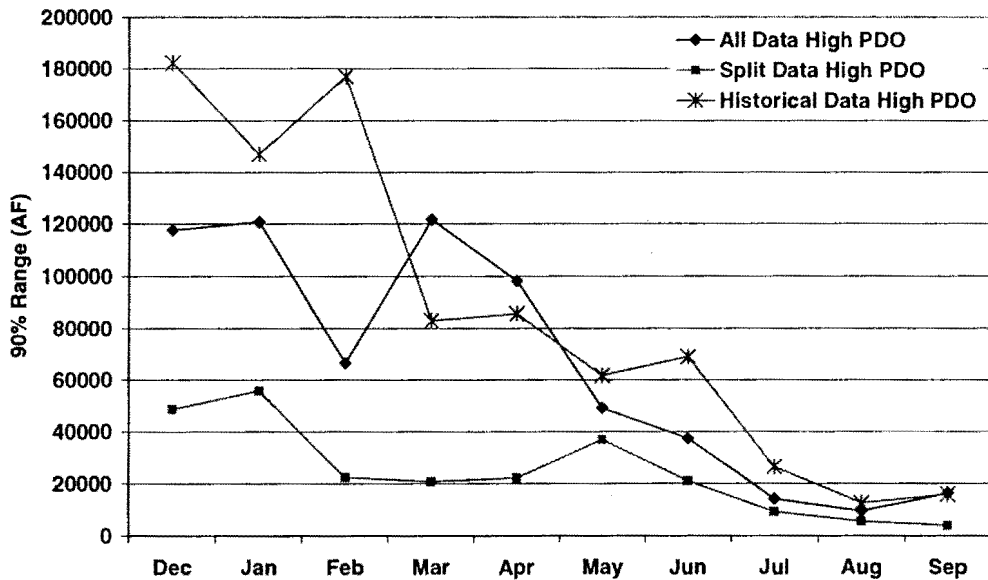


Figure 36: Variability by month for the December-September volume during the warm phase of the PDO (high PDO index).

Sandy River nr Marmot (Jan - Sep Volume), Cool PDO Phase

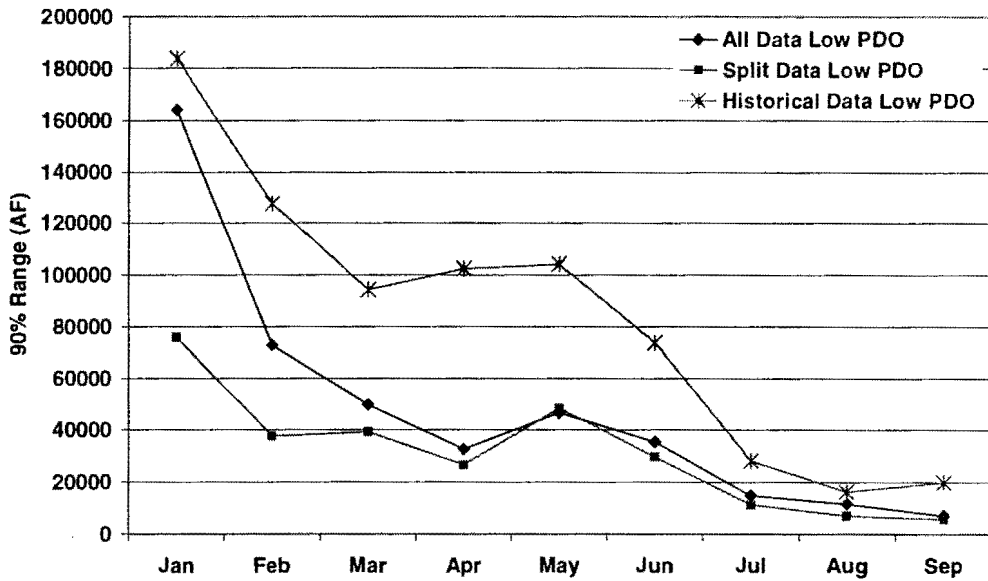


Figure 37: Variability by month for the January-September volume during the cool phase of the PDO (low PDO index).

Sandy River nr Marmot (Jan - Sep Volume), Warm PDO Phase

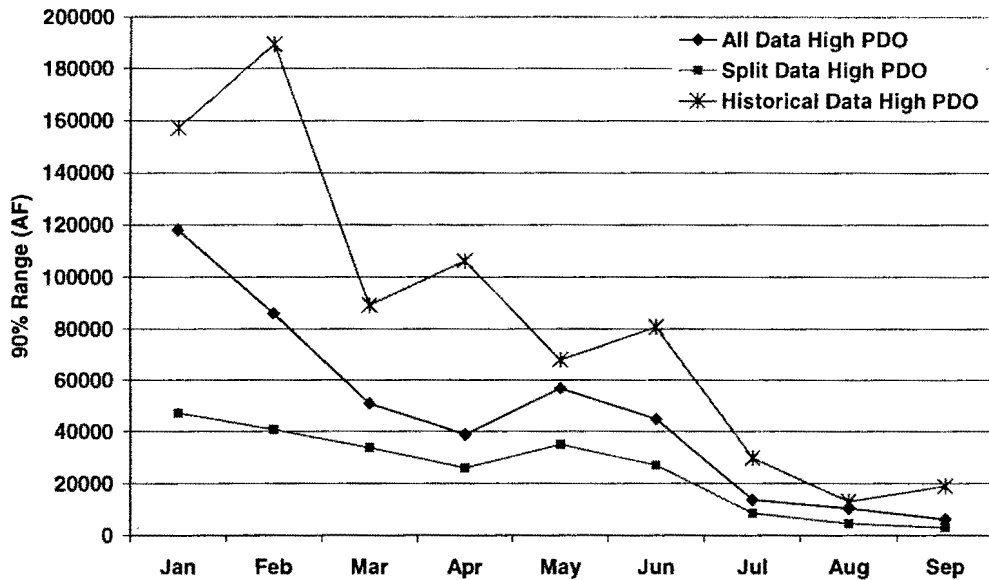


Figure 38: Variability by month for the January-September volume during the warm phase of the PDO (high PDO index).

The consistency of the results from the November volume on suggests that the split disaggregation scheme generate less variable results as compared to history or the control (all data). All ten sets of equations show results similar to those shown above.

In general, the results displayed in Figures 20 – 39 show that disaggregating the seasonal volume forecasts produced for both the warm and cool phases of the PDO and mixing them results in better average estimates of monthly flow and less overall variability than the disaggregated seasonal volumes produced by the all data forecasts.

Chapter 5

Forecasting Process and aY2K Forecast

To this point, the volume forecast and disaggregated forecasts have been compared based on their performance over the historical record. The ultimate verification of mixing forecasts based on the PDO is to produce a forecast for the water year of 2000 (October '99 to September '00) with parameters fit to the data preceding the forecast or prior to October 1999.

5.1 Methods

The process of actually producing a real-time forecast is slightly different than verifying one method versus others. The steps required to produce a forecast are as follows:

1. Fit parameters of the warm and cool PDO equations from the historical data (as described in Chapters 3 and 4).
2. Fit the parameters of the warm and cool PDO disaggregation equations from the historical volume and monthly streamflow data.
3. Generate 100 warm and cool PDO forecasts with the equations fit in (1) and the appropriate predictors.
4. Disaggregate the 100 warm and cool PDO forecasts with their respective disaggregation models fit in (3).
5. Mix the disaggregated sequences in proportion to the probability of the PDO.

In general, the first two steps are only required when initially fitting the models or when the models are updated with additional data. Steps 3 through 5 are carried out each time a forecast is produced.

5.2 Results

The forecasts are mixed in step (5) using the PDO index to predict the state of the PDO as was done in Chapters 3 and 4. Since the previous three years (1997-1999) of the PDO index were on average positive, 89 of the warm sequences were mixed with 11 of the cool sequences to produce the mixed forecast shown in Figure 39. The mixed forecast shown in Figure 39 is the mixed disaggregated October to September seasonal volume produced with data prior to October 1999.

The forecast is very near the median value in October, January, and March through May. The most variable months are historically November, December, and January. Of those the November and December values were close to within the 50% range of median, and the January value was nearly perfect. Overall, the forecast is good with all the values falling within 90% confidence bounds and most of the values falling within the 50% confidence bounds.

Year 2000 split model forecast for the Oct - Sep volume, Sandy River nr Marmot

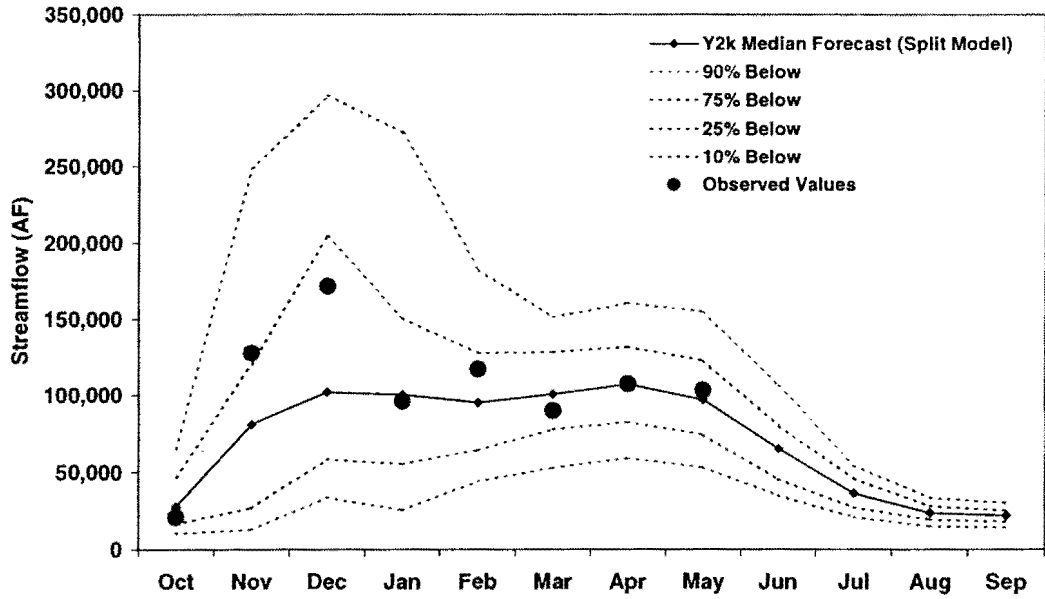


Figure 39: Year 2000 Forecast made by mixing warm and cool PDO disaggregated seasonal streamflow volume forecasts.

Chapter 6

Summary and Conclusions

The split sample analysis performed on the Sandy, Rogue, and Skykomish rivers revealed that the phase of the PDO as defined by Mantua et al. (1997) conditions the response of these rivers to the ENSO as measured by the SOI, which is important in long lead seasonal streamflow volume forecasting in the western U.S. The analysis found that in cool phases of the PDO, the correlation of the SOI and streamflow in the Sandy, Rogue, and Skykomish rivers is much higher than the same correlation during the corresponding warm phase of the PDO. In fact, the correlation of the SOI and streamflow in the Rogue and Skykomish was not significantly different from zero at the 5% level, when all the data were used and in the years when the PDO is in the warm phase. However, during the years when the PDO is in the cool phase the relationship is quite significant. The consistency of the findings across the study basins suggests a consistent relationship in at least the Washington and Oregon Cascades watersheds.

A new seasonal streamflow volume forecasting approach using existing statistical methods was proposed to make use of this new information. The purpose is to generate sequences of monthly data for use in water resource management and as input to a decision support system for river/reservoir operation. The new model was verified by cross validation and by a forecast model developed on all the data without consideration of the PDO. The mixing of the split seasonal volume

forecasts resulted in more accurate mean forecasts with more consistent estimates of variability. Further, the mixing of the split forecasts better estimated the flows in the upper terciles, however improvements in lower tercile estimates are inconclusive. The control or all data equations underestimated the high flows during the cool PDO phase and overestimated the flows during the warm PDO phase. While the mixed split forecasts were still slightly biased, the new model produced results that were significantly better than the control.

The importance of the mixed split forecasts is evident even more so from the results in Chapter 4. Here it was shown that the split forecasts produced consistent mean monthly forecasts with less variability in each of the early forecasts with the exception of the October-September forecast, which made use of the historical mean and standard deviation as a substitute for the warm PDO phase forecast. The results suggest that not only can seasonal streamflow volume forecasts be extended reliably back to the October-September volume, but further that the disaggregation of those forecasts into monthly values can be accomplished to provide additional information to water supply management.

The method was applied to forecast monthly streamflow for the Sandy River for water year 2000 as a verification of the split forecast method. Only the November volume is outside the 50% confidence limits. Further, many of the volumes are very near the median forecast. The June through September 2000 volumes have not yet been observed at the completion of this thesis.

Although it has been shown that mixing disaggregated volume forecasts based on the PDO improves the skill of seasonal streamflow forecasting, the process of splitting the historical data by the phases of the PDO and fitting forecast equations for both phases is data intensive. Many forecast sites will not have data over a long enough period of record to adequately fit forecast equations for both phases of the PDO. Even though some forecast sites will not support the new forecast methods described here, integration of the PDO and disaggregation into statistical streamflow forecasting will benefit water supply management in the future.

References

1. Cassidy, J.J. and D.P. Lettenmaier eds. (1985) A critical assessment of forecasting in water quality goals in western water resources management. American Water Resources Association, Bethesda, MD.
2. Cayan, D.R., and D.H. Petersen (1989) The influence of North Pacific atmospheric circulation on streamflow in the west. Geophysical Monograph 55, American Geophysical Union, pp. 375-397.
3. Cayan, D.R., L.G. Riddle, and E. Aguado (1993) The influence of precipitation and temperature on seasonal streamflow in California. Water Resources Research, 29(4), pp. 1127-1140.
4. Cheiw, F.H.S., Piechota, T.C., Dracup, J.A. and T.A. McMahon. (1998) El Niño/Southern Oscillation and Australian rainfall, streamflow and drought: links and potential for forecasting, Journal of Hydrology, Vol. 204, No. 1-4, pp. 138-149.
5. Croley, Thomas E. II. (1996) Using NOAA's new climate outlooks in operational hydrology. Journal of Hydrologic Engineering, 1(3), pp. 93-102.
6. CRWMG (1993) Review of runoff forecasting in the Columbia River and Pacific slope basins. Report to the Northwest Power Planning Council, October 1993.
7. Garen, David C. (1992) Improved techniques in regression-based streamflow volume forecasting. Journal of Water Resources Planning and Management, 118(6), pp. 654-670.
8. Garen, David C. (1998) ENSO indicators and long-lead climate outlooks: Usage in seasonal streamflow volume forecasting in the western United States, American Geophysical Union Fall Meeting, San Francisco, CA.
9. Garen, David C., Johnson, Gregory L., and Clayton L. Hanson. (1994) Mean areal precipitation for daily hydrologic modeling in mountainous regions. Water Resources Bulletin, 30(3), pp. 481-491.
10. Haan, C.T. (1977) Statistical methods in hydrology. Iowa State University Press, Ames, Iowa.

11. Haan, C.T., and D.M. Allen (1972) Comparison of multiple regression and principle component regression for predicting water yields in Kentucky. Water Resources Research, 8(6), 1593-1596.
12. Hamlet, Alan F. and D.P. Lettenmaier (1999) Columbia River streamflow forecasting based on ENSO and PDO climate signals. Journal of Water Resources Planning and Management, 125(6), pp. 333-341.
13. Helsel, D. and Hirsch, R.M. (1992) Statistical methods in water resources. Elsevier Science, Amsterdam.
14. Kahya, Ercan, and J.A. Dracup (1993) U.S. streamflow patterns in relation to the El Nino/Southern Oscillation. Water Resources Research, 29(8), pp. 2491-2503.
15. Koch, R.W. and A.R. Fisher (2000) Effects of inter-annual and decadal-scale climate variability on winter and spring streamflow in western Oregon and Washington, Proceedings of the Western Snow Conference, Port Angeles, WA, to be published.
16. Koch, R.W. and D. Buller (1993) Forecasting seasonal streamflow: Columbia River at The Dalles. Report prepared for the Bonneville Power Administration, June 1993.
17. Koch, R.W., C.F. Buzzard and D.M. Johnson. (1991) Variation in snow water equivalent and streamflow in relation to the El Niño/Southern Oscillation, Proceedings of the Western Snow Conference, Juneau, Alaska, pp. 37-48.
18. Lane, William L. (1982) Corrected parameter estimates for disaggregation schemes. Statistical Analysis of Rainfall and Runoff. V.P. Singh ed. Water Resource Publications: Littleton, Colorado. pp. 505 – 529.
19. Leung, L.R., A.F. Hamlet, D.P. Lettenmaier, and A. Kuman (1999) Simulations of the ENSO hydroclimatic signals in the Pacific Northwest Columbia River Basin. Bulletin of the American Meteorological Society, 80(11), pp. 2313-2329.
20. Lettenmaier, D.P. and D.C. Garen (1979) Evaluation of streamflow forecasting methods. Proceedings of the Western Snow Conference, Sparks, NV, pp. 48-55.
21. Liang, Xu, D.P. Lettenmaier, E.F. Wood, and S.J. Burges (1994) A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research, 99(D7), pp. 14,415-14,428.

22. Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., and R.C.Francis. (1997) A Pacific inter-decadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society, 78(6),pp. 1069 -1079.
23. Mantua, N.J. (1999) The Pacific Decadal Oscillation, to appear in the Encyclopedia of Global Environmental Change (available at www.atmos.washington.edu/~mantua/REPORTS/PDO/PDO_egec.htm).
24. McCabe, Gregory, and Michael D. Dettinger. (1998) Decadal variability in the strength of ENSO teleconnections with precipitation in the Western United States. PACLIM Conference Proceedings, pp. 69-78.
25. McCuen, R.H. (1985) Statistical methods for engineers. Prentice-Hall, Englewood Cliffs, N.J.
26. Mejia, Jose M., and Jean Rousselle (1976) Disaggregation models in hydrology revisited. Water Resources Research, 12 (2), 185 – 186.
27. Modini, G.C. (2000) Long-lead precipitation outlook augmentation of the multi-variate linear regression streamflow forecasts. Proceedings of the Western Snow Conference, Port Angeles, WA, to be published.
28. Pei, Daniel, Stephen J. Burges, and Jerry R. Stedinger (1987) Runoff volume forecasts conditioned on a total seasonal runoff forecast. Water Resources Research, 23 (1), pp. 9 – 14.
29. Piechota, Thomas C., and J.A. Dracup (1996) Drought and regional hydrologic variation in the United States: Association with the El Nino-Southern Oscillation. Water Resources Research, 32(5), pp. 1359-1373.
30. Piechota, Thomas C., J.A. Dracup, and R.G. Fovell (1997) Western U.S. streamflow and atmospheric circulation patterns during El Nino-Southern Oscillation. Journal of Hydrology, 201, pp. 249-271.
31. Rasmusson, Eugene M. (1985) El Nino and variations in climate. American Scientist, 73, pp. 168-177.
32. Redmond, K.T. and R.W. Koch. (1991) Surface climate and streamflow variability in the Western United States and their relationship to large-scale circulation indices. Water Resources Research, 27(9), pp. 2381-2399.

33. Salas, J.D., J. W. Delleur, V. yevjevich, and W. L. Lane (1980) Applied modeling of hydrologic time series. Littleton, Colorado: Water Resource Publications.
34. Speers, D.D. and J.D. Versteeg (1982) Runoff forecasting for reservoir operations-The past and future. Proceedings of the Western Snow Conference, Reno, NV, pp. 149-156.
35. Stedinger, Jery R., Daniel Pei, and Timothy A. Cohn (1985) A condensed disaggregation model for incorporating parameter uncertainty into monthly reservoir simulations. Water Resources Research, 21 (5), 665 – 675.
36. Trenberth, Kevin E. (1997) The definition of El Nino. Bulletin of the American Meteorological Society, 78(12), pp. 2771-2777.
37. Twedt, T.M., J.C. Schaake, Jr. and E.L. Peck (1977) National weather service extended streamflow prediction. Proceedings of the Western Snow Conference, pp. 52-57.
38. Valencia, Dario, and John C. Schaake Jr. (1973) Disaggregation processes in stochastic hydrology. Water Resources Research, 9 (3), 580-585.
39. Williams, Kenneth L. (1979) Interagency coordination in water supply forecasting: Progress and opportunities. Proceedings of the Western Snow Conference, Reno, NV, pp. 112-115.

Appendix

Forecast Equations

Variable	Period	Description	Oct - Sep	Nov - Sep	Dec - Sep	Jan - Sep	Feb - Sep	Mar - Sep	Apr - Sep	May - Sep	Jun - Sep	Jul - Sep
Snow Water Equivalent	jan	Clear Lake Snotel				9966						
	feb	Clear Lake Snotel										
	feb	Red Hill Snotel					2455					
	mar	Clackamas Lake Snotel						6607				
	mar	Clear Lake Snotel										
	mar	Red Hill Snotel										
	apr	Clackamas Lake Snotel							4515	2162	1177	451
	apr	Clear Lake Snotel										
	apr	Red Hill Snotel										
Monthly Volumes of Precipitation	may	Clear Lake Snotel								2985	1625	622
	Oct	Headworks, Portland OR		17098								
	Nov	Headworks, Portland OR			9956							
	Dec	Headworks, Portland OR										
	Jan	Headworks, Portland OR										
	Feb	Headworks, Portland OR										
	Mar	Headworks, Portland OR										
	Apr	Headworks, Portland OR										
	May	Headworks, Portland OR										
	Jun	Headworks, Portland OR										
	Oct	Government Camp, OR										
	Nov	Government Camp, OR										
	Dec	Government Camp, OR										
	Jan	Government Camp, OR										
	Feb	Government Camp, OR										
	Mar	Government Camp, OR							7104	4877	2654	1017
	Apr	Government Camp, OR										
	May	Government Camp, OR										
Jun	Government Camp, OR											
SOI	jun-sep	SOI	81247	69781	57763	51588	40133	19339				
		Intercept	994023	833218	744079	662007	524387	386834	247526	164092	104922	68854

Table A1: Variables used in the 10 control forecast equations. Dependent variables are in the far left column and the location of the variables is located in the column headings.

Variable	Period	Description	Oct - Sep	Nov - Sep	Dec - Sep	Jan - Sep	Feb - Sep	Mar - Sep	Apr - Sep	May - Sep	Jun - Sep	Jul - Sep
Snow Water Equivalent	jan	Clear Lake Snotel				5413						
	feb	Clear Lake Snotel					2907					
	feb	Red Hill Snotel					985					
	mar	Clackamas Lake Snotel						6839				
	mar	Clear Lake Snotel										
	mar	Red Hill Snotel										
	apr	Clackamas Lake Snotel							5056	3005	1532	492
	apr	Clear Lake Snotel										
	apr	Red Hill Snotel										
Monthly Volumes of Precipitation	may	Clear Lake Snotel										
	Oct	Headworks, Portland OR		21211	12595							1136
	Nov	Headworks, Portland OR			7454	4247						
	Dec	Headworks, Portland OR										
	Jan	Headworks, Portland OR					3226					
	Feb	Headworks, Portland OR										
	Mar	Headworks, Portland OR										
	Apr	Headworks, Portland OR										
	May	Headworks, Portland OR										
	Jun	Headworks, Portland OR										2206
	Oct	Government Camp, OR				4919						
	Nov	Government Camp, OR										
	Dec	Government Camp, OR										
	Jan	Government Camp, OR					2462					
	Feb	Government Camp, OR										
	Mar	Government Camp, OR										
	Apr	Government Camp, OR										
	May	Government Camp, OR										
Jun	Government Camp, OR											
SOI	jun-sep	SOI										
		Intercept		748702	626931	556145	438009	358023	285871	203709	128212	59233

Table A2: Variables used in the 10 warm PDO forecast equations. Dependent variables are in the far left column and the location of the variables is located in the column headings.

Variable	Period	Description	Oct - Sep	Nov - Sep	Dec - Sep	Jan - Sep	Feb - Sep	Mar - Sep	Apr - Sep	May - Sep	Jun - Sep	Jul - Sep
Snow Water Equivalent	jan	Clear Lake Snotel				6851						
	feb	Clear Lake Snotel										
	feb	Red Hill Snotel					1603					
	mar	Cleckamas Lake Snotel						3174				
	mar	Clear Lake Snotel										
	mar	Red Hill Snotel						1281				
	apr	Cleckamas Lake Snotel										
	apr	Clear Lake Snotel										
	apr	Red Hill Snotel							1964			
	may	Clear Lake Snotel								5275	3193	780
Monthly Volumes of Precipitation	Oct	Headworks, Portland OR										
	Nov	Headworks, Portland OR										
	Dec	Headworks, Portland OR										
	Jan	Headworks, Portland OR										
	Feb	Headworks, Portland OR						3179				
	Mar	Headworks, Portland OR										
	Apr	Headworks, Portland OR										
	May	Headworks, Portland OR										
	Jun	Headworks, Portland OR										
	Oct	Government Camp, OR										
	Nov	Government Camp, OR			13712	10877	7183					
	Dec	Government Camp, OR										
	Jan	Government Camp, OR										
	Feb	Government Camp, OR										
	Mar	Government Camp, OR							10132	9562	5787	1869
	Apr	Government Camp, OR										
	May	Government Camp, OR										
	Jun	Government Camp, OR										
SOI	jun-sep	SOI	100132	98334	75221	52249	33077					
		Intercept	1055376	1008045	750961	588995	485528	383359	208556	132376	75646	55424

Table A3: Variables used in the 10 cool PDO forecast equations. Dependent variables are in the far left column and the location of the variables is located in the column headings.